Final Technical Report to the National Institute of Justice on Award Number 2011-MU-MU-K402

Maximizing mtDNA Testing Potential with the Generation of High-Quality mtGenome Reference Data

Authors:

Rebecca S. Just[a,b], Walther Parson[c,d] and Jodi A. Irwin[a,b,e]

[a]Armed Forces DNA Identification Laboratory, 115 Purple Heart Dr., Dover AFB, DE, 19902
[b]American Registry of Pathology, 9210 Corporate Blvd., Suite 120, Rockville, MD, 20850
[c]Institute of Legal Medicine, Innsbruck Medical University, Müllerstrasse 44, Innsbruck, Austria
[d]Penn State Eberly College of Science, 517 Thomas Building, University Park, PA, 16802
[e]Federal Bureau of Investigation, 2501 Investigation Parkway, Quantico, VA 22135

Recipient Organization:

American Registry of Pathology
9210 Corporate Blvd., Suite 120
Rockville, MD 20850

Project Period:

April 1, 2011 - March 31, 2015

Abstract

Forensic mitochondrial DNA (mtDNA) testing requires high quality reference population data for estimating the rarity of questioned haplotypes and, in turn, the strength of the mtDNA evidence. Though novel methods that quickly and easily recover mtDNA coding region data are becoming increasingly available, the appropriate reference data and database tools required for their routine application in forensic casework have been lacking. The primary goal of this National Institute of Justice funded grant was to address these deficiencies by: 1) increasing the large-scale availability of high quality entire mitochondrial genome (mtGenome) reference population data, and 2) improving the information technology infrastructure required to access/search mtGenome data and employ them in forensic casework.

To meet the first of these objectives, we developed a Sanger-based sequencing strategy which could be performed in high-throughput fashion on robotic instrumentation. Using this strategy and an intensive, multi-step data review process, we produced 588 full mitochondrial genome haplotypes, from anonymized, randomly sampled blood serum specimens representing three U.S. population groups (African Americans, U.S. Caucasians and U.S. Hispanics), that meet all current forensic data quality standards. Nearly complete resolution of the haplotypes was achieved with full mtGenome sequences for all three populations, and comparisons to published control region (CR) datasets demonstrated that the databases we have developed are as representative as the reference data on which haplotype frequency estimates presently rely.

To achieve the second objective, we modified the existing structure of the European DNA Profiling Group mtDNA Population database (EMPOP) to both store and query full mtGenome reference data. In addition, we further improved the utility of the database for forensic applications by the addition of a number of new features. These additions include software that performs

automated mtDNA haplogroup estimations for both full and partial mtGenome sequences, updated population structure schemes for all mtDNA data currently housed in the database, and various tools that permit both searches and visualization of the geographic distribution of mtDNA haplogroups, sequences and individual sequence variants.

The full mtGenome population reference data we have produced, and the tools for their storage and use that we have developed in the new version of EMPOP (EMPOP3), will provide a solid foundation for the generation of complete mtGenome haplotype frequency estimates for forensic applications. The thoroughly vetted data can serve as a standard against which the quality and features of future mtGenome datasets (especially those developed via next generation sequencing) may be evaluated, and the extensive data review methods we applied can be used as a model for future mtGenome databasing initiatives. Our successful use of a semi-automated processing strategy on forensic-like samples provides laboratories with practical insight into the feasibility of producing complete mtGenome data in a routine casework environment, and our detailed empirical data on amplification success rates across a range of DNA input quantities will also be useful moving forward as PCR-based strategies for mtDNA enrichment are considered for targeted next generation sequencing workflows. Modifications to the EMPOP database will not only permit the query of full mtGenome datasets to assess haplotype frequencies, but will also provide substantially enhanced functionality for analyses of both complete and partial mtDNA sequences. Overall, the products we have delivered will provide the essential groundwork for expanded use of mtDNA typing in both the short and long term, as next generation sequencing methods begin to be employed in forensic casework practice.

Acknowledgements

A number of key scientists made enormous contributions of both time and expertise to this grant project, and their dedication and exceptional quality of work was critical to its successful completion. We are grateful to Melissa Scheible, Kimberly Andreaggi, Spence Fast, Alexander Röck, Elizabeth Lyons, Jocelyn Bush, Jennifer Higginbotham, Michelle Peck, Joseph Ring, and Gabriela Huber. We also thank Catarina Xavier, Christina Strobl, Toni Diegoli, Martin Bodner, Liane Fendt, Petra Kralj, Simone Nagl, Daniela Niederwieser, and Bettina Zimmerman for contributing their time and considerable skill.

We also wish to recognize the contributions of a long list of scientific colleagues, and technical and administrative personnel across multiple institutions who have been involved in or have supported the work in various ways. These individuals include Lara Adams, Thomas Callaghan, James Canik, Lanelle Chisholm, Richard Coughlin, Michael Cummings, Arne Dür, COL Louis Finelli, Constance Fisher, Shairose Lalani, Odile Loreille, Charla Marshall, Timothy McMahon, Jon Norris, Anthony Onorato, Eric Pokorak, CAPT Edward Reedy, Lt Col Laura Regan, James Ross, Richard Scheithauer and Cynthia Thomas.

Finally, we wish to express our sincere gratitude to our NIJ program managers, Minh Nguyen and Chad Ernst, who have been incredibly supportive and helpful throughout the project.

Table of Contents

Executive Summary

Note: Portions of the text as well as figures and tables throughout this report have been taken from the authors' published papers on our rationale, methods and research findings related to this project. Specifically, text, figures and tables from Irwin et al., Röck et al., Lyons et al., Just et al. and Just et al. [1-5] were used in the preparation of this report.

Mitochondrial DNA (mtDNA) testing in the forensic context requires appropriate, high quality population reference data for estimating the rarity of questioned haplotypes, which is necessary to assess the strength of mtDNA evidence. Since 2003 the Armed Forces DNA Identification Laboratory (AFDIL) has been systematically generating mtDNA data to augment available reference population data and ultimately to improve the framework upon which forensic mtDNA typing is based. However, these data, and indeed all available forensic mtDNA reference databases, have only included information from the control region (CR). While this CR information has obviously strengthened the foundation upon which current mtDNA identification efforts are based, these data have not adequately prepared the field for the advancements in mtDNA typing technologies – especially next generation sequencing (NGS) - that will soon facilitate the acquisition of entire mitochondrial genome (mtGenome) information from forensic specimens.

The lack of appropriate, randomly-sampled and high quality entire mtGenome reference data, and a means to access and use those data for forensic purposes, has been the focus of this National Institute of Justice (NIJ) funded grant. The intent of this project was to: 1) increase the large-scale availability of high quality entire mtDNA genome reference population data, and 2) improve the information technology infrastructure required to access/search mtGenome data and employ them in forensic casework. The specific goals and objectives of this large-scale databasing effort were the development of 550 complete, high quality mtGenomes spanning three U.S. population groups (African American, U.S. Caucasian and U.S. Hispanic), and database structure

and query modifications to the publicly-available European DNA Profiling Group mtDNA Population database (EMPOP; www.empop.org) [6].

To assure the generation of the highest quality mtGenome profiles, we designed an optimized, Sanger-based PCR and sequencing strategy, as well as a laboratory processing workflow in which nearly all pipetting steps – from initial sample placement through sequence detection – are performed robotically [3]. Amplification of the full mtGenome in eight fragments was targeted to facilitate sample processing in 96-well plate format, a strategy that permits 11 samples (plus the appropriate negative controls) to be amplified simultaneously. An established primer set which captures the complete CR in an 1198 bp fragment [7] was utilized, and seven new, overlapping coding region amplicons were designed. PCR and sequencing primers were designed using published mtGenome sequences and position-specific relative substitution rates [8] to reduce the potential for both reaction failures and non-specific binding. The final 135 sequencing reaction strategy produces redundant and overlapping forward and reverse sequence coverage across the complete mtGenome. Developmental validation of the protocol and high-throughput workflow indicated a high rate of success in the first pass of automated processing on DNA inputs ranging from 10 pg to 1 ng, and on samples representing a range of mtDNA haplogroups. Overall, this optimized, highly automated data production strategy reduces data generation costs, hands-on laboratory time and – most importantly - opportunities for human error by substantially decreasing the number of manual production steps and the extent of sample reprocessing necessary to construct complete mtGenome haplotypes [3]. In addition to its utility for the production of forensic-quality full mtGenome data by Sanger-based methods, the PCR strategy also has clear application as a straight-forward method to enrich samples for mtDNA for massively parallel sequencing.

For the databasing effort, we applied the strategy to 625 anonymized, randomly-sampled blood serum specimens from the Department of Defense Serum Repository [9]. As cellular components have been removed from blood serum by centrifugation, the only DNA remaining is cell-free or residual – and thus for a set of blood serum extracts quantified prior to amplification, we found that DNA quantities averaged just 15 pg/µL. Due to the high incidence of multiple amplification failures observed when DNA input quantities for PCR were less than 10 pg, we modified our sample handling strategy for these low DNA quantity specimens to maintain a high rate of throughput. For instance, we doubled the enzyme input for PCR for two of the amplicons in some instances (such as when evidence of inhibition had been previously observed); and samples that exhibited a single PCR failure during the initial, automated processing were manually reamplified to obtain PCR product that could be carried through to sequencing, whereas samples for which more than one of the eight target mtGenome regions failed to amplify were typically not processed further.

To meet all current standards for mtDNA data quality [10,11], we devised a rigorous, multi-step data review process for application to our mtGenome data. The data review workflow is a version of the review strategy described by Irwin *et al.* [7], modified for complete mtGenome data considering a) the use of a multi-amplicon PCR strategy, and b) the dearth of automated tools available to perform quality control of full mtGenome haplotypes and datasets. In the initial assembly of the data, sequences were aligned relative to the revised Cambridge Reference Sequence (rCRS; [12,13]) following phylogenetic alignment rules [10,11,14]. Review of the raw data was performed by at least three distinct scientists at the two laboratories involved in this project: AFDIL and the Institute of Legal Medicine, Innsbruck Medical University (GMI), curator of the EMPOP database [6]. The mtGenome haplotypes developed by each laboratory were

compared electronically, and any discrepancies were investigated by visual inspection of the raw data (and in a few instances, by generation of additional sequence data). Completed haplotypes were assigned a haplogroup using the automated software, EMMA [2], and were examined in comparison to PhyloTree [15] haplogroup motifs to confirm phylogenetic consistency across the eight amplicons and enable a review of all private mutations (substitutions and indels not associated with the haplogroup definition). In addition, all point heteroplasmies (PHP) were screened with reference to nuclear mitochondrial pseudogenes (NUMTs; [16]) to ensure that NUMT contamination was not represented in the final profiles. In accordance with best practice guidelines [10,11], all data transfer steps into internal databases and between laboratories were performed electronically.

The results we present on the databasing initiative in this final report pertain both to the data generation effort and the haplotypes themselves. We analyzed data processing metrics to glean useful information related to performance of the PCR protocol and the capture of mtGenome data from low DNA quantity specimens, and developed review metrics to provide insight that could benefit future data production efforts [4]. For the mtGenome haplotypes, we analyzed the datasets with respect to a) the increased resolution that could be achieved by complete mtGenome typing in comparison to CR sequencing, b) the statistical weight of a matching novel haplotype, c) published CR datasets for the same three U.S. populations, and d) features (such as the pattern and distribution of heteroplasmy) observed across the full mtGenome [5]. Finally, we also include results related to the EMPOP back-end database infrastructure and new features for the web interface that will substantially improve the functionality and utility of the database for forensic mtDNA investigations.

A subset of the blood serum extracts used for the databasing initiative were quantified using a qPCR assay targeting a 143 bp mtDNA fragment [17,18]. Among these 242 samples, 13.6% exhibited at least one PCR failure during the first-pass automated processing. The vast majority of the PCR failures, and all but one instance in which multiple mtGenome target regions for the same sample failed to amplify, occurred when DNA input quantities were less than 10 pg. At input quantities equal to or greater than 10 pg, more than 99% of PCR reactions were successful. For a large majority (70.9%) of a set of 433 low DNA quantity samples that were carried through to sequencing and examined with respect to sequence coverage, the first pass of automated data generation produced redundant sequence coverage across the entire mtGenome and no manual repeat sequencing was necessary. Among the samples in this set for which manual reprocessing was performed, nearly 50% required only one or two additional sequencing reactions to achieve complete sequence coverage across the entirety of the molecule. When sequence coverage was examined with respect to PCR input for a set of 230 blood serum extracts, nearly all samples that required more than two manual sequencing reactions had input quantities less than 50 pg. When inputs for PCR met or exceeded 50 pg, the average number of additional sequences required was 0.33 - which equates to only one manual reaction for every three mtGenome haplotypes. Similarly, an examination of sequence failures (defined as two or more of the four signal intensities below 100 relative fluorescent units) revealed that sequence failure rates were three times higher when PCR products were developed from less than 50 pg of input DNA.

Overall, the detailed processing data indicate that DNA input quantities which meet or exceed 10 pg should result in the highest rate of successful amplification using the PCR strategy we developed; and for Sanger-based sequencing, input concentrations for PCR of 50 pg or greater will reduce the opportunities for human error inherent in manual sample reprocessing and

minimize the cost to generate each mtGenome haplotype. More generally, though, the PCR and sequencing success rates demonstrate that it is feasible to generate forensic-quality complete mtGenome haplotypes in a production (*i.e.*, high-throughput) environment from forensic-like (low template) specimens, and, further, provide clear evidence that amplicons exceeding 2,000 bp can regularly be recovered from very low DNA quantity specimens. The detailed information we include in this report on both PCR and Sanger sequencing success rates across a range of qPCR-measured mtDNA quantities should help to inform future Sanger and NGS-based mtGenome databasing efforts, and may be useful to forensic practitioners when attempting to determine the specific mtDNA amplicons, assays or markers to pursue when DNA quantities are known and case sample extract volumes are limited.

Ultimately, we developed a total of 588 complete mtGenome haplotypes from three U.S. populations: African American (n=170), U.S. Caucasian (n=263) and U.S. Hispanic (n=155). Nearly complete resolution of the haplotypes was achieved with full mtGenome sequencing: a total of 580 distinct haplotypes were observed when all heteroplasmies and length variants were considered. When cytosine insertions at positions 309, 573 and 16193 were ignored, 90.3%, 97.0% and 98.8% of the haplotypes per dataset (U.S. Hispanic, U.S. Caucasian and African American, respectively) were unique.

To investigate the increase in mtDNA lineage discrimination achieved by sequencing the full mtGenome, we developed summary statistics for each population sample using data from hypervariable region (HV) 1 alone, HV1 and HV2 in combination, the full CR, and the complete mtGenome. For the African American population sample, for example, the increase in the number of unique haplotypes that would be detected by HV1 and HV2 sequencing compared to HV1 sequencing alone is 13.2%; and moving from HV1 and HV2 typing to complete CR sequencing

would increase the number of unique haplotypes detected by 8.3%. In comparison to CR sequencing, complete mtGenome sequencing would increase the number of singletons by 29.2% for this population sample – well more than double the increase seen by moving either from HV1 alone to HV1/HV2, or from HV1/HV2 to the full CR.

Given the substantially greater haplotype resolution with full mtGenome data, we also investigated the likelihood ratios (LRs) that would be calculated for a previously unobserved full mtGenome haplotype using two different methods: Clopper-Pearson [19] (currently recommended for use in U.S. laboratories [10]), and the "kappa method" published by Brenner [20], which incorporates the proportion of the database composed of singletons (haplotypes observed only once) into the LR calculation. The analysis clearly demonstrated that for novel mtGenome haplotypes, the Clopper-Pearson approach will result in LRs that are smaller than CR-based LRs, due entirely to the disparity in the size of current mtGenome versus CR databases and despite the distinctly lower random match probability when the complete mtGenome is typed. Thus, until larger full mtGenome databases are available, there would seem to be some benefit in statistical approaches that take into account both database size and database.

On the basis of the EMMA [2] analyses and comparisons to Build 16 of PhyloTree [15], 393 distinct named haplogroups were assigned to the 588 haplotypes. Across the three population samples, all major known haplogroups were represented except L4, L5, L6, O, P, Q, S and Z. Based on the assigned haplogroups, the haplotypes were classified into one of four broad biogeographic ancestry categories: African, East Asian, West Eurasian and Native American. As has been previously reported [21], self-identified ancestry was highly correlated with maternal biogeographic ancestry for the African American and U.S. Caucasian populations. For the African American dataset, the vast majority of haplotypes (90.0%) were assigned to haplogroups L0, L1,

L2 and L3; and similarly, 94.7% of the U.S. Caucasian haplotypes in the population sample are of West Eurasian ancestry. By contrast, while the majority (60.0%) of the U.S. Hispanic population sample was comprised of Native American lineages, West Eurasian and African maternal ancestries were also represented in substantial proportions (25.8% and 12.3% of haplotypes, respectively). Using the biogeographic ancestry proportions, we compared the population samples to six previously published CR-based datasets (two large datasets for each population) to investigate their consistency. For the African American and U.S. Caucasian populations, the proportion of haplotypes reflecting the predominant ancestry was not statistically significantly different between this and previous studies. However, for the U.S. Hispanic population, the differing proportions of Native American haplotypes across the three population samples were significant (p=0.007). Specifically, the proportion of Native American haplotypes in the U.S. Hispanic mtGenome data differed significantly from that reported by Allard *et al.* [22] (p=0.008), even after Bonferroni correction for multiple tests. This is most likely due to differences in geographic sampling, which will reflect the substantial regional variation in the Native American component of a U.S. "Hispanic" population sample [7]. Along these lines, the proportion of haplotypes representing Native American maternal ancestry in a recently published Southwest Hispanic population sample from Texas (71.7%; [23]) is highly similar to the frequency of Native American haplotypes (70.8%) in the Allard *et al.* study [22].

The incidence and pattern of both length heteroplasmy (LHP) and point heteroplasmy (PHP) observed in the CR in our 588 haplotypes was generally consistent with previous studies [24-26]. In the coding region, LHP associated with insertions and deletions (indels) was detected at five of the 18 coding region positions at which indels were present in the haplotypes. LHP was also observed in nearly 90% of individuals in an 8 base pair (bp) poly-A tract around nucleotide

positions 12418-12425. In all cases the heteroplasmy was very minor (and likely only detectible due to the generally very high quality the sequence data), and the majority variant matched the rCRS (eight adenine residues). Overall, PHP was observed in 23.8% of individuals across the complete mtGenome. 102 total PHPs were detected in the coding region, and all were unique in the dataset (none were observed in more than a single individual). An examination of these coding region heteroplasmies combined with PHPs reported in three other recent studies ([23,26,27]; approximately 200 PHP in total from approximately 1100 entire mtGenome haplotypes across the four reports), and with respect to coding region substitution rates [8], indicated both 1) an absence of coding region hotspots for PHP, and 2) no correlation between position-specific substitution and heteroplasmy rates in the coding region. This pattern of heteroplasmy is distinctly different than what has been observed in the CR [25], and is likely due at least in part to selection. Further examination of the coding region PHPs detected in our haplotypes in comparison to previous reports on coding region substitution patterns [28-31] also pointed to purifying selection, as both nonsynonymous to synonymous ratios and the estimated pathogenicity of the heteroplasmic variation in protein coding genes are closer to a neutral model of sequence evolution than has been reported for complete substitutions.

Modifications to the EMPOP database to facilitate the use of mtGenome reference data for forensic purposes included updates to both data handling and data search capabilities. The database infrastructure was modified to enable a) the import and export of mtGenome data, b) the storage of mtGenome data, and c) search string functionality [32]. Web interface features that were adapted to enable queries of the full mtGenome data stored in EMPOP included search features that permit the input of full mtGenome data as either differences from the rCRS or as text strings, and search parameters that allow users to select data ranges, inclusion or exclusion of specific indels, and

options for handling heteroplasmies. In addition to these mtGenome-specific modifications, extensive additional functionality was added to EMPOP that relates to examinations of both partial and complete mtGenome haplotypes. Among other features, these additions included a new scheme for geographic/metapopulation characterization that integrates linguistic information; visual displays of haplotype/haplogroup distributions on a global heat map; and the incorporation of the EMMA [2] software to enable automated haplogroup estimations and genetic distance-based costs (based on nucleotide position-specific fluctuation rates) for comparisons between searched and database haplotypes. We expect that this new version of EMPOP, dubbed EMPOP3, will be publicly released in the second quarter of 2015.

The full mtGenome population reference data we have delivered, and the tools for their storage and use that we have developed in EMPOP3, will provide a solid foundation for the generation of complete mtGenome haplotype frequency estimates for forensic applications. Though additional reference population databasing efforts will be required over time to increase mtGenome dataset sizes and refine haplotype frequency estimates, the sequences generated here are a critical start. Not only do they represent individuals from across the United States, but since the sequences themselves have been developed via well-established Sanger technology and with strict data generation and review protocols, we fully expect these data to serve as the quality benchmark against which new mtGenome reference population datasets will be measured. The care taken and safeguards employed in the generation of our haplotypes were instituted specifically to address and rectify issues identified in earlier, CR-based datasets ([33-35]). Given these lessons learned, the overall process we employed should also serve as a model for future NGS-based mtGenome data production, and the data themselves should serve as a critical frame of reference for the development of robust and reliable NGS data analysis and interpretation guidelines.

Our successful use of a semi-automated processing strategy on forensic-like samples provides laboratories with practical insight into the feasibility of producing complete mtGenome data in a routine casework environment, and conclusively demonstrates that large mtDNA fragments can regularly be recovered from very low DNA quantity specimens. The processing metrics we detail in this report on amplification success rates across a range of DNA input quantities should additionally be useful moving forward as PCR-based strategies for mtDNA enrichment are considered for targeted NGS workflows. This practical information gleaned from our experience generating the mtGenome reference data is, in fact, already being put to use to develop complete mtGenome profiles from single hair shafts, using a combination of PCR-based enrichment (using the 8-amplicon strategy we developed) and Illumina sequencing [36].

In addition to providing both actionable information and the necessary infrastructure to perform near-term mtGenome sequencing in casework, the 588 haplotypes developed here are providing, for the first time, an empirical understanding of the true value of mtGenome data for practical mtDNA casework purposes. While it has been well-understood that haplotype resolution increases substantially when mtDNA analysis is extended from the HV regions to the entire CR, it is only now, and as a result of this project, understood that by extending analysis to the entire mtGenome nearly full resolution of mtDNA haplotypes can be achieved. Moving forward, this information should substantially aid the forensic community as it considers whether full mtGenome typing should be performed in casework practice, a critical question that will need to be addressed in short order given the state of investigation into NGS methods for full mtGenome typing on forensic specimens.

Perhaps even more important than the near-term benefits to forensic casework that the project deliverables may bring, this effort lays crucial groundwork for long-term changes in the

way forensic cases are handled. As stated in our original proposal, the overarching objective of this project was to establish the data foundation required for expanded mtDNA-based forensic capabilities. Given the technological advancements in mtDNA sequencing that have been brought on by NGS, it is very possible that mtDNA data will be sought in most or all future cases – even when STR profiles may be recoverable. Complete mtGenome data developed from more samples should lead to more profiles stored in crime, offender and missing persons databases, which should, in turn, lead to more solved crimes and more identified human remains. The products we have delivered here are essential components of what will be needed to make this long-term vision a reality.

Main Body

Note: Portions of the text as well as figures and tables throughout this report have been taken from the authors' published papers on our rationale, methods and research findings related to this project. Specifically, text, figures and tables from Irwin et al., Röck et al., Lyons et al., Just et al. and Just et al. [1-5] were used in the preparation of this report.

I. Introduction

MtDNA testing has long been a mainstay in forensics, primarily as a result of its utility on a variety of evidentiary specimen types that harbor little or no nuclear DNA. Although data from multiple autosomal short tandem repeats (STRs) are generally preferred in forensic investigations because they can provide much higher discriminatory power than mtDNA data, many sample types are too degraded to yield informative STR profiles. As a result, mtDNA sequencing is regularly performed on samples such as shed hairs, fingernails and severely degraded skeletal elements due to the relative abundance of intact mtDNA (present in several thousand copies per cell) as compared to nuclear DNA in these types of remains [37-43]. MtDNA data is also useful in forensic investigations for the same reasons it is useful in studies of human evolutionary history – not only is it highly variable and thus useful in distinguishing individuals, but it is also maternally inherited without recombination. In forensic testing, this is valuable because all maternal relatives possess identical mtDNA haplotypes in the absence of mutation, and thus even distant relatives can provide reference material for mtDNA comparisons. This is extremely beneficial in missing persons cases when references from immediate family members are unavailable for standard STR testing.

MtDNA typing in forensic casework has historically focused on the two HV regions of the non-coding CR [39-41,44]. These approximately 600 bases have the highest average substitution rate in the mtGenome, and thus present the greatest opportunity for inter-individual differentiation while minimizing data generation effort. It is the case, however, that examination of these 600

bases alone limits the power of forensic mtDNA testing in general, leading to situations in which HV1 and HV2 data do not provide sufficient discriminatory information to resolve distinct maternal lineages. Further resolution is often obtained by increasing the range of data analyzed to additional portions of the CR (*e.g.*, with a sample of Austrians, analysis of the entire CR reduces the random match probability from 0.011 to 0.008) [45-47], and in recent years complete CR profiles have become the forensic standard for population reference data [10,11,48,49]. Yet, many individuals will remain indistinguishable despite complete CR data. In those cases, variation in the mtDNA coding region can be targeted [50-54].

It has been shown that mtDNA coding region data can be useful in a number of forensic casework situations. For instance, it has been valuable in: resolving multiple casualty cases where more than one reference family shared the same mtDNA CR haplotype [55,56]; sorting and re-association of commingled remains [56]; increasing statistical support when exclusionary references were unavailable [57]; mtDNA haplogroup typing for rapid screening of casework specimens [58-60]; and assessing maternal bio-geographic ancestry as an investigative tool [61,62]. Additionally, coding region information has been strategically targeted in cases for which extremely limited evidentiary material is available following standard and, in these situations, non-distinguishing CR testing. In order to preserve the little remaining evidence for analyses likely to provide resolution, coding regions from the relevant reference samples were first investigated to identify sites that distinguished the reference lineages. These *case-specific* discriminatory sites were then directly typed on the remaining evidence material to ultimately establish identity [63].

Still, even in these very specific casework scenarios, it has generally been impractical to Sanger sequence large portions of the mtGenome. The cost and effort required to obtain even partial CR profiles by Sanger-based methods from case specimens is substantial (especially in

comparison to standard STR typing), in part because mtDNA sequence data is usually sought when the genetic material is severely limited and/or compromised. Numerous short amplicons with adequate overlap among them, significant sequence coverage over each amplicon to ensure sufficient data quality, and highly redundant data analysis and review are required to produce CR haplotypes. Generation of Sanger-based coding region data for resolution of specific cases has therefore been not only prohibitively laborious for most practicing forensic laboratories, but also limited by the availability of sufficient evidentiary material. As a result, the forensic methods to access coding region data have historically involved either the optimization of a published assay or in-house development of sequencing or single nucleotide polymorphism (SNP) typing protocols that minimize effort and sample consumption ([64-66], for example).

Few commercial off-the-shelf products have been available for the generation of coding region data, and those that have been evaluated for forensic use have limited utility due to sample quantity requirements and other issues related to data quality standards required for forensic application [67,68]. However, new massively parallel sequencing technologies hold great potential for efforts to expand forensic mtDNA typing beyond current capabilities. Since the first such technology was introduced in 2005 [69], NGS has transformed genetic data generation in many fields of research, including ancient DNA (for an overview of some ancient DNA studies that have used NGS, see Table 1 in Knapp and Hofreiter [70]; and for a review of the application of NGS to mtDNA sequencing in particular, see Ho and Gilbert [71] and Paijmans *et al.* [72]). The advantages of NGS in comparison to traditional Sanger-type sequencing that have been exploited for analyses of ancient samples also have clear relevance to the low DNA quantity and/or quality specimens to which mtDNA typing is often applied in forensics, and these technologies will likely make their way into forensic practice in the very near term. Recent studies have demonstrated both that 1)

NGS can effectively recover complete mtGenome profiles even from highly damaged and degraded forensic samples [73,74], and 2) that full mtGenome sequencing by NGS may be cost-effective in comparison to methods currently used by the forensic community for mtDNA data generation [23]. While much further work remains before NGS-based protocols (whether for mtGenome or nuclear genome typing) can be fully validated for forensic use and routinely applied to forensic casework specimens, the ongoing research into NGS for forensic application in many laboratories [23,73-86] clearly indicates the direction in which the field is moving. Thus, it seems that in the very near future, routine use of mtDNA coding region data in forensic casework will depend primarily upon the availability of high quality entire mtGenome population reference data that can be used to assess the rarity of mtGenome haplotypes, rather than the ease with which coding region data can be generated from forensic specimens.

GenBank has been the only repository of complete mtGenomes that is regularly updated with new information. Although it contains a growing number of complete sequences, the available data are an imperfect substitute for a forensic reference database. Most of the sequences available in GenBank have not been produced as randomly sampled, unrelated individuals that are representative of particular population groups. For those populations that are represented, the datasets tend to be inconsistent in terms of the associated metadata required for their use in the forensic context. Further, because GenBank data are neither curated nor quality control checked, many sequences contain errors that may not only obscure precise estimates of mtDNA substitution rates (as required for likelihood calculations; [87]), but, more importantly, may also confound estimates of mtDNA haplotype frequencies. Finally, the tools available for GenBank searches are not the most useful for practical casework application. Search parameters that are specific to forensic mtDNA queries, including specific reference populations, inclusion/exclusion of

polycytosine indels, and pre-defined sequence ranges, are unavailable and difficult to accommodate in the BLAST [88] interface. Even novel tools that support the access and handling of GenBank mtDNA sequence data (*e.g.*, MitoVariome [89]) fail to address specific alignment issues in length variant regions that are relevant to sequence comparisons in forensic casework [14].

Efforts have been underway for several years to improve and expand publicly-available forensic mtDNA CR data sets, and nearly 35,000 haplotypes representing more than 60 populations are now available in EMPOP [6]. While these data have substantially strengthened the foundation upon which current mtDNA identification efforts are based, they do not adequately prepare the field for the new technological advancements in mtDNA typing that are beginning to facilitate access to coding region information in the most difficult forensic specimens.

On this background, the work we proposed for this grant-funded initiative aimed to address these deficiencies by increasing the large-scale availability of high quality entire mtDNA genome reference population data, and improving the information technology infrastructure required to access/search mtGenome data and employ them in forensic casework. Our primary goals and objectives were:

1. The production of 550 high quality entire mitochondrial genomes from U.S. populations groups. We aimed to sequence 150 U.S. “Caucasians”, 150 U.S. “Hispanics” and 150 U.S. African Americans, plus 100 additional mtGenomes from one or more of these U.S. populations groups, and make those data available to the forensic and academic communities for casework and other applications.
2. Database structure and query modifications to the publicly-available EMPOP database [6]. The EMPOP database is currently the world’s leading reference mtDNA population database

for forensic applications, but the data structure and interface required modification to both accommodate entire mtGenome sequences and permit queries of entire mtGenome sequence data.

II. Materials and Methods

MtGenome sequencing protocol development and developmental validation

The databasing portion of this grant project required a robust PCR and Sanger sequencing strategy that could be applied in high-throughput fashion to produce complete mtGenome haplotypes which meet the highest data quality expectations. As developmental validation of the protocol we designed for this effort was required prior to its use for databasing, both the development and developmental validation of the protocol are described in this Materials and Methods section of this report, rather than in the Results section.

Amplification of the full mtGenome in eight fragments was targeted to facilitate sample processing in 96-well plate format, a strategy that permits 11 samples (plus the appropriate negative controls) to be PCR-amplified simultaneously (Figure 1; all report figures can be found in Appendix A). An established primer set which amplifies the complete CR in an 1198 bp fragment [7] was utilized, and the development of seven new overlapping amplicons to span the coding region is described below.

Given the need for a robust mtGenome assay that could be applied with equal efficacy across samples representing diverse mtDNA haplogroups, the potential for primer binding site mutations was given careful consideration in the design of coding region amplification primers. To this end, a "global alignment" developed from 193 complete or coding region only mtDNA sequences was used to assess regions of the mtGenome appropriate for primer placement. The

alignment consisted of mtGenome sequences from most major named mtDNA haplogroups (six sequences each from haplogroups A, B, C, D, E, F, G, H, I, J, K, L0, L1, L2, L3, L4, L5, M, N, P, Q, R, R0, S, T, U, V, W, X, Y, and Z; and four and three sequences from haplogroups L6 and O, respectively) sampled at random using GenBank accession numbers available on the PhyloTree mtDNA phylogeny [15]. In addition, published mtGenome substitution rate data drawn from 2196 complete mtGenomes [8] were used to develop a substitution rate histogram by nucleotide position (not shown). In combination, the global alignment and substitution rate graph were used to identify haplogroup-specific mutations and overall highly polymorphic positions and/or regions which could potentially interfere with proper primer annealing during PCR.

Initially, 22 coding region amplification primers employed for earlier mtGenome sequencing at AFDIL [52,90-93] were evaluated for use in the new protocol. The global alignment and substitution rate histogram described above were used to assess the potential for primer binding site mutations, and the web-based Primer3 program [94] was used to examine primer characteristics such as melting temperature, GC content, and self-complementarity. Based on the criteria applied all previously used primers were disqualified from further use, most due to potential primer binding site issues. Thus, seven new coding region amplicons were designed using the global alignment and substitution rate histogram. Within bp ranges deemed acceptable (by virtue of a high degree of sequence conservation, a lack of haplogroup-defining mutations, and sufficient overlap with neighboring amplicons), specific primer sequences were selected using the default settings in Primer3 [94]. The PCR primer sequences are given in Appendix C. The average overlap between amplicons is 210 bp, with a minimum overlap of 71 bp (between Amplicons 7 and 8) and a maximum overlap of 338 bp.

Considerations given highest priority in the design of the mtGenome sequencing strategy were 1) the desire to develop high-resolution sequence coverage in both the forward and reverse directions across as much of the molecule as feasible, and 2) a protocol that would be amenable to high-throughput processing on automated liquid-handling instrumentation. For the CR, the sequencing approach described by [7] was adopted. Coding region primers previously utilized by AFDIL for mtGenome sequencing [52,90-93] were evaluated using the global alignment, substitution rate histogram, and Primer3 software [94], as described above. In addition, the typical quality of the sequence data produced by 74 of these primers was assessed by inspection of 2237 previously-generated sequence electropherograms, and only primers which routinely produced data with sufficient signal and minimal noise were considered for further use. As a result of these examinations, 20 sequencing primers were maintained for use in the new protocol. Ninety-nine new primers were selected in the same manner as described above for the PCR primers, with old and new coding region sequencing primers spaced at intervals designed to produce overlapping, high-resolution forward and reverse sequence coverage across the genome.

The final, 8-amplicon mtGenome strategy is depicted in Figures 2 and 3. The number of sequencing primers per amplicon ranges from 16 to 18, and the strategy produces 135 sequences from 127 unique primers. The resulting redundant sequence coverage across the complete mtGenome is demonstrated in Figure 4. Sequencing primers and their sources (published or new) are listed in Appendix D.

To assess the sensitivity of the amplification protocol, PCR was performed in duplicate for a range of positive control (9947A) DNA input quantities (300 pg, 100 pg, 25 pg, 10 pg, 5 pg, 2.5 pg, 1.0 pg, 0.5 pg 0.25 pg, and 0.1 pg). The PCR products were quantified using the QIAxcel Advanced system (QIAGEN Inc., Valencia, CA) and the resulting values were normalized with

respect to amplicon size to enable direct comparison. Figure 5 displays a box and whisker plot of the normalized amplification product concentrations, reported in ng/µL per 1000 bp, at each DNA input concentration. The long whiskers (highly variable product concentrations) at each DNA input level reflect the range of sensitivities of the eight primer pairs, however all regions were successfully amplified down to 10 pg of input DNA. Beginning at 5 pg input DNA a few amplification failures were observed, and below 1 pg input successful amplification was sporadic and limited to a few high-efficiency primer pairs.

To evaluate the performance of the protocol on a variety of haplotypes, 11 anonymous, high quality population samples from ten distinct mtDNA haplogroups (A, B, C, D, H, U, K, L1, L2 and L3) were amplified in duplicate and sequenced using an automated, high-throughput process. The DNA input for PCR varied by sample, and ranged from approximately 0.1-1.5 ng. Trimming and assembly of the raw electropherograms for replicate samples was performed by separate individuals according to laboratory standard guidelines for data quality in terms of background to noise ratio and peak resolution. Sequence coverage across the molecule was assessed in terms of a) redundant and bi-directional coverage, b) the degree of additional manual re-processing that would be required to develop complete replicate coverage, and c) the correlation between sequence coverage and sequence distance from the rCRS. The final haplotypes for each sample were compared to control data (complete mtGenome profiles previously developed from the same sample extracts [92]).

High quality sequence data (as defined by signal to noise ratio) was developed from most primers for most samples in a single pass with the automated processing. As Figure 6 depicts, on average 99.87% (s.d. 0.23%) of the mtGenome was covered by at least 2 sequences, and 99.07% (s.d. 0.67%) of the mtGenome had both forward and reverse sequence coverage when small

regions with unidirectional coverage due to length heteroplasmy in HV1 and HV2 were ignored. The number of manual resequencing reactions that would be required to achieve redundant coverage ranged from zero to two (Figure 7), with approximately one resequencing reaction required for every two complete haplotypes. Considering that 135 sequences were generated for each sample, this equates to a 0.32% resequencing rate. A weak but non-significant correlation was observed between mtGenome coverage and sequence distance from the rCRS, with a mere 1-3% of the variance in mtGenome coverage attributed to sequence distance (data not shown). In all cases the final haplotype matched the haplotype previously developed for each sample.

Though amplification of nuclear insertions of mtDNA (NUMTs) is unlikely when sufficient mtDNA is present in a sample [95], the reference assembly of the complete human genome was nonetheless queried using PrimerBLAST [88] for the seven coding region amplification primer pairs. Any close sequence matches (defined as 75% or greater overall similarity for both primers, with no more than one mismatch in the 3′ most 5 bp) that could potentially amplify a fragment similar in size to the authentic mitochondrial target (less than 500 bp difference) were further evaluated. For each potentially amplifiable nuclear genome region, the percentage similarity to modern mtDNA was assessed by aligning the NCBI reference sequence to the rCRS in Sequencher version 4.8 (Gene Codes Corporation, Ann Arbor, MI). When the nuclear genome sequence region could not be aligned to the rCRS due to high dissimilarity, the percentage similarity was noted as being less than 60%.

Using the described criteria, 13 potentially amplifiable regions of the nuclear genome were identified (Table 1; all numbered tables can be found in Appendix B). Of these, only three had a sequence similarity to the rCRS greater than 90%. For the two Chromosome 1 regions with greater than 98% sequence similarity to Amplicon 3 (2291 bp) and Amplicon 4 (2511 bp), the

Chromosome 1 sequence differed from the rCRS sequence at 35 and 36 nucleotide positions, respectively. This region in Chromosome 1 corresponds to a described NUMT approximately 5842 bp in length [96].

Sequencing artifacts (*i.e.*, small regions of compression and/or unusual peak morphology) due to region-specific sequence motifs were reproducibly observed in both the positive control samples sequenced during protocol development and the developmental validation on population samples. Typically, each artifact was observed in a single sequence direction, and the severity of the artifact varied by primer distance from the artifact. An example of a sequencing artifact is shown in Figure 8.

The standard sequencing protocol used at AFDIL for high-throughput generation of mtDNA population data includes one-quarter the recommended volume of BigDye® Terminator v1.1 Ready Reaction Mix (Life Technologies, Applied Biosystems, Foster City, CA) and replaces 25% of the dITP-containing BigDye® with dGTP BigDye® Terminator v1.1 Ready Reaction Mix ([93]; reaction volumes are specified in the "MtGenome data generation" subsection below). The addition of dGTP BigDye® was originally implemented to assist the sequencing of difficult templates, specifically GC-rich regions or polycytosine tracts, in the reverse direction. For this protocol, dGTP BigDye® was eliminated from forward sequencing reactions to reduce the number of artifacts produced in those sequences (Figure 8). All remaining artifacts (nearly all of which occurred in the reverse direction) that were consistent and reproducible across multiple samples, and with replicate sequencing, were cataloged. In practice in our laboratory, this catalog is referenced during assembly and analysis of mtGenome sequences, and known artifacts are annotated in the assembled contig.

It is worth noting here that these types of sequencing artifacts are typically only apparent and recognizable as such because the quality of the sequence data produced using this protocol is generally pristine. With even a small amount of noise in the sequence data, many of these artifacts would not be evident. In general, the artifacts do not confound data interpretation, as they are typically minor and apparent in only one sequencing direction. Nevertheless, when previously uncatalogued artifacts were encountered during data production, we noted the affected bases as ambiguous and resequenced the region to confirm that the authentic sequence was represented in the consensus sequence for the region.

EMMA development

A brief overview of the development of the maximum-likelihood based automated haplogrouping software, EMMA, is included here. Further details on the published mtGenome sequences that were examined and utilized for EMMA, the calculation of fluctuation rates, the assignment algorithm and its parameters, and the developmental validation of EMMA can be found in Röck *et al*. [2].

The phylogenetic tree in PhyloTree [15] represents known global mtDNA variation by defining haplogroups and their signature mutations. For EMMA development, an R script [97] was designed that transforms the tree into a list of hypothetical haplotypes carrying the signature mutations of the respective haplogroups (tree nodes) as differences to the rCRS. As these haplotypes are inferred rather than observed in the real world they are herein referred to as virtual haplotypes. The defined haplogroup motifs in PhyloTree are based on a database of published mtGenomes (http://www.phylotree.org/mtDNA_seqs.htm) that were downloaded from GenBank and evaluated for their application within EMMA. Incomplete haplotypes, haplotypes in which

errors were identified, and haplotypes containing ten or more ambiguity designations in the FASTA string were excluded. All FASTA strings from non-excluded mtGenomes were translated into rCRS-coded haplotypes using SAM [32] and subsequently checked with in-house software to harmonize alignment. The quality filtering of the sequences ultimately resulted in a database of 14,990 full mtGenomes, stored with their accession numbers and version. In conjunction with the 3925 virtual PhyloTree motifs, these 18,915 virtual and real mtGenomes form the basis for haplogroup estimation.

Haplogrouping of mtDNA data in rCRS-based format requires consistent alignment and notation of sequences following a phylogenetic approach [14] in order to assess the stability of mutations in defined haplogroups. We refer to this mutational (in)stability as a fluctuation rate. The weighting scheme presented for the string-search method in Röck *et al*. [32] was updated by assessing the stability of mutations within the mtDNA CR among 19,171 full control region haplotypes for which raw lane data were available. Haplogroups were manually assigned to all sequences in this dataset following the classification outlined in PhyloTree Builds 12 through 15 [15]. Subsequently, the sequences were grouped into discernable CR haplogroup clusters (*i.e.*, clusters of haplogroups that can be confidently determined based on CR motifs), and fluctuation rates were calculated for different categories of CR mutations using the formula and conditions specified in Röck *et al*. [2]. The fluctuation rates developed for the CR mutations were expanded to the coding region using the number of occurrences of coding region mutations reported by Soares *et al.* [8].

With regards to the assignment algorithm, a test profile is compared to every database profile with the same or a larger reading frame (sequence span) than the test profile. As the number of positions can be large (*e.g.*, about 16.6 thousand for full mtGenomes), and the computational

work can become too high, the ranking of the base profiles is determined by minimizing the costs for the change from the base profile symbol to the test profile symbol using the fluctuation rates. In the output of the algorithm only base profiles with the lowest and second lowest costs are presented, and a tolerance value is used to cluster the optimal and suboptimal profiles.

Specimens and sampling

The samples used for the databasing initiative were anonymized blood serum specimens from the Department of Defense Serum Repository (DoDSR; [9]). The 175 African-American, 275 U.S. Caucasian, and 175 U.S. Hispanic samples initially targeted for processing were selected randomly from specimens in the DoDSR collection. Specimens were received with only state and self-reported population/ethnicity information. The research was reviewed by the U.S. Army Medical Research and Materiel Command's Office of Research Protections, Institutional Review Board Office for applicability of human subjects protection, and it was determined to not meet the definition of human subjects research according to federal regulations (see Department of Defense regulation 32 CFR 219.102(f)).

MtGenome data generation

Blood serum specimens were robotically transferred from the tubes in which they were received to 96-well plates. Genomic DNA was extracted from 100 µl of blood serum using the QIAamp 96 DNA Blood Kit (QIAGEN, Valencia, CA), and a combination of robotic pipetting and manual centrifugation. DNA was eluted from the silica columns using either 100 µl or 200 µl of TE buffer (10mM Tris and 0.5mM EDTA), and the eluate was evaporated to eliminate any potential ethanol carryover. DNA extracts were resuspended in 100 µl of either UV-irradiated

deionized water or TE buffer. Some, but not all, DNA extracts were quantified prior to PCR, using an mtDNA quantitative PCR (qPCR) assay [18] adapted from Niederstätter *et al.* [17].

PCR (using the primers listed in Appendix C) was performed in a 50 µL total reaction volume, generally using 5 µL GeneAmp 10X PCR Buffer I (Life Technologies, Applied Biosystems), 4 µL GeneAmp® dNTP blend 10 mM (Life Technologies, Applied Biosystems), 2 µL of each 10 µM amplification primer, 3 µL DNA extract, 0.5 µL (2.5 units) AmpliTaq Gold® DNA Polymerase (Life Technologies, Applied Biosystems), and 33.5 µL deionized water. When qPCR results indicated DNA quantities less than 10 pg/µL, extract input for PCR was doubled from 3 µl to 6 µl. In some cases, such as when specimens from the same extract plate had previously exhibited evidence of inhibition, or to improve first-pass processing success for one or two of the eight mtGenome region targets with the poorest amplification efficiency among the lowest DNA quantity specimens, polymerase inputs were doubled from 2.5 to 5 units. Thermal cycling parameters were: 96°C hold for 10 minutes; 40 cycles of 94°C for 15 seconds, 56°C for 30 seconds, 72°C for 2.5 minutes; and a 72°C hold for 7 minutes.

High-throughput amplification of the mtGenome (for the initial, automated sample processing) was performed on a liquid-handling instrument (MICROLAB® STARlet, Hamilton Robotics, Reno, NV), utilizing single-use, pre-made tubes of amplification master mix (containing all amplification reagents except enzyme) for each amplicon. The master mix was pre-aliquoted and stored in labeled 1.7 mL tubes at −20°C, with the polymerase added prior to PCR reaction set-up. Any necessary repeat amplification was also performed using the pre-made master mixes, and manual sample transfer steps were witnessed (by at least one, and sometimes two additional scientists) to ensure correct sample placement.

Amplification success was assessed by capillary electrophoresis. PCR products were injected directly from the 96-well amplification plate on a QIAxcel Advanced instrument (QIAGEN Inc.), and sizing of the products was performed using the QX alignment marker 50 bp/5 kb and the QX DNA size marker 250 bp-4 kb (QIAGEN Inc.). To maintain a high rate of throughput and minimize manual reprocessing, extracts for which only a single region failed to amplify were re-amplified manually prior to sequencing, whereas samples for which more than one fragment failed to amplify were typically abandoned and not processed further.

Purification of PCR products prior to sequencing was performed enzymatically, using 10 µL Exonuclease I and 5 µL recombinant Shrimp Alkaline Phosphotase (Affymetrix, USB, Cleveland, OH) per 50 µL PCR product. For purification of a full 96-well plate, a master mix of 1100 µL Exonuclease I and 550 µL recombinant Shrimp Alkaline Phosphotase was prepared, and 15 µL of the master mix was manually pipetted to each sample PCR product. Negative controls in column 12 of the 96-well plate were not purified. Thermal cycling conditions were 37°C for 20 minutes followed by 90°C for 20 minutes.

Each mtGenome was initially sequenced in a total of 135 reactions using 127 unique primers. The sequencing primers used for each of the eight mtGenome amplicons are listed in Appendix D. Sequencing reactions included 8 µL deionized water; 6 µL dilution buffer (400 mmol/l TRIS, 10 mmol/l $MgCl_2$, pH 9.0); either 2 µL BigDye® v1.1 (Life Technologies, Applied Biosystems) for forward sequencing reactions, or 1.5 µL BigDye® v1.1 and 0.5 µL dGTP BigDye® v1.1 (Life Technologies, Applied Biosystems) for reverse sequencing reactions; 2 µL sequencing primer at 10 µM; and 2 µL PCR product for a total reaction volume of 20 µL. Thermal cycling conditions were as follows: 96°C hold for 1 minute, followed by 25 cycles of 96°C for 15 seconds, 50°C for 5 seconds, and 60°C for 2 minutes.

For high-throughput sequencing of 11 amplified samples at a time, all pipetting steps were performed on a liquid-handling instrument (MICROLAB® STARplus, Hamilton Robotics) using a master mix of sequencing reagents and pre-made, single-use primer plates. The robotically prepared primer plates included 50 µL of each 10 µM primer according to the plate layouts in Appendix E. Sequencing plate maps (eight forward and eight reverse, for a total of 16) for the high-throughput processing are shown in Appendix F. Manual resequencing was performed when the first pass robotic processing did not produce complete sequence coverage (defined as at least two strands of sequence data) across the entire mtGenome. In most instances the reprocessing involved manual sequencing from the original PCR products to fill in small gaps in the sequence coverage. However, when multiple new sequences from the same genome region were required, the sample extract was sometimes re-amplified to produce a better quality PCR product. Reaction set-up for repeat sequencing was performed manually, and sample placement was witnessed by at least one additional scientist.

Sequence product purification was performed via gel filtration. For our high-throughput process, Performa DTR V3 96-well short plates (Edge Biosystems, Gaithersburg, MD) were used, and purification steps were performed in two eight-plate batches. Performa plates were first manually centrifuged at 850 x g for two minutes to remove some liquid (as per the manufacturer's recommendation), then pipetting from the sequencing product plates to the prepared Performa plates was performed robotically. Filtration into new, barcoded 96-well plates was accomplished by centrifugation for 5 minutes at 850 g.

Purified sequence products were evaporated by heated vacuum centrifugation, then resuspended in 10 µL Hi-Di™ Formamide (Life Technologies, Applied Biosystems). Sequence detection was performed by capillary electrophoresis on a 3730 Genetic Analyzer (Life

Technologies, Applied Biosystems) using a 50 cm array, the FastSeq instrument protocol (FastSeq50_POP7 run module) and the SeqAnalysis Fast analysis protocol (Basecaller_3730POP7RR) with the default instrument settings. Post-detection, raw signal data were initially processed on the 3730 Genetic Analyzer computer using Sequencing Analysis v5.3.1 (Life Technologies, Applied Biosystems) with the spacing parameters set to 12.0.

MtGenome data review

The data review workflow employed for this project is a version of the review strategy described by Irwin *et al.* [7] modified for complete mtGenome data developed using a multi-amplicon strategy. The workflow was developed in accordance with the current Scientific Working Group on DNA Analysis Methods (SWGDAM) and International Society for Forensic Genetics (ISFG) best practice guidelines for forensic mtDNA data development [10,11]. Data review was performed by at least three distinct scientists across two different laboratories: AFDIL, and GMI, curator of the EMPOP database [6]. In detail, the steps were as follows:

Initial assembly, trimming and review of the raw sequence data for each sample was performed in Sequencher version 4.8 or 5.0 (Gene Codes Corporation, Ann Arbor, MI). Sequences were aligned to the rCRS following phylogenetic alignment rules [10,11,14]. In cases of LHP, a single dominant variant was identified (as per recommendations for mtDNA data interpretation [11,14]). With regard to PHP, an mtGenome position was deemed heteroplasmic only if specific criteria were met upon visual review of the raw sequence data:

1) If the minor sequence variant was readily visible (that is, a distinct peak of normal morphology with white space beneath it could be seen in the trace data without changing the

chromatogram view in Sequencher to examine the signal closer to the baseline) in all forward and reverse sequences covering the position, a PHP was called.

2) If the minor sequence variant was readily visible in some but not all sequences, data closer to the baseline were inspected for each sequence. If the baseline view demonstrated that the minor variant was substantially higher than any sequence background/noise in a) the majority of the sequences, and b) both forward and reverse sequences, a PHP was called.

When heteroplasmy was suspected but not confirmed according to the above criteria, additional sequence data were generated for the sample/region to clarify the presence or absence of heteroplasmy.

Once each sample haplotype was complete (that is, every mtGenome position had at least two strands of high-resolution sequence coverage), a list of differences from the rCRS was prepared manually, and a variance report was electronically exported from Sequencher.

Each mtGenome haplotype contig generated during the initial analysis of the raw data was reviewed on a position-by-position basis by a second scientist. A list of differences from the rCRS was generated manually and compared to the list generated at the prior stage, and any discrepancies were resolved to the satisfaction of both reviewers. A variance report was again exported from Sequencher, and compared to the manually-prepared lists of differences from the rCRS to ensure full agreement across all paper and electronic records. In addition, sequences present in the final sample contig were visually examined to confirm that all sequences had the same sample identifier (*i.e.*, that no sequences from a different sample were mistakenly included).

The Sequencher variance reports were then electronically imported into the custom software Laboratory Information Systems Applications (LISA; Future Technologies Inc., Fairfax VA). For each sample, the imported record was compared to the handwritten list of differences

from the rCRS prepared in the previous data analysis stage to confirm that the database record was consistent with the paper record. In addition, all coding region indels, PHPs and transversions in each electronic profile were visually confirmed by re-review of the raw data at the relevant positions. To confirm the database haplotypes, a second scientist again reviewed each electronic record in comparison to the previously-generated lists of differences from the rCRS, and checked that the correct sequence coverage range (1-16,569 base pairs) was associated with each profile.

Given the multi-amplicon PCR protocol used for data generation in this project, each mtGenome haplotype was evaluated for phylogenetic feasibility as a quality control measure. Haplotypes were first assigned a preliminary haplogroup, and subsequently compared to the then-current version of PhyloTree (Build 14 or 15, depending on the dates on which different subsets of the data were checked) [15] to assess each difference from the rCRS. The raw data for each sample were re-reviewed to confirm a) any expected mutations (based on the preliminary haplogroup) that were lacking, b) all private mutations (mutations not part of the haplogroup definition), and c) all PHPs and transversions.

Following the phylogenetic check, Sequencher project files, variance reports and all raw data for each sample were electronically transferred to EMPOP for tertiary raw data review. At EMPOP, each mtGenome haplotype contig was again reviewed on a position-by-position basis, and edits to the project files were made as warranted. A variance report of differences from the rCRS was exported from Sequencher and imported into a local database.

EMPOP and AFDIL-generated variance reports for each haplotype were electronically compared in the local database at EMPOP. Any discrepancies between the haplotypes were reported to AFDIL; and for those samples with discrepancies, the raw data were re-examined by both laboratories for the positions in question. In a few cases, sample re-processing was performed

at this stage to clarify the haplotypes. The sample haplotypes were considered finalized once both EMPOP and AFDIL were in agreement, and all relevant files had been corrected at AFDIL and re-sent to EMPOP.

Haplogroups were assigned to each mtGenome haplotype using EMMA [2] and Build 16 of PhyloTree [15]. These automated assignments were then compared to the preliminary haplogroups assigned at the phylogenetic check stage, and any discrepancies were evaluated in detail. In all cases, the EMMA-estimated haplogroup was the final haplogroup assigned to the sample.

Finally, the mtGenome haplotypes were further reviewed with respect to the assigned haplogroups and NUMT sequences. All indels relative to the rCRS in the completed haplotypes were reviewed to assess correct placement according to phylogenetic alignment rules [10,11,14] and PhyloTree Build 16 [15], given the assigned haplogroup. All PHPs in the final haplotypes were compared to a list of positions at which two specific NUMTs (on Chromosomes 1 and 5, and possessing greater than 90% similarity to modern human mtDNA; see Table 3 in Lyons *et al*. [3]) differ from the rCRS. Any haplotypes with PHPs that occurred at one of these positions were re-reviewed by careful examination of the raw data to ensure that the PHP was not due to co-detection of a NUMT (which would be expected to present as multiple mixed positions within the amplicon in question [3]).

With regards to data transfer and corrections, all data transfer steps into internal databases and between laboratories were performed electronically. When changes were made to haplotypes at AFDIL after the initial transfer of sample files to EMPOP, all relevant sample files were re-sent to EMPOP for complete replacement (that is, no manual changes were made to haplotypes at EMPOP).

MtGenome data analyses

Summary statistics (number of haplotypes, number of unique haplotypes, random match probability, haplotype diversity and power of discrimination) for multiple regions of the mtGenome (HV1 only; HV1 and HV2 in combination; the complete CR; and the full mtGenome) were based on pairwise comparisons of each of the three populations in the LISA custom software. Cytosine insertions at nucleotide positions 309, 573 and 16193 were ignored for the analyses, and PHPs were treated as differences.

Estimations of broad scale maternal biogeographic ancestry (African, East Asian, West Eurasian or Native American) were based on the haplogroups assigned to each haplotype. For the few haplogroup M, N and U lineages which have overlapping present day distributions in certain geographic regions (North Africa, southern Europe and the Near East), assignment to one of the ancestry categories was made on the basis of the geographic distribution of the same or closely related lineages in global populations represented in a beta version of the EMPOP3 database [98].

Pairwise comparisons of the haplotypes representing each population and biogeographic ancestry group were performed for a) the full mtGenome, and b) with comparisons restricted to the CR, in the LISA custom software. Cytosine insertions at nucleotide positions 309, 573 and 16193 were ignored for the analyses.

Statistical calculations to assess significance were performed either in Microsoft Office Excel 2010, or, for Chi-Square tests of independence (for comparisons of differing proportions), using the calculator spreadsheet available from http://udel.edu/~mcdonald/statchiind.html [99].

Likelihood ratios (LRs) were developed using two methods: the "exact" method for confidence intervals (Clopper-Pearson) [19] and the "kappa method" [20]. Clopper-Pearson 95% confidence intervals were calculated using HaploCALc version 1.8 by Steven Myers

(steven.myers@doj.ca.gov). LR calculations using the one-tailed confidence interval used the standard formula LR = *x*/*y*, where *x* represents the probability that the questioned and known haplotypes represent the same maternal lineage, and *y* is the probability that the questioned sample will match an unrelated (or only randomly related) haplotype in the database. The value used for *x* was always 1, and the value used for *y* was the one-tailed 95% confidence limit. LR calculations for the kappa method used equation 6 from Brenner [20]: $LR_{\kappa} = n/(1-\kappa)$, where κ represents the proportion of haplotypes in the population sample that are singletons (haplotypes observed only once), and *n* represents the size of the population sample.

III. Results

Data production and review

For the subset of blood serum extracts that were quantified (via a qPCR assay using a 143 bp mtDNA target [17,18]; 242 of the extracts in total), DNA quantities ranged from 0.00 to 777.64 pg/µL with an average of 14.91 pg/µL (s.d. 53.79). Thirty-three of these samples (13.6%) exhibited at least one amplification failure during the first-pass automated processing (Figure 9). The vast majority (86.6%) of the amplification failures, and all but one instance in which multiple regions for the same sample failed to amplify, occurred when DNA input quantities were less than 10 pg. The average DNA quantity for samples with multiple amplification failures was 1.00 pg/µL (s.d. 0.80). At DNA input quantities equal to or greater than 10 pg, 99.4% of amplifications were successful.

As described in the Materials and Methods section, samples that exhibited a single PCR failure during the initial, automated processing were manually reamplified to obtain PCR product that could be carried through to sequencing, whereas samples for which more than one of the eight

target mtGenome regions failed to amplify were typically abandoned and not processed beyond amplification. Out of a total of 625 samples that were attempted, 37 were dropped due to PCR failure in two or more of the eight mtGenome target regions. As Figure 10 demonstrates, among the first 242 quantified samples processed, all 12 samples dropped due to multiple PCR failures had PCR DNA input quantities less than 10 pg/µL [4]. But, as PCR failures can occur due to primer binding site mutations, and those mutations may be haplogroup or lineage-specific, we explored the extent of PCR failure across all 588 completed haplotypes in relation to the PCR strategy employed.

An examination of the incidence and pattern of PCR failure among samples with primer binding region mutations indicates that such mutations are unlikely to have biased the final datasets for any of the three population samples. A total of 52 sequence variants, representing 34 distinct mutations, were found across the 16 primer binding regions. Primer binding region mutations were found in 46 of the 588 completed samples (7.8%), and overall had the potential to impact primer binding in 1.1% of the initial eight high-throughput PCR reactions performed per sample (a total of 4704 PCR reactions). Yet, manual reamplification (due to near or complete PCR failure) was required in only eight of the 52 instances in which a mutation was later found in a PCR primer binding region; and thus sequence variation in primer binding regions potentially caused PCR failure in only 1.4% of samples and 0.2% of amplifications. Further, as Figure 11 demonstrates, the position of the mutation relative to the 3' end of the primer was highly variable in these eight instances of reamplification, and thus the mutation may not have been the reason for the PCR failure in all eight cases. Among the 46 samples which were carried through to sequencing and later found to possess mutations in primer binding regions, five (8.9%) exhibited a mutation in

more than one of the 16 primer binding regions, yet only three PCR failures (of ten potentially affected reactions) were observed among these five samples.

Given the wide variety of mtDNA haplogroups represented by the 588 haplotypes reported in this study (see below), and the low DNA quantities for the first 12 samples abandoned [4], the very low overall incidence of reamplification among samples with known primer binding region mutations suggests that 1) PCR failure due to haplogroup-specific mutations when using the Lyons *et al.* [3] primers is likely to be quite infrequent, and 2) few, if any, of the abandoned samples exhibited multiple PCR failures due to primer binding region mutations. It is therefore unlikely that the PCR or sample handling strategy introduced any particular bias into the datasets we developed.

As described in the Materials and Methods, manual reprocessing was also performed when the first pass robotic processing did not produce sufficient sequence coverage. For a large majority (70.9%) of a set of 433 low DNA quantity samples that were examined with respect to sequence coverage, the first pass of automated data generation produced complete sequence coverage across the entire mtGenome and no manual reprocessing was necessary. For 13.2% and 6.2% of the samples, respectively, minimal (defined as one or two additional sequencing reactions) or moderate (three to nine additional sequencing reactions) reprocessing was required to achieve the desired sequence coverage (Figure 12). For 9.7% of samples more extensive reprocessing (ten or more manual sequencing reactions) was performed, and usually included complete re-amplification of one or more regions of the genome.

Initial results utilizing an earlier version of the amplification strategy made clear that some of the exceptionally low template blood serum specimens required extensive reprocessing for Amplicons 2 and 6 in particular. For instance, among the 40 samples with PCR inputs less than 10

pg processed using the initial Amplicon 2 PCR primers, 12 samples (30.0%) required reamplification and resequencing of that amplicon; and among the 29 samples with PCR inputs less than 10 pg processed using the initial Amplicon 6 PCR primers, 11 samples (37.9%) required reamplification and resequencing of the fragment. To increase the first pass success rates for these two amplicons, the PCR primer sets were redesigned early on in the databasing project. To assess success rates using the published strategy [3], all blood serum samples amplified prior to the PCR primer redesign were reconsidered without the Amplicon 2 or 6 reprocessing requirements. This reduced the number of samples which required moderate or extensive manual sequencing from 15.9% to 10.2%, with only 20 of 433 samples (5.5%) requiring extensive reprocessing (Figure 12).

The extent of manual sequencing required was also examined in comparison to PCR input DNA quantity for a set of 230 extracts (the 242 quantified extracts referenced above, minus the 12 samples which were not processed beyond amplification due to multiple amplification failures; Figure 13). All nine samples which required extensive manual reprocessing and nearly all samples which required moderate manual sequencing had PCR input DNA quantities less than 50 pg. For the nine samples with DNA inputs less than 50 pg which required extensive reprocessing, most of the initial sequence data quality issues were caused by a failure of the post-amplification enzymatic purification which necessitated reamplification and complete manual resequencing of the fragment. Among the 43 samples with input DNA quantities greater than 50 pg, only one sample required more than two manual reactions to achieve complete mtGenome sequence coverage, and the total number of additional sequences required was only 14 – which equates to approximately one manual sequencing reaction for every three haplotypes.

In addition to the more qualitative assessments of sequencing success described above, we also performed a quantitative evaluation of sequencing failure rates in comparison to input DNA

quantity. For a qPCR-quantified set of 185 samples with no amplification failures, Sequence Scanner v 1.0 (Life Technologies, Applied Biosystems) was used to capture the electrophoretic signal intensities for 21,601 sequencing products detected on the 3730 genetic analyzer. For these data, we defined a failed sequence as one with at least two of the four signal intensities below 100 relative florescence units (RFUs). To reflect the published protocol [3], sequences generated from PCR products developed using the initial Amplicon 2 and 6 primer sets (discussed above) were excluded from the analysis.

A scatter plot of the percentage of failed sequences at each PCR input DNA quantity is displayed in Figure 14. For samples for which PCR DNA inputs were less than 50 pg, the average sequence failure rate was 2.51% (s.d. of 0.05), which equates to approximately three failed sequences per sample. Among samples for which PCR DNA inputs were greater than 50 pg, the average sequence failure rate was 0.82% (s.d. of 0.02); and only one of these 39 samples had a sequence failure rate greater than 5.0%. The picture provided by these data is highly similar to that developed from the reprocessing data (Figure 13). These two complementary analyses demonstrate that, using the published protocol [3] with the minor amplification modifications and sample handling strategy described here and in Just *et al.* [4], sequencing was largely successful but variable when PCR input DNA quantities were less than 50 pg, and nearly always successful when DNA input quantities exceeded 50 pg.

Sequencing success/failure was also investigated in relation to QIAxcel-measured amplification product concentration. For the 2677 sequencing reactions performed from PCR product concentrations below 2.00 ng/µl/1000bp included in the analysis, a clear relationship between sequencing failure and product concentration only emerged when the data were broadly categorized. Both the percentage of failed sequences (defined by electrophoretic signal intensities,

as described above) and the resequencing rate (calculated by comparing the number of manual sequences required to the number of sequences produced in the initial automated processing) were higher when PCR product concentrations were below 1.00 ng/µl/1000bp as compared to product concentrations in the 1.01-2.00 ng/µl/1000bp range. When product concentrations were greater than 1.00 ng/µl/1000bp, the resequencing rate was only 0.37%. However, the more obvious trend observed across all of these lower amplification product concentrations was that sequencing failure was highly amplicon-specific. More than 90% of the 198 sequences with low signal intensities resulted from two target regions: Amplicon 4, with 68.0% of the sequencing failures, and Amplicon 6, with 25.1% of the sequencing failures.

To summarize the performance of the processing strategy with respect to the 433 low DNA quantity samples examined, we calculated an overall resequencing rate: the number of manual sequences required in comparison to the 135 sequences generated per sample as part of the initial automated processing. When all manual sequence reprocessing was considered the resequencing rate was 2.84%. However, when data from Amplicons 2 and 6 prior to their redesign was excluded to reflect the published protocol design [3], the resequencing rate was 1.20%. This latter value reflects an average of 1.59 manual sequencing reactions required per sample to develop a complete, forensic-quality mtGenome haplotype from a successfully amplified, low template extract.

During the course of the data generation, a small number of samples were noted as mixtures at the first stage of data analysis, and were initially removed from the sample set. Re-examination of the raw data for four of these samples at a later date revealed 1) a larger number of mixed positions than would be expected from a mixture of two modern mtDNA lineages, 2) mixed data at positions that did not correspond to any known modern mtDNA haplogroup motif, and 3) presence of the mixed data in some PCR target regions but not others. Additionally, one further

sample that was investigated for a potential very low level PHP indicated the same features when new sequence data was produced and examined very close to the baseline. For each of the amplicon regions in these five (total) samples in which mixed data were observed, a consensus sequence of the minority molecule was generated and compared to two NUMT sequences with high similarity to modern mtDNA (Table 2). In each case, the minority consensus sequences matched the Chromosome 1 or Chromosome 5 nuclear DNA reference assembly [100-103] at 96-99% of the 2290-2758 nucleotide positions compared, indicating that a NUMT sequence had been co-amplified with the authentic mtDNA target. Further testing on these samples indicated that a new DNA extraction from the serum specimen could produce sequence data exclusively from the intended mtDNA target region. Thus, for four of the five samples, new sample extracts were used to produce new PCR products for the effected regions, and the new sequence data for each of those regions was used to replace the initial NUMT-contaminated data. These four samples are represented among the final haplotypes. One affected sample, which was not reprocessed due to limited material, was removed from the sample set entirely.

The formalized data review process we employed for this databasing initiative (see the Materials and Methods section) included an electronic comparison of the haplotypes independently developed by AFDIL and EMPOP from the raw sequence data. Across the 588 haplotypes compared, 27 discrepancies in 23 samples were identified, a non-concordance rate of 4.6%. While a few discrepancies were due to raw data editing differences (two instances) or indel misalignments (six instances), the majority of these discrepancies (19 in total, or 70%) were due to differences in the heteroplasmies called in the AFDIL versus the EMPOP analysis. For three of these samples manual reprocessing was performed to generate additional data to determine whether a low-level PHP was or was not present.

In one of these instances, inspection of the new sequence data close to the baseline indicated that NUMT co-amplification was likely present in the affected region, and was the source of the appearance of a very low-level PHP at a single nucleotide position (sample mtGHispTX0043 in Table 2). In the same manner as described above, the affected mtGenome region was reamplified and resequenced in its entirety from a new extraction, and clean sequence data with no indication of NUMT contamination nor PHP was obtained. The identification of probable NUMT co-amplification in the initial processing of this sample – which resulted in sequence data which appeared to contain an extremely low PHP at only a single position – prompted visual re-inspection of the raw data for all mtGenome regions containing PHPs that occurred at positions that could be variant due to one of the two NUMTs previously detected. No further NUMT contamination was found.

In addition to the differences found upon cross-check of the haplotypes, two further indel misalignments were later identified during additional review of the datasets. In both instances the original alignment of the sequence data was inconsistent with phylogenetic alignment rules and the current mtDNA phylogeny [10,11,14,15]. In one case, a haplotype with 2885 2887del 2888del was incorrectly aligned as 2885del 2886del 2887; and in the second case, a haplotype with 292.1A 292.2T was incorrectly aligned as 291.1T 291.2A. For these two haplotypes the indels were misaligned by both AFDIL and EMPOP, and thus no discrepancy was identified as part of the concordance check. The identification of these two misalignments prompted a thorough review of all 2767 indels present in the 588 haplotypes, and no additional misalignments were found.

Figure 15 provides a breakdown of the 29 total data review issues identified during our data production effort. The results of the concordance check and the two additional indel misalignments identified later both 1) underscore the need for multiple reviews of mtDNA

sequence data to ensure correct haplotypes are reported, and 2) highlight a need for an automated method for checking regions of the mtGenome prone to indels prior to dataset publication and inclusion in a database. EMPOP includes a software tool that evaluates CR indel placement and is routinely employed to examine CR datasets prior to their inclusion in the database. Until a similar tool is developed to reliably assess complete mtGenome haplotypes, all indels in complete mtGenome datasets should be reviewed in relation to the current knowledge regarding the human mtDNA phylogeny prior to publication.

Database composition and statistics

In total, 588 complete mtGenome haplotypes were generated from three U.S. populations: African American (n=170), U.S. Caucasian (n=263) and U.S. Hispanic (n=155). The number of samples per U.S. state/territory for each population is given in Appendix G.

The 580 distinct mtGenome haplotypes that were observed are presented in Appendices H-J, and are available in GenBank (accession numbers KM101569-KM102156) and in our publication [5], and will soon be available in EMPOP3 [98]. Summary statistics for each population are given in Table 3. Across the entire mtGenome, 168 of 170 (98.8%) African American haplotypes, 255 of 263 (97.0%) U.S. Caucasian haplotypes, and 140 of 155 (90.3%) U.S. Hispanic haplotypes were unique in the respective datasets when cytosine insertions at positions 309, 573 and 16193 were ignored. With regard to the summary statistics, the additional value added by sequencing the complete mtGenome is most powerfully demonstrated by comparing the information gleaned from the subsets of the molecule historically targeted for forensic typing. For example, for the African American population sample, the increase in the number of unique haplotypes that would be detected by HV1 and HV2 sequencing compared to

HV1 sequencing alone is 13.2%; and moving from HV1 and HV2 typing to complete CR sequencing would increase the number of unique haplotypes detected by 8.3%. In comparison to CR sequencing, complete mtGenome sequencing would increase the number of singletons by 29.2% for this population sample – well more than double the increase seen by moving either from HV1 alone to HV1/HV2, or from HV1/HV2 to the full CR. In the U.S. Caucasian population sample set, full mtGenome sequencing increased the number of unique haplotypes by more than 100% when compared to the resolution achieved by examination of HV1 alone.

Given the substantially higher degree of haplotype resolution with full mtGenome sequences in comparison to smaller portions of the molecule, we investigated the LRs that would be calculated for previously unobserved haplotypes when considering HV1/HV2 alone, the CR and the complete mtGenome using two different methods: Clopper-Pearson [19] and the "kappa method" published by Brenner [20]. Confidence interval calculations with the Clopper-Pearson "exact" method use the cumulative probability from a binomial distribution given the number of observations of interest and a sample size; and thus for previously unobserved haplotypes in a database, Clopper-Pearson 95% confidence intervals (either one-tailed or two-tailed) and the resulting LRs will depend entirely on the size of the reference population sample. By contrast, as Brenner's kappa method uses the proportion of singletons (haplotypes observed only once) in a population sample to approximate the rarity of a new haplotype, the calculated LR for a previously unobserved mtDNA haplotype will depend both on database size and the portion of the molecule targeted (as Table 3 demonstrates that the proportion of singletons will be greater as the size of the targeted mtDNA region increases).

In comparison to the Clopper-Pearson one-tailed method (currently recommended for use in U.S. laboratories [10]), LRs developed using the kappa method ranged from 8 to 14-fold higher

across our three population samples when only HV1 and HV2 were considered, and from 13 to 18-fold higher when the full CR was considered (Table 4). When the numbers of singletons across the entire mtGenome were used, LRs developed by the kappa method were 31 to 254-fold higher in comparison to the Clopper-Pearson method using a 1-tailed 95% upper confidence limit. While the most conservative haplotype frequency estimate may be preferred for some purposes, it is clear from these results that LR calculations using the Clopper-Pearson method negate some of the benefits of the increased resolution achieved by typing the complete mtGenome. Until larger full mtGenome databases are available, Clopper-Pearson based LRs developed for previously unobserved mtGenome haplotypes will be reduced in comparison to even shared haplotypes based on smaller subsets of the molecule given the size of current CR databases (for example, 2823 African American CR haplotypes are presently available in EMPOP, Release 11 [6]). That is, despite the clearly smaller likelihood of encountering a matching mtGenome haplotype versus a matching CR haplotype (for example) among randomly-selected individuals (Table 3), Clopper-Pearson LRs for full mtGenome haplotypes will, for the time being, be smaller due to database size alone.

On the basis of the EMMA [2] analyses and comparisons to Build 16 of PhyloTree [15], 393 distinct named haplogroups were assigned to the 588 haplotypes reported in this study (Appendices H-J). Across the three population samples, all major haplogroups were represented except L4, L5, L6, O, P, Q, S and Z. The frequency of each major haplogroup by population is given in Table 5, and Table 6 details the specific haplogroups present in each population at greater than 5.0%. The level of phylogenetic resolution of the haplogroups in the latter table was selected for more direct comparison to previous, CR-based mtDNA studies; however more highly resolved

haplogroup categorizations are included where the frequencies also exceed 5.0%. These data provide a snapshot of the predominant lineages found in each of the population samples.

Based on the assigned haplogroups, the 588 mtGenome haplotypes were classified into one of four broad biogeographic ancestry categories: African, East Asian, West Eurasian and Native American (Figure 16). As has been previously reported [21], self-identified ancestry was highly correlated with maternal biogeographic ancestry for the African American and U.S. Caucasian populations. For the African American dataset, the vast majority of haplotypes (90.0%) were assigned to haplogroups L0, L1, L2 and L3; whereas only 2.4%, 4.7% and 2.9% of the haplotypes represent East Asian, West Eurasian and Native American ancestry, respectively. Similarly, 94.7% of the U.S. Caucasian haplotypes in this population sample are of West Eurasian ancestry, with only minor contributions from African, East Asian and Native American lineages (0.8%, 1.9% and 2.7%, respectively). By contrast, while the majority (60.0%) of the U.S. Hispanic population sample was comprised of Native American lineages, West Eurasian and African maternal ancestries were represented in substantial proportions (25.8% and 12.3% of haplotypes, respectively).

Comparisons between the population samples reported here and previously published CR-based datasets were made on the basis of biogeographic ancestry proportions, as these can typically be ascertained for most haplotypes given CR data alone. Table 7 provides the ancestry percentages for the current study as well as for two previous studies for each of the three U.S. population groups [21,22,104-107]. For the African American and U.S. Caucasian populations, the proportion of haplotypes reflecting the predominant ancestry is not statistically significantly different between this and previous studies. However, for the U.S. Hispanic population, the differing proportions of Native American haplotypes across three population samples (this study, Saunier *et al.* [106] and

Allard *et al.* [22]) are significant (p=0.007). Specifically, the proportion of Native American haplotypes in the U.S. Hispanic population sample reported here differs significantly from that reported in the Allard *et al.* [22] study (p=0.008), even after Bonferroni correction for multiple tests. This is most likely due to differences in geographic sampling, which will reflect the substantial regional differences in the Native American component of a U.S. Hispanic population sample [7]. Along these lines, the proportion of haplotypes representing Native American maternal ancestry in a recently published Southwest Hispanic population sample from Texas (71.7%; [23]) is highly similar to the frequency of Native American haplotypes (70.8%) in the Allard *et al.* study [22].

In addition to comparisons based on inferred maternal biogeographic ancestry, we also compared the haplotype distribution for the African American population sample reported in this study to that described by Salas *et al.* [108] in their analysis of an FBI dataset [109]. When using the same haplogroup categories and level of phylogenetic resolution, the composition of our African American sample (Figure 17) is nearly identical to Figure 1 in Salas *et al. [108]*, and reflects the predominantly West African, west-central African and southwestern African origins of the mtDNA lineages present in U.S. haplotypes of recent African descent reported by the authors and in other studies [110-112].

The composition of the African American, U.S. Caucasian and U.S. Hispanic populations, and the extent of the diversity within each of the ancestry groups that contribute to them, are reflected in pairwise comparisons performed for a) each population sample and b) all samples ascribed to each of the four biogeographic ancestry categories. Figure 18 displays histograms of pairwise comparisons for both the full mtGenome and the CR only, for each of the three populations and three of the four ancestry groups, plotted by the proportion of comparisons

performed to normalize for the differing sample sizes. The average number of pairwise differences for each of these sets of comparisons are reported in Table 8. When the entire mtGenome was considered, the U.S. Caucasian population sample (Figure 18b) and the haplotypes of West Eurasian ancestry (Figure 18e) had asymmetrical bimodal pairwise distributions, with the first, smaller peak representing the comparisons between recently diverged lineages in the dataset, and the second, larger peak representing the comparisons between more distantly related haplotypes. When these same analyses were performed with the comparison restricted to the CR (Figures 18h and 18k), the distributions were unimodal and Poisson-like (though still significantly different from a Poisson distribution; $p < 0.0001$ for both). For the U.S. Hispanic dataset, Figure 2c displays an asymmetrical bimodal distribution similar to the U.S. Caucasians, but with a substantial tail to the right that represents comparisons to and between the African ancestry haplotypes present in the population sample. The Native American ancestry comparisons (Figures 18f and 18l) are sharply bimodal and more symmetrical, reflecting the origins of Native Americans and the genetic distance between the haplotypes in this sample set (primarily, haplogroups A and B from macrohaplogroup N, and haplogroups C and D from macrohaplogroup M). The comparisons between these haplotypes based on the CR alone (Figure 18l) are the only CR pairwise distribution that closely mirrors the shape of the distribution based on the full mtGenome. In contrast to the other sample sets, comparisons of both the African American population sample and the African ancestry lineages for the complete mtGenome resulted in multimodal distributions (Figure 18a and 18d) and high average pairwise numbers of differences (Table 8). In comparison to the U.S. Caucasian and U.S. Hispanic populations, fewer of the African American haplotypes are highly similar to one another across the entire mtGenome, and a much greater number are genetically very distant. Consistent with results from previous studies of African American population

samples [23,108,110-112], the distributions for these two comparisons underscore the extensive mtDNA diversity that exists within the African ancestry component of U.S. populations.

Features of the mtGenome data: Heteroplasmy

LHP in the CR has been well-characterized in a previous study [25] with a much larger sample size than we have developed here, and the observed incidence of LHP across the complete CR in our dataset is generally consistent with previous reports (see Table 9). However, a few observations from our data are worth noting. Overall, we observed LHP in hypervariable region 1 (HV1) in 17.5% of individuals. Consistent with earlier examinations [25,113,114], LHP in HV1 was observed in every sample in which a transition at position 16189 resulted in a homopolymer of nine or more cytosine residues, and no LHP was observed when seven or fewer cytosine residues were present. Among the 13 samples in which some combination of transitions and insertions in HV1 resulted in a homopolymer consisting of exactly eight cytosines, eight samples had detectible LHP. In the remaining five samples, LHP was either not present or was too minor to distinguish from sequence background/noise. The incidence of HV1 LHP across all 588 samples in this study is significantly higher ($p=0.001$) than the 5.0% recently described for a set of 101 western European individuals [26]. When our data were considered by population, though, the observed frequency of HV1 LHP varied significantly ($p<0.00001$), with a high of 25.2% in the U.S. Hispanic population, and a low of 9.1% in the U.S. Caucasian population (Table 9). This latter value is relatively consistent with the data reported by Ramos *et al.* [26]; and the differences we observed by population are largely explained by a) the nucleotide state at position 16189 (C or T), and b) the presence or absence of a homopolymer with at least eight cytosine residues, when these factors are considered by a major haplogroup (see Figure 19).

LHP in the 523-524 AC repeat region was clearly apparent (readily observed above sequence background and/or noise upon initial inspection of the raw data) in 5.3% of the samples in our dataset. The majority (65%) of instances occurred in samples with at least six dinucleotide repeats, and all 13 haplotypes with seven or more AC repeats had clear LHP. This result is consistent with a previous report on LHP in the AC repeat region, which found "pronounced" AC repeat LHP in 4.3% of samples, and generally in individuals with six or more dinucleotide repeats [25]. In addition to the LHP observed in this and the three other expected regions (in HV1 around position 16193, in HV2 around position 309, and in HV3 around position 573), a single sample exhibited one further LHP in the CR, at position 463. This haplotype has T to C transitions at positions 454, 455 and 460, resulting in a 10 bp cytosine homopolymer. Overall, across the 588 haplotypes, 374 individuals (63.6%) exhibited CR LHP, and 87 individuals (14.8%) possessed LHP in more than one portion of the CR.

LHP associated with indels in the coding region was observed in 11 instances across our three datasets (1.9% of samples), at five of the 18 coding region positions at which indels were found (Table 10). In four individuals, a T to C transition at position 961 resulted in a 10 bp polycytosine tract, and all four of these haplotypes exhibited LHP at position 965. Similarly, a T to C transition at position 8277 resulted in a 7 bp polycytosine stretch in three individuals; and in two of these, cytosine insertions (two or three) and LHP were observed. In the third individual, no additional cytosines were present, and no LHP could be detected. LHP was also observed in one sample at position 8287, due to a T to C transition at 8286 and cytosine insertions that resulted in a 12 bp cytosine homopolymer. At position 5899, no LHP was detected when only a single cytosine was inserted, but LHP was observed in the three samples with six or more C insertions. And finally,

one sample had LHP of the 8281-8289 9 bp insertion. In this individual at least two length variants were detected, and the majority molecule was two 9 bp insertions.

In addition to the LHP observed at coding region positions with indels relative to the rCRS, 88.8% of samples had detectible LHP around position 12425. Positions 12418-12425 are an 8 bp polyadenine tract, and a mixture of molecules in this region has been previously described (in a report on mtDNA heteroplasmy from NGS data [27], and in multiple cancer studies as reviewed in Lee *et al*. [115]). In our Sanger data, LHP in this region generally appeared as a mixture of two molecules consisting of seven or eight adenine residues (see Figure 20 for an example). In all cases the majority molecule matched the rCRS (eight adenines), and the LHP was generally minor enough that it did not impact sequence coverage (*i.e.*, in most cases, sequences did not need to be trimmed). Among most of the 66 individuals in which LHP at 12425 was not identified or could not be confidently called, nearly all sequences in the region had noise (*i.e.*, background) to the extent that the very low level LHP typically observed at 12425 would be obscured or difficult to detect. However, for two of the samples, a transition at position 12425 appears to have prevented LHP.

The frequency of point heteroplasmy (PHP) in the 588 haplotypes was also examined (findings are summarized in Tables 11 and 12). Across the entire mtGenome, a total of 166 PHPs, in 140 individuals (23.8%) were identified. Twenty-five samples (4.3%) exhibited more than one PHP (24 samples had two PHPs, and one had three PHPs); and of the individuals with PHP, 17.9% had multiple PHPs. The incidence of PHP across the entire mtGenome varied significantly between the three populations ($p = 0.029$). However, when pairwise comparisons of the populations were performed, only the comparison between the African American and U.S. Hispanic populations was significant after Bonferroni correction for multiple tests ($p = 0.007992$),

and the differences between populations were not significant when the CR and coding region PHPs were considered separately. In a large study of more than 5000 individuals, Irwin *et al.* [25] found significant variation in the incidence of CR PHP between multiple populations, and postulated the differences might be due to the differing mtDNA lineages comprising each of the populations. As Figure 16 demonstrates, there is certainly extreme variation in the composition of each of the three U.S. populations described here. Consistent with a recent study of heteroplasmy in complete mtGenomes [26], though, no significant differences in the frequency of PHP by haplogroup across the entire mtGenome were observed in our data, even when statistical analysis was restricted to the 11 major haplogroups with greater than five PHPs (see Table 13 for the incidence of PHP by haplogroup). Similarly, no significant differences by haplogroup were observed when PHPs in the CR and the coding region were considered separately. In the case of the present study and the results reported by Ramos *et al.* [26], it may be that the numbers of samples with PHP on a per-haplogroup basis are simply too small to detect any non-random differences.

A complete list of the mtGenome positions at which PHP was detected is given in Table 14. The 64 PHPs observed in the CR were found in 58 of the 588 individuals (9.9%), at 44 different positions. For a majority of these positions (75%), PHP was observed in just one individual. Eight positions (18%) were heteroplasmic in two individuals (one of these positions, 228, was observed as both 228R and 228K); and three positions – 189, 152 and 16093 – were heteroplasmic in four, five and six individuals, respectively. Several previous examinations of PHP in the CR have indicated that both 16093 and 152 may be hotspots for heteroplasmy [24-26,43,116]. However, to our knowledge a high observed incidence of PHP at position 189 has only been reported in muscle tissue samples associated with increased age [117,118], and in association with increased BMI and insulin resistance [119] (this excludes the data reported by He *et al.* [120], which has been shown

to be problematic [121]), though position 189 is recognized as one of the faster mutating sites in the mtGenome [8,122-125]. In our data, PHP at 189 occurred on varied haplotypic backgrounds (haplogroups L3b1a4, U5a1d1, J1c3 and H1ag1), and in two of the three populations. Visually estimated percentages of the minor molecule across the four samples with 189 PHP ranged from 5-15%. In all four cases the variant nucleotide was most clearly apparent in the reverse sequences covering the position, but was confirmed by at least one (though typically more than one) forward sequence. In three of the four cases of PHP at 189, the majority molecule matched the rCRS. No age or health-related information was available for the anonymized blood serum specimens used for the current study.

A total of 102 PHPs were observed in the coding region. Nine individuals exhibited more than one coding region PHP, and thus the total number of individuals with coding region PHP was 93 (15.8%). However, each PHP was unique in the dataset (observed in only a single individual). The absence of coding region PHPs detected in more than one individual is consistent with the recent analysis by Ramos *et al.* [26], which found 21 unique coding region PHPs among 101 individuals. Among the 24 coding region PHPs reported by Li *et al.* [27], one was shared by more than one individual; however this PHP (3492M) is unlikely to be authentic in either individual, given 1) the very low incidence of transversion-type PHPs reported by Ramos *et al.* [26] and observed in this study (see below), 2) the very low frequency of substitution at position 3492 (observed just once, and as a transition, among the more than 2000 mtGenomes analyzed by Soares *et al.* [8]), 3) the identification (by the authors themselves) of position 3492 as a sequencing error hot spot, and 4) the coverage dip observed in this region in multiple mtGenome sequencing studies ([23,85,126]; R. Just, unpublished data; and W. Parson, unpublished data) using Illumina platforms (Illumina, Inc., San Diego, CA). In a slight departure from the absence of authentic shared PHPs

in the datasets reported by Ramos *et al.* [26], Li *et al.* [27] and in this study, the haplotypes recently published by King *et al.* [23] included three shared PHPs (at positions 1438, 2083, and 8994) among the 58 total coding region PHPs detected (using an 18% threshold) in 283 individuals.

When 203 coding region PHPs (from the 1103 total mtGenomes published by Ramos *et al.* [26], Li *et al.* [27] (minus the 3492M PHPs), King *et al.* [23] and reported in this study) were considered in combination, only five additional PHPs were observed in more than one individual (see Table 15). All five of these positions had low relative substitution rates (1 to 3) among the 2196 complete mtGenome sequences previously analyzed in a phylogenetic framework by Soares *et al.* [8]. In fact, of the 102 coding region PHPs in our data, only two occurred at positions among the 15 fastest evolving sites in the coding region (and only four among the 50 fastest sites), while nearly half (44%) occurred at positions invariant among the >2000 published mtGenomes included the Soares *et al.* analysis [8] (see Table 14). In combination, these studies suggest that the distribution of heteroplasmy (which should more closely reflect mutation rates than does complete substitution) in the coding region is not consistent with the gamma-distributed relative substitution rates reported for the region [8]. This finding is in contrast to the general correlation (with a few exceptions) between heteroplasmic hotspots and mutation/substitution hotspots in the CR [25]. The seeming difference between the observed relative heteroplasmy and substitution rates on a position-by-position basis in the coding region has several possible explanations, including selection (at multiple potential levels, *e.g.*, individual, population, etc.), nucleotide state stability/mutability (that may be sequence context dependent), and genetic drift. These factors, alone and in combination, have been previously suggested to explain the difference between phylogenetic and pedigree substitution rates in the CR [127-129], departures from the correlation between observed relative substitution and heteroplasmy rates by position in the CR [24,25,116]

and patterns of substitution ([8,28,29,125], among others) and heteroplasmy [26,130] in the coding region.

In a substantial departure from the above-mentioned studies regarding heteroplasmy across the mtGenome, a very recent examination of mtDNA sequences from 1,085 individuals using high coverage depth NGS data and an ~1% heteroplasmy detection threshold found 4342 total PHPs at 2531 mtDNA positions (of 13,659 positions examined), of which only 69.42% were observed in just a single individual [131]. Relying on the same relative substitution rates published by Soares *et al.* [8] referenced above, Ye *et al.* [131] reported a positive correlation between relative substitution rates and heteroplasmy rates (R^2 = 0.3702). However, coding region heteroplasmies were not separated from CR heteroplasmies for that analysis, and an association between substitution and heteroplasmy hotspots has been previously described for the CR [25]. When we applied the same analysis to all 166 PHPs detected in our study (64 and 102 PHPs in the CR and coding region, respectively), a similar positive correlation was observed (R^2 = 0.3003, r = 0.5480; see Figure 21a) despite the clear lack of correlation between relative substitution rates and heteroplasmy rates among the coding region PHPs in this study. When the same regression analysis was performed using only the 3547 *coding region* PHPs reported by Ye *et al.* [131], a much weaker positive correlation between relative substitution rates and heteroplasmy rates was observed (R^2 = 0.1076, r = 0.3280; see Figure 21b).

Additionally, further examination of the PHPs reported by Ye *et al.* [131] indicated that some may be due to mixtures between distinct individuals/samples, rather than true intraindividual mtDNA variation [132]. For example, among the 71 PHPs reported for sample HG00740, nearly all of the positions are diagnostic for two distinct mtDNA haplogroups (L1b1a1a and B2b3a; according to Build 16 of PhyloTree [15]). Similar issues were observed among the PHPs described

in another recent report on human mtGenome heteroplasmy [133]. In that paper, nearly all of the 20 PHPs given for sample NA12248 (for example) can be ascribed to one of two haplogroups (U5b2a2b or H1e), and few PHPs that would be expected from a mixture of two samples representing those haplogroups are absent. These findings cast some doubt on the veracity of the incidence and pattern of heteroplasmy reported in the Ye *et al.* [131] and Sosa *et al.* [133] studies, and thus the conclusions those authors have drawn from the data.

Among the PHPs observed in the CR in our study, all but two (97%) were transition-type (purine to purine, or pyrimidine to pyrimidine) PHPs; and of these, approximately two-thirds were pyrimidine transitions while one-third were purine transitions (Tables 12 and 14). The 1.6:1 pyrimidine to purine ratio for PHPs in the CR is consistent both with earlier analyses of CR heteroplasmy [25,134] and with the approximately 1.3:1 pyrimidine to purine ratio in the nucleotide composition for the region. Only one of the 102 PHPs in the coding region was a transversion-type change, indicating an even more extreme bias toward transition-type heteroplasmies than has been previously reported [26,130]. And in contrast to the CR, more of the coding region PHPs were purine (59%) versus pyrimidine (41%) transitions, despite a pyrimidine to purine ratio (in terms of average overall nucleotide composition for the coding region) that is nearly identical to the CR. The same phenomenon has been observed in previous studies of both substitution and heteroplasmy in the coding region [26,30].

Figure 22 displays the proportion of PHPs observed by mtGenome region in our data; and Figure 23 details both the proportion of positions within each coding region gene at which PHP was observed, and the portion of that variation that would lead to synonymous and nonsynonymous changes to the amino acid if the observed mutations were fixed. In our data, the highest rate of PHP was observed in *ATP8* (four PHPs observed across 207 total positions). The lowest rate of

PHP was seen in *ND3*, with heteroplasmy observed at one of 346 possible positions, followed closely by 12S rRNA. Consistent with previous reports on coding region substitutions [28,30], the highest rate of nonsynonymous variation in our heteroplasmy data was observed in *ATP6*, where six of seven PHPs would result in amino acid changes if the mutations were to become fixed. This 1:0.17 nonsynonymous to synonymous ratio exceeds the gene with the next highest ratio (*CYTB*, 1:0.6) more than 3-fold. However, *ATP8*, with the highest overall rate of PHP in this study, and previously reported to have a high rate of nonsynonymous *substitution* [30], had one of the lowest nonsynonymous to synonymous heteroplasmy ratios at 1:3. With regard to codon position, 87% of the 76 PHPs in protein-coding genes were observed in first or third positions, whereas only ten were observed in the second codon position (see Table 14). However, all first codon position PHPs we detected were nonsynonymous changes. Approximately twice as many PHPs occurred in third versus first codon positions, and the first to second to third position ratio for PHPs was 2.2:1:4.5.

Overall, the nonsynonymous to synonymous change ratio for the 76 PHPs detected in protein-coding genes in our study was 1:1.4, a value that is in close agreement with a recent report on coding region heteroplasmy [26]. Our ratio is both closer to a neutral model of sequence evolution and significantly different from some previous examinations of patterns of coding region substitution in protein coding genes (1:2.32 from Elson *et al.* [28], p=0.035; and 1:2.5 from Kivisild *et al.* [29], p=0.013), but is not significantly different from the overall ratio determined from an evaluation of >5000 published mtGenomes by Pereira *et al.* (1:1.97, [30]). However, the ratio from our data *was* significantly different from the nonsynonymous to synonymous ratio those authors reported for the substitutions with frequencies at 0.1% or greater in the dataset (1:2.69, p=0.006).

In addition to calculations of overall nonsynonymous to synonymous change ratios, examinations of protein-coding gene substitutions in previous studies have also found 1) a higher proportion of nonsynonymous variation, and 2) higher pathogenicity scores for nonsynonymous substitutions, in younger versus older branches in the human mtDNA phylogeny and other species ([8,28,29,31,135], among multiple others), both of which provide further evidence that selection is acting to remove deleterious mutations from the mtGenome over time. When we compared the average pathogenicity scores (based on MutPred values [136] reported by Pereira *et al*. in their tables S1 and S3 [31]) for a) all possible nonsynonymous substitutions across the mtGenome, b) the 60 nonsynonymous PHPs detected in our haplotypes and reported in three recent studies [23,26,27], and c) the nonsynonymous substitutions evaluated by Pereira *et al*. [31] for mtDNA haplogroup L, M and N trees, the results again indicated that heteroplasmic changes appear closer to a neutral model of sequence evolution than do complete substitutions (Figure 24). While the difference between the average pathogenicity scores for heteroplasmies versus all possible substitutions was statistically significant ($p=0.01$), the average pathogenicity score for the PHPs was also significantly higher ($p=0.0001$) than the average for the haplogroup L, M and N substitutions with rho values of zero (that is, the mutations observed at the tips of the trees) reported by Pereira *et al*. In other words, the heteroplasmic variants in our study have greater potential for deleterious effect than the most recently acquired complete substitutions in the haplogroup L, M and N lineages analyzed by the authors. Given the relative evolutionary timescales for heteroplasmy versus the fixation of new mutations, these comparisons between heteroplasmic changes and complete substitutions in protein-coding genes across both close and distant human mtDNA lineages thus also appear to provide some further support for the role of purifying selection in the evolution of the mtDNA coding region.

EMPOP database modifications

Modifications to the EMPOP database to accommodate entire mtGenome data and enable its use by register users included updates to both data handling and data search capabilities. On the back end, the database infrastructure was updated to 1) permit the storage of full mtGenome data, 2) allow for traceability of full mtGenome data and changes in the database, 3) enable mtGenome data import and export (by users with administrative privileges only) in several different formats (*e.g.*, SeqScape, Sequencher, EMPOP, FASTA and binary matrix to permit statistical analyses), and 4) enable string search functionality [32] of entire mtGenome query sequences against entire mtGenome reference datasets. The mtGenome-specific web interface features that were developed and will be available to registered users include 1) search features that permit the input of full mtGenome data as either differences from the rCRS or text strings, and 2) search parameters that permit users to select the data range to be searched, the inclusion or exclusion of specific insertions and/or deletions, and different options for handling heteroplasmic positions.

In addition to the above modifications made specifically to permit the storage and query of full mtGenome data, extensive additional functionality was added to both the back end of the database and the EMPOP web interface that relates to examinations of both partial and complete mtGenome haplotypes. A new scheme for geographic/metapopulation characterization (that includes linguistic information) of the stored datasets was developed to sort database matches according to forensically-relevant criteria. Queried datasets are visually represented on a global map (Figure 25); and a tabular summary of haplotype frequency estimates for database matches are displayed in an interactive format that includes sortable columns and expandable continental origin and metapopulation categories (Figure 26). The matching haplotypes can also be examined in detail, in a further results table that includes the haplogroup affiliation for the matches to provide

phylogenetic background information that may be useful in the forensic context (Figure 27). From this table, clickable links imbedded in the haplogroup classification open a new browser window that displays a heat map of the geographic distribution of the haplogroup among the mtDNA haplotypes stored in the database (Figure 28). In a further table, neighbors to haplotype matches in the database can also be examined (Figure 29). Neighboring haplotypes are displayed by sequence distance and costs (based on the fluctuation rates developed as described in the Materials and Methods section and Röck et al. [2]), and can be sorted by the various columns.

The maximum likelihood based software program, EMMA [2], was also incorporated in EMPOP3 to permit automated haplogroup estimates for mtDNA sequences. When queries of the database do not result in any matches, the details displayed for the searched haplotype include the two highest ranking (lowest cost) haplogroup estimates, as well as details on the individual profiles and virtual haplotypes in the database that fall within those rankings (Figure 30).

Finally, the new tools in EMPOP3 additionally include a dynamic haplogroup browser which enables users to surf the mtDNA phylogenetic tree (updated based on the current build of PhyloTree [15]) and search the database mtGenome sequences for specific mutations and sets of mutations (Figure 31). Clicking on a haplogroup link within the tree results in a heat map that displays the geographic distribution of the haplogroup, as well as the number of sequences belonging to that haplogroup that are contained in the EMPOP database. When particular mutations are searched, the haplogroups for which the motifs include those mutations are displayed in a table.

A beta version of EMPOP3 that included all of these features was sent to a number of mtDNA practitioners and current EMPOP users in the U.S. and abroad in early 2014. EMPOP3 is

presently being finalized based on the beta tester feedback, and we expect to publicly release the new version of the database in the second quarter of 2015 [98].

IV. Conclusions

Discussion of findings

Though the mtGenome amplification and sequencing protocol we developed for this project can be performed manually (with, we must emphasize, abundant attention paid at pipetting steps to prevent sample misplacement), the strategy was specifically designed to be implemented on liquid handling instruments to facilitate high-throughput data generation. At AFDIL, all pre-PCR pipetting steps (including sample placement, extraction and PCR reaction set-up) are performed in 96-well plate format on a bench top liquid handling robot; amplification product detection is performed directly from the 96-well plate on a capillary electrophoresis instrument; and, with the exception of the addition of enzymes for post-PCR purification (which, due to high viscosity, are pipetted manually into the sample plate to reduce reagent waste and cost), all post-PCR pipetting steps are performed robotically.

Based on the observation of some PCR failures with positive control DNA and blood serum extracts when DNA concentrations were low, and given the extent of sample reprocessing necessary at various PCR input DNA quantities with the blood serum specimens, we suggest an input DNA concentration for PCR of 50 pg or greater when possible for this Sanger sequencing protocol. Further, due to the increased noise (a result of excessive electrophoretic signal) observed in some sequences during the development of the protocol when DNA inputs for PCR were high, we recommend that highly concentrated sample extracts be diluted so that PCR input does not exceed 1 ng. Though high quality Sanger data has been developed from higher and substantially

lower DNA inputs using this protocol, inputs between 50 pg and 1 ng should ensure consistent amplification success and the production of high quality Sanger sequence data across all amplicons in the first pass of sample processing. Following these DNA input guidelines will accordingly reduce the opportunities for human error inherent in manual sample reprocessing and minimize the cost to generate each mtGenome haplotype.

Regarding the potential for amplification of portions of the human nuclear genome, it seems highly improbable that a NUMT sequence would be represented in any complete mtGenome haplotypes developed using this protocol. Amplification of a nuclear genome sequence alone (in place of the target mtDNA) is extremely unlikely given the abundance of mtDNA relative to nuclear DNA in human cells, and could reasonably only be expected to occur if mtDNA was nearly or completely absent in a DNA extract [95]. In the unlikely case that a NUMT was amplified in place of the mtDNA target, any close inspection of the data (which would reveal an excess number of differences from the rCRS; unusual insertions, deletions, and/or transversions; etc.) or attempt to assign a haplogroup to the mtGenome profile would readily indicate a problem.

A more likely scenario with an overall low DNA quantity sample is co-amplification of a NUMT with the authentic mtDNA, which could occur when by chance the mtDNA primers encounter a close-match nuclear DNA target during the early cycles of PCR. While we did not encounter this during protocol development, a few samples did exhibit evidence of NUMT co-amplification during our initial processing of the blood serum specimens. In four of the five samples, the NUMT contamination presented as a clear mixture in the sequence data for the affected amplicons, as the high number of positions at which two bases were observed could not reasonably be explained by mtDNA heteroplasmy. In the fifth case, the NUMT co-amplification was only detected upon close scrutiny of a position identified as a possible very low-level PHP.

This latter case highlighted the need to review all PHPs detected in full mtGenome haplotypes to ensure authentic heteroplasmy – rather than NUMT contamination – is reported. In all five instances of NUMT co-amplification that we detected, however, repeat PCR and complete resequencing of the affected amplicons following re-extraction produced sequence data with no evidence of NUMT contamination.

In addition to a robust laboratory protocol and, preferably, automated rather than manual sample processing, a well-considered data analysis workflow that includes proper procedures for data interpretation and handling is essential to the generation of high quality, error-free mtDNA data for forensic genetic or other purposes. For the development of complete mtGenome haplotypes we recommend adoption of the best practice alignment, nomenclature and reporting guidelines outlined for the production of mtDNA CR data for forensics [10,44,137]. We also recommend review of the raw electropherogram data by *at least* two scientists and fully electronic data transfer, as described in Irwin *et al.* and Salas *et al.* [7,138]. Further, with the use of a multi-amplicon protocol such as the one we developed, and especially if any manual processing must be performed, we suggest additional post-data production checks to confirm that each complete mtGenome haplotype represents data from a single sample. All of these steps were performed in our review of the mtGenome haplotypes we report here.

The high-throughput amplification and sequencing strategy we developed regularly produces redundant sequence coverage across the entire mtGenome in the first pass of automated data generation. The described workflow, especially when implemented on robotic instrumentation, reduces both the cost of mtGenome sequencing and the opportunities for human error by decreasing the extent of manual sample processing/reprocessing required. As the amplification and sequencing primers were carefully selected based on highly conserved regions

of the mtGenome, the protocol works equally well on samples originating from diverse mtDNA haplogroups. Overall, this protocol should facilitate more rapid production of the complete mtGenome population reference data needed for future forensic applications, and, when combined with the adoption of best-practice data review and interpretation strategies, ensure that the data sets are of the highest quality possible.

The minor modifications we made to the protocol for its application to the blood serum specimens, combined with the sample handling strategy applied to the production of our mtGenome haplotypes, reliably produced high quality data from very low DNA quantity specimens in the first pass of automated data generation, and most samples did not require any manual reprocessing to generate complete mtGenome haplotypes. Amplification was successful 99.4% of the time when DNA inputs were greater than 10 pg, and no PCR failures were observed at inputs greater than 50 pg. Sequencing success – assessed both in terms of sequencing failure (determined by electrophoretic signal) and the amount of reprocessing required to generate a complete haplotype – was variable but generally still high when PCR DNA input quantities were less than 50 pg. At PCR inputs exceeding 50 pg, an average of 0.82% of sequencing reactions failed and only one manual sequencing reaction was required for every three haplotypes. At QIAxcel-measured PCR product concentrations less than 2 ng/µl/1000 bp, more than 90% of the sequencing failures were observed in two target regions (Amplicons 4 and 6).

The amplification and sequencing success rates reported here demonstrate that it is feasible to generate forensic-quality complete mtGenome haplotypes in a routine casework environment from forensic-like (low template) specimens. The development of this large, thoroughly evaluated data set from blood serum samples provides clear evidence that amplicons exceeding 2,000 bp can regularly be recovered from very low DNA quantity specimens; and the data also provide detailed

information on both PCR and Sanger sequencing success rates across a range of qPCR-measured mtDNA quantities. The processing metrics detailed here may thus be useful to forensic practitioners when attempting to determine the specific mtDNA amplicons, assays or markers to pursue when DNA quantities are known and case sample extract volumes are limited. Additionally, the data provide an indication of the first-pass amplification success rates that could be expected with low DNA quantity specimens in a high-throughput environment if the PCR strategy was applied as an enrichment method for targeted NGS of mtDNA.

The 588 complete mtGenome haplotypes that we report here were developed according to current best-practice guidelines in forensics for the generation and review of mtDNA population reference data [10,11]. The use of a robust amplification and sequencing strategy, primarily robotic sample handling, electronic data transfer, adherence to phylogenetic alignment rules [10,11,14] with reference to the current mtDNA phylogeny [15], repeated reviews of the raw data, and the inclusion of multiple quality control measures ensure that these haplotypes meet the highest data quality standards and are appropriate for forensic use. In terms of data review, though two laboratories highly accustomed to examining mtDNA sequence data were involved in this databasing effort (AFDIL and EMPOP), a small number of haplotype discrepancies (most regarding heteroplasmies) were encountered when the raw data reviews were compared. In addition, two alignments that did not adhere to the mtDNA phylogeny and were overlooked by *both* laboratories were later found upon screening all >2000 indels in the 588 haplotypes. While typically very easily resolved by re-review of the raw data, these discrepancies and misalignments (all fully corrected in the final haplotypes reported here) once again highlight the importance of incorporating multiple levels of quality control in the review of mtDNA population reference data generated for forensic purposes. However, we note that the efficacy of automated processing

combined with a rigorous review strategy in preventing errors with our multi-amplicon protocol was evident from the absence of problems detected at the stage of phylogenetic data evaluation. That is, despite performing more than 5,000 PCR reactions and nearly 80,000 Sanger sequencing reactions to develop the mtGenome haplotypes, no instances of sample swaps (*i.e.*, "artificial recombination" [34,137]) were identified.

The biogeographic ancestry proportions inferred from the full mtGenome haplotypes are consistent with previously-published mtDNA CR datasets for the same three U.S. populations, thus demonstrating that the population samples reported here are as representative as the reference population data on which current haplotype frequency estimates rely. The single exception was the Native American ancestry component of the U.S. Hispanic population sample, which differed significantly between this and one previous study [22]. This is likely explained by geographic sampling differences between the earlier study and the U.S.-wide population sample we report here.

On average, full mtGenome sequencing increased the proportion of unique haplotypes in each population sample by 19.3% over what would have been achieved with CR sequencing, by 32.5% over HV1/HV2 sequencing, and by 68.3% over HV1 sequencing. Though these resolution improvements and the overall paucity of shared mtGenome haplotypes in each population sample (in both this and another recent study [23]) clearly reveal the discriminatory power of complete mtGenome typing among randomly-sampled individuals, the development of LRs using the currently-recommended [10] Clopper-Pearson method for 95% confidence interval calculations [19] will largely negate this advantage (in terms of describing the statistical weight of a match for a novel haplotype) until full mtGenome databases are substantially larger. Because of this, and the anticipated movement from CR-only sequencing to typing greater portions of the mtGenome in

forensic practice, the question of how best to capture and convey this additional discriminatory information arises. For the specific scenarios presented here, there would seem to be some benefit in statistical approaches that take into account both database size and database *composition*.

As the haplotypes reported here are based on high quality Sanger sequence data with minimal noise, these 588 profiles permit the most extensive insight to date into the heteroplasmy observed across a large set of randomly-sampled, population based complete mtDNAs developed to forensic standards. The incidence of PHP across the entire mtGenome that we detected - 23.8% of individuals - is strikingly similar to the PHP frequency described in two previous analyses [26,27]. This PHP rate is substantially lower than the incidence of heteroplasmy reported in recent NGS studies using bioinformatics methods (and in one case, a detection threshold close to 1%) [131,133]; yet those higher heteroplasmy rates are questionable due to errors detected in at least some of the data [132]. A far greater proportion of individuals exhibited LHP in our study than has been previously reported [26], in largest part due to 1) the LHP we detected in the 12418-12425 adenine homopolymer, and 2) the differences between the populations examined. When PHP and LHP are considered in combination, nearly all individuals (96.4%) in this study were heteroplasmic. Though our data – even when considered in combination with previous studies - provide only a preliminary look at coding region heteroplasmy (versus the extent of information now available on mtDNA CR heteroplasmy), comparisons between coding region heteroplasmy and substitution patterns seem to provide additional support for selection as a mechanism of human mtGenome evolution.

The complete mtGenome databases representing the African American, U.S. Caucasian and U.S. Hispanic populations that we have developed will be publicly available from multiple sources. They are included in this report; are available for download from GenBank; are available

in a primary publication [5]; and will be available for query using forensic tools and parameters in EMPOP3 [98]). These extensively vetted and thoroughly examined Sanger-based population reference data provide not only a solid foundation for the generation of haplotype frequency estimates, but can also serve as a benchmark for the evaluation of future mtGenome data developed for forensic purposes. This includes comparative examination of the features (*e.g.*, variable positions, indels, and heteroplasmy) of not only datasets developed as additional population reference data, but also single mtGenome haplotypes – especially those generated using NGS technologies and protocols new to forensics – from casework specimens.

The modifications to EMPOP described here will be available in the upcoming version of the database, EMPOP3, with expected release in the second quarter of 2015 [98]. Beyond the addition of mtGenome-specific functionality that will permit both the storage of mtGenome data as well as queries of the data by registered users, the additional features that will be available in EMPOP3 will enhance the user experience and the utility of the database for forensic mtDNA investigations. To enable automated haplogrouping of mtDNA sequences (including both partial and complete mtGenomes), EMMA [2], which estimates mitochondrial DNA haplogroups using a maximum likelihood approach, has been implemented in EMPOP3. Haplogroup estimates using this software rely on both a comprehensive and curated database of 14,990 full mtGenomes and 3925 virtual haplotypes from PhyloTree [15], as well as a database of nearly 20,000 manually haplogrouped CR haplotypes that were used to determine the fluctuation rate of mtDNA mutations (and are incorporated in EMMA and EMPOP3 to develop cost values for differences between the queried and backbone/database mtDNA profiles). Thus, EMMA-based assignments will be more accurate than any existing publicly-accessible haplogrouping tools, all of which rely solely on virtual haplotypes. The newly developed characterization scheme for geographic/metapopulation

classification of the database haplotypes permits the evaluation of query results according to forensic-specific criteria, and the dynamic pages and visual/geographic displays of the queried populations and match results details also greatly improve the functionality of the database.

To date, the EMPOP database contains nearly 35,000 profiles from more than 60 global populations. The new features developed for EMPOP3, the inclusion of forensic-quality full mtGenome population data for the first time in *any* publicly-accessible forensic mtDNA database, and the continuing high quality standards for dataset acceptance by EMPOP will further solidify EMPOP as an exceptionally useful and reliable mtDNA population reference database for forensic purposes.

Implications for policy and practice

As DNA typing methods are increasingly employed on evermore small, damaged and degraded specimens, and as DNA based missing persons, mass fatality, criminalistics and intelligence applications expand, more and more forensic laboratories will turn to mtDNA testing for its obvious benefits. Historically, the forensic DNA community has understandably favored the use of multiple short tandem repeats because they provide much higher discriminatory power than data from the mtGenome. However, there remain many specimens for which nuclear DNA is too degraded or too scarce to yield forensically informative STR profiles. In addition, there are numerous cases involving questions of extended, or complex, kinship that cannot be resolved with autosomal STR data alone. It is in these scenarios that mtDNA testing will be required and that data outside the control region will be important to case resolution.

In our original proposal, we suggested that there were three critical points to bear in mind for mtDNA based applications to criminal, missing persons, mass fatality and intelligence casework:

1. Forensic genetic applications that require mtDNA testing are growing rapidly due to expanding missing persons and disaster victim identification efforts, the benefits of lineage marker data for the resolution of complex kinship scenarios, and increasing interest in the recovery of DNA from small/damaged evidentiary material
2. Forensically-useful assays that target entire mitochondrial DNA data have already been developed; and new technologies permitting access to mtGenome data from the most damaged specimens are emerging rapidly
3. *Reference population data which adhere to strict forensic standards and are necessary for proper forensic interpretation of mtGenome evidence do not yet exist.*

As predicted, the simultaneous yet somewhat independent initiatives of #1 and #2 have continued to converge over the past four years. A handful of very recent studies have demonstrated that mtGenome data can now be developed for forensic purposes via NGS technologies [23,80,85,86], and entire mtGenomes have been recovered from some of the most difficult forensic samples [36,73,74] These developments have made point #3 an increasingly obvious practical limitation to the implementation and utility of complete mtGenome data in forensics. Fortunately, as a result of the current NIJ funded project, reliable and representative reference mtGenome population data that did not exist three years ago are finally available to the forensic practitioner community. In addition, the database infrastructure and algorithms required to handle and search these data will soon be available at empop.org [6].

Though additional reference population databasing efforts will be required over time to increase mtGenome dataset sizes and refine haplotype frequency estimates, the sequences generated here are a critical start. Not only do they represent individuals from across the United States, but since the sequences themselves have been developed via well-established Sanger

technology and with strict data generation and review protocols, we fully expect these data to serve as the quality benchmark against which new mtGenome reference population datasets will be measured. New datasets will very likely be developed with NGS technologies for which routine laboratory workflows, data analysis and interpretation guidelines, and data quality checks are still not well-defined. Moving forward, it would behoove the forensic community to address these outstanding issues related to NGS carefully in order to avoid a repeat of the past, when shortcomings of Sanger-based reference mtDNA *control region* databases became the focus of both a high profile scientific controversy and a courtroom challenge to the admissibility of mtDNA evidence [33-35,139]. The care taken and safeguards employed in the generation of this and other recent Sanger-based CR datasets ([140-145] for example) were instituted specifically to address and rectify issues identified in earlier datasets. Given these lessons learned, the overall process described here should serve as a model for future NGS data production, and the data themselves should serve as a critical frame of reference for the development of robust and reliable NGS data analysis and interpretation guidelines.

With regard to near-term direct implications to U.S. policy and practice, the following can be said of this project:

1. High-quality mtGenome data developed to forensic standards are now available
    a. For assessment of mtGenome profile rarity
    b. To serve as a model workflow for the generation of new mtGenome data/datasets
    c. To serve as a data quality benchmark for the development of formal analysis and interpretation guidelines for NGS-derived mtGenome data

2. These haplotypes, and the database infrastructure to support searches and reporting of mtGenome data, will soon be available in EMPOP3 [98]. The EMPOP database is supported by both SWGDAM and the ISFG DNA Commission [10,11]

In addition to providing the necessary infrastructure to perform near-term mtGenome sequencing in casework, the 588 haplotypes developed here are providing, for the first time, an empirical understanding of the true value of mtGenome data for practical mtDNA casework purposes. While it has been well-understood that by extending mtDNA analysis from the HV regions to the entire CR, haplotype resolution increases substantially (from 64-76%, depending on the population, to 74-80%, using data from the current study), it is only now, and as a result of this project, understood that by extending analysis to the entire mtGenome nearly full resolution of mtDNA haplotypes can be achieved. To put this in practical terms, we can use the estimated 40,000 unidentified human remains stored in medical examiner facilities across the United States as an example [146]. With HV1/HV2 data alone, we would expect to encounter a common mtDNA type in approximately 20% of these cases (based on U.S. Caucasians). Now, with complete mtGenome haplotypes, and given what we have learned as a result of this project, we can expect to encounter a shared haplotype in approximately 1% of these cases (based on the U.S. Caucasian data presented here). This same improvement in resolution will, of course, hold true in other forensic applications as well – *e.g.*, criminalistics, mass disaster, intelligence, etc.

Perhaps even more important than the near-term benefits to forensic casework that the project deliverables may bring, this effort lays critical groundwork for long-term changes in the way forensic cases are handled. As stated in our original proposal, the overarching objective of this project was to establish – in a proactive manner - the long-term data foundation required for

expanded mtDNA based forensic capabilities. Given the technological advancements in mtDNA sequencing that are introduced by NGS (which include cost-effective sequencing for even small forensic labs, the straightforward recovery of entire mtGenomes, the resolution of mtDNA mixtures, and the retrieval of mtDNA from "Low Template" or "touch" DNA samples), it is very possible that mtDNA data will be sought in most or all future cases – even when STR profiles may be recoverable.

First, MPS removes one of the last practical barriers to routine mtDNA data production in a forensic setting. MtDNA typing has historically required the application of laboratory processes, and data analysis tools and strategies, entirely distinct from what is used for nuclear DNA typing. With MPS, mtDNA and nuclear DNA typing will entail highly similar workflows, thus virtually eliminating the logistical differences in data production for the marker types. Because of this, commercial assays and data analysis packages that simultaneously recover data from multiple marker systems will soon be a reality. Kits which combine nuclear STR markers and forensically-relevant single nucleotide polymorphisms are already in production [147,148], and a joint nuclear DNA and mtDNA MPS workflow aiming to maximize the information gleaned with a "forensic genomic" approach is also advertised [149]. In addition, a combined nuclear and mtDNA kit will soon be commercially available [150].

Though data from multiple marker systems will not be needed in all situations (direct comparisons between evidence when full STR profiles are recovered, for instance), some common scenarios are likely to benefit greatly from concurrent development of nuclear and mtDNA data. Cases involving touch or low DNA template samples – those that have traditionally been addressed with a "Low Template" or "Low Copy Number" approach to autosomal STR typing – will clearly benefit from the addition of mtGenome information. Similarly, in those criminal, intelligence and

missing persons investigations where kinship and relatedness are in question, mtDNA data will continue to offer significant additional resolution. Further, the availability of data from multiple markers should assist in determining the number of persons represented in mixed contributor samples, with mtDNA information potentially distinguishing individuals unresolvable by limited STR data alone.

On a larger and longer-term scale, mtDNA data (and complete mtGenome data to be specific) recovered from more samples should lead to more crime scene profile comparisons in local and national DNA databases. At present, most local and national STR databases have minimum marker or allele number requirements for both uploading profiles and searching profiles. The addition of full mtGenome data to otherwise incomplete STR profiles will increase the number of profiles that can both uploaded and searched in these databases. More stored crime scene profiles will lead to more crime-offender, crime-crime and missing persons comparisons which, in turn, should lead to more solved crimes and more identified human remains.

The final, and perhaps greatest, barrier to the near-term implementation of full mtGenome sequencing in forensic practice concerns the medico-legal-ethics implications of retrieving genetic data with potential medical significance. The community has touched briefly on this topic in the past with regard to typing greater portions of the mtGenome [1,52,151,152], but it is clear that a more in-depth examination of the pertinent issues will be needed in the very near future. While a small number of mtDNA coding region mutations have long been confirmed as disease-associated (for instance, with Leber Hereditary Optic Neuropathy [153,154] and Leigh disease [155]), many others have been suspected (see the extensive list at http://www.mitomap.org [156]). And indeed, it is possible that the mtDNA control region sequence data routinely developed for forensic casework purposes may eventually be shown to have some medical relevance, given the

astronomical increase in the amount of genetic data now available in biomedical databases and the resulting capability to examine linkage disequilibrium and disease correlation to a greater degree than ever before. Yet, while the incidental recovery of information with medical significance in the course of identity testing is clearly not desirable, a cost-benefit analysis may well determine that the greater good achieved by complete mtGenome sequencing is worth the effort that will be required to establish new policies and systems for protecting privacy in forensic mtDNA practice. Genetic privacy and several related socio-ethical issues, such as informed consent, the retention and use of specimens, the safeguarding of genetic data and data sharing, familial relationships, and incidental findings in genetic testing are increasingly matters of examination (by both forensic practitioners and non-scientists) in relation to the societal benefits of DNA-based human identification efforts and crime resolution [157-168]. In the broader genetics testing and research communities some very similar questions regarding the handling of genetic information have become a subject of intense focus in the "genomic era" ([169-176], for example), and it is possible that some clear best-practice models may emerge from those discussions that could be adopted – or, more likely, adapted - for forensic practice.

With nearly all of the foundational pieces now in place to support full mtGenome typing in forensics, it is our hope that the larger community of forensic mtDNA practitioners and researchers both in the U.S. and abroad will begin to address these critical medico-legal-ethical questions, and work toward developing 1) a consensus regarding full mtGenome sequencing on known specimens, and 2) guidelines for the use and storage of full mtGenome data in forensic casework, that can eventually be adopted as policy for U.S.-based DNA testing laboratories. In recognition of the need for further conversation concerning the use of mtGenome coding region data in human identification cases, SWGDAM recently convened a group of experts (internal and

external to the forensic community) to discuss the responsible handling and use of potentially sensitive data, the advancement of public understanding regarding risks/benefits of mtGenome typing, and approaches to fostering public dialogue and trust [177]. If these and other bioethical/legal considerations can be effectively addressed, and if cost, throughput, ease of recovery, sample consumption, and ease of analysis are no longer restrictions to the development of complete mtDNA data with NGS, there would seem to be no downside to mtDNA typing in most or all future forensic cases – only benefit.

Implications for further research

Future work on the mtGenome haplotypes reported here should include an evaluation of the data in a phylogenetic framework to assess position-specific mtDNA substitution rates for the coding region. The relative substitution rates for positions across the full mtGenome developed by Soares *et al.* [8] are an extremely useful resource and have been widely cited in mtDNA studies since their publication (and were utilized in our study for both the development of EMMA [2] and as a point of comparison for patterns of coding region heteroplasmy). However, though these rates were determined from a large set of mtGenomes (more than 2000), some of the published datasets that were utilized may contain errors [87]; and such errors – especially if systematic within particular datasets – may impact the relative substitution rates developed from analyses of the data. Given the extremely high standards to which the haplotypes reported in this work were developed, it would be worthwhile to use these data to estimate relative substitution rates, and to compare the rates to those published by Soares *et al.* [8]. Refinement of mtDNA substitution rates will provide a more thorough understanding of patterns of mutation in the coding region, which will in turn assist in the interpretation of mixed contributor samples, heteroplasmy and DNA damage as they are encountered in forensic casework.

In addition to clarifying coding region substitution rates, a phylogenetic analysis of the haplotypes (perhaps in combination with other published full mtGenome datasets that are likely to be free of errors) may result in greater resolution of portions of the human mtDNA phylogeny. As Appendices H-J demonstrate, while the haplogroup assignments for the majority of haplotypes are highly refined, some assigned haplogroups (*e.g.*, A2) are rather coarse by comparison. Those assignments, developed using the automated, maximum likelihood-based program developed by EMPOP that relies on both published mtGenomes and Build 16 of PhyloTree [2,15], generally reflect the current knowledge of the human mtDNA phylogeny, which is based on most of the complete mtGenome sequences published (or publicly available) to date (more than 20,000 as of 19 Feb 2014; www.phylotree.org [15]). Thus analyses of haplotypes representing less highly resolved branches of the tree may result in further insight into the molecular evolution of those lineages and the population histories they represent.

Essential to the validation of mtDNA population reference datasets for forensic use are quality control checks of the data [11]. For mtDNA CR datasets, two leading forensic journals (Forensic Science International: Genetics and the International Journal of Legal Medicine) require quality evaluation of the data (performed by EMPOP) prior to manuscript acceptance [11,49,178]. EMPOP performs these examinations using a variety of applications developed in-house or modified for forensic-specific and automated use [11]. Yet, the majority of these are specific for CR (or partial CR) data and cannot yet be applied to complete mtGenome datasets. It is in part for this reason that the haplotypes developed in this study were repeatedly and thoroughly examined; no tools yet exist that would permit an automated evaluation of the data quality. Thus, modification of the tools currently employed by EMPOP, or the development of new applications, will be required to adequately assess the quality of future full mtGenome datasets in a more automated

fashion (especially given the expected increase in the volume of mtGenome data produced as forensic reference data). To this end, EMPOP aims to utilize the high quality and thoroughly vetted mtGenome haplotypes we have produced as a reference dataset to develop both a screening method for indel placement according to phylogenetic alignment rules [10,11,14], as well as population-specific filters for quasi median network analyses [6,11,179,180] of full mtGenome data. Presumably, these filters and a network analysis application could be available in the future on the EMPOP website to enable researchers to evaluate their own datasets, as is presently the case for CR data with the current version of EMPOP [11,180].

Looking ahead, it is evident that larger mtGenome databases representing additional population groups (both U.S. and global) will be needed in forensics. As the examination of LRs in this study demonstrates, use of the presently-proscribed [10] Clopper-Pearson "exact" method for 95% confidence interval calculations [19] will result in LRs for novel full mtGenome haplotypes that are smaller than CR haplotype LRs due entirely to the disparity in database sizes, despite the distinctly lower random match probability when the complete mtGenome is typed. And clearly, the databases reported in this study will not be appropriate for the development of haplotype frequency estimates when questioned specimens are believed to originate from non-U.S. population groups [11].

As many of these mtGenome databases will likely be developed using NGS protocols and technologies (given the cost and throughput advantages in comparison to Sanger sequencing [23]), an obvious area for continued research is in the development and validation of both laboratory protocols and data analysis pipelines that produce mtDNA profiles that meet or exceed current standards for Sanger data. Here, the seemingly best approach is generation of NGS-based haplotypes from the same specimens for which high quality Sanger data are available (or can be

developed concurrently) so that direct comparisons between the data types can be performed. Such work is already underway in multiple forensic laboratories ([80,126,126,181], among others), and will likely continue at least until (if not beyond) the time when recognized authoritative bodies for forensic genetics (*e.g.*, SWGDAM, ISFG [10,11]) publish guidelines for the use of NGS technologies for mtDNA data generation for both reference population samples and casework specimens.

In line with these efforts, a question that naturally arises from the analyses of mtGenome heteroplasmy rates presented both here and in our recent analysis of a published study [132] is: at what minor variant frequency can authentic mtDNA heteroplasmy be *reliably* detected in NGS data? In addition to some notably problematic reports on mtDNA heteroplasmy from outside the field of forensics [120,131], a few forensic studies have also described detection of variant nucleotides at very low frequencies when data resulting from artificially mixed reference-quality samples have been examined (<5%; [76,85]). However, no published study has yet convincingly demonstrated consistent and reliable heteroplasmy or mixture detection - *accurately distinguished from any errors or other artifacts present* – at these frequency levels from NGS data, and authentic casework material has not yet been examined in this regard. Given the complexities associated with the forensic specimens to which mtDNA typing is often applied, it seems possible that further research may demonstrate that thresholds for accurate heteroplasmy detection will never achieve the sub 5% levels across the molecule that have been promised in some studies (due both to template limitations and the extreme sensitivity of NGS methods to contamination); and thus that the application of NGS data in forensic practice may not be substantially more sensitive to heteroplasmy detection than the high quality Sanger data that can be developed at present. Regardless of any research efforts aimed at determining the lowest possible minor variant detection

threshold for forensics, though, it is important to note that mtDNA heteroplasmy detection guidelines for NGS data for forensic purposes will likely need to be determined via internal validation studies on a by-laboratory, by-platform and by-protocol basis, in the same way that stutter filters and relative fluorescent unit thresholds are presently determined for STR typing in forensics [182,183]. Whatever the intent of the research (tending toward basic or applied), additional and more focused investigations into mtDNA heteroplasmy detection (and, very similarly, the deconvolution of mixed contributor samples – which has implications beyond forensic identity testing [76]) from NGS data are clearly warranted.

If the forensic community ultimately determines that the entire mtGenome can typed for identification purposes, it seems it would also be worthwhile to revisit the interpretation rules presently applied to match comparisons of mtDNA data. Current U.S. interpretation guidelines [10] dictate that matching nucleotide states at all typed positions between a questioned specimen and a known sample be reported as "cannot be excluded", a mismatch at one position be described as "inconclusive", and a mismatch at two positions as an "exclusion" (all in reference to the mtDNA having originated from the same maternal lineage). These rules are applied without regard to the particular variant mtDNA positions in question, despite substitution rate heterogeneity being a well-known feature of the human mtGenome [8,40,52,122-124,127,184,185]. Available information indicates that some coding region positions have relative substitution rates on par with some of the fastest-evolving CR positions; and on the opposite end of the spectrum, that many coding region positions have never been observed as variant. For example, when the entire mtGenome is considered, three coding region positions - 709, 11914 and 5460 – rank among the 20 mtDNA sites with the greatest number of observed changes [8], yet nearly 65% of protein-coding gene positions were completely invariant in an analysis of more than 5000 complete

mtGenomes [30]. Given the high relative substitution rate of some coding region positions, future mtDNA match comparisons based on larger portions of the molecule may require the forensic mtDNA community to evaluate and consider interpretation guidelines that incorporate rate information in a statistical framework, rather than simply counting the number of differences between two mtDNA profiles, to accurately assess the weight of such sequence differences and the likelihood of observation of more than one complete substitution between generations. Clearly, such an approach would also permit more nuanced interpretation of a sequence difference observed between samples at a single mtDNA position when that position has an extremely low substitution rate. The incorporation of position-specific substitution rates into LR calculations for the interpretation of mtDNA match comparisons has been previously proposed [186,187], but not yet implemented.

With highly sensitive and more cost-effective typing methodologies on the horizon, the forensic mtDNA field is truly maturing and approaching its fullest potential. Though there is a clear need for further research to support the continued use of mtDNA for forensic purposes, especially as relates to the implementation of NGS-based protocols, the work reported here lays important pieces of the essential foundation for extension of typing efforts to the full mtGenome.

V. References

[1] J.A. Irwin, W. Parson, M.D. Coble, R.S. Just, mtGenome reference population databases and the future of forensic mtDNA analysis, Forensic.Sci.Int.Genet. 5 (2011) 222-225.

[2] A.W. Röck, A. Dur, M. van Oven, W. Parson, Concept for estimating mitochondrial DNA haplogroups using a maximum likelihood approach (EMMA), Forensic.Sci.Int.Genet. 7 (2013) 601-609.

[3] E.A. Lyons, M.K. Scheible, K. Sturk-Andreaggi, J.A. Irwin, R.S. Just, A high-throughput Sanger strategy for human mitochondrial genome sequencing, BMC Genomics. 14 (2013) 881-2164-14-881.

[4] R.S. Just, M.K. Scheible, S.A. Fast, K. Sturk-Andreaggi, J.L. Higginbotham, E.A. Lyons, et al., Development of forensic-quality full mtGenome haplotypes: Success rates with low template specimens, Forensic.Sci.Int.Genet. 10 (2014) 73-79.

[5] R.S. Just, M.K. Scheible, S.A. Fast, K. Sturk-Andreaggi, A.W. Röck, J.M. Bush, et al., Full mtGenome reference data: Development and characterization of 588 forensic-quality haplotypes, Forensic Sci.Int.Genet. 14 (2015) 141-155.

[6] W. Parson, A. Dür, EMPOP--a forensic mtDNA database, Forensic.Sci.Int.Genet. 1 (2007) 88-92.

[7] J.A. Irwin, J.L. Saunier, K.M. Strouss, K.A. Sturk, T.M. Diegoli, R.S. Just, et al., Development and expansion of high-quality control region databases to improve forensic mtDNA evidence interpretation, Forensic.Sci.Int.Genet. 1 (2007) 154-157.

[8] P. Soares, L. Ermini, N. Thomson, M. Mormina, T. Rito, A. Rohl, et al., Correcting for purifying selection: an improved human mitochondrial molecular clock, Am.J.Hum.Genet. 84 (2009) 740-759.

[9] Serum specimens from the Department of Defense Serum Repository: The Armed Forces Health Surveillance Center, U.S. Department of Defense, Silver Spring, MD [November 8, 2010; August 1, 2011; and October 20, 2011].

[10] Scientific Working Group on DNA Analysis Methods (SWGDAM), Interpretation Guidelines for Mitochondrial DNA Analysis by Forensic DNA Testing Laboratories, (2013) http://swgdam.org/SWGDAM%20mtDNA_Interpretation_Guidelines_APPROVED_073013.pdf

[11] W. Parson, L. Gusmão, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, et al., DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing, Forensic Sci.Int.Genet. 13 (2014) 134-142.

[12] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, et al., Sequence and organization of the human mitochondrial genome, Nature. 290 (1981) 457-465.

[13] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat.Genet. 23 (1999) 147.

[14] H.J. Bandelt, W. Parson, Consistent treatment of length variants in the human mtDNA control region: a reappraisal, Int.J.Legal Med. 122 (2008) 11-21.

[15] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum.Mutat. 30 (2009) E386-94.

[16] T. Tsuzuki, H. Nomiyama, C. Setoyama, S. Maeda, K. Shimada, Presence of mitochondrial-DNA-like sequences in the human nuclear DNA, Gene. 25 (1983) 223-229.

[17] H. Niederstätter, S. Kochl, P. Grubwieser, M. Pavlic, M. Steinlechner, W. Parson, A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA, Forensic.Sci.Int.Genet. 1 (2007) 29-34.

[18] T.M. Diegoli, M.D. Coble, H. Niederstätter, O.M. Loreille, T.J. Parsons, The Use of a Mitochondrial DNA-Specific qPCR Assay to Assess Degradation and Inhibition. Presented at Mid-Atlantic Association of Forensic Scientists Annual Meeting, Washington, DC, May 2007.

[19] C.J. Clopper, E.S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika. 26 (1934) 404-413.

[20] C.H. Brenner, Fundamental problem of forensic mathematics--the evidential value of a rare haplotype, Forensic.Sci.Int.Genet. 4 (2010) 281-291.

[21] O. Lao, P.M. Vallone, M.D. Coble, T.M. Diegoli, M. van Oven, K.J. van der Gaag, et al., Evaluating self-declared ancestry of U.S. Americans with autosomal, Y-chromosomal and mitochondrial DNA, Hum.Mutat. 31 (2010) E1875-93.

[22] M.W. Allard, D. Polanskey, M.R. Wilson, K.L. Monson, B. Budowle, Evaluation of variation in control region sequences for Hispanic individuals in the SWGDAM mtDNA data set, J.Forensic Sci. 51 (2006) 566-573.

[23] J.L. King, B.L. LaRue, N. Novroski, M. Stoljarova, S.B. Seo, X. Zeng, et al., High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq, Forensic Sci.Int.Genet. 12 (2014) 128-135.

[24] C. Santos, B. Sierra, L. Alvarez, A. Ramos, E. Fernández, R. Nogués, et al., Frequency and pattern of heteroplasmy in the control region of human mitochondrial DNA, J.Mol.Evol. 67 (2008) 191-200.

[25] J.A. Irwin, J.L. Saunier, H. Niederstätter, K.M. Strouss, K.A. Sturk, T.M. Diegoli, et al., Investigation of heteroplasmy in the human mitochondrial DNA control region: a synthesis of observations from more than 5000 global population samples, J.Mol.Evol. 68 (2009) 516-527.

[26] A. Ramos, C. Santos, L. Mateiu, M. Gonzalez Mdel, L. Alvarez, L. Azevedo, et al., Frequency and pattern of heteroplasmy in the complete human mitochondrial genome, PLoS One. 8 (2013) e74636.

[27] M. Li, A. Schönberg, M. Schaefer, R. Schroeder, I. Nasidze, M. Stoneking, Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes, Am.J.Hum.Genet. 87 (2010) 237-249.

[28] J.L. Elson, D.M. Turnbull, N. Howell, Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection, Am.J.Hum.Genet. 74 (2004) 229-238.

[29] T. Kivisild, P. Shen, D.P. Wall, B. Do, R. Sung, K. Davis, et al., The role of selection in the evolution of human mitochondrial genomes, Genetics. 172 (2006) 373-387.

[30] L. Pereira, F. Freitas, V. Fernandes, J.B. Pereira, M.D. Costa, S. Costa, et al., The diversity present in 5140 human mitochondrial genomes, Am.J.Hum.Genet. 84 (2009) 628-640.

[31] L. Pereira, P. Soares, P. Radivojac, B. Li, D.C. Samuels, Comparing phylogeny and the predicted pathogenicity of protein variations reveals equal purifying selection across the global human mtDNA diversity, Am.J.Hum.Genet. 88 (2011) 433-439.

[32] A. Röck, J. Irwin, A. Dür, T. Parsons, W. Parson, SAM: String-based sequence search algorithm for mitochondrial DNA database queries, Forensic.Sci.Int.Genet. 5 (2011) 126-132.

[33] H.J. Bandelt, A. Salas, C. Bravi, Problems in FBI mtDNA database, Science. 305 (2004) 1402-1404.

[34] H.J. Bandelt, A. Salas, S. Lutz-Bonengel, Artificial recombination in forensic mtDNA population databases, Int.J.Legal Med. 118 (2004) 267-273.

[35] Y.G. Yao, C.M. Bravi, H.J. Bandelt, A call for mtDNA data quality control in forensic science, Forensic Sci.Int. 141 (2004) 1-6.

[36] W. Parson, G.E. Huber, L. Moreno, M. Madel, M.D. Brandhagen, S. Nagl, et al., Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples, Forensic.Sci.Int. Genet. 15 (2015) 8-15.

[37] S. Pääbo, R.G. Higuchi, A.C. Wilson, Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology, J.Biol.Chem. 264 (1989) 9709-9712.

[38] M.N. Hochmeister, B. Budowle, U.V. Borer, U. Eggmann, C.T. Comey, R. Dirnhofer, Typing of deoxyribonucleic acid (DNA) extracted from compact bone from human remains, J.Forensic Sci. 36 (1991) 1649-1661.

[39] M.M. Holland, D.L. Fisher, L.G. Mitchell, W.C. Rodriquez, J.J. Canik, C.R. Merril, et al., Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War, J.Forensic Sci. 38 (1993) 542-553.

[40] M.M. Holland, T.J. Parsons, Mitochondrial DNA sequence analysis - validation and use for forensic casework, Forensic Sci.Rev. 11 (1999).

[41] M.R. Wilson, D. Polanskey, J. Butler, J.A. DiZinno, J. Replogle, B. Budowle, Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts, BioTechniques. 18 (1995) 662-669.

[42] A.R. Eisenberg, J.M. Moore, Mitochondrial DNA analysis at the FBI laboratory, Forensic.Sci. Comm. 1 (1999) http://www.fbi.gov/hq/lab/fsc/backissu/july1999/dnalist.htm.

[43] T. Melton, K. Nelson, Forensic mitochondrial DNA analysis: two years of commercial casework experience in the United States, Croat.Med.J. 42 (2001) 298-303.

[44] A. Carracedo, W. Bär, P. Lincoln, W. Mayr, N. Morling, B. Olaisen, et al., DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing, Forensic Sci.Int. 110 (2000) 79-85.

[45] S. Lutz, H.J. Weisser, J. Heizmann, S. Pollak, Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany, Int.J.Legal Med. 111 (1998) 67-77.

[46] S. Lutz, H. Wittig, H.J. Weisser, J. Heizmann, A. Junge, N. Dimo-Simonin, et al., Is it possible to differentiate mtDNA by means of HVIII in samples that cannot be distinguished by sequencing the HVI and HVII regions? Forensic Sci.Int. 113 (2000) 97-101.

[47] A. Brandstätter, H. Niederstätter, M. Pavlic, P. Grubwieser, W. Parson, Generating population data for the EMPOP database - an overview of the mtDNA sequencing and data evaluation processes considering 273 Austrian control region sequences as example, Forensic Sci.Int. 166 (2007) 164-175.

[48] A. Carracedo, J.M. Butler, L. Gusmão, W. Parson, L. Roewer, P.M. Schneider, Publication of population data for forensic purposes, Forensic.Sci.Int.Genet. 4 (2010) 145-147.

[49] A. Carracedo, J.M. Butler, L. Gusmão, A. Linacre, W. Parson, L. Roewer, et al., New guidelines for the publication of genetic population data, Forensic.Sci.Int.Genet. 7 (2013) 217-220.

[50] S.D. Lee, Y.S. Lee, J.B. Lee, Polymorphism in the mitochondrial cytochrome B gene in Koreans. An additional marker for individual identification, Int.J.Legal Med. 116 (2002) 74-78.

[51] H. Andréasson, A. Asp, A. Alderborn, U. Gyllensten, M. Allen, Mitochondrial sequence analysis for forensic identification using pyrosequencing technology, BioTechniques. 32 (2002) 124-6, 128, 130-3.

[52] M.D. Coble, R.S. Just, J.E. O'Callaghan, I.H. Letmanyi, C.T. Peterson, J.A. Irwin, et al., Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians, Int.J.Legal Med. 118 (2004) 137-146.

[53] B. Quintáns, V. Alvarez-Iglesias, A. Salas, C. Phillips, M.V. Lareu, A. Carracedo, Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing, Forensic Sci.Int. 140 (2004) 251-257.

[54] S. Sigurdsson, M. Hedman, P. Sistonen, A. Sajantila, A.C. Syvänen, A microarray system for genotyping 150 single nucleotide polymorphisms in the coding region of human mitochondrial DNA, Genomics. 87 (2006) 534-542.

[55] K.A. Sturk, M.D. Coble, S.M. Barritt, T.J. Parsons, R.S. Just, The application of mtDNA SNPs to a forensic case, Forensic Science International: Genetics Supplement Series. 1 (2008) 295-297.

[56] R.S. Just, M.D. Leney, S.M. Barritt, C.W. Los, B.C. Smith, T.D. Holland, et al., The use of mitochondrial DNA single nucleotide polymorphisms to assist in the resolution of three challenging forensic cases, J.Forensic Sci. 54 (2009) 887-891.

[57] J.A. Irwin, S.M. Edson, O. Loreille, R.S. Just, S.M. Barritt, D.A. Lee, et al., DNA identification of "Earthquake McGoon" 50 years postmortem, J.Forensic Sci. 52 (2007) 1115-1118.

[58] A. Brandstätter, T.J. Parsons, W. Parson, Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups, Int.J.Legal Med. 117 (2003) 291-298.

[59] W. Parson, A. Brandstätter, H. Niederstätter, P. Grubwieser, R. Scheithauer, Unravelling the mystery of Nanga Parbat, Int.J.Legal Med. 121 (2007) 309-310.

[60] V. Alvarez-Iglesias, J.C. Jaime, A. Carracedo, A. Salas, Coding region mitochondrial DNA SNPs: targeting East Asian and Native American haplogroups, Forensic.Sci.Int.Genet. 1 (2007) 44-55.

[61] T.M. Nelson, R.S. Just, O. Loreille, M.S. Schanfield, D. Podini, Development of a multiplex single base extension assay for mitochondrial DNA haplogroup typing, Croat.Med.J. 48 (2007) 460-472.

[62] S. Kohnemann, H. Pfeiffer, Application of mtDNA SNP analysis in forensic casework, Forensic.Sci.Int.Genet. 5 (2011) 216-221.

[63] R.S. Just, O.M. Loreille, J.E. Molto, D.A. Merriwether, S.R. Woodward, C. Matheson, et al., Titanic's unknown child: the critical role of the mitochondrial DNA coding region in a re-identification effort, Forensic.Sci.Int.Genet. 5 (2011) 231-235.

[64] P.M. Vallone, R.S. Just, M.D. Coble, J.M. Butler, T.J. Parsons, A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome, Int.J.Legal Med. 118 (2004) 147-157.

[65] M. Allen, H. Andréasson, Mitochondrial D-loop and coding sequence analysis using pyrosequencing, Methods Mol.Biol. 297 (2005) 179-196.

[66] A. Salas, B. Quintáns, V. Alvarez-Iglesias, SNaPshot typing of mitochondrial DNA coding region variants, Methods Mol.Biol. 297 (2005) 197-208.

[67] P.M. Vallone, J.P. Jakupciak, M.D. Coble, Forensic application of the Affymetrix human mitochondrial resequencing array, Forensic.Sci.Int.Genet. 1 (2007) 196-198.

[68] R.S. Just, A.M. Lehrmann, K.E. Harris, M.D. Coble, Comparison of the mitoSEQr Resequencing Sets to Standard mtDNA Sequencing Protocols. Presented at English-Speaking Working Group (ESWG) Meeting of the International Society for Forensic Genetics (ISFG), Sinaia, Romania, June 2008.

[69] M. Margulies, M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, et al., Genome sequencing in microfabricated high-density picolitre reactors, Nature. 437 (2005) 376-380.

[70] M. Knapp, M. Hofreiter, Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives, Genes. 1 (2010) 227-243.

[71] S.Y. Ho, M.T. Gilbert, Ancient mitogenomics, Mitochondrion. 10 (2010) 1-11.

[72] J.L. Paijmans, M.T. Gilbert, M. Hofreiter, Mitogenomic analyses from ancient DNA, Mol.Phylogenet.Evol. 69 (2013) 404-416.

[73] O. Loreille, H. Koshinsky, V.Y. Fofanov, J.A. Irwin, Application of next generation sequencing technologies to the identification of highly degraded unknown soldiers' remains, Forensic Sci.Int.Genet.Suppl.Ser. 3 (2011) e540-e541.

[74] J.E. Templeton, P.M. Brotherton, B. Llamas, J. Soubrier, W. Haak, A. Cooper, et al., DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification, Investig.Genet. 4 (2013) 26-2223-4-26.

[75] M. Mikkelsen, E. Rockenbauer, A. Wächter, L. Fendt, B. Zimmermann, W. Parson, et al., Application of full mitochondrial genome sequencing using 454 GS FLX pyrosequencing, Forensic Sci. Int.: Genet. Suppl. Series. 2 (2009) 518-519.

[76] M.M. Holland, M.R. McQuillan, K.A. O'Hanlon, Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy, Croat.Med.J. 52 (2011) 299-313.

[77] J. Irwin, R. Just, M. Scheible, O. Loreille, Assessing the potential of next generation sequencing technologies for missing persons identification efforts, Forensic Sci.Int.Genet.Suppl.Ser. 3 (2011) e447-e448.

[78] C. Van Neste, F. Van Nieuwerburgh, D. Van Hoofstat, D. Deforce, Forensic STR analysis using massive parallel sequencing, Forensic.Sci.Int.Genet. 6 (2012) 810-818.

[79] D.M. Bornman, M.E. Hester, J.M. Schuetter, M.D. Kasoji, A. Minard-Smith, C.A. Barden, et al., Short-read, high-throughput sequencing technology for STR genotyping, BioTechniques. 0 (2012) 1-6.

[80] W. Parson, C. Strobl, G. Huber, B. Zimmermann, S.M. Gomes, L. Souto, et al., Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM), Forensic.Sci.Int.Genet. 7 (2013) 543-549.

[81] E. Rockenbauer, S. Hansen, M. Mikkelsen, C. Børsting, N. Morling, Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing, Forensic.Sci.Int.Genet. 8 (2014) 68-72.

[82] J. Weber-Lehmann, E. Schilling, G. Gradl, D.C. Richter, J. Wiehler, B. Rolf, Finding the needle in the haystack: differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing, Forensic.Sci.Int.Genet. 9 (2014) 42-46.

[83] B.J. Bintz, G.B. Dixon, M.R. Wilson, Simultaneous Detection of Human Mitochondrial DNA and Nuclear-Inserted Mitochondrial-origin Sequences (NumtS) using Forensic mtDNA Amplification Strategies and Pyrosequencing Technology, J.Forensic Sci. 59 (2014) 1064-1073.

[84] M.K. Scheible, O. Loreille, R.S. Just, J.A. Irwin, Short tandem repeat typing on the 454 platform: Strategies and considerations for targeted sequencing of common forensic markers, Forensic Sci.Int.Genet. 12 (2014) 107-119.

[85] J.A. McElhoe, M.M. Holland, K.D. Makova, M.S. Su, I.M. Paul, C.H. Baker, et al., Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq, Forensic.Sci.Int.Genet. 13C (2014) 20-29.

[86] M. Mikkelsen, R. Frank-Hansen, A.J. Hansen, N. Morling, Massively parallel pyrosequencing of the mitochondrial genome with the 454 methodology in forensic genetics, Forensic.Sci.Int. Genet. 12C (2014) 30-37.

[87] Y.G. Yao, A. Salas, I. Logan, H.J. Bandelt, mtDNA data mining in GenBank needs surveying, Am.J.Hum.Genet. 85 (2009) 929-33; author reply 933.

[88] NCBI, Primer-BLAST: Finding primers specific to your PCR template (using Primer3 and BLAST), http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?.

[89] Y.S. Lee, W.Y. Kim, M. Ji, J.H. Kim, J. Bhak, MitoVariome: a variome database of human mitochondrial DNA, BMC Genomics. 10 Suppl 3 (2009) S12-2164-10-S3-S12.

[90] B.C. Levin, H. Cheng, D.J. Reeder, A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection, Genomics. 55 (1999) 135-146.

[91] B.C. Levin, K.A. Holland, D.K. Hancock, M. Coble, T.J. Parsons, L.J. Kienker, et al., Comparison of the complete mtDNA genome sequences of human cell lines--HL-60 and GM10742A--from individuals with pro-myelocytic leukemia and leber hereditary optic neuropathy, respectively, and the inclusion of HL-60 in the NIST human mitochondrial DNA standard reference material--SRM 2392-I, Mitochondrion. 2 (2003) 387-400.

[92] R.S. Just, T.M. Diegoli, J.L. Saunier, J.A. Irwin, T.J. Parsons, Complete mitochondrial genome sequences for 265 African American and U.S. "Hispanic" individuals, Forensic.Sci.Int.Genet. 2 (2008) e45-8.

[93] T.J. Parsons, Mitochondrial DNA Genome Sequencing and SNP Assay Development for Increased Power of Discrimination, Armed Forces DNA Identification Laboratory, Rockville, Maryland, 2006, http://www.ncjrs.gov/pdffiles1/nij/grants/213502.pdf.

[94] S. Rozen, H.J. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, in: Krawetz S., Misener S. (Eds.), Bioinformatics Methods and Protocols: Methods in Molecular Biology, Humana Press, Totowa, NJ, 2000, pp. 365-386.

[95] A. Goios, L. Prieto, A. Amorim, L. Pereira, Specificity of mtDNA-directed PCR-influence of NUclear MTDNA insertion (NUMT) contamination in routine samples and techniques, Int.J.Legal Med. 122 (2008) 341-345.

[96] C. Herrnstadt, W. Clevenger, S.S. Ghosh, C. Anderson, E. Fahy, S. Miller, et al., A novel mitochondrial DNA-like sequence in the human nuclear genome, Genomics. 60 (1999) 67-77.

[97] R Core Team, R: A language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012.

[98] W. Parson, A.W. Röck, EMPOP 3 NGS mitochondrial databasing, Oral presentation at the DNA in Forensics Conference, May 14-16, 2014, Brussels, Belgium.

[99] J.H. McDonald, Handbook of Biological Statistics (2nd ed.), Sparky House Publishing, Baltimore, MD, 2009, pp. 57-63.

[100] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, et al., Initial sequencing and analysis of the human genome, Nature. 409 (2001) 860-921.

[101] International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome, Nature. 431 (2004) 931-945.

[102] J. Schmutz, J. Martin, A. Terry, O. Couronne, J. Grimwood, S. Lowry, et al., The DNA sequence and comparative analysis of human chromosome 5, Nature. 431 (2004) 268-274.

[103] S.G. Gregory, K.F. Barlow, K.E. McLay, R. Kaul, D. Swarbreck, A. Dunham, et al., The DNA sequence and biological annotation of human chromosome 1, Nature. 441 (2006) 315-321.

[104] M.W. Allard, D. Polanskey, K. Miller, M.R. Wilson, K.L. Monson, B. Budowle, Characterization of human control region sequences of the African American SWGDAM forensic mtDNA data set, Forensic Sci.Int. 148 (2005) 169-179.

[105] V.F. Gonçalves, F. Prosdocimi, L.S. Santos, J.M. Ortega, S.D. Pena, Sex-biased gene flow in African Americans but not in American Caucasians, Genet.Mol.Res. 6 (2007) 256-261.

[106] J.L. Saunier, J.A. Irwin, R.S. Just, J. O'Callaghan, T.J. Parsons, Mitochondrial control region sequences from a U.S. "Hispanic" population sample, Forensic.Sci.Int.Genet. 2 (2008) e19-23.

[107] T.M. Diegoli, J.A. Irwin, R.S. Just, J.L. Saunier, J.E. O'Callaghan, T.J. Parsons, Mitochondrial control region sequences from an African American population sample, Forensic.Sci.Int.Genet. 4 (2009) e45-52.

[108] A. Salas, A. Carracedo, M. Richards, V. Macaulay, Charting the ancestry of African Americans, Am.J.Hum.Genet. 77 (2005) 676-680.

[109] K.L. Monson, K.W.P. Miller, M.R. Wilson, J.A. Dizzino, B. Budowle, The mtDNA population database: An integrated software and database resource for forensic comparisons, Forensic Sci Comm. 4 (2002) no 2.

[110] A. Salas, M. Richards, M.V. Lareu, R. Scozzari, A. Coppa, A. Torroni, et al., The African diaspora: mitochondrial DNA and the Atlantic slave trade, Am.J.Hum.Genet. 74 (2004) 454-465.

[111] B. Ely, J.L. Wilson, F. Jackson, B.A. Jackson, African-American mitochondrial DNAs often match mtDNAs found in multiple African ethnic groups, BMC Biol. 4 (2006) 34.

[112] K. Stefflova, M.C. Dulik, J.S. Barnholtz-Sloan, A.A. Pai, A.H. Walker, T.R. Rebbeck, Dissecting the within-Africa ancestry of populations of African descent in the Americas, PLoS One. 6 (2011) e14495.

[113] H.Y. Lee, U. Chung, J.E. Yoo, M.J. Park, K.J. Shin, Quantitative and qualitative profiling of mitochondrial DNA length heteroplasmy, Electrophoresis. 25 (2004) 28-34.

[114] T. Melton, Mitochondrial DNA heteroplasmy, Forensic Science Review. 16 (2004) 1-20.

[115] H.C. Lee, K.H. Huang, T.S. Yeh, C.W. Chi, Somatic alterations in mitochondrial DNA and mitochondrial dysfunction in gastric cancer progression, World J.Gastroenterol. 20 (2014) 3950-3959.

[116] L.A. Tully, T.J. Parsons, R.J. Steighner, M.M. Holland, M.A. Marino, V.L. Prenger, A sensitive denaturing gradient-Gel electrophoresis assay reveals a high frequency of heteroplasmy

in hypervariable region 1 of the human mtDNA control region, Am.J.Hum.Genet. 67 (2000) 432-443.

[117] C.D. Calloway, R.L. Reynolds, G.L. Herrin Jr, W.W. Anderson, The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age, Am.J.Hum.Genet. 66 (2000) 1384-1397.

[118] C. Theves, C. Keyser-Tracqui, E. Crubezy, J.P. Salles, B. Ludes, N. Telmon, Detection and quantification of the age-related point mutation A189G in the human mitochondrial DNA, J.Forensic Sci. 51 (2006) 865-873.

[119] T. Andrew, C.D. Calloway, S. Stuart, S.H. Lee, R. Gill, G. Clement, et al., A twin study of mitochondrial DNA polymorphisms shows that heteroplasmy at multiple sites is associated with mtDNA variant 16093 but not with zygosity, PLoS One. 6 (2011) e22332.

[120] Y. He, J. Wu, D.C. Dressman, C. Iacobuzio-Donahue, S.D. Markowitz, V.E. Velculescu, et al., Heteroplasmic mitochondrial DNA mutations in normal and tumour cells, Nature. 464 (2010) 610-614.

[121] H.J. Bandelt, A. Salas, Current next generation sequencing technology may not meet forensic standards, Forensic.Sci.Int.Genet. 6 (2012) 143-145.

[122] S. Meyer, G. Weiss, A. von Haeseler, Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA, Genetics. 152 (1999) 1103-1110.

[123] M. Stoneking, Hypervariable sites in the mtDNA control region are mutational hotspots, Am.J.Hum.Genet. 67 (2000) 1029-1032.

[124] K. Strouss, Relative evolutionary rate estimation for sites in the mtDNA control region, Masters Thesis, George Washington University, Washington, DC, 2006.

[125] N. Howell, J.L. Elson, C. Howell, D.M. Turnbull, Relative rates of evolution in the coding and control regions of African mtDNAs, Mol.Biol.Evol. 24 (2007) 2213-2221.

[126] J. Irwin, L. Moreno, T. Callaghan, Evaluating next generation sequencing technologies for expanded mitochondrial DNA identification capabilities at the FBI lab, Oral presentation at the DNA in Forensics Conference, May 14-16, 2014, Brussels, Belgium.

[127] T.J. Parsons, D.S. Muniec, K. Sullivan, N. Woodyatt, R. Alliston-Greiner, M.R. Wilson, et al., A high observed substitution rate in the human mitochondrial DNA control region, Nat.Genet. 15 (1997) 363-368.

[128] N. Howell, C.B. Smejkal, D.A. Mackey, P.F. Chinnery, D.M. Turnbull, C. Herrnstadt, The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates, Am.J.Hum.Genet. 72 (2003) 659-670.

[129] C. Santos, R. Montiel, B. Sierra, C. Bettencourt, E. Fernandez, L. Alvarez, et al., Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal), Mol.Biol.Evol. 22 (2005) 1490-1505.

[130] G. Avital, M. Buchshtav, I. Zhidkov, J. Tuval Feder, S. Dadon, E. Rubin, et al., Mitochondrial DNA heteroplasmy in diabetes and normal adults: role of acquired and inherited mutational patterns in twins, Hum.Mol.Genet. 21 (2012) 4214-4224.

[131] K. Ye, J. Lu, F. Ma, A. Keinan, Z. Gu, Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals, Proc.Natl.Acad.Sci.U.S.A. 111 (2014) 10654-10659.

[132] R.S. Just, J.A. Irwin, W. Parson, Questioning the prevalence and reliability of mitochondrial DNA heteroplasmy from massively parallel sequencing data, Proc.Natl.Acad.Sci.U.S.A. 111 (2014) E4546-4547.

[133] M.X. Sosa, I.K. Sivakumar, S. Maragh, V. Veeramachaneni, R. Hariharan, M. Parulekar, et al., Next-generation sequencing of human mitochondrial reference genomes uncovers high heteroplasmy frequency, PLoS Comput.Biol. 8 (2012) e1002737.

[134] L. Forster, P. Forster, S. Lutz-Bonengel, H. Willkomm, B. Brinkmann, Natural radioactivity and human mitochondrial DNA mutations, Proc.Natl.Acad.Sci.U.S.A. 99 (2002) 13950-13954.

[135] P. Soares, D. Abrantes, T. Rito, N. Thomson, P. Radivojac, B. Li, et al., Evaluating purifying selection in the mitochondrial DNA of various mammalian species, PLoS One. 8 (2013) e58993.

[136] B. Li, V.G. Krishnan, M.E. Mort, F. Xin, K.K. Kamati, D.N. Cooper, et al., Automated inference of molecular mechanisms of disease from amino acid substitutions, Bioinformatics. 25 (2009) 2744-2750.

[137] W. Parson, H.J. Bandelt, Extended guidelines for mtDNA typing of population data in forensic science, Forensic.Sci.Int.Genet. 1 (2007) 13-19.

[138] A. Salas, A. Carracedo, V. Macaulay, M. Richards, H.J. Bandelt, A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics, Biochem.Biophys.Res.Commun. 335 (2005) 891-899.

[139] F.A. Kaestle, R.A. Kittles, R.L. Roth, E.J. Ungvarsky, Database limitations on the evidentiary value of forensic mitochondrial DNA evidence, Am Crim Law Rev. 43 (2006) 53-88.

[140] J.A. Irwin, A. Ikramov, J. Saunier, M. Bodner, S. Amory, A. Röck, et al., The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns, Int.J.Legal Med. 124 (2010) 195-204.

[141] M. Scheible, M. Alenizi, K. Sturk-Andreaggi, M.D. Coble, S. Ismael, J.A. Irwin, Mitochondrial DNA control region variation in a Kuwaiti population sample, Forensic.Sci.Int.Genet. 5 (2011) e112-3.

[142] M. Bodner, B. Zimmermann, A. Röck, A. Kloss-Brandstatter, D. Horst, B. Horst, et al., Southeast Asian diversity: first insights into the complex mtDNA structure of Laos, BMC Evol.Biol. 11 (2011) 49-2148-11-49.

[143] L. Fendt, G. Huber, A.W. Röck, B. Zimmermann, M. Bodner, R. Delport, et al., Mitochondrial DNA control region data from indigenous Angolan Khoe-San lineages, Forensic.Sci.Int.Genet. 6 (2012) 662-663.

[144] L. Fendt, A. Röck, B. Zimmermann, M. Bodner, T. Thye, F. Tschentscher, et al., MtDNA diversity of Ghana: a forensic and phylogeographic view, Forensic.Sci.Int.Genet. 6 (2012) 244-249.

[145] R. Aboukhalid, K. Sturk-Andreaggi, M. Bouabdellah, D. Squalli, J.A. Irwin, S. Amzazi, Mitochondrial DNA control region variation from samples of the Moroccan population, Int.J.Legal Med. 127 (2013) 757-759.

[146] N. Ritter, Missing persons and unidentified remains: The nation's silent mass disaster, National Institute of Justice Journal. 256 (2007) 2-7.

[147] Illumina, ForenSeq Signature DNA Prep Kit Data Sheet, Pub No. 1470-2014-003. (13 January 2015).

[148] Life Tecnologies, HID-ION AmpliSeq Identity Panel, 2014.

[149] Illumina, Targeted Next-Generation Sequencing for Forensic Genomics, Pub No. 770-2012-034. (18 February 2013).

[150] L. Downey, J. Patel, S. Hermanson, L. Steffen, C. Sprecher, R.S. McLaren, et al., Development of PowerSeq Systems for Forensic Identification Using Next Generation Sequencing, American Academy of Forensic Sciences Annual Meeting. (2015).

[151] B. Budowle, U. Gyllensten, R. Chakraborty, M. Allen, Forensic analysis of the mitochondrial coding region and association to disease, Int.J.Legal Med. 119 (2005) 314-315.

[152] M.D. Coble, P.M. Vallone, R.S. Just, T.M. Diegoli, B.C. Smith, T.J. Parsons, Effective strategies for forensic analysis in the mitochondrial DNA coding region, Int.J.Legal Med. 120 (2006) 27-32.

[153] R.P. Erickson, Leber's optic atrophy, a possible example of maternal inheritance, Am.J.Hum.Genet. 24 (1972) 348-349.

[154] D.C. Wallace, G. Singh, M.T. Lott, J.A. Hodge, T.G. Schurr, A.M. Lezza, et al., Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy, Science. 242 (1988) 1427-1430.

[155] P.L. Adams, R.N. Lightowlers, D.M. Turnbull, Molecular analysis of cytochrome c oxidase deficiency in Leigh's syndrome, Ann.Neurol. 41 (1997) 268-270.

[156] MITOMAP: A Human Mitochondrial Genome Database, 2013, http://www.mitomap.org.

[157] C.H. Asplen, American Society of Law, Medicine & Ethics Report: The Non-Forensic Use of Biological Samples Taken for Forensic Purposes:An International Perspective, American Society of Law, Medicine & Ethics; Boston, MA, 2004, pp. 1-24.

[158] R. Williams, P. Johnson, Inclusiveness, effectiveness and intrusiveness: issues in the developing uses of DNA profiling in support of criminal investigations, J.Law Med.Ethics. 33 (2005) 545-558.

[159] Nuffield Council on Bioethics, The forensic use of bioinformation: Ethical issues, Nuffield Council on Bioethics, London, UK, 2007, pp. 1-139.

[160] E. Zadok, Legislative and Ethical Questions regarding DNA and Other Forensic "Biometric" Databases, Lecture Notes in Computer Science. 6005 (2010) 27-39.

[161] L.S. Parker, A.J. London, J.D. Aronson, Incidental findings in the use of DNA to identify human remains: an ethical assessment, Forensic.Sci.Int.Genet. 7 (2013) 221-229.

[162] J. Kim, S.H. Katsanis, Brave New World of human-rights DNA collection, Trends Genet. 29 (2013) 329-332.

[163] L. Caenazzo, P. Tozzo, Retaining the Genetic Profile of Innocent People: A Difficult Balance Between Respecting the Individual's Privacy and Public Security, Synesis: A Journal of Science, Technology, Ethics, and Policy. 4 (2013) 14.

[164] V. Marchese, N. Cerri, L. Caenazzo, Italian National Forensic DNA Database in an European perspective, Forensic Science International: Genetics Supplement Series. 4 (2013) e246-e247.

[165] M. Cox, P. Jones, Ethical considerations in the use of DNA as a contribution toward the determination of identification in historic cases: considerations from the Western front, New Genet.Soc. 33 (2014) 295-312.

[166] J.L. Scully, Naming the dead: DNA-based identification of historical remains as an act of care, New Genet.Soc. 33 (2014) 313-332.

[167] E. Haimes, V. Toom, Hidden in full sight: kinship, science and the law in the aftermath of the Srebrenica genocide, New Genet.Soc. 33 (2014) 277-294.

[168] V.B. Penchaszadeh, Ethical, legal and social issues in restoring genetic identity after forced disappearance and suppression of identity in Argentina, J.Community Genet. (2015).

[169] Presidential Commission for the Study of Bioethical Issues, Privacy and progress in whole genome sequencing, Washington, DC, 2012, pp. 1-140.

[170] Presidential Commission for the Study of Bioethical Issues, Anticipate and Communicate: Ethical Management of Incidental and Secondary Findings in the Clinical, Research, and Direct-to-Consumer Contexts, Washington, DC, 2013, pp. 1-146.

[171] J.E. McEwen, J.T. Boyer, K.Y. Sun, Evolving approaches to the ethical management of genomic data, Trends Genet. 29 (2013) 375-382.

[172] B.M. Knoppers, A. Thorogood, R. Chadwick, The Human Genome Organisation: towards next-generation ethics, Genome Med. 5 (2013) 38.

[173] R.C. Green, J.S. Berg, W.W. Grody, S.S. Kalia, B.R. Korf, C.L. Martin, et al., ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing, Genet.Med. 15 (2013) 565-574.

[174] C.G. van El, M.C. Cornel, P. Borry, R.J. Hastings, F. Fellmann, S.V. Hodgson, et al., Whole-genome sequencing in health care: recommendations of the European Society of Human Genetics, Eur.J.Hum.Genet. 21 (2013) 580-584.

[175] D.N. Paltoo, L.L. Rodriguez, M. Feolo, E. Gillanders, E.M. Ramos, J.L. Rutter, et al., Data use under the NIH GWAS data sharing policy and future directions, Nat.Genet. 46 (2014) 934-938.

[176] Nuffield Council on Bioethics, The collection, linking and use of data in biomedical research and health care: ethical issues; Nuffield Council on Bioethics, London, England, 2015, pp. 1-198.

[177] A. Onorato, Scientific Analysis on DNA Analysis Methods (SWGDAM) January 2015 Report, Quantico, VA, 2015, pp. 1-13.

[178] W. Parson, L. Roewer, Publication of population data of linearly inherited DNA markers in the International Journal of Legal Medicine, Int.J.Legal Med. 124 (2010) 505-509.

[179] B. Zimmermann, A. Röck, G. Huber, T. Kramer, P.M. Schneider, W. Parson, Application of a west Eurasian-specific filter for quasi-median network analysis: Sharpening the blade for mtDNA error detection, Forensic.Sci.Int.Genet. 5 (2011) 133-137.

[180] B. Zimmermann, A.W. Röck, A. Dur, W. Parson, Improved visibility of character conflicts in quasi-median networks with the EMPOP NETWORK software, Croat.Med.J. 55 (2014) 115-120.

[181] R.S. Just, M. Whitten, M. Li, E.A. Lyons, J.P. Ross, O.M. Loreille, et al., Development of Forensic-Quality Mitochondrial DNA Data on the Illumina Platform. Presented at Sequencing, Finishing, and Analyis in the Future Meeting, Santa Fe, NM, June 2012.

[182] M.D. Coble, Stutter, Mixture Interpretation Workshop, 21st International Symposium on Human Identification, San Antonio, Texas. (2010).

[183] Scientific Working Group on DNA Analysis Methods (SWGDAM), Validation guidelines for DNA analysis methods, 2012, http://swgdam.org/SWGDAM_Validation_Guidelines_APPROVED_Dec_2012.pdf.

[184] M. Hasegawa, A. Di Rienzo, T.D. Kocher, A.C. Wilson, Toward a more accurate time scale for the human mitochondrial DNA tree, J.Mol.Evol. 37 (1993) 347-354.

[185] J. Wakeley, Substitution rate variation among sites in hypervariable region 1 of human mitochondrial DNA, J.Mol.Evol. 37 (1993) 613-623.

[186] T.J. Parsons, Evolutionary rates in mtDNA sequences: Forensic applications and implications, Course lecture, pre-Congress workshop on Mitochondrial DNA, 22nd Congress of the International Society of Forensic Genetics, Copenhagen, Denmark. (2007).

[187] T.J. Parsons, K.M. Strouss, Site specific relative evolutionary rates in the mtDNA control region by phylogenetic analysis of a large population dataset, Oral presentation, DNA in Forensics conference, Ancona, Italy. (2008).

[188] M.S. Lee, B.C. Levin, MitoAnalyzer, 2000, http://www.cstl.nist.gov/biotech/strbase/mitoanalyzer.html.

VI. Dissemination of research findings

Published and accepted papers

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Röck AW, Bush JM, Higginbotham JL, Peck MA, Ring JD, Huber GE, Xavier C, Strobl C, Lyons EA, Diegoli TM, Bodner M, Fendt L, Kralj P, Nagl S, Niederwieser D, Zimmermann B, Parson W, Irwin JA. *Full mtGenome reference data: Development and characterization of 588 forensic-quality haplotypes*. Forensic Sci. Int. Genet. 14 (2015) 141-155.

Parson W, Gusmão L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, Parsons TJ. *DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing*, Forensic Sci. Int. Genet. 13C (2014) 134-142.

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Röck AW, Huber GE, Nagl S, Strobl C, Zimmerman B, Parson W, Irwin JA. *Development of forensic-quality full mtGenome haplotypes: Success rates with low template specimens*. Forensic Sci. Int. Genet. 10 (2014) 73-79.

Lyons EA, Scheible MK, Sturk-Andreaggi K, Irwin JA, Just RS. *A high-throughput Sanger strategy for human mitochondrial genome sequencing*. BMC Genomics.14 (2013) 881.

Sturk-Andreaggi K, Diegoli T, Just R, Irwin J. *Evaluation of automatable silica-based extraction methods for low quantity samples*. Forensic Sci. Int. Genet. Supplement Series 3 (2011) e504-505.

Presentations

Just R, Parson W, Irwin J. *Laying the foundation for forensic use of the complete mitochondrial genome through a collaborative databasing initiative.* Oral presentation at the 2015 NIJ Grantees Meeting, at the American Academy of Forensic Sciences Annual Meeting. Orlando, FL.

Just, RS. *Insights from examinations of full mitochondrial genome data, including the pattern and prevalence of heteroplasmy.* Oral presentation to AFDIL staff, December, 2014. Dover, DE.

Just R, Irwin J, Parson W. *The prevalence, pattern and reliability of mitochondrial genome heteroplasmy: Insights from recent studies.* Oral presentation at the Northeastern Academy of Forensic Sciences Annual Meeting, November, 2014. Hershey, PA.

Parson W. *Massively parallel mitochondrial DNA sequencing in forensic genetics.* Oral presentation at the International Symposium on Human Identification, September, 2014. Phoenix, AZ.

Just RS, Parson W, Irwin JA. *The full mitochondrial genome for forensics: Development and characterization of reference population data, and an examination of heteroplasmy rates and patterns.* Oral presentation to the FBI Casework and Support Units, September, 2014. Quantico, VA.

Parson W. *EMPOP 3: NGS mitochondrial databasing*. Oral presentation to the FBI Casework and Support Units, June, 2014. Quantico, VA.

Parson W. *The impact of massively parallel sequencing technologies for mtDNA analysis in forensic genetics*. Oral presentation at Forensica 2014: 4th International Conference on Forensic Genetics, May, 2014. Prague, the Czech Republic.

Just RS, Fast SA, Scheible MK, Sturk-Andreaggi K, Bush JM, Peck MA, Ring JD, Higginbotham JL, Lyons EA, Diegoli TM, Rock AW, Huber GE, Nagl S, Strobl C, Zimmermann B, Parson W, Irwin JA. *Full mtGenome reference population data: Development and evaluation of 588 forensic-quality haplotypes*. Poster presentation at the DNA in Forensics Meeting, May, 2014. Brussels, Belgium.

Parson W, Rock AW. *EMPOP 3: NGS mitochondrial databasing*. Oral presentation at the DNA in Forensics Meeting, May, 2014. Brussels, Belgium.

Parson W. *EMPOP*. Oral presentation at the EDNAP Meeting, April, 2014. Tbilisi, Georgia.

Just RS, Fast SA, Scheible MK, Sturk-Andreaggi K, Bush JM, Peck MA, Ring JD, Higginbotham JL, Lyons EA, Diegoli TM, Rock AW, Huber GE, Nagl S, Strobl C, Zimmermann B, Parson W, Irwin JA. *NIJ mtGenome grant project: Products, analyses and outcomes*. Oral presentation at AFDIL CE Day, March, 2014. Dover, DE.

Just RS, Fast SA, Scheible MK, Sturk-Andreaggi K, Bush JM, Peck MA, Ring JD, Higginbotham JL, Lyons EA, Diegoli TM, Rock AW, Huber GE, Nagl S, Strobl C, Zimmermann B, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data: Project introduction*. Oral presentation at AFDIL CE Day, March, 2014. Dover, DE.

Just RS, Fast SA, Scheible MK, Sturk-Andreaggi K, Higginbotham JL, Bush JM, Peck MA, Ring JD, Rock AW, Huber GE, Nagl S, Strobl C, Zimmermann B, Lyons EA, Diegoli TM, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Oral presentation at the American Academy of Sciences Annual Meeting, February, 2014. Seattle, WA.

Fast SA, Scheible MK, Irwin JA, Just RS. *Co-amplification of NUMTs in low DNA quantity specimens.* Oral presentation at the American Academy of Sciences Annual Meeting, February, 2014. Seattle, WA.

Parson W. *Mitochondrial DNA analysis using next generation sequencing technologies*. Oral presentation at the American Academy of Sciences Annual Meeting, February, 2014. Seattle, WA.

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Röck AW, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Oral presentation at the Institute for Legal Medicine, Innsbruck Medical University, December, 2013. Innsbruck, Austria.

Parson W. *Forensic exploration of the mitochondrial phylogeny from Eve's daughters to Mozart.* Oral presentation at the 5th Asian Forensic Science Network Annual Meeting and Symposium, November, 2013. Singapore.

Scheible M, Fast S, Andreaggi K, Lyons E, Diegoli T, Higginbotham J, Irwin J, Just R. *Generation of high-quality mtGenome reference data: Maximizing the typing success of high quality, low quantity specimens.* Poster presentation at Promega's International Symposium on Human Identification, October, 2013. Atlanta, GA.

Parson W. *Analytical concepts and new developments in forensic molecular genetics.* Oral presentation at the Life Technologies Asia Pacific Japan 2013 Genetic Analysis Summit, October, 2013.

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Röck AW, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Poster presentation at The 25th Annual World Congress of the International Society of Forensic Genetics, September, 2013. Melbourne, Australia.

Parson W, Huber G, Niederstatter H, Mayr-Eduardoff M, Strobl C, Rock AW, Bodner M. *Next generation mtGenome sequencing of forensic samples with the Ion Torrent PGM*. Oral presentation at 25th Annual World Congress of the International Society of Forensic Genetics, September, 2013. Melbourne, Australia.

Higginbotham JL, Diegoli TM, Scheible MK, Fast SA, Irwin JA, Just RS. *The use of a custom mt-qPCR assay for routine amplification of low template database samples*. Oral presentation at the Mid-Atlantic Association of Forensic Scientists Annual Meeting, May, 2013. Roanoke, VA.

Peck MA, Sturk-Andreaggi K, Just RS, Irwin JA. *Detection and forensic implication of large scale deletions of the mitochondrial coding region*. Oral presentation at the Mid-Atlantic Association of Forensic Scientists Annual Meeting, May, 2013. Roanoke, VA.

Parson W. *Next generation mtGenome sequencing using the Ion Torrent PGM.* Oral presentation at the Joint Conference of HGM 2013 and 21st International Congress of Genetics, April, 2013. Singapore.

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Röck AW, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Poster presentation at the American Academy of Forensic Sciences Annual Meeting, February, 2013. Washington, DC.

Sturk-Andreaggi K, Just R, Scheible M, Fast S, Irwin J. *Forensic considerations for automating a large-scale mtGenome databasing project*. Poster presentation at the Society for Laboratory Automation and Screening Annual Conference, January, 2013. Orlando, FL.

Just RS, Scheible MK, Fast SA, Sturk-Andreaggi K, Higginbotham JL, Lyons EA, Bush JM, Peck MA, Ring JD, Diegoli TM, Röck AW, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Oral presentation to the FBI Mitochondrial DNA Unit, December, 2012. Quantico, VA.

Strobl C, Huber G, Lagace R, Langit R, Wootton S, Hennessy L, Parson W. *Next generation mtGenome sequencing for forensic purposes using the Ion Torrent PGM*. Oral presentation at the DNA in Forensics Meeting, September, 2012. Innsbruck, Austria.

Just RS, Scheible MK, Fast SA, Daniels-Higginbotham J, Lyons EA, Sturk-Andreaggi K, Diegoli TM, Röck AW, Parson W, Irwin. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Poster presentation at the NIJ conference, June, 2012. Crystal City, VA.

Just RS, Whitten M, Li M, Lyons EA, Ross JP, Loreille OM, Irwin JA. *Development of forensic-quality mitochondrial DNA data on the Illumina platform*. Poster presentation at the Sequencing, Finishing, and Analysis in the Future, June, 2012. Santa Fe, NM.

Just RS, Scheible MK, Lyons EA, Sturk-Andreaggi K, Diegoli TM, Fast SA, Daniels-Higginbotham J, Rock AW, Parson W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data*. Oral presentation for the Forensic Research and Development Seminar Series: Research in Mass Disaster Identification Tools, May, 2012. Webinars (3).

Just R, Loreille O, Diegoli T, Sturk-Andreaggi A, Scheible M, Fast S, Daniels-Higginbotham J, Irwin J. *New technologies and ongoing AFDIL research efforts to advance the POW/MIA accounting mission.* Oral presentation at the JPAC-MAKRI Joint Symposium at the American Academy of Forensic Sciences Annual Meeting, February, 2012. Atlanta, GA

Just RS, Scheible MK, Lyons EA, Sturk-Andreaggi K, Diegoli TM, Daniels-Higginbotham J, Rock AW, Parson W, Irwin J. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data.* Oral presentation at the NIJ Symposium, Forensics Without Borders – How NIJ Funded Research Impacts the World, at the American Academy of Forensic Sciences Annual Meeting, February, 2012. Atlanta, GA.

Sturk-Andreaggi K, Diegoli T, Scheible M, Just R, Irwin J. *Evaluation of high-throughput extraction methods for low quality samples*. Poster presentation at The 24th World Conference of the International Society for Forensic Genetics, September, 2011. Vienna, Austria.

Just RS, Scheible MR, Lyons EA, Sturk-Andreaggi KA, Diegoli TM, Rock A, Parsons W, Irwin JA. *Maximizing mtDNA testing potential with the generation of high-quality mtGenome reference data.* Poster presentation at the NIJ Conference, June, 2011. Washington, DC.

Appendices

Appendix A. Figures

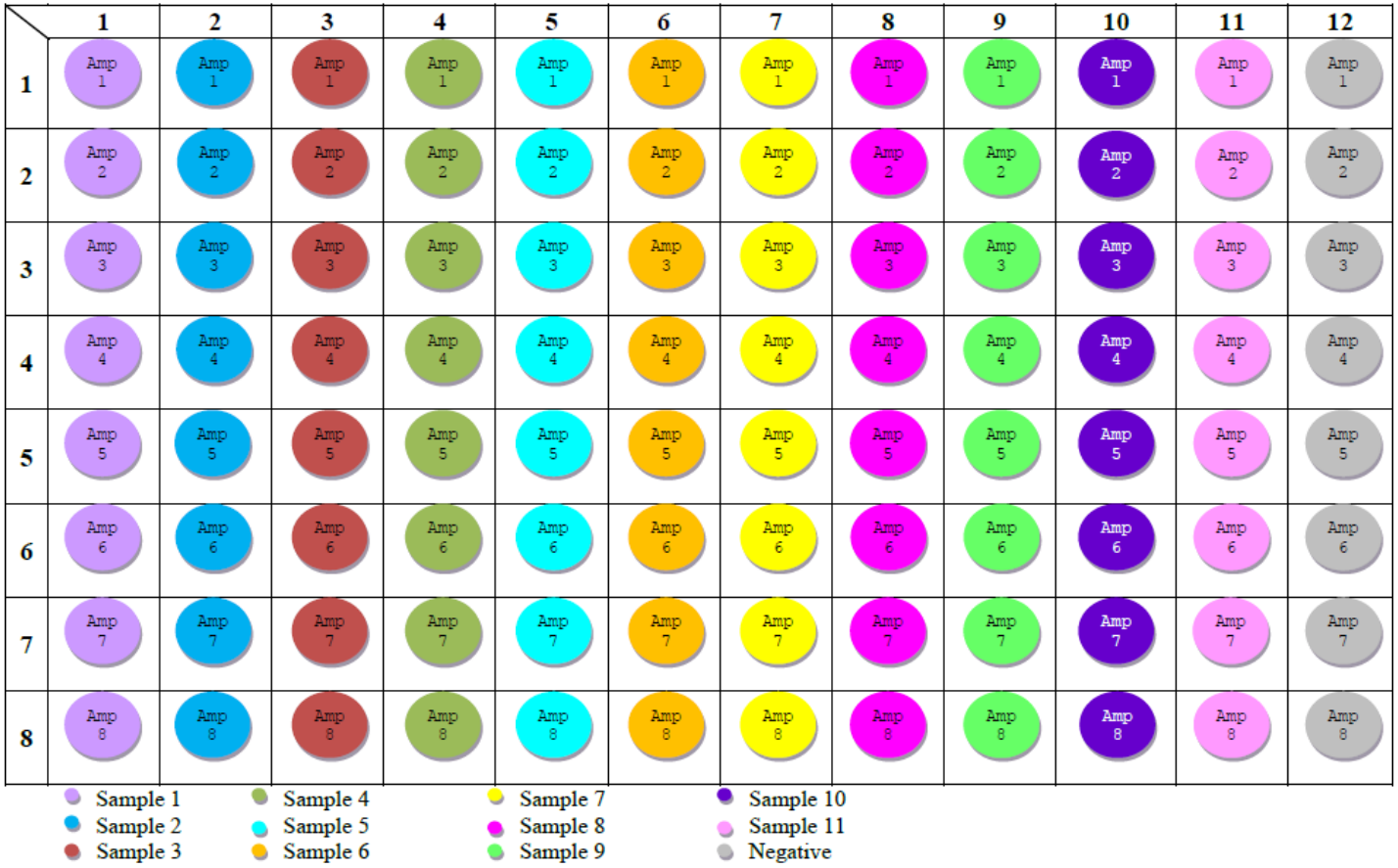


Figure 1. 96-well amplification plate map
The figure displays the strategy applied to permit simultaneous amplification of 11 samples in a 96-well plate. Samples are organized by columns (Sample 1 in column 1, Sample 2 in column 2, etc.), and each row represents one of the eight mtGenome amplicons (Amplicon 1 in row 1, Amplicon 2 in row 2, etc.). Negative controls for each target fragment are amplified in column 12.

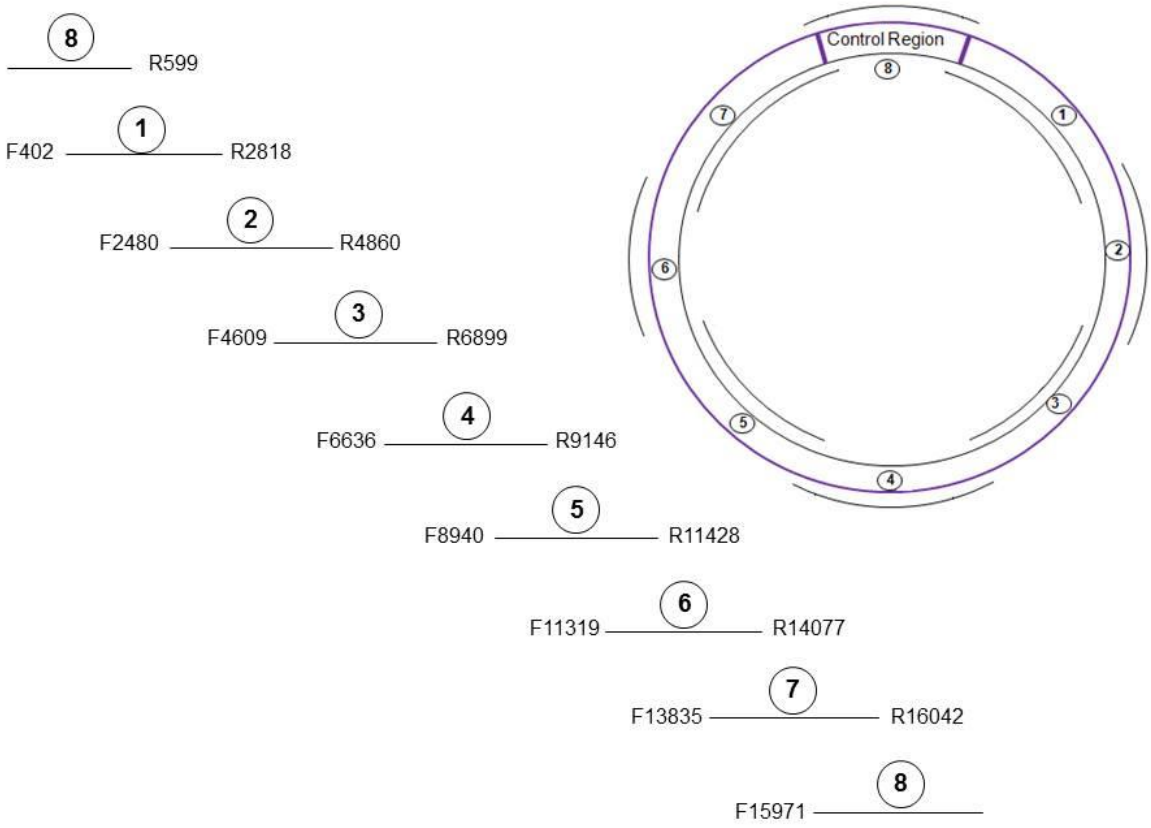
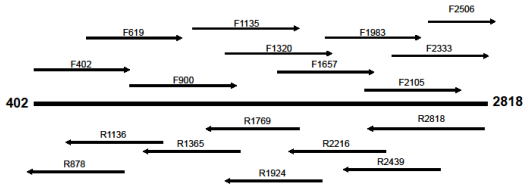


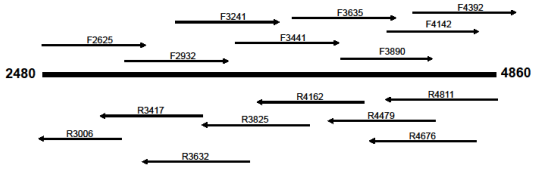
Figure 2. Amplification of the mtGenome in eight fragments
Positioning of the eight, overlapping target fragments around the circular mtGenome, along with the primers used to amplify each region, is depicted.

Figure 3. Organization of sequencing primers by amplicon (next page)
Approximate positioning and coverage of each of the 127 sequencing primers used to generate 135 sequences across the mtGenome. Amplicon start and end points (in terms of nucleotide position) are given, and forward sequences are represented above the template strand while reverse sequences are listed below. The number of sequencing primers per amplicon ranges from 16 to 18. The sequencing strategy used for the CR (Amplicon 8) was adopted wholesale from Irwin *et al.* [7] and uses eight distinct primers to produce 16 sequences.

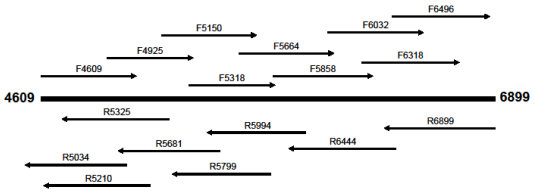
Amplicon 1



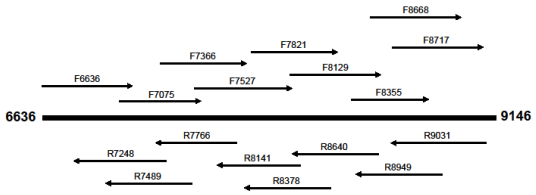
Amplicon 2



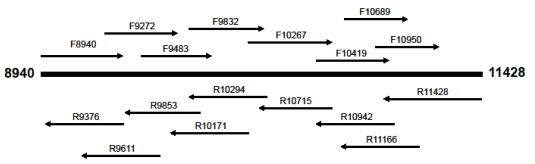
Amplicon 3



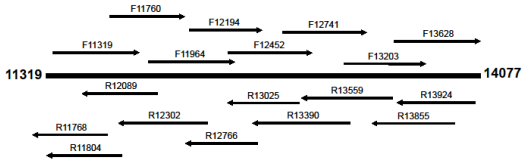
Amplicon 4



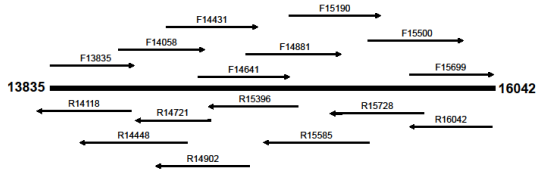
Amplicon 5



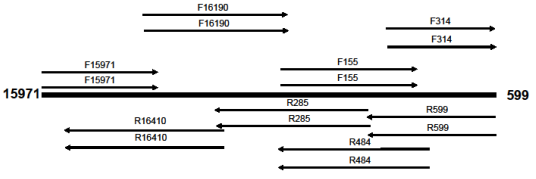
Amplicon 6



Amplicon 7



Amplicon 8 (CR)



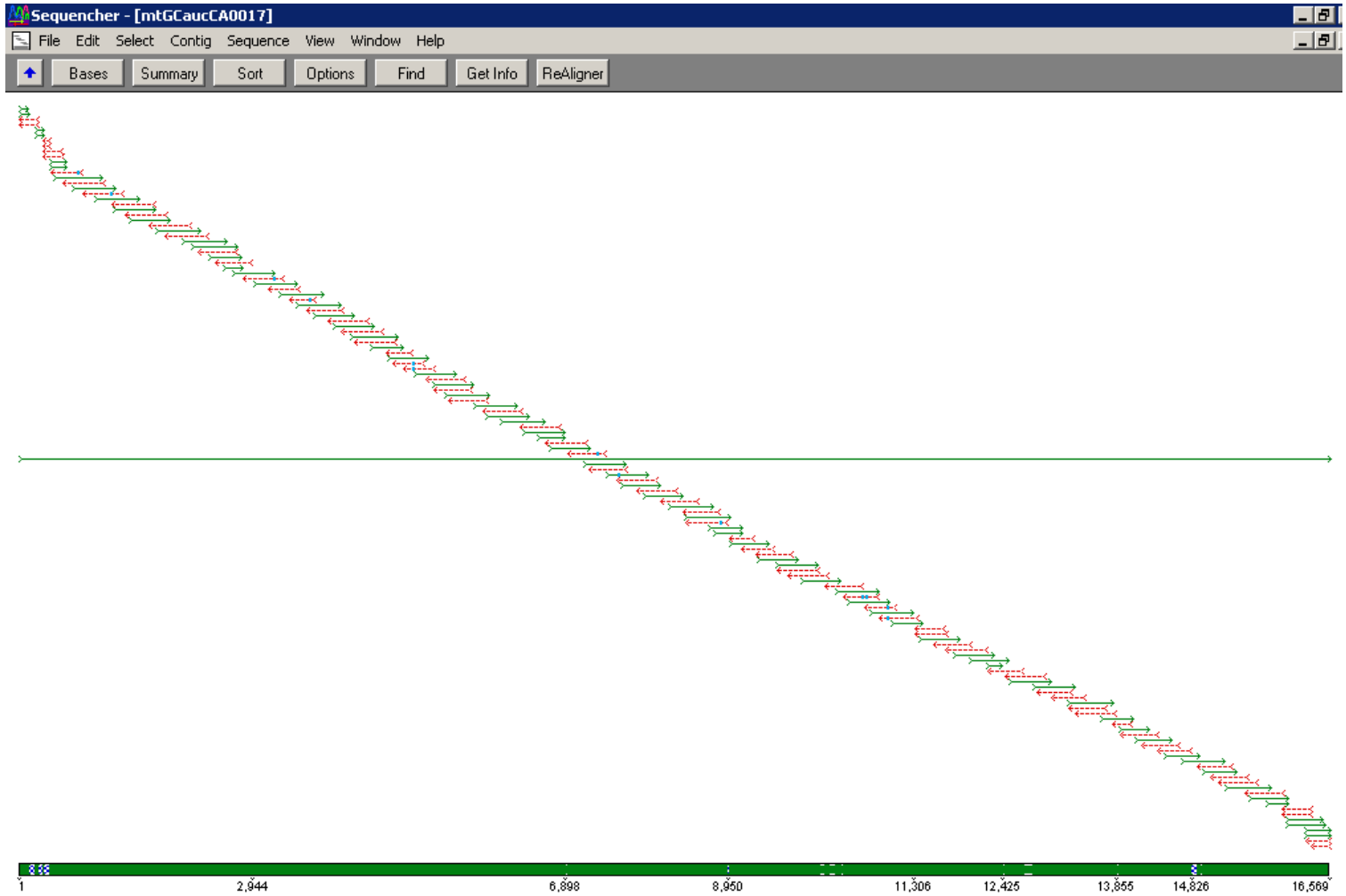


Figure 4. MtGenome sequence coverage
This Sequencher (Gene Codes Corporation) screen capture demonstrates the typical mtGenome sequence coverage that results from the 135-sequence strategy. Individual forward sequences are denoted in green, and reverse sequences are represented in red. The data come from a population sample processed using the protocol and selected at random, and no reprocessing was required to achieve complete coverage across the mtGenome. Small regions with replicate but unidirectional coverage (three in the CR due to polycytosine stretches and length heteroplasmy, and two in the coding region, totaling 294 bp) are indicated by blue hashing in the coverage bar.

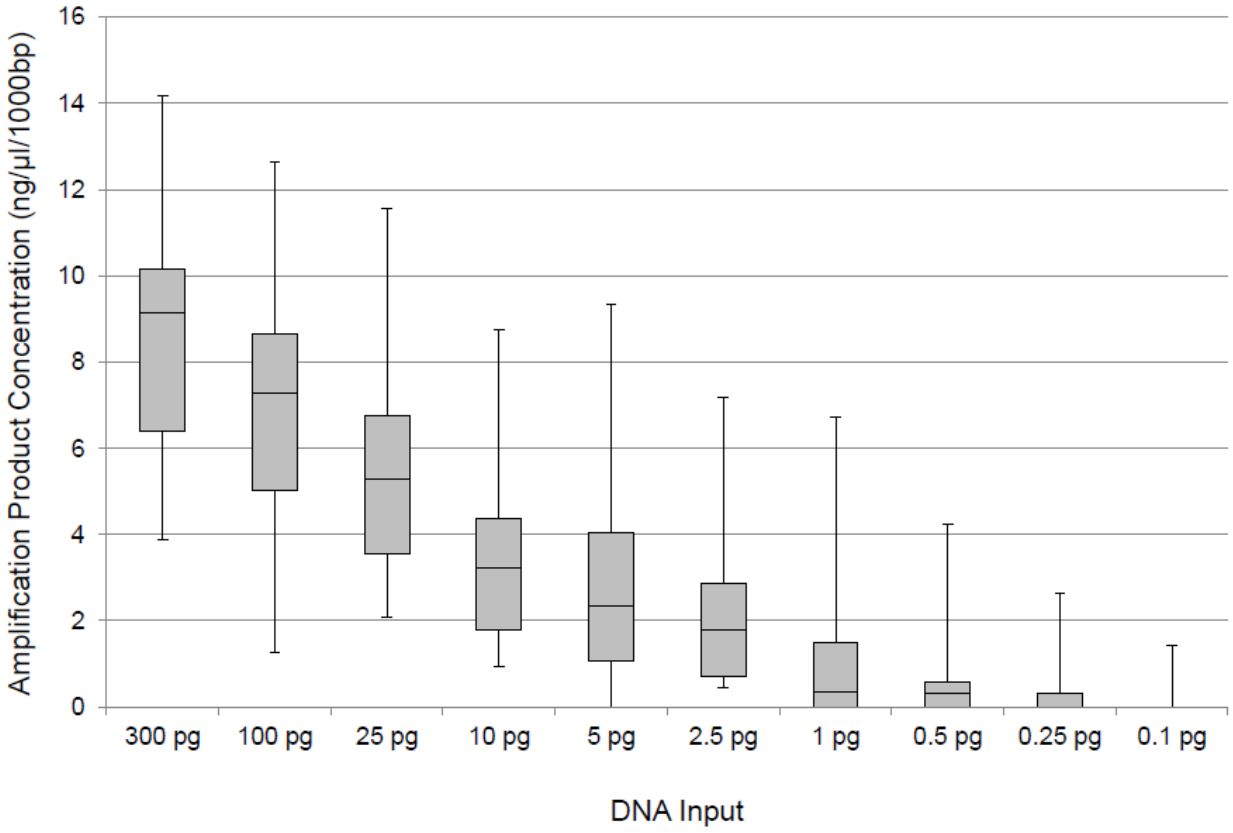


Figure 5. Amplification sensitivity with positive control DNA
Results from duplicate amplifications of all eight mtGenome fragments with ten input levels of positive control DNA (9947A). Amplification product concentrations of the target fragments were measured by automated injection on a QIAxcel Advanced capillary electrophoresis instrument (QIAGEN Inc.), and were normalized to ng/µL per 1000 bp (given the variable sizes of the eight target fragments). The wide range of product concentrations represented in the long "whiskers" reflects the differing amplification efficiencies of the eight primer pairs; however, PCR product was produced for all regions down to 10 pg input DNA.

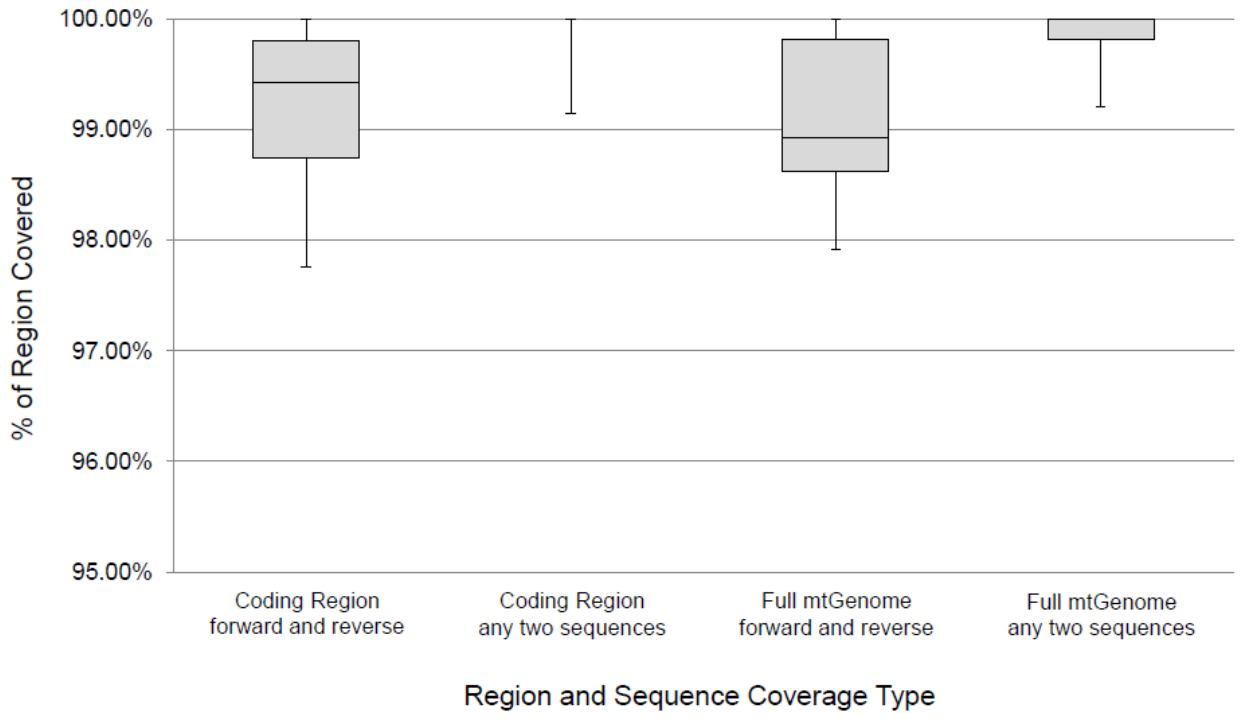


Figure 6. Sequence coverage
Percentage of the coding region or full mtGenome with redundant sequence coverage following a single pass of automated data generation for 11 population samples, representing a range of mtDNA haplogroups, processed in duplicate. One sample, for which all sequence data in a single direction for a single amplicon was unusable and sourced to instrument failure, was removed from the analysis as an outlier; and small regions of unidirectional sequence coverage due to length heteroplasmy in HV1 and HV2 in some samples were ignored. On average across the 22 samples, high quality forward and reverse coverage was produced for 99% of the mtGenome.

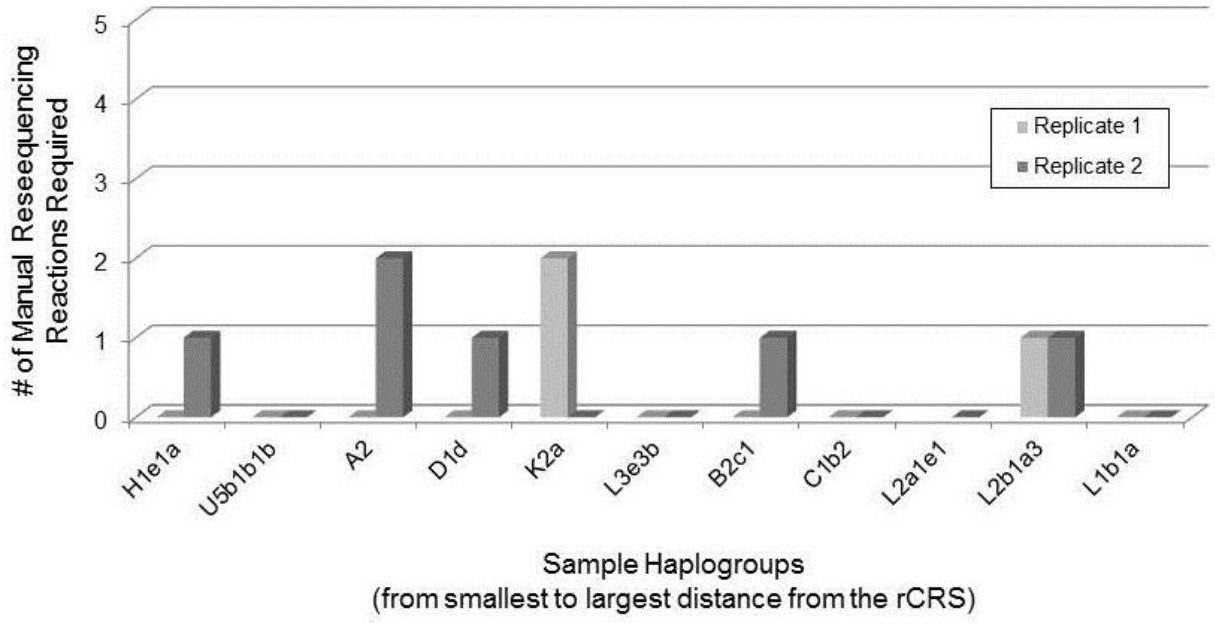


Figure 7. Reprocessing required
The number of manual resequencing reactions that would be required to achieve complete double stranded coverage for the 22 population samples (duplicate processing of 11 distinct samples) ranged from zero to two. This equates to approximately one resequencing reaction for every two mtGenomes processed. Replicate 1 of the sample representing haplogroup L2a1e1 was not included in the analysis due sequence failures resulting from instrument failure.

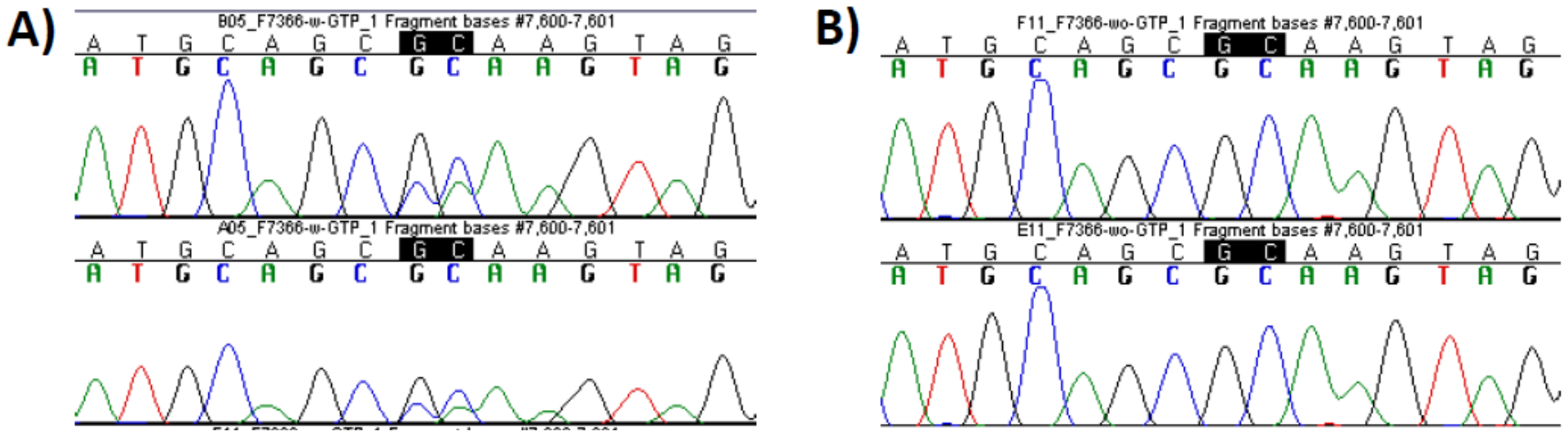


Figure 8. Sequencing artifacts
A) Screen captures of forward sequences aligned in Sequencher (Gene Codes Corporation) showing compression and unusual peak morphology around nucleotide position 7600. **B)** When the same primer (F7366) was used without the addition of dGTP BigDye® (Life Technologies, Applied Biosystems), the sequencing artifacts were no longer apparent.

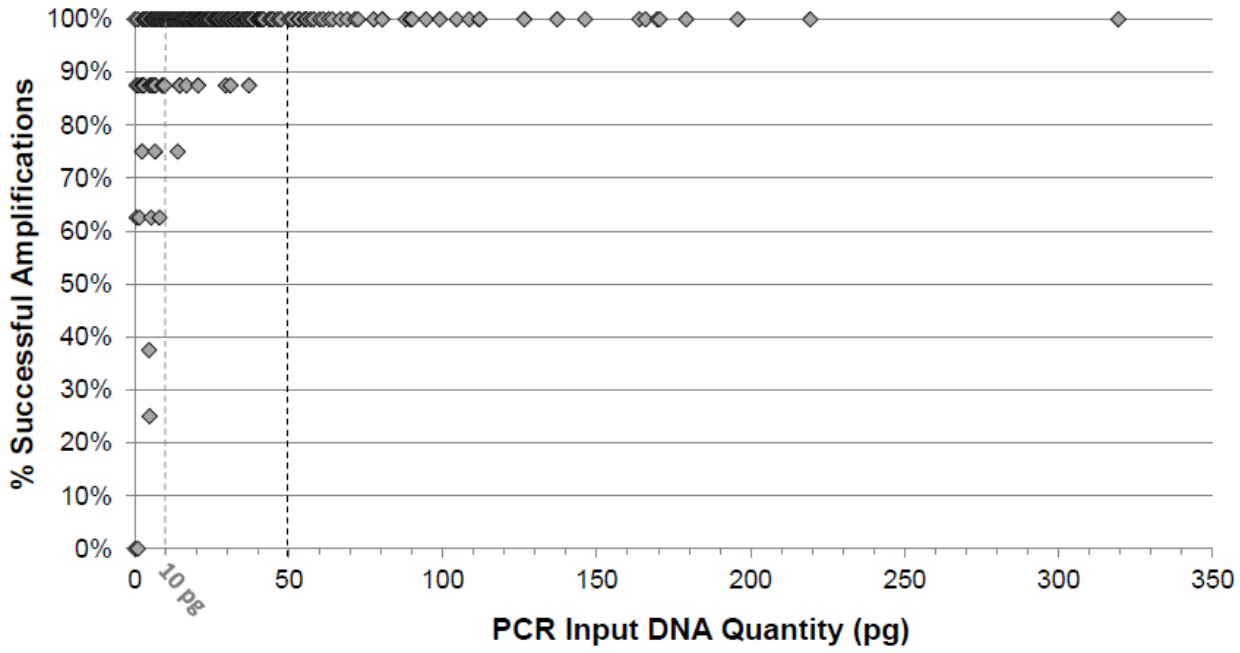


Figure 9. Amplification success with blood serum specimen extracts
Quantified using a modified mtDNA qPCR assay, 242 blood serum extract DNA concentrations ranged from 0.00-777.64 pg/µL. First pass amplification success rates (the percentage of eight amplicons which successfully amplified) for PCR inputs ranging from 0-319 pg are shown. Two samples, representing PCR inputs of 815 and 2333 pg respectively, and for which all eight amplicons were successfully amplified, are not in included in the plot. No amplification failures were observed with PCR inputs greater than 50 pg. In only one instance was more than one amplification failure observed when PCR inputs were greater than 10 pg.

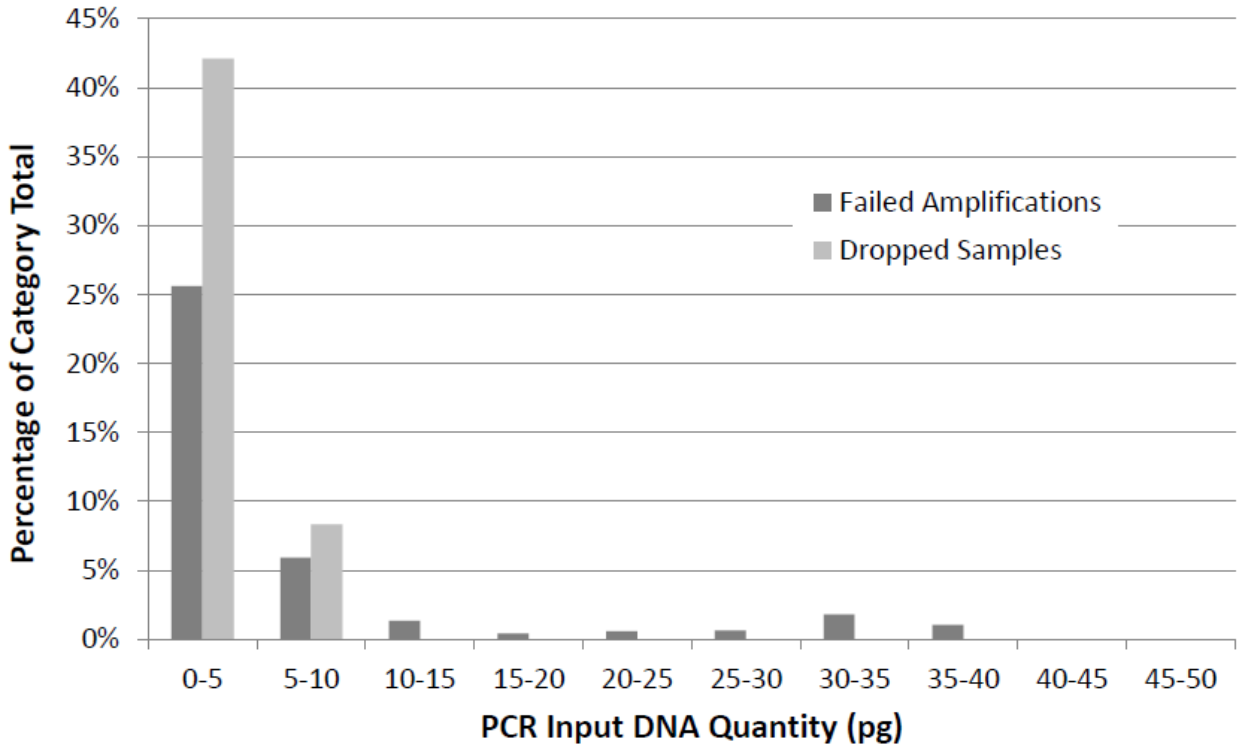


Figure 10. Amplification failures and dropped samples by DNA input quantity
Among 242 blood serum specimens that were quantified prior to amplification, 25.63% of amplifications failed and 42.11% of samples were dropped when PCR inputs were in the range of 0-5 pg. At PCR inputs between 5 and 10 pg, 5.90% of amplifications failed and 8.33% of samples were dropped.

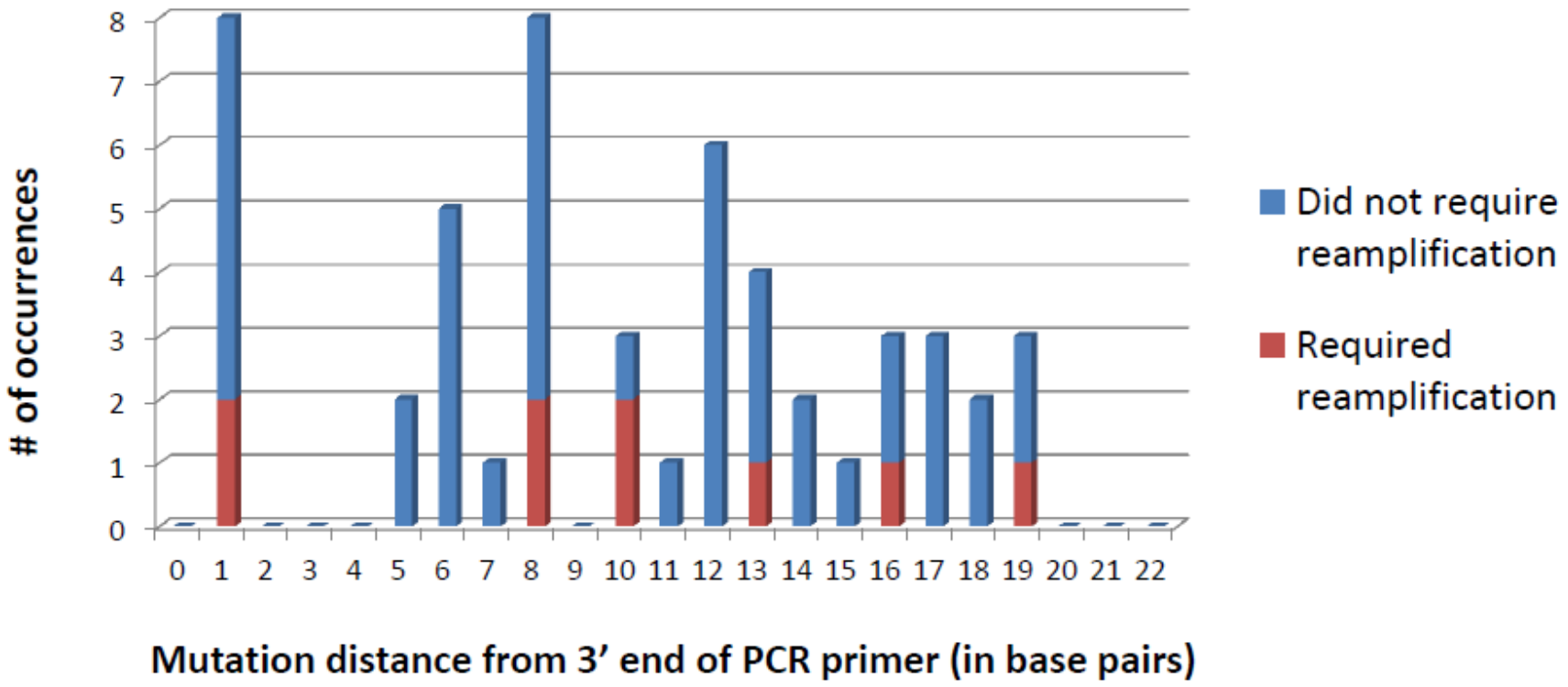


Figure 11. Positional distribution and incidence of PCR failure among samples with mutations in primer binding regions
Fifty-two total mutations in 588 samples were identified across the 16 PCR primer binding regions. In the majority of these cases (82.7%), the first pass of automated processing produced sufficient PCR product. Among the eight samples with mutations in primer binding regions that required manual reamplification due to near or complete PCR failures, the position of the mutation relative to the 3' end of the primer varied from 1 to 19 bp.

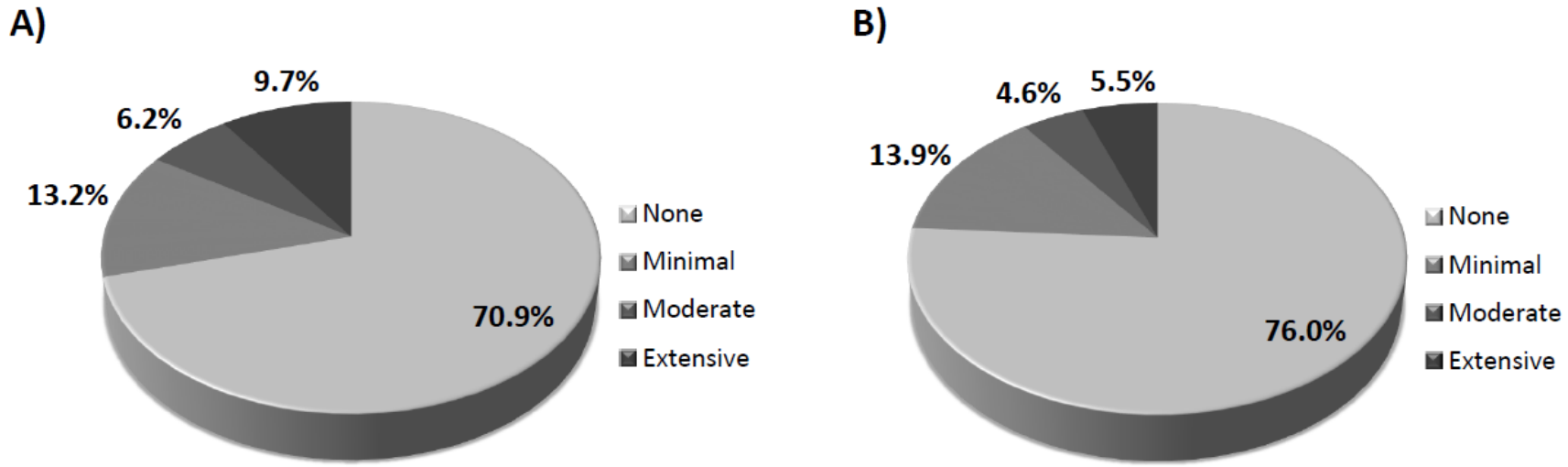


Figure 12. Extent of manual reprocessing
Manual reprocessing was performed when the first pass robotic processing did not result in complete sequence coverage across the entire mtGenome. Here, we have categorized the amount of reprocessing required as none, minimal (defined as one or two additional sequencing reactions), moderate (three to nine additional sequences), or extensive (ten or more additional sequences). Panel A demonstrates the reprocessing required when the first 433 samples sequenced were considered. The reprocessing displayed in Panel B ignores the reprocessing performed for Amplicons 2 and 6 prior to their redesign, and thus reflects the published amplification primer sets [3].

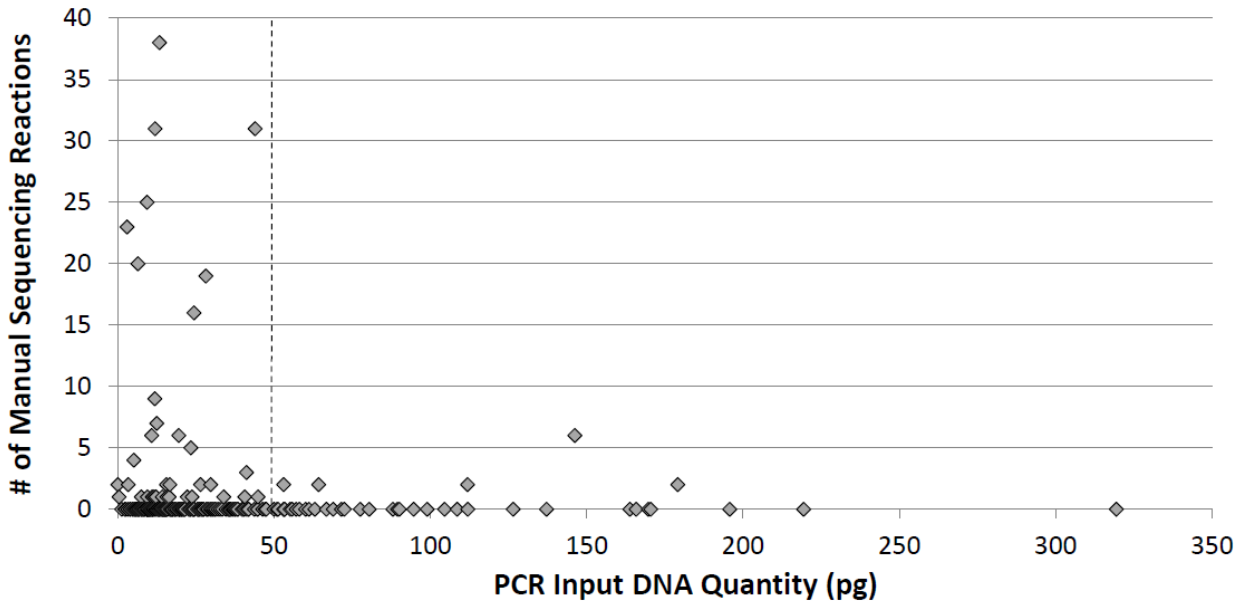


Figure 13. Manual sequencing by PCR input DNA quantity
Generally speaking, when PCR DNA input quantities exceeded 50 pg very little manual resequencing was required to achieve complete mtGenome haplotypes with the desired sequence coverage. For a set of 230 samples quantified prior to amplification, only one sample with a PCR input greater than 50 pg needed more than two additional sequences. Of the samples from this set which required more than 15 additional sequences, most were due to post-amplification enzymatic purification failure for one target region, which necessitated reamplification and complete resequencing of the fragment.

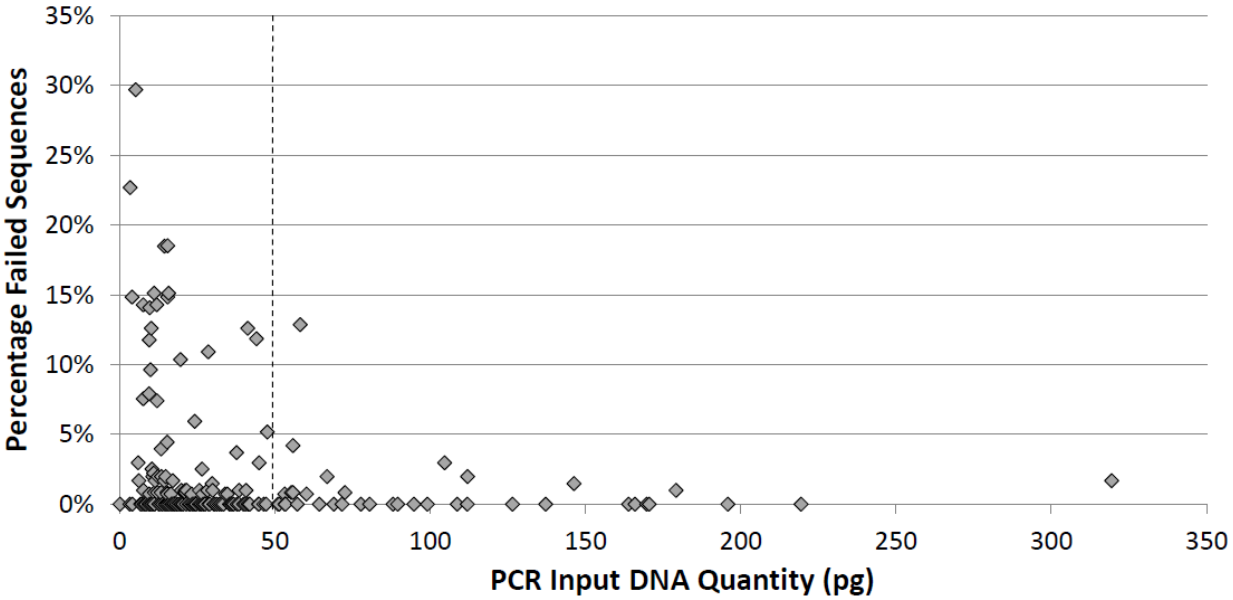


Figure 14. Sequence failures by PCR input DNA quantity
When 21,601 sequencing reactions were examined, sequencing success was more variable and sequencing failure more common at PCR DNA inputs below 50 pg. Overall, the sequencing failure rate for PCR inputs below 50 pg was 2.51%. When PCR inputs exceeded 50 pg, the sequencing failure rate was 0.82%, and only one of 39 samples had a sequencing failure rate greater than 5.0%.

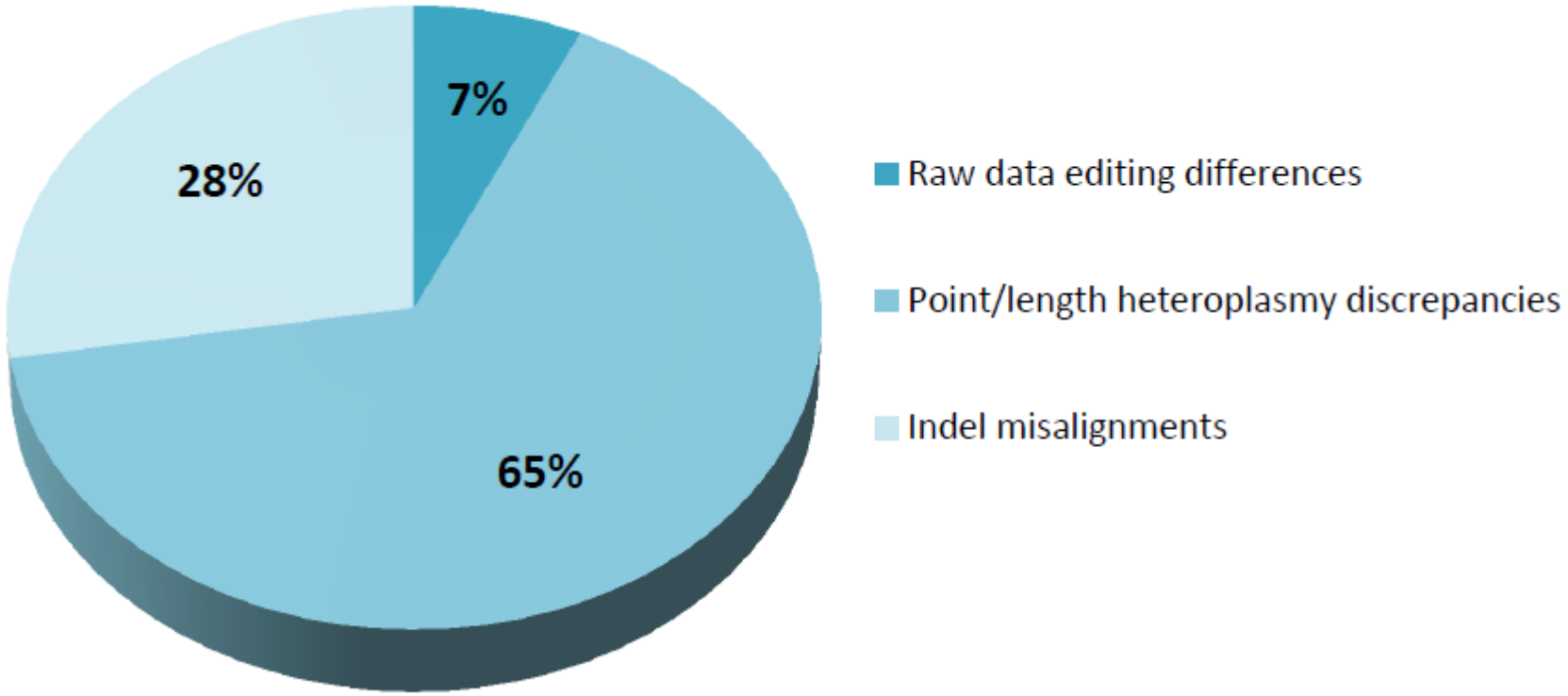


Figure 15. Review differences identified in the 588 haplotypes
A total of 29 discrepancies were identified across the 588 haplotypes. 27 of the occurrences were instances of non-concordance between the haplotypes independently generated by AFDIL and EMPOP. The majority of those discrepancies (19 instances) were related to point or length heteroplasmy. Two discrepancies resulted from raw data editing differences, while six discrepancies were due to different alignments of indels between the two laboratories. Two additional indel misalignments were detected in a later quality control check.

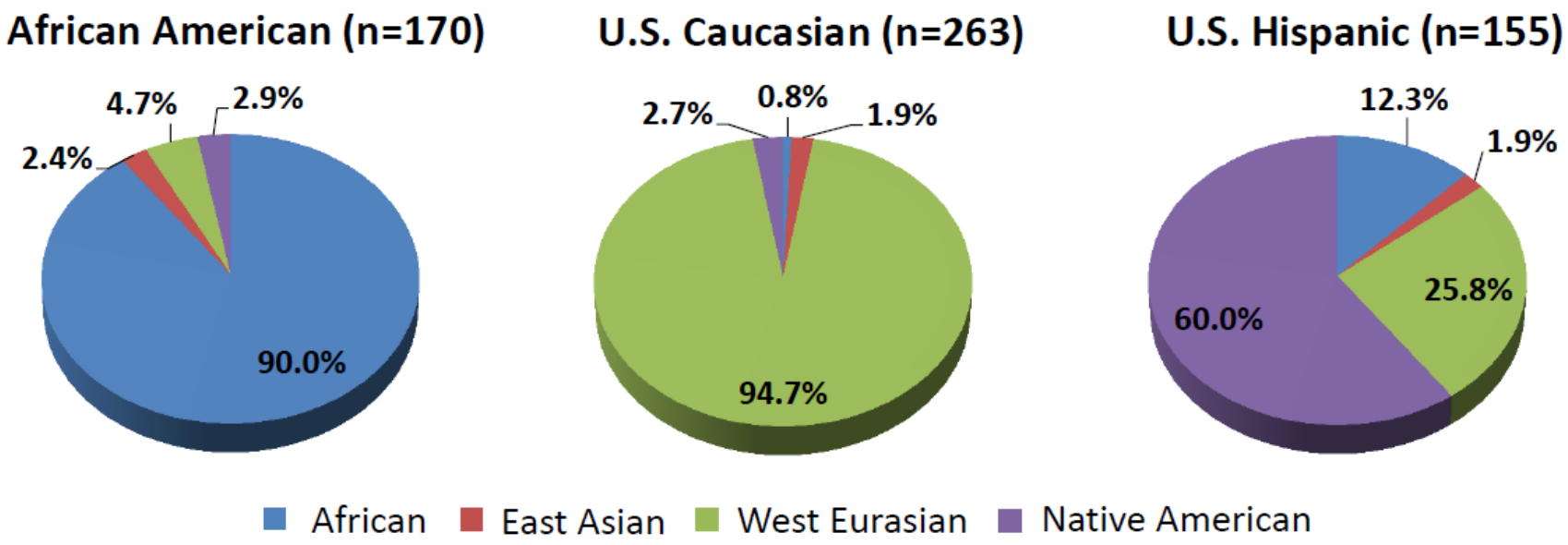


Figure 16. Biogeographic ancestry proportions in each of the three U.S. population group samples
Haplotypes for each population were assigned to one of four broad biogeographic ancestry categories (African, East Asian, West Eurasian and Native American) on the basis of EMMA [2] estimated haplogroups using PhyloTree build 16 [15].

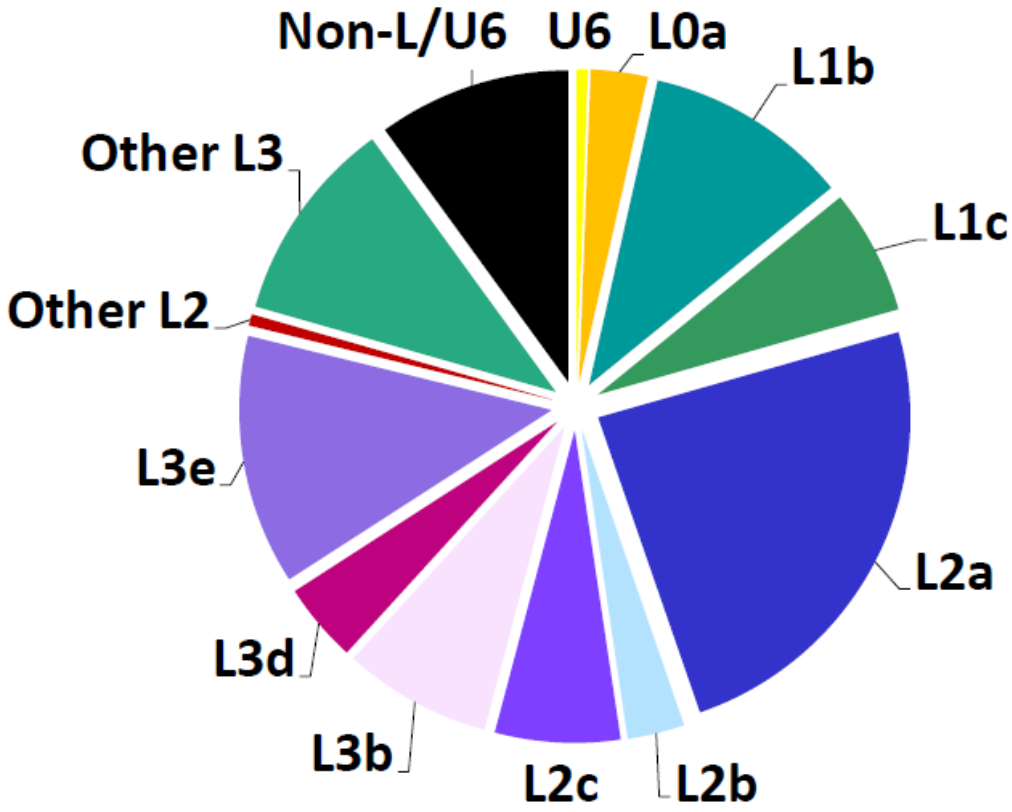
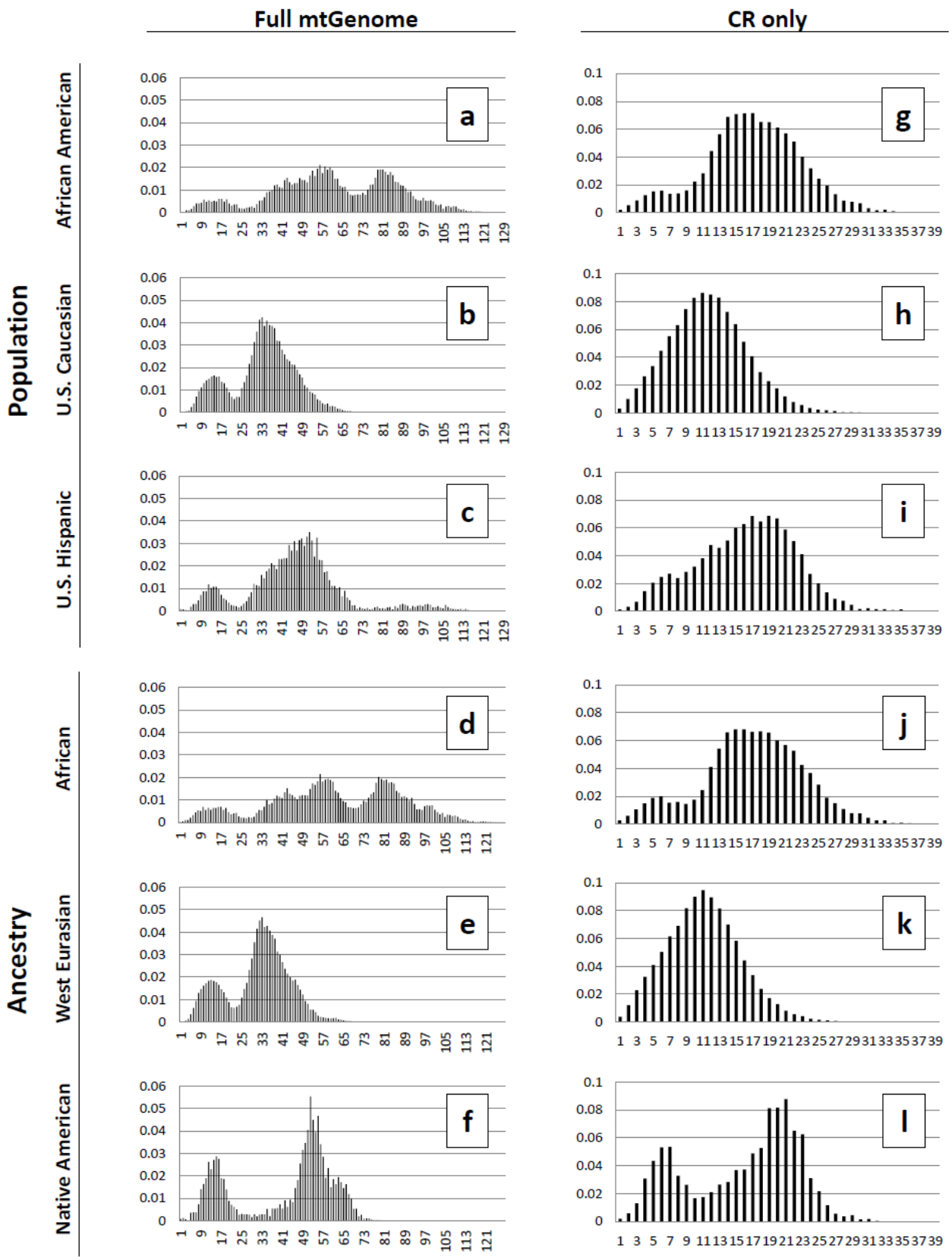


Figure 17. Haplogroup composition of the African American population sample
For the African ancestry haplotypes present in our African American population sample, haplogroups were assigned at a level of phylogenetic resolution that would permit direct comparison to Figure 1 in Salas et al. [108]. The images are strikingly similar.

Figure 18. Haplotype pairwise comparisons (next page)
Pairwise comparisons of the haplotypes were performed for each of the three populations and three of the four biogeographic ancestry groups (African, West Eurasian and Native American). Comparisons for the biogeographic ancestry groups utilized all haplotypes assigned to the ancestry group, regardless of population. The y-axis indicates the proportion of comparisons performed (to normalize for differing sample sizes), and the x-axis represents the number of differences. Histograms on the left side of the figure (panels a through f) represent comparisons performed using the complete mtGenome; whereas for the comparisons on the right side of the figure (g through l), the data compared were restricted to the CR. For all analyses, length insertions at positions 309, 573 and 16193 were ignored.



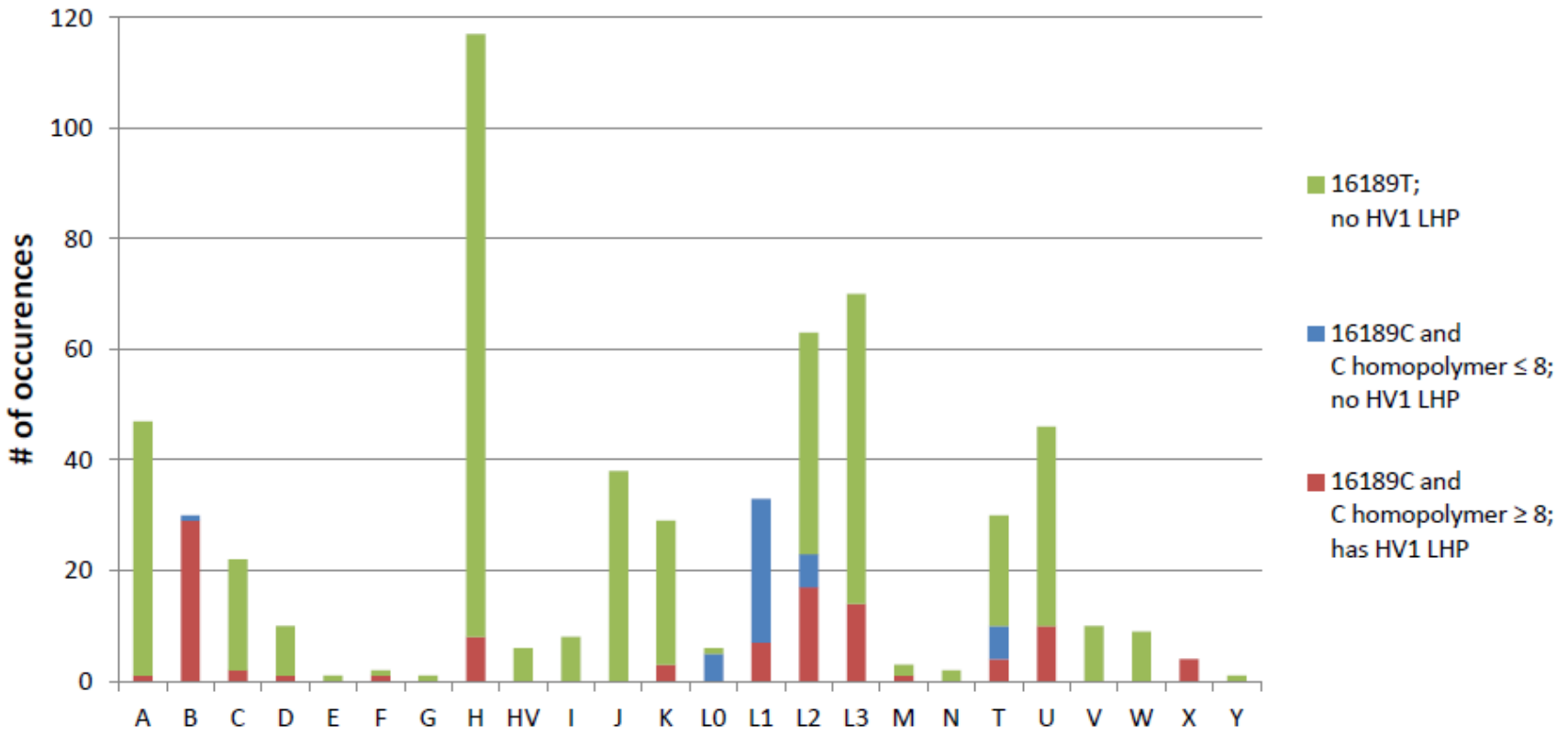
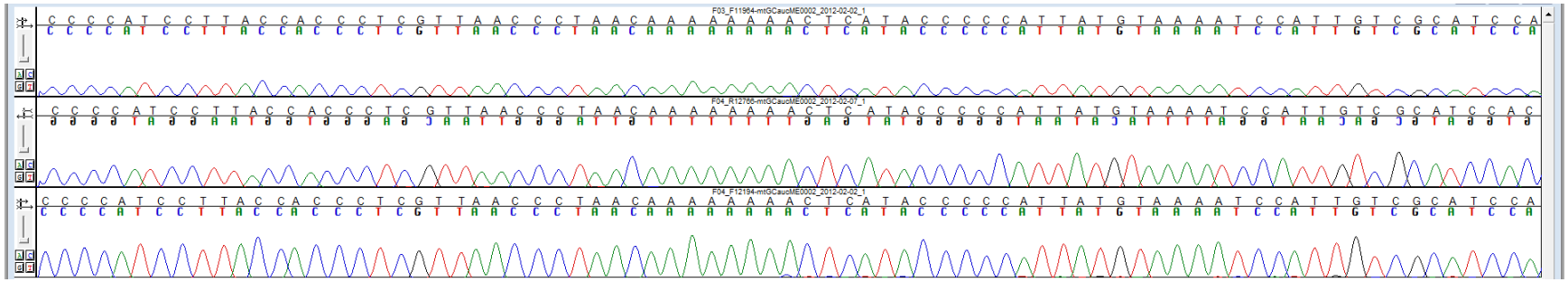


Figure 19. Length heteroplasmy in HV1 by major haplogroup
The observed difference in the frequency of LHP in HV1 by U.S. population in this study (Table 9) is largely explained by the presence or absence of a) the T to C transition at position 16189, and b) a resulting cytosine homopolymer at least 8 bp or longer. Nearly all haplogroup B individuals (the majority of which are from the U.S. Hispanic population sample) and a substantial proportion of haplogroup L1, L2 and L3 individuals (all but two of which are from either the African American or U.S. Hispanic population sample) meet both conditions, while overall fewer samples with West Eurasian haplogroups have the 16189 transition.

A)



B)

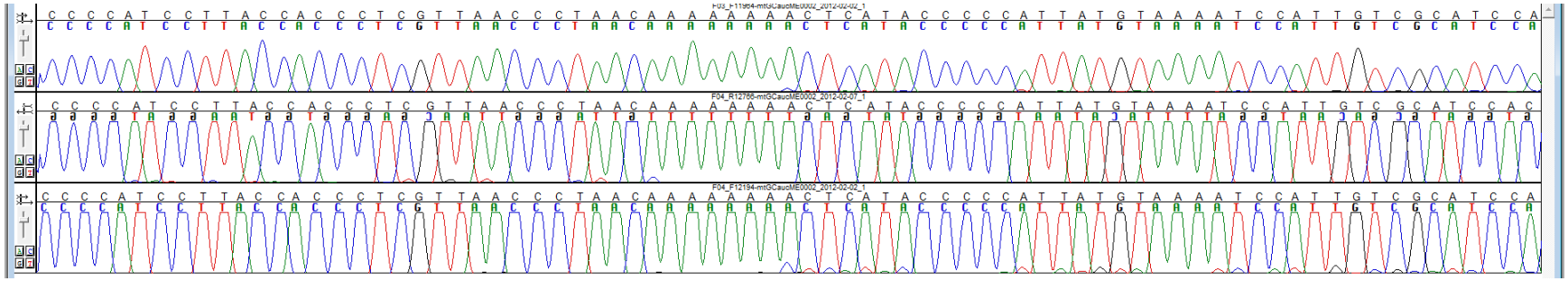


Figure 20. Example of length heteroplasmy in the 12418-12425 adenine homopolymer
LHP in the coding region around position 12425 was detected in almost 90% of samples. Generally a mixture of two molecules (containing seven and eight adenine residues) was observed, and in all cases, the majority molecule matched the rCRS (eight adenines). The LHP was typically very minor, and frequently was only clearly apparent in one sequence in normal view (panel A; see the bottom trace produced using a forward primer). However, when peaks are "pulled up" (panel B), evidence of the LHP can be seen in all three sequences clearly above noise/background (top and bottom traces in panel B were produced using forward primers, and the middle trace was produced using a reverse primer).

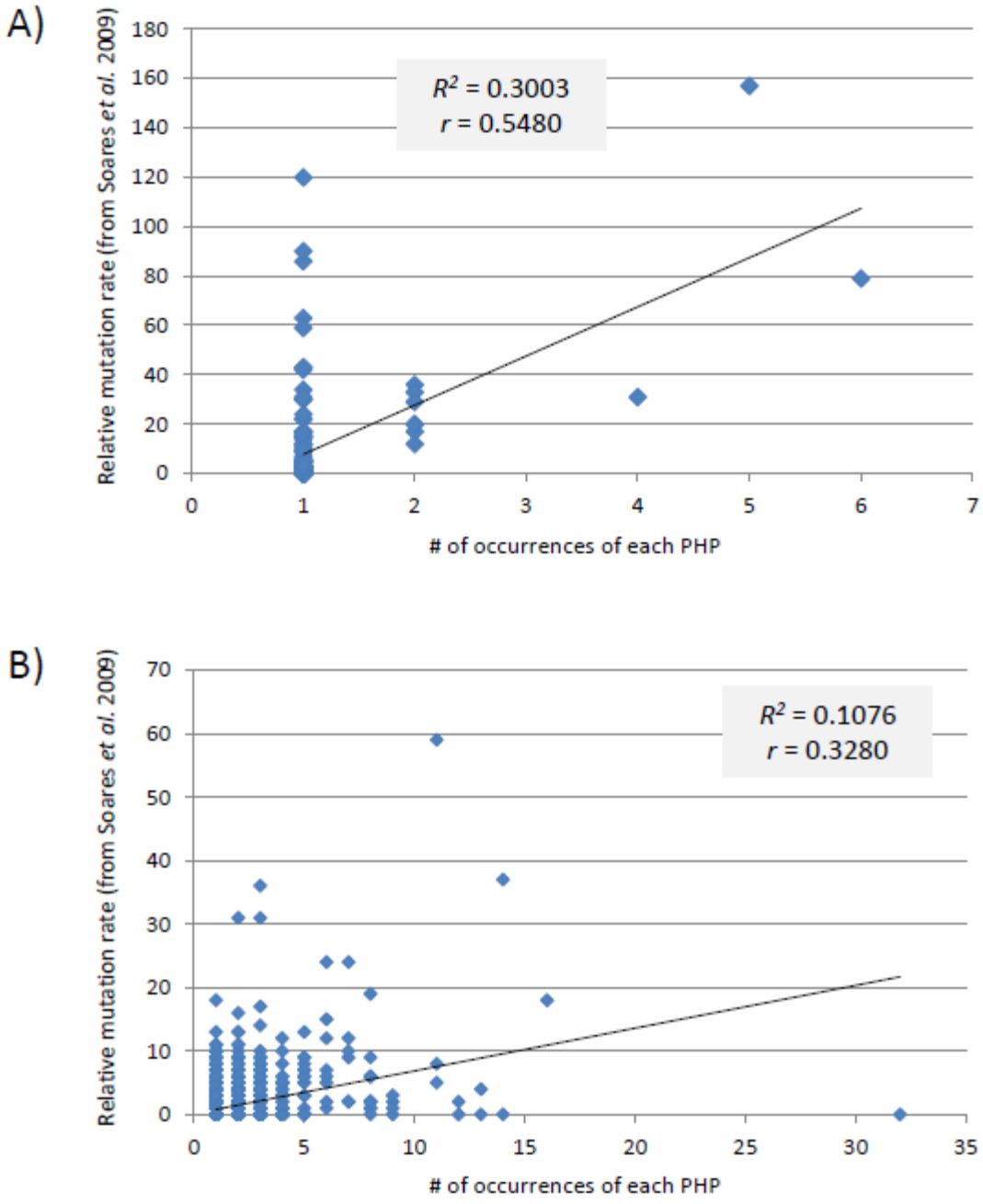


Figure 21. Correlation between PHP observations and relative mutation rates
Plots reflect the relative substitution rate for each PHP observed a) across the entire mtGenome in our study (166 total PHPs) and b) only the coding region PHPs in Ye *et al.* (3547 in total) [131]. The R^2 value for all PHPs across the full mtGenome in this study (a) is similar to the R^2 value reported by Ye *et al.* (compare to their Figure S6B), and the strength of the positive correlation between relative substitution rates and heteroplasmy rates is reduced when the Ye *et al.* analysis was repeated without the CR PHPs.

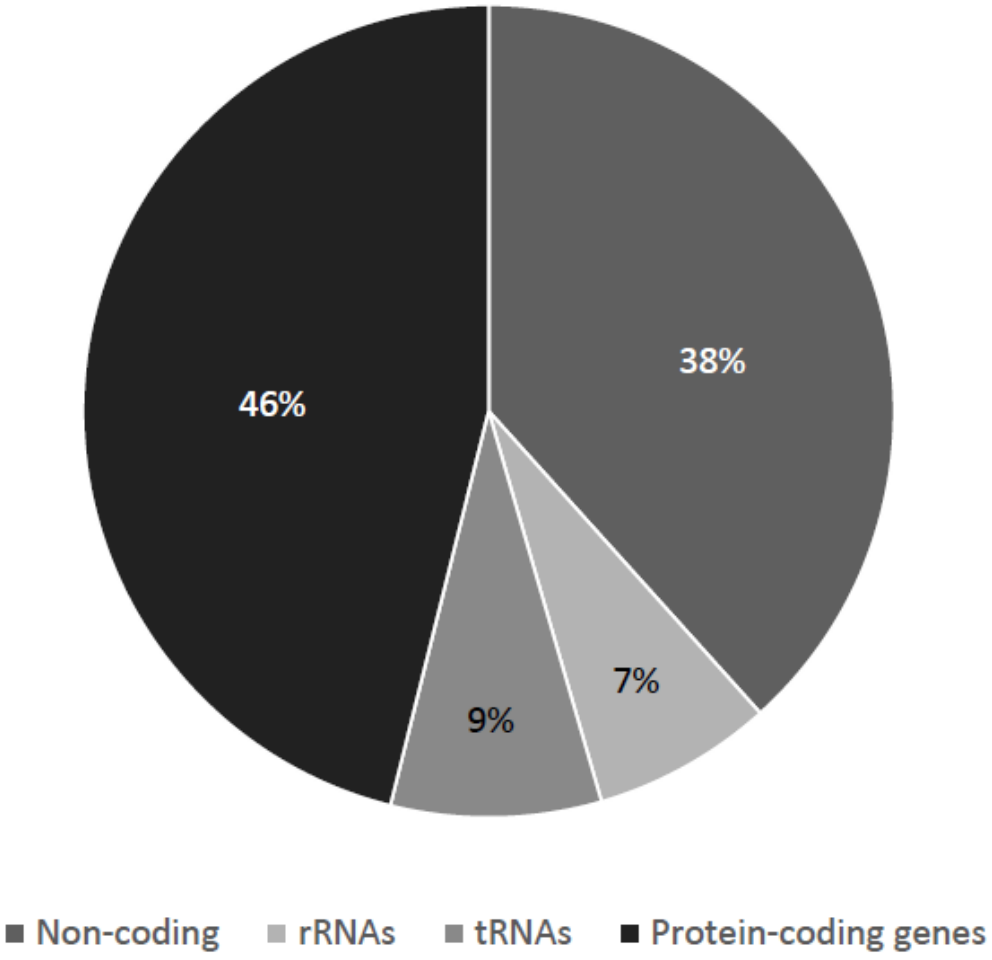


Figure 22. Point heteroplasmies by mtDNA region type
PHPs across all samples were categorized into four regions: non-coding, rRNAs, tRNAs, and protein-coding genes. All PHPs in non-coding regions were found in the CR (that is, no PHPs were observed in the small intergenic non-coding regions).

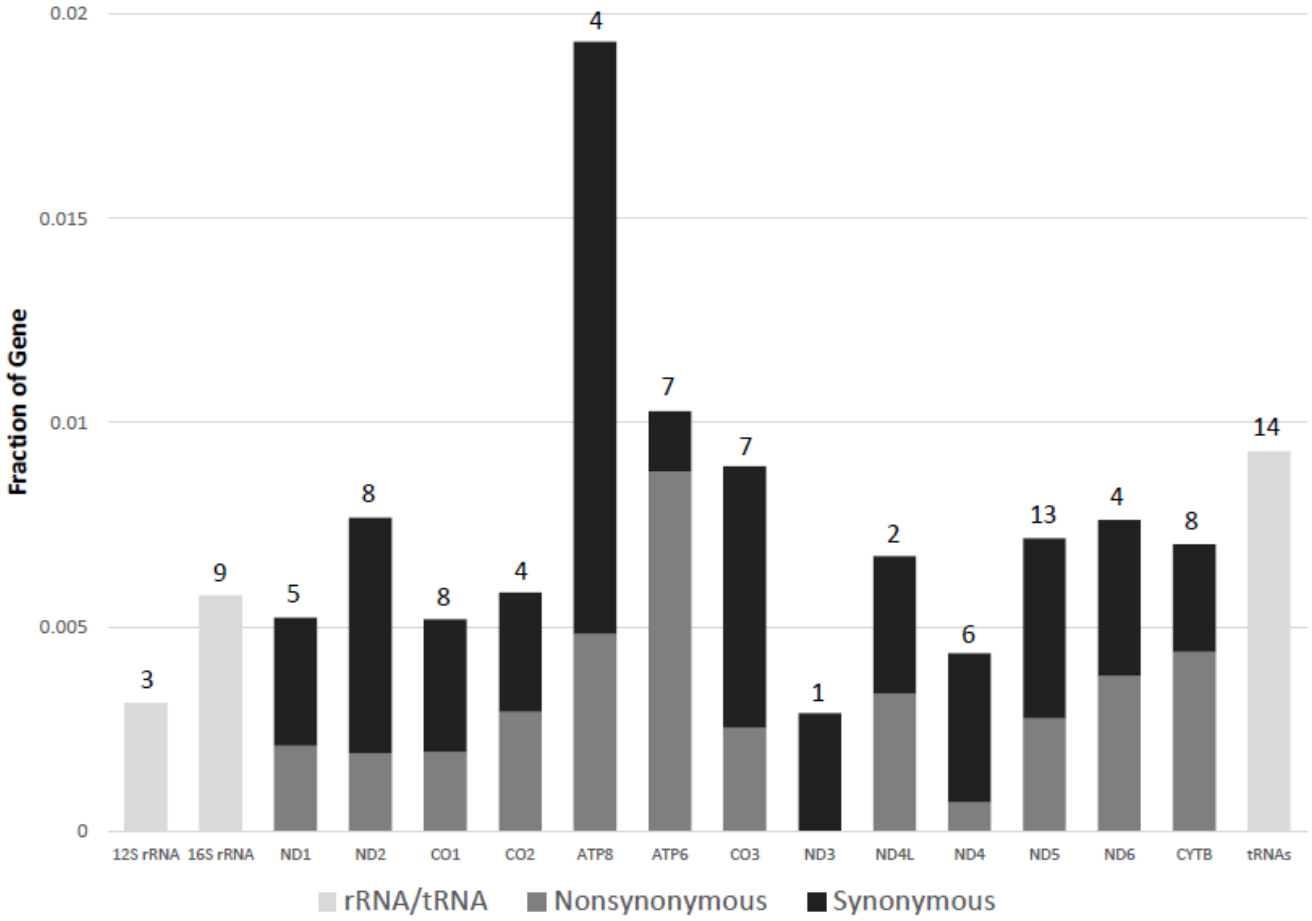


Figure 23. Point heteroplasmy proportions by gene
PHPs across all protein-coding genes plus the two rRNAs and all tRNAs (combined) were plotted by the fraction of potential positions (size of the gene) at which PHPs were observed. Thus, the height of each bar in the histogram indicates the relative rate of mutation observed for each gene. The actual number of PHPs observed for each gene are indicated above the bars. The mutations in the 13 protein-coding genes were categorized as to synonymous or nonsynonymous amino acid changes if the mutations were to become fixed.

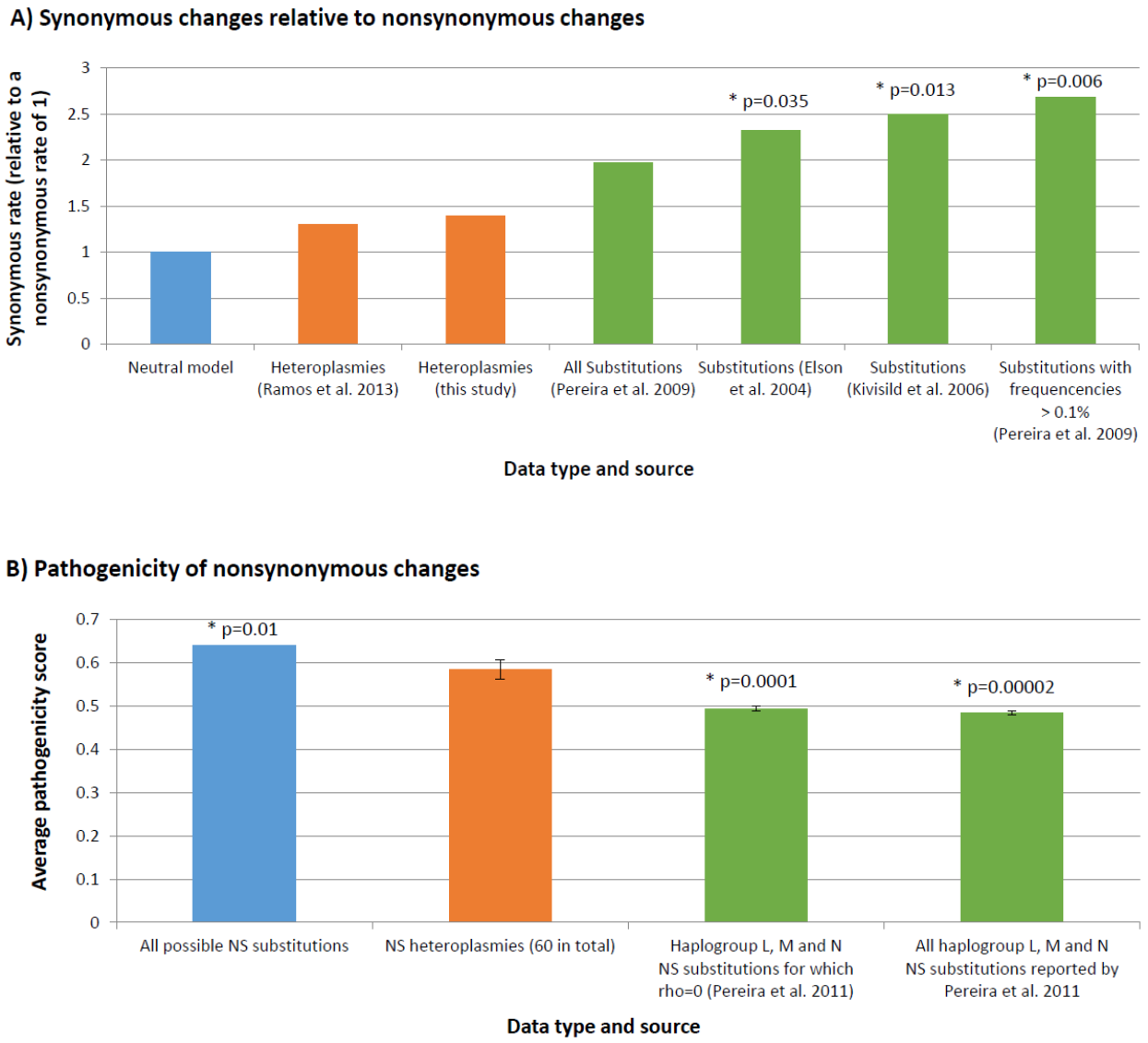


Figure 24. Synonymous change rates and pathogenicity scores for heteroplasmies versus complete substitutions

Panel A displays the synonymous change rates (relative to a nonsynonymous change rate of 1) for heteroplasmies (from this study and Ramos *et al.* [26]) and substitutions analyzed in previous studies [28-30]. Panel B displays the average pathogenicity scores (based on MutPred values [136] reported by Pereira *et al.* [31]) for all possible substitutions, the 60 total nonsynonymous heteroplasmies detected in our haplotypes and reported in three previous studies [23,26,27], and complete substitutions analyzed by Pereira *et al.* [31]. In each panel, asterisks represent statistically significant differences from the heteroplasmies (orange bars).

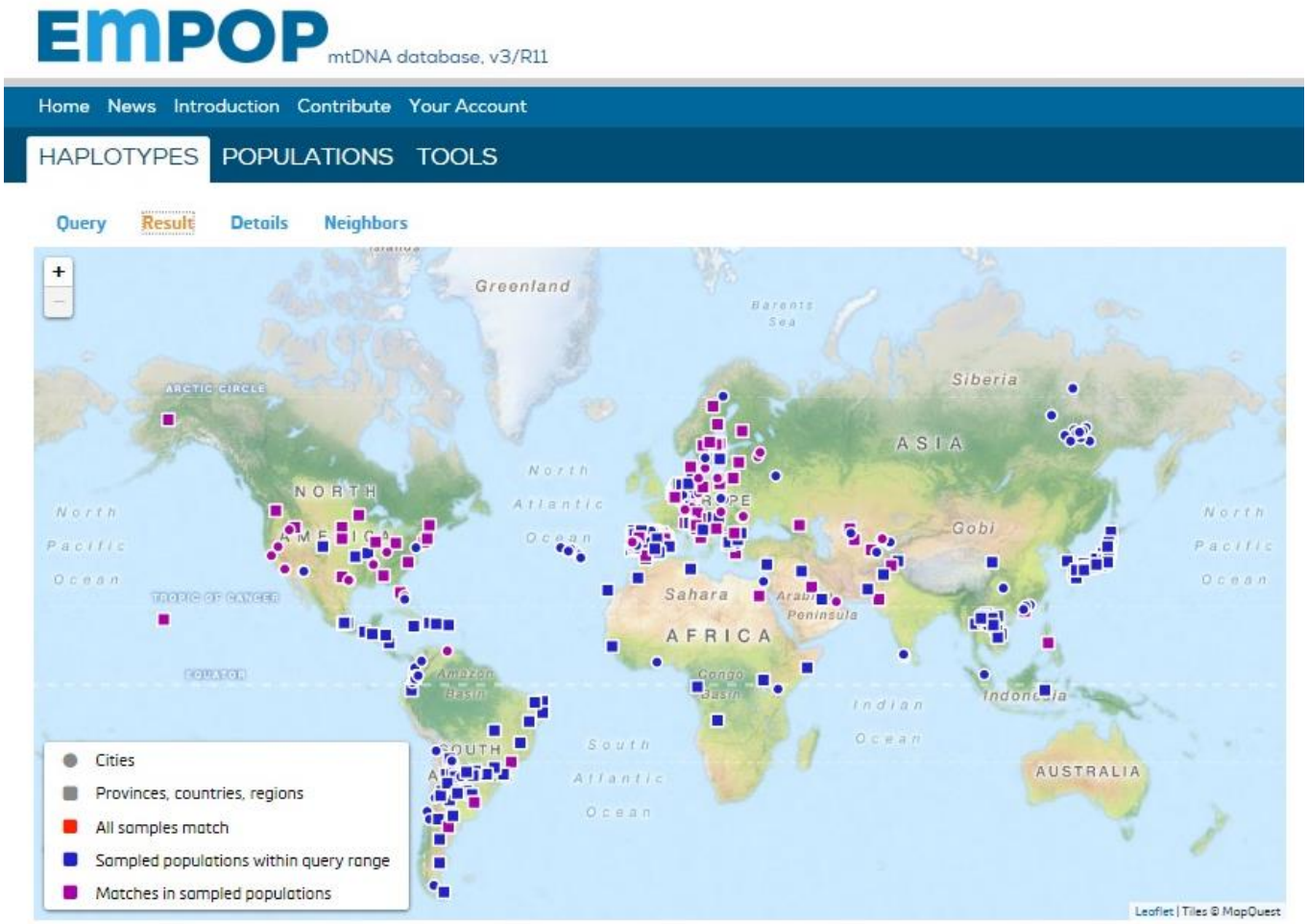
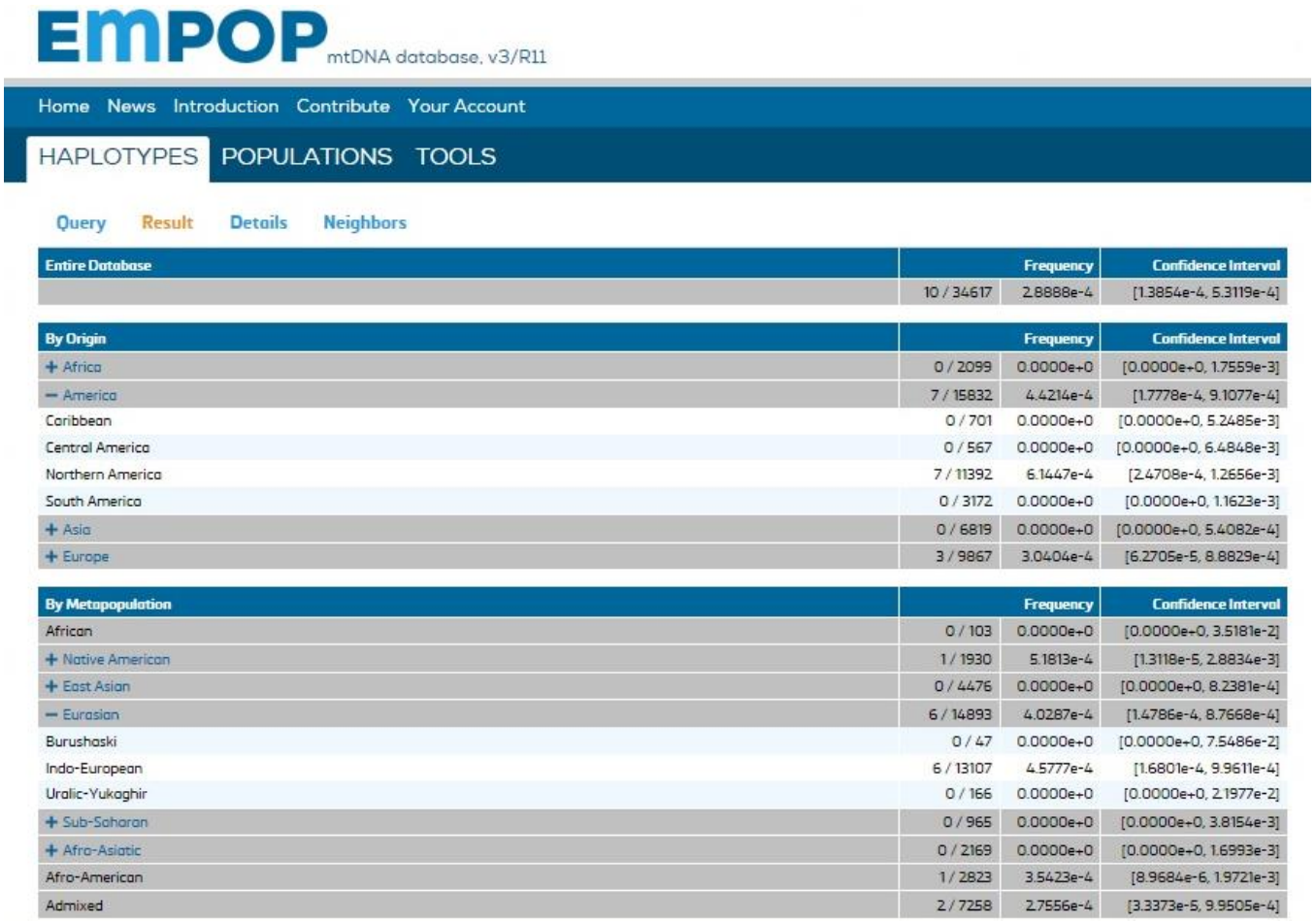


Figure 25. Map display of queried datasets
This screen shot from EMPOP3 shows the visual representation of the datasets queried for a given haplotype. Both sampled populations as well as haplotype matches are displayed.



EMPOP mtDNA database, v3/R11

Home News Introduction Contribute Your Account

HAPLOTYPES POPULATIONS TOOLS

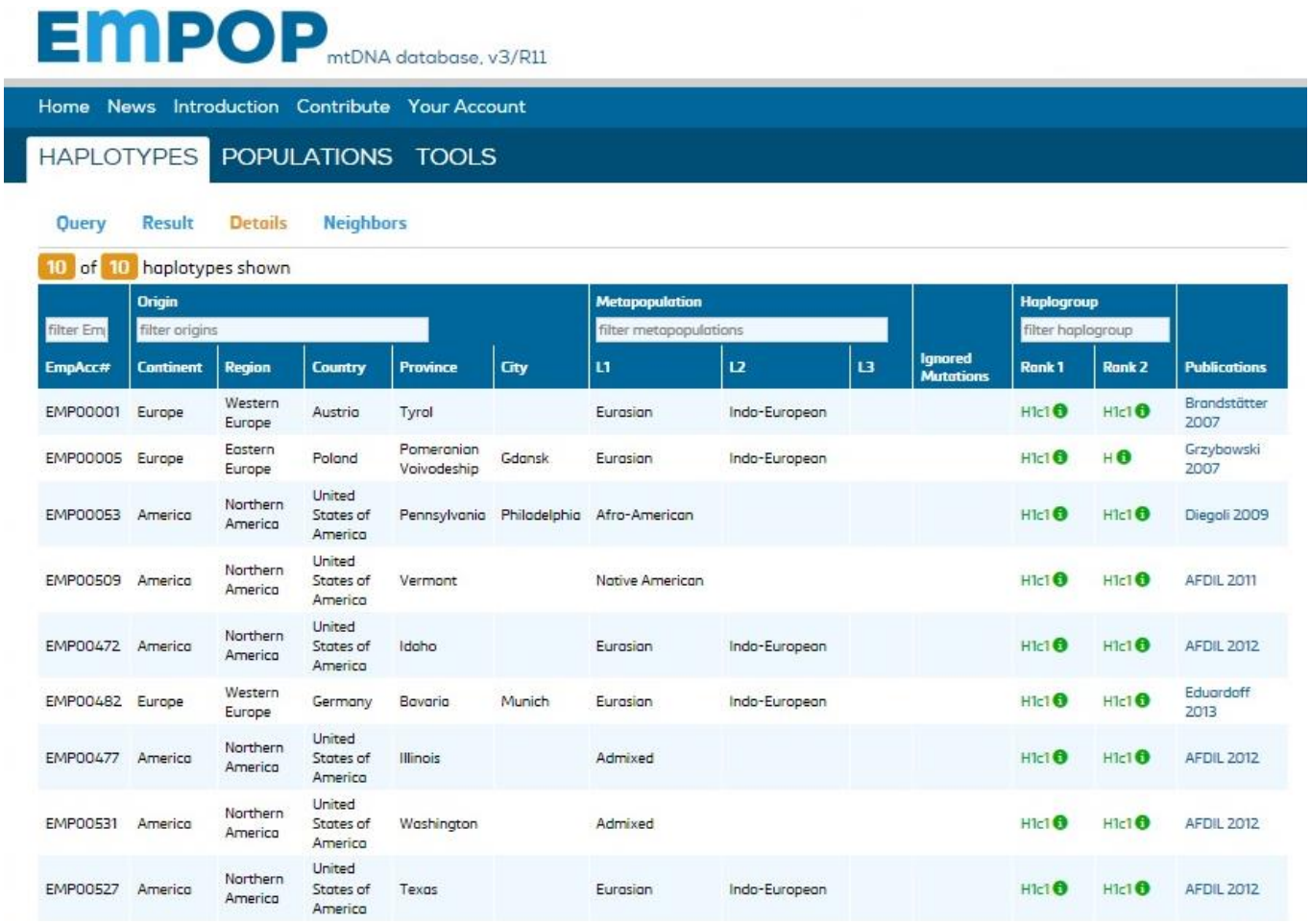
Query Result Details Neighbors

Entire Database		Frequency	Confidence Interval
	10 / 34617	2.8888e-4	[1.3854e-4, 5.3119e-4]

By Origin		Frequency	Confidence Interval
+ Africa	0 / 2099	0.0000e+0	[0.0000e+0, 1.7559e-3]
− America	7 / 15832	4.4214e-4	[1.7778e-4, 9.1077e-4]
Caribbean	0 / 701	0.0000e+0	[0.0000e+0, 5.2485e-3]
Central America	0 / 567	0.0000e+0	[0.0000e+0, 6.4848e-3]
Northern America	7 / 11392	6.1447e-4	[2.4708e-4, 1.2656e-3]
South America	0 / 3172	0.0000e+0	[0.0000e+0, 1.1623e-3]
+ Asia	0 / 6819	0.0000e+0	[0.0000e+0, 5.4082e-4]
+ Europe	3 / 9867	3.0404e-4	[6.2705e-5, 8.8829e-4]

By Metapopulation		Frequency	Confidence Interval
African	0 / 103	0.0000e+0	[0.0000e+0, 3.5181e-2]
+ Native American	1 / 1930	5.1813e-4	[1.3118e-5, 2.8834e-3]
+ East Asian	0 / 4476	0.0000e+0	[0.0000e+0, 8.2381e-4]
− Eurasian	6 / 14893	4.0287e-4	[1.4786e-4, 8.7668e-4]
Burushaski	0 / 47	0.0000e+0	[0.0000e+0, 7.5486e-2]
Indo-European	6 / 13107	4.5777e-4	[1.6801e-4, 9.9611e-4]
Uralic-Yukaghir	0 / 166	0.0000e+0	[0.0000e+0, 2.1977e-2]
+ Sub-Saharan	0 / 965	0.0000e+0	[0.0000e+0, 3.8154e-3]
+ Afro-Asiatic	0 / 2169	0.0000e+0	[0.0000e+0, 1.6993e-3]
Afro-American	1 / 2823	3.5423e-4	[8.9684e-6, 1.9721e-3]
Admixed	2 / 7258	2.7556e-4	[3.3373e-5, 9.9505e-4]

Figure 26. Tabular format for haplotype query result
Screen capture of EMPOP3 displaying the number of matches, haplotype frequency estimates and 95% confidence intervals for datasets categorized by continental origin and metapopulation. Origin and metapopulation categories are expandable to display forensically-relevant region and linguistic subcategories.



	Origin					Metapopulation				Haplogroup		
filter Em	filter origins					filter metapopulations				filter haplogroup		
EmpAcc#	Continent	Region	Country	Province	City	L1	L2	L3	Ignored Mutations	Rank 1	Rank 2	Publications
EMP00001	Europe	Western Europe	Austria	Tyrol		Eurasian	Indo-European			H1c1	H1c1	Brandstätter 2007
EMP00005	Europe	Eastern Europe	Poland	Pomeranian Voivodeship	Gdansk	Eurasian	Indo-European			H1c1	H	Grzybowski 2007
EMP00053	America	Northern America	United States of America	Pennsylvania	Philadelphia	Afro-American				H1c1	H1c1	Diegoli 2009
EMP00509	America	Northern America	United States of America	Vermont		Native American				H1c1	H1c1	AFDIL 2011
EMP00472	America	Northern America	United States of America	Idaho		Eurasian	Indo-European			H1c1	H1c1	AFDIL 2012
EMP00482	Europe	Western Europe	Germany	Bavaria	Munich	Eurasian	Indo-European			H1c1	H1c1	Eduardoff 2013
EMP00477	America	Northern America	United States of America	Illinois		Admixed				H1c1	H1c1	AFDIL 2012
EMP00531	America	Northern America	United States of America	Washington		Admixed				H1c1	H1c1	AFDIL 2012
EMP00527	America	Northern America	United States of America	Texas		Eurasian	Indo-European			H1c1	H1c1	AFDIL 2012

Figure 27. Haplotype query result details
Detailed information on haplotype matches in the database, including clickable Rank 1 and Rank 2 haplogroups for each match, are displayed in dynamic format in EMPOP3. Categories can be filtered, and the various columns can be sorted.

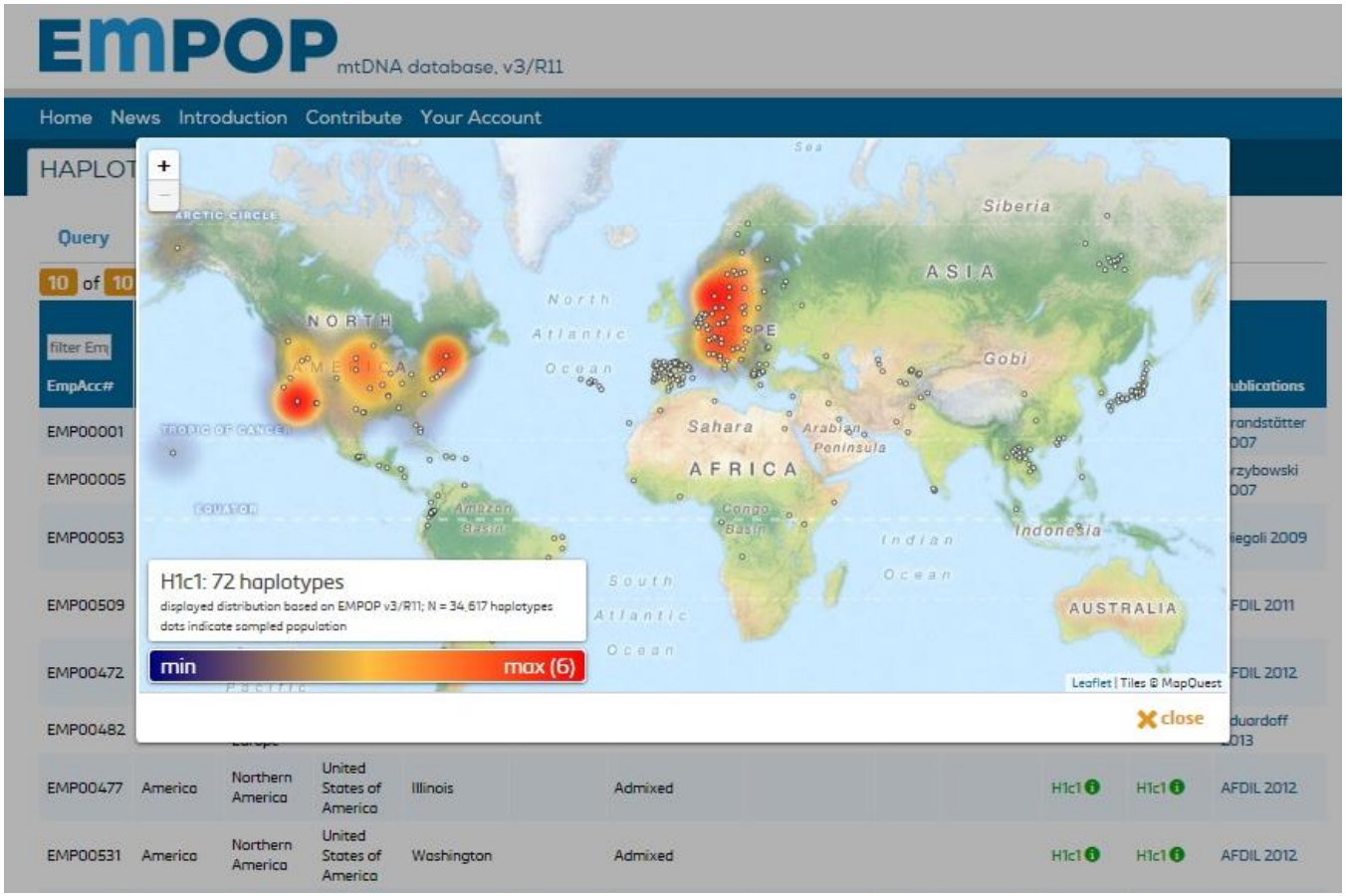
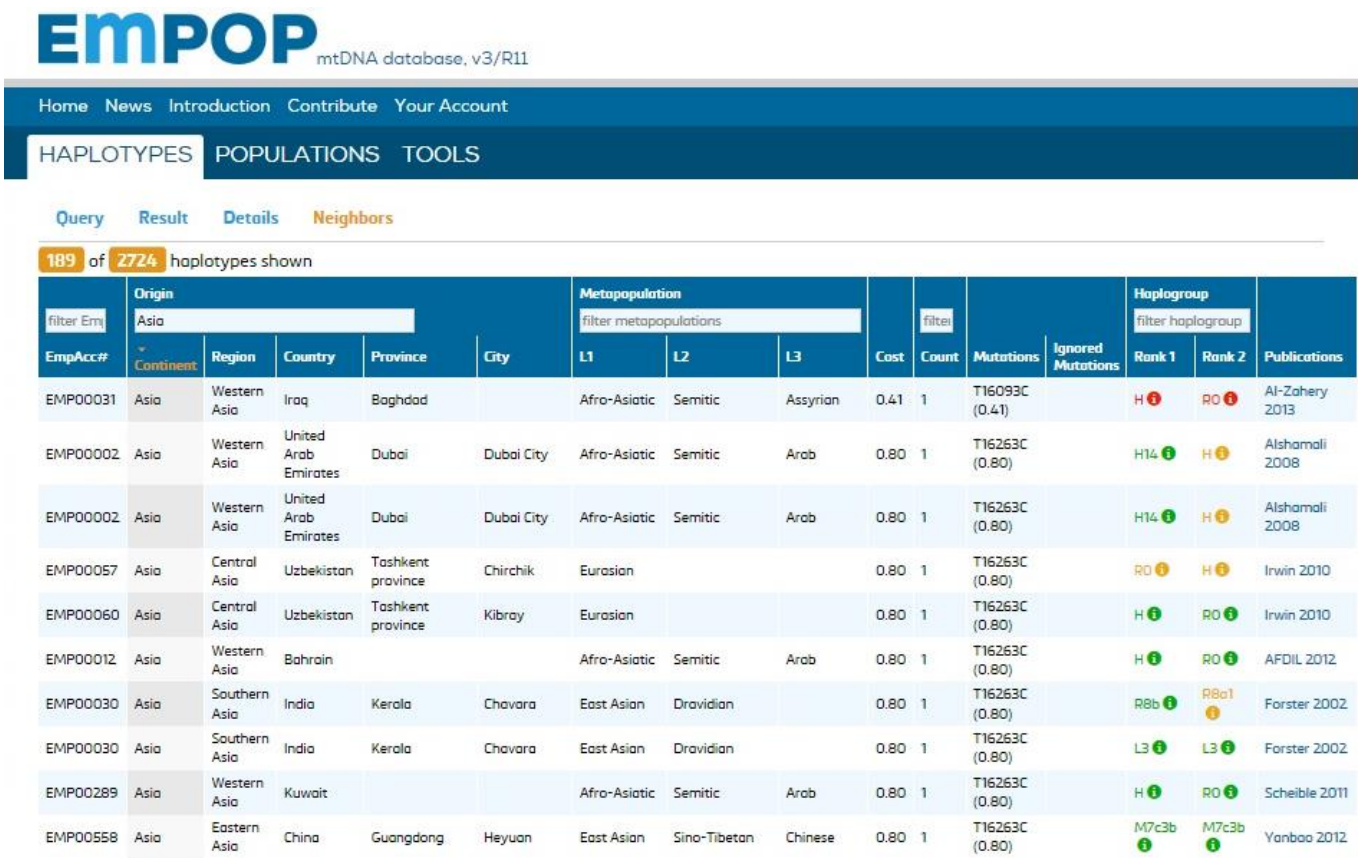


Figure 28. Pop-up heatmap display
Clicking on a haplogroup link in the details tab of the query results will display a pop-up heat map showing the global distribution of the haplogroup based on the haplotypes stored in EMPOP.



EMPOP mtDNA database, v3/R11

Home News Introduction Contribute Your Account

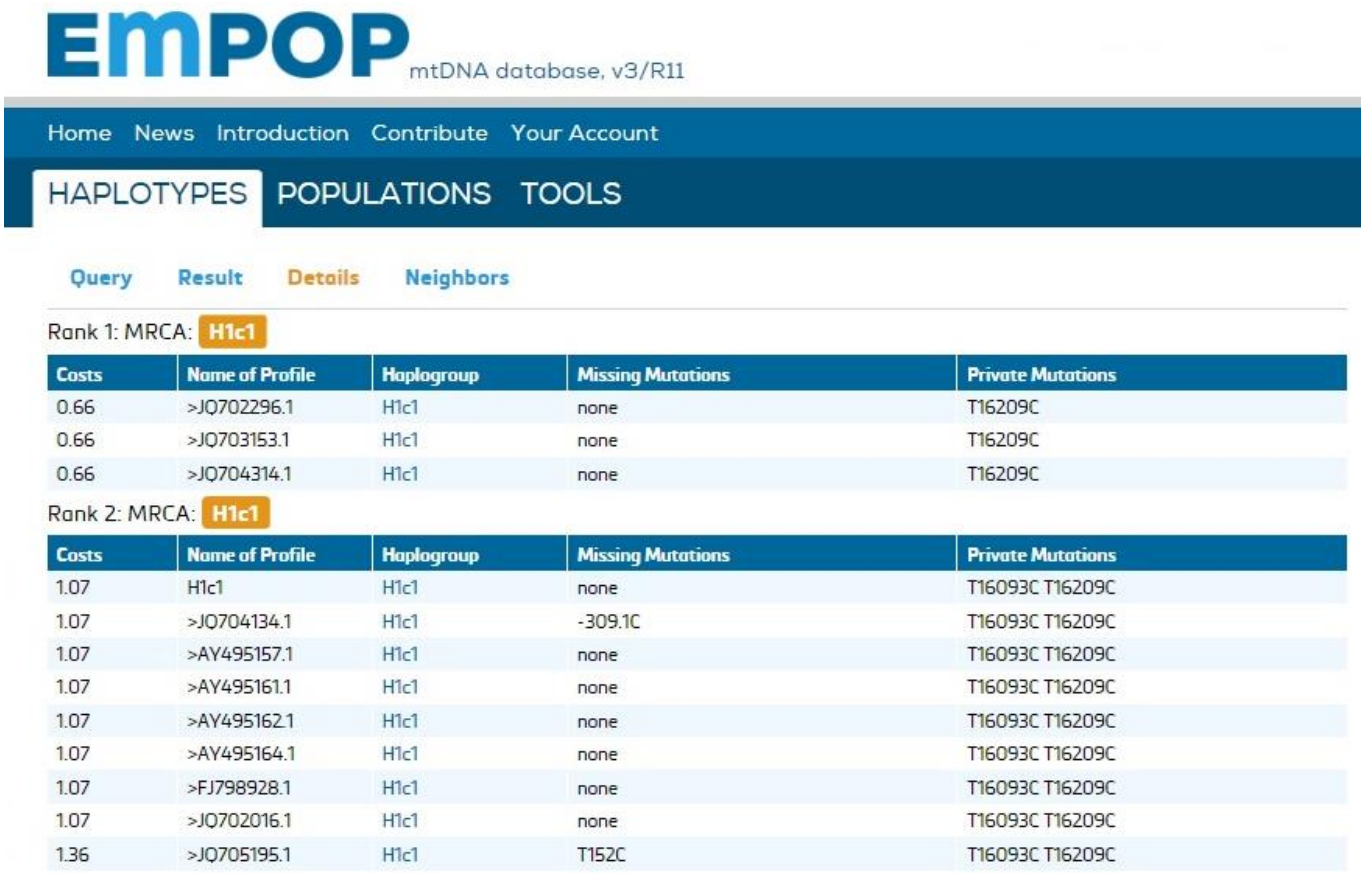
HAPLOTYPES POPULATIONS TOOLS

Query Result Details Neighbors

189 of 2724 haplotypes shown

	Origin					Metapopulation							Haplogroup		
filter Em	Asia					filter metapopulations				filter			filter haplogroup		
EmpAcc#	Continent	Region	Country	Province	City	L1	L2	L3	Cost	Count	Mutations	Ignored Mutations	Rank 1	Rank 2	Publications
EMP00031	Asia	Western Asia	Iraq	Baghdad		Afro-Asiatic	Semitic	Assyrian	0.41	1	T16093C (0.41)		H	R0	Al-Zahery 2013
EMP00002	Asia	Western Asia	United Arab Emirates	Dubai	Dubai City	Afro-Asiatic	Semitic	Arab	0.80	1	T16263C (0.80)		H14	H	Alshamali 2008
EMP00002	Asia	Western Asia	United Arab Emirates	Dubai	Dubai City	Afro-Asiatic	Semitic	Arab	0.80	1	T16263C (0.80)		H14	H	Alshamali 2008
EMP00057	Asia	Central Asia	Uzbekistan	Tashkent province	Chirchik	Eurasian			0.80	1	T16263C (0.80)		R0	H	Irwin 2010
EMP00060	Asia	Central Asia	Uzbekistan	Tashkent province	Kibray	Eurasian			0.80	1	T16263C (0.80)		H	R0	Irwin 2010
EMP00012	Asia	Western Asia	Bahrain			Afro-Asiatic	Semitic	Arab	0.80	1	T16263C (0.80)		H	R0	AFDIL 2012
EMP00030	Asia	Southern Asia	India	Kerala	Chavara	East Asian	Dravidian		0.80	1	T16263C (0.80)		R8b	R8a1	Forster 2002
EMP00030	Asia	Southern Asia	India	Kerala	Chavara	East Asian	Dravidian		0.80	1	T16263C (0.80)		L3	L3	Forster 2002
EMP00289	Asia	Western Asia	Kuwait			Afro-Asiatic	Semitic	Arab	0.80	1	T16263C (0.80)		H	R0	Scheible 2011
EMP00558	Asia	Eastern Asia	China	Guangdong	Heyuan	East Asian	Sino-Tibetan	Chinese	0.80	1	T16263C (0.80)		M7c3b	M7c3b	Yanbao 2012

Figure 29. Neighbors to haplotype matches
In EMPOP3, detailed information on database haplotypes which neighbor the searched haplotype can also be examined.



EMPOP mtDNA database, v3/R11

Home News Introduction Contribute Your Account

HAPLOTYPES POPULATIONS TOOLS

Query Result Details Neighbors

Rank 1: MRCA: H1c1

Costs	Name of Profile	Haplogroup	Missing Mutations	Private Mutations
0.66	>JQ702296.1	H1c1	none	T16209C
0.66	>JQ703153.1	H1c1	none	T16209C
0.66	>JQ704314.1	H1c1	none	T16209C

Rank 2: MRCA: H1c1

Costs	Name of Profile	Haplogroup	Missing Mutations	Private Mutations
1.07	H1c1	H1c1	none	T16093C T16209C
1.07	>JQ704134.1	H1c1	-309.1C	T16093C T16209C
1.07	>AY495157.1	H1c1	none	T16093C T16209C
1.07	>AY495161.1	H1c1	none	T16093C T16209C
1.07	>AY495162.1	H1c1	none	T16093C T16209C
1.07	>AY495164.1	H1c1	none	T16093C T16209C
1.07	>FJ798928.1	H1c1	none	T16093C T16209C
1.07	>JQ702016.1	H1c1	none	T16093C T16209C
1.36	>JQ705195.1	H1c1	T152C	T16093C T16209C

Figure 30. Data displayed when no matching haplotypes are found in the database
When a haplotype search results in no matching profiles in EMPOP, the highest ranking haplogroups are displayed, along with details on the lowest cost database haplotypes assigned to each haplogroup.

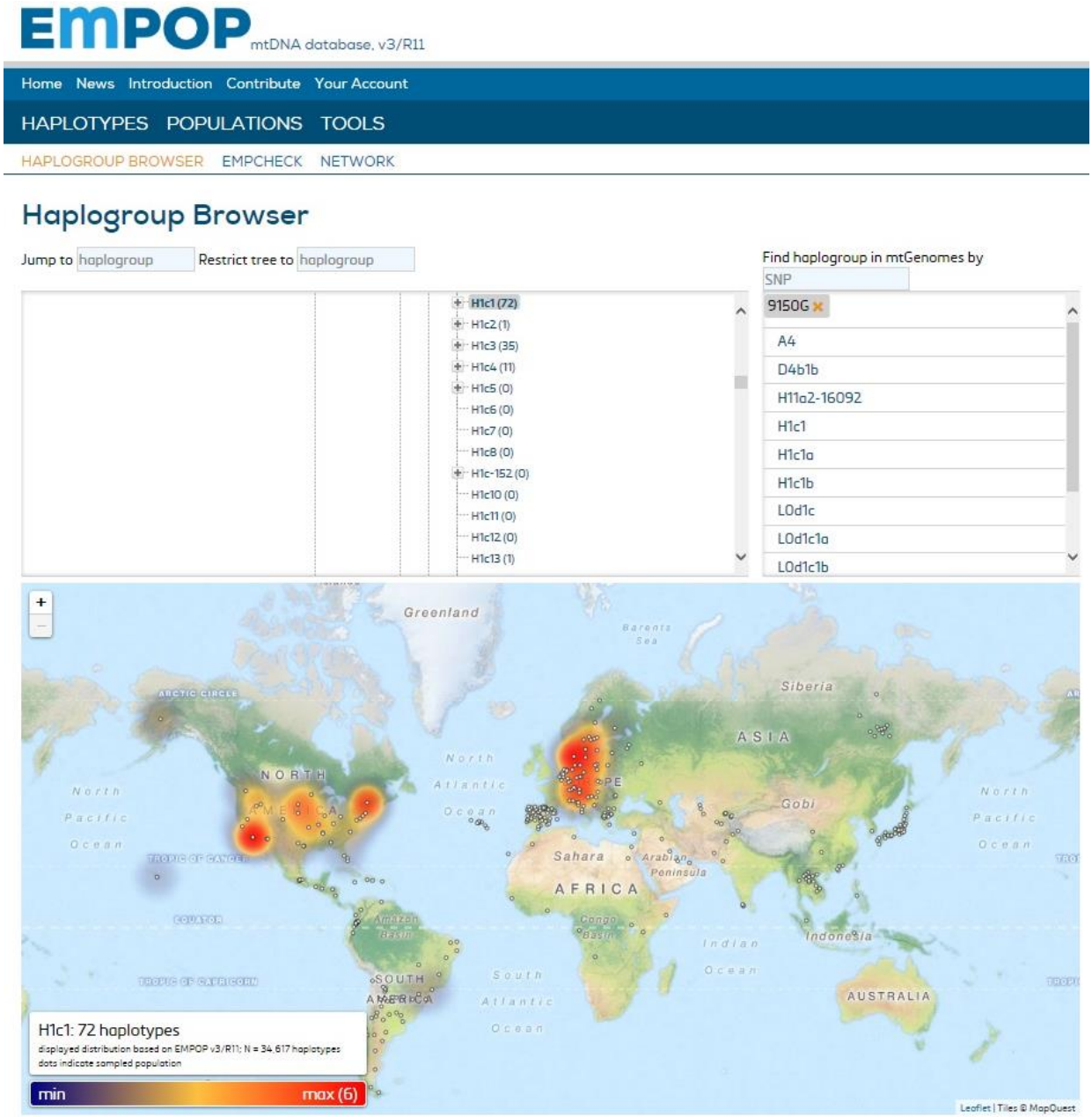


Figure 31. Haplogroup browser
The dynamic haplogroup browser in EMPOP3 enables users to examine portions of the mtDNA phylogenetic tree by searching for named haplogroups, individual mutations, or sets of mutations. Search results are displayed tabularly, and clicking on a haplogroup link displays a heat map for global distribution of the haplogroup among the profiles stored in the database.

Appendix B. Tables

Amplicon Number	Number of nuclear genome matches	Nuclear genome location	GenBank Accession Number	Priming region similarity (forward; reverse)	Size difference from mtDNA target (in bp)	Sequence similarity to rCRS
1	1	Chr 11 (9510471–9507925)	NT_009237.18	76.2%; 76.2%	170	< 60%
2	1	Chr 6 (53426379–53424135)	NT_025741.15	84.2%; 78.9%	118	< 60%
3	1	Chr 1 (43791–46082)	NT_004350.19	100%; 100%	1	98.47%
4	2	Chr 1 (45818–48327)	NT_004350.19	95.5%; 100%	1	98.57%
		Chr 5 (7704044–7701575)	NT_034772.6	95.5%; 90.9%	1	88.53%
5	2	Chr 7 (6847950–6850445)	NT_033968.6	84.0%; 88.9%	7	75.16%
		Chr 2 (10717704–10720523)	NT_022135.16	80.0%; 88.9%	331	63.40%
6	2	Chr 5 (42577018–42574300)	NT_034772.6	95.5%; 100%	0	94.02%
		Chr 5 (7699358–7696640)	NT_034772.6	90.5%; 90.0%	0	88.97%
7	4	Chr 5 (2218206–2220412)	NT_034772.6	95.2%; 90.0%	1	87.27%
		Chr 5 (2218173–2220412)	NT_034772.6	81.0%; 90.0%	32	87.19%
		Chr 17 (13111672–13109937)	NT_010718.16	80.0%; 85.0%	453	< 60%
		Chr 7 (5831689–5829543)	NT_033968.6	81.0%; 75.0%	41	< 60%

Table 1. Amplicon PrimerBLAST results
Coding region amplification primer pairs were queried against the reference assembly of the complete human genome using PrimerBLAST [88] and results which met specific similarity criteria were noted. For these 13 regions of the nuclear genome which are potentially amplifiable using the coding region PCR primer pairs listed in Appendix C, the nuclear genome sequence was aligned to the rCRS to determine a percentage sequence similarity. For Amplicon 7, the two Chromosome 5 matches represent slightly different primer binding sites within the same portion of the chromosome.

Sample	mtDNA Amplicon	mtDNA Nucleotide Range	mtDNA Coverage	nDNA Chromosome	nDNA Nucleotide Range	Amplified NUMT Similarity to GRCh38 Reference Assembly
mtGHispCA0019	3	4609-6899	4609-6898	1	629779-632070	99.91%
mtGHispDE0001	3	4609-6899	4616-6893	1	629779-632070	99.56%
mtGHispFL0005	3	4609-6899	4616-6889	1	629779-632070	99.96%
	4	6636-9146	6650-9145	1	631806-634315	99.44%
	6	11319-14077	11320-14074	5	134927477-134924719	96.81%
mtGHispTX0015	6	11319-14077	11323-14074	5	134927477-134924719	97.82%
mtGHispTX0043	6	11319-14077	11323-14071	5	134927477-134924719	96.69%

Table 2. Summary of NUMT co-amplification

NUMT co-amplification was detected in five samples. In samples mtGHispCA0019, mtGHispDE0001, mtGHispFL0005 and mtGHispTX0015, the contamination was detected during initial review of the raw sequencing data. The co-amplification event in mtGHispTX0043 was not detected until cross-check of the AFDIL and EMPOP haplotypes, when additional data was sought to investigate a potential very low level PHP. Four of the five affected samples were re-extracted and reprocessed to remove the NUMT co-amplification, and were utilized in the final haplotype database. One sample, mtGHispTX0015, was not reprocessed due to limited sample material, and is not represented in the database. The percent similarity of the co-amplified NUMT to the human nuclear genome reference assembly was determined by BLAST [88] results using the number of matching nucleotides. Differences in percent similarity are due to differences in visibility of co-amplified sequences in the affected samples.

African American (n=170)					Percentage Increase			
	HV1	HV1/HV2	CR	mtG	HV1 to HV1/HV2	HV1/HV2 to CR	CR to mtG	HV1 to mtG
# Haplotypes	124	140	148	169	12.9%	5.7%	14.2%	71.5%
# Unique Haplotypes	106	120	130	168	13.2%	8.3%	29.2%	58.5%
Random Match Probability	1.38%	0.92%	0.78%	0.60%				
Haplotype Diversity	0.9920	0.9967	0.9981	0.9999				
Power of Discrimination	99.20%	99.67%	99.81%	99.99%				
U.S. Caucasian (n=263)					Percentage Increase			
	HV1	HV1/HV2	CR	mtG	HV1 to HV1/HV2	HV1/HV2 to CR	CR to mtG	HV1 to mtG
# Haplotypes	151	200	229	259	32.5%	14.5%	13.1%	71.5%
# Unique Haplotypes	122	170	211	255	39.3%	24.1%	20.9%	109.0%
Random Match Probability	2.75%	0.96%	0.60%	0.39%				
Haplotype Diversity	0.9762	0.9942	0.9978	0.9999				
Power of Discrimination	97.62%	99.42%	99.78%	99.99%				
U.S. Hispanic (n=155)					Percentage Increase			
	HV1	HV1/HV2	CR	mtG	HV1 to HV1/HV2	HV1/HV2 to CR	CR to mtG	HV1 to mtG
# Haplotypes	119	134	141	147	12.6%	5.2%	4.3%	23.5%
# Unique Haplotypes	102	121	130	140	18.6%	7.4%	7.7%	37.3%
Random Match Probability	1.27%	0.90%	0.79%	0.72%				
Haplotype Diversity	0.9937	0.9974	0.9986	0.9992				
Power of Discrimination	99.37%	99.74%	99.86%	99.92%				

Table 3. Summary statistics

Summary statistics were calculated for each of the three U.S. populations for several regions historically targeted for forensic typing: HV1 alone, HV1 and HV2 in combination, the entire CR, and the full mtGenome. Haplotype diversity was calculated as (1-Random Match Probability)*((n-1)/n). The "Percentage Increase" (in the number of distinct haplotypes and the number of haplotypes unique in each population) represents the gain in information moving from one portion of the molecule to a larger portion of the molecule. For instance, in the African American population, the number of haplotypes observed increases from 124 when only HV1 is examined to 140 when HV1 and HV2 are examined in combination. The 16 additional haplotypes observed with HV1/HV2 data represents an increase of 12.9% when compared to the number of haplotypes observed with HV1 alone. These percentage increases are given for each successively larger portion of the molecule and for the smallest portion of the molecule (HV1) versus the entire molecule.

	Clopper-Pearson				Brenner kappa								
	1-tailed		2-tailed		HV1/HV2			CR			Full mtGenome		
	95% CI	LR	Upper 95% CI	LR	Singletons	κ	LR_κ	Singletons	κ	LR_κ	Singletons	κ	LR_κ
This study													
African American (n=170)	0.0175	57	0.0215	47	120	0.7059	578	130	0.7647	723	168	0.9882	14450
U.S. Caucasian (n=263)	0.0113	88	0.0139	72	170	0.6464	744	211	0.8023	1330	255	0.9696	8646
U.S. Hispanic (n=155)	0.0191	52	0.0235	43	121	0.7806	707	130	0.8387	961	140	0.9032	1602
King et al. 2014													
Texas African American (n=87)	0.0338	30	0.0415	24	76	0.8736	688				85	0.9770	3785
Texas Caucasian (n=83)	0.0354	28	0.0435	23	77	0.9277	1148				83	>0.99*	9222
Texas Hispanic (n=113)	0.0262	38	0.0321	31	96	0.8496	751				111	0.9823	6384

*As modeled in Brenner 2010 to avoid κ=1

Table 4. Likelihood ratios for unobserved haplotypes using two different methods

Clopper-Pearson 95% confidence intervals [19] and Brenner's "kappa method" [20] were used to calculate LRs for a haplotype not present in the database, for a) both the three population samples reported in this study and the three population samples reported by King *et al.* [23], and b) given different portions of the mtGenome. As defined by Brenner [20], κ refers to the proportion of singletons (unique haplotypes) in the database. The number of singletons in the King *et al. [23]* datasets were obtained from their Table 1 (no full CR values were reported).

Haplogroup	African American	U.S. Caucasian	U.S. Hispanic
A	**1.8%**	**1.1%**	**26.5%**
A2 (NA)	*1.2%*	*0.8%*	*26.5%*
A5, A10 (EA)	*0.6%*	*0.4%*	
B	**0.6%**	**1.5%**	**16.1%**
B2 (NA)	*0.6%*	*1.1%*	*15.5%*
B4 (EA)		*0.4%*	*0.6%*
C	**0.6%**	**0.8%**	**12.3%**
C1b, C1c, C4c (NA)	*0.6%*	*0.8%*	*12.3%*
D	**0.6%**		**5.8%**
D1, D4h3 (NA)			*5.8%*
D4e (EA)	*0.6%*		
E			**0.6%**
F	**0.6%**	**0.4%**	
G		**0.4%**	
H	**1.8%**	**36.5%**	**11.6%**
HV		**2.3%**	
I		**2.3%**	**1.3%**
J		**13.7%**	**1.3%**
K	**1.2%**	**8.0%**	**3.9%**
L0	**2.9%**		**0.6%**
L1	**17.1%**		**2.6%**
L2	**34.1%**	**0.8%**	**1.9%**
L3	**34.7%**		**7.1%**
M	**1.2%**	**0.4%**	
M1 (WE/AF)	*0.6%*		
M7 (EA)	*0.6%*	*0.4%*	
N	**0.6%**	**0.4%**	
N1a (WE/AF)		*0.4%*	
N1b (WE)	*0.6%*		
T		**9.9%**	**2.6%**
U	**0.6%**	**14.8%**	**3.9%**
U2, U3, U4, U5 (WE)		*13.7%*	*3.9%*
U6a3c (WE/AF)	*0.6%*		
U6a7a (WE/AF)		*0.8%*	
U7a (WE)		*0.4%*	
V	**1.2%**	**3.0%**	
W		**2.7%**	**1.3%**
X	**0.6%**	**1.1%**	
X2b, X2c, X2i (WE)		*1.1%*	
X2a (NA)	*0.6%*		
Y			**0.6%**

Table 5. Haplogroup frequencies by population

Population	Haplogroup	% of haplotypes
African American	L1b	10.6%
	L1c	6.5%
	L2a	24.1%
	L2a1	*24.1%*
	L2a1a	*5.9%*
	L2a1c	*6.5%*
	L2a1f	*5.3%*
	L2c	6.5%
	L3b	7.6%
	L3e	12.9%
	L3e2	*7.6%*
	L3f	8.8%
	L3f1b	*8.8%*
U.S. Caucasian	H1	13.3%
	J1	10.3%
	J1c	*8.7%*
	K1	6.5%
	T2	8.0%
	U5	10.6%
	U5a	*5.7%*
U.S. Hispanic	A2	26.5%
	B2	15.5%
	C1	12.3%
	C1b	*9.7%*
	D1	5.2%
	H1	5.8%
	L3	7.1%

Table 6. Most common haplogroups by population
For each U.S. population sample, major haplogroups and subhaplogroups (in italics) detected at frequencies greater than 5.0% are listed.

African American	**This study (n=170)**	**Diegoli *et al.* 2009 (n=248)**	**Allard *et al.* 2005 (n=1148)**
African	90.0%	93.1%	91.6%
East Asian	2.4%	1.1%	*
West Eurasian	4.7%	4.3%	5.1%
Native American	2.9%	0.7%	*
U.S. Caucasian	**This study (n=263)**	**Gonçalves *et al.* 2007 (n=1387)**	**Lao *et al.* 2010 (n=245)**
African	0.8%	0.9%	**
East Asian	1.9%	*	**
West Eurasian	94.7%	96.9%	96.7%
Native American	2.7%	*	**
U.S. Hispanic	**This study (n=155)**	**Saunier *et al.* 2008 (n=128)**	**Allard *et al.* 2006 (n=686)**
African	12.3%	14.8%	11.8%
East Asian	1.9%	1.6%	-
West Eurasian	25.8%	22.7%	17.8%
Native American	60.0%	60.9%	70.8%†

*Cannot be adequately separated based on the data presented in the papers
**Not reported
†Significantly different from the proportion reported in this study

Table 7. Biogeographic ancestry proportions for each U.S. population from this study and previous CR-based studies

The maternal biogeographic ancestry proportions inferred for each of the three U.S. populations based on full mtGenome data (this study) and CR data (previous studies). When the proportion of haplotypes assigned to the predominant biogeographic ancestry for each population group (highlighted rows in the table) were compared, only the frequency of Native American haplotypes in the U.S. Hispanic population sample in our study versus the Allard et al. [22] data differed significantly (p=0.007).

Groups Compared	Full mtGenome	CR Only
By Population		
African American pairwise	60.20	15.91
U.S. Caucasian pairwise	32.50	10.69
U.S. Hispanic pairwise	45.47	15.20
African American and U.S. Caucasian	58.45	15.91
African American and U.S. Hispanic	60.90	17.42
U.S. Caucasian and U.S. Hispanic	42.56	14.27
By Ancestry		
African pairwise	60.38	16.01
West Eurasian pairwise	30.69	10.08
Native American pairwise	41.17	14.74
African and West Eurasian	60.16	16.14
African and Native American	64.70	19.08
West Eurasian and Native American	44.33	15.60

Table 8. Average pairwise number of differences
The average pairwise number of differences for both the mtGenome and CR only are listed. Calculations were performed for each of the three U.S. population groups, three ancestry groups (African, West Eurasian, and Native American), as well as combinations of those groups. Sample sizes for each group were: African American (170), U.S. Caucasian (263), U.S. Hispanic (155), African (174), West Eurasian (297), Native American (105).

	Santos *et al.* 2008 (n=210)	**Irwin *et al.* 2009 (n=5015)**	**Ramos *et al.* 2013 (n=101)**	**This study (n=588)**	***African American (n=170)***	***U.S. Caucasian (n=263)***	***U.S. Hispanic (n=155)***
HV1 LHP	17.1%	15%	5.0%	17.7%	*24.1%*	*9.1%*	*25.2%*
HV2 LHP	64.8%	45%	38.6%	53.7%	*52.4%*	*50.6%*	*60.6%*
HV3 LHP	*	3%	5.0%	3.2%	*2.4%*	*4.2%*	*2.6%*
AC repeat	*	4.3%**	3.0%	5.3%	*3.5%*	*6.5%*	*4.5%*

*Region not analyzed or data not repoı
**Percentage reflects AC repeat LHP described as "pronounced". The authors report that the majority of samples (greater than 70%) exhibited some degree of LHP in this region.

Table 9. Frequency of LHP in the CR from this and recent studies
The percentage of samples with LHP are given for this study and three recent studies of heteroplasmy that described rates for more than one portion of the CR. Some statistical comparisons were performed for HV1 LHP frequencies, and significant differences were found a) between the Ramos et al. [26] data and the rate across all 588 of our haplotypes ($p=0.0001$), and b) across the three U.S populations reported in the current study ($p<0.00001$).

Indels relative to the rCRS*	Number of Individuals	Instances of associated LHP
595.1A	1	
960 del	1	
960.XC	5	
965.XC	4	4
2156.1A	3	
2232.1A	3	
2395 del	13	
2887-2888 del	1	
3307.1A	1	
4317 del	1	
5752 del	1	
5899 del	1	
5752 del	1	
5899 del	1	
5899.XC	12	3
8278.XC	2	2
8287.XC	1	1
8281-8289 9bp del	39	
8289.X 9bp ins	7	1
12241 del	1	
15944 del	32	

*Excludes 3107 del

Table 10. Coding region indels
Across all 588 haplotypes, indels were detected at 18 different positions in the coding region. At three of these 18 positions (960, 5899 and 8289), both insertions and deletions were observed. LHP was detected at five of the 18 positions, in 11 total instances. While observation of an indel in multiple individuals does not necessarily imply multiple occurrences of insertion or deletion at the position (as some indels are primarily or exclusively haplogroup-associated), the number of observations provides some indication of how frequently each indel might be observed in a population sample.

	All Haplotypes	African American	U.S. Caucasian	U.S. Hispanic
n (individuals)	588	170	263	155
# of PHP	166	65	68	33
# (%) of individuals with PHP	140 (23.8%)	51 (30.0%)	62 (23.6%)	27 (17.4%)
# (%) of individuals with >1 PHP	25 (4.3%)	13 (7.6%)	6 (2.3%)	6 (3.9%)
% of individuals with PHP that have >1 PHP	*17.9%*	*25.5%*	*9.7%*	*22.2%*
# (%) of individuals with 2 PHP	24 (4.1%)	12 (7.1%)	6 (2.3%)	6 (3.9%)
# (%) of individuals with 3 PHP	1 (0.2%)	1 (0.6%)	0 (0.0%)	0 (0.0%)
# (%) of individuals with CR PHP	64 (9.9%)	28 (13.5%)	24 (8.8%)	12 (7.7%)
# (%) of individuals with coding region PHP	102 (15.8%)	37 (19.4%)	44 (15.6%)	21 (12.3%)

Table 11. Point heteroplasmy statistics across all 588 samples and by population
PHP statistics were calculated for all 588 haplotypes and for each of the three U.S. populations. A total of 140 individuals (23.8%) had at least one PHP; and among those individuals with PHP, 25 (17.9%) had more than one PHP. Thus, across the entire mtGenome, multiple PHPs were seen within one individual 4.3% of the time. The highest number of PHPs observed within a single individual was three.

	CR	coding region
# of PHP	64	102
# (%) of individuals with PHP	58 (9.9%)	93 (15.8%)
# of positions at which PHP was observed	44	102
# of PHP observed in >1 individual	10*	0
% of individuals with >1 PHP in the region	0.85%	1.53%
# (%) of PHPs that represented transitions	62 (96.9%)	101 (99.0%)
# (%) of PHPs that were pyrimidine-pyrimidine	*38 (59.4%)*	*41 (40.2%)*
# (%) of PHPs that were purine-purine	*24 (37.5%)*	*60 (58.8%)*
Ratio of pyrimidine to purine PHPs	*1.6:1*	*0.7:1*
# (%) of PHPs that represented transversions	2 (3.1%)	1 (1.0%)
Ratio of transition to transversion PHPs	31:1	101:1

*Both 228K and 228R were observed; the total number of *positions* at which PHP was observed in >1 individual is 11

Table 12. Point heteroplasmy statistics by region
PHP statistics were calculated for the CR and the coding region. The number, percentage and ratio of transitions (separated by type) and transversions are listed for each region of the molecule.

	n (hg)	# PHP	# and % samples w/ PHP		# and % samples w/ >1 PHP		# CR PHP	# and % samples w/ CR PHP		# codR PHP	# and % samples w/ codR PHP	
H	117	34	31	26.5%	3	2.6%	13	13	11.1%	21	20	17.1%
L3	70	28	21	30.0%	6	8.6%	8	6	8.6%	20	17	24.3%
L2	63	22	17	27.0%	5	7.9%	12	10	15.9%	10	10	15.9%
A	47	9	7	14.9%	2	4.3%	4	3	6.4%	5	5	10.6%
U	46	14	12	26.1%	2	4.3%	4	4	8.7%	10	8	17.4%
J	38	9	7	18.4%	2	5.3%	3	2	5.3%	6	6	15.8%
L1	33	7	6	18.2%	1	3.0%	3	3	9.1%	4	3	9.1%
B	30	6	6	20.0%	0	0.0%	2	2	6.7%	4	4	13.3%
T	30	6	6	20.0%	0	0.0%	0	0	0.0%	6	6	20.0%
K	29	7	6	20.7%	1	3.4%	3	3	10.3%	4	4	13.8%
C	22	7	6	27.3%	1	4.5%	3	3	13.6%	4	4	18.2%
D	10	6	4	40.0%	2	20.0%	1	1	10.0%	5	3	30.0%
V	10	3	3	30.0%	0	0.0%	3	3	30.0%	0	0	0.0%
W	9	3	3	33.3%	0	0.0%	2	2	22.2%	1	1	11.1%
I	8	1	1	12.5%	0	0.0%	0	0	0.0%	1	1	12.5%
HV	6	1	1	16.7%	0	0.0%	1	1	16.7%	0	0	0.0%
L0	6	1	1	16.7%	0	0.0%	1	1	16.7%	0	0	0.0%
X	4	2	2	50.0%	0	0.0%	1	1	25.0%	1	1	25.0%
M	3	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%
F	2	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%
N	2	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%
E	1	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%
G	1	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%
Y	1	0	0	0.0%	0	0.0%	0	0	0.0%	0	0	0.0%

Table 13. Point heteroplasmy by major haplogroup

Table 14. All 166 point heteroplasmies observed across 588 haplotypes (following pages)
PHPs are organized by the number of occurrences, then by position. Both 228K and 228R were observed, and are highlighted pink. For each heteroplasmy, the gene or region, codon position and type of change (synonymous versus nonsynonymous) were determined using the web-based tool MitoAnalyzer [188].

Occurrences	Position	IUPAC Code	Region/Gene	Codon Position	Syn/NonSyn	Relative Substitution Rate (from Soares *et al.* 2009)
6	16093	Y	CR			79
5	152	Y	CR			157
4	189	R	CR			31
2	143	R	CR			17
2	198	Y	CR			20
2	207	R	CR			36
2	234	R	CR			12
2	16126	Y	CR			20
2	16192	Y	CR			33
2	16256	Y	CR			29
1	64	Y	CR			22
1	73	R	CR			11
1	150	Y	CR			63
1	153	R	CR			15
1	199	Y	CR			30
1	200	R	CR			30
1	204	Y	CR			43
1	228	K	CR			0
1	228	R	CR			15
1	251	R	CR			0
1	279	Y	CR			5
1	385	R	CR			5
1	482	Y	CR			5
1	513	R	CR			22
1	629	Y	tRNA			3
1	709	R	12S rRNA			59
1	794	Y	12S rRNA			2
1	870	Y	12S rRNA			2
1	1617	Y	tRNA			0
1	1632	Y	tRNA			0
1	1692	R	16S rRNA			3
1	1806	Y	16S rRNA			0
1	1958	R	16S rRNA			0
1	2707	R	16S rRNA			1
1	2784	R	16S rRNA			0
1	2887	Y	16S rRNA			2
1	2905	R	16S rRNA			1
1	3196	R	16S rRNA			0
1	3206	Y	16S rRNA			2
1	3316	R	ND1	1	NonSyn	10
1	3645	Y	ND1	3	Syn	2
1	3705	R	ND1	3	Syn	5
1	4083	Y	ND1	3	Syn	0
1	4129	R	ND1	1	NonSyn	3
1	4316	R	tRNA			0
1	4561	Y	ND2	2	NonSyn	2
1	4638	R	ND2	1	NonSyn	0
1	4646	Y	ND2	3	Syn	3
1	4748	Y	ND2	3	Syn	0
1	4973	Y	ND2	3	Syn	1
1	5099	Y	ND2	3	Syn	0
1	5147	R	ND2	3	Syn	24
1	5177	R	ND2	3	Syn	3

Occurrences	Position	IUPAC Code	Region/Gene	Codon Position	Syn/NonSyn	Relative Substitution Rate (from Soares *et al.* 2009)
1	5604	Y	tRNA			0
1	5887	Y	tRNA			0
1	6481	Y	CO1	2	NonSyn	0
1	6510	R	CO1	1	NonSyn	0
1	6581	R	CO1	3	Syn	2
1	6626	Y	CO1	3	Syn	0
1	7202	R	CO1	3	Syn	3
1	7286	Y	CO1	3	Syn	0
1	7388	W	CO1	3	Syn	0
1	7428	R	CO1	1	NonSyn	0
1	7498	R	tRNA			2
1	7543	R	tRNA			1
1	7594	Y	CO2	3	Syn	1
1	7642	R	CO2	3	Syn	3
1	7673	R	CO2	1	NonSyn	4
1	8075	R	CO2	1	NonSyn	2
1	8348	R	tRNA			1
1	8464	Y	ATP8	3	Syn	0
1	8503	Y	ATP8	3	Syn	2
1	8521	R	ATP8	3	Syn	3
1	8531	R	ATP8/ATP6	1/2	NonSyn/NonSyn	1
1	8587	R	ATP6	1	NonSyn	1
1	8745	R	ATP6	3	Syn	0
1	8903	Y	ATP6	2	NonSyn	0
1	8950	R	ATP6	1	NonSyn	3
1	9025	R	ATP6	1	NonSyn	1
1	9122	Y	ATP6	2	NonSyn	0
1	9242	R	CO3	3	Syn	2
1	9377	R	CO3	3	Syn	5
1	9689	R	CO3	3	Syn	0
1	9746	R	CO3	3	Syn	0
1	9837	R	CO3	1	NonSyn	0
1	9947	R	CO3	3	Syn	3
1	9967	Y	CO3	2	NonSyn	0
1	10018	R	tRNA			0
1	10259	R	ND3	3	Syn	0
1	10644	R	ND4L	1	NonSyn	0
1	10754	R	ND4L	3	Syn	2
1	10972	R	ND4	3	Syn	2
1	11431	Y	ND4	3	Syn	1
1	11893	R	ND4	3	Syn	0
1	11908	R	ND4	3	Syn	3
1	12007	R	ND4	3	Syn	12
1	12071	Y	ND4	1	NonSyn	0
1	12145	Y	tRNA			0
1	12202	Y	tRNA			0
1	12382	R	ND5	1	NonSyn	0
1	12594	Y	ND5	3	Syn	0
1	12654	R	ND5	3	Syn	3
1	12904	R	ND5	1	NonSyn	1
1	13327	R	ND5	1	NonSyn	1
1	13434	R	ND5	3	Syn	5
1	13473	R	ND5	3	Syn	0
1	13477	R	ND5	1	NonSyn	4
1	13506	Y	ND5	3	Syn	1
1	13656	Y	ND5	3	Syn	6
1	13884	R	ND5	3	Syn	1
1	13952	R	ND5	2	NonSyn	0
1	13965	Y	ND5	3	Syn	2
1	14208	Y	ND6	1	NonSyn	0

Occurrences	Position	IUPAC Code	Region/Gene	Codon Position	Syn/NonSyn	Relative Substitution Rate (from Soares *et al.* 2009)
1	14305	R	ND6	3	Syn	7
1	14384	R	ND6	2	NonSyn	2
1	14581	Y	ND6	3	Syn	2
1	15043	R	CYTB	3	Syn	9
1	15080	R	CYTB	1	NonSyn	0
1	15213	Y	CYTB	2	NonSyn	0
1	15260	R	CYTB	1	NonSyn	0
1	15289	Y	CYTB	3	Syn	2
1	15565	Y	CYTB	3	Syn	3
1	15774	Y	CYTB	2	NonSyn	2
1	15785	Y	CYTB	1	NonSyn	0
1	15934	R	tRNA			0
1	15994	R	tRNA			0
1	16069	Y	CR			2
1	16092	Y	CR			17
1	16129	R	CR			86
1	16147	Y	CR			5
1	16169	S	CR			12
1	16172	Y	CR			42
1	16189	Y	CR			90
1	16233	R	CR			2
1	16265	R	CR			5
1	16266	Y	CR			17
1	16278	Y	CR			43
1	16286	Y	CR			5
1	16291	Y	CR			34
1	16293	R	CR			17
1	16309	R	CR			14
1	16311	Y	CR			120
1	16320	Y	CR			15
1	16325	Y	CR			16
1	16390	R	CR			31
1	16400	Y	CR			3
1	16497	R	CR			9

PHP	Source Data	Soares *et al.* 2009 relative substitution rate	Gene/Region	Syn/NonSyn
Observed in two individuals in a single study				
1438R	2 from King *et al.* 2014	10	12S	
2083Y	2 from King *et al.* 2014	1	16S	
3492M*	2 from Li *et al.* 2010	0	ND1	NonSyn
8994R	2 from King *et al.* 2014	6	ATP6	Syn
Observed in two individuals from different studies				
2887Y	1 from Ramos *et al.* 2013 and 1 from this study	2	16S	
5177R	1 from King *et al.* 2014 and 1 from this study	3	ND2	Syn
7754R	1 from King *et al.* 2014 and 1 from Li *et al.* 2010	1	CO2	NonSyn
9025R	1 from King *et al.* 2014 and 1 from this study	1	ATP6	NonSyn
12654R	1 from Li *et al.* 2010 and 1 from this study	3	ND5	Syn

* Likely not authentic

Table 15. Coding region point heteroplasmies observed in more than one individual
No coding region PHPs were found in multiple individuals in our study. When coding region PHP data from more than 1000 individuals ([23,26,27], plus the current study) was combined, only nine coding region PHPs were observed in more than one individual. One of these PHPs (3492M) is likely not authentic.

Appendix C. Amplification primers

Amplicon Number	Amplicon Size	Primer Name	Primer Sequence	Source
1	2417	F402	ATCTTTTGGCGGTATGCACTTT	New
		R2818	GCCCCAACCGAAATTTTTAAT	New
2	2381	F2480	AAATCTTACCCCGCCTGTTT	New
		R4860	GAAGAAGCAGGCCGGATGT	New
3	2291	F4609	AAATAAACCCTCGTTCCACAGA	New
		R6899	CATATTGCTTCCGTGGAGTGTG	New
4	2511	F6636	ATTCTTATCCTACCAGGCTTCG	New
		R9146	GCGACAGCGATTTCTAGGATAG	New
5	2489	F8940	CCCCATACTAGTTATTATCGAAACC	New
		R11428	GGCTTCGACATGGGCTTT	New
6	2759	F11319	CAAACTCCTGAGCCAACAACTT	New
		R14077	TTTGGGTTGAGGTGATGATG	New
7	2208	F13835	CAGCCCTAGACCTCAACTACC	New
		R16042	CTGCTTCCCCATGAAAGAAC	[93]
8	1198	F15971	TTAACTCCACCATTAGCACC	[7]
		R599	TTGAGGAGGTAAGCTACATA	[7]

Appendix D. Sequencing primers

Amplicon Number	Primer Name	Sequence	Source
1	F402	ATCTTTTGGCGGTATGCACTTT	New
	F619	TTAGACGGGCTCACATCACC	[93]
	R878	CCAACCCTGGGGTTAGTATAGC	New
	F900	CGGTCACACGATTAACCCAAG	New
	F1135	CCAGAACACTACGAGCCACA	New
	R1136	GGCGAGCAGTTTTGTTGATT	New
	F1320	GACGTTAGGTCAAGGTGTAGCC	New
	R1365	TAGCCCATTTCTTGCCACCT	New
	F1657	CTTGACCGCTCTGAGCTAAAC	[90,91]
	R1769	GCCAGGTTTCAATTTCTATCG	[90,91]
	R1924	AGGTAGCTCGTCTGGTTTCG	New
	F1983	TAGAGGCGACAAACCTACCG	New
	F2105	GAGGAACAGCTCTTTGGACAC	[90,91]
	R2216	TGTTGAGCTTGAACGCTTTCTT	New
	F2333	GCATAAGCCTGCGTCAGAT	New
	R2439	ATGCCTGTGTTGGGTTGAC	New
	F2506	AACATCACCTCTAGCATCACCA	New
	R2818	GCCCCAACCGAAATTTTTAAT	New
2	F2625	CTGTATGAATGGCTCCACGAG	New
	F2932	GGGATAACAGCGCAATCCTAT	New
	R3006	ATGTCCTGATCCAACATCGAG	[90,91]
	F3241	AGAGCCCGGTAATCGCATAA	New
	R3417	GGGGCCTTTGCGTAGTTGTA	New
	F3441	ACTACAACCCTTCGCTGACG	[90,91]
	R3632	GAGGTGGCTAGAATAAATAGGAGGC	New
	F3635	GCCTAGCCGTTTACTCAATCC	[90,91]
	R3825	TCAGAGGTGTTCTTGTGTTGTGAT	New
	F3890	GAACCCCCTTCGACCTTG	New
	R4162	TGAGTTGGTCGTAGCGGAATC	[90,91]
	F4142	GATTCCGCTACGACCAACT	New
	F4392	CCCATCCTAAAGTAAGGTCAGC	[90,91]
	R4479	GGGGATTAATTAGTACGGGAAGG	New
	R4676	GATTATGGATGCGGTTGCTT	New
	R4811	TCAGAAGTGAAAGGGGGCTAT	New
3	F4609	AAATAAACCCTCGTTCCACAGA	New
	F4925	CCTTCTCCTCACTCTCTCAATC	New
	R5034	ATCCTATGTGGGTAATTGAGGA	New
	F5150	CCTACTACTATCTCGCACCTGAA	New
	R5210	GGTGGATGGAATTAAGGGTGT	New
	F5318	CACCATCACCCTCCTTAACC	[90,91]
	R5325	TGATGGTGGCTATGATGGTG	New
	R5681	GTGGGTTTAAGTCCCATTGGT	New
	F5664	AATGGGACTTAAACCCACAAA	New

	F5858	TTTACAGTCCAATGCTTCACTC	New
	R5799	TGCAAATTCGAAGAAGCAG	New
	R5994	TGCCTAGGACTCCAGCTCAT	[19]
	F6032	GCCAGGCAACCTTCTAGGTA	New
	F6318	CCTGGAGCCTCCGTAGACCT	New
	R6444	TTTGGTATTGGGTTATGGCAG	New
	F6496	CTCTCCCAGTCCTAGCTGCTG	New
	R6899	CATATTGCTTCCGTGGAGTGTG	New
4	F6636	ATTCTTATCCTACCAGGCTTCG	New
	F7075	GTATGGGGATAAGGGGTGTA	[90,91]
	R7248	TGGTGTATGCATCGGGGTAGT	New
	F7366	CCTCCATAAACCTGGAGTGA	New
	R7489	TGGCTTGAAACCAGCTTTG	[93]
	F7527	GAAAAACCATTTCATAACTTTGTCA	New
	R7766	TTTCCTGAGCGTCTGAGATGT	New
	F7821	CATCCCTACGCATCCTTTACAT	New
	F8129	ACCACTTTCACCGCTACACG	New
	R8141	CGGTGAAAGTGGTTTGGTTTA	New
	F8355	TTTACAGTGAAATGCCCCAAC	New
	R8378	TTAGTTGGGGCATTTCACTGT	New
	R8640	GATGAGATATTTGGAGGTGGG	New
	F8668	TGACTAATCAAACTAACCTCAAAACA	New
	F8717	AAGGACGAACCTGATCTCTTATACT	New
	R8949	TAGTATGGGGATAAGGGGTGTA	New
	R9031	GGTGGCCTGCAGTAATGTTAG	New
5	F8940	CCCCATACTAGTTATTATCGAAACC	New
	F9272	CTCAGCCCTCCTAATGACCTC	New
	R9376	CATTGGTATATGGTTAGTGTGTTGG	New
	F9483	TTCTTCGCAGGATTTTTCTGA	New
	R9611	GGATGTGTTTAGGAGTGGGACT	New
	R9853	GTGAGGAAAGTTGAGCCAATAA	New
	F9832	TTATTGGCTCAACTTTCCTCAC	New
	R10171	TAGAAAAATCCACCCCTTACGA	New
	F10267	CCCTCCTTTTACCCCTACCAT	New
	R10294	AGGGCTCATGGTAGGGGTAA	New
	F10419	AACAAAACGAATGATTTCGACTC	New
	F10689	GGCCTAGCCCTACTAGTCTCAA	New
	R10715	CGTAGTCTAGGCCATATGTGTTG	[93]
	R10942	TAGGGGGTCGGAGGAAAAG	New
	F10950	CCCTCCTAATACTAACTACCTGACTC	New
	R11166	CATCGGGTGATGATAGCCAAG	[90,91]
	R11428	GGCTTCGACATGGGCTTT	New
6	F11319	CAAACTCCTGAGCCAACAACTT	New
	F11760	ACGAACGCACTCACAGTCG	[90,91]
	R11768	TGCGTTCGTAGTTTGAGTTTG	New
	R11804	GAAGTCCTTGAGAGAGGATTATGA	New
	F11964	TCACAGCCCTATACTCCCTCT	New
	R12089	TGGGGGATAGGTGTATGAACA	New
	F12194	CCCCTTATTTACCGAGAAAGC	New

	R12302	GCCTAAGACCAATGGATAGCT	New
	F12452	TTGTCGCATCCACCTTTATT	New
	F12741	CAACCTATTCCAACTGTTCATCG	New
	R12766	AGCCGATGAACAGTTGGAATA	New
	R13025	TGGAGACCTAATTGGGCTGA	New
	F13203	AGTCTGCGCCCTTACACAAA	New
	R13390	TGTTAAGGTTGTGGATGATGGA	New
	R13559	GCTCAGGCGTTTGTGTATGAT	New
	F13628	CTAACAGGTCAACCTCGCTTC	New
	R13855	GGTAGTTGAGGTCTAGGGCTGTT	New
	R13924	GGTAGAATCCGAGTATGTTGGAG	New
7	F13835	CAGCCCTAGACCTCAACTACC	New
	F14058	CATCATCACCTCAACCCAAA	New
	R14118	TGGGAAGAAGAAAGAGAGGAAG	[90,91]
	F14431	TGCCTCAGGATACTCCTCAAT	New
	R14448	GAGGAGTATCCTGAGGCATGG	New
	F14641	ACCCACACTCAACAGAAACAAA	New
	R14721	CGATGGTTTTTCATATCATTGG	New
	F14881	CACCACAGGACTATTCCTAGCC	New
	R14902	GGCTAGGAATAGTCCTGTGGTG	New
	F15190	CTTACTATCCGCCATCCCATA	New
	R15396	TTATCGGAATGGGAGGTGATTC	[90,91]
	F15500	GACCCAGACAATTATACCCTAGCC	New
	R15585	ATTGTGTAGGCGAATAGGAAATA	New
	F15699	GCCCACTAAGCCAATCACTT	[93]
	R15728	GGAGTCAATAAAGTGATTGGCTTAG	New
	R16042	CTGCTTCCCCATGAAAGAAC	[93]
8	F15971	TTAACTCCACCATTAGCACC	[7]
	F16190	CCCCATGCTTACAAGCAAGT	[7]
	F155	TATTTATCGCACCTACGTTC	[7]
	F314	CCGCTTCTGGCCACAGCACT	[7]
	R16410	GAGGATGGTGGTCAAGGGA	[7]
	R285	GTTATGATGTCTGTGTGGAA	[7]
	R484	TGAGATTAGTAGTATGGGAG	[7]
	R599	TTGAGGAGGTAAGCTACATA	[7]

Appendix E. Forward and reverse primer plate maps for high-throughput processing

A)

	1	2	3	4	5	6	7	8	9	10	11	12
A	F402	F619	F900	F1320	F1657	F1983	F2105	F2333	F2506			
B	F2625	F2932	F3241	F3441	F3635	F3890	F4142	F4392	F1135			
C	F4609	F4925	F5150	F5318	F5664	F5858	F6032	F6318	R3632			
D	F6636	F7075	F7366	F7527	F7821	F8129	F8355	F8717	F6496			
E	F8940	F9272	F9483	F9832	F10267	F10419	F10689	F10950	F8668			
F	F11319	F11760	F11964	F12194	F12452	F12741	F13203	F13628				
G	F13835	F14058	F14431	F14641	F14881	F15190	F15500	F15699				
H	F15971a	F15971b	F16190a	F16190b	F155a	F155b	F314a	F314b				

B)

	1	2	3	4	5	6	7	8	9	10	11	12
A	R878	R1136	R1365	R1769	R1924	R2216	R2439	R2818	R11428			
B	R3006	R3417		R3825	R4162	R4479	R4676	R4811	R13924			
C	R5034	R5210	R5325	R5681	R5799	R5994	R6444	R6899	R11804			
D	R7248	R7489	R7766	R8141	R8378	R8640	R8949	R9031				
E	R9376	R9611	R9853	R10171	R10294	R10715	R10942	R11166				
F	R11768	R12089	R12302	R12766	R13025	R13390	R13559	R13855				
G	R14118	R14448	R14721	R14902	R15396	R15585	R15728	R16042				
H	R16410a	R16410b	R285a	R285b	R484a	R484b	R599a	R599b				

Appendix F. Sequencing plate maps for high-throughput processing

A) Forward

PLATE 1	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind1-F402	Ind1-F619	Ind1-F900	Ind1-F1320	Ind2-F402	Ind2-F619	Ind2-F900	Ind2-F1320	Ind3-F402	Ind3-F619	Ind3-F900	Ind3-F1320
B	Ind1-F2625	Ind1-F2932	Ind1-F3241	Ind1-F3441	Ind2-F2625	Ind2-F2932	Ind2-F3241	Ind2-F3441	Ind3-F2625	Ind3-F2932	Ind3-F3241	Ind3-F3441
C	Ind1-F4609	Ind1-F4925	Ind1-F5150	Ind1-F5318	Ind2-F4609	Ind2-F4925	Ind2-F5150	Ind2-F5318	Ind3-F4609	Ind3-F4925	Ind3-F5150	Ind3-F5318
D	Ind1-F6636	Ind1-F7075	Ind1-F7366	Ind1-F7527	Ind2-F6636	Ind2-F7075	Ind2-F7366	Ind2-F7527	Ind3-F6636	Ind3-F7075	Ind3-F7366	Ind3-F7527
E	Ind1-F8940	Ind1-F9272	Ind1-F9483	Ind1-F9832	Ind2-F8940	Ind2-F9272	Ind2-F9483	Ind2-F9832	Ind3-F8940	Ind3-F9272	Ind3-F9483	Ind3-F9832
F	Ind1-F11319	Ind1-F11760	Ind1-F11964			Ind2-F11760	Ind2-F11964			Ind3-F11760	Ind3-F11964	Ind3-F12194
G	Ind1-F13835	Ind1-F14058	Ind1-F14431	Ind1-F14641	Ind2-F13835	Ind2-F14058	Ind2-F14431	Ind2-F14641	Ind3-F13835	Ind3-F14058	Ind3-F14431	Ind3-F14641
H	Ind1-F15971	Ind1-F15971	Ind1-F16190	Ind1-F16190	Ind2-F15971	Ind2-F15971	Ind2-F16190	Ind2-F16190	Ind3-F15971	Ind3-F15971	Ind3-F16190	Ind3-F16190

PLATE 2	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind4-F402	Ind4-F619	Ind4-F900	Ind4-F1320	Ind5-F402	Ind5-F619	Ind5-F900	Ind5-F1320	Ind6-F402	Ind6-F619	Ind6-F900	Ind6-F1320
B	Ind4-F2625	Ind4-F2932	Ind4-F3241	Ind4-F3441	Ind5-F2625	Ind5-F2932	Ind5-F3241	Ind5-F3441	Ind6-F2625	Ind6-F2932	Ind6-F3241	Ind6-F3441
C	Ind4-F4609	Ind4-F4925	Ind4-F5150	Ind4-F5318	Ind5-F4609	Ind5-F4925	Ind5-F5150	Ind5-F5318	Ind6-F4609	Ind6-F4925	Ind6-F5150	Ind6-F5318
D	Ind4-F6636	Ind4-F7075	Ind4-F7366	Ind4-F7527	Ind5-F6636	Ind5-F7075	Ind5-F7366	Ind5-F7527	Ind6-F6636	Ind6-F7075	Ind6-F7366	Ind6-F7527
E	Ind4-F8940	Ind4-F9272	Ind4-F9483	Ind4-F9832	Ind5-F8940	Ind5-F9272	Ind5-F9483	Ind5-F9832	Ind6-F8940	Ind6-F9272	Ind6-F9483	Ind6-F9832
F	Ind4-F11319	Ind4-F11760	Ind4-F11964			Ind5-F11760	Ind5-F11964			Ind6-F11760	Ind6-F11964	Ind6-F12194
G	Ind4-F13835	Ind4-F14058	Ind4-F14431	Ind4-F14641	Ind5-F13835	Ind5-F14058	Ind5-F14431	Ind5-F14641	Ind6-F13835	Ind6-F14058	Ind6-F14431	Ind6-F14641
H	Ind4-F15971	Ind4-F15971	Ind4-F16190	Ind4-F16190	Ind5-F15971	Ind5-F15971	Ind5-F16190	Ind5-F16190	Ind6-F15971	Ind6-F15971	Ind6-F16190	Ind6-F16190

PLATE 3	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind7-F402	Ind7-F619	Ind7-F900	Ind7-F1320	Ind8-F402	Ind8-F619	Ind8-F900	Ind8-F1320	Ind9-F402	Ind9-F619	Ind9-F900	Ind9-F1320
B	Ind7-F2625	Ind7-F2932	Ind7-F3241	Ind7-F3441	Ind8-F2625	Ind8-F2932	Ind8-F3241	Ind8-F3441	Ind9-F2625	Ind9-F2932	Ind9-F3241	Ind9-F3441
C	Ind7-F4609	Ind7-F4925	Ind7-F5150	Ind7-F5318	Ind8-F4609	Ind8-F4925	Ind8-F5150	Ind8-F5318	Ind9-F4609	Ind9-F4925	Ind9-F5150	Ind9-F5318
D	Ind7-F6636	Ind7-F7075	Ind7-F7366	Ind7-F7527	Ind8-F6636	Ind8-F7075	Ind8-F7366	Ind8-F7527	Ind9-F6636	Ind9-F7075	Ind9-F7366	Ind9-F7527
E	Ind7-F8940	Ind7-F9272	Ind7-F9483	Ind7-F9832	Ind8-F8940	Ind8-F9272	Ind8-F9483	Ind8-F9832	Ind9-F8940	Ind9-F9272	Ind9-F9483	Ind9-F9832
F	Ind7-F11319	Ind7-F11760	Ind7-F11964			Ind8-F11760	Ind8-F11964	Ind8-F12194	Ind9-F11319	Ind9-F11760	Ind9-F11964	Ind9-F12194
G	Ind7-F13835	Ind7-F14058	Ind7-F14431	Ind7-F14641	Ind8-F13835	Ind8-F14058	Ind8-F14431	Ind8-F14641	Ind9-F13835	Ind9-F14058	Ind9-F14431	Ind9-F14641
H	Ind7-F15971	Ind7-F15971	Ind7-F16190	Ind7-F16190	Ind8-F15971	Ind8-F15971	Ind8-F16190	Ind8-F16190	Ind9-F15971	Ind9-F15971	Ind9-F16190	Ind9-F16190

PLATE 4	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind10-F402	Ind10-F619	Ind10-F900	Ind10-F1320	Ind11-F402	Ind11-F619	Ind11-F900	Ind11-F1320	Ind1-F1657	Ind1-F1983	Ind1-F2105	Ind1-F2333
B	Ind10-F2625	Ind10-F2932	Ind10-F3241	Ind10-F3441	Ind11-F2625	Ind11-F2932	Ind11-F3241	Ind11-F3441	Ind1-F3635	Ind1-F3890	Ind1-F4142	Ind1-F4392
C	Ind10-F4609	Ind10-F4925	Ind10-F5150	Ind10-F5318	Ind11-F4609	Ind11-F4925	Ind11-F5150	Ind11-F5318	Ind1-F5664	Ind1-F5858	Ind1-F6032	Ind1-F6318
D	Ind10-F6636	Ind10-F7075	Ind10-F7366	Ind10-F7527	Ind11-F6636	Ind11-F7075	Ind11-F7366	Ind11-F7527	Ind1-F7821	Ind1-F8129	Ind1-F8355	Ind1-F8717
E	Ind10-F8940	Ind10-F9272	Ind10-F9483	Ind10-F9832	Ind11-F8940	Ind11-F9272	Ind11-F9483	Ind11-F9832	Ind1-F10267	Ind1-F10419	Ind1-F10689	Ind1-F10950
F	Ind10-F11319	Ind10-F11760	Ind10-F11964	Ind10-F12194	Ind11-F11319	Ind11-F11760	Ind11-F11964	Ind11-F12194	Ind1-F12452	Ind1-F12741	Ind1-F13203	Ind1-F13628
G	Ind10-F13835	Ind10-F14058	Ind10-F14431	Ind10-F14641	Ind11-F13835	Ind11-F14058	Ind11-F14431	Ind11-F14641	Ind1-F14881	Ind1-F15190	Ind1-F15500	Ind1-F15699
H	Ind10-F15971	Ind10-F15971	Ind10-F16190	Ind10-F16190	Ind11-F15971	Ind11-F15971	Ind11-F16190	Ind11-F16190	Ind1-F155	Ind1-F155	Ind1-F314	Ind1-F314

PLATE 5	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind2-F1657	Ind2-F1983	Ind2-F2105	Ind2-F2333	Ind3-F1657	Ind3-F1983	Ind3-F2105	Ind3-F2333	Ind4-F1657	Ind4-F1983	Ind4-F2105	Ind4-F2333
B	Ind2-F3635	Ind2-F3890	Ind2-F4142	Ind2-F4392	Ind3-F3635	Ind3-F3890	Ind3-F4142	Ind3-F4392	Ind4-F3635	Ind4-F3890	Ind4-F4142	Ind4-F4392
C	Ind2-F5664	Ind2-F5858	Ind2-F6032	Ind2-F6318	Ind3-F5664	Ind3-F5858	Ind3-F6032	Ind3-F6318	Ind4-F5664	Ind4-F5858	Ind4-F6032	Ind4-F6318
D	Ind2-F7821	Ind2-F8129	Ind2-F8355	Ind2-F8717	Ind3-F7821	Ind3-F8129	Ind3-F8355	Ind3-F8717	Ind4-F7821	Ind4-F8129	Ind4-F8355	Ind4-F8717
E	Ind2-F10267	Ind2-F10419	Ind2-F10689	Ind2-F10950	Ind3-F10267	Ind3-F10419	Ind3-F10689	Ind3-F10950	Ind4-F10267	Ind4-F10419	Ind4-F10689	Ind4-F10950
F	Ind2-F12452	Ind2-F12741	Ind2-F13203			Ind3-F12741	Ind3-F13203			Ind4-F12741	Ind4-F13203	Ind4-F13628
G	Ind2-F14881	Ind2-F15190	Ind2-F15500	Ind2-F15699	Ind3-F14881	Ind3-F15190	Ind3-F15500	Ind3-F15699	Ind4-F14881	Ind4-F15190	Ind4-F15500	Ind4-F15699
H	Ind2-F155	Ind2-F155	Ind2-F314	Ind2-F314	Ind3-F155	Ind3-F155	Ind3-F314	Ind3-F314	Ind4-F155	Ind4-F155	Ind4-F314	Ind4-F314

PLATE 6	1	2	3	4	5	6	7	8	9	10	11	12
A	Ind5-F1657	Ind5-F1983	Ind5-F2105	Ind5-F2333	Ind6-F1657	Ind6-F1983	Ind6-F2105	Ind6-F2333	Ind7-F1657	Ind7-F1983	Ind7-F2105	Ind7-F2333
B	Ind5-F3635	Ind5-F3890	Ind5-F4142	Ind5-F4392	Ind6-F3635	Ind6-F3890	Ind6-F4142	Ind6-F4392	Ind7-F3635	Ind7-F3890	Ind7-F4142	Ind7-F4392

C	Ind5-F5664	Ind5-F5858	Ind5-F6032	Ind5-F6318	Ind6-F5664	Ind6-F5858	Ind6-F6032	Ind6-F6318	Ind7-F5664	Ind7-F5858	Ind7-F6032	Ind7-F6318
D	Ind5-F7821	Ind5-F8129	Ind5-F8355	Ind5-F8717	Ind6-F7821	Ind6-F8129	Ind6-F8355	Ind6-F8717	Ind7-F7821	Ind7-F8129	Ind7-F8355	Ind7-F8717
E	Ind5-F10267	Ind5-F10419	Ind5-F10689	Ind5-F10950	Ind6-F10267	Ind6-F10419	Ind6-F10689	Ind6-F10950	Ind7-F10267	Ind7-F10419	Ind7-F10689	Ind7-F10950
F	Ind5-F12452	Ind5-F12741	Ind5-F13203	Ind5-F13628	Ind6-F12452	Ind6-F12741	Ind6-F13203	Ind6-F13628	Ind7-F12452	Ind7-F12741	Ind7-F13203	Ind7-F13628
G	Ind5-F14881	Ind5-F15190	Ind5-F15500	Ind5-F15699	Ind6-F14881	Ind6-F15190	Ind6-F15500	Ind6-F15699	Ind7-F14881	Ind7-F15190	Ind7-F15500	Ind7-F15699
H	Ind5-F155	Ind5-F155	Ind5-F314	Ind5-F314	Ind6-F155	Ind6-F155	Ind6-F314	Ind6-F314	Ind7-F155	Ind7-F155	Ind7-F314	Ind7-F314

PLATE 7

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind8-F1657	Ind8-F1983	Ind8-F2105	Ind8-F2333	Ind9-F1657	Ind9-F1983	Ind9-F2105	Ind9-F2333	Ind10-F1657	Ind10-F1983	Ind10-F2105	Ind10-F2333
B	Ind8-F3635	Ind8-F3890	Ind8-F4142	Ind8-F4392	Ind9-F3635	Ind9-F3890	Ind9-F4142	Ind9-F4392	Ind10-F3635	Ind10-F3890	Ind10-F4142	Ind10-F4392
C	Ind8-F5664	Ind8-F5858	Ind8-F6032	Ind8-F6318	Ind9-F5664	Ind9-F5858	Ind9-F6032	Ind9-F6318	Ind10-F5664	Ind10-F5858	Ind10-F6032	Ind10-F6318
D	Ind8-F7821	Ind8-F8129	Ind8-F8355	Ind8-F8717	Ind9-F7821	Ind9-F8129	Ind9-F8355	Ind9-F8717	Ind10-F7821	Ind10-F8129	Ind10-F8355	Ind10-F8717
E	Ind8-F10267	Ind8-F10419	Ind8-F10689	Ind8-F10950	Ind9-F10267	Ind9-F10419	Ind9-F10689	Ind9-F10950	Ind10-F10267	Ind10-F10419	Ind10-F10689	Ind10-F10950
F	Ind8-F12452	Ind8-F12741	Ind8-F13203	Ind8-F13628	Ind9-F12452	Ind9-F12741	Ind9-F13203	Ind9-F13628	Ind10-F12452	Ind10-F12741	Ind10-F13203	Ind10-F13628
G	Ind8-F14881	Ind8-F15190	Ind8-F15500	Ind8-F15699	Ind9-F14881	Ind9-F15190	Ind9-F15500	Ind9-F15699	Ind10-F14881	Ind10-F15190	Ind10-F15500	Ind10-F15699
H	Ind8-F155	Ind8-F155	Ind8-F314	Ind8-F314	Ind9-F155	Ind9-F155	Ind9-F314	Ind9-F314	Ind10-F155	Ind10-F155	Ind10-F314	Ind10-F314

PLATE 8

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind11-F1657	Ind11-F1983	Ind11-F2105	Ind11-F2333	Ind1-F2506	Ind2-F6496	Ind4-F1135	Ind5-F8668	Ind7-R3632	Ind9-F2506	Ind10-F6496	EMPTY
B	Ind11-F3635	Ind11-F3890	Ind11-F4142	Ind11-F4392	Ind1-F1135	Ind2-F8668	Ind4-R3632	Ind6-F2506	Ind7-F6496	Ind9-F1135	Ind10-F8668	EMPTY
C	Ind11-F5664	Ind11-F5858	Ind11-F6032	Ind11-F6318	Ind1-R3632	Ind3-F2506	Ind4-F6496	Ind6-F1135	Ind7-F8668	Ind9-R3632	Ind11-F2506	EMPTY
D	Ind11-F7821	Ind11-F8129	Ind11-F8355	Ind11-F8717	Ind1-F6496	Ind3-F1135	Ind4-F8668	Ind6-R3632	Ind8-F2506	Ind9-F6496	Ind11-F1135	EMPTY
E	Ind11-F10267	Ind11-F10419	Ind11-F10689	Ind11-F10950	Ind1-F8668	Ind3-R3632	Ind5-F2506	Ind6-F6496	Ind8-F1135	Ind9-F8668	Ind11-R3632	EMPTY
F	Ind11-F12452	Ind11-F12741	Ind11-F13203	Ind11-F13628	Ind2-F2506	Ind3-F6496	Ind5-F1135	Ind6-F8668	Ind8-R3632	Ind10-F2506	Ind11-F6496	EMPTY
G	Ind11-F14881	Ind11-F15190	Ind11-F15500	Ind11-F15699	Ind2-F1135	Ind3-F8668	Ind5-R3632	Ind7-F2506	Ind8-F6496	Ind10-F1135	Ind11-F8668	EMPTY
H	Ind11-F155	Ind11-F155	Ind11-F314	Ind11-F314	Ind2-R3632	Ind4-F2506	Ind5-F6496	Ind7-F1135	Ind8-F8668	Ind10-R3632	EMPTY	EMPTY

B) Reverse

PLATE 1

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind1-R878	Ind1-R1136	Ind1-R1365	Ind1-R1769	Ind2-R878	Ind2-R1136	Ind2-R1365	Ind2-R1769	Ind3-R878	Ind3-R1136	Ind3-R1365	Ind3-R1769
B	Ind1-R3006	Ind1-R3417	EMPTY	Ind1-R3825	Ind2-R3006	Ind2-R3417	EMPTY	Ind2-R3825	Ind3-R3006	Ind3-R3417	EMPTY	Ind3-R3825
C	Ind1-R5034	Ind1-R5210	Ind1-R5325	Ind1-R5681	Ind2-R5034	Ind2-R5210	Ind2-R5325	Ind2-R5681	Ind3-R5034	Ind3-R5210	Ind3-R5325	Ind3-R5681
D	Ind1-R7248	Ind1-R7489	Ind1-R7766	Ind1-R8141	Ind2-R7248	Ind2-R7489	Ind2-R7766	Ind2-R8141	Ind3-R7248	Ind3-R7489	Ind3-R7766	Ind3-R8141
E	Ind1-R9376	Ind1-R9611	Ind1-R9853	Ind1-R10171	Ind2-R9376	Ind2-R9611	Ind2-R9853	Ind2-R10171	Ind3-R9376	Ind3-R9611	Ind3-R9853	Ind3-R10171
F	Ind1-R11768	Ind1-R12089	Ind1-R12302			Ind2-R12089	Ind2-R12302			Ind3-R12089	Ind3-R12302	Ind3-R12766
G	Ind1-R14118	Ind1-R14448	Ind1-R14721	Ind1-R14902	Ind2-R14118	Ind2-R14448	Ind2-R14721	Ind2-R14902	Ind3-R14118	Ind3-R14448	Ind3-R14721	Ind3-R14902
H	Ind1-R16410	Ind1-R16410	Ind1-R285	Ind1-R285	Ind2-R16410	Ind2-R16410	Ind2-R285	Ind2-R285	Ind3-R16410	Ind3-R16410	Ind3-R285	Ind3-R285

PLATE 2

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind4-R878	Ind4-R1136	Ind4-R1365	Ind4-R1769	Ind5-R878	Ind5-R1136	Ind5-R1365	Ind5-R1769	Ind6-R878	Ind6-R1136	Ind6-R1365	Ind6-R1769
B	Ind4-R3006	Ind4-R3417	EMPTY	Ind4-R3825	Ind5-R3006	Ind5-R3417	EMPTY	Ind5-R3825	Ind6-R3006	Ind6-R3417	EMPTY	Ind6-R3825
C	Ind4-R5034	Ind4-R5210	Ind4-R5325	Ind4-R5681	Ind5-R5034	Ind5-R5210	Ind5-R5325	Ind5-R5681	Ind6-R5034	Ind6-R5210	Ind6-R5325	Ind6-R5681
D	Ind4-R7248	Ind4-R7489	Ind4-R7766	Ind4-R8141	Ind5-R7248	Ind5-R7489	Ind5-R7766	Ind5-R8141	Ind6-R7248	Ind6-R7489	Ind6-R7766	Ind6-R8141
E	Ind4-R9376	Ind4-R9611	Ind4-R9853	Ind4-R10171	Ind5-R9376	Ind5-R9611	Ind5-R9853	Ind5-R10171	Ind6-R9376	Ind6-R9611	Ind6-R9853	Ind6-R10171
F	Ind4-R11768	Ind4-R12089	Ind4-R12302			Ind5-R12089	Ind5-R12302			Ind6-R12089	Ind6-R12302	Ind6-R12766
G	Ind4-R14118	Ind4-R14448	Ind4-R14721	Ind4-R14902	Ind5-R14118	Ind5-R14448	Ind5-R14721	Ind5-R14902	Ind6-R14118	Ind6-R14448	Ind6-R14721	Ind6-R14902
H	Ind4-R16410	Ind4-R16410	Ind4-R285	Ind4-R285	Ind5-R16410	Ind5-R16410	Ind5-R285	Ind5-R285	Ind6-R16410	Ind6-R16410	Ind6-R285	Ind6-R285

PLATE 3

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind7-R878	Ind7-R1136	Ind7-R1365	Ind7-R1769	Ind8-R878	Ind8-R1136	Ind8-R1365	Ind8-R1769	Ind9-R878	Ind9-R1136	Ind9-R1365	Ind9-R1769
B	Ind7-R3006	Ind7-R3417	EMPTY	Ind7-R3825	Ind8-R3006	Ind8-R3417	EMPTY	Ind8-R3825	Ind9-R3006	Ind9-R3417	EMPTY	Ind9-R3825
C	Ind7-R5034	Ind7-R5210	Ind7-R5325			Ind8-R5210	Ind8-R5325	Ind8-R5681	Ind9-R5034	Ind9-R5210	Ind9-R5325	Ind9-R5681
D	Ind7-R7248	Ind7-R7489	Ind7-R7766	Ind7-R8141	Ind8-R7248	Ind8-R7489	Ind8-R7766	Ind8-R8141	Ind9-R7248	Ind9-R7489	Ind9-R7766	Ind9-R8141

E	Ind7-R9376	Ind7-R9611	Ind7-R9853	Ind7-R10171	Ind8-R9376	Ind8-R9611	Ind8-R9853	Ind8-R10171	Ind9-R9376	Ind9-R9611	Ind9-R9853	Ind9-R10171
F	Ind7-R11768	Ind7-R12089	Ind7-R12302	Ind7-R12766	Ind8-R11768	Ind8-R12089	Ind8-R12302	Ind8-R12766	Ind9-R11768	Ind9-R12089	Ind9-R12302	Ind9-R12766
G	Ind7-R14118	Ind7-R14448	Ind7-R14721			Ind8-R14448	Ind8-R14721	Ind8-R14902	Ind9-R14118	Ind9-R14448	Ind9-R14721	Ind9-R14902
H	Ind7-R16410	Ind7-R16410	Ind7-R285	Ind7-R285	Ind8-R16410	Ind8-R16410	Ind8-R285	Ind8-R285	Ind9-R16410	Ind9-R16410	Ind9-R285	Ind9-R285

PLATE 4

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind10-R878	Ind10-R1136	Ind10-R1365	Ind10-R1769	Ind11-R878	Ind11-R1136	Ind11-R1365	Ind11-R1769	Ind1-R1924	Ind1-R2216	Ind1-R2439	Ind1-R2818
B	Ind10-R3006	Ind10-R3417	EMPTY	Ind10-R3825	Ind11-R3006	Ind11-R3417	EMPTY	Ind11-R3825	Ind1-R4162	Ind1-R4479	Ind1-R4676	Ind1-R4811
C	Ind10-R5034	Ind10-R5210	Ind10-R5325	Ind10-R5681	Ind11-R5034	Ind11-R5210	Ind11-R5325	Ind11-R5681	Ind1-R5799	Ind1-R5994	Ind1-R6444	Ind1-R6899
D	Ind10-R7248	Ind10-R7489	Ind10-R7766	Ind10-R8141	Ind11-R7248	Ind11-R7489	Ind11-R7766	Ind11-R8141	Ind1-R8378	Ind1-R8640	Ind1-R8949	Ind1-R9031
E	Ind10-R9376	Ind10-R9611	Ind10-R9853	Ind10-R10171	Ind11-R9376	Ind11-R9611	Ind11-R9853	Ind11-R10171	Ind1-R10294	Ind1-R10715	Ind1-R10942	Ind1-R11166
F	Ind10-R11768	Ind10-R12089	Ind10-R12302	Ind10-R12766	Ind11-R11768	Ind11-R12089	Ind11-R12302	Ind11-R12766	Ind1-R13025	Ind1-R13390	Ind1-R13559	Ind1-R13855
G	Ind10-R14118	Ind10-R14448	Ind10-R14721	Ind10-R14902	Ind11-R14118	Ind11-R14448	Ind11-R14721	Ind11-R14902	Ind1-R15396	Ind1-R15585	Ind1-R15728	Ind1-R16042
H	Ind10-R16410	Ind10-R16410	Ind10-R285	Ind10-R285	Ind11-R16410	Ind11-R16410	Ind11-R285	Ind11-R285	Ind1-R484	Ind1-R484	Ind1-R599	Ind1-R599

PLATE 5

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind2-R1924	Ind2-R2216	Ind2-R2439	Ind2-R2818	Ind3-R1924	Ind3-R2216	Ind3-R2439	Ind3-R2818	Ind4-R1924	Ind4-R2216	Ind4-R2439	Ind4-R2818
B	Ind2-R4162	Ind2-R4479	Ind2-R4676	Ind2-R4811	Ind3-R4162	Ind3-R4479	Ind3-R4676	Ind3-R4811	Ind4-R4162	Ind4-R4479	Ind4-R4676	Ind4-R4811
C	Ind2-R5799	Ind2-R5994	Ind2-R6444	Ind2-R6899	Ind3-R5799	Ind3-R5994	Ind3-R6444	Ind3-R6899	Ind4-R5799	Ind4-R5994	Ind4-R6444	Ind4-R6899
D	Ind2-R8378	Ind2-R8640	Ind2-R8949	Ind2-R9031	Ind3-R8378	Ind3-R8640	Ind3-R8949	Ind3-R9031	Ind4-R8378	Ind4-R8640	Ind4-R8949	Ind4-R9031
E	Ind2-R10294	Ind2-R10715	Ind2-R10942	Ind2-R11166	Ind3-R10294	Ind3-R10715	Ind3-R10942	Ind3-R11166	Ind4-R10294	Ind4-R10715	Ind4-R10942	Ind4-R11166
F	Ind2-R13025	Ind2-R13390	Ind2-R13559	Ind2-R13855	Ind3-R13025	Ind3-R13390	Ind3-R13559	Ind3-R13855	Ind4-R13025	Ind4-R13390	Ind4-R13559	Ind4-R13855
G	Ind2-R15396	Ind2-R15585	Ind2-R15728	Ind2-R16042	Ind3-R15396	Ind3-R15585	Ind3-R15728	Ind3-R16042	Ind4-R15396	Ind4-R15585	Ind4-R15728	Ind4-R16042
H	Ind2-R484	Ind2-R484	Ind2-R599	Ind2-R599	Ind3-R484	Ind3-R484	Ind3-R599	Ind3-R599	Ind4-R484	Ind4-R484	Ind4-R599	Ind4-R599

PLATE 6

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind5-R1924	Ind5-R2216	Ind5-R2439	Ind5-R2818	Ind6-R1924	Ind6-R2216	Ind6-R2439	Ind6-R2818	Ind7-R1924	Ind7-R2216	Ind7-R2439	Ind7-R2818
B	Ind5-R4162	Ind5-R4479	Ind5-R4676	Ind5-R4811	Ind6-R4162	Ind6-R4479	Ind6-R4676	Ind6-R4811	Ind7-R4162	Ind7-R4479	Ind7-R4676	Ind7-R4811
C	Ind5-R5799	Ind5-R5994	Ind5-R6444	Ind5-R6899	Ind6-R5799	Ind6-R5994	Ind6-R6444	Ind6-R6899	Ind7-R5799	Ind7-R5994	Ind7-R6444	Ind7-R6899
D	Ind5-R8378	Ind5-R8640	Ind5-R8949	Ind5-R9031	Ind6-R8378	Ind6-R8640	Ind6-R8949	Ind6-R9031	Ind7-R8378	Ind7-R8640	Ind7-R8949	Ind7-R9031
E	Ind5-R10294	Ind5-R10715	Ind5-R10942	Ind5-R11166	Ind6-R10294	Ind6-R10715	Ind6-R10942	Ind6-R11166	Ind7-R10294	Ind7-R10715	Ind7-R10942	Ind7-R11166
F	Ind5-R13025	Ind5-R13390	Ind5-R13559	Ind5-R13855	Ind6-R13025	Ind6-R13390	Ind6-R13559	Ind6-R13855	Ind7-R13025	Ind7-R13390	Ind7-R13559	Ind7-R13855
G	Ind5-R15396	Ind5-R15585	Ind5-R15728	Ind5-R16042	Ind6-R15396	Ind6-R15585	Ind6-R15728	Ind6-R16042	Ind7-R15396	Ind7-R15585	Ind7-R15728	Ind7-R16042
H	Ind5-R484	Ind5-R484	Ind5-R599	Ind5-R599	Ind6-R484	Ind6-R484	Ind6-R599	Ind6-R599	Ind7-R484	Ind7-R484	Ind7-R599	Ind7-R599

PLATE 7

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind8-R1924	Ind8-R2216	Ind8-R2439	Ind8-R2818	Ind9-R1924	Ind9-R2216	Ind9-R2439	Ind9-R2818	Ind10-R1924	Ind10-R2216	Ind10-R2439	Ind10-R2818
B	Ind8-R4162	Ind8-R4479	Ind8-R4676	Ind8-R4811	Ind9-R4162	Ind9-R4479	Ind9-R4676	Ind9-R4811	Ind10-R4162	Ind10-R4479	Ind10-R4676	Ind10-R4811
C	Ind8-R5799	Ind8-R5994	Ind8-R6444	Ind8-R6899	Ind9-R5799	Ind9-R5994	Ind9-R6444	Ind9-R6899	Ind10-R5799	Ind10-R5994	Ind10-R6444	Ind10-R6899
D	Ind8-R8378	Ind8-R8640	Ind8-R8949	Ind8-R9031	Ind9-R8378	Ind9-R8640	Ind9-R8949	Ind9-R9031	Ind10-R8378	Ind10-R8640	Ind10-R8949	Ind10-R9031
E	Ind8-R10294	Ind8-R10715	Ind8-R10942	Ind8-R11166	Ind9-R10294	Ind9-R10715	Ind9-R10942	Ind9-R11166	Ind10-R10294	Ind10-R10715	Ind10-R10942	Ind10-R11166
F	Ind8-R13025	Ind8-R13390	Ind8-R13559	Ind8-R13855	Ind9-R13025	Ind9-R13390	Ind9-R13559	Ind9-R13855	Ind10-R13025	Ind10-R13390	Ind10-R13559	Ind10-R13855
G	Ind8-R15396	Ind8-R15585	Ind8-R15728	Ind8-R16042	Ind9-R15396	Ind9-R15585	Ind9-R15728	Ind9-R16042	Ind10-R15396	Ind10-R15585	Ind10-R15728	Ind10-R16042
H	Ind8-R484	Ind8-R484	Ind8-R599	Ind8-R599	Ind9-R484	Ind9-R484	Ind9-R599	Ind9-R599	Ind10-R484	Ind10-R484	Ind10-R599	Ind10-R599

PLATE 8

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	Ind11-R1924	Ind11-R2216	Ind11-R2439	Ind11-R2818	Ind1-R11428	Ind3-R11804	Ind6-R13924	Ind9-R11428	Ind11-R11804	EMPTY	EMPTY	EMPTY
B	Ind11-R4162	Ind11-R4479	Ind11-R4676	Ind11-R4811	Ind1-R13924	Ind4-R11428	Ind6-R11804	Ind9-R13924	EMPTY	EMPTY	EMPTY	EMPTY
C	Ind11-R5799	Ind11-R5994	Ind11-R6444	Ind11-R6899	Ind1-R11804	Ind4-R13924	Ind7-R11428	Ind9-R11804	EMPTY	EMPTY	EMPTY	EMPTY
D	Ind11-R8378	Ind11-R8640	Ind11-R8949	Ind11-R9031	Ind2-R11428	Ind4-R11804	Ind7-R13924	Ind10-R11428	EMPTY	EMPTY	EMPTY	EMPTY
E	Ind11-R10294	Ind11-R10715	Ind11-R10942	Ind11-R11166	Ind2-R13924	Ind5-R11428	Ind7-R11804	Ind10-R13924	EMPTY	EMPTY	EMPTY	EMPTY
F	Ind11-R13025	Ind11-R13390	Ind11-R13559	Ind11-R13855	Ind2-R11804			Ind10-R11804	EMPTY	EMPTY	EMPTY	EMPTY
G	Ind11-R15396	Ind11-R15585	Ind11-R15728	Ind11-R16042	Ind3-R11428	Ind5-R11804	Ind8-R13924	Ind11-R11428	EMPTY	EMPTY	EMPTY	EMPTY
H	Ind11-R484	Ind11-R484	Ind11-R599	Ind11-R599	Ind3-R13924	Ind6-R11428	Ind8-R11804	Ind11-R13924	EMPTY	EMPTY	EMPTY	EMPTY

Appendix G. Number and proportion of complete samples by U.S. state/territory

State/Territory	Abbreviation	African American (n=170) number	African American (n=170) proportion	U.S. Caucasian (n=263) number	U.S. Caucasian (n=263) proportion	U.S. Hispanic (n=155) number	U.S. Hispanic (n=155) proportion
Alabama	AL	6	0.035	1	0.004	0	0.000
Alaska	AK	0	0.000	0	0.000	0	0.000
American Samoa	AS	0	0.000	0	0.000	0	0.000
Arizona	AZ	0	0.000	3	0.011	8	0.052
Arkansas	AR	0	0.000	3	0.011	1	0.006
California	CA	5	0.029	18	0.068	39	0.252
Colorado	CO	2	0.012	3	0.011	1	0.006
Connecticut	CT	0	0.000	6	0.023	0	0.000
Delaware	DE	2	0.012	0	0.000	1	0.006
District of Columbia	DC	3	0.018	0	0.000	0	0.000
Federated States of Micronesia	FM	0	0.000	0	0.000	0	0.000
Florida	FL	11	0.065	17	0.065	10	0.065
Georgia	GA	15	0.088	3	0.011	2	0.013
Guam	GU	0	0.000	0	0.000	0	0.000
Hawaii	HI	0	0.000	0	0.000	3	0.019
Idaho	ID	0	0.000	4	0.015	0	0.000
Illinois	IL	7	0.041	12	0.046	4	0.026
Indiana	IN	4	0.024	8	0.030	3	0.019
Iowa	IA	0	0.000	4	0.015	1	0.006
Kansas	KS	1	0.006	1	0.004	2	0.013
Kentucky	KY	1	0.006	4	0.015	0	0.000
Louisiana	LA	12	0.071	3	0.011	0	0.000
Maine	ME	0	0.000	2	0.008	0	0.000
Marshall Islands	MH	0	0.000	0	0.000	0	0.000
Maryland	MD	8	0.047	7	0.027	2	0.013
Massachusetts	MA	1	0.006	7	0.027	1	0.006
Michigan	MI	6	0.035	11	0.042	1	0.006
Minnesota	MN	0	0.000	7	0.027	1	0.006
Mississippi	MS	6	0.035	2	0.008	0	0.000
Missouri	MO	5	0.029	6	0.023	0	0.000
Montana	MT	0	0.000	2	0.008	0	0.000
Nebraska	NE	1	0.006	3	0.011	0	0.000
Nevada	NV	1	0.006	1	0.004	2	0.013
New Hampshire	NH	0	0.000	2	0.008	0	0.000
New Jersey	NJ	3	0.018	8	0.030	4	0.026
New Mexico	NM	0	0.000	0	0.000	3	0.019
New York	NY	9	0.053	21	0.080	7	0.045
North Carolina	NC	15	0.088	4	0.015	1	0.006
North Dakota	ND	0	0.000	0	0.000	0	0.000
Northern Mariana Islands	MP	0	0.000	0	0.000	0	0.000
Ohio	OH	1	0.006	23	0.087	0	0.000
Oklahoma	OK	0	0.000	6	0.023	0	0.000
Oregon	OR	0	0.000	2	0.008	1	0.006
Palau	PW	0	0.000	0	0.000	0	0.000
Pennsylvania	PA	3	0.018	19	0.072	1	0.006
Puerto Rico	PR	0	0.000	0	0.000	8	0.052
Rhode Island	RI	0	0.000	3	0.011	0	0.000
South Carolina	SC	11	0.065	3	0.011	0	0.000
South Dakota	SD	0	0.000	3	0.011	0	0.000
Tennessee	TN	4	0.024	7	0.027	1	0.006
Texas	TX	8	0.047	0	0.000	41	0.265
Utah	UT	0	0.000	1	0.004	0	0.000
Vermont	VT	0	0.000	1	0.004	0	0.000
Virgin Islands	VI	2	0.012	0	0.000	0	0.000
Virginia	VA	13	0.076	11	0.042	3	0.019
Washington	WA	3	0.018	0	0.000	3	0.019
West Virginia	WV	0	0.000	3	0.011	0	0.000
Wisconsin	WI	1	0.006	8	0.030	0	0.000
Wyoming	WY	0	0.000	0	0.000	0	0.000

Appendix H. African American haplotypes

# of Haplotypes	Haplogroup	Sample Name(s)	Haplotype (as differences from the rCRS)
2	V	mtGAfrMD0003 mtGAfrVA0006	263G 309.1C 315.1C 750G 1438G 2706G 4580A 4769G 5902C 7028T 8860G 15326G 15904T 16298C
1	A2	mtGAfrIN0004	73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4129R 4248C 4769G 4824G 7028T 7805A 8027A 8265C 8794T 8860G 10271T 11719A 12007A 12285C 12705T 13020C 14766T 15326G 16111T 16183C 16189C 16193.1C 16223T 16290T 16319A 16362C
1	A2n	mtGAfrNC0013	63A 64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 709A 750G 1438G 1736G 2706G 3849A 4248C 4769G 4824G 5105C 7028T 8027A 8794T 8860G 9344T 10700G 11719A 12007A 12705T 14766T 15326G 16145A 16223T 16290T 16319A 16362C
1	A5a3	mtGAfrKS0001	73G 207R 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2156.1A 2706G 4248C 4655A 4769G 4824G 7028T 8563G 8794T 8860G 9909C 11536T 11647T 11719A 12705T 12909G 14766T 15326G 16093Y 16187T 16223T 16290T 16319A
1	B2e	mtGAfrAL0007	73G 263G 309.1C 309.2C 315.1C 499A 524.1A 524.2C 750G 827G 982G 1438G 2706G 3547G 4769G 4820A 4977C 6119T 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11419C 11719A 13590A 14049T 14308C 14766T 15326G 15535T 16182C 16183C 16189C 16193.1C 16217C 16362C 16519C
1	C1c	mtGAfrCA0001	73G 249- 263G 290- 291- 315.1C 489C 750G 1438G 1888A 2706G 3552A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 9689R 10398G 10400T 10873C 11242T 11719A 11914A 12705T 13263G 14207A 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15930A 16223T 16298C 16325C 16327T
1	D4e2	mtGAfrDC0003	73G 152Y 263G 315.1C 489C 750G 1438G 2706G 3010A 3593C 4769G 4883T 5178A 7028T 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11215T 11719A 12705T 14668T 14766T 14783C 15043A 15301A 15326G 15874G 16223T 16362C
1	F1b1a1a	mtGAfrGA0010	73G 152C 249- 263G 309.1C 309.2C 315.1C 523- 524- 750G 1438G 2706G 3970T 4732G 4769G 5049T 5147A 6392C 6962A 7028T 8860G 10310A 10609C 10976T 11719A 12406A 12633T 12882T 13928C 14476A 14766T 15326G 15954G 16129A 16182C 16183C 16189C 16193.1C 16232A 16249C 16304C 16311C 16344T 16519C
1	H	mtGAfrNY0007	263G 315.1C 750G 1438G 2360C 4769G 6902G 8860G 15326G 16093Y
1	H39	mtGAfrWA0003	239C 263G 309.1C 309.2C 315.1C 750G 1438G 4769G 4890G 8860G 12145Y 15326G 16299G 16519C
1	H6a1b4	mtGAfrVA0004	239C 263G 315.1C 750G 1438G 3915A 4727G 4769G 5567C 8860G 9380A 10589A 15326G 16249C 16362C 16482G
1	K1b1c	mtGAfrTN0003	73G 94A 199C 263G 309.1C 315.1C 750G 1189C 1438G 1811G 2706G 3337A 3480G 4769G 5913A 6935T 7028T 8860G 9055A 9698C 9962A 10289G 10398G 10550G 11299C 11467G 11719A 11914A 12308G 12372A 14167T 14766T 14798C 15326G 15946T 16224C 16311C 16325C 16519C
1	K2a5a1	mtGAfrVI0001	73G 146C 152C 263G 309.1C 315.1C 324T 709A 750G 1438G 1811G 2706G 3480G 4561C 4769G 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11491G 11719A 12308G 12372A 14167T 14766T 14798C 15326G 15803A 16148T 16224C 16235G 16311C 16519C
1	L0a1a2	mtGAfrGA0001	64T 93G 185A 189G 236C 247A 263G 315.1C 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245C 2706G 2758A 2885C 3516A 3594T 3866C 4104G 4312T 4586C 4769G 5096C 5231A 5442C 5460A 5603T 6185C 7028T 7146G 7256T 7521A 8428T 8468T 8566G 8655T 8701G 8860G 9042T 9347G 9540C 9755A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 11176A 11641G 11719A 11914A 12007A 12705T 12720G

			13105G 13276G 13506T 13650T 14308C 14766T 15136T 15326G 15431A 16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16291Y 16311C 16320T 16357C 16519C
1	L0a1a2	mtGAfrMD0006	64T 93G 185A 189G 200G 247A 263G 315.1C 514T 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245C 2706G 2758A 2885C 3516A 3594T 3866C 4104G 4312T 4586C 4769G 5096C 5231A 5442C 5460A 5603T 6185C 7028T 7146G 7256T 7521A 8428T 8468T 8566G 8655T 8701G 8860G 9042T 9347G 9540C 9755A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 10942G 11176A 11641G 11719A 11914A 12007A 12705T 12720G 13105G 13276G 13506T 13650T 14308C 14766T 15136T 15326G 15431A 16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T 16362C
1	L0a1a2	mtGAfrNC0015	64T 93G 185A 189G 200G 236C 247A 263G 309.1C 315.1C 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245C 2706G 2758A 2885C 3516A 3594T 3866C 4104G 4312T 4586C 4769G 5096C 5231A 5442C 5460A 5603T 6185C 7028T 7146G 7256T 7521A 8428T 8468T 8521G 8552C 8566G 8655T 8701G 8860G 9042T 9347G 9540C 9755A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 11176A 11641G 11719A 11914A 12007A 12705T 12720G 13105G 13276G 13506T 13650T 14308C 14766T 15136T 15326G 15431A 16129A 16148T 16168T 16172C 16187T 16188G 16223T 16230G 16311C 16320T
1	L0a1b2	mtGAfrMS0005	93G 95C 185A 189G 236C 247A 263G 315.1C 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245G 2706G 2758A 2885C 3516A 3594T 4104G 4312T 4586C 4769G 5096C 5231A 5442C 5460A 5563A 5603T 5911T 6185C 7028T 7146G 7256T 7521A 7711C 8428T 8468T 8566G 8655T 8701G 8860G 8950A 9042T 9347G 9540C 9755A 9804A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 11176A 11641G 11719A 11914A 12007A 12705T 12720G 13105G 13276G 13506T 13650T 14007G 14106C 14308C 14766T 15136T 15326G 15431A 16093C 16129A 16148T 16168T 16172C 16187T 16188G 16189C 16223T 16230G 16278T 16311C 16320T
1	L0a2a1b	mtGAfrCO0001	64T 93G 152C 189G 199C 236C 247A 263G 309.1C 315.1C 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245G 2706G 2758A 2885C 3516A 3594T 3834A 4104G 4312T 4586C 4769G 5147A 5231A 5442C 5460A 5558G 5603T 5711G 6185C 6257A 7028T 7146G 7256T 7521A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8428T 8460G 8468T 8566G 8655T 8701G 8860G 9042T 9347G 9540C 9755A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 11143T 11172G 11176A 11641G 11719A 11914A 12007A 12705T 12720G 13105G 13276G 13506T 13650T 14182C 14308C 14755G 14766T 15136T 15326G 15431A 16093C 16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T 16519C
1	L1b1a	mtGAfrFL0009	73G 146C 152C 182T 185T 195C 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 13965Y 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16172C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a	mtGAfrTX0003	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13563G 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16311C 16519C
1	L1b1a10	mtGAfrNC0003	73G 151T 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C

			14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16311C 16519C
1	L1b1a14	mtGAfrSC0003	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 3930T 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12512T 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16256T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a17	mtGAfrNC0007	73G 146C 152C 182T 185T 189G 195C 234R 247A 263G 309.1C 315.1C 357G 523- 524- 573.1C 573.2C 573.3C 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7859A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 15629C 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16311C 16519C
1	L1b1a3	mtGAfrAL0001	73G 152C 182T 185T 189G 195C 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 3828G 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506Y 13650T 13789C 13880A 13980A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a3	mtGAfrAL0002	73G 150T 152C 182T 185T 189G 195C 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 13980A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a3a	mtGAfrKY0001	73G 152C 182T 185T 189G 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 3762G 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 7915T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 13980A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a3a	mtGAfrMD0002	73G 152C 182T 185T 189G 195C 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 7915T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 13980A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a3a	mtGAfrVA0012	73G 152C 182T 185T 189G 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7747T 7867T 7915T 8248G 8468T 8655T 8701G 8860G 9540C 10084C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 13980A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C

1	L1b1a4	mtGAfrMI0003	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 5894G 6260A 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8790A 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16114A 16126C 16187T 16189C 16223T 16264T 16265G 16270T 16278T 16293G 16311C 16519C
1	L1b1a4	mtGAfrWA0002	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8790A 8860G 9540C 10172A 10398G 10688A 10810C 10873C 11719A 12477C 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16114A 16126Y 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16399G 16519C
1	L1b1a6	mtGAfrIN0001	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3385G 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 9755A 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14110C 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a6	mtGAfrMO0004	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3385G 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 9755A 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14110C 14178C 14203G 14560A 14766T 14769G 15115C 15271T 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a7	mtGAfrIN0002	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6378C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 15406A 16126C 16189C 16223T 16264T 16278T 16293G 16311C 16519C
1	L1b1a7	mtGAfrNY0004	73G 152C 182T 185T 195C 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6378C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7543G 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16311C 16519C
1	L1b1a7	mtGAfrPA0002	73G 152C 182T 185T 195C 198T 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6378C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9300A 9540C 9855G 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13650T 13789C 13831A 13880A 14178C 14203G 14560A 14766T 14769G 15115C 15326G 15706G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b1a7a	mtGAfrIL0004	73G 152C 182T 185T 195C 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C

			6378C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7854C 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14053G 14178C 14203G 14560A 14766T 14769G 15115C 15326G 16126C 16148T 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1c1b	mtGAfrNC0005	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A 523- 524- 750G 769A 825A 1018A 1291C 1438G 2395- 2706G 2758A 2885C 3594T 3666A 3796T 3843G 4104G 4219A 4688C 4769G 5553C 5876G 5951G 6071C 7028T 7055G 7146G 7256T 7389C 7498R 7521A 8027A 8277C 8278.1C 8278.2C 8468T 8619T 8655T 8701G 8860G 9072G 9467C 9540C 9861C 10084C 10321C 10398G 10586A 10589A 10688A 10810C 10873C 11719A 11899C 12681C 12705T 12810G 13105G 13485G 13506T 13650T 13789C 14000A 14148G 14178C 14393G 14560A 14766T 14911T 15326G 15934R 16086C 16129A 16187T 16189C 16223T 16241G 16278T 16291T 16294T 16311C 16519C
1	L1c1c	mtGAfrCA0004	73G 152C 182T 186A 189C 195C 247A 249- 263G 297G 315.1C 316A 385R 750G 769A 825A 1018A 1438G 2232.1A 2395- 2706G 2758A 2885C 2887- 2888- 3208T 3594T 3666A 3796T 3834A 3843G 4104G 4712T 4767G 4769G 5951G 6071C 6267A 7028T 7055G 7146G 7256T 7389C 7521A 8027A 8387A 8389G 8468T 8655T 8701G 8860G 9072G 9233C 9540C 10321C 10398G 10586A 10595C 10688A 10810C 10873C 11335T 11719A 12705T 12810G 12879C 13105G 13485G 13506T 13650T 13789C 14000A 14148G 14178C 14560A 14766T 14911T 15326G 16172C 16189C 16223T 16278T 16311C 16360T 16519C 16527T
1	L1c2a1a	mtGAfrMI0004	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A 750G 769A 825A 1018A 1420C 1438G 2156.1A 2395- 2706G 2758A 2885C 3594T 3666A 4104G 4769G 5585A 5899.1C 5951G 6071C 6150A 6253C 7028T 7055G 7076G 7146G 7256T 7337A 7389C 7521A 7744C 8027A 8214C 8251A 8468T 8655T 8701G 8784G 8860G 8877C 9072G 9540C 10321C 10398G 10586A 10688A 10792G 10793T 10810C 10873C 11654G 11719A 12049T 12705T 12810G 13105G 13149G 13212T 13281C 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14812T 14911T 15016T 15326G 15784C 16071T 16129A 16145A 16187T 16189C 16213A 16223T 16234T 16265C 16278T 16286G 16294T 16311C 16360T 16519C 16527T
1	L1c2a2	mtGAfrVA0007	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A 709A 750G 769A 825A 942G 1018A 1420C 1438G 1531T 2156.1A 2395- 2706G 2758A 2885C 3438A 3594T 3666A 4104G 4769G 5899.1C 5899.2C 5899.3C 5899.4C 5899.5C 5899.6C 5899.7C 5899.8C 5899.9C 5951G 6071C 6150A 6253C 6990T 7028T 7055G 7070T 7076G 7146G 7160T 7256T 7301G 7337A 7389C 7521A 8027A 8468T 8655T 8784G 8860G 8877C 9072G 9540C 10321C 10398G 10586A 10688A 10792G 10793T 10810C 10873C 11654G 11719A 12049T 12705T 12810G 13105G 13149G 13437C 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14668T 14766T 14911T 15016T 15326G 15784C 16129A 16187T 16189C 16223T 16265C 16278T 16286G 16288C 16294T 16311C 16360T 16519C 16527T
1	L1c2b1a	mtGAfrTX0005	73G 151T 152C 182T 186A 189C 195C 198T 204C 247A 263G 297G 315.1C 316A 508G 750G 769A 825A 1018A 1438G 1709A 2220G 2395- 2706G 2758A 2885C 3594T 3666A 4104G 4221T 4598C 4769G 5580C 5899.1C 5899.2C 5899.3C 5899.4C 5899.5C 5899.6C 5899.7C 5899.8C 5951G 6071C 6150A 6253C 6767G 7028T 7055G 7076G 7146G 7256T 7337A 7389C 7521A 8027A 8296G 8468T 8655T 8701G 8784G 8860G 8877C 9072G 9540C 10031C 10321C 10398G 10586A 10688A 10792G 10793T 10810C 10873C 11164G 11654G 11719A 12049T 12705T 12771A 12810G 12865G 13105G 13149G 13485G 13506T 13650T 13789C 14000A 14071G 14178C 14560A 14687G 14766T 14911T 15326G 16093C 16129A 16187T 16189C 16223T 16265C 16278T 16286G 16294T 16311C 16360T 16362C 16519C 16524G 16527T
1	L1c2b1c	mtGAfrLA0008	73G 152C 182T 186A 189C 195C 198T 247A 263G 297G 309.1C 315.1C 316A 633G 723G 750G 769A 825A 1018A 1438G 2395- 2706G 2758A 2885C 3421A 3594T 3666A 3777C 4104G 4769G 5580C 5894C 5899.1C 5899.2C 5899.3C 5899.4C 5899.5C 5899.6C 5951G 6071C 6150A 6253C 7028T 7055G 7076G 7146G 7256T 7337A 7389C 7521A 8027A 8468T 8655T 8701G 8784G 8860G 8877C 9072G 9540C 10031C 10071C 10321C

			10398G 10586A 10688A 10792G 10793T 10810C 10873C 11164G 11654G 11719A 12049T 12705T 12810G 13105G 13149G 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15326G 15672C 16129A 16187T 16189C 16223T 16265C 16274A 16278T 16286A 16292T 16294T 16311C 16360T 16519C 16527T
1	L1c3a	mtGAfrFL0002	73G 151T 152C 182T 186A 189C 247A 263G 309.1C 315.1C 316A 523- 524- 750G 769A 825A 1018A 1438G 2395- 2706G 2758A 2885C 3594T 3666A 4104G 4769G 5951G 6071C 6221A 6260A 6791G 6917A 7028T 7146G 7256T 7389C 7498A 7521A 7789A 8027A 8468T 8655T 8701G 8860G 9072G 9540C 9966A 10398G 10586A 10688A 10810C 10873C 11302T 11719A 12019T 12501A 12705T 12810G 12906T 13105G 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15226G 15326G 15905C 15978T 16129A 16182C 16183C 16189C 16215G 16223T 16278T 16294T 16311C 16360T 16368C 16519C
1	L1c3a	mtGAfrNJ0001	73G 151T 152C 182T 186A 189C 247A 263G 315.1C 316A 523- 524- 750G 769A 825A 990C 1018A 1438G 2395- 2619G 2706G 2758A 2885C 3594T 3666A 4104G 4769G 4977C 5951G 6071C 6128T 6221A 6260A 6917A 7028T 7146G 7256T 7389C 7498A 7521A 7789A 8027A 8468T 8655T 8701G 8860G 9072G 9540C 9966A 10398G 10586A 10688A 10810C 10873C 11302T 11719A 12019T 12375C 12501A 12705T 12810G 13105G 13365T 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15226G 15326G 15905C 15978T 16093C 16129A 16183C 16189C 16193.1C 16215G 16223T 16278T 16294T 16311C 16360T 16519C
1	L1c3a1a	mtGAfrSC0001	73G 151T 152C 182T 186A 189C 247A 263G 309.1C 315.1C 316A 523- 524- 750G 769A 825A 1018A 1438G 2387C 2395- 2706G 2758A 2885C 3105G 3594T 3666A 4104G 4769G 5951G 6071C 6221A 6260A 6917A 7028T 7146G 7256T 7389C 7498A 7521A 7789A 8027A 8468T 8655T 8701G 8860G 9072G 9540C 9966A 10398G 10586A 10688A 10810C 10873C 11302T 11719A 12019T 12501A 12616C 12705T 12810G 13105G 13204A 13485G 13506T 13593T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15226G 15326G 15905C 15978T 16093C 16129A 16183C 16189C 16193.1C 16215G 16223T 16278T 16294T 16311C 16360T 16519C
1	L1c3b1b	mtGAfrPA0001	73G 151T 152C 182T 186A 189C 247A 263G 309.1C 315.1C 316A 523- 524- 629C 750G 769A 825A 1018A 1438G 1686G 1888C 2083C 2283T 2395- 2706G 2758A 2885C 3210T 3434G 3594T 3666A 4104G 4755C 4769G 5951G 6071C 6182A 6221A 6917A 7028T 7146G 7256T 7389C 7521A 8027A 8251A 8417T 8468T 8655T 8701G 8860G 9072G 9540C 10398G 10586A 10688A 10810C 10873C 11302T 11719A 12400G 12542T 12705T 12810G 13105G 13182C 13485G 13506T 13650T 13789C 13981T 14000A 14178C 14560A 14766T 14794T 14911T 15226G 15326G 15905C 15978T 16017C 16129A 16163G 16187T 16189C 16223T 16278T 16293G 16294T 16311C 16360T 16519C
1	L1c5	mtGAfrLA0004	73G 151T 152C 182T 186A 189C 247A 263G 291T 297G 315.1C 316A 523- 524- 709A 750G 769A 825A 1018A 1438G 2395- 2706G 2758A 2885C 3594T 3666A 4104G 4769G 5390G 5951G 6071C 7028T 7055G 7146G 7256T 7389C 7521A 7762A 8027A 8143C 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8468T 8655T 8701G 8860G 9072G 9540C 9899C 10398G 10586A 10688A 10810C 10873C 11150A 11719A 12425G 12630A 12705T 12810G 13105G 13359A 13368A 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15326G 15449C 15553A 15941C 16114G 16129A 16189C 16223T 16261T 16278T 16311C 16360T 16519C
1	L2a1	mtGAfrNY0001	73G 143A 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4316G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8788T 8860G 9221G 9540C 10115C 10398G 10873C 11404G 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16223T 16278T 16294T 16390A
1	L2a1a	mtGAfrMS0001	73G 146C 152C 195C 263G 309.1C 309.2C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 6626Y 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15629C 15784C 16223T 16278T 16294T 16309G 16390A 16497R 16519C

1	L2a1a1	mtGAfrMO0003	73G 146C 152C 195C 263G 309.1C 315.1C 524.1A 524.2C 524.3A 524.4C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4161T 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 11963A 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15391T 15629C 15784C 16223T 16278T 16294T 16309G 16368C 16390A 16519C
1	L2a1a1	mtGAfrNC0011	73G 146C 152C 195C 263G 309.1C 309.2C 315.1C 467T 524.1A 524.2C 524.3A 524.4C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 6152C 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8616A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15391T 15629C 15784C 16223T 16278T 16294T 16309G 16368C 16390A 16519C
1	L2a1a1	mtGAfrNY0006	73G 146C 152C 195C 263G 309.1C 315.1C 524.1A 524.2C 524.3A 524.4C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 6152C 7028T 7175C 7256T 7274T 7521A 7771G 8078A 8188G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15391T 15629C 15784C 16223T 16278T 16294T 16309G 16368C 16390A 16519C
1	L2a1a1	mtGAfrVA0009	73G 146C 152C 195C 263G 292.1A 292.2T 315.1C 524.1A 524.2C 524.3A 524.4C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 6152C 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11248G 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15391T 15629C 15784C 16129A 16223T 16278T 16294T 16309G 16368C 16390A 16519C
1	L2a1a2	mtGAfrSC0010	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2232.1A 2259T 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10870T 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15629C 15784C 16093C 16223T 16278T 16286T 16294T 16309G 16390A 16519C
1	L2a1a2	mtGAfrTN0004	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7562G 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15629C 15784C 16066G 16223T 16278T 16286T 16294T 16309G 16390A 16519C
1	L2a1a2a1a	mtGAfrMD0005	73G 146C 152C 195C 198T 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4659A 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10454C 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15211T 15244G 15301A 15326G 15421G 15629C 15784C 16223T 16278T 16286T 16294T 16309G 16390A 16519C
1	L2a1a2a1a	mtGAfrTX0002	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10454C 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15211T 15244G 15301A 15326G 15421G 15629C 15784C 16092C 16223T 16278T 16286Y 16294T 16309G 16390A 16519C
1	L2a1a3	mtGAfrMD0004	73G 143A 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 3918A 4104G 4769G 5285G 7028T 7175C 7256T 7274T 7521A 7753T 7771G 8206A 8701G 8860G 9007G 9221G 9540C 10115C 10398G 10873C 11399C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15244G 15301A 15326G 15629C 15747C 15784C 16093C 16223T 16278T 16294T 16309G 16390A 16519C

1	L2a1b	mtGAfrLA0001	73G 146C 152C 195C 198T 207R 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8347G 8701G 8860G 9221G 9540C 10115C 10143A 10398G 10873C 11719A 11914A 11944C 12241- 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16187T 16189C 16223T 16231C 16278T 16294T 16309G 16390A
1	L2a1b	mtGAfrLA0011	73G 143A 146C 152C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9422G 9540C 10115C 10143A 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16166- 16189C 16223T 16278T 16294T 16309G 16325C 16390A
1	L2a1b1	mtGAfrFL0004	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 7830A 8206A 8701G 8860G 9221G 9540C 10115C 10143A 10398G 10873C 11719A 11914A 11944C 12130C 12357G 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15735T 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A 16519C
1	L2a1b1	mtGAfrGA0013	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 7830A 8206A 8701G 8860G 9221G 9540C 10115C 10143A 10398G 10873C 11719A 11914A 11944C 12130C 12357G 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15735T 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A
1	L2a1b1	mtGAfrOH0001	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5583T 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10143A 10398G 10873C 11719A 11914A 11944C 12130C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15735T 15784C 16189C 16192Y 16223T 16278T 16294T 16309G 16390A
1	L2a1c	mtGAfrMI0005	73G 143A 146C 152C 195C 263G 309.1C 315.1C 523- 524- 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8155A 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11404G 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 15884A 16223T 16278T 16294T 16309G 16390A
1	L2a1c	mtGAfrNC0010	73G 143A 146C 152C 195C 200G 263G 309.1C 315.1C 523- 524- 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 3852T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16223T 16278T 16294T 16309G 16390A
1	L2a1c	mtGAfrTX0008	73G 143A 146C 195C 263G 309.1C 315.1C 523- 524- 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4561C 4769G 5774C 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 9947R 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15476A 15784C 16140C 16223T 16278T 16294T 16309G 16390A
1	L2a1c2	mtGAfrVA0005	73G 143A 146C 152C 195C 263G 309.1C 315.1C 644G 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7299G 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 10903T 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 15924G 16213A 16223T 16278T 16294T 16309G 16390A 16519C
1	L2a1c2a	mtGAfrMO0005	73G 143A 146C 152C 195C 263G 309.1C 315.1C 513A 750G 769A 1018A 1438G 2083C 2416C 2706G 2789T 3010A 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 9968A 10115C 10398G 10873C 10903T 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14007G 14249A 14566G 14766T 15301A 15326G 15784C 15924G 16193T 16213A 16223T 16239T 16278T 16294T 16309G 16390A 16519C
1	L2a1c3a1	mtGAfrGA0007	73G 143A 146C 152C 195C 198T 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2702A 2706G 2789T 3010A 3594T 4104G 4769G 5843G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 7858T 8206A 8701G 8860G 9221G 9540C 9932A 10115C 10398G 10873C 11350G 11719A 11914A 11944C 12693G 12705T 13590A 13650T

			13653T 13803G 14374C 14566G 14766T 15115C 15301A 15326G 15394C 15784C 16223T 16278T 16290T 16294T 16309G 16390A 16519C
1	L2a1c3a1	mtGAfrVA0002	73G 143A 146C 152C 195C 198T 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 5843G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 7858T 8206A 8473C 8701G 8860G 9221G 9540C 9932A 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15115C 15283C 15301A 15326G 15784C 16223T 16278T 16290T 16294T 16309G 16390A 16519C
1	L2a1c3b	mtGAfrSC0008	73G 143A 146C 152C 195C 263G 309.1C 315.1C 455.1T 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9096C 9221G 9540C 9545G 9932A 10115C 10873C 11719A 11914A 11944C 12693G 12705T 12950G 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16223T 16278T 16294T 16355T 16390A
1	L2a1c4a1	mtGAfrWA0001	73G 143A 146C 152C 195C 263G 309.1C 315.1C 523- 524- 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 5252A 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10208C 10398G 10873C 11719A 11914A 11923G 11944C 12172G 12354C 12693G 12705T 13590A 13650T 13803G 13952R 14566G 14766T 15301A 15326G 15784C 16223T 16278T 16294T 16309G 16390A 16519C
1	L2a1c5	mtGAfrIL0006	73G 143A 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 5747G 6663G 6713T 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8701G 8860G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13260C 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16129A 16223T 16278T 16294T 16309G 16390A
1	L2a1c5	mtGAfrMO0001	73G 143A 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3010A 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8538C 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12651A 12693G 12705T 13260C 13590A 13650T 13803G 13886C 14566G 14766T 15301A 15326G 15784C 16129A 16223T 16278T 16294T 16309G 16390A
1	L2a1e	mtGAfrSC0007	73G 143A 146C 152C 195C 198T 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3495A 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8790A 8860G 9221G 9230C 9540C 9833C 10115C 10398G 10873C 11719A 11914A 11944C 12630A 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15601C 15758G 15784C 16051G 16223T 16278T 16294T 16309R 16390A 16519C
1	L2a1f	mtGAfrAL0005	73G 146C 152C 195C 198Y 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4083Y 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11176A 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16223T 16232T 16278T 16294T 16309G 16390A
1	L2a1f	mtGAfrGA0004	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8289.10C 8289.11C 8289.12C 8289.13C 8289.14C 8289.15T 8289.16C 8289.17T 8289.18A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11137C 11719A 11914A 11944C 12079T 12693G 12705T 13590A 13650T 13788A 13803G 14566G 14766T 15301A 15326G 15784C 16183C 16189C 16193.1C 16223T 16278T 16294T 16309G 16390A 16519C
1	L2a1f	mtGAfrGA0006	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1118G 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 7837C 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A 16519C

1	L2a1f	mtGAfrGA0012	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7740G 7771G 7853A 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16223T 16278T 16294T 16309G 16390A
1	L2a1f	mtGAfrLA0006	73G 146C 152C 195C 263G 309.1C 315.1C 513A 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12603T 12693G 12705T 13590A 13650T 13803G 13884R 14566G 14766T 15301A 15326G 15784C 16184T 16189C 16191.1C 16192T 16223T 16278T 16294T 16309G 16390A
1	L2a1f	mtGAfrNC0001	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13708A 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16223T 16270T 16278T 16294T 16309G 16390A
1	L2a1f	mtGAfrPA0003	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A 16519C
1	L2a1f1	mtGAfrNV0001	73G 146C 152C 195C 263G 309.1C 315.1C 731G 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13759A 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A
1	L2a1f2	mtGAfrGA0014	73G 146C 152C 195C 263G 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 5663T 7028T 7175C 7256T 7274T 7521A 7771G 7870C 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11692T 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14766T 15301A 15326G 15784C 16093C 16189C 16192T 16223T 16278T 16294T 16309G 16390A
1	L2a1i	mtGAfrMD0001	73G 143A 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 10891G 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13716G 13803G 14566G 14766T 15229C 15301A 15326G 15784C 16183C 16189C 16192T 16223T 16278T 16294T 16362C 16390A
1	L2a1l2	mtGAfrNY0002	73G 146C 152C 195C 263G 309.1C 315.1C 534T 750G 769A 1018A 1438G 2416C 2706G 2789T 3325A 3594T 4104G 4769G 5366A 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A
1	L2a1m1a	mtGAfrFL0008	73G 143R 146C 152C 153G 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8553T 8701G 8860G 9221G 9254G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13326C 13590A 13650T 13803G 13884G 14566G 14766T 15301A 15326G 15784C 16189Y 16223T 16278T 16294T 16309G 16390A
1	L2a1n	mtGAfrAL0003	73G 143A 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4317- 4769G 5147A 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16265G 16270T 16278T 16294T 16390A
1	L2b1b	mtGAfrCO0002	73G 146C 150T 152C 182T 183G 195C 198T 204C 263G 315.1C 385G 418T 523- 524- 750G 769A 1018A 1438G 1442A 1706T 2332T 2358G 2416C 2706G 3594T 4104G 4158G 4370C 4767G 4769G 5027T 5147A 5331A 5460A 5814C 6026A 6629G 6713T 7028T 7256T 7521A 7624A 8080T 8206A 8350G 8387A 8701G 8860G 9221G 9540C

			10115C 10398G 10586C 10828C 10873C 11151T 11719A 11944C 12236A 12705T 12948G 13590A 13650T 13924T 14059G 14305R 14766T 15110A 15217A 15301A 15326G 16114A 16129A 16213A 16223T 16278T 16362C 16390A 16519C
1	L2b2	mtGAfrIL0003	73G 146C 150T 152C 182T 183G 195C 198T 204C 263G 309.1C 315.1C 750G 769A 1018A 1438G 1442A 1706T 2332T 2358G 2416C 2706G 3594T 4104G 4158G 4370C 4767G 4769G 5027T 5331A 5814C 6614C 6713T 6806G 7028T 7256T 7521A 7624A 8080T 8206A 8387A 8503C 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11119T 11719A 11944C 12236A 12705T 12948G 13590A 13650T 13980A 14020A 14059G 14766T 15110A 15217A 15301A 15326G 15604T 16114A 16129A 16213A 16223T 16278T 16390A
1	L2b2	mtGAfrVA0008	73G 146C 150T 152C 182T 195C 198T 204C 263G 315.1C 513A 750G 769A 1018A 1438G 1442A 1706T 2332T 2358G 2416C 2706G 3196A 3594T 4104G 4158G 4370C 4767G 4769G 4823C 5027T 5258G 5331A 5814C 6614C 6681C 6713T 6806G 7028T 7256T 7521A 7624A 8080T 8206A 8387A 8503C 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 12948G 13590A 13650T 14059G 14766T 15110A 15217A 15301A 15318T 15326G 16114A 16129A 16213A 16223T 16265G 16278T 16311C 16368C 16390A 16497G
1	L2b2a	mtGAfrLA0012	73G 146C 150T 152C 182T 195C 198T 204C 263G 315.1C 709A 750G 769A 1018A 1438G 1442A 1706T 2332T 2358G 2416C 2706G 3594T 4104G 4158G 4370C 4767G 4769G 5027T 5331A 5814C 6614C 6713T 6806G 7028T 7256T 7521A 7624A 8080T 8206A 8387A 8503C 8701G 8790A 8860G 9221G 9350G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 12948G 13590A 13650T 13966G 14059G 14407T 14766T 15110A 15217A 15301A 15326G 16114A 16129A 16213A 16223T 16278T 16354T 16390A
1	L2b2a	mtGAfrTX0001	73G 146C 150T 152C 182T 195C 198T 204C 263G 309.1C 315.1C 524.1A 524.2C 709A 750G 769A 1018A 1438G 1442A 1706T 2332T 2358G 2416C 2706G 3594T 4104G 4158G 4370C 4767G 4769G 5027T 5331A 5814C 6614C 6713T 6806G 7028T 7256T 7521A 7624A 8080T 8206A 8387A 8503C 8701G 8790A 8860G 9221G 9350G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 12948G 13590A 13650T 13966G 14059G 14407T 14766T 15110A 15217A 15301A 15326G 16114A 16129A 16213A 16223T 16260T 16278T 16354T 16390A
1	L2c	mtGAfrFL0006	64T 73G 93G 146C 150T 182T 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3316A 3594T 4104G 4769G 6150A 7028T 7256T 7521A 7624A 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12523G 12705T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15317A 15326G 15497A 15849T 15978T 16223T 16278T 16390A
1	L2c	mtGAfrFL0007	73G 93G 95C 146C 150T 152C 182T 195C 198T 263G 309.1C 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 7754A 8020A 8206A 8269A 8701G 8860G 9221G 9540C 9591A 9746R 10115C 10398G 10873C 11074T 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14311C 14750T 14766T 15110A 15217A 15301A 15326G 15849T 16093C 16223T 16278T 16390A 16519C
1	L2c	mtGAfrLA0002	73G 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3591A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8425G 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11020G 11719A 11944C 12236A 12705T 12771A 13368A 13590A 13650T 13928C 13958C 14569A 14766T 15110A 15217A 15301A 15326G 15849T 16223T 16278T 16390A 16519C
1	L2c	mtGAfrMS0003	64T 73G 93G 146C 150T 182T 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 6150A 7028T 7256T 7521A 7624A 8206A 8701G 8860G 9221G 9540C 9935C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15317A 15326G 15497A 15849T 15978T 16223T 16278T 16390A
1	L2c	mtGAfrTX0007	73G 93G 146C 150T 152C 182T 195C 198T 263G 309.1C 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 6150A 7028T 7256T 7521A 7624A 7856G 8206A

			8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13759A 13928C 13958C 14766T 15110A 15217A 15301A 15326G 15497A 16184T 16223T 16278T 16294T 16390A
1	L2c2	mtGAfrGA0002	73G 93G 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1040C 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8348R 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15326G 15849T 16093C 16223T 16264T 16278T 16390A
1	L2c2	mtGAfrIL0007	73G 93G 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1040C 1438G 1442A 2332T 2416C 2706G 3196R 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8701G 8860G 8967T 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15326G 15849T 16223T 16264T 16278T 16390A
1	L2c2a	mtGAfrSC0009	73G 93G 146C 150T 152C 182T 195C 198T 263G 309.1C 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1040C 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14766T 15043A 15110A 15217A 15301A 15326G 15849T 16093C 16223T 16264T 16278T 16390A
1	L2c2a	mtGAfrVA0001	73G 93G 146C 150T 152C 182T 195C 263G 309.1C 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1040C 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8701G 8843C 8860G 9221G 9293T 9462C 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14208Y 14766T 15043A 15110A 15217A 15301A 15326G 15849T 16093Y 16223T 16278T 16390A
1	L2c2b1b	mtGAfrCA0003	73G 93G 146C 150T 152C 182T 183G 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1040C 1438G 1442A 2332T 2416C 2706G 3200A 3594T 3769T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8567C 8701G 8772C 8860G 9063G 9221G 9540C 10115C 10398G 10790C 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15313C 15326G 15749T 15849T 16223T 16264T 16278T 16311C 16390A
1	L2c4	mtGAfrNC0004	73G 89C 93G 146C 150T 152C 182T 195C 198Y 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 5471A 7028T 7256T 7521A 7624A 8206A 8701G 8860G 9055A 9221G 9540C 10115C 10398G 10586A 10873C 11719A 11944C 12236A 12705T 13440T 13590A 13650T 13928C 13958C 14766T 15110A 15217A 15301A 15326G 15849T 16223T 16256Y 16278T 16390A 16519C
1	L2d1a	mtGAfrMS0006	73G 146C 150T 195C 263G 315.1C 456T 750G 769A 870T 1018A 1438G 2159C 2332T 2416C 2706G 3254A 3434G 3594T 3693A 4104G 4769G 6231T 7028T 7256T 7521A 8206A 8701G 8856A 8860G 9007G 9221G 9540C 9554A 9941G 10115C 10398G 10700G 10873C 10955T 11353C 11719A 11944C 12705T 13590A 13650T 14766T 14845T 15263T 15301A 15326G 15458C 15703G 15777C 16129A 16183C 16189C 16193.1C 16278T 16300G 16311C 16354T 16390A 16399G 16519C
1	L3b	mtGAfrMS0004	73G 189G 263G 315.1C 523- 524- 750G 1438G 2706G 3450T 4769G 5773A 6221C 7028T 8251A 8537G 8701G 8860G 9007G 9449T 9540C 10086G 10398G 10873C 11719A 12705T 13105G 13914A 14182C 14766T 15301A 15311G 15326G 15824G 15944- 16048A 16124C 16223T 16278T 16362C
1	L3b	mtGAfrNJ0002	73G 189G 263G 315.1C 523- 524- 593C 750G 1438G 2706G 3450T 4769G 5201C 5773A 6221C 6261A 6620C 7028T 8701G 8860G 9007G 9449T 9540C 10086G 10398G 10873C 11719A 12705T 13105G 13914A 14182C 14766T 15301A 15311G 15326G 15824G 15944- 16048A 16124C 16163G 16223T 16278T 16362C
1	L3b1a	mtGAfrNC0008	73G 263G 315.1C 523- 524- 750G 1438G 2706G 3450T 4769G 5099Y 5773A 6221C 7028T 8701G 8772C 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13105G 13914A 14766T 15301A 15311G 15326G 15434T 15824G 15944- 16124C 16223T 16278T 16301T 16362C 16519C

1	L3b1a10	mtGAfrCA0002	73G 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3450T 4659A 4769G 5773A 5978G 6221C 6510R 7028T 7859A 8701G 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13105G 13914A 14766T 15301A 15311G 15326G 15824G 15944- 16124C 16189C 16223T 16271C 16278T 16362C 16519C 16527T
1	L3b1a1a	mtGAfrNY0005	73G 263G 315.1C 523- 524- 750G 794Y 1438G 2706G 3450T 4769G 5773A 6221C 7028T 8701G 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 11800G 12705T 13105G 13914A 14766T 15301A 15311G 15326G 15824G 15883A 15944- 16124C 16223T 16278T 16362C 16519C
1	L3b1a4	mtGAfrMI0006	73G 189R 263G 315.1C 513A 523- 524- 750G 1710C 2706G 3450T 3645C 4769G 5773A 6221C 7028T 8701G 8860G 9449T 9540C 9605T 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13050G 13105G 13443C 13914A 14766T 15301A 15311G 15326G 15824G 15944- 16124C 16223T 16255A 16278T 16311C 16362C 16519C
1	L3b1a6	mtGAfrLA0007	73G 152C 263G 315.1C 523- 524- 750G 1438G 2706G 3385G 3450T 4769G 5255T 5773A 6221C 7028T 8155A 8701G 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 11914A 13105G 13658T 13914A 14766T 15217A 15301A 15326G 15824G 15944- 16124C 16223T 16278T 16362C 16519C
1	L3b1a6	mtGAfrMS0002	73G 152C 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3385G 3450T 4769G 5255T 5773A 6221C 7028T 8155A 8701G 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13105G 13914A 14766T 15301A 15311G 15314A 15326G 15824G 15944- 16124C 16223T 16278T 16362C 16519C
1	L3b1b	mtGAfrNJ0003	73G 152C 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3450T 4062C 4769G 5773A 6221C 6891G 7028T 8701G 8860G 9079G 9449T 9540C 10086G 10373A 10398G 10873C 11719A 12705T 13105G 13914A 14766T 15301A 15311G 15326G 15824G 15944- 16124C 16223T 16278T 16362C
1	L3b2	mtGAfrDC0001	73G 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3420T 3450T 3645Y 4769G 5231A 5773A 6221C 6917A 7028T 8701G 8860G 9299G 9449T 9540C 10086G 10398G 10640C 10873C 11719A 12705T 13105G 13914A 14766T 14950T 15301A 15311G 15326G 15824G 15944- 16124C 16223T 16271C 16278T 16362C 16527T
1	L3b2	mtGAfrGA0003	73G 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3420T 3450T 4769G 5231A 5773A 6221C 6917A 7028T 8701G 8860G 9299G 9449T 9540C 10086G 10398G 10640C 10873C 11719A 12705T 13105G 13914A 14766T 14950T 15301A 15311G 15326G 15824G 15944- 16124C 16223T 16271C 16278T 16362C 16527T
1	L3b2b	mtGAfrDE0002	73G 146C 263G 315.1C 523- 524- 750G 1438G 2706G 3420T 3450T 4769G 5773A 6221C 7028T 8701G 8860G 9053A 9067G 9449T 9540C 10086G 10398G 10640C 10873C 11719A 12705T 13105G 13914A 14766T 15301A 15311G 15326G 15550T 15824G 15944- 16124C 16223T 16278T 16362C 16527T
1	L3b3	mtGAfrFL0011	73G 185A 189G 263G 315.1C 523- 524- 750G 1438G 2706G 3450T 4769G 5773A 6221C 6527G 7028T 8701G 8860G 9007G 9449T 9540C 10086G 10398G 10873C 11719A 12705T 13105G 13914A 13934T 14182C 14766T 15301A 15311G 15326G 15824G 15944- 16048A 16124C 16223T 16278T 16362C 16519C
1	L3d1'2'3'4'5'6	mtGAfrMD0008	73G 93G 151T 152C 195C 263G 309.1C 315.1C 523- 524- 750G 921C 1438G 2706G 4769G 5147A 5187T 7028T 7424G 7594Y 8618C 8701G 8860G 8950R 9540C 9554A 10398G 10873C 11719A 12705T 13105G 13886C 14188G 14284T 14766T 15301A 15326G 15949A 16124C 16223T 16257T 16519C
1	L3d1'2'3'4'5'6	mtGAfrTX0006	73G 151T 152C 263G 309.1C 315.1C 523- 524- 750G 921C 1438G 2706G 4769G 5147A 5187T 7028T 7424G 8618C 8701G 8860G 9540C 9554A 10398G 10873C 11719A 12705T 13105G 13886C 14188G 14284T 14766T 15301A 15326G 15949A 16124C 16189C 16223T 16257T 16519C
1	L3d1a1	mtGAfrNY0003	73G 152C 263G 315.1C 523- 524- 750G 870Y 921C 1438G 1503A 2706G 3796G 4048A 4769G 5147A 6680C 7028T 7424G 7648T 8618C 8701G 8860G 9540C 10398G 10463C 10873C 11719A 11887A 12705T 13105G 13886C 14284T 14766T 15061G 15301A 15326G 16124C 16223T 16300G 16319A
1	L3d1a1a	mtGAfrVA0003	73G 150T 152C 263G 315.1C 523- 524- 750G 921C 1438G 1503A 2706G 4048A 4203G 4769G 5147A 5471A 5899.1C 6339G 6680C 7028T 7424G 7648T 8618C 8701G 8860G 9540C 10398G 10640C 10873C 10915C 11719A 11887A 12705T 13105G 13886C 14284T 14766T 15301A 15326G 16124C 16223T 16319A

1	L3d1b2	mtGAfrLA0013	73G 150T 152C 195C 263G 315.1C 523- 524- 750G 921C 1438G 2706G 4553C 4769G 5046A 5147A 6680C 7028T 7424G 7609C 8618C 8701G 8860G 9540C 10398G 10873C 11719A 12705T 13105G 13886C 14284T 14766T 15301A 15326G 15514C 16124C 16223T 16292T
1	L3d1b3a	mtGAfrAL0004	73G 146C 152C 263G 315.1C 523- 524- 750G 921C 1438G 2706G 4769G 5046A 5147A 6446A 6641C 6680C 7028T 7424G 8618C 8701G 8860G 9540C 10398G 10873C 11719A 12705T 13105G 13886C 14284T 14634C 14766T 15110A 15301A 15326G 15748C 16124C 16171G 16223T 16256T
1	L3d1d	mtGAfrVA0013	73G 152C 263G 309.1C 315.1C 523- 524- 750G 921C 1438G 2706G 4769G 5147A 6680C 7028T 7424G 7765G 8618C 8701G 8709T 8860G 8945C 9151G 9540C 10398G 10873C 11719A 12654R 12705T 13105G 13632G 13886C 14128G 14284T 14766T 15301A 15326G 16124C 16223T 16256T 16368C 16519C
1	L3e1	mtGAfrLA0005	73G 150T 200G 263G 309.1C 315.1C 595.1A 750G 1438G 2352C 2416C 2706G 4769G 6221C 6587T 7028T 8701G 8860G 9540C 9932A 10398G 10819G 10873C 11719A 12705T 14152G 14212C 14766T 15301A 15326G 15670C 15942C 16189C 16223T 16327T 16343T
1	L3e1a1a	mtGAfrTN0001	73G 150T 185A 189G 200G 263G 315.1C 750G 1438G 2352C 2706G 3438A 4769G 6221C 6587T 7028T 8650T 8701G 8745R 8860G 9540C 10398G 10819G 10873C 11719A 12007R 12705T 14152G 14212C 14766T 15301A 15326G 15670C 15942C 16169T 16185T 16223T 16311C 16327T
1	L3e1e	mtGAfrCA0005	73G 150T 189G 200G 263G 309.1C 315.1C 750G 1438G 2352C 2706G 4769G 6221C 6587T 7028T 8701G 8860G 9389G 9540C 10370C 10398G 10819G 10873C 11719A 12705T 14152G 14212C 14753T 14766T 15301A 15326G 15670C 15942C 16223T 16309G 16327T
1	L3e2a1b	mtGAfrMI0001	73G 150T 195C 198T 263G 315.1C 750G 1438G 2352C 2706G 4769G 4823C 6413C 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14212C 14766T 14869A 14905A 15301A 15326G 16223T 16311C 16320T 16519C
1	L3e2a1b	mtGAfrSC0002	73G 150T 195C 198T 263G 315.1C 750G 942G 1438G 2352C 2706G 4769G 4823C 6413C 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14212C 14766T 14869A 14905A 15301A 15326G 16223T 16320T 16519C
1	L3e2a1b1	mtGAfrFL0010	73G 150T 195C 198T 263G 315.1C 499A 750G 1438G 2352C 2706G 4769G 4823C 6413C 7028T 7852A 8701G 8860G 9006G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14212C 14766T 14869A 14905A 15301A 15326G 16086C 16223T 16320T 16399G 16519C
1	L3e2a1b1	mtGAfrGA0015	73G 150T 195C 198T 263G 315.1C 499A 750G 1438G 2352C 2706G 4769G 4823C 6413C 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12545T 12705T 13105G 13958C 14212C 14766T 14869A 14905A 15301A 15326G 16223T 16320T 16399G 16519C
1	L3e2a1b1	mtGAfrIN0003	73G 150T 152Y 195C 198T 251A 263G 315.1C 499A 750G 1438G 2352C 2706G 4769G 4823C 5508C 6413C 6480A 7028T 8701G 8860G 9156G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14212C 14364A 14766T 14869A 14905A 15301A 15326G 16093C 16223T 16311C 16320T 16399G 16519C
1	L3e2a1b1	mtGAfrMI0002	73G 150T 195C 198T 263G 315.1C 499A 709A 723G 750G 1438G 2352C 2706G 4769G 4823C 6413C 6680C 7028T 8701G 8764A 8860G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14180C 14212C 14323A 14766T 14869A 14905A 15301A 15310C 15326G 15734A 16093C 16223T 16293T 16320T 16399G 16519C
1	L3e2a1b3	mtGAfrIL0001	73G 150T 195C 198T 263G 309.1C 315.1C 750G 1438G 1888A 2352C 2706G 4769G 4823C 6413C 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11699G 11719A 12705T 13105G 13889A 14212C 14750G 14766T 14869A 14905A 15301A 15326G 16086C 16223T 16320T 16519C
1	L3e2b	mtGAfrFL0005	73G 150T 152C 195C 263G 309.1C 315.1C 750G 1438G 2352C 2706G 4769G 7028T 7746G 8701G 8860G 9531G 9540C 10398G 10819G 10873C 11719A 12705T 14212C 14766T 14905A 15301A 15326G 16172C 16183C 16189C 16193.1C 16223T 16320T 16519C

1	L3e2b	mtGAfrGA0005	73G 150T 195C 263G 309.1C 315.1C 750G 1438G 2352C 2706G 4769G 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 14212C 14766T 14905A 15301A 15326G 15930A 16172C 16183C 16189C 16193.1C 16223T 16320T 16519C
1	L3e2b	mtGAfrGA0009	73G 150T 195C 263G 315.1C 750G 769A 1438G 2352C 2706G 4164G 4769G 6663G 7028T 8701G 8860G 9302T 9540C 10398G 10819G 10873C 11719A 12705T 14212C 14766T 14905A 15301A 15326G 16172C 16182C 16183C 16189C 16223T 16320T 16519C
1	L3e2b	mtGAfrSC0011	73G 146C 150T 195C 263G 315.1C 335G 750G 769A 1438G 2352C 2706G 4769G 7028T 8701G 8860G 9540C 10018G 10398G 10819G 10873C 11719A 12705T 14212C 14364A 14766T 14905A 15301A 15326G 16172C 16183C 16189C 16193.1C 16223T 16320T 16519C
1	L3e2b	mtGAfrTX0004	73G 150T 195C 263G 315.1C 750G 1438G 2352C 2706G 4769G 5186G 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11431Y 11719A 11963A 12705T 14212C 14766T 14905A 15301A 15326G 16172C 16183C 16189C 16193.1C 16223T 16320T 16519C
1	L3e2b	mtGAfrVI0002	73G 150T 152C 195C 263G 315.1C 750G 1438G 2352C 2706G 4769G 5899.1C 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 13528G 14212C 14766T 14905A 15301A 15326G 16172C 16183C 16189C 16193.1C 16223T 16320T 16519C
1	L3e3a	mtGAfrNC0002	73G 150T 152C 195C 263G 309.1C 315.1C 499A 523- 524- 573.1C 573.2C 573.3C 573.4C 573.5C 1438G 2000T 2352C 2706G 4769G 5262A 6524C 7028T 8701G 8860G 9540C 9554A 10286G 10398G 10667C 10816G 10819G 10873C 11719A 12397G 12705T 13101C 14212C 14766T 15301A 15326G 16223T 16265T 16311C 16519C
1	L3e3b	mtGAfrGA0011	73G 150T 195C 263G 315.1C 523- 524- 1438G 2000T 2352C 2706G 4655A 4769G 5262A 6524C 7028T 8701G 8860G 9440T 9540C 9554A 10398G 10667C 10816G 10819G 10873C 11719A 12248G 12705T 13101C 13116T 13197T 13651G 14212C 14766T 15245A 15301A 15326G 15812A 16093Y 16148T 16223T 16265T 16311C 16519C
1	L3e3b	mtGAfrNC0014	73G 150T 195C 263G 315.1C 523- 524- 1438G 2000T 2352C 2706G 4655A 4769G 5262A 6524C 7028T 8701G 8860G 9540C 9554A 10398G 10667C 10816G 10819G 10873C 11719A 12248G 12705T 13101C 13197T 13651G 14212C 14766T 15301A 15326G 15812A 16093C 16223T 16265T 16519C
1	L3e3b3	mtGAfrTN0002	73G 150T 185A 189G 263G 315.1C 523- 524- 1438G 2000T 2352C 2706G 4655A 4769G 5262A 6524C 7028T 8701G 8860G 9540C 9554A 10398G 10667C 10816G 10819G 10873C 11719A 12248G 12705T 13101C 13197T 13651G 14212C 14750G 14766T 15043R 15301A 15326G 15812A 16183C 16189C 16193.1C 16223T 16265T 16465T
1	L3e3b3	mtGAfrVA0011	73G 150T 263G 315.1C 523- 524- 980C 1438G 2000T 2352C 2706G 4655A 4769G 5262A 6524C 7028T 8701G 8860G 9540C 9554A 10398G 10667C 10816G 10819G 10873C 11016A 11719A 12248G 12705T 13101C 13197T 13651G 14212C 14766T 15301A 15326G 15812A 16189C 16223T 16265T 16465T
1	L3e4a	mtGAfrIL0005	73G 150T 263G 309.1C 315.1C 523- 524- 709A 1438G 2352C 2706G 3316A 3915A 4769G 5262A 5584G 7028T 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8701G 8860G 9540C 10398G 10819G 10873C 11257T 11719A 12705T 12771A 13749T 14212C 14766T 15301A 15326G 16051G 16169S 16223T 16264T 16265R 16266Y 16519C
1	L3f1b	mtGAfrFL0003	73G 150T 189G 200G 207A 263G 309.1C 315.1C 750G 1438G 1719A 1822C 2706G 3396C 4218C 4769G 5601T 6650T 7028T 7379A 7819A 7852A 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10670T 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16209C 16223T 16274A 16292T 16311C 16519C
1	L3f1b	mtGAfrIL0002	73G 150T 189G 200G 207A 263G 309.1C 315.1C 750G 1406C 1438G 1822C 2706G 3396C 4218C 4769G 5601T 7028T 7819A 8075R 8503Y 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10670T 10873C 11440A 11719A 12705T 13263G 14766T 14769G 15301A 15326G 15514C 15944- 16209C 16223T 16235G 16274A 16292T 16311C 16519C

1	L3f1b	mtGAfrLA0010	73G 150T 189G 200G 215G 263G 315.1C 750G 1438G 1822C 2706G 3396C 3777C 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10873C 11440A 11719A 12705T 13879C 14766T 14769G 15301A 15326G 15514C 15944- 16140C 16189C 16209C 16223T 16292T 16311C 16519C
1	L3f1b	mtGAfrNC0012	73G 150T 189G 200G 207A 263G 315.1C 523- 524- 750G 1438G 1822C 1958R 2706G 3396C 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15244G 15301A 15326G 15514C 15944- 16209C 16223T 16266T 16274A 16292T 16311C 16519C
1	L3f1b1a	mtGAfrDC0002	73G 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 4136G 4218C 4769G 5601T 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12705T 13464T 14766T 14769G 15301A 15326G 15514C 15944- 16051G 16129A 16153A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a	mtGAfrDE0001	73G 189G 200G 204C 263G 309.1C 309.2C 315.1C 523- 524- 750G 1438G 1822C 2706G 3396C 4218C 4769G 5601T 6722A 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12678C 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16129A 16183C 16189C 16193.1C 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a	mtGAfrFL0001	73G 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 4218C 4769G 5601T 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16129A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a	mtGAfrNC0009	73G 189G 200G 207A 263G 315.1C 750G 1438G 1822C 2706G 3396C 4218C 4769G 5601T 5899.1C 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16129A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a	mtGAfrSC0004	73G 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 4218C 4769G 5601T 6581R 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16129A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a	mtGAfrSC0006	73G 189G 200G 263G 315.1C 750G 1438G 1822C 2706G 3396C 4218C 4769G 5601T 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 11722C 12705T 14766T 14769G 15301A 15326G 15461C 15514C 15944- 16129A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b1a1	mtGAfrNE0001	73G 189G 200G 263G 272G 309.1C 315.1C 482Y 750G 1438G 1822C 2706G 3396C 4218C 4743A 4769G 5601T 7028T 7819A 8410T 8527G 8701G 8860G 8932T 9540C 9950C 10070T 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15080R 15301A 15326G 15514C 15944- 16129A 16209C 16223T 16292T 16295T 16311C 16519C
1	L3f1b4	mtGAfrMA0001	73G 150T 189G 200G 263G 309.1C 309.2C 315.1C 750G 1438G 1822C 2706G 3396C 3505G 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 8937A 9540C 9950C 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15629C 15944- 16209C 16223T 16311C 16519C
1	L3f1b4a1	mtGAfrNY0009	73G 150T 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 3505G 3705R 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8799G 8860G 8932T 9540C 9950C 10398G 10873C 11440A 11719A 12705T 13167G 14766T 14769G 15301A 15326G 15514C 15944- 16209C 16223T 16311C 16519C
1	L3f1b4c	mtGAfrMD0007	73G 150T 152C 189G 200G 234R 263G 315.1C 750G 1438G 1822C 2706G 3396C 3505G 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15565Y 15944- 16209C 16218T 16223T 16291T 16292T 16295T 16311C 16519C
1	L3f1b4c	mtGAfrSC0005	73G 150T 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 3505G 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 9540C 9950C 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15944- 16209C 16218T 16223T 16292T 16311C 16352C 16519C

1	L3h1b1a	mtGAfrNC0006	73G 189C 195C 263G 315.1C 523- 524- 750G 1438G 1719A 2706G 2831A 3777C 4388G 4769G 5300T 6756C 7028T 7055G 7861C 8701G 8767G 8860G 9509C 9540C 9575A 9827T 10044G 10289G 10398G 10873C 11563T 11590G 11719A 11963A 12705T 14410A 14766T 15301A 15326G 16126C 16179T 16215G 16223T 16256A 16284G 16311C
1	L3k1	mtGAfrMO0002	73G 150T 152C 235G 263G 315.1C 494T 735G 750G 1438G 2706G 3918A 4313C 4769G 6620G 7028T 7428R 8649G 8701G 8860G 9007T 9329A 9467C 9540C 10398G 10819G 10873C 11719A 12705T 13135A 13477A 13542G 13862G 13992T 14766T 15301A 15314A 15326G 16075C 16129A 16223T 16355T
1	M1a1d	mtGAfrVA0010	73G 150T 189G 195C 263G 309.1C 315.1C 489C 523- 524- 750G 813G 1438G 2706G 3421A 3705A 4769G 6261A 6446A 6671C 6680C 7028T 8270T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8701G 8860G 9540C 10398G 10400T 10873C 11719A 12346T 12403T 12705T 12950C 14110C 14766T 14783C 15043A 15301A 15326G 15672C 16129A 16189C 16223T 16249C 16311C 16359C 16519C
1	M7c1c3	mtGAfrLA0009	73G 146C 199C 263G 309.1C 315.1C 489C 523- 524- 750G 1438G 2706G 3606G 4071T 4769G 4850T 5442C 6455T 7028T 8701G 8860G 9540C 9824C 10398G 10400T 10873C 11665T 11719A 12091C 12705T 13708A 14766T 14783C 15043A 15236G 15301A 15326G 16223T 16295T 16362C 16519C
1	N1b1b1	mtGAfrNY0008	73G 152C 263G 309.1C 315.1C 750G 1438G 1598A 1703T 1719A 2639T 2706G 3921A 4735A 4769G 4917G 4960T 5471A 7028T 8251A 8472T 8836G 8860G 10238C 11719A 11928G 12092T 12501A 12705T 12822G 13129T 13710G 14581C 14766T 15326G 16145A 16176A 16223T 16390A 16519C
1	U6a3c	mtGAfrGA0008	73G 146C 185A 263G 291.1A 315.1C 523- 524- 750G 960- 1438G 1809C 2706G 3348G 4769G 5147A 5554A 6182A 7028T 7805A 8860G 11272G 11467G 11719A 12308G 12372A 12954C 14179G 14766T 15326G 15380G 15790T 16172C 16183C 16189C 16193.1C 16219G 16278T 16519C
1	X2a1a	mtGAfrWI0001	73G 143A 153G 195C 200G 204Y 263G 309.1C 309.2C 309.3C 315.1C 573.1C 750G 1438G 1719A 2393T 2706G 3552C 4769G 6113G 6221C 6371T 7028T 8860G 8913G 11719A 12397G 12705T 13966G 14470C 14502C 14766T 15326G 16093C 16182C 16183C 16189C 16193.1C 16193.2C 16223T 16278T 16357C 16519C

Appendix I. U.S. Caucasian haplotypes

# of Haplotypes	Haplogroup	Sample Name(s)	Haplotype (as differences from the rCRS)
2	H4a1a2a	mtGCaucMA0002 mtGCaucPA0001	263G 315.1C 523- 524- 750G 1438G 3992T 4024G 4769G 5004C 8269A 8860G 9123A 10124C 13708A 14365T 14582G 14956C 15326G
2	J1c3b	mtGCaucNJ0006 mtGCaucPA0020	73G 185A 263G 295T 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 10398G 11251G 11719A 12612G 13708A 13934T 14766T 14798C 15326G 15367T 15452A 16069T 16126C
1	A10	mtGCaucNY0006	73G 235G 263G 315.1C 523- 524- 544T 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 5393C 7028T 7468T 8794T 8860G 9948A 10094T 11719A 12705T 14766T 15326G 16223T 16227C 16290T 16311C 16319A 16519C
1	A2q	mtGCaucCA0004	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 523- 524- 663G 750G 1438G 1736G 1746G 2706G 3633C 4248C 4769G 4824G 4961G 4976G 5899.1C 7028T 8027A 8794T 8860G 8946T 9122Y 9893T 11719A 12007A 12705T 14154G 14766T 15326G 16111T 16129A 16209C 16223T 16290T 16319A 16362C 16519C
1	A2w	mtGCaucIL0003	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 7124G 8027A 8794T 8860G 11016A 11719A 12007A 12705T 14766T 15326G 16111T 16223T 16290T 16319A 16362C
1	B2a1a1	mtGCaucCA0009	73G 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10813T 10895G 11177T 11719A 12729G 13135A 13590A 15326G 15535T 16111T 16183C 16189C 16193.1C 16217C 16483A
1	B2a4a	mtGCaucCO0002	73G 215G 228A 263G 309.1C 309.2C 315.1C 499A 750G 827G 1393A 1438G 2706G 3547G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8817G 8860G 9770C 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16092C 16111T 16129A 16182C 16183C 16189C 16193.1C 16217C 16320Y 16483A 16519C
1	B2s	mtGCaucTN0005	73G 263G 310C 314- 315- 499A 750G 827G 930A 1438G 2706G 3547G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8567C 8860G 9950C 11177T 11719A 12616C 13590A 13740C 14766T 15326G 15535T 16152C 16182C 16183C 16189C 16193.1C 16217C 16325C 16519C
1	B4c2	mtGCaucCA0013	44.1C 73G 263G 309.1C 315.1C 524.1A 524.2C 750G 1119C 1438G 2706G 4769G 5108C 6932G 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 11719A 14088C 14178C 14209G 14766T 15326G 15346A 16147T 16183C 16184A 16189C 16217C 16235G 16309G 16519C
1	C1c4	mtGCaucCA0001	73G 200G 204C 214G 249- 263G 290- 291- 309.1C 315.1C 489C 750G 1438G 1888A 2706G 3552A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9156G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15930A 16223T 16256T 16274A 16298C 16325C 16327T
1	C4c1b	mtGCaucMI0011	73G 249- 263G 309.1C 315.1C 489C 750G 1243C 1438G 2232.1A 2706G 3552A 4715G 4769G 6026A 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 11969A 12705T 13263G 14208C 14318C 14433T 14766T 14783C 15043A 15148A 15204C 15301A 15326G 15487T 16203G 16223T 16298C 16327T 16362C 16519C
1	F1a4a1	mtGCaucCA0010	73G 146C 152C 249- 263G 309.1C 309.2C 315.1C 521- 522- 523- 524- 750G 1438G 2706G 3970T 4086T 4769G 5985A 6392C 6719C 6962A 7028T 8277C 8278.1C 8278.2C 8278.3C 8860G 9053A 9548A 10310A 10609C 11719A 12406A 12882T 13422G 13434G 13759A 13928C 14766T 15326G 15445C 16129A 16172C 16294T 16304C 16362C 16519C

1	G2a1	mtGCaucCA0019	73G 263G 315.1C 489C 523- 524- 709A 750G 1438G 2706G 4769G 4833G 5108C 5601T 7028T 7600A 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8701G 8860G 9377G 9540C 9575A 10398G 10400T 10873C 11719A 12705T 13563G 14200C 14569A 14766T 14783C 15043A 15301A 15326G 16086C 16172C 16223T 16227G 16278T 16362C
1	H	mtGCaucFL0012	195C 263G 315.1C 523- 524- 750G 1438G 4769G 8860G 10658G 15326G 16311C 16519C
1	H	mtGCaucMD0007	73G 263G 315.1C 750G 1438G 4769G 8860G 11134G 15326G 16223T
1	H	mtGCaucPA0012	263G 315.1C 750G 1438G 3221G 4769G 5460A 8860G 9299G 15326G 15565C 16519C
1	H1	mtGCaucFL0003	263G 315.1C 750G 1438G 3010A 4769G 8860G 11293T 14502C 15326G 16519C
1	H1	mtGCaucKY0003	263G 315.1C 750G 1438G 2746C 3010A 4769G 8860G 13708A 13746T 15261A 15326G 15924G 16093C 16189C 16193.1C 16193.2C 16368C 16519C
1	H1	mtGCaucMA0005	263G 309.1C 309.2C 315.1C 750G 1438G 1462A 3010A 4769G 8860G 15326G 16126C 16189C 16193.1C 16193.2C 16293C 16400T 16519C
1	H1	mtGCaucNY0009	263G 315.1C 750G 1438G 3010A 4769G 5774C 8860G 15326G 16519C
1	H1	mtGCaucOH0009	263G 315.1C 750G 1438G 3010A 4769G 8860G 15326G 15785Y 16239T 16519C
1	H1	mtGCaucWI0004	263G 309.1C 315.1C 750G 1438G 1692R 3010A 4769G 6253C 8860G 15326G 16519C
1	H10e1a	mtGCaucPA0005	152C 263G 279Y 309.1C 315.1C 750G 1438G 4769G 8860G 10410C 12284A 13830C 14470A 15326G 16093C 16221T 16266T 16519C
1	H11a	mtGCaucMD0008	195C 263G 309.1C 309.2C 315.1C 750G 961G 1438G 4769G 8448C 8860G 13759A 15326G 16293G 16311C
1	H11a1	mtGCaucNY0023	195C 263G 315.1C 750G 961G 1438G 4769G 5021C 8448C 8860G 8898T 13759A 15326G 16183C 16189C 16193.1C 16278T 16293G 16311C 16519C
1	H11a2a	mtGCaucSC0002	195C 263G 315.1C 750G 961G 1438G 4769G 8448C 8860G 13759A 14587G 15326G 16092C 16140C 16293R 16311C
1	H13a1a4	mtGCaucNJ0003	263G 309.1C 315.1C 750G 1438G 2259T 4745G 4769G 8381G 8860G 13680T 14872T 15326G
1	H14a2	mtGCaucVA0009	263G 315.1C 750G 1438G 1462A 4769G 6182A 7645C 8227C 8860G 10217G 15326G 16256T 16352C
1	H14b	mtGCaucPA0016	263G 309.1C 315.1C 523- 524- 750G 1438G 3197C 4012G 4769G 7403G 7645C 8860G 10217G 14887G 15326G 16192T 16519C
1	H16d	mtGCaucVA0010	152C 263G 315.1C 750G 1438G 4769G 8860G 10394T 14155T 14869C 15326G 16519C
1	H18	mtGCaucFL0008	263G 309.1C 309.2C 315.1C 524.1A 524.2C 750G 1438G 4769G 5585A 6152C 8860G 13708A 14364A 15326G 16311C 16519C
1	H1a	mtGCaucOH0018	73G 263G 315.1C 750G 1438G 3010A 4769G 8601G 8860G 15326G 16162G 16519C
1	H1a	mtGCaucVA0012	73G 263G 309.1C 315.1C 750G 1438G 3010A 4769G 5460A 8860G 15326G 16162G 16519C
1	H1a	mtGCaucWI0001	73G 263G 315.1C 750G 1018A 1438G 3010A 4769G 5093C 8860G 15326G 16162G 16519C
1	H1a1	mtGCaucNC0002	73G 263G 315.1C 750G 1438G 3010A 4769G 6365C 8860G 15326G 16162G 16209C 16519C
1	H1ac	mtGCaucPA0017	73G 263G 309.1C 315.1C 750G 1438G 3010A 4769G 8860G 11447A 11893G 14979G 15326G 16189C 16519C
1	H1ag	mtGCaucSD0003	152C 263G 309.1C 315.1C 750G 789C 1438G 3010A 4769G 5752- 8860G 14869A 15326G 16519C
1	H1ag1	mtGCaucLA0003	263G 315.1C 750G 1438G 3010A 4769G 6272G 8860G 10259R 14869A 15326G 16267T 16519C
1	H1ag1	mtGCaucPA0010	189R 263G 309.1C 315.1C 750G 1438G 2885C 3010A 4769G 6272G 7153C 8860G 14869A 15326G 16519C
1	H1ak1	mtGCaucOH0001	195C 263G 309.1C 309.2C 315.1C 750G 1438G 3010A 3666A 4769G 5393C 8860G 12094T 15326G 16519C
1	H1au1b	mtGCaucIL0004	263G 315.1C 750G 1438G 1719A 3010A 4769G 8860G 11515A 15326G 16148T 16519C
1	H1b	mtGCaucID0003	140T 263G 315.1C 750G 1438G 3010A 4769G 8860G 15326G 16189C 16356C 16362C 16519C
1	H1ba	mtGCaucFL0011	263G 309.1C 315.1C 750G 1438G 3010A 4769G 6569T 8860G 10101C 15326G 16270T
1	H1c1	mtGCaucAR0001	263G 315.1C 477C 750G 1438G 1632Y 3010A 4769G 8860G 9150G 15326G 16093C 16209C 16263C 16519C
1	H1c1	mtGCaucCT0004	263G 315.1C 477C 750G 1438G 3010A 4769G 8860G 9150G 15326G 16263C 16519C

1	H1c1	mtGCaucGA0002	263G 315.1C 477C 750G 1438G 3010A 4769G 8531R 8860G 9150G 15326G 16263C 16519C
1	H1c1	mtGCaucGA0003	263G 309.1C 315.1C 477C 750G 1438G 3010A 4769G 8860G 9150G 15326G 16263C 16519C
1	H1c2	mtGCaucFL0004	44.1C 263G 315.1C 477C 750G 1438G 3010A 4769G 8860G 12858T 15326G 16519C
1	H1c2	mtGCaucNH0001	263G 309.1C 309.2C 315.1C 477C 750G 1193C 1438G 3010A 4769G 5249C 8860G 12858T 15326G 16390R 16519C
1	H1c3	mtGCaucTN0001	152C 195C 257G 263G 309.1C 315.1C 477C 750G 1438G 3010A 4769G 8473C 8860G 8931G 15326G 16519C
1	H1c4b	mtGCaucOH0023	263G 315.1C 477C 750G 1438G 3010A 4769G 8860G 13759A 15326G 16311C 16355T 16400Y 16519C
1	H1e1b1a	mtGCaucID0004	263G 309.1C 315.1C 453C 750G 1438G 3010A 4769G 5460A 8023C 8512G 8860G 13590A 15326G 15692G 16519C
1	H1e3	mtGCaucWI0007	93G 199C 207A 263G 309.1C 315.1C 709R 750G 960.1C 1438G 2885C 3010A 4769G 5460A 8860G 15326G 15346A 16093C 16519C
1	H1j	mtGCaucMA0003	263G 309.1C 315.1C 750G 1438G 3010A 4733C 4769G 8860G 13911G 14025C 15326G 16519C
1	H1j3	mtGCaucCO0001	263G 315.1C 750G 1438G 3010A 4733C 4769G 5249C 8860G 15326G 16519C
1	H1n	mtGCaucLA0002	146C 263G 309.1C 309.2C 315.1C 750G 1438G 2098A 3010A 4769G 8860G 15326G 16519C
1	H1n2	mtGCaucRI0002	64Y 146C 152C 200G 263G 309.1C 315.1C 750G 1438G 2098A 3010A 4769G 5483C 8277C 8860G 11227T 13431T 15244G 15326G 16519C
1	H1q	mtGCaucPA0007	263G 315.1C 750G 1438G 3010A 4769G 4859C 8860G 15326G 15954G 16188G 16519C
1	H1q	mtGCaucSC0001	263G 309.1C 315.1C 750G 1438G 3010A 4769G 4859C 4973Y 6481Y 8860G 9948A 10586A 15301A 15326G 16093C 16519C
1	H1u	mtGCaucFL0014	263G 309.1C 315.1C 523- 524- 750G 1438G 3010A 3483A 4769G 8573A 8860G 9923T 13708A 15326G 16266T 16320T 16519C
1	H23	mtGCaucNY0017	263G 315.1C 750G 1438G 4769G 8860G 10211T 15326G 16318G 16519C
1	H24a	mtGCaucCT0005	263G 315.1C 750G 1438G 3333T 4769G 7598A 8860G 13473R 15326G 15394C 16519C
1	H24a	mtGCaucIL0009	263G 315.1C 750G 1438G 3333T 4769G 7083G 8860G 13105G 15326G 16293G 16519C
1	H24a	mtGCaucMT0002	263G 315.1C 750G 1438G 3333T 4769G 8860G 15326G 16293G 16519C
1	H27a	mtGCaucPA0011	152C 263G 309.1C 315.1C 750G 1438G 4769G 8860G 9391T 11719A 15326G 16093C 16129A 16316G 16519C
1	H2a2a1a	mtGCaucIL0010	309.1C 315.1C 9126A 14279A 15314A
1	H2a2a1b	mtGCaucMN0001	9299G
1	H2a5	mtGCaucNY0005	263G 309.1C 315.1C 750G 6179A 8860G 13708A 13962A 15326G 16519C
1	H3	mtGCaucWV0002	150T 152C 263G 315.1C 750G 1438G 4769G 6776C 8251A 8556C 8860G 15326G 16519C
1	H31a	mtGCaucMA0008	72G 146C 195C 263G 309.1C 315.1C 750G 1438G 4769G 7930T 8860G 10771G 11893R 15326G 16319A 16519C
1	H33c	mtGCaucNY0007	263G 309.1C 315.1C 573.1C 573.2C 709A 750G 1438G 4769G 8860G 10373A 11447A 14569A 15326G 16188T 16519C
1	H3ap	mtGCaucCA0017	146C 263G 309.1C 315.1C 750G 1438G 4769G 6776C 8464Y 8860G 10915C 15314A 15326G 16114T 16519C
1	H3g	mtGCaucID0002	152C 263G 315.1C 750G 1438G 4769G 6776C 8860G 10754C 12362T 15326G 16519C
1	H3k1a	mtGCaucIN0003	152C 263G 315.1C 750G 1438G 3591A 4769G 6776C 8860G 11590G 12217G 14687G 15326G 16519C
1	H3k1a	mtGCaucIN0007	152C 263G 315.1C 750G 1438G 3591A 4561Y 4769G 6776C 8860G 11590G 12217G 14687G 15326G 16519C
1	H3m	mtGCaucIN0001	195C 263G 315.1C 750G 1438G 4263G 4769G 6776C 8860G 13656Y 14501G 15326G 16311C 16519C
1	H47a	mtGCaucCA0003	152C 263G 315.1C 523- 524- 750G 1438G 4769G 8756C 8860G 8986G 9530C 12633T 13020A 15326G
1	H4a1a1a	mtGCaucMA0001	73G 263G 309.1C 315.1C 523- 524- 750G 1438G 3992T 4024G 4769G 5004C 8269A 8860G 9123A 10044G 14365T 14582G 15326G

1	H4a1a4b	mtGCaucCT0006	195C 263G 315.1C 523- 524- 750G 1438G 3992T 4024G 4769G 5004C 8269A 8860G 9123A 9300A 12642G 14365T 14569A 14582G 15326G 15884A
1	H56	mtGCaucOH0020	152Y 263G 315.1C 750G 1438G 4769G 8850G 8860G 11788T 15326G 16519C
1	H56	mtGCaucWV0001	263G 309.1C 315.1C 750G 1438G 4769G 8860G 11788T 15326G 16519C
1	H5a1	mtGCaucCA0018	186T 263G 315.1C 456T 523- 524- 750G 1438G 4336C 4769G 8860G 12634G 15326G 15833T 16304C
1	H5a1	mtGCaucCT0001	152Y 263G 315.1C 456T 523- 524- 750G 1438G 4336C 4769G 6722A 7642A 8860G 9103C 15326G 15833T 16304C
1	H5a1f	mtGCaucNY0001	263G 309.1C 315.1C 456T 523- 524- 750G 961C 965.1C 1438G 4336C 4736C 4769G 8860G 12535T 15326G 15833T 16223T 16304C
1	H5a1g1	mtGCaucSD0002	263G 315.1C 444G 456T 523- 524- 750G 1438G 4336C 4769G 5082C 8860G 9804A 15326G 15833T 16114T 16172C 16304C 16311C
1	H5a1q	mtGCaucFL0013	263G 309.1C 315.1C 456T 523- 524- 750G 1438G 4336C 4769G 4916G 8860G 12151G 13879C 14771A 15326G 15833T 16304C
1	H5a2	mtGCaucIL0002	263G 309.1C 315.1C 456T 750G 1438G 4336C 4769G 5839T 8860G 15326G 16304C
1	H5a6	mtGCaucAR0003	152C 263G 315.1C 456T 750G 1438G 4336C 4769G 5319G 8563G 8860G 15326G 16304C
1	H5a6	mtGCaucMD0006	152C 263G 315.1C 340T 456T 750G 1438G 4336C 4769G 5319G 7388W 8563G 8860G 13708A 15326G 16261T 16304C 16497G
1	H5b	mtGCaucTN0002	152C 263G 315.1C 456T 750G 1438G 4769G 5471A 8860G 10238C 13285G 15326G 16129A 16304C
1	H5b1	mtGCaucVA0011	146C 195C 263G 309.1C 309.2C 315.1C 456T 750G 1438G 1822C 4002C 4769G 5471A 5492C 8860G 14497G 15326G 16304C
1	H5b1	mtGCaucWI0003	146C 152C 195C 263G 309.1C 315.1C 456T 750G 1438G 4769G 5471A 8860G 13581C 14497G 15326G 16304C
1	H5b2	mtGCaucOH0010	263G 315.1C 327T 456T 750G 1438G 4769G 5471A 8860G 9948A 12864C 15326G 15493T 16304C
1	H5k	mtGCaucME0001	263G 315.1C 456T 750G 1438G 2626C 2887Y 4769G 8020A 8860G 15326G 16304C 16519C
1	H6a1a	mtGCaucGA0001	239C 263G 315.1C 750G 1438G 3548C 3915A 4727G 4769G 5048C 8860G 9380A 10166C 11253C 15326G 16362C 16482G
1	H6a1a2b1	mtGCaucOH0019	239C 263G 309.1C 315.1C 750G 1438G 3705A 3915A 4727G 4769G 5979A 7202G 8860G 9380A 9773T 9818T 11253C 11662C 15326G 16209C 16362C
1	H6a1a3a	mtGCaucKY0001	239C 263G 309.1C 315.1C 750G 827G 1438G 3915A 4727G 4769G 5785C 8860G 9380A 11253C 15099C 15326G 16362C 16482G
1	H6a1b2	mtGCaucNJ0005	239C 263G 309.1C 309.2C 315.1C 750G 1438G 3915A 4727G 4769G 8860G 9254G 9380A 10589A 14305A 15326G 16169T 16362C 16482G 16519C
1	H6a1b3	mtGCaucLA0001	204C 239C 263G 315.1C 750G 1438G 3915A 4727G 4769G 8860G 9380A 10589A 15326G 16193T 16219G 16319A 16362C 16482G
1	H6a1b3	mtGCaucMI0010	26T 146C 204C 239C 263G 309.1C 315.1C 750G 1438G 3915A 4727G 4769G 8860G 9380A 10589A 11969A 15326G 16193T 16219G 16362C 16482G
1	H6a1b4	mtGCaucFL0010	239C 263G 315.1C 750G 1438G 3915A 4727G 4769G 8860G 9380A 10589A 15326G 16249C 16362C 16482G
1	H6a2a	mtGCaucNC0001	239C 263G 309.1C 315.1C 750G 1438G 3915A 4769G 8860G 9380A 11155T 15326G 15940C 16362C 16482G 16519C
1	H6c	mtGCaucMI0012	152C 239C 263G 315.1C 750G 1438G 4769G 6869T 8860G 9804A 15326G 16362C 16482G 16527T
1	H7	mtGCaucRI0001	73G 263G 315.1C 750G 1438G 4769G 4793G 8860G 15326G 16519C
1	H7	mtGCaucWI0006	263G 315.1C 524.1A 524.2C 750G 1438G 4769G 4793G 8587R 8860G 15326G 16355T 16519C
1	H72	mtGCaucNY0021	263G 315.1C 750G 1438G 4769G 6647G 8860G 13785T 15326G 15927A 16093C 16311C 16519C
1	H7c1	mtGCaucNY0014	263G 309.1C 315.1C 750G 1438G 4769G 4793G 5601T 6296A 8860G 15326G 15758G 16093C 16265G 16519C

1	HV	mtGCaucNC0003	263G 309.1C 315.1C 750G 1438G 2706G 4769G 7028T 8860G 9801A 10205T 10920T 15326G 15514C 16311C
1	HV	mtGCaucWI0005	146C 263G 315.1C 750G 1438G 2706G 4769G 7028T 8860G 10389C 15326G 15902G 16234T 16271C 16311C
1	HV0	mtGCaucOK0002	72C 195C 263G 309.1C 315.1C 750G 1438G 2706G 4688C 4769G 7004G 7028T 8860G 15326G 16166G 16298C
1	HV0a1	mtGCaucNE0001	263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 4769G 7028T 8860G 10196T 15326G 15904T 16126C 16298C
1	HV0b	mtGCaucCT0003	72C 195C 198T 263G 309.1C 309.2C 315.1C 750G 1438G 2706G 4769G 7028T 8860G 15326G 16298C 16519C
1	HV16	mtGCaucID0001	150Y 263G 309.1C 315.1C 750G 1438G 2706G 4769G 5581G 7028T 8860G 12492T 15326G 16311C
1	I1a1b	mtGCaucAR0002	73G 199C 203A 204C 250C 263G 315.1C 455.1T 573.1C 573.2C 573.3C 573.4C 573.5C 750G 1438G 1719A 2706G 3447G 3990T 4529T 4769G 6734A 7028T 8251A 8616T 8860G 9947A 10034C 10238C 10398G 10915C 11719A 12501A 12705T 13780G 14182C 14766T 15043A 15326G 15924G 16129A 16172C 16223T 16311C 16391A 16519C
1	I2	mtGCaucCA0002	73G 152C 199C 204C 207A 250C 263G 309.1C 315.1C 573.1C 573.2C 573.3C 573.4C 573.5C 750G 1438G 1719A 2706G 4529T 4769G 7028T 8119C 8251A 8557A 8860G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15758G 15924G 16129A 16223T 16391A
1	I2d	mtGCaucNJ0004	73G 152C 199C 204C 207A 250C 263G 309.1C 315.1C 573.1C 573.2C 573.3C 573.4C 573.5C 750G 1438G 1719A 2706G 4529T 4769G 6480A 7028T 8251A 8860G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15758G 15924G 16129A 16223T 16391A 16519C
1	I3a	mtGCaucOH0004	73G 152C 199C 204C 207A 239C 246C 250C 263G 309.1C 315.1C 524.1A 524.2C 524.3A 524.4C 573.1C 573.2C 573.3C 573.4C 573.5C 573.6C 750G 1438G 1719A 2706G 4529T 4769G 7028T 8251A 8860G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15924G 16086C 16129A 16223T 16319A 16391A 16519C
1	I3c	mtGCaucIA0004	73G 152C 199C 204C 207A 239C 250C 263G 309.1C 315.1C 524.1A 524.2C 524.3A 524.4C 573.1C 573.2C 573.3C 750G 1438G 1719A 2628C 2706G 4529T 4769G 7028T 8251A 8860G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14766T 15043A 15326G 15924G 16129A 16223T 16391A 16519C
1	I5c	mtGCaucOH0017	73G 199C 204C 250C 263G 315.1C 524.1A 524.2C 573.1C 573.2C 573.3C 573.4C 750G 1438G 1719A 2044G 2706G 4529T 4769G 5471A 7028T 8251A 8269A 8860G 9025R 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14233G 14766T 15043A 15119A 15326G 15924G 16129A 16169T 16223T 16391A 16519C
1	J1b	mtGCaucNE0002	73G 195C 263G 295T 309.1C 309.2C 315.1C 462T 489C 750G 1438G 2706G 3010A 3338C 4216C 4769G 6962A 7028T 8269A 8860G 10398G 10685A 10873C 11251G 11719A 12127A 12591T 12612G 13708A 13933G 14766T 15326G 15452A 16069T 16126C 16145A 16222T 16261T 16519C
1	J1b1a1	mtGCaucIL0007	73G 242T 263G 295T 315.1C 462T 489C 750G 1438G 2158C 2706G 3010A 3254T 4216C 4769G 5460A 7028T 8269A 8557A 8860G 10237C 10398G 11251G 11719A 12007A 12612G 13632G 13708A 13879C 14766T 15326G 15452A 16069T 16093C 16126C 16145A 16172C 16222T 16261T
1	J1b1a1	mtGCaucMT0003	73G 242T 263G 295T 315.1C 462T 489C 750G 1438G 2158C 2706G 3010A 4216C 4769G 5319G 5460A 7028T 8269A 8557A 8860G 10398G 11251G 11719A 12007A 12612G 13708A 13879C 14766T 15326G 15452A 16069T 16126C 16145A 16172C 16222T 16261T
1	J1b1a1	mtGCaucNC0004	73G 146C 242T 263G 295T 309.1C 315.1C 462T 489C 750G 1438G 2158C 2706G 3010A 4216C 4769G 5460A 7028T 8269A 8557A 8860G 10398G 11251G 11719A 12007A 12612G 13708A 13753C 13879C 14470C 14766T 15326G 15452A 16069T 16126C 16129R 16145A 16172C 16222T 16261T 16311Y
1	J1c1	mtGCaucIL0012	73G 185A 263G 295T 309.1C 315.1C 462T 482C 489C 709A 750G 1438G 2706G 3010A 3394C 4216C 4769G 7028T 8387A 8860G 9205C 10398G 11251G 11719A 11810T 11908R 12612G 13708A 14443T 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c1	mtGCaucMN0006	73G 185A 228A 263G 295T 315.1C 462T 482C 489C 750G 1438G 2706G 3010A 3394C 4216C 4769G 7028T 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16215G

1	J1c12	mtGCaucTN0006	73G 185A 189G 228A 263G 295T 309.1C 315.1C 462T 489C 750G 789C 1438G 2706G 2905G 3010A 4084A 4216C 4769G 6224T 7028T 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16261T
1	J1c1a	mtGCaucIN0002	73G 228A 263G 295T 315.1C 462T 482C 489C 750G 1438G 2706G 3010A 3394C 4216C 4769G 7028T 8860G 9635C 10398G 11116C 11251G 11623T 11719A 12612G 13708A 13899C 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c1b	mtGCaucCA0016	73G 185A 228A 263G 295T 309.1C 315.1C 462T 482C 489C 750G 1438G 2706G 3010A 3394C 4216C 4769G 7028T 7184G 8860G 10398G 11251G 11719A 12612G 12696C 13708A 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c1b	mtGCaucNY0010	73G 228A 263G 295T 315.1C 462T 482C 489C 750G 1438G 2706G 3010A 3394C 3548C 4216C 4769G 7028T 7184G 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c1b1a	mtGCaucFL0015	73G 185A 210G 228A 263G 295T 309.1C 315.1C 462T 482C 489C 750G 870T 1438G 2706G 3010A 3394C 4216C 4769G 5773A 6040G 7028T 7184G 8860G 10398G 10463C 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c1b1a	mtGCaucMI0009	73G 185A 228A 263G 295T 315.1C 462T 482C 489C 750G 1438G 2706G 3010A 3197C 3394C 4216C 4769G 5773A 6912C 7028T 7184G 8860G 9548A 10398G 10463C 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16272G 16519C
1	J1c2	mtGCaucCA0011	73G 185A 188G 228A 263G 295T 315.1C 462T 489C 523- 524- 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16261T 16519C
1	J1c2a1	mtGCaucKY0004	73G 185A 188G 228A 263G 295T 309.1C 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 6293C 7028T 7245G 8839A 8860G 9181G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 15909G 16069T 16126C 16519C
1	J1c2c	mtGCaucMO0001	73G 146C 185A 188G 228A 263G 295T 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 10398G 10685A 11251G 11719A 12612G 13281C 13708A 13933G 14766T 14798C 15326G 15452A 16069T 16126C 16519C
1	J1c2o	mtGCaucMI0002	73G 185A 188G 228A 263G 295T 309.1C 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16163G 16519C
1	J1c3	mtGCaucMI0003	73G 185A 189R 263G 295T 309.1C 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4316R 4769G 7028T 8860G 10398G 11251G 11719A 12565A 12612G 13708A 13934T 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c3a2	mtGCaucMS0001	73G 228A 263G 295T 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 9548A 9836C 10398G 11251G 11719A 12612G 13708A 13934T 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c3c	mtGCaucOK0006	73G 185A 228A 263G 295T 315.1C 462T 489C 750G 962A 1438G 3010A 4216C 4769G 6956C 7028T 8860G 10398G 11251G 11719A 12382R 12612G 13708A 13934T 14766T 14798C 15326G 15452A 16069T 16126C 16222T 16235G
1	J1c3f	mtGCaucME0002	73G 228A 263G 295T 315.1C 462T 489C 709A 750G 1438G 2706G 3010A 4216C 4769G 7028T 8227C 8860G 10398G 10845T 11251G 11719A 12477C 12612G 13708A 13934T 14766T 14798C 15326G 15452A 16063C 16069T 16093C 16126C 16311C
1	J1c5	mtGCaucTN0004	73G 185A 263G 295T 315.1C 462T 489C 573.1C 573.2C 573.3C 573.4C 750G 1438G 2706G 3010A 4216C 4769G 5198G 7028T 8860G 10398G 11251G 11719A 12612G 13708A 14311C 14766T 14798C 15326G 15452A 16069T 16126C

1	J1c5a1	mtGCaucAL0001	73G 185A 263G 295T 315.1C 462T 489C 750G 1438G 2387C 2706G 3010A 4216C 4769G 5198G 7028T 8860G 10192T 10398G 10598G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 15553A 16069T 16126C
1	J1c5a1	mtGCaucIL0001	73G 185A 228A 263G 295T 309.1C 315.1C 462T 489C 750G 1438G 2387C 2706G 3010A 4216C 4769G 5198G 6497C 7028T 8860G 10192T 10398G 10598G 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C
1	J1c7	mtGCaucNY0008	73G 228A 263G 295T 315.1C 462T 489C 750G 1438G 2706G 3010A 4216C 4769G 6554T 6734A 7028T 8860G 10398G 11251G 11719A 12127A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16092C 16126C 16261T
1	J1c8a	mtGCaucIA0003	73G 185A 228A 263G 295T 315.1C 462T 489C 524.1A 524.2C 750G 1438G 2706G 3010A 4216C 4769G 7028T 8860G 10084C 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16319A
1	J2a1a1	mtGCaucOH0008	73G 150T 152C 189G 195C 214G 215G 263G 295T 309.1C 315.1C 319C 489C 513A 750G 1438G 2706G 4216C 4769G 7028T 7476T 7789A 8860G 10398G 10499G 11251G 11377A 11719A 12612G 13708A 13722G 14133G 14766T 15257A 15326G 15452A 16069T 16145A 16231C 16261T
1	J2a1a1a	mtGCaucAZ0002	73G 150T 152C 195C 215G 263G 295T 310.1T 315.1C 319C 489C 513A 750G 1438G 1617Y 1850C 2706G 4216C 4769G 7028T 7476T 7789A 8860G 10398G 10499G 11251G 11377A 11719A 12612G 13708A 13722G 14133G 14766T 15257A 15326G 15452A 16069T 16126C 16145A 16231C 16261T
1	J2a1a1a2	mtGCaucIA0002	73G 150T 152C 195C 215G 263G 295T 310.1T 315.1C 319C 489C 513A 750G 1438G 1850C 2706G 3447G 4216C 4769G 7028T 7476T 7789A 8860G 10398G 10499G 11251G 11377A 11719A 12612G 13708A 13722G 14133G 14766T 15257A 15326G 15452A 16069T 16093C 16126C 16145A 16231C 16261T
1	J2a1a1a2	mtGCaucVA0003	73G 150T 152C 195C 215G 263G 295T 310.1T 315.1C 319C 489C 513A 750G 1438G 1850C 2706G 3447G 4216C 4769G 7028T 7476T 7789A 8860G 10398G 10499G 11251G 11377A 11719A 12612G 13708A 13722G 14133G 14766T 15257A 15326G 15452A 16069T 16126C 16145A 16231C 16261T
1	J2b1a	mtGCaucIN0005	73G 150T 152C 263G 295T 309.1C 315.1C 489C 709A 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 9909C 10172A 10398G 10972R 11251G 11719A 12612G 13708A 14766T 15257A 15326G 15452A 15812A 16069T 16126C 16193T 16278T 16519C
1	J2b1a	mtGCaucOH0016	73G 150T 152C 263G 295T 315.1C 489C 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 10172A 10321C 10398G 11251G 11719A 12612G 13708A 14766T 15110A 15257A 15326G 15452A 15812A 16069T 16126C 16193T 16278T
1	J2b1a	mtGCaucVA0008	73G 150T 152C 263G 295T 309.1C 315.1C 489C 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 10172A 10398G 11084G 11251G 11719A 12612G 13708A 14766T 14774T 15257A 15326G 15452A 15812A 16069T 16126C 16193T 16278T
1	J2b1a1a	mtGCaucIN0004	73G 150T 152C 263G 295T 315.1C 489C 750G 1438G 2706G 4216C 4769G 5228G 5633T 7028T 7476T 8860G 10172A 10398G 11251G 11719A 12612G 13708A 14569A 14766T 15257A 15326G 15452A 15812A 16069T 16193T 16278T 16362C
1	J2b1a6	mtGCaucNE0003	73G 150T 263G 295T 309.1C 315.1C 489C 750G 1438G 2706G 4216C 4769G 5633T 7028T 7476T 8860G 9016G 9494G 10172A 10398G 11251G 11719A 12612G 13708A 14766T 15257A 15326G 15452A 15662G 15812A 16069T 16126C 16193T 16278T 16519C
1	K1a1	mtGCaucNJ0001	73G 263G 315.1C 497T 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 7825T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11719A 11914A 12308G 12372A 13434G 14167T 14766T 14798C 15326G 16093C 16224C 16311C 16344T 16390A 16519C
1	K1a3a1	mtGCaucSC0003	73G 263G 309.1C 315.1C 497T 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 7559G 8440G 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11719A 12308G 12372A 12397G 13117G 13212T 14167T 14766T 14798C 15326G 16093C 16224C 16311C 16519C

1	K1a4	mtGCaucIL0005	73G 251R 263G 315.1C 497T 524.1A 524.2C 524.3A 524.4C 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11485C 11719A 11914A 12308G 12372A 14167T 14766T 14798C 15326G 16224C 16249C 16311C 16519C
1	K1a4a1	mtGCaucFL0007	73G 263G 309.1C 309.2C 315.1C 497T 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 4769G 6260A 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11485C 11719A 11840T 12308G 12372A 13740C 14167T 14766T 14798C 15326G 16224C 16311C 16519C
1	K1a4a1a	mtGCaucIL0006	73G 195C 263G 315.1C 497T 513R 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 4295G 4769G 6260A 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11485C 11719A 11840T 12308G 12372A 12904R 13740C 14167T 14766T 14798C 15326G 15884A 16224C 16311C 16519C
1	K1a4a1a2	mtGCaucOH0003	73G 263G 309.1C 315.1C 497T 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 3553T 4295G 4769G 5508C 6260A 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11485C 11719A 11840T 12308G 12372A 13740C 14167T 14766T 14798C 15326G 15884A 16224C 16245T 16266T 16311C 16519C
1	K1a4a1a3	mtGCaucNY0016	73G 152C 195C 263G 315.1C 497T 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 4295G 4769G 5177A 6260A 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11485C 11719A 11840T 12308G 12372A 13740C 14167T 14766T 14798C 15326G 15884A 16224C 16311C 16519C
1	K1a4a1f	mtGCaucKS0001	73G 152C 263G 315.1C 325T 497T 750G 1189C 1438G 1811G 2706G 3480G 4769G 6260A 7028T 8860G 9055A 9698C 10029G 10398G 10550G 11299C 11467G 11485C 11719A 11840T 12308G 12372A 13740C 14167T 14766T 14798C 15326G 16224C 16311C 16519C
1	K1a4c1	mtGCaucOK0003	73G 263G 315.1C 497T 524.1A 524.2C 524.3A 524.4C 524.5A 524.6C 524.7A 524.8C 750G 1189C 1250T 1438G 1811G 2706G 3480G 4769G 5264T 7028T 8860G 9055A 9698C 10398G 10410C 10550G 11299C 11467G 11485C 11719A 11782T 12308G 12372A 12612G 13098G 13827G 14167T 14581Y 14766T 14798C 15326G 16169T 16224C 16246T 16311C 16519C
1	K1a4c1	mtGCaucOK0005	73G 263G 315.1C 497T 524.1A 524.2C 524.3A 524.4C 524.5A 524.6C 750G 1189C 1250T 1438G 1811G 2706G 3480G 4769G 5264T 7028T 8860G 9055A 9698C 10398G 10410C 10550G 11299C 11467G 11485C 11719A 11782T 12308G 12372A 12612G 13098G 13827G 14167T 14766T 14798C 15326G 16169T 16224C 16246T 16311C 16519C
1	K1a4d	mtGCaucMN0002	73G 263G 315.1C 497T 524.1A 524.2C 723C 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9698C 10398G 10550G 11071T 11299C 11467G 11485C 11719A 12308G 12372A 14167T 14766T 14798C 15326G 15355A 16093C 16224C 16311C 16519C
1	K1b2a1	mtGCaucMO0003	73G 146C 195C 263G 315.1C 524.1A 524.2C 750G 1189C 1438G 1811G 2706G 3480G 4769G 5913A 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11539T 11719A 12308G 12372A 12738G 12771A 13759A 14167T 14766T 14798C 15326G 16224C 16311C 16519C
1	K1c1	mtGCaucOH0007	73G 146C 152C 263G 309.1C 315.1C 498- 523- 524- 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9093G 9698C 10398G 10550G 11299C 11377A 11467G 11719A 12308G 12372A 14167T 14751T 14766T 14798C 15326G 16224C 16311C 16519C
1	K1c1	mtGCaucTN0007	73G 146C 152C 214G 263G 309.1C 315.1C 498- 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9093G 9698C 10398G 10550G 11299C 11377A 11467G 11719A 12308G 12372A 13375G 14167T 14766T 14798C 15326G 16224C 16311C 16519C 16527T
1	K1c1b	mtGCaucCT0002	73G 146C 152C 263G 315.1C 498- 750G 1189C 1438G 1811G 2706G 3480G 4769G 6389T 7028T 8860G 9055A 9093G 9698C 10398G 10550G 11299C 11377A 11467G 11719A 12308G 12372A 14167T 14766T 14798C 15326G 15900C 16224C 16311C 16519C
1	K1c2	mtGCaucMD0005	73G 146C 152C 263G 309.1C 315.1C 498- 750G 1189C 1438G 1811G 2706G 3480G 4769G 5973A 6468T 6620C 7028T 7046G 8860G 9055A 9698C 10248C 10398G 10550G 11299C 11362G 11467G 11719A 12308G 12372A 12759T 12834G 14002G 14040A 14167T 14766T 14798C 15326G 16224C 16311C 16320T 16519C

1	K1c2	mtGCaucPA0003	73G 146C 152C 263G 309.1C 315.1C 498- 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9006G 9055A 9698C 10398G 10550G 11299C 11467G 11719A 12308G 12372A 14002G 14040A 14167T 14766T 14798C 15326G 16224C 16311C 16320T 16519C
1	K2a3	mtGCaucMA0006	73G 146C 152C 263G 309.1C 315.1C 709A 750G 1438G 1811G 2706G 3480G 4561C 4748Y 4769G 6750T 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 12308G 12372A 13293T 14167T 14766T 14798C 15326G 16224C 16311C 16519C
1	K2a6	mtGCaucCO0003	73G 146C 152C 263G 294.1T 315.1C 709A 750G 1438G 1811G 2308G 2706G 3480G 4561C 4769G 7028T 7286Y 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 12308G 12372A 14167T 14305A 14766T 14798C 15326G 16224C 16311C 16327T 16519C
1	K2b1a1a	mtGCaucNY0011	73G 146C 153G 263G 309.1C 315.1C 750G 1438G 1811G 2217T 2706G 3480G 4769G 5231A 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 11869A 12308G 12372A 13135A 14037G 14167T 14766T 14798C 15326G 15484G 16222T 16224C 16270T 16311C 16519C
1	K2b1a1a	mtGCaucSD0001	73G 146C 263G 315.1C 750G 1438G 1811G 2217T 2706G 3480G 4769G 5231A 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 11869A 12308G 12372A 13135A 14037G 14167T 14766T 14798C 15326G 16222T 16224C 16270T 16311C 16519C
1	L2a1c1	mtGCaucMO0004	73G 143R 146C 152C 195C 198T 263G 315.1C 750G 769A 930A 1018A 1438G 2416C 2706G 2789T 3010A 3308C 3594T 4104G 4769G 6663G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8604C 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14001G 14566G 14766T 15301A 15326G 15784C 16086C 16223T 16278T 16294T 16309G 16390A
1	L2a1l1b	mtGCaucMO0006	73G 143A 146C 152C 195C 263G 315.1C 534T 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12408C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 15880G 16093C 16189C 16192T 16223T 16278T 16294T 16309G 16390A
1	M7c1c3	mtGCaucOH0005	73G 146C 199C 207A 263G 309.1C 315.1C 489C 523- 524- 750G 1438G 2706G 3606G 4071T 4769G 4850T 5442C 6455T 7028T 8701G 8860G 9540C 9824C 10398G 10400T 10873C 11665T 11719A 12091C 12705T 13896C 14766T 14783C 15043A 15236G 15301A 15326G 16223T 16295T 16362C 16519C
1	N1a1a1a2	mtGCaucMN0003	73G 152C 189G 199C 204C 207A 263G 315.1C 573.1C 573.2C 573.3C 573.4C 573.5C 669C 750G 1438G 1719A 2702A 2706G 3336C 4769G 5315G 7028T 8485A 8860G 8901G 10238C 10398G 11719A 12501A 12705T 12810G 13780G 14766T 15043A 15184C 15299C 15326G 16086C 16147A 16223T 16248T 16320T 16355T 16519C
1	T1a1	mtGCaucIN0008	73G 152C 195C 263G 315.1C 709A 750G 1438G 1888A 2706G 4104G 4216C 4769G 4917G 7028T 7372C 8697A 8860G 9899C 10463C 11251G 11719A 12633A 13368A 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16163G 16186T 16189C 16294T 16304C 16519C
1	T1a1	mtGCaucMI0007	73G 152C 183G 195C 215G 263G 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 7028T 8697A 8860G 9899C 10463C 11251G 11719A 12633A 13368A 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16163G 16186T 16189C 16294T 16519C
1	T1a1	mtGCaucPA0018	73G 150T 195C 263G 309.1C 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 7028T 8697A 8860G 9377R 9899C 10463C 11251G 11539T 11719A 12633A 13368A 14180C 14766T 14905A 15326G 15452A 15607G 15766G 15928A 16126C 16163G 16186T 16189C 16294T 16519C
1	T1a1b	mtGCaucWI0008	73G 152C 195C 263G 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 7028T 8697A 8860G 9899C 10143A 10463C 11251G 11719A 12633A 13368A 14281T 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16163G 16186T 16189C 16294T 16519C

1	T1a1c	mtGCaucFL0001	73G 152C 195C 263G 315.1C 709A 750G 1438G 1888A 2706G 4216C 4664T 4769G 4917G 7028T 8697A 8860G 9120G 9899C 10463C 11251G 11719A 12633A 13368A 14766T 14905A 15236G 15326G 15452A 15607G 15928A 15965G 16126C 16163G 16186T 16189C 16294T 16519C
1	T2a	mtGCaucNY0003	73G 195C 263G 309.1C 315.1C 709A 750G 1438G 2706G 4216C 4464A 4769G 4917G 7028T 8697A 8860G 10463C 11251G 11719A 11812G 13368A 13965C 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16519C
1	T2a1a	mtGCaucFL0009	73G 263G 309.1C 315.1C 709A 750G 1438G 1888A 2706G 2850C 4216C 4769G 4917G 6632C 7022C 7028T 8697A 8860G 10463C 11251G 11719A 11812G 13368A 13965C 14233G 14687G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16519C
1	T2a1a	mtGCaucPA0019	73G 263G 315.1C 709A 750G 1438G 1888A 2706G 2850C 4216C 4769G 4917G 7022C 7028T 8697A 8860G 9242R 10463C 11251G 11719A 11812G 13368A 13965C 14233G 14687G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16519C
1	T2a1a2	mtGCaucVA0005	73G 263G 315.1C 709A 750G 1438G 1888A 2706G 2850C 4216C 4688C 4769G 4917G 7022C 7028T 8697A 8860G 10463C 11251G 11719A 11812G 13368A 13965C 14233G 14687G 14766T 14905A 15213Y 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16519C
1	T2a1b	mtGCaucNY0015	73G 263G 315.1C 709A 750G 1438G 1888A 2706G 3394C 3591A 4216C 4769G 4917G 7028T 8697A 8772C 8860G 10463C 11251G 11719A 11812G 13368A 13965C 14233G 14687G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16287T 16294T 16296T 16324C 16362C 16519C
1	T2b	mtGCaucMS0002	73G 263G 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 9469T 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2b	mtGCaucOH0022	73G 263G 309.1C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14587G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2b	mtGCaucUT0001	73G 263G 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 7262G 8697A 8860G 9531G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2b13	mtGCaucPA0013	73G 195C 263G 309.1C 309.2C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 11251G 11719A 11812G 12202Y 13368A 14233G 14766T 14861A 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2b13b	mtGCaucNJ0002	73G 263G 309.1C 309.2C 315.1C 573.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 7269A 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14861A 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16362C 16519C
1	T2b3	mtGCaucPA0006	73G 151T 263G 309.1C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4561C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 10750G 11251G 11299C 11719A 11812G 12280G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2b3e	mtGCaucMD0002	73G 151T 263G 309.1C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 10750G 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16187T 16294T 16296T 16304C 16519C
1	T2b4b	mtGCaucNJ0008	73G 152C 263G 309.1C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 9254G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16104T 16126C 16294T 16304C 16519C

1	T2b5a1	mtGCaucOR0001	3C 73G 263G 315.1C 573.1C 709A 750G 930A 1438G 1888A 2706G 3826C 4216C 4769G 4917G 5147A 5201C 7028T 8504C 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16465T 16519C
1	T2b6a	mtGCaucOH0021	73G 263G 309.1C 315.1C 458T 709A 750G 930A 1438G 1709A 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 9300A 10463C 11251G 11533T 11719A 11812G 12007A 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	T2c1a	mtGCaucMN0005	73G 146C 234G 263G 315.1C 573.1C 573.2C 573.3C 573.4C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 6261A 6975C 7028T 8455T 8697A 8860G 8903Y 10463C 10822T 11251G 11719A 11812G 12123T 13368A 13973T 14233G 14766T 14905A 15115C 15326G 15452A 15607G 15928A 16126C 16189C 16265G 16292T 16294T 16519C
1	T2c1d1	mtGCaucMI0001	73G 146C 152C 263G 279C 315.1C 709A 750G 1438G 1888A 2706G 3206Y 4216C 4769G 4917G 5187T 6261A 7028T 7873T 8697A 8860G 10463C 10822T 11251G 11719A 11812G 11914A 13368A 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16292T 16294T 16519C
1	T2d2	mtGCaucMD0004	73G 195C 263G 309.1C 309.2C 315.1C 709A 750G 1438G 1888A 2706G 4113A 4216C 4769G 4917G 5471A 6445T 7028T 7961C 8697A 8860G 9210G 9615C 10463C 11251G 11719A 11812G 12408C 13260C 13368A 14233G 14323A 14605G 14766T 14905A 15326G 15452A 15607G 15784C 15928A 16126C 16294T 16296T 16368C 16519C
1	T2f	mtGCaucOH0014	73G 263G 309.1C 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 7028T 8270T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8697A 8860G 10463C 11251G 11272G 11719A 11812G 13368A 14233G 14766T 14905A 15314A 15326G 15452A 15607G 15928A 16126C 16189C 16294T 16296T 16519C
1	T2f1a1	mtGCaucIL0011	73G 195C 263G 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 5277C 5426C 6489A 7028T 8270T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15028A 15043A 15326G 15346A 15452A 15607G 15928A 16126C 16182C 16183C 16189C 16193.1C 16294T 16296T 16298C 16519C
1	T2f1a1	mtGCaucOH0012	73G 195C 263G 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 5277C 5426C 6489A 7028T 8270T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8697A 8860G 10463C 11251G 11719A 11812G 13368A 14233G 14766T 14905A 15028A 15043A 15326G 15452A 15607G 15650A 15928A 16126C 16182C 16183C 16189C 16294T 16296T 16298C 16519C
1	U2e1a1	mtGCaucPA0009	73G 152C 217C 263G 309.1C 309.2C 315.1C 340T 508G 524.1A 524.2C 750G 1438G 1811G 2706G 3116T 3720G 4769G 5390G 5426C 6045T 6152C 7028T 8860G 10127G 10876G 11197T 11467G 11719A 11732C 12308G 12372A 13020C 13734C 14766T 15326G 15907G 16051G 16129C 16183C 16189C 16193.1C 16362C 16519C
1	U2e3a	mtGCaucOH0015	73G 152C 217C 263G 309- 315.1C 394T 508G 524.1A 524.2C 524.3A 524.4C 575T 750G 1438G 1811G 2706G 3170A 3720G 4769G 5390G 5426C 6045T 6152C 7028T 8860G 10876G 11467G 11719A 12308G 12372A 13020C 13734C 14766T 15326G 15721C 15907G 16051G 16129C 16181G 16182C 16183C 16189C 16260T 16356C 16362C 16519C
1	U3b2	mtGCaucNY0004	73G 150T 152C 263G 309.1C 309.2C 315.1C 523- 524- 750G 1438G 1811G 2706G 4188G 4640A 4769G 5004C 7028T 7660C 8860G 9656C 11467G 11719A 12308G 12372A 13474C 13743C 14139G 14766T 15289Y 15326G 15454C 15930A 15944- 16343G
1	U4a1a1	mtGCaucFL0006	73G 152C 195C 263G 309.1C 315.1C 499A 524.1A 524.2C 750G 961C 965.1C 965.2C 1438G 1811G 2706G 4646C 4769G 5899.1C 5999C 6047G 7028T 8167C 8818T 8860G 11332T 11467G 11719A 12308G 12372A 12618A 12937G 14620T 14766T 15326G 15693C 16134T 16356C 16519C

1	U4a1a1	mtGCaucIL0013	73G 152C 195C 263G 309.1C 315.1C 499A 750G 961C 965.1C 965.2C 1438G 1811G 2706G 4646C 4769G 5999C 6047G 7028T 8818T 8860G 11332T 11467G 11719A 12308G 12372A 12937G 14384R 14620T 14766T 15326G 15693C 16134T 16356C 16519C
1	U4b1b1	mtGCaucPA0004	73G 152C 195C 263G 315.1C 499A 750G 1438G 1811G 2706G 4646C 4769G 5999C 6047G 7028T 7705C 8308G 8860G 9389G 10819G 11332T 11339C 11467G 11719A 12308G 12348T 12372A 13528G 13565T 14620T 14766T 15326G 15373G 15693C 15758G 16356C 16390A 16519C
1	U4b2a	mtGCaucRI0003	73G 195C 263G 309.1C 315.1C 499A 524.1A 524.2C 750G 1438G 1811G 2706G 4646C 4769G 5999C 6047G 7028T 7673R 7705C 8860G 11332T 11467G 11719A 12308G 12372A 14620T 14766T 15260R 15326G 15693C 15883A 16136C 16356C 16519C
1	U4c1a	mtGCaucCA0014	73G 150T 195C 263G 315.1C 499A 524.1A 524.2C 524.3A 524.4C 750G 1438G 1811G 2706G 4646C 4769G 4811G 5604Y 5999C 6047G 6146G 7028T 8860G 9070G 10907C 11009C 11332T 11467G 11719A 12308G 12372A 14620T 14766T 14866T 15326G 15693C 16179T 16356C 16519C
1	U5a1a1	mtGCaucFL0002	73G 263G 315.1C 750G 1438G 1700C 2706G 3197C 4769G 5495C 7028T 8860G 9477A 11467G 11719A 12308G 12372A 13617C 14364A 14766T 14793G 15218G 15326G 15924G 16256T 16270T 16399G 16519C
1	U5a1a1	mtGCaucNY0022	73G 263G 315.1C 524.1A 524.2C 524.3A 524.4C 750G 1438G 1700C 2706G 3197C 4769G 5495C 7028T 8860G 9477A 11467G 11719A 11914A 12308G 12372A 13617C 14766T 14793G 15218G 15326G 15924G 16184T 16256T 16270T 16399G
1	U5a1b	mtGCaucCA0006	73G 263G 315.1C 750G 1438G 2706G 3197C 4615G 4769G 7028T 8860G 9477A 9667G 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15218G 15326G 16192T 16256T 16270T 16399G
1	U5a1b1	mtGCaucVT0001	73G 263G 315.1C 750G 1438G 2706G 3197C 4769G 7028T 8860G 8865A 9477A 9667G 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15218G 15326G 15404C 16192T 16256T 16270T 16291T 16399G
1	U5a1d1	mtGCaucIN0006	73G 189R 234G 263G 309.1C 315.1C 750G 1438G 2706G 3027C 3197C 4769G 5263T 7028T 8860G 9477A 11467G 11719A 12308G 12372A 13002A 13617C 14766T 14793G 14870G 15218G 15326G 16192T 16256T 16270T 16399G
1	U5a1f1	mtGCaucCA0008	73G 195C 263G 315.1C 750G 1438G 2706G 3197C 4313C 4769G 5585A 6023A 7028T 7403G 7569G 7933G 8860G 9477A 11467G 11719A 12308G 12372A 13617C 13781C 14766T 14793G 15218G 15326G 16192T 16256T 16270T 16278Y 16298C 16311C 16399G
1	U5a1f1	mtGCaucNY0013	73G 195C 263G 315.1C 750G 1438G 2706G 3197C 4313C 4769G 5585A 6023A 7028T 7403G 7569G 8659G 8860G 9024G 9477A 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15218G 15326G 16192T 16256T 16270T 16311C 16399G
1	U5a1i	mtGCaucPA0002	73G 263G 315.1C 750G 1438G 2706G 3197C 4769G 4796T 7028T 8772A 8860G 9477A 11467G 11719A 12103A 12308G 12372A 13617C 14003T 14766T 14793G 14893G 14971C 15218G 15326G 16093C 16192T 16256T 16270T 16399G
1	U5a2a1	mtGCaucMO0002	73G 263G 309.1C 315.1C 750G 1438G 2706G 3197C 4769G 7028T 8860G 9187C 9477A 11467G 11719A 12308G 12372A 13617C 13827G 13928C 14766T 14793G 15326G 16114A 16192Y 16256T 16270T 16294T 16526A
1	U5a2a1b	mtGCaucIA0001	73G 309.1C 315.1C 750G 1438G 2706G 3197C 4769G 7028T 8860G 9115G 9477A 11467G 11719A 12308G 12372A 13015C 13617C 13827G 13928C 14766T 14793G 15326G 16114A 16192T 16256T 16270T 16294T 16526A
1	U5a2b1c	mtGCaucVA0006	73G 263G 315.1C 750G 960.1C 1438G 2706G 3197C 4769G 7028T 8860G 9477A 9548A 11467G 11719A 12308G 12372A 13617C 13999T 14766T 14793G 15326G 15380G 15774Y 15903G 16192T 16256T 16270T 16526A

1	U5a2c3a	mtGCaucCA0015	73G 140T 263G 309.1C 315.1C 573.1C 573.2C 750G 1438G 2706G 3197C 4769G 7028T 7960C 8860G 9477A 10619T 10709C 11465C 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15326G 16256T 16270T 16311C 16526A
1	U5a2c4	mtGCaucOH0006	73G 263G 315.1C 493G 750G 1438G 2706G 3197C 3531A 4769G 7028T 8860G 9477A 10619T 10644R 11467G 11719A 12308G 12372A 13617C 14766T 14793G 15326G 16192T 16256T 16270T 16526A
1	U5a2d1a	mtGCaucOH0002	73G 263G 309.1C 315.1C 523- 524- 750G 1438G 2706G 3197C 3750T 4769G 7028T 7843G 7978T 8104C 8860G 9148C 9477A 11107T 11467G 11719A 12308G 12372A 13617C 14577G 14766T 14793G 15326G 16256T 16270T 16526A
1	U5a2d1a	mtGCaucOH0013	73G 263G 315.1C 523- 524- 750G 1438G 2706G 3197C 3750T 4769G 7028T 7807T 7843G 7978T 8104C 8860G 9148C 9477A 11107T 11467G 11719A 12308G 12372A 13617C 14577G 14766T 14793G 15326G 16256T 16270T 16290T 16526A
1	U5b1	mtGCaucMN0007	73G 150T 263G 315.1C 750G 1438G 2581G 2706G 3197C 4769G 5656G 7028T 7768G 8860G 9438A 9477A 11467G 11719A 12092T 12308G 12372A 12930C 13617C 14182C 14766T 15326G 16182C 16183C 16189C 16270T 16519C
1	U5b1b1	mtGCaucMO0005	73G 146C 150T 263G 315.1C 750G 1438G 2706G 3197C 4769G 5656G 7028T 7385G 7768G 8860G 9477A 10927C 11467G 11719A 12308G 12372A 12618A 13617C 14182C 14766T 15326G 16093C 16183C 16189C 16193.1C 16270T
1	U5b1c	mtGCaucNH0002	73G 150T 263G 309.1C 315.1C 750G 1438G 2486C 2706G 3197C 4769G 5656G 5824A 7028T 7768G 8860G 9477A 11467G 11719A 12308G 12372A 12726T 13617C 14182C 14766T 15191C 15326G 16189C 16270T 16311C
1	U5b1c1	mtGCaucCA0007	73G 150T 263G 309.1C 315.1C 519G 750G 1438G 2706G 3197C 4769G 5656G 7028T 7768G 8860G 9477A 11467G 11719A 12308G 12372A 13617C 14178C 14182C 14766T 15191C 15326G 16189C 16270T 16311C 16336A 16526A
1	U5b2a1a1	mtGCaucMI0006	73G 150T 263G 309.1C 315.1C 750G 896G 1438G 1721T 2706G 3197C 4732G 4769G 7028T 7674C 7768G 8860G 9477A 11467G 11719A 12308G 12372A 13617C 13637G 14182C 14766T 15326G 15511C 16192T 16311C
1	U5b2a1a1	mtGCaucMN0004	73G 150T 263G 315.1C 750G 896G 1438G 1721T 2706G 3197C 4732G 4769G 7028T 7768G 8860G 9477A 11467G 11719A 12308G 12372A 12715G 13617C 13637G 14182C 14384A 14766T 15326G 15511C 16239T 16311C
1	U5b2a2b1	mtGCaucOK0004	73G 150T 263G 309.1C 315.1C 750G 1438G 1721T 2706G 2757G 3197C 3212T 4732G 4769G 4843T 7028T 7768G 8074G 8860G 9477A 11467G 11719A 12308G 12372A 13617C 13637G 14182C 14766T 14956C 15326G 16189C 16191.1C 16192T 16270T 16398A
1	U5b2b3	mtGCaucNY0002	73G 150T 263G 315.1C 517T 750G 1438G 1721T 2706G 2755G 3197C 4769G 5899.1C 7028T 7543R 7768G 8542C 8860G 9477A 9727T 10018R 11467G 11653G 11719A 12308G 12372A 12634G 13617C 13630G 13637G 14182C 14766T 15326G 15905C 16192T 16224C 16270T 16362C 16519C
1	U5b2b3a1a	mtGCaucCA0005	73G 150T 263G 279C 315.1C 517T 750G 1438G 1721T 2706G 2755G 3197C 3338C 4769G 5261A 7028T 7768G 8860G 9477A 9494G 11467G 11653G 11719A 12308G 12372A 12634G 13617C 13630G 13637G 14182C 14766T 15326G 15905C 16114T 16224C 16270T
1	U5b2c2	mtGCaucNY0019	73G 150T 263G 309.1C 309.2C 315.1C 723G 750G 960.1C 1438G 1721T 2706G 3197C 3861G 4769G 5836G 7028T 7768G 8860G 9477A 10262G 11467G 11719A 12308G 12372A 13017G 13617C 13637G 14182C 14766T 15326G 16192T 16270T
1	U5b2c2	mtGCaucPA0014	73G 150T 263G 309.1C 315.1C 723G 750G 960.1C 1438G 1721T 2706G 3197C 3861G 4769G 5836G 7028T 7768G 8860G 9477A 10262G 11467G 11719A 12308G 12372A 13017G 13617C 13637G 14182C 14766T 15326G 16192T 16270T

1	U5b3a2	mtGCaucAZ0003	73G 150T 263G 315.1C 750G 1438G 2706G 3197C 4769G 7028T 7226A 7768G 8860G 9477A 9967Y 10978G 11467G 11719A 12308G 12372A 13617C 14182C 14766T 14803T 15326G 16192T 16235G 16270T 16304C 16465T
1	U5b3g	mtGCaucVA0007	73G 150T 228A 263G 315.1C 750G 1438G 2706G 3197C 4639C 4769G 5147A 7028T 7226A 7768G 8860G 9477A 10335C 11467G 11719A 12308G 12372A 13617C 14020C 14182C 14766T 15326G 15672C 16270T 16304C 16311C
1	U6a7a1a	mtGCaucMA0004	73G 152C 263G 315.1C 750G 794A 1193C 1438G 1692T 2672G 2706G 3348G 4769G 5120G 5471A 7028T 7805A 8473C 8860G 11467G 11719A 11929C 12308G 12372A 14179G 14766T 15043A 15326G 15530C 15632T 16172C 16219G 16278T
1	U6a7a2a	mtGCaucTN0003	73G 152C 263G 309.1C 315.1C 750G 794A 1193C 1438G 1692T 2706G 3348G 4769G 5471A 7028T 7805A 8473C 8860G 11467G 11719A 11941G 12308G 12372A 14034C 14179G 14766T 15043A 15326G 15530C 15632T 16172C 16219G
1	U7a	mtGCaucNJ0007	73G 151T 152C 263G 315.1C 523- 524- 571A 750G 980C 1438G 1811G 2706G 3741T 4769G 5360T 5390G 7028T 8137T 8286C 8287.1C 8287.2C 8287.3C 8287.4C 8287.5C 8684T 8860G 10142T 11467G 11719A 12236A 12308G 12372A 13500C 14569A 14766T 15326G 16309G 16318C 16519C
1	V	mtGCaucKY0002	72C 263G 315.1C 750G 1438G 2706G 4580A 4769G 7028T 8860G 15326G 15904T
1	V	mtGCaucNV0001	72C 228K 263G 309.1C 309.2C 315.1C 750G 1438G 2706G 4580A 4769G 7028T 8860G 15326G 15904T 16180G 16298C
1	V	mtGCaucOK0001	72C 204C 263G 309.1C 309.2C 315.1C 750G 1438G 2706G 4580A 4769G 7028T 8860G 11197T 15326G 15904T 15930A 16298C
1	V11	mtGCaucPA0008	72C 93G 263G 309.1C 315.1C 750G 1438G 2706G 4580A 4769G 7028T 8251A 8860G 9974T 11620G 11761T 15326G 15904T 16298C 16519C
1	V1a1	mtGCaucVA0002	72C 263G 309.1C 315.1C 750G 1438G 2706G 4580A 4639C 4769G 5263T 7028T 8022C 8860G 8869G 15326G 15904T 16233R 16298C
1	V1a1	mtGCaucWI0002	263G 309.1C 309.2C 315.1C 750G 1438G 2706G 4580A 4639C 4769G 5263T 7028T 8860G 8869G 9948A 15326G 15904T 16298C
1	V7	mtGCaucMI0008	93G 204C 263G 309.1C 315.1C 750G 1438G 2706G 2757G 4580A 4769G 7028T 7444A 8860G 9367C 15326G 15904T 16266T 16298C
1	V7a	mtGCaucOR0002	72C 93G 263G 309.1C 309.2C 315.1C 750G 1438G 2706G 4580A 4769G 7028T 7444A 8860G 11899C 15326G 15894A 15904T 16093Y 16153A 16298C
1	W3a1a3	mtGCaucWV0003	73G 189G 194T 195C 204C 207A 263G 309.1C 315.1C 709A 750G 1243C 1406C 1438G 2706G 3505G 4769G 5046A 5460A 6267A 7028T 7151T 8251A 8860G 8994A 11674T 11719A 11947G 12414C 12705T 13263G 14766T 15109C 15326G 15784C 15884C 16223T 16292T 16519C
1	W3a1c	mtGCaucFL0016	73G 189G 194T 195C 199C 204C 207A 263G 309.1C 315.1C 709A 750G 1243C 1406C 1438G 2706G 3505G 4769G 5046A 5460A 7028T 7269A 7853A 8251A 8860G 8994A 9716C 11674T 11719A 11947G 12414C 12705T 13263G 14766T 15326G 15784C 15884C 16223T 16292T 16519C
1	W3a1c	mtGCaucOH0011	73G 152Y 189G 194T 195C 199C 204C 207A 263G 309.1C 315.1C 709A 750G 1243C 1406C 1438G 2706G 3505G 4769G 5046A 5460A 7028T 7269A 7853A 8251A 8860G 8994A 9716C 11674T 11719A 11947G 12414C 12705T 13263G 14766T 15326G 15784C 15884C 16223T 16292T 16519C
1	W4a1	mtGCaucAZ0001	73G 143A 189G 192C 194T 195C 196C 204C 207A 263G 315.1C 709A 750G 1243C 1438G 2706G 3505G 3531A 4769G 5046A 5460A 7028T 8251A 8860G 8994A 11674T 11719A 11947G 12414C 12705T 14766T 15326G 15884C 15924G 16223T 16519C

1	W5a1a	mtGCaucFL0017	73G 189G 194T 195C 200R 204C 207A 263G 309.1C 315.1C 709A 750G 1243C 1438G 2706G 3505G 4363C 4769G 5046A 5460A 6528T 7028T 8251A 8860G 8994A 10097G 10410C 11674T 11719A 11947G 12414C 12705T 14766T 15326G 15775G 15884C 16223T 16258G 16292T 16362C 16519C
1	W5a1a	mtGCaucNY0020	73G 189G 194T 195C 204C 207A 263G 315.1C 709A 750G 1243C 1438G 2706G 3505G 4363C 4769G 5046A 5460A 6528T 7028T 8251A 8860G 8994A 10097G 10410C 11674T 11719A 11947G 12414C 12705T 14766T 15326G 15775G 15884C 16166G 16223T 16292T 16362C 16519C
1	W6a	mtGCaucVA0001	73G 189G 194T 195C 204C 207A 263G 309.1C 315.1C 524.1A 524.2C 709A 750G 1243C 1438G 2706G 3505G 3939T 4093G 4769G 5046A 5460A 7028T 8251A 8610C 8614C 8860G 8994A 11674T 11719A 11947G 12414C 12705T 13722G 14766T 15326G 15781T 15884C 16192T 16223T 16292T 16325C 16519C
1	X2b	mtGCaucMI0004	73G 153G 189G 195C 225A 226C 263G 309.1C 309.2C 315.1C 524.1A 524.2C 750G 1438G 1719A 2706G 4769G 6221C 6371T 7028T 8393T 8860G 8910T 11719A 12705T 13708A 13966G 14470C 14766T 15326G 15927A 16183C 16189C 16193.1C 16223T 16278T 16311C 16519C
1	X2c1	mtGCaucFL0005	73G 153G 195C 225A 227G 263G 315.1C 750G 1438G 1719A 2706G 4769G 6221C 6371T 7028T 8705C 8860G 11719A 12705T 13966G 14470C 14766T 15326G 16183C 16189C 16193.1C 16223T 16255A 16278T 16519C
1	X2i	mtGCaucMD0001	73G 153G 195C 225A 263G 309.1C 315.1C 750G 1438G 1719A 2706G 4769G 6221C 6371T 7028T 8521R 8860G 8966C 11719A 12705T 13966G 14470C 14766T 15326G 16183C 16189C 16193.1C 16223T 16248T 16278T 16519C

Appendix J. U.S. Hispanic haplotypes

# of Haplotypes	Haplogroup	Sample Name(s)	Haplotype (as differences from the rCRS)
3	K2a8	mtGHispKS0002 mtGHispTX0029 mtGHispTX0036	73G 146C 152C 207A 263G 315.1C 709A 750G 1438G 1811G 2706G 3480G 4561C 4769G 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 12308G 12372A 14167T 14766T 14798C 15326G 16182C 16183C 16189C 16224C 16311C 16519C
2	A2h1	mtGHispCA0038 mtGHispIN0003	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1598A 1736G 1888A 2706G 4248C 4769G 4824G 7028T 7702A 8027A 8794T 8860G 11719A 12007A 12705T 12811C 14766T 15326G 16111T 16223T 16290T 16319A 16335G 16526A
2	C1b2	mtGHispIL0002 mtGHispVA0003	73G 249- 290- 291- 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4242T 4715G 4769G 7013A 7028T 7196A 8584A 8701G 8860G 9540C 9545G 9557T 10398G 10400T 10873C 11719A 11914A 12454A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16223T 16298C 16325C 16327T 16519C
2	L3e1e2	mtGHispPR0003 mtGHispTX0005	90A 97A 106- 107- 108- 109- 110- 111- 150T 189G 200G 263G 315.1C 750G 1438G 2352C 2706G 4562G 4769G 6221C 6587T 7028T 8701G 8860G 9098C 9540C 10370C 10398G 10819G 10873C 11719A 12705T 14152G 14212C 14766T 15301A 15326G 15670C 15942C 16223T 16327T
1	A2	mtGHispCA0006	64T 73G 146C 153G 199C 234G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 3826C 4248C 4769G 4824G 5147A 7028T 8027A 8794T 8860G 9311C 11453A 11719A 12007A 12705T 14551G 14766T 15326G 15355A 16111T 16216G 16223T 16290T 16319A 16362C 16391A 16519C
1	A2	mtGHispCA0009	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11221G 11719A 12007A 12705T 14766T 15326G 16111T 16223T 16290T 16362C 16519C
1	A2	mtGHispCA0010	64T 73G 146C 153G 235G 263G 310C 315- 523- 524- 663G 750G 1438G 1736G 2706G 3394C 4248C 4769G 4824G 5634G 6260A 7028T 7853A 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 16111T 16193T 16223T 16290T 16319A 16362C 16519C
1	A2	mtGHispCA0017	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 5899.1C 7028T 8027A 8794T 8860G 8932T 9893T 11719A 12007A 12609C 12705T 14154G 14766T 15326G 16111T 16223T 16290T 16319A 16362C 16518A 16519C
1	A2	mtGHispCA0022	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 596C 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6638C 7028T 7226A 8027A 8623G 8794T 8860G 11719A 12007A 12705T 12825C 14662G 14766T 15326G 16111T 16131C 16134T 16223T 16290T 16319A 16357C 16362C
1	A2	mtGHispCA0026	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 5114G 5177R 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 16111T 16145A 16223T 16290T 16319A 16362C
1	A2	mtGHispCA0030	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 9039A 11719A 12007A 12705T 14766T 15326G 16095T 16111T 16223T 16290T 16319A 16362C
1	A2	mtGHispDE0001	73G 146C 153G 179C 235G 263G 315.1C 385G 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8794T 8860G 9947A 11719A 12007A 12705T 14766T 15326G 16111T 16218T 16223T 16290T 16319A 16362C 16519C
1	A2	mtGHispFL0009	64T 73G 146C 152C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6032A 7028T 8027A 8794T 8860G 11719A 11914A 12007A 12705T 13740C 14766T 15326G 16111T 16223T 16290T 16311C 16319A 16362C 16519C

1	A2	mtGHispHI0001	64T 73G 146C 152C 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 7389C 7852A 8027A 8461A 8794T 8860G 8975C 9755A 11719A 12007A 12408C 12705T 14766T 15326G 16111T 16124C 16223T 16290T 16319A 16362C
1	A2	mtGHispMI0001	64T 73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 16223T 16256T 16290T 16319A 16362C
1	A2	mtGHispPA0001	73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11653G 11719A 12007A 12705T 13731G 14766T 15326G 16111T 16223T 16290T 16319A 16362C
1	A2	mtGHispPR0006	73G 146C 153G 179C 207A 235G 263G 315.1C 385G 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 9947A 11719A 12007A 12705T 14766T 15326G 16111T 16223T 16290T 16319A 16362C 16519C
1	A2	mtGHispTX0018	64T 73G 146C 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8598C 8794T 8860G 11719A 12007A 12705T 14766T 15258G 15326G 16069Y 16111T 16129A 16223T 16265T 16287T 16294T 16319A 16320T 16362C
1	A2	mtGHispTX0022	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11221G 11719A 12007A 12705T 14766T 15326G 16111T 16223T 16290T 16362C 16519C
1	A2	mtGHispTX0035	64T 73G 146C 153G 235G 247A 263G 309.1C 315.1C 374G 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12373G 12705T 14766T 15326G 16111T 16223T 16278T 16290T 16319A 16362C 16519C
1	A2	mtGHispTX0043	64T 73G 146C 235G 263G 315.1C 523- 524- 663G 720C 750G 1438G 1736G 2706G 4248C 4370C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 15793T 16093C 16111T 16223T 16290T 16319A 16362C 16519C
1	A2ac	mtGHispFL0002	64T 73G 146C 153G 235G 249G 263G 309.1C 309.2C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 9064A 9377G 11719A 12007A 12705T 14766T 15074C 15326G 16111T 16213A 16223T 16290T 16319A 16362C
1	A2ad	mtGHispFL0005	64T 73G 146C 153G 189.1A 194T 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8093C 8471T 8794T 8860G 9837R 11719A 12007A 12705T 14766T 15326G 16111T 16187T 16223T 16290T 16300G 16319A 16362C 16519C
1	A2am	mtGHispTX0023	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6253C 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15229C 15314A 15326G 16111T 16129A 16223T 16290T 16319A 16362C
1	A2am	mtGHispTX0031	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6253C 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15106A 15229C 15314A 15326G 16111T 16129A 16223T 16290T 16319A 16362C
1	A2f	mtGHispTX0003	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 523- 524- 663G 676A 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 12940A 14766T 15326G 16111T 16223T 16290T 16319A 16362C
1	A2f3	mtGHispPR0002	64T 73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 5585A 7028T 8027A 8794T 8860G 9156G 11719A 11914A 12007A 12705T 12940A 14275T 14766T 15028T 15323A 15326G 16111T 16223T 16234T 16290T 16319A 16362C 16519C
1	A2h1	mtGHispTX0025	64T 73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1598A 1736G 1888A 2706G 4248C 4769G 4824G 7028T 7702A 8027A 8047C 8794T 8860G 11719A 12007A 12705T 12811C 14766T 15326G 16111T 16223T 16290T 16319A 16335G 16526A

1	A2h1	mtGHispTX0028	64T 73G 146C 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1598A 1736G 1888A 2706G 4248C 4769G 4824G 7028T 7702A 8027A 8794T 8860G 11719A 12007A 12705T 12811C 14766T 15326G 16111T 16223T 16290T 16319A 16335G 16526A
1	A2i	mtGHispWA0003	64T 73G 94A 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 960.1C 960.2C 1438G 1736G 2706G 3307.1A 3308C 4248C 4769G 4824G 5165T 6527G 6620C 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14280G 14470C 14766T 15326G 15386T 16111T 16223T 16290T 16319A 16325C 16362C 16519C
1	A2j	mtGHispTX0030	64T 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C 385G 523- 524- 663G 750G 1193C 1438G 1736G 2706G 4248C 4769G 4824G 6307G 7028T 8027A 8794T 8860G 10595C 11548G 11719A 12007A 12561A 12705T 14766T 15326G 16111T 16223T 16290T 16319A 16362C
1	A2l	mtGHispWA0002	64T 73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 15629C 16223T 16290T 16319A 16362C
1	A2m	mtGHispCO0001	73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 4856C 6689T 7028T 7245G 8027A 8794T 8860G 8947T 8995A 9039A 10274C 11719A 11914A 12007A 12705T 13135A 14530C 14766T 15172A 15326G 15784C 16104T 16129A 16223T 16240G 16290T 16319A 16362C 16449T 16451T 16455A
1	A2o	mtGHispCA0016	73G 146C 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 3972G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 16223T 16290T 16319A 16362C
1	A2o	mtGHispTX0047	64T 73G 146C 153G 235G 263G 315.1C 523- 524- 663G 723C 750G 1007A 1438G 1736G 2706G 3972G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 11719A 12007A 12705T 13032G 14766T 15326G 16223T 16319A 16362C
1	A2t	mtGHispCA0027	73G 146C 153G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 1842G 2071C 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 9053A 11404G 11719A 12007A 12705T 14766T 15236G 15326G 16111T 16223T 16290T 16319A 16356C 16362C
1	A2t	mtGHispOR0001	73G 146C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 1842G 2071C 2706G 4248C 4769G 4824G 7028T 8027A 8794T 8860G 9053A 11404G 11719A 12007A 12705T 14766T 15236G 15326G 16111T 16223T 16290T 16319A 16356C 16362C
1	A2u1	mtGHispCA0024	64T 73G 146C 153R 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 1806Y 2706G 4248C 4769G 4824G 7028T 7702A 8027A 8794T 8860G 11719A 12007A 12705T 12906T 14766T 15326G 16111T 16136C 16147T 16223T 16257T 16290T 16319A 16344T 16362C
1	A2v1	mtGHispAZ0006	64T 73G 146C 152C 153G 235G 263G 309.1C 309.2C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6491A 7028T 7051C 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 15679G 16111T 16223T 16239T 16290T 16319A 16362C 16519C
1	A2v1	mtGHispTX0020	64T 73G 146C 152C 153G 235G 263G 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 6491A 7028T 7051C 8027A 8794T 8860G 11719A 12007A 12705T 14766T 15326G 15679G 16111T 16223T 16239T 16290T 16319A 16362C 16519C
1	A2w1	mtGHispMD0002	64T 73G 146C 153G 214G 235G 263G 309.1C 309.2C 315.1C 523- 524- 573.1C 573.2C 573.3C 573.4C 663G 750G 1438G 1736G 2706G 4248C 4769G 4824G 7028T 7124G 8027A 8572A 8794T 8860G 11016A 11719A 12007A 12366G 12705T 13681G 14693G 14766T 15326G 16111T 16187T 16223T 16290T 16311C 16319A 16362C
1	A2w1	mtGHispNM0005	64T 73G 146C 153G 263G 309.1C 315.1C 523- 524- 573.1C 573.2C 573.3C 573.4C 663G 750G 1005C 1438G 1736G 2706G 4248C 4769G 4824G 6221C 7028T 7124G 8027A 8794T 8860G 8896A 10907C 11016A 11719A 12007A 12705T 14766T 15326G 16111T 16187T 16223T 16290T 16319A 16362C

1	A2z	mtGHispIL0004	73G 146C 152C 153G 214G 235G 263G 309.1C 315.1C 523- 524- 663G 750G 1438G 1736G 2706G 2836T 3744G 4248C 4769G 4824G 6632C 7028T 8027A 8794T 8860G 11719A 12007A 12194T 12705T 14766T 15326G 16083T 16111T 16223T 16256T 16290T 16319A 16362C
1	B2	mtGHispCA0004	73G 263G 309.1C 315.1C 499A 750G 827G 1438G 2706G 3547G 4231G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11719A 13105G 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16278T 16519C
1	B2	mtGHispCA0020	73G 152C 263G 309.1C 315.1C 499A 750G 827G 1438G 2706G 3547G 4512A 4769G 4820A 4977C 5964C 6245G 6473T 7028T 8269A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9210G 9950C 11177T 11719A 11854C 11914A 13590A 14323A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16293G 16319A 16519C
1	B2	mtGHispCA0039	73G 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 5471A 5483C 5777A 6473T 7028T 7260T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10978G 11177T 11719A 13590A 14766T 15326G 15535T 16182C 16183C 16189C 16217C 16259T 16519C
1	B2	mtGHispCA0041	73G 263G 307- 308- 309- 315.1C 499A 750G 827G 1438G 2352C 2706G 3547G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9186T 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16278T 16456A 16519C
1	B2a	mtGHispTX0027	73G 263G 309.1C 309.2C 315.1C 499A 524.1A 524.2C 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 5054A 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10795G 11177T 11719A 13590A 14766T 15326G 15535T 16111T 16183C 16189C 16193.1C 16217C 16483A 16519C
1	B2b	mtGHispTX0034	73G 263G 309.1C 315.1C 499A 524.1A 524.2C 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 6473T 6755A 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11719A 13477R 13590A 14766T 15326G 15535T 16092C 16182C 16183C 16189C 16193.1C 16217C 16223T 16519C
1	B2b2	mtGHispTX0004	55C 63C 64T 73G 195C 263G 309.1C 315.1C 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 6473T 6755A 7028T 8152A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9083C 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16129A 16145A 16183C 16187T 16189C 16217C 16319A 16323C 16519C
1	B2b3a	mtGHispPR0008	73G 152C 263G 271T 309.1C 309.2C 315.1C 454C 455C 460C 463- 499A 750G 827G 1438G 2706G 3547G 3918A 4012G 4232C 4769G 4820A 4977C 6473T 6755A 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9531G 9950C 11177T 11719A 13590A 13708A 14766T 15326G 15535T 15784C 16092C 16182C 16183C 16189C 16193.1C 16193.2C 16217C 16249C 16312G 16344T 16519C
1	B2c2a	mtGHispTX0037	73G 146C 263G 309.1C 309.2C 315.1C 499A 523- 524- 750G 827G 1438G 2706G 3547G 4755C 4769G 4820A 4977C 6473T 7028T 7241G 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8702T 8860G 9950C 10133G 11177T 11719A 13590A 14757C 14766T 15326G 15535T 16182C 16183C 16189C 16217C 16319A 16519C
1	B2c2b	mtGHispAZ0004	73G 146C 152C 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 4755C 4769G 4820A 4977C 6473T 7028T 7241G 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9682C 9950C 11177T 11719A 13590A 13661G 14757C 14766T 15326G 15535T 16182C 16183C 16189C 16193.1C 16217C 16295T 16324C 16519C
1	B2c2b	mtGHispFL0004	73G 146C 152C 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 4755C 4769G 4820A 4977C 6473T 7028T 7241G 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9682C 9950C 11177T 11719A 13590A 13661G 14757C 14766T 15326G 15535T 16182C 16183C 16189C 16217C 16295T 16519C

1	B2d	mtGHispNJ0003	73G 263G 309.1C 309.2C 315.1C 498- 499A 750G 827G 1438G 2706G 3547G 4122G 4123G 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 8875C 9682C 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16218T 16335G 16519C
1	B2f	mtGHispIN0001	73G 263G 315.1C 499A 750G 827G 1438G 2706G 3547G 3796G 3996T 4769G 4820A 4977C 6473T 7028T 7202R 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9212T 9950C 10535C 11177T 11719A 13590A 13833G 14766T 15326G 15535T 16183C 16189C 16217C 16519C
1	B2f	mtGHispTX0050	73G 263G 309.1C 315.1C 499A 524.1A 524.2C 750G 827G 1438G 2706G 3547G 3796G 3996T 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10535C 11177T 11719A 13590A 13833G 14766T 15326G 15535T 16093C 16183C 16189C 16193.1C 16217C 16519C
1	B2g1	mtGHispTX0026	73G 114G 146C 263G 315.1C 499A 709A 750G 827G 1002T 1438G 2706G 3547G 3766C 4769G 4820A 4977C 6164T 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16298C 16519C
1	B2g2	mtGHispCA0018	73G 114G 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 3766C 4769G 4820A 4977C 6040G 6164T 6473T 7028T 7340A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11647T 11719A 11875C 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16519C
1	B2g2	mtGHispTX0021	73G 94A 114G 263G 309.1C 315.1C 499A 750G 827G 1438G 2706G 3547G 3766C 4769G 4820A 4937C 4977C 6164T 6473T 7028T 7340A 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11623T 11647T 11719A 11875C 13590A 14470C 14766T 15326G 15535T 15994R 16148T 16183C 16189C 16193.1C 16217C 16519C
1	B2j	mtGHispFL0010	73G 131C 183G 263G 291.1A 315.1C 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 5270T 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10754R 11177T 11719A 12950G 13590A 14766T 15326G 15535T 15924G 16156A 16166G 16183C 16189C 16193.1C 16217C 16519C
1	B2m	mtGHispCA0031	73G 263G 309.1C 315.1C 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 5585A 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 11177T 11719A 13590A 14766T 15326G 15535T 15766G 16164G 16183C 16189C 16193.1C 16217C 16325Y 16519A
1	B2o	mtGHispAZ0007	73G 159C 263G 315.1C 499A 750G 827G 1438G 2706G 2804G 3547G 4769G 4820A 4977C 6473T 6647C 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8292A 8860G 9950C 11177T 11719A 13590A 13879A 14766T 15326G 15535T 16092C 16182C 16183C 16189C 16217C 16519C
1	B2o	mtGHispCA0013	73G 159C 263G 296T 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 2804G 3547G 4769G 4820A 4977C 6473T 6647C 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9950C 10601C 11177T 11719A 13590A 14766T 15326G 15535T 16092C 16182C 16183C 16189C 16193.1C 16193.2C 16217C 16519C
1	B2q	mtGHispAZ0003	73G 146C 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 2706G 3547G 4047C 4769G 4820A 4977C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9287A 9950C 11177T 11719A 12633T 13590A 14766T 15326G 15535T 16183C 16189C 16193.1C 16217C 16278T 16519C
1	B2r	mtGHispTX0019	73G 152C 263G 309.1C 309.2C 315.1C 499A 750G 827G 1438G 1664A 1888A 2010C 2706G 3547G 4769G 4820A 4977C 5899- 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9615C 9950C 11177T 11719A 13590A 14766T 15326G 15535T 16182C 16183C 16189C 16193.1C 16217C 16519C
1	B2t	mtGHispTX0010	73G 263G 310C 315- 499A 750G 827G 1438G 2706G 3547G 4769G 4820A 4977C 5786C 6473T 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8772C 8860G 9950C 10792G 11177T 11719A 13590A 14766T 15244G 15326G 15535T 15884A 16183C 16189C 16193.1C 16217C 16357C 16467T 16519C
1	B4a1a	mtGHispHI0003	73G 146C 263G 309.1C 309.2C 315.1C 523- 524- 750G 793T 1438G 1842G 2706G 4769G 5465C 6719C 7028T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8860G 9123A 10238C 11719A 12239T 14766T 15326G 15746G 16182C 16183C 16189C 16217C 16261T 16519C

1	C1b	mtGHispCA0028	73G 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 5351G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 9804A 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14560A 14766T 14783C 15043A 15301A 15326G 15487T 16183C 16189C 16193.1C 16223T 16298C 16325C 16327T 16359C
1	C1b	mtGHispTX0012	73G 195C 249- 263G 290- 291- 309.1C 309.2C 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 7028T 7196A 7211A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11152C 11719A 11914A 12071Y 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16092Y 16223T 16243C 16298C 16325C 16327T
1	C1b	mtGHispTX0014	9A 73G 188G 199Y 249- 263G 290- 291- 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 5057T 7028T 7196A 7337A 8155A 8584A 8685A 8701G 8860G 9368G 9380A 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 13701G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16223T 16298C 16325C 16327T 16465T
1	C1b	mtGHispTX0024	73G 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 1438G 1717C 2706G 3335C 3394C 3552A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15941C 16223T 16325C 16327T
1	C1b10	mtGHispCA0001	73G 146C 152C 249- 263G 290- 291- 309.1C 315.1C 385G 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 6284G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14482T 14766T 14783C 15043A 15301A 15326G 15487T 15622C 16129A 16172C 16223T 16298C 16311C 16325C 16327T 16519C
1	C1b11	mtGHispTX0011	72C 73G 194T 249- 263G 290- 291- 315.1C 489C 493G 523- 524- 750G 2706G 3552A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15670C 16223T 16295T 16298C 16325C 16327T
1	C1b14	mtGHispCA0025	73G 228R 249- 263G 290- 291- 315.1C 489C 493G 523- 524- 709A 750G 1438G 2706G 3552A 4715G 4769G 5894G 6872G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10397G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 13953C 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16172C 16181G 16223T 16298C 16325C 16327T
1	C1b4	mtGHispPR0007	73G 143A 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 2706G 3552A 4167T 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14524G 14766T 14783C 15043A 15301A 15326G 15487T 16086C 16183C 16189C 16193.1C 16223T 16278T 16298C 16325C 16327T
1	C1b7a	mtGHispCA0003	60C 73G 249- 263G 290- 291- 309.1C 309.2C 315.1C 489C 493G 523- 524- 750G 1117G 1310T 1438G 2706G 3552A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 13635C 14318C 14766T 14783C 15043A 15301A 15326G 15470C 15487T 16223T 16298C 16311C 16325C 16327T 16463G
1	C1b7a	mtGHispCA0015	60.1T 71- 73G 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 1310T 1438G 2706G 3552A 4715G 4769G 7028T 7196A 8251A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 13635C 14318C 14766T 14783C 15043A 15301A 15326G 15470C 15487T 16223T 16298C 16311C 16325C 16327T 16390A
1	C1b7a	mtGHispCA0034	73G 249- 263G 290- 291- 309.1C 315.1C 489C 493G 512C 523- 524- 750G 1438G 2706G 3552A 4695C 4715G 4769G 5671T 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11800G 11914A 12705T 13263G 13635C 14318C 14766T 14783C 15043A 15301A 15326G 15470C 15487T 16223T 16298C 16311C 16325C 16327T 16519C
1	C1b9	mtGHispNM0004	73G 198T 210G 247A 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 6297C 7028T 7196A 8047C 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A

			11914A 12705T 13263G 13434R 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16223T 16298C 16325C 16327T
1	C1b9	mtGHispTX0017	73G 198T 249- 263G 290- 291- 309.1C 315.1C 489C 493G 523- 524- 750G 1438G 2706G 3552A 4715G 4769G 5887Y 6297C 7028T 7196A 8047C 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 16223T 16298C 16325C 16327T
1	C1c	mtGHispFL0001	73G 249- 263G 290- 291- 293C 315.1C 489C 750G 1438G 1888A 2706G 3552A 4715G 4769G 5333C 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15930A 16223T 16298C 16325C 16327T
1	C1c	mtGHispNY0001	73G 249- 263G 290- 291- 293C 315.1C 489C 750G 1189C 1438G 1888A 2706G 3552A 4715G 4769G 5333C 7028T 7196A 8584A 8701G 8860G 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15930A 16223T 16298C 16325C 16327T
1	C1c	mtGHispTX0038	71.1G 73G 249- 263G 290- 291- 309.1C 315.1C 489C 750G 1438G 1888A 2366A 2706G 3552A 4506G 4715G 4769G 7028T 7196A 8584A 8701G 8860G 8994A 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12705T 13153G 13263G 14318C 14766T 14783C 15043A 15064G 15301A 15326G 15487T 15930A 16126C 16298C 16325C 16327T 16519C
1	C1c6	mtGHispAZ0002	73G 249- 263G 290- 291- 309.1C 315.1C 489C 750G 1438G 1888A 2706G 3552A 3693A 4715G 4769G 7028T 7196A 8584A 8701G 8860G 9230C 9540C 9545G 10398G 10400T 10873C 11719A 11914A 12414C 12705T 13105G 13263G 14318C 14766T 14783C 15043A 15301A 15326G 15487T 15930A 16153A 16223T 16298C 16325C
1	D1	mtGHispCA0008	73G 263G 315.1C 489C 750G 961C 965.1C 965.2C 1438G 2092T 2706G 3010A 3834A 4769G 4883T 5178A 6254G 7028T 8414T 8701G 8860G 8871G 9540C 9591A 10188G 10398G 10400T 10873C 11719A 12705T 14668T 14766T 14783C 15043A 15301A 15326G 15984C 16223T 16325C 16362C
1	D1	mtGHispCA0019	73G 263G 309.1C 309.2C 315.1C 489C 750G 1438G 2092T 2706G 3010A 4769G 4883T 5178A 7028T 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11719A 12705T 14668T 14766T 14783C 15043A 15301A 15326G 16223T 16325C 16362C
1	D1	mtGHispNM0002	73G 263G 292C 309.1C 315.1C 489C 750G 1438G 2092T 2706G 3010A 4769G 4883T 5178A 7028T 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11719A 12705T 14668T 14766T 14783C 15043A 15301A 15326G 16223T 16325C 16362C 16526A
1	D1	mtGHispNY0007	55.1T 57C 59C 64T 73G 131C 143A 189G 207A 263G 309.1C 309.2C 315.1C 489C 750G 1438G 2092T 2706G 3010A 3699G 4769G 4883T 5178A 5655C 7028T 8701G 8860G 9540C 10097G 10398G 10400T 10873C 11719A 12705T 14668T 14766T 14783C 15043A 15301A 15326G 16183C 16189C 16194C 16195C 16223T 16292T 16325C 16362C
1	D1	mtGHispTX0007	73G 263G 309.1C 315.1C 489C 750G 1438G 2092T 2706G 3010A 4769G 4883T 5178A 7028T 8222C 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11719A 12705T 13488C 14668T 14766T 14783C 15043A 15301A 15326G 16223T 16325C 16362C 16519C
1	D1h1	mtGHispIL0003	73G 204C 263G 309.1C 315.1C 489C 750G 1438G 2092T 2706G 3010A 4769G 4883T 5178A 7028T 7861C 8414T 8701G 8860G 9095C 9540C 10398G 10400T 10873C 11719A 12594Y 12705T 13327R 13635C 14668T 14766T 14783C 15043A 15301A 15326G 16092C 16223T 16256T 16274A 16325C 16362C
1	D1i	mtGHispTX0049	73G 263G 315.1C 417A 489C 750G 1438G 2092T 2706G 3010A 3438A 4769G 4883T 5178A 5237A 7028T 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11719A 12280G 12705T 14668T 14766T 14783C 15043A 15301A 15326G 16223T 16325C 16362C
1	D1i2	mtGHispCA0014	73G 263G 315.1C 417A 489C 551G 709A 750G 1438G 2092T 2706G 3010A 3316R 4769G 4883T 5178A 7028T 7642R 8414T 8701G 8860G 9540C 10398G 10400T 10873C 11314G 11719A 12280G 12705T 14668T 14766T 14783C 15001C 15043A 15301A 15326G 15877A 16223T 16274A 16325C 16362C 16368C

1	D4h3a	mtGHispCA0021	73G 263G 315.1C 489C 523- 524- 750G 1438G 2706G 3010A 3336C 3396C 3644C 3927G 4025T 4092A 4646Y 4769G 4883T 5048C 5178A 5480G 6285A 7028T 8414T 8701G 8713G 8860G 8946G 9458T 9540C 10398G 10400T 10873C 11719A 12705T 13135A 14668T 14766T 14783C 15043A 15301A 15326G 15734A 16223T 16241G 16263C 16301T 16342C 16362C
1	E2a1	mtGHispCA0032	73G 195C 263G 315.1C 489C 750G 1438G 2706G 3027C 3705A 4491A 4769G 7028T 7598A 8440G 8701G 8730G 8860G 9080G 9254G 9540C 10398G 10400T 10873C 11719A 12705T 13626T 14766T 14783C 15043A 15178G 15301A 15326G 16051G 16223T 16362C 16390A 16519C
1	H	mtGHispNJ0004	73R 263G 315.1C 629Y 750G 1438G 4769G 8860G 15326G 16093C 16223T
1	H1	mtGHispVA0002	263G 315.1C 750G 1438G 3010A 3421A 4769G 5075C 8860G 15326G 16519C
1	H10e1	mtGHispTX0041	263G 309.1C 315.1C 750G 1438G 4769G 8860G 13830C 14470A 15326G 16221T 16519C
1	H1a	mtGHispCA0005	73G 263G 309.1C 315.1C 750G 1438G 3010A 4769G 8860G 15326G 16162G 16519C
1	H1am	mtGHispNV0004	263G 309.1C 315.1C 750G 1438G 3010A 4763A 4769G 8860G 15326G 16519C
1	H1b	mtGHispWA0001	93G 263G 315.1C 315.2C 750G 1438G 3010A 4769G 8860G 15326G 16093C 16189C 16193.1C 16193.2C 16356C 16519C
1	H1b1	mtGHispKS0001	263G 315.1C 523- 524- 750G 1438G 3010A 3796G 4769G 8860G 15326G 16183C 16189C 16193.1C 16356C 16362C 16519C
1	H1c1	mtGHispIL0001	263G 315.1C 477C 750G 1438G 3010A 4769G 5147R 5945T 8860G 9150G 15326G 16147Y 16263C 16519C
1	H1c21	mtGHispCA0036	263G 309.1C 315.1C 477C 750G 1438G 3010A 4638R 4767G 4769G 8860G 15326G 16189C 16519C
1	H1e1a6	mtGHispCA0037	150T 263G 315.1C 750G 1438G 3010A 4769G 5460A 5964C 8512G 8860G 14902T 15326G 16256Y 16519C
1	H1i1	mtGHispVA0001	152C 263G 315.1C 750G 1438G 2784R 3010A 4769G 6237A 8860G 9300A 15326G 16519C
1	H2a2a1	mtGHispMA0001	315.1C 3107-
1	H2a2b	mtGHispGA0002	263G 309.1C 309.2C 315.1C 8860G 15326G 16291T
1	H3	mtGHispTN0001	263G 315.1C 750G 1438G 4769G 6776C 8860G 12950G 13928C 15326G 16126C 16145A 16166G 16519C
1	H3	mtGHispTX0040	263G 315.1C 750G 1438G 4769G 6776C 8860G 14900A 15326G 16519C
1	H3ap	mtGHispTX0032	263G 309.1C 309.2C 315.1C 750G 1438G 4769G 6590C 6776C 8860G 10199T 10915C 15326G 16042A 16519C
1	H5a3b	mtGHispCA0033	263G 309.1C 309.2C 315.1C 456T 513A 750G 1438G 4336C 4769G 8860G 12648G 15326G 15884A 16093Y 16304C
1	H7	mtGHispTX0009	263G 315.1C 750G 1438G 4769G 4793G 8860G 15326G 16249C 16519C
1	I2	mtGHispAZ0001	73G 152C 199C 204C 207A 250C 263G 309.1C 315.1C 573.1C 573.2C 573.3C 573.4C 750G 1438G 1719A 2706G 4529T 4769G 5973A 7028T 8251A 8860G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14560A 14766T 15043A 15326G 15758G 15924G 16129A 16223T 16391A 16519C
1	I5a2	mtGHispTX0039	73G 199C 204C 250C 263G 309.1C 315.1C 385G 573.1C 573.2C 573.3C 573.4C 573.5C 750G 1438G 1719A 2706G 3615G 3918A 4529T 4769G 5074C 5096C 7028T 8251A 8742G 8860G 9254G 10034C 10238C 10398G 11719A 12501A 12705T 13780G 14088C 14233G 14766T 15043A 15326G 15924G 16092C 16129A 16148T 16223T 16354T 16391A 16519C
1	J1b2	mtGHispCA0002	73G 263G 295T 315.1C 462T 489C 750G 1438G 1733T 2706G 3010A 4216C 4769G 6719C 7028T 8269A 8860G 10398G 11251G 11719A 12612G 13708A 14766T 14927G 15326G 15452A 16069T 16126C 16145A 16222T 16261T
1	J1c8a2	mtGHispFL0006	73G 185A 228A 263G 295T 315.1C 462T 489C 750G 1438G 2706G 2707R 3010A 4216C 4769G 7028T 7598A 8860G 10084C 11251G 11719A 12612G 13708A 14766T 14798C 15326G 15452A 16069T 16126C 16265C 16319A
1	K1a10a	mtGHispMD0001	73G 150T 195C 263G 315.1C 497T 524.1A 524.2C 524.3A 524.4C 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11719A 12308G 12372A 14167T 14766T 14798C 15326G 16048A 16093C 16172Y 16224C 16291T 16311C 16519C

1	K1a1b1d	mtGHispTX0033	73G 263G 315.1C 497T 750G 1189C 1438G 1811G 2706G 3480G 4769G 7028T 8860G 9055A 9698C 10398G 10550G 11299C 11467G 11470G 11719A 11914A 12308G 12372A 14167T 14388G 14569A 14766T 14798C 15326G 15924G 16092C 16223T 16224C 16311C 16519C
1	K2b1a	mtGHispAZ0008	73G 146C 263G 315.1C 524.1A 524.2C 750G 1438G 1811G 2217T 2706G 3480G 4769G 5231A 7028T 8860G 9055A 9698C 9716C 10550G 11299C 11467G 11719A 11869A 12308G 12372A 13135A 14037G 14167T 14766T 14798C 15326G 16213A 16224C 16311C 16519C
1	L0a2a1a1	mtGHispMN0001	64T 93G 152C 189G 236C 247A 263G 315.1C 523- 524- 750G 769A 825A 1018A 1048T 1438G 2245G 2706G 2758A 2885C 3516A 3594T 4104G 4312T 4586C 4598C 4769G 5147A 5231A 5442C 5460A 5603T 5711G 6185C 6257A 7028T 7146G 7256T 7521A 8264T 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8428T 8460G 8468T 8566G 8655T 8701G 8860G 9042T 9178A 9347G 9540C 9755A 9818T 10398G 10589A 10664T 10688A 10810C 10873C 10915C 11143T 11172G 11176A 11641G 11719A 11830C 11914A 12007A 12705T 12720G 13105G 13276G 13506T 13650T 14182C 14308C 14755G 14766T 15136T 15326G 15431A 16148T 16172C 16187T 16188G 16189C 16223T 16230G 16311C 16320T 16519C
1	L1b1a7a	mtGHispNY0008	73G 152C 182T 185T 195C 228A 247A 263G 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 2352C 2706G 2758A 2768G 2885C 3308C 3594T 3666A 3693A 4104G 4769G 5036G 5046A 5393C 5655C 6378C 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12519C 12705T 13105G 13506T 13650T 13789C 13880A 14053G 14178C 14203G 14560A 14766T 14769G 15115C 15224T 15326G 16126C 16187T 16189C 16223T 16264T 16270T 16278T 16293G 16311C 16519C
1	L1b2a	mtGHispGA0001	73G 146C 152C 182T 185T 189G 247A 263G 309.1C 315.1C 357G 523- 524- 709A 710C 750G 769A 825A 1018A 1738C 1842G 2352C 2706G 2758A 2768G 2885C 3308C 3339G 3594T 3666A 3693A 4104G 4755C 4769G 6548T 6827C 6989G 7028T 7055G 7146G 7256T 7389C 7521A 7867T 8248G 8468T 8655T 8701G 8860G 9540C 10398G 10688A 10810C 10873C 11719A 12171G 12519C 12705T 12891T 13105G 13506T 13650T 13789C 13893G 14178C 14323A 14560A 14766T 14769G 15115C 15326G 16111T 16126C 16187T 16189C 16223T 16239T 16270T 16278T 16293G 16311C 16519C
1	L1c2b1b	mtGHispNJ0005	73G 151T 152C 182T 186A 189C 195C 198T 247A 263G 297G 315.1C 316A 513A 750G 769A 825A 1018A 1438G 2220G 2395- 2706G 2758A 2885C 3202C 3594T 3666A 4104G 4769G 5087C 5814C 5899.1C 5951G 6071C 6150A 6253C 6480A 7028T 7055G 7076G 7146G 7256T 7337A 7389C 7521A 8027A 8468T 8655T 8701G 8784G 8860G 8877C 9072G 9108G 9540C 10031C 10321C 10398G 10586A 10688A 10792G 10793T 10810C 10873C 11164G 11654G 11719A 11963A 12049T 12669T 12705T 12810G 13105G 13149G 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14587G 14766T 14911T 15326G 15924G 16129A 16189C 16214T 16223T 16265C 16278T 16286A 16291T 16294T 16311C 16360T 16519C 16527T
1	L1c5	mtGHispCA0040	73G 151T 152C 182T 186A 189C 247A 263G 291T 297G 315.1C 316A 523- 524- 709A 750G 769A 825A 1018A 1438G 2395- 2706G 2758A 2885C 3594T 3666A 4104G 4769G 5390G 5951G 6071C 7028T 7055G 7146G 7256T 7389C 7521A 7762A 8027A 8143C 8281- 8282- 8283- 8284- 8285- 8286- 8287- 8288- 8289- 8468T 8655T 8701G 8860G 9072G 9540C 9899C 10398G 10586A 10688A 10810C 10873C 11150A 11719A 12425G 12630A 12705T 12810G 13105G 13359A 13368A 13485G 13506T 13650T 13789C 14000A 14178C 14560A 14766T 14911T 15326G 15449C 15553A 15941C 16114G 16129A 16187T 16189C 16223T 16261T 16278T 16311C 16360T 16519C
1	L2a1f	mtGHispNV0001	73G 146C 152C 195C 263G 309.1C 315.1C 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5581G 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9248T 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14566G 14766T 15301A 15326G 15784C 16189C 16192T 16223T 16278T 16294T 16309G 16390A 16519C

1	L2a1l2	mtGHispNY0003	73G 143A 146C 152C 195C 263G 315.1C 534T 750G 769A 1018A 1438G 2416C 2706G 2789T 3594T 4104G 4769G 5366A 7028T 7175C 7256T 7274T 7521A 7771G 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11914A 11944C 12693G 12705T 13590A 13650T 13803G 14110C 14566G 14766T 15301A 15326G 15784C 16189C 16223T 16278T 16294T 16309G 16390A
1	L2c	mtGHispPR0004	73G 93G 95C 146C 150T 152C 182T 195C 198T 263G 315.1C 325T 523- 524- 680C 709A 750G 769A 1018A 1438G 1442A 2332T 2416C 2706G 3200A 3594T 4104G 4769G 7028T 7256T 7521A 7624A 8206A 8701G 8860G 9221G 9540C 10115C 10398G 10873C 11719A 11944C 12236A 12705T 13590A 13650T 13928C 13958C 14750T 14766T 15110A 15217A 15301A 15326G 15849T 16180G 16223T 16278T 16390A 16519C
1	L3b1a	mtGHispPR0005	73G 152C 263G 315.1C 523- 524- 750G 1438G 2706G 3450T 4769G 5773A 6221C 7028T 8547C 8701G 8772C 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13105G 13914A 13926C 14766T 15301A 15311G 15326G 15434T 15824G 15944- 16124C 16223T 16278T 16362C 16519C
1	L3b1a10	mtGHispPR0001	73G 263G 315.1C 523- 524- 750G 1438G 2706G 3450T 4659A 4769G 5773A 6221C 7028T 7859A 8701G 8860G 9449T 9540C 10086G 10373A 10398G 10873C 11002G 11719A 12705T 13105G 13914A 14182C 14323A 14766T 15301A 15311G 15326G 15824G 15944- 16124C 16223T 16278T 16362C 16519C
1	L3d1'2'3'4'5'6	mtGHispTX0016	73G 152C 263G 309.1C 315.1C 523- 524- 750G 921C 1438G 2706G 3504C 4619C 4769G 5147A 7028T 7424G 8618C 8701G 8860G 9540C 10398G 10873C 11503T 11719A 12705T 13105G 13368A 13886C 13966G 14284T 14766T 15301A 15326G 15434T 16124C 16148T 16223T 16311C 16362C 16519C
1	L3d2b	mtGHispNY0004	73G 152C 199C 263G 309.1C 315.1C 523- 524- 750G 921C 1438G 2163G 2706G 4688C 4769G 5147A 7028T 7424G 8618C 8701G 8860G 9540C 10398G 10873C 11150A 11719A 12705T 13105G 13886C 14272T 14284T 14584C 14766T 15301A 15326G 16111T 16124C 16223T
1	L3e1d1	mtGHispFL0008	73G 150T 152C 189G 200G 263G 309.1C 315.1C 750G 1438G 2352C 2706G 4769G 6221C 6587T 7028T 8701G 8703T 8860G 9300A 9540C 10398G 10819G 10873C 11176A 11719A 12705T 12738C 14152G 14212C 14766T 15301A 15326G 15670C 15942C 16176T 16223T 16256T 16327T
1	L3e2a1b	mtGHispFL0007	73G 150T 195C 198T 263G 315.1C 750G 793T 1438G 2352C 2706G 4769G 4823C 6413C 7028T 8011T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 13105G 14212C 14766T 14869A 14905A 15301A 15319T 15326G 16223T 16320T 16519C
1	L3e2b	mtGHispNY0006	73G 150T 195C 263G 315.1C 750G 1438G 2352C 2706G 4769G 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11719A 12705T 14212C 14766T 14905A 15287C 15301A 15326G 16172C 16183C 16189C 16193.1C 16223T 16301T 16320T 16519C
1	L3e4a	mtGHispCA0035	73G 150T 263G 315.1C 523- 524- 1438G 2352C 2706G 3915A 4769G 5262A 5584G 7028T 8701G 8860G 9540C 10398G 10819G 10873C 11257T 11719A 12705T 13368A 13749T 14212C 14662G 14766T 15301A 15326G 16051G 16093C 16223T 16264T 16311C 16519C
1	L3f1b4	mtGHispTX0045	73G 150T 189G 200G 263G 309.1C 315.1C 750G 1438G 1822C 2706G 3396C 3505G 4218C 4769G 5601T 7028T 7819A 8527G 8701G 8860G 8932T 8937A 9540C 9950C 10398G 10873C 11440A 11719A 12705T 14766T 14769G 15301A 15326G 15514C 15629C 15944- 16209C 16223T 16311C 16519C
1	T1a1	mtGHispNC0001	73G 152C 195C 263G 309.1C 315.1C 709A 750G 1438G 1888A 2706G 4216C 4769G 4917G 6845T 7028T 8697A 8860G 9899C 10463C 11251G 11719A 12633A 13368A 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16163G 16186T 16189C 16294T 16519C
1	T2a1b1a	mtGHispNJ0001	73G 263G 315.1C 709A 750G 1438G 1888A 2141C 2396T 2706G 4216C 4769G 4917G 7028T 8563G 8697A 8860G 9117C 10463C 11251G 11719A 11812G 13145A 13368A 13965C 13966G 14233G 14687G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16324C 16519C
1	T2b19b	mtGHispIA0001	73G 195C 263G 309.1C 315.1C 523- 524- 709A 750G 930A 1438G 1664A 1888A 2706G 4216C 4769G 4917G 4944G 5147A 7028T 7859A 8697A 8860G 10463C 11251G 11719A 11812G 13368A 13681G 13928C 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16248T 16294T 16296T 16304C 16519C

1	T2b3b	mtGHispAR0001	73G 152C 263G 309.1C 315.1C 709A 750G 930A 1438G 1888A 2706G 4216C 4769G 4917G 5147A 7028T 8697A 8860G 10463C 10750G 11251G 11719A 11812G 13368A 13722G 14233G 14766T 14905A 15326G 15452A 15607G 15928A 16126C 16294T 16296T 16304C 16519C
1	U5a2a1d	mtGHispFL0003	73G 263G 309.1C 315.1C 750G 1438G 2706G 3197C 4232C 4655A 4769G 7028T 8020A 8860G 9477A 11467G 11719A 11893G 12308G 12372A 13617C 13827G 13928C 14766T 14793G 15326G 16114A 16126Y 16171G 16192T 16256T 16270T 16294T 16526A
1	U5a2b3a	mtGHispAZ0005	73G 150T 263G 309.1C 315.1C 455- 750G 1438G 2706G 3197C 4769G 7028T 8860G 9477A 9548A 11467G 11719A 12308G 12372A 13246C 13351T 13617C 14684T 14766T 14793G 15326G 16168T 16192T 16256T 16270T 16526A
1	U5b1c2	mtGHispCA0007	73G 146C 150T 263G 309.1C 315.1C 516T 750G 1438G 2706G 3197C 4769G 5656G 7028T 7768G 8860G 9110C 9477A 11467G 11569C 11719A 12308G 12372A 13617C 14182C 14766T 15191C 15326G 16174T 16189C 16270T 16311C
1	U5b1c2b	mtGHispIN0002	73G 150T 263G 315.1C 516T 750G 1438G 2706G 3197C 4769G 5656G 6341T 7028T 7768G 8860G 9477A 11467G 11719A 12308G 12372A 13194A 13617C 14182C 14766T 15191C 15326G 16174T 16189C 16192T 16270T 16311C
1	U5b2b	mtGHispCA0011	73G 150T 263G 309.1C 315.1C 750G 1438G 1721T 2706G 3197C 4769G 5460A 7028T 7768G 8860G 9477A 10908C 11467G 11653G 11719A 11923G 12308G 12372A 12634G 13617C 13630G 13637G 14182C 14766T 15326G 16270T
1	U5b3	mtGHispCA0012	73G 150T 228A 263G 315.1C 338T 750G 1438G 2706G 3197C 3203G 3710T 4248C 4769G 7028T 7226A 7768G 8860G 9477A 11467G 11719A 12308G 12372A 13617C 13926C 14182C 14766T 15326G 16192T 16270T 16304C 16526A
1	W1	mtGHispHI0002	73G 119C 189G 195C 204C 207A 263G 315.1C 709A 750G 1243C 1438G 2706G 2905R 3505G 3795T 4769G 5046A 5460A 5495C 7028T 7864T 8251A 8860G 8994A 11674T 11719A 11947G 12414C 12705T 14766T 15326G 15884C 16223T 16292T 16519C
1	W3a1b	mtGHispNY0009	73G 146C 189G 194T 195C 204C 207A 263G 309.1C 315.1C 709A 750G 1243C 1406C 1438G 2706G 3505G 4769G 5046A 5460A 7028T 8251A 8860G 8994A 9611T 10245C 11674T 11719A 11947G 12414C 12705T 13263G 14766T 15326G 15784C 15884C 16223T 16290T 16292T 16519C
1	Y2a1a	mtGHispTX0008	73G 263G 309.1C 315.1C 482C 750G 1438G 2706G 2856T 4769G 5147A 5417A 6941C 7028T 7859A 8289.1C 8289.2C 8289.3C 8289.4C 8289.5C 8289.6T 8289.7C 8289.8T 8289.9A 8392A 8860G 10398G 11299C 11719A 12161C 12705T 14178C 14693G 14766T 14914G 15244G 15326G 16126C 16231C 16311C