Developing a DNA Microarray for Human Ancestry and Migration

An honors thesis for the Department of Chemistry

Jennifer Lynn Torpey

Tufts University, 2007

Acknowledgements:

Thesis Advisor: Dr. David R. Walt

Thesis Committee: Dr. Krishna Kumar
Dr. Mitch McVey

I would like to thank the Howard Hughes Medical Institute for supporting this project and the members of the Walt Laboratory for their guidance, assistance and patience: Ryan Hayman, Dr. Dave Rissin, Ragnhild Whitaker, Tim Blicharz, Stacey Watkins, Dr. Chris Lafratta, Kyle Bake, Chak Chow, Matt Aernecke, Dr. Hans-Heiner Gorris, Dr. Alexei Faustov, and Dr .Xin Yu. I would also like to thank my family and friends for their unconditional love and support.

Table of Contents: **Page number:**

Acknowledgements ii

Table of Contents iii

Chapter 1. Introduction **1**

1.1 The Human Genome Project and Mitochondrial DNA 1
1.2 Human Evolution and Migration 4
1.3 Phylogenetic Networks: Implications for Genealogy and Human Migration 7
1.4 mtDNA Databases 7
1.5 Genealogical Analyses 9
1.6 Howard Hughes Medical Institute Research Project 10
1.7 DNA Microarrays 12
1.8 The Illumina GoldenGate™ Assay 12

Chapter 2. Materials and Methods **16**

2.1 Materials 16
2.2 Illumina Array Design 17
2.3 Sample Collection 17
2.4 DNA Extraction 18
2.5 Mitochondrial DNA Amplification 19
2.6 Mitochondrial DNA Identification 20
2.7 Illumina Array Protocol 21
2.8 Statistical Analysis 21

Chapter 3. Results and Discussion **29**

3.1 Illumina Array Design 29
3.2 Estimated Maternal Background 53
3.3 Identification of mtDNA 54
3.4 mtDNA Amplification 55

3.5 SNP Calls 57
3.6 FamilyTree® DNA Results 66

Chapter 4. Future Work **68**

Appendix **70**

References **75**

Chapter 1: INTRODUCTION

1.1 The Human Genome Project and Mitochondrial DNA

The entire human genome contains approximately 3 billion nucleotides whose sequence was recently completed in 2003 by the Human Genome Project. The full implications of this project will be manifested in important projects throughout the natural and social science disciplines within the upcoming decades. Several of the most exciting research topics involving genetic components include molecular medicine [1], energy sources and environmental applications [2], disease risk assessment [3], DNA forensics [4, 5], bioprocessing, bioarchaeology, anthropology, and human evolution and migration [6]. The goals of the Human Genome Project were first to identify the approximately 20,000-30,000 genes coded within the DNA and then determine the sequence of the nucleotide bases that comprise the entire genome. This information has proven to be critical for understanding many biological pathways and creates an entirely new avenue for nucleotide-specific scientific investigation. In the Genome Project it was determined that 99.9% of the nucleotides which make up the human genome are conserved, leaving only a small number of differences that characterize individuals [7]. It is these differences that research scientists now aim to identify and classify to further understand the molecular basis of human life and evolution, and recognize the root of genetic disease and phylogeny.

Full nuclear genome identification and analysis is only one part of a larger-scale human DNA investigation that has been underway for decades. The mitochondrial genome has been the molecule of choice for genetic analyses prior to the dawn of the 21st century and continues to be a hot topic in molecular biology, specifically related to

determining human evolution and migration patterns. This genome, whose host organelle is essential for cellular energy production, is a closed circular DNA structure containing approximately 16,000 nucleotides that encompass 37 known genes [8]. In addition to its size, this genome was chosen for population studies for several reasons. The mitochondrial genome is independently replicated and regulated, and holds its own genetic code different from that of nuclear DNA. Mitochondrial (mt) DNA is particularly well-suited for genetic genealogical analyses, in comparison to nuclear DNA, because of its high rate of mutation, high cellular concentration, maternal mode of inheritance, and apparent lack of recombination [9, 10]. These last two features are especially ideal for population genetics because they allow genomic mutations to be tracked, enabling correlations between related individuals from different generations and those from varied geographic locations. The high rate of mutation, which has been attributed to the lack of mitochondrial histones and a high concentration of oxidative radicals [11], provides an increased resolution for analyzing more recent evolutionary events since mtDNA mutates 5-10 times faster than nuclear DNA [12]. These positive features of mtDNA are juxtaposed with those of nuclear DNA which is present in approximately two copies per cell and transmitted equally, with recombination, from both parents [13] making it an inferior platform for evolutionary analysis.

These positive aspects of mtDNA (high cellular concentration, high rate of mutation, maternal mode of inheritance and lack of recombination) have been highly debated as the popularity of using this molecule for genomic analyses has continued to increase. Initially, a correlation between linkage disequilibrium (LD) and the distance between variable sites suggested recombination had occurred in the mtDNA genome

[14]. However, further analyses maintained that there was no evidence of recombination and the use of the LD measure was criticized, since it did not take allele frequency into account and used limited portions of the mtDNA sequence [15]. Despite the confirmation of heteroplasmy amongst the hundreds of mtDNA molecules per cell, recombination among these molecules has not been convincingly demonstrated [11]. At the present time, the positive attributes of the mitochondrial genome surpass those of chromosomal DNA. Thus, mtDNA is a useful material for genetic analysis that contains distinct variations which are desirable for population genetics studies [10].

Human genetic analysis for genealogical purposes began with a landmark study by W.M. Brown in 1980 focused on mtDNA sequence information from 21 individuals of diverse ethnic backgrounds and yielded an outline for the disentanglement of human history. Points of interest within the mitochondrial genome were analyzed by restriction fragment length polymorphism (RFLP) patterns and compared for potentially useful point mutations. Comparing these sequences suggested the existence of a single common ancestor, which Brown estimated lived in the surprisingly recent past at ~180,000 years ago [8]. With the influx of whole genome sequencing technology, mainly from techniques developed for the Human Genome Project, the ability to obtain entire sequences became much more cost-effective and feasible. Initial studies of the mitochondrial genome focused on RFLP patterns of either purified genomic or mitochondrial DNA, followed by sequencing of select regions, mainly the D-loop or control region, using PCR and rapid sequencing technology. These analyses were limited to certain portions of the mitochondrial genome and mainly focused on the first hypervariable segment (HVR I or HVS I) since this region, and later HVR II, was

designated as a mutational hotspot. The development of high-throughput sequencing in recent years has made it more common to analyze whole mitochondrial genome sequences rather than RFLP patterns or just the sequence of HVR I. A debate currently exists over the worthiness of the additional cost and effort of this sequencing since phylogenetic trees based on whole genome sequences are similar to those generated using just the HVR I and RFLP patterns [10]. However, whole genome sequencing is still gaining popularity as recent data claim that the D-loop, containing 30% of the polymorphisms within the genome, only comprises 7% of the genome itself. If an analysis were carried out on this region alone, many regions of interest would be disregarded. RFLP patterns themselves are limited by the recognition sites available to the restriction enzymes, so even those patterns that use the entire mitochondrial genome present another opportunity to potentially overlook valuable polymorphisms [11]. It has also been suggested that RFLP patterns do not show as strong of support for major branches within phylogenetic trees, resulting in ambiguous structures [11].

1.2 Human Evolution and Migration

There are two conflicting philosophies regarding the study of population genomics and the origin of modern humans, namely, the "multi-regional" theory, and the "Out of Africa" theory. The "multi-regional" theory, which is primarily based on fossil evidence, postulates that the transition to anatomically modern humans occurred concurrently in different parts of the world from a widely-distributed progenitor species. This fossil evidence has been found in and outside of Africa demonstrating continuity in its transition from ancient to modern human forms [11]. The other theory, nicknamed

"Out of Africa," suggests that the single common ancestor who gave birth to the modern human race existed approximately 100,000 to 200,000 years ago. This theory assumes that this "mitochondrial Eve" originated in Sub-Saharan Africa and subsequently passed on her heritable information throughout the world with little genetic mixing [13, 16]. Substitution rates, inside and out of the control region, have been used to estimate the date of this most recent common ancestor (MRCA).

The offspring of "mitochondrial Eve" and the many subsequent generations spread out across the globe following along certain paths that are now being identified through genetic analyses. Along this journey, mutations occurred within these individuals' genomes that were passed down from one generation to the next leaving a trail of human migration. These mutations, called single nucleotide polymorphisms (SNPs), are found within the mitochondrial genome and have been correlated with certain haplogroups based on the phylogenetic networks generated from sequence analyses. These haplogroups are ancestral designations based on the theoretical migration of the human population away from its origin in sub-Saharan Africa. A collaborative effort between National Geographic, IBM, and the Waitt Foundation called The Genographic Project is now underway to more specifically define these haplogroups and trace each branch's origin by collecting thousands of mitochondrial genome sequences to create a hypothetical map of human migration. The project aims to answer the following questions: "How did we, each of us, end up where we are?" and "Why do we appear in such a wide array of different colors and features [17]?" According to Project Director and explorer-in-residence Dr. Spencer Wells, "The greatest history book ever written is the one hidden in our DNA [17]." With the establishment of ten research laboratories

throughout the globe, the project explorers aim to collect genetic samples from the world's remaining indigenous and traditional peoples whose ethnic and genetic identities are isolated. Mixing populations within the global 21st century melting pot will irreversibly scramble the genetic puzzle. The origins of some of these branches are still under investigation; however, a hypothetical map of human migration has been created using whole genome sequencing results. The haplogroups are described by their geographic origins and locations based on information present in the DNA from different ethnic groups, and their proposed migration patterns in combination with anthropological and archaeological data. Each haplogroup is assigned a letter (A, B, C etc.) and branches within these haplogroups are given numbers, and additional letters if necessary (A1, B5b, etc.).

Haplogroups tend to be continent-specific; however, as the migratory map unfolds, new subgroups are being continually uncovered throughout the world. On a broad scale, haplogroups in the L category represent African descent, haplogroups A, B, C, and D characterize Central Asia and the Americas, haplogroup M migrates through South Asia and Oceania, and haplogroups U, V, W, K, J and I are found in Europe [17]. In general, there have been nine European, seven Asian, and three African haplogroups characterized by certain mutations within the mtDNA genome [18]. As the subgroups within these haplogroups become more clearly defined, individuals can learn of their ancestral origins and it is hoped this information will lead to a more advanced understanding of the complexity and diversity that developed into the modern human race. While this knowledge provides information for the development of anthropological studies, it can also manifest itself in geological and ecological manners, for example,

where and when did humans migrate between continents, by what paths, and what did they take with them or leave behind? All of these aspects of the history of the human race can be augmented by studying mitochondrial DNA. As generation after generation moved out and away from its origins in sub-Saharan Africa, mutations that arose within their genomes were passed down, leaving a treasure map for population geneticists.

1.3 Phylogenetic Networks: Implications for Genealogy and Human Migration

A conceptual definition of a phylogenetic tree is a leaf-labeled "tree" diagram that represents the evolutionary history of a set of taxa; however, it is difficult to fit complex evolutionary scenarios into these models. In order to visualize and explore evolutionary relationships, different types of phylogenetic networks were developed to represent and analyze the data in different ways [19]. These networks have been essential to the identification of certain haplogroups and their subgroups since the nodal arrangement of these networks creates an organized manner in which to view connections between different sequences from known populations [18, 20]. Many large-scale experiments have been carried out in recent years to develop these networks by sequencing hundreds of mitochondrial genomes from individuals from specific populations [18, 21-28].

1.4 mtDNA Databases

There are several databases currently available that provide valuable information for those studying the mitochondrial genome. The three most popular databases are the mtDB: Human Mitochondrial Database [29], the HmtDB: Human Mitochondrial Genomic Resource [30], and MITOMAP [31], all providing an organized fashion to

easily view the polymorphisms currently under investigation and those which have been published in various articles. The usefulness of these databases stems from their ability to house many sequences that scientists can easily locate and reference. In order to make comparisons between mitochondrial genomes to select useful polymorphisms, it is important to use as many complete sequences as possible. The error potential is evident in the first edition of the Cambridge reference sequence [32] which contained several mistakes and rare polymorphisms that were later corrected [33]. With the compilation of numerous whole genome sequences, thousands of polymorphic sites have been identified, providing an essential resource for population geneticists as well as forensic and disease scientists. Pathological mtDNA mutations have been identified in relation to certain diseases such as Leigh syndrome, Leber hereditary optic neuropathy, ocular myopathies including Kearns-Sayre syndrome and chronic progressive external ophthalmoplegia, Pearson marrow/pancreas syndrome and adult-onset diabetes mellitus and deafness [34]. In addition, mtDNA typing has become routine in forensic biology to analyze bones, hair shafts, teeth, and other biological samples [5, 35]. Because of the multitude of different SNPs present in every individual, it becomes necessary to identify the specific information, if any, contained in each polymorphism. The sequencing of the entire mitochondrial genome is providing an overwhelming amount of information, and comparisons of the differences between individuals is uncovering the information hidden within the genome.

1.5 Genealogical Analyses

Interest in genealogical studies has consistently permeated our society especially as each generation creates another layer of ancestral complexity. Prior to the 1980s, individual's looking to trace their ancestral roots had to rely on anthropological and archaeological information in order to determine certain relations between families and cultures [36]. Currently, the genetic portion of genealogy has taken over as evidenced by the multitude of commercial and academic companies advertising individual genetic analysis for ancestry determination. One such company whose goal is to provide ancestral information based on genetic analysis is FamilyTree DNA®. This company, titled "America's first genealogy-driven DNA testing service," offers several genetic assessments including maternal and paternal DNA analyses using DNA from the Y chromosome and the mitochondrial genome, respectively. This company is working in conjunction with the Genographic Project to process the multitude of incoming samples. In addition, this company will provide the same service, for a fee, to any individual who would like his or her genetic information. The haplogroups of these individuals can be determined by sequencing mitochondrial genomes extracted from buccal swab kits. FamilyTree DNA® also provides, upon request, contact information for anyone who shows similar genomic markers. By comparing genetic and genealogical information from the sequences within this company's database, individuals are able to determine their ancestral origins.

The next step for the development of genealogical assays will likely involve a transition to rapid and cost-effective whole-genome genotyping. High-throughput mtDNA SNP genotyping appears to be the solution to screen and select the most

important polymorphisms. In this thesis, I will describe the compilation of mtDNA sequences, the development of an existing extension-ligation assay format, and discuss the results from the first round of screening human mtDNA SNPs.

1.6 Howard Hughes Medical Institute Research Project

In April of 2006, Dr. David R. Walt was awarded a professorship from the Howard Hughes Medical Institute (HHMI). This non-profit organization was founded in 1953 with the purpose of funding the advancement of biomedical research and understanding, as Hughes once said, "the genesis of life itself" [37, 38]. The HHMI is currently the second largest private foundation and the largest devoted to biological and medical research with an endowment of nearly $15 billion [37]. By providing funds directly to individual investigators, rather than funding proposed projects, the HHMI encourages further exploration and modification of a project as it unfolds. This freedom to follow their "scientific" intuition allows these investigators to continually remain on the cutting edge of their respective fields.

The overview of the project proposed by Dr. Walt aims at integrating advanced research technology, used in upper-level research facilities, into the science education curriculum of undergraduate and, eventually, K-12 institutions. The project has four main aspects, most important of which is enhancing the enthusiasm for scientific research. Recipe-style laboratory procedures inherently decrease the sense of excitement of novel experimentation and potential for unexpected discoveries. Thus, participation in ongoing biochemical and bioanalytical research projects in the Walt Laboratory will be encouraged either through individual projects undertaken by graduate or undergraduate

students, or through participation in an integrated and customized science curriculum in an undergraduate laboratory.

In this multi-faceted, multi-year project, a DNA microarray will be developed to identify single nucleotide polymorphisms (SNPs) associated with human ancestry and migration within the mitochondrial genome. This genotyping experiment will be implemented in an Organic Chemistry laboratory and therefore has significant restrictions. Currently, the use of microarrays as an educational tool for undergraduate institutions is impracticable due to time and cost limitations. The Walt Laboratory is equipped with the necessary tools, instruments, and expertise to carry out these experiments. These resources clearly exceed the capabilities of most teaching laboratories. In order to bring this technology to a more cost-effective level, one component of the HHMI project will be to modify the chemistry accordingly.

Current avenues of investigation for the adaptation of DNA microarrays include the use of silica gel and polysaccharides as the solid support to which DNA will be attached. Surface chemistry is being developed to chemically modify these substrates for optimal oligonucleotide probe coupling [39-42]. The eventual goal is to attach hundreds of DNA probes to a surface and elucidate genetic information based on fluorescent or colorimetric signals produced from successful hybridizations of mtDNA [42]. Upcoming elements of this project include fine tuning of these protocols and the development of educational materials to aid students in understanding the chemical and biological processes upon which these experiments are based. In addition, students from the Computer Science department will be recruited to generate automated programs for efficient analysis of the data output from these microarrays.

1.7 DNA Microarrays

DNA microarrays are small, solid supports to which many oligonucleotide sequences can be immobilized at specific locations. The first use of DNA microarrays was the novel examination of multiple genes in parallel [43], and the technique has exponentially expanded to become the standard tool for genotyping and gene expression experiments in molecular biology and clinical diagnostics [44]. Oligonucleotides (oligos), or short segments of DNA, are designed to specifically hybridize to complementary sequences of target DNA being analyzed [45]. The four major applications for DNA microarrays within the biotechnology industry are gene expression, genotyping, infectious disease detection, and comparative genomic hybridization [46]. Following the completion of the Human Genome Project, SNP genotyping has become increasingly widespread. In order to meet the demands of the industry, several high-throughput SNP genotyping assays for comprehensive genome-wide genetic association studies have been developed [47, 48].

1.8 The Illumina GoldenGate™ Assay

The platform used for this project was the 16-sample, slide-based Sentrix® BeadChip from Illumina Inc [49]. By relying on the random self-assembly of quantitatively-pooled bead libraries into patterned microwell substrates, the Illumina genotyping products are able to provide one of the highest density arrays available. Each bead within this array contains hundreds of thousands of covalently-attached oligonucleotide probes whose position within the array can be decoded following a sequential hybridization-based procedure [47, 50]. The highly-multiplexed genotyping

format relies on an allele-specific extension, adapter ligation, and amplification assay based on the GoldenGate^{TM} protocol (see Figure 1) [47]. The extension-ligation assay relies on three oligos per SNP locus. Two of these oligos are specific to the allele site (ASOs) and the third, called the locus-specific oligo (LSO) binds 1-20 base pairs downstream of the target locus. This third oligo also contains an "address" sequence which binds to its complement on a particular bead type [47]. In the first step of the procedure, the genomic DNA is activated by denaturation and incubated with the assay oligonucleotides, hybridization buffer, and paramagnetic beads. Following this hybridization, a polymerase with a high specificity for a 3'-match extends the ASO. In order to get this extension, an exact 3' match must occur at the SNP locus. The gap is then sealed with DNA ligase to create a PCR template easily amplified using the universal primer regions. This template now includes the address sequence of the LSO and extended sequence of the ASO with the information regarding the genotype present at the SNP site. Three universal primers are used – one for each of the oligos. The primers that recognize the ASOs are labeled with fluorescent dyes, Cy3 and Cy5, with each color corresponding to one of the genotyping calls. Following PCR thermal-cycling, the single-stranded, dye-labeled DNA products are hybridized to their respective bead type address sequences. Illumina's BeadArray Reader then simultaneously scans the array at two different wavelengths, 550 and 630 nm, decodes the bead identity using map files, and outputs image files with raw fluorescence intensities [47]. Genotype calls for each SNP locus are then generated using BeadStudio.

Using single nucleotide polymorphisms (SNPs) chosen from a thorough literature review and resources from several online databases exclusively dedicated to the storage

and organization of mitochondrial DNA genomic sequences [29, 30], 1056 SNPs were chosen for this array and their primer sequences were generated using the Revised Cambridge Reference Sequence (rCRS) [33]. Because the identification of a single point mutation was desired, the GoldenGate™ assay provided the ideal platform.

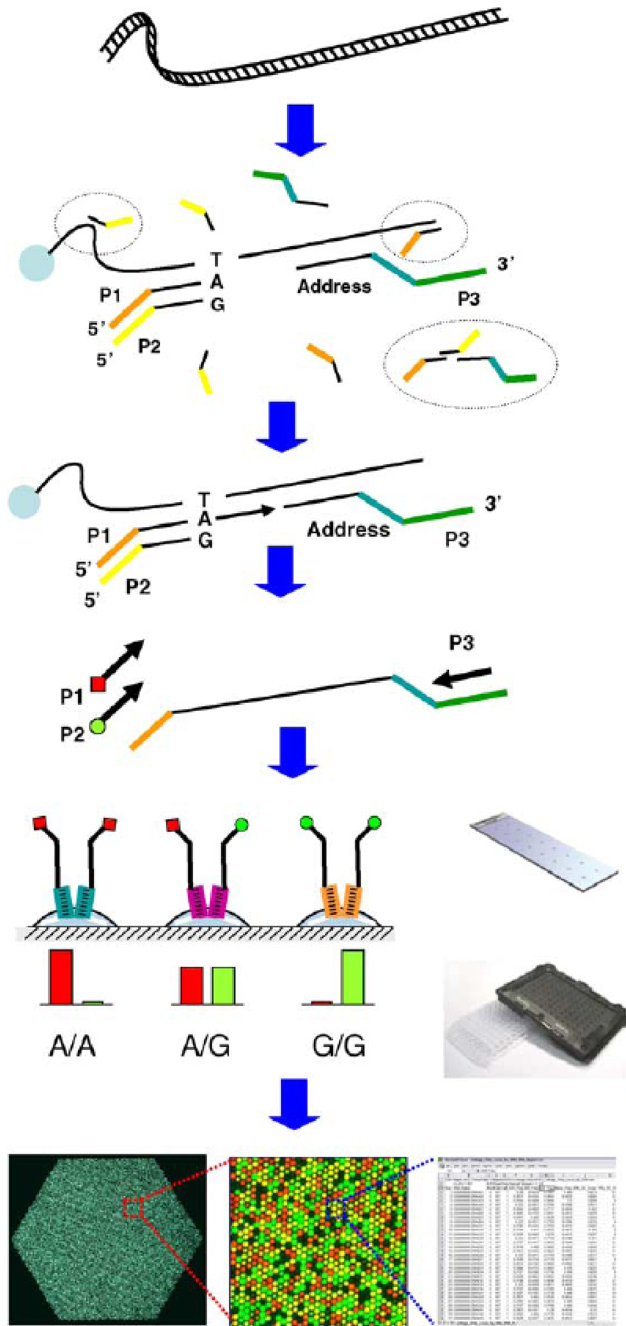

Figure 1 – Schematic Representation of Illumina GoldenGate™ Assay [47]

Three oligos are designed for each SNP locus investigated: two allele-specific (ASOs) and one locus-specific (LSO). Hybridization to the target sequence occurs with the LSO and complementary ASO. The ASO is extended with DNA polymerase and ligated to the LSO. The extension-ligation product undergoes PCR with fluorescent-labeled primers and the intensity of the product is measured to generate a SNP call at the examined locus.

Chapter 2: MATERIALS and METHODS

2.1 Materials

Tris-ethylenediaminetetraacetic acid (TE) buffer (pH 8.33) (Sigma, St. Louis, MO) was prepared by 1:100 dilution with 18 MΩ Millipore water from A10 Gradient System (Millipore, Billerica, MA) and filtered through vacuum filtration columns (Millipore, Billerica, MA). 95% ethanol was prepared by diluting 200 proof ethanol, molecular-biology grade (Sigma, St. Louis, MO) with Millipore water. The heat block used was the TruTemp™ Microheating System (SciGene, Sunnyvale, CA). Synthetic oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and diluted to 100 μM with Millipore water. For the PCR reactions, 2X Universal MasterMix (Promega, Madison, WI) were used and the samples run on the PTC-100™ Peltier Thermal Cycler (MJ Research, Inc, Waltham, MA). For the DNA purification, Oragene® Self-Collection Kits (DNA Genotek, Ottawa, Ontario) were used, and the DNA amplification was carried out using the Qiagen REPLI-g mtDNA Amplification Kit (Qiagen, Valencia, CA). For the gel electrophoresis, Invitrogen 2% agarose E-Gels (Invitrogen, Carlsbad, CA) were used with the TrackIt™ 100 bp DNA Ladder (Invitrogen, Carlsbad, CA). The DNA microarrays were processed using the Illumina protocol for Single Use DNA (SUD) (Illumina Inc., San Diego, CA) using the Sentrix® BeadChips (Illumina Inc., San Diego, CA). BeadStation®, hybridization ovens, BeadStudio® and ancillary equipment were obtained from Illumina Inc. (San Diego, CA). For heat-sealing of the reaction plates, Thermo-Seal Heat Sealing Foil Sheets (ABgene, Rochester, NY) were used with the Combi Thermo-Sealer (ABgene, Rochester,

NY). 0.1 N sodium hydroxide (Fisher Scientific, Pittsburgh, PA) and 2-propanol Ultra grade for molecular biology (Fluka, Switzerland) were used.

2.2 Illumina® Array Design

The Illumina® oligonucleotide pool was designed with primer sequences for 1056 single nucleotide polymorphisms (SNPs) of interest along the mitochondrial genome. Two main databases, the mtDB: Human Mitochondrial Genome Database [29] and the HmtDB: Human Mitochondrial Genomic Resource [30], were used to select these SNPs. Each sequence was derived from the Revised Cambridge Reference Sequence (rCRS) [33]. Forty bases on each side of the SNP were selected on the uploaded GenBank sequence (gi: 115315570) and transferred into the BioEdit® program. Each SNP was assigned an ID number based on its position along the rCRS.

After generation of the final array primers by Illumina, the sequences were scored for theoretical accuracy and specificity. Primer design was optimized using OligoDesigner® Software, which evaluated the sequences flanking each SNP based on repeated regions within the genome, palindromic sequences, GC and AT content and neighboring polymorphisms [47].

2.3 Sample Collection

Saliva samples were collected from 13 members of the Walt Laboratory in the Chemistry Department at Tufts University. The volunteers were asked to provide their estimated maternal background for as many past generations as possible.

2.4 DNA Extraction

Saliva samples were collected using the DNA Genotek Oragene™ DNA Self-Collection Kit. The DNA present in these saliva samples comes from buccal epithelial cells and white blood cells found in the mouth. DNA collection with this device is purportedly equal to or better than DNA collection from blood when used for downstream applications such as PCR and genotyping [51]. These kits are especially convenient because the Oragene™ solution stabilizes the DNA for long-term storage at room temperature, and allows for multiple freeze-thaw cycles without degradation.

Participants were instructed to rinse their mouths with water, wait one minute and then allow approximately two milliliters of saliva to flow into the collection vessel within 30 minutes. Once the appropriate amount of saliva had been collected, the vials were sealed to release the Oragene™ solution, approximately 2 mL, and mixed by inversion. The total volume of the solution was approximately 4 mL. The samples were processed using the protocol outlined by DNA Genotek, with several variations. In total, 15 samples and one negative control, with TE buffer only, were processed. The 15 samples included one from each participant collected on 3/31/07, one from RDW collected on 4/3/07, one from DR collected on 3/9/07 and one from SMW collected on 3/6/07. The samples were stored at room temperature.

The collection vials were incubated overnight at 50°C in a water bath. Sample RDW 4-3-07 was the only exception and was incubated at 50°C for two hours, however, this sample still fit the protocol, which calls for an incubation of at least one hour. Each saliva solution was then separated into four 1.5 mL microfuge tubes in 1 mL aliquots. 40µL of the Oragene™ Purifier reagent was added to each tube. The sample became

turbid, likely the result of the precipitation of "impurities and inhibitors", according to the DNA Genotek protocol. In the sample purification performed on 3/31/07, the purifier reagent was added and the samples were left to sit at room temperature for 30 minutes before proceeding to the next step. The samples were then incubated on ice for ten minutes. After ten minutes the samples were spun at 15,000 x g for ten minutes. Without disturbing the pellet, the supernatant was collected from the four microcentrifuge tubes and transferred into one 15 mL Falcon conical tube which produced, on average, 4 mL of solution. The 3-4 mL volume of supernatant for each sample was estimated and then diluted 1:1 with 95% ethanol. The samples were mixed gently by inversion. It was recommended that the solution stand for ten minutes at room temperature to allow full precipitation of the DNA, however, after a failed first attempt to collect DNA using these kits, the procedure was modified, and the samples were set overnight at room temperature. The samples were centrifuged at room temperature for ten minutes at 1,100 x g. Without disturbing the pellets, the supernatant was transferred to another 15 mL conical tube and the DNA pellet was re-suspended in 500 μL of TE buffer and stored at room temperature. This DNA purification method was selected based on the efficacy of the purification protocol, the ease of sample collection and the ample theoretical yield of genomic DNA (ca. 110 μg).

2.5 Mitochondrial DNA Amplification

A 10 μL aliquot of the template DNA rehydrated in TE buffer from the purification step was added to nuclease-free water to a final volume of 20 μL. REPLI-g mt Reaction Buffer (29 μL) was added and the solution was mixed by vortexing and was

then pulse-centrifuged. The samples were then incubated at 75°C for 5 min in a heat block and allowed to cool to room temperature (15-25°C) for approximately 10-15 min. During this step, the REPLI-g Midi DNA Polymerase was thawed on ice, and 1 μL was added to each reaction tube. The samples were mixed by vortexing and were then pulse-centrifuged. The samples were incubated at 33°C on a heat block overnight. Following this incubation, the polymerase was inactivated by heating the sample to 65°C for 3 min. The samples were stored at -20°C.

2.6 Mitochondrial DNA Identification

The mitochondrial DNA was identified by PCR and gel electrophoresis. Two primer pairs were selected to amplify 800-900 bp regions of the mitochondrial genome [52]. Each polymerase chain reaction (PCR) mixture was prepared using Promega 2X Universal Master Mix in a final reaction volume of 50 μL. Stock 100 μM solutions of the primer pairs were prepared by diluting the oligonucleotides in Millipore water. A 7.5 μL aliquot of the forward and reverse primers was added to the reaction for a final concentration of 15 μM. The PTC-100 Peltier thermal cycler was run with the following program: initial denaturation at 94°C for one minute, followed by 35 cycles of denaturation at 94°C for 30 seconds, primer annealing for 45 seconds at 50°C, and primer extension at 72°C for two minutes. Following these cycles, a final extension was carried out for five minutes at 72°C. Following PCR, 15 μL aliquots of each sample were loaded into the gels and run for 30 minutes at 60 V or until the yellow band of the DNA ladder was 1 cm from the bottom of the gel. PCR product length was compared to a 10 μL aliquot of Invitrogen TrackIt™ 100 bp DNA ladder.

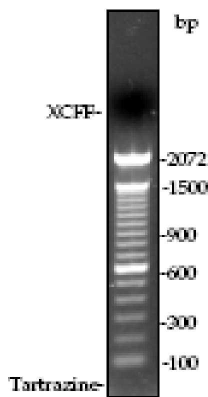

Figure 2: Invitrogen TrackIt™ 100 bp ladder

2.7 Illumina® Array Protocol

Two 16-array Sentrix® BeadChip silica slides were used in parallel for a total of 32 samples. Two samples from each individual were processed: one following the amplification step, and one unamplified sample. In addition to these 26 samples and two negative controls containing TE buffer only, four additional samples from different time points were processed as a comparison to fill all 32 available spots: RDW 4-3-07, DR 3-9-07, DR 3-9-07 amplified and SMW 3-6-07. The published protocol for this array was followed as indicated with one modification [53]: in the "Precip SUD" section for step 9, the inverted plate was tapped gently on the desk rather than centrifuged at 8 x g.

2.8 Statistical Analysis

Illumina BeadStudio™ software was used to analyze the data generated from the GoldenGate® Assay. This software compares the intensities of the two fluorescent channels from each sample and makes a "SNP Call" based on these results. For example, there are three choices for a diploid system: homozygous AA, heterozygous AB, and

homozygous BB. Polar coordinates are used to generate more accurate clustering analyses, because they display an easier manner of comparing sample data than Cartesian coordinates. Because the mitochondrial genome is haploid, the clustering had to be done manually, so the data were exported to Microsoft Excel for further analysis. The manual SNP calls were made based on their theta values and were assigned an A or B call, with θ less than 0.2 designated A and θ greater than 0.8 designated B. For each SNP, A and B represent different nucleotides depending on the reference base, the type of mutation, and which strand was used to generate the primer. Using SNPs with well-defined haplogroups (represented in Table 1 from HmtDB: Human Mitochondrial Database), the SNP calls for 461 haplogroup-specific sites were analyzed and haplogroup designations were assigned.

Table 1 – HmtDB Haplogroup SNPs

haplogroup_code	position	SNP_type	haplogroup_code	position	SNP_type
L0	1048	T	L0k	11296	T
L0	4312	T	L0k	11299	C
L0	6185	C	L0k	11653	G
L0	9755	A	L0k	13590	A
L0	11914	A	L0k	13819	C
L0	12007	A	L0k	13928	C
L0	3516	A	L0k	14020	C
L0	5442	C	L0k	14182	C
L0	9042	T	L0k	14371	C
L0	9347	G	L0k	14374	C
L0	10589	A	L0d	1438	G
L0	10664	T	L0d	4232	C
L0	10915	C	L0d	8152	A
L0	13276	G	L0d	8251	A
L0f	1719	A	L0d	12121	C
L0f	2789	T	L0d	15930	G
L0f	7148	C	L0d	498	T
L0f	10143	A	L0d	3756	G

L0f	11287	C	L0d	6815	C
L0f	13708	A	L0d	8113	A
L0f	14109	T	L0d	12720	G
L0f	15852	C	L0d	15466	C
L0a	5460	A	L0d	15941	C
L0a	8428	T	L1	14560	A
L0a	11176	A	L1	3666	A
L0a	5231	A	L1	7055	G
L0a	12720	G	L1	7389	C
L0a	14308	C	L1	13789	C
L0a	8566	G	L1	14178	C
L0k	4907	C	L1b	710	C
L0k	5460	A	L1b	1438	G
L0k	7257	G	L1b	2352	C
L0k	8994	A	L1b	3308	C
L0k	10920	T	L1b	3693	A
L0k	12070	A	L1b	15115	C
L0k	13020	C	L1b	709	A
L0k	850	C	L1b	1738	C
L0k	1243	C	L1b	5036	G
L0k	2836	T	L1b	5046	A
L0k	4542	T	L1b	5655	C
L0k	5811	G	L1b	6548	T
L0k	6938	T	L1b	6827	C
L0k	8911	C	L1b	6939	C
L0k	9136	G	L1b	7867	T
L0k	9755	A	L1b	8248	G
L0k	10499	G	L1b	12519	C
L0k	10876	G	L1b	13880	A
L0k	10939	T	L1b	14203	G
haplogroup_code	**position**	**SNP_type**	**haplogroup_code**	**position**	**SNP_type**
L1b	14769	G	L2b	8080	T
L1c	6071	C	L2b	8387	A
L1c	10586	A	L2b	12948	G
L1c	12810	G	L2b	14059	G
L1c	2395	G	L2c	680	C
L1c	5951	G	L2c	3200	A
L1c	8027	A	L2c	13958	C
L1c	9072	G	L2c	15849	T
L1c	13485	G	L2d	3963	T
L1c	14000	C	L0/L1/L2	3594	T
L1c	14911	T	L0/L1/L2	4104	G
L5	8152	A	L0/L1/L2	7256	T
L5	8155	A	L0/L1/L2	7521	A

L5	3423	C	L0/L1/L2	13650	T
L5	12432	T	L4	3918	A
L2	2416	C	L4	6260	A
L2	8206	A	L4	8104	C
L2	9221	G	L4	12609	C
L2	10115	C	L4	9855	G
L2	13590	A	L4	13470	G
L2a	2789	T	L/M	8701	G
L2a	7175	C	L/M	10398	G
L2a	7274	T	L/M	9540	C
L2a	7771	G	L/M	10873	C
L2a	11914	A	L/M	15301	A
L2a	13803	G	L3b	3450	T
L2a	14566	G	L3b	5773	A
L2a1	12693	G	L3b	6221	C
L2a1	15784	C	L3b	9449	T
L2a2	709	A	L3b	10086	G
L2a2	6752	G	L3b	13914	A
L2a2	9932	A	L3b	15311	G
L2a2	15939	T	L3b	15824	G
L2b/c	1442	A	L3b	15940	C
L2b/c	2332	T	L3d	5147	A
L2b/c	7624	A	L3d	7424	G
L2b/c	12236	A	L3d	8618	C
L2b/c	15110	A	L3d	13886	C
L2b/c	15217	A	L3d	14284	T
L2b	1706	T	L3f	1822	C
L2b	2358	G	L3f	3396	C
L2b	4158	G	L3f	4218	C
L2b	4370	C	L3f	5601	T
L2b	4767	G	L3f	7819	A
L2b	5027	T	L3f	8527	G
L2b	5331	A	L3f	8932	T
L2b	5814	C	L3f	9950	C
L2b	6713	T	L3f	11440	A
haplogroup_code	position	SNP_type	haplogroup_code	position	SNP_type
L3f	14769	G	C	11914	A
L3f	15514	C	C	13263	G
L3f	15940	C	C	14318	C
L3h	1719	A	Z	6752	G
L3h	4388	G	Z	9090	C
L3h	5300	T	Z	15784	C
L3h	7861	C	M8a	6179	A
L3h	9509	C	M8a	8684	T

<table>
<tr><td>L3h</td><td>9575</td><td>A</td><td>M8a</td><td>14470</td><td>C</td></tr>
<tr><td>L3h</td><td>11590</td><td>G</td><td>M7</td><td>6455</td><td>T</td></tr>
<tr><td>L3e</td><td>10819</td><td>G</td><td>M7</td><td>9824</td><td>C</td></tr>
<tr><td>L3e</td><td>2352</td><td>C</td><td>M7a</td><td>2626</td><td>C</td></tr>
<tr><td>L3e</td><td>14212</td><td>C</td><td>M7a</td><td>2772</td><td>T</td></tr>
<tr><td>L3i</td><td>10819</td><td>G</td><td>M7a</td><td>4386</td><td>C</td></tr>
<tr><td>L3i</td><td>7645</td><td>C</td><td>M7a</td><td>4958</td><td>C</td></tr>
<tr><td>L3i</td><td>10679</td><td>G</td><td>M7a</td><td>12771</td><td>A</td></tr>
<tr><td>L3i</td><td>11260</td><td>C</td><td>M7a</td><td>14364</td><td>A</td></tr>
<tr><td>L3i</td><td>13800</td><td>A</td><td>M7b/c</td><td>4071</td><td>T</td></tr>
<tr><td>L3i</td><td>13967</td><td>T</td><td>M7b</td><td>4048</td><td>A</td></tr>
<tr><td>L3i</td><td>14060</td><td>C</td><td>M7b</td><td>4164</td><td>G</td></tr>
<tr><td>L3i</td><td>15758</td><td>G</td><td>M7b</td><td>5351</td><td>G</td></tr>
<tr><td>M</td><td>489</td><td>C</td><td>M7b</td><td>5460</td><td>A</td></tr>
<tr><td>M</td><td>10400</td><td>T</td><td>M7b</td><td>6680</td><td>C</td></tr>
<tr><td>M</td><td>14783</td><td>C</td><td>M7b</td><td>7684</td><td>C</td></tr>
<tr><td>M</td><td>15043</td><td>A</td><td>M7b</td><td>7853</td><td>A</td></tr>
<tr><td>M1</td><td>6446</td><td>A</td><td>M7b</td><td>12405</td><td>T</td></tr>
<tr><td>M1</td><td>6680</td><td>C</td><td>M7b</td><td>12811</td><td>C</td></tr>
<tr><td>M1</td><td>12403</td><td>T</td><td>M7c</td><td>4071</td><td>T</td></tr>
<tr><td>M1</td><td>14110</td><td>C</td><td>M7c</td><td>4850</td><td>A</td></tr>
<tr><td>M2a</td><td>447</td><td>G</td><td>M7c</td><td>5442</td><td>C</td></tr>
<tr><td>M2a</td><td>1780</td><td>C</td><td>M7c</td><td>11665</td><td>T</td></tr>
<tr><td>M2a</td><td>8396</td><td>G</td><td>M7c</td><td>12091</td><td>C</td></tr>
<tr><td>M2a</td><td>8502</td><td>G</td><td>Q</td><td>4117</td><td>C</td></tr>
<tr><td>M3</td><td>4580</td><td>A</td><td>Q</td><td>5843</td><td>G</td></tr>
<tr><td>D</td><td>4883</td><td>T</td><td>Q</td><td>8790</td><td>A</td></tr>
<tr><td>D</td><td>5178</td><td>A</td><td>Q</td><td>12940</td><td>A</td></tr>
<tr><td>D4</td><td>8414</td><td>T</td><td>Q</td><td>13500</td><td>C</td></tr>
<tr><td>D4</td><td>14468</td><td>C</td><td>Q1</td><td>5460</td><td>A</td></tr>
<tr><td>D1</td><td>2092</td><td>T</td><td>Q1</td><td>8964</td><td>T</td></tr>
<tr><td>D5</td><td>1107</td><td>C</td><td>Q1</td><td>14025</td><td>C</td></tr>
<tr><td>D5</td><td>5301</td><td>G</td><td>Q2</td><td>2768</td><td>G</td></tr>
<tr><td>D5</td><td>10397</td><td>G</td><td>Q2</td><td>4335</td><td>T</td></tr>
<tr><td>M8</td><td>4715</td><td>G</td><td>M12</td><td>8793</td><td>C</td></tr>
<tr><td>M8</td><td>7196</td><td>A</td><td>M12</td><td>12771</td><td>A</td></tr>
<tr><td>M8</td><td>8584</td><td>A</td><td>M12</td><td>15098</td><td>G</td></tr>
<tr><td>M8</td><td>15487</td><td>G</td><td>M12</td><td>4508</td><td>T</td></tr>
<tr><td>C</td><td>3552</td><td>A</td><td>M12</td><td>8251</td><td>A</td></tr>
<tr><td>C</td><td>9545</td><td>G</td><td>M12</td><td>9156</td><td>G</td></tr>
<tr><th>haplogroup_code</th><th>position</th><th>SNP_type</th><th>haplogroup_code</th><th>position</th><th>SNP_type</th></tr>
<tr><td>N/M/L</td><td>12705</td><td>T</td><td>B5b</td><td>15927</td><td>A</td></tr>
<tr><td>A</td><td>663</td><td>G</td><td>B5b</td><td>15851</td><td>G</td></tr>
<tr><td>A</td><td>1736</td><td>G</td><td>R11</td><td>709</td><td>A</td></tr>
</table>

A	4248	C	R11	8277	C
A	4824	G	R11	8278+CC	ins
A	8794	T	R11	10031	C
S	8404	C	R11	10398	G
N9	5417	A	R11	11061	T
N9a	5231	A	R11	12950	G
N9a	12358	G	R11	13681	G
N9a	12372	A	F	3970	T
W	15884	C	F	6392	C
W	8251	A	F	10310	A
W	709	A	F	13928	C
W	1243	C	R12	9644	G
W	3505	G	R12	14305	A
W	5048	C	R12	15884	A
W	5460	G	P	10118	C
W	8994	A	P	15607	G
W	11674	T	P1a	3203	G
W	11947	G	P1a	3882	A
W	12414	C	P1a	4122	G
N1/X	1719	A	P1a	8859	T
N1	10238	C	P1a	14890	G
N1	12501	A	U	11467	G
X	6221	C	U	12308	G
X	6371	T	U	12372	A
X	13966	G	U1a	4991	A
X	14470	C	U1a	14364	A
B	8281-8289	del	U1	15148	A
B4a	5465	C	U1a	2218	T
B4a	9123	A	U1a	7581	C
B4a	10238	C	U1	12879	C
B4b	499	A	U1	13104	G
B4b	827	G	U1	14070	G
B4b	4820	A	U1	15954	C
B4b	13590	A	U	1811	G
B4b	15535	T	U2e	3116	T
B5b	709	A	U2e	3720	G
B5b	1598	A	U2e	5390	G
B5b	8584	A	U2e	5426	G
B5b	8829	T	U2e	6045	T
B5b	9950	C	U2e	6152	C
B5b	10398	G	U2e	10876	G
B5b	12361	G	U2e	11197	T
B5b	15223	T	U2e	11732	C
B5b	15508	T	U2e	13020	C

B5b	15662	G	U2e	13734	C
haplogroup_code	position	SNP_type	haplogroup_code	position	SNP_type
U2e	15907	G	H2	4769	A
U3a	4703	C	H3	6776	C
U3a	6518	T	H7	4793	G
U3a	9266	A	H5	456	T
U3a	10506	G	H5a	4336	C
U3a	13934	T	H6a	9380	A
U3	14139	G	H6a	3815	G
U3	15454	C	JT	4216	C
U3b	4188	G	JT	11251	G
U3b	4640	T	JT	15452	A
U3b	9656	C	J	489	C
U3b	13743	C	J	10398	G
U4	4646	C	J	12612	G
U4	6047	G	J	13708	A
U4	11332	T	J1	462	T
U4	14260	G	J2	7476	T
U4	15693	C	J2	15257	A
U6	3348	G	T	709	A
U7	980	C	T	1888	A
U7	5360	T	T	4917	G
U7	10084	C	T	8697	A
U7	10142	T	T	10463	C
U8/K	9698	C	T	13368	A
K	3480	G	T	14905	A
K	9055	A	T	15607	G
K	11299	C	T	15928	A
K	14167	T	T1	12633	A
K	14798	C	T2	11812	G
U5	3197	C	T2	14233	G
U5	9477	A			
U5	13617	C			
U5a	14793	G			
U5b	7768	G			
U5b	14182	C			
U9	6368	T			
HV	14766	C			
pre HV	11719	G			
(pre HV)1	2442	C			
(pre HV)1	3847	C			
(pre HV)1	13188	T			
H	2706	A			
H	7028	C			

HV1	8014	G
HV1	15218	G
V	4580	A
V	15904	T
H1	3010	A
H2	1438	A

Table 1: HmtDB Haplogroup-specific SNPs. The SNP position is in reference to the rCRS and the SNP type is the polymorphism that must be present for the indicated haplogroup designation.

Chapter 3: RESULTS and DISCUSSION

3.1 Illumina Array Design

The Illumina® oligonucleotide pool was designed with primer sequences for 1056 single nucleotide polymorphisms (SNPs) of interest along the mitochondrial genome. These SNPs were chosen based on their correlations with known haplogroups or certain populations. Two main databases, the mtDB: Human Mitochondrial Genome Database[29] and the HmtDB: Human Mitochondrial Genomic Resource[30], contain the majority of the available polymorphism genotypes associated with human migration and were used to select the SNPs that would provide the most useful genealogical information. Each sequence was derived from the Revised Cambridge Reference Sequence (rCRS) [33]. Table 2 shows the chosen SNPs presented with the reference base and mutation on either side of a backslash presented in brackets. Most mutations were transitions (A←→G, C←→T) however there were also several transversions (A←→T, A←→C, C←→G, G←→T).

Table 2: HHMI Illumina Microarray SNP Primer Sequences

SNP_ID#	Sequence
1438A	tatgaaacttaagggtcgaaggtggatttagcagtaaact[A/G]agagtagagtgcttagttgaacagggccctgaagcgcgta
14766C	atttcaactacaagaacaccaatgaccccaatacgcaaaa[C/T]taaccccctaataaaattaattaaccactcattcatcgac
3594C	ttctactatgaacccccctccccatacccaaccccctggt[C/T]aacctcaacctaggcctcctatttattctagccacctcta
4769A	ccaatactaccaatcaatactcatcattaataatcataat[A/G]gctatagcaataaaactaggaatagccccctttcacttct
7028C	cactagacatcgtactacacgacacgtactacgttgtagc[C/T]cacttccactatgtcctatcaataggagctgtatttgcca
A10005G	tgtctccatctattgatgagggtcttactcttttagtata[A/G]atagtaccgttaacttccaattaactagttttgacaacat
A10044G	aaatagtaccgttaacttccaattaactagttttgacaac[A/G]ttcaaaaaagagtaataaacttcgccttaattttaataat
A10086G	tcaaaaaagagtaataaacttcgccttaattttaataatc[A/G]acaccctcctagccttactactaataattattacattttg
A10097G	taataaacttcgccttaattttaataatcaacaccctcct[A/G]gccttactactaataattattacattttgactaccacaac
A10283G	tcttattatttgatctagaaattgccctccttttacccct[A/G]ccatgagccctacaaacaactaacctgccactaatagttat
A10289G	tatttgatctagaaattgccctccttttacccctaccatg[A/G]gccctacaaacaactaacctgccactaatagttatgtcat
A10358G	tagttatgtcatccctcttattaatcatcatcctagccct[A/G]agtctggcctatgagtgactacaaaaaggattagactgaa
A10398G	aagtctggcctatgagtgactacaaaaaggattagactga[A/G]ccgaattggtatatagtttaaacaaaacgaatgatttcga
A10411G	gagtgactacaaaaaggattagactgaaccgaattggtat[A/G]tagtttaaacaaaacgaatgatttcgactcattaaattat
A1041G	ccagttgacacaaaatagactacgaaagtggctttaacat[A/G]tctgaacacacaatagctaagacccaaactgggattagat

A10499G	atatttaccaaatgcccctcatttacataaatattatact[A/G]gcatttaccatctcacttctaggaatactagtatatcgct
A10506G	ccaaatgcccctcatttacataaatattatactagcattt[A/G]ccatctcacttctaggaatactagtatatcgctcacacct
A10550G	tctcacttctaggaatactagtatatcgctcacacctcat[A/G]tcctccctactatgcctagaaggaataatactatcgctgt
A10598G	tactatgcctagaaggaataatactatcgctgttcattat[A/G]gctactctcataaccctcaacacccactccctcttagcca
A10658G	acacccactccctcttagccaatattgtgcctattgccat[A/G]ctagtctttgccgcctgcgaagcagcggtgggcctagccc
A10700G	tagtctttgccgcctgcgaagcagcggtgggcctagccct[A/G]ctagtctcaatctccaacacatatggcctagactacgtac
A10750G	atctccaacacatatggcctagactacgtacataacctaa[A/G]cctactccaatgctaaaactaatcgtcccaacaattatat
A10792G	ctactccaatgctaaaactaatcgtcccaacaattatatt[A/G]ctaccactgacatgactttccaaaaaacacataatttgaa
A10876G	acacaaccacccacagcctaattattagcatcatccctct[A/G]ctatttttaaccaaatcaacaacaacctatttagctgtt
A10899G	attagcatcatccctctactatttttaaccaaatcaaca[A/G]caacctatttagctgttccccaaccttttcctccgacccc
A10978G	ccctaacaaccccctcctaatactaactacctgactcct[A/G]cccctcacaatcatggcaagccaacgccacttatccagtg
A11002G	taactacctgactcctacccctcacaatcatggcaagcca[A/G]cgccacttatccagtgaaccactatcacgaaaaaaactct
A11020G	ccctcacaatcatggcaagccaacgccacttatccagtga[A/G]ccactatcacgaaaaaaactctacctctctatactaatct
A11065G	tatcacgaaaaaaactctacctctctatactaatctccct[A/G]caaatctccttaattataacattcacagccacagaactaa
A1117G	tagataccccactatgcttagccctaaacctcaacagtta[A/G]atcaacaaaactgctcgccagaacactacgagccacagct
A11251G	tctacaccctagtaggctcccttcccctactcatcgcact[A/G]atttacactcacaacaccctaggctcactaaacattctac
A11293G	tttacactcacaacaccctaggctcactaaacattctact[A/G]ctcactctcactgcccaagaactatcaaactcctgagcca
A11404G	cttttatagtaaagatacctctttacggactccacttatg[A/G]ctccctaaagcccatgtcgaagcccccatcgctgggtcaa
A1148G	caacagttaaatcaacaaaactgctcgccagaacactacg[A/G]gccacagcttaaaactcaaaggacctggcggtgcttcata
A11560G	caaaacacatagcctaccccttccttgtactatccctatg[A/G]ggcataattataacaagctccatctgcctacgacaaacag
A11590G	tatccctatgaggcataattataacaagctccatctgcct[A/G]cgacaaacagacctaaaatcgctcattgcatactcttcaa
A11593T	ccctatgaggcataattataacaagctccatctgcctacg[A/T]caaacagacctaaaatcgctcattgcatactcttcaatca
A11653G	tcattgcatactcttcaatcagccacatagccctcgtagt[A/G]acagccattctcatccaaaccccctgaagcttcaccggcg
A11654G	tcattgcatactcttcaatcagccacatagccctcgtagta[A/G]cagccattctcatccaaaccccctgaagcttcaccggcg
A11800G	acgaacgcactcacagtcgcatcataatcctctctcaagg[A/G]cttcaaactctactcccactaatagctttttgatgacttc
A11812G	acagtcgcatcataatcctctctcaaggacttcaaactct[A/G]ctcccactaatagctttttgatgacttctagcaagcctcg
A11884G	caagcctcgctaacctcgccttacccccactattaacct[A/G]ctgggagaactctctgtgctagtaaccacgttctcctgat
A11928G	ggagaactctctgtgctagtaaccacgttctcctgatcaa[A/G]tatcactctcctacttacaggactcaacatactagtcaca
A11947G	taaccacgttctcctgatcaaatatcactctcctacttac[A/G]ggactcaacatactagtcacagccctatactccctctaca
A11950G	ccacgttctcctgatcaaatatcactctcctacttacagg[A/G]ctcaacatactagtcacagccctatactccctctacatat
A11959G	cctgatcaaatatcactctcctacttacaggactcaacat[A/G]ctagtcacagccctatactccctctacatatttaccacaa
A12030G	tttaccacaacacaatggggctcactcacccaccacatta[A/G]caacataaaaccctcattcacacgagaaaacaccctcatg
A12234G	ccccttatttaccgagaaagctcacaagaactgctaactc[A/G]tgcccccatgtctaacaacatggctttctcaacttttaaa
A12248G	agaaagctcacaagaactgctaactcatgcccccatgtct[A/G]acaacatggctttctcaacttttaaaggataacagctatc
A12308G	ttttaaaggataacagctatccattggtcttaggccccaa[A/G]aattttggtgcaactccaaataaaagtaataaccatgcac
A12366G	aaataaaagtaataaccatgcacactactataaccaccta[A/G]ccctgacttccctaattccccccatccttaccaccctcgt
A12507G	ttattatcagtctcttccccacaacaatattcatgtgcct[A/G]gaccaagaagttattatctcgaactgacactgagccacaa
A12570G	actgacactgagccacaacccaaacaacccagctctccct[A/G]agcttcaaactagactacttctccataatattcatccctg
A12579G	gagccacaacccaaacaacccagctctccctaagcttcaa[A/G]ctagactacttctccataatattcatccctgtagcattgt
A12612G	gcttcaaactagactacttctccataatattcatccctgt[A/G]gcattgttcgttacatggtccatcatagaattctcactgt
A12634G	cataatattcatccctgtagcattgttcgttacatggtcc[A/G]tcatagaattctcactgtgatatataaactcagacccaaac
A12642G	tcatccctgtagcattgttcgttacatggtccatcataga[A/G]ttctcactgtgatatataaactcagacccaaacattaatc
A12654G	cattgttcgttacatggtccatcatagaattctcactgtg[A/G]tatataaactcagacccaaacattaatcagttcttcaaat
A12672G	ccatcatagaattctcactgtgatatataaactcagaccc[A/G]aacattaatcagttcttcaaatatctactcatcttcctaa
A12693G	gatatataaactcagacccaaacattaatcagttcttcaa[A/G]tatctactcatcttcctaattaccatactaatcttagtta
A12720G	atcagttcttcaaatatctactcatcttcctaattaccat[A/G]ctaatcttagttaccgctaacaacctattccaactgttca
A12753G	ttaccatactaatcttagttaccgctaacaacctattcca[A/G]ctgttcatcggctgagagggcgtaggaattatatccttct

A12768G	tagttaccgctaacaacctattccaactgttcatcggctg[A/G]gagggcgtaggaattatatccttcttgctcatcagttgat
A12810G	agggcgtaggaattatatccttcttgctcatcagttgatg[A/G]tacgcccgagcagatgccaacacagcagccattcaagcaa
A12822G	ttatatccttcttgctcatcagttgatgatacgcccgagc[A/G]gatgccaacacagcagccattcaagcaatcctatacaacc
A12948G	tacactccaactcatgagacccacaacaaatagcccttct[A/G]aacgctaatccaagcctcaccccactactaggcctcctcc
A12978G	tagcccttctaaacgctaatccaagcctcaccccactact[A/G]ggcctcctcctagcagcagcaggcaaatcagcccaattag
A1299G	taccgccatcttcagcaaaccctgatgaaggctacaaagt[A/G]agcgcaagtacccacgtaaagacgttaggtcaaggtgtag
A13050G	cccaattaggtctccacccctgactcccctcagccataga[A/G]ggccccaccccagtctcagccctactccactcaagcacta
A13068G	cctgactcccctcagccatagaaggccccaccccagtctc[A/G]gccctactccactcaagcactatagttgtagcaggaatct
A13074G	tcccctcagccatagaaggccccaccccagtctcagccct[A/G]ctccactcaagcactatagttgtagcaggaatcttcttac
A1309G	ttcagcaaaccctgatgaaggctacaaagtaagcgcaagt[A/G]cccacgtaaagacgttaggtcaaggtgtagcccatgaggt
A13101C	cagtctcagccctactccactcaagcactatagttgtagc[A/C]ggaatcttcttactcatccgcttccacccccctagcagaaa
A13117G	tccactcaagcactatagttgtagcaggaatcttcttactc[A/G]tccgcttccacccccctagcagaaaatagcccactaatcc
A13152G	tactcatccgcttccacccccctagcagaaaatagcccact[A/G]atccaaactctaacactatgcttaggcgctatcaccactc
A13221G	ctatcaccactctgttcgcagcagtctgcgcccttacaca[A/G]aatgacatcaaaaaaatcgtagccttctccacttcaagtc
A13263G	atgacatcaaaaaaatcgtagccttctccacttcaagtca[A/G]ctaggactcataatagttacaatcggcatcaaccaaccac
A13269G	tcaaaaaaatcgtagccttctccacttcaagtcaactagg[A/G]ctcataatagttacaatcggcatcaaccaaccacacctag
A13276G	aatcgtagccttctccacttcaagtcaactaggactcata[A/G]tagttacaatcggcatcaaccaaccacacctagcattcct
A13278G	tcgtagccttctccacttcaagtcaactaggactcataat[A/G]gttacaatcggcatcaaccaaccacacctagcattcctgc
A13350G	cattcctgcacatctgtacccacgccttcttcaaagccat[A/G]ctatttatgtgctccgggtccatcatccacaaccttaaca
A13395G	ttatgtgctccgggtccatcatccacaaccttaacaatga[A/G]caagatattcgaaaaataggaggactactcaaaaccatac
A13422G	accttaacaatgaacaagatattcgaaaaataggaggact[A/G]ctcaaaaccatacctctcacttcaacctccctcaccattg
A13470G	ccatacctctcacttcaacctccctcaccattggcagcct[A/G]gcattagcaggaatacctttcctcacaggtttctactcca
A13485G	caacctccctcaccattggcagcctagcattagcaggaat[A/G]cctttcctcacaggtttctactccaaagaccacatcatcg
A13512G	cattagcaggaatacctttcctcacaggtttctactccaa[A/G]gaccacatcatcgaaaccgcaaacatatcatacacaaacg
A13629G	tgacaagcgcctatagcactcgaataattcttctcaccct[A/G]acaggtcaacctcgcttccccacccttactaacattaacga
A13637G	gcctatagcactcgaataattcttctcaccctaacaggtc[A/G]acctcgcttccccacccttactaacattaacgaaaataac
A13681G	tcgcttccccacccttactaacattaacgaaaataacccc[A/G]ccctactaaaccccattaaacgcctggcagccggaagcct
A13710G	aaaataaccccaccctactaaaccccattaaaggcctggc[A/G]gccggaagcctattcgcaggatttctcattactaacaaca
A13780G	tactaacaacatttcccccgcatcccccttccaaacaaca[A/G]tccccctctacctaaaactcacagccctcgctgtcacttt
A13803G	ccccccttccaaacaacaatccccctctacctaaaactcac[A/G]gccctcgctgtcactttcctaggacttctaacagccctag
A1382C	catgaggtggcaagaaatgggctacattttctaccccaga[A/C]aactacgatagcccttatgaaacttaagggtcgaaggtgg
A13857G	ctttcctaggacttctaacagccctagacctcaactacct[A/G]accaacaaacttaaaataaaatccccactatgcacatttt
A13887G	tcaactacctaaccaacaaacttaaaataaaatccccact[A/G]tgcacattttatttctccaacatactcggattctaccctа
A13933G	attttatttctccaacatactcggattctaccctagcatc[A/G]cacaccgcacaatcccctatctaggccttcttacgagcca
A13942G	ctccaacatactcggattctaccctagcatcacacaccgc[A/G]caatcccctatctaggccttcttacgagccaaaacctgcc
A13966G	tagcatcacacaccgcacaatcccctatctaggccttctt[A/G]cgagccaaaacctgcccctactcctcctagacctaacctg
A14007G	cgagccaaaacctgcccctactcctcctagacctaacctg[A/G]ctagaaaagctattacctaaaacaatttcacagcaccaaa
A14037G	acctaacctgactagaaaagctattacctaaaacaatttc[A/G]cagcaccaaatctccacctccatcatcacctcaacccaaa
A14053G	aaagctattacctaaaacaatttcacagcaccaaatctcc[A/G]cctccatcatcacctcaacccaaaaaggcataattaaact
A14059G	attacctaaaacaatttcacagcaccaaatctccacctcc[A/G]tcatcacctcaacccaaaaaggcataattaaactttactt
A14091G	ccacctccatcatcacctcaacccaaaaaggcataattaa[A/G]ctttacttcctctctttcttcttcccactcatcctaaccc
A14118G	aaggcataattaaactttacttcctctctttcttcttccc[A/G]ctcatcctaaccctactcctaatcacataacctattcccc
A14152G	cttcccactcatcctaaccctactcctaatcacataacct[A/G]ttcccccgagcaatctcaattacaatatatacaccaacaa
A14179G	aatcacataacctattcccccgagcaatctcaattacaat[A/G]tatacaccaacaaacaatgttcaaccagtaactactacta
A14185G	ataacctattcccccgagcaatctcaattacaatatatac[A/G]ccaacaaacaatggtcaaccagtaactactactaatcaac
A14203G	caatctcaattacaatatatacaccaacaaacaatgttca[A/G]ccagtaactactactaatcaacgcccataatcatacaaag
A14233G	acaatgttcaaccagtaactactactaatcaacgcccata[A/G]tcatacaaagcccccgcaccaataggatcctcccgaatca

A14370G	acaaccaccaccccatcatactctttcacccacaggacga[A/G]tcctacctccatcgctaaccccactaaaacactcaccaag
A1438G	tatgaaacttaagggtcgaaggtggatttagcagtaaact[A/G]agagtagagtgcttagttgaacagggccctgaagcgcgta
A14417G	ctccatcgctaaccccactaaaacactcaccaagacctca[A/G]cccctgacccccatgcctcaggatactcctcaatagccat
A14582G	cagaataataacacacccgaccacaccgctaacaatcaat[A/G]ctaaaccccataaataggagaaggcttagaagaaaaccc
A14587G	ataataacacacccgaccacaccgctaacaatcaatactaa[A/G]cccccataaataggagaaggcttagaagaaaaccccaca
A14687G	actcaacagaaacaaagcatacatcattattctcgcacgg[A/G]ctacaaccacgaccaatgatatgaaaaaccatcgttgtat
A14696G	aaacaaagcatacatcattattctcgcacggactacaacc[A/G]cgaccaatgatatgaaaaaccatcgttgtatttcaactac
A14750G	gaaaaaccatcgttgtatttcaactacaagaacaccaatg[A/G]ccccaatacgcaaaattaaccccctaataaaattaattaa
A14793G	ccaatacgcaaaactaaccccctaataaaattaattaacc[A/G]ctcattcatcgacctccccaccccatccaacatctccgca
A14890G	cactccttggcgcctgcctgatcctccaaatcaccacagg[A/G]ctattcctagccatgcactactcaccagacgcctcaaccg
A14978G	atcaatcgcccacatcactcgagacgtaaattatggctga[A/G]tcatccgctaccttcacgccaatggcgcctcaatattctt
A15208G	gaggggccacagtaattacaaacttactatccgccatccc[A/G]tacattgggacagacctagttcaatgaatctgaggaggcta
A15218G	agtaattacaaacttactatccgccatcccatacattggg[A/G]cagacctagttcaatgaatctgaggaggctactcagtaga
A1524G	cgcccgtcaccctcctcaagtatacttcaaaggacattta[A/G]ctaaaacccctacgcatttatatagaggagacaagtcgta
A15274G	gaatctgaggaggctactcagtagacagtcccaccctcac[A/G]cgattctttacctttcacttcatcttgcccttcattattg
A15311G	tcacacgattctttacctttcacttcatcttgcccttcatt[A/G]ttgcagccctagcaacactccacctcctattcttgcacga
A15326G	ctttcacttcatcttgcccttcattattgcagccctagca[A/G]cactccacctcctattcttgcacgaaacgggatcaaacaa
A1555G	ggacatttaactaaaacccctacgcatttatatagaggag[A/G]caagtcgtaacatggtaagtgtactggaaagtgcacttgg
A15562G	tagccaaccccttaaacacccctccccacatcaagcccgaatg[A/G]tatttcctattcgcctacacaattctccgatccgtccct a
A15607G	tcctattcgcctacacaattctccgatccgtccctaacaa[A/G]ctaggaggcgtccttgccctattactatccatcctcatcc
A15679G	tcctcatcctagcaataatccccatcctccatatatccaa[A/G]caacaaagcataatatttcgcccactaagccaatcacttt
A15758G	ttattgactcctagccgcagacctcctcattctaacctga[A/G]tcggaggacaaccagtaagctacccttttaccatcattgg
A15766G	tcctagccgcagacctcctcattctaacctgaatcggagg[A/G]caaccagtaagctacccttttaccatcattggacaagtag
A15775G	cagacctcctcattctaacctgaatcggaggacaaccagt[A/G]agctacccttttaccatcattggacaagtagcatccgtac
A15805G	gacaaccagtaagctacccttttaccatcattggacaagt[A/G]gcatccgtactatacttcacaacaatcctaatcctaatac
A15824G	ttttaccatcattggacaagtagcatccgtactatacttc[A/G]caacaatcctaatcctaataccaactatctccctaattga
A15907G	acaaaatactcaaatgggcctgtccttgtagtataaacta[A/G]tacaccagtcttgtaaaccggagatgaaaacctttttcca
A15924G	gcctgtccttgtagtataaactaatacaccagtcttgtaa[A/G]ccggagatgaaaacctttttccaaggacaaatcagagaaa
A16051G	aaactattctctgttctttcatggggaagcagatttgggt[A/G]ccacccaagtattgactcacccatcaacaaccgctatgta
A16066G	ctttcatggggaagcagatttgggtaccacccaagtattg[A/G]ctcacccatcaacaaccgctatgtatttcgtacattactg
A16146G	gccagccaccatgaatattgtacggtaccataaatacttg[A/G]ccacctgtagtacataaaaacccaatccacatcaaaaccc
A16182C	cttgaccacctgtagtacataaaaacccaatccacatcaa[A/C]acccccctccccatgcttacaagcaagtacagcaatcaacc
A16183C	ttgaccacctgtagtacataaaaacccaatccacatcaaa[A/C]ccccctccccatgcttacaagcaagtacagcaatcaaccc
A16203G	aaaacccaatccacatcaaaacccccctccccatgcttaca[A/G]gcaagtacagcaatcaaccctcaactatcacacatcaact
A16219G	caaaacccccctccccatgcttacaagcaagtacagcaatc[A/G]accctcaactatcacacatcaactgcaactccaaagccac
A16230G	ccccatgcttacaagcaagtacagcaatcaaccctcaact[A/G]tcacacatcaactgcaactccaaagccacccctcacccac
A1625G	agtgcacttggacgaaccagagtgtagcttaacacaaagc[A/G]cccaacttacacttaggagatttcaacttaacttgaccgc
A16300G	ccctcacccactaggataccaacaaacctacccaccctta[A/G]cagtacatagtacataaagccatttaccgtacatagcaca
A16335G	ccttaacagtacatagtacataaagccatttaccgtacat[A/G]gcacattacagtcaaatcccttctcgtccccatggatgac
A1736G	caccttactaccagacaaccttagccaaaccatttaccca[A/G]ataaagtataggcgatagaaattgaaacctggcgcaatag
A1811G	caatagatatagtaccgcaagggaaagatgaaaaattata[A/G]ccaagcataatatagcaaggactaacccctataccttctg
A1842G	aaaattataaccaagcataatatagcaaggactaacccct[A/G]taccttctgcataatgaattaactagaaataactttgcaa
A188G	atcctattatttatcgcacctacgttcaatattacaggcg[A/G]acatacttactaaagtgtgttaattaattaatgcttgtag
A189G	tcctattatttatcgcacctacgttcaatattacaggcga[A/G]catacttactaaagtgtgttaattaattaatgcttgtagg
A1900G	attaactagaaataactttgcaaggagagccaaagctaag[A/G]cccccgaaaccagacgagctacctaagaacagctaaaaga
A1943G	cccgaaaccagacgagctacctaagaacagctaaaagagc[A/G]cacccgtctatgtagcaaaatagtgggaagatttataggt
A1978G	agagcacacccgtctatgtagcaaaatagtgggaagattt[A/G]taggtagaggcgacaaacctaccgagcctggtgatagctg

A2045G	ctggtgatagctggttgtccaagatagaatcttagttca[A/G]ctttaaatttgcccacagaaccctctaaatccccttgtaa
A2109G	tctaaatccccttgtaaatttaactgttagtccaaagagg[A/G]acagctctttggacactaggaaaaaaccttgtagagagag
A2220G	aagcagccaccaattaagaaagcgttcaagctcaacaccc[A/G]ctacctaaaaaatcccaaacatataactgaactcctcaca
A2245G	tcaagctcaacacccactacctaaaaaatcccaaacatat[A/G]actgaactcctcacacccaattggaccaatctatcaccct
A2246G	caagctcaacacccactacctaaaaaatcccaaacatata[A/G]ctgaactcctcacacccaattggaccaatctatcaccctа
A227G	gaacatacttactaaagtgtgttaattaattaatgcttgt[A/G]ggacataataataacaattgaatgtctgcacagccacttt
A2294G	cctcacacccaattggaccaatctatcaccctatagaaga[A/G]ctaatgttagtataagtaacatgaaaacattctcctccgc
A2308G	ggaccaatctatcaccctatagaagaactaatgttagtat[A/G]agtaacatgaaaacattctcctccgcataagcctgcgtca
A2315G	tctatcaccctatagaagaactaatgttagtataagtaac[A/G]tgaaaacattctcctccgcataagcctgcgtcagattaaa
A2355G	atgaaaacattctcctccgcataagcctgcgtcagattaa[A/G]acactgaactgacaattaacagcccaatatctacaatcaa
A2358G	aaaacattctcctccgcataagcctgcgtcagattaaaac[A/G]ctgaactgacaattaacagcccaatatctacaatcaacca
A2581G	cgcctgcccagtgacacatgtttaacggccgcggtaccct[A/G]accgtgcaaaggtagcataatcacttgttccttaaatagg
A2755G	agacgagaagaccctatggagctttaatttattaatgcaa[A/G]cagtacctaacaaacccacaggtcctaaactaccaaacct
A2768G	ctatggagctttaatttattaatgcaaacagtacctaaca[A/G]acccacaggtcctaaactaccaaacctgcattaaaaattt
A2851G	gttggggcgacctcggagcagaacccaacctccgagcagt[A/G]catgctaagacttcaccagtcaaagcgaactactatactc
A297G	acagccactttccacacagacatcataacaaaaaatttcc[A/G]ccaaaccccccctcccccgcttctggccacagcacttaaa
A3105G	tctgagttcagaccggagtaatccaggtcggtttctatct[A/G]ccttcaaattcctccctgtacgaaaggacaagagaaataa
A3145G	accttcaaattcctccctgtacgaaaggacaagagaaata[A/G]ggcctacttcacaaagcgccttcccccgtaaatgatatca
A3203G	gccttcccccgtaaatgatatcatctcaacttagtattat[A/G]cccacacccacccaagaacagggtttgttaagatggcaga
A3221G	tatcatctcaacttagtattatacccacacccacccaaga[A/G]cagggtttgttaagatggcagagcccggtaatcgcataaa
A3348G	tacccatggccaacctcctactcctcattgtacccattct[A/G]atcgcaatggcattcctaatgcttaccgaacgaaaaattc
A3384G	ttctaatcgcaatggcattcctaatgcttaccgaacgaaa[A/G]attctaggctatatacaactacgcaaaggccccaacgtgg
A3397G	ggcattcctaatgcttaccgaacgaaaaattctaggctat[A/G]tacaactacgcaaaggccccaacgtggtaggcccctacgg
A3434G	tatatacaactacgcaaaggccccaacgtggtaggcccct[A/G]cgggctactacaacccttcgctgacgccataaaactcttc
A3447G	gcaaaggccccaacgtggtaggcccctacgggctactaca[A/G]cccttcgctgacgccataaaactcttcaccaaagagcccc
A3480G	tactacaacccttcgctgacgccataaaactcttcaccaa[A/G]gagcccctaaaacccgccacatctaccatcaccctctaca
A3505G	aaaactcttcaccaaagagcccctaaaacccgccacatct[A/G]ccatcaccctctacatcaccgccccgaccttagctctcac
A357G	cttctggccacagcacttaaacacatctctgccaaacccc[A/G]aaaacaaagaaccctaacaccagcctaaccagatttcaaa
A3606G	ccccccctccccatacccaaccccctggtcaacctcaacct[A/G]ggcctcctatttattctagccacctctagcctagccgttt
A3615G	catacccaaccccctggtcaacctcaacctaggcctcct[A/G]tttattctagccacctctagcctagccgtttactcaatcc
A3714G	caaactcaaactacgccctgatcggcgcactgcgagcagt[A/G]gcccaaacaatctcatatgaagtcaccctagccatcattc
A3720G	caaactacgccctgatcggcgcactgcgagcagtagccca[A/G]acaatctcatatgaagtcaccctagccatcattctactat
A3729G	ccctgatcggcgcactgcgagcagtagcccaaacaatctc[A/G]tatgaagtcaccctagccatcattctactatcaacattac
A373G	ttaaacacatctctgccaaaccccaaaaacaaagaaccct[A/G]acaccagcctaaccagatttcaaattttatcttttggcgg
A3756G	cccaaacaatctcatatgaagtcaccctagccatcattct[A/G]ctatcaacattactaataagtggctcctttaacctctcca
A3759G	aaacaatctcatatgaagtcaccctagccatcattctact[A/G]tcaacattactaataagtggctcctttaacctctccaccc
A3796G	actatcaacattactaataagtggctcctttaacctctcc[A/G]cccttatcacaacacaagaacacctctgattactcctgcc
A3816G	gtggctcctttaacctctccacccttatcacaacacaaga[A/G]cacctctgattactcctgccatcatgacccttggccataa
A3927G	accgaacccccttcgaccttgccgaaggggagtccgaact[A/G]gtctcaggcttcaacatcgaatacgccgcaggccccttcg
A3981G	acatcgaatacgccgcaggccccttcgccctattcttcat[A/G]gccgaatacacaaacattattataataaacaccctcacca
A4012G	attcttcatagccgaatacacaaacattattataataaac[A/G]ccctcaccactacaatcttcctaggaacaacatatgacgc
A4024G	cgaatacacaaacattattataataaacaccctcaccact[A/G]caatcttcctaggaacaacatatgacgcactctcccctga
A4122G	ttgtcaccaagaccctacttctaacctccctgttcttatg[A/G]attcgaacagcatacccccgattccgctacgaccaactca
A4129G	caagaccctacttctaacctccctgttcttatgaattcga[A/G]cagcatacccccgattccgctacgaccaactcatacacct
A4158G	tatgaattcgaacagcatacccccgattccgctacgacca[A/G]ctcatacacctcctatgaaaaaacttcctaccactcaccc
A4225G	cctaccactcaccctagcattacttatatgatatgtctcc[A/G]tacccattacaatctccagcattccccctcaaacctaaga
A4295G	caaacctaagaaatatgtctgataaaagagttactttgat[A/G]gagtaaataataggagcttaaaccccccttatttctaggac

A4310G	tgtctgataaaagagttactttgatagagtaaataatagg[A/G]gcttaaaccccttatttctaggactatgagaatcgaacc
A4388G	acccatccctgagaatccaaaattctccgtgccacctatc[A/G]caccccatcctaaagtaaggtcagctaaataagctatcgg
A4435G	atcctaaagtaaggtcagctaaataagctatcgggcccat[A/G]ccccgaaaatgttggttatacccttcccgtactaattaat
A4506G	actaattaatcccctggcccaacccgtcatctactctacc[A/G]tctttgcaggcacactcatcacagcgctaagctcgcactg
A4529T	ccgtcatctactctaccatctttgcaggcacactcatcac[A/T]gcgctaagctcgcactgatttttacctgagtaggcctag
A4562G	ctcatcacagcgctaagctcgcactgatttttacctgagt[A/G]ggcctagaaataaacatgctagcttttattccagttctaa
A4727G	ctatcctcttcaacaatatactctccggacaatgaaccat[A/G]accaatactaccaatcaatactcatcattaataatcataa
A4732G	ctcttcaacaatatactctccggacaatgaaccataacca[A/G]tactaccaatcaatactcatcattaataatcataatagct
A4767G	aaccaatactaccaatcaatactcatcattaataatcata[A/G]tagctatagcaataaaactaggaatagccccctttcactt
A4769G	ccaatactaccaatcaatactcatcattaataatcataat[A/G]gctatagcaataaaactaggaatagccccctttcacttct
A4811G	ctatagcaataaaactaggaatagccccctttcacttctg[A/G]gtcccagaggttacccaaggcacccctctgacatccggcc
A4917G	agcccccatctcaatcatataccaaatctctccctcacta[A/G]acgtaagccttctcctcactctctcaatcttatccatcat
A4976G	ctctctcaatcttatccatcatagcaggcagttgaggtgg[A/G]ttaaaccagacccagctacgcaaaatcttagcatactcct
A5036G	gcaaaatcttagcatactcctcaattacccacataggatg[A/G]ataatagcagttctaccgtacaaccctaacataaccattc
A508G	ccatactactaatctcatcaatacaacccccgcccatcct[A/G]cccagcacacacacaccgctgctaaccccataccccgaac
A5198G	tatctcgcacctgaaacaagctaacatgactaacaccctt[A/G]attccatccaccctcctctccctaggaggcctgcccccgc
A5240G	ttccatccaccctcctctccctaggaggcctgcccccgct[A/G]accggctttttgcccaaatgggccattatcgaagaattca
A5285G	gctttttgcccaaatgggccattatcgaagaattcacaaa[A/G]aacaatagcctcatcatccccaccatcatagccaccatca
A5319G	cacaaaaaacaatagcctcatcatccccaccatcatagcc[A/G]ccatcaccctccttaacctctacttctacctacgcctaat
A5390G	tacgcctaatctactccacctcaatcacactactccccat[A/G]tctaacaacgtaaaaataaaatgacagtttgaacatacaa
A5498G	tcgcccttaccacgctactcctacctatctccccttttat[A/G]ctaataatcttatagaaatttaggttaaatacagaccaag
A5656G	actgaacgcaaatcagccactttaattaagctaagccctt[A/G]ctagaccaatgggacttaaacccacaaacacttagttaac
A5836G	tgcaattcaatatgaaaatcacctcggagctggtaaaaag[A/G]ggcctaacccctgtctttagatttacagtccaatgcttca
A5951G	ccgaccgttgactattctctacaaaccacaaagacattgg[A/G]acactatacctattattcggcgcatgagctggagtcctag
A5984G	acattggaacactatacctattattcggcgcatgagctgg[A/G]gtcctaggcacagctctaagcctccttattcgagccgagc
A6047G	tccttattcgagccgagctgggccagccaggcaaccttct[A/G]ggtaacgaccacatctacaacgttatcgtcacagcccatg
A6146G	tcatagtaatacccatcataatcggaggctttggcaactg[A/G]ctagttcccctaataatcggtgcccccgatatggcgtttc
A6266G	ctctcctactcctgctcgcatctgctatagtggaggccgg[A/G]gcaggaacaggttgaacagtctaccctcccttagcaggga
A6281G	tcgcatctgctatagtggaggccggagcaggaacaggttg[A/G]acagtctaccctcccttagcagggaactactcccaccctg
A6575G	cagaccgcaacctcaacaccaccttcttcgaccccgccgg[A/G]ggaggagaccccattctataccaacacctattctgatttt
A6599G	tcttcgaccccgccggaggaggagaccccattctatacca[A/G]cacctattctgatttttcggtcaccctgaagtttatattc
A6629G	ttctataccaacacctattctgatttttcggtcaccctga[A/G]gtttatattcttatcctaccaggcttcggaataatctccca
A663G	acgggctcacatcaccccataaacaaataggtttggtcct[A/G]gcctttctattagctcttagtaagattacacatgcaagca
A6663G	ccctgaagtttatattcttatcctaccaggcttcggaata[A/G]tctcccatattgtaacttactactccggaaaaaaagaacc
A6989G	ccgtaggtggcctgactggcattgtattagcaaactcatc[A/G]ctagacatcgtactacacgacacgtactacgttgtagccc
A7076G	actatgtcctatcaataggagctgtatttgccatcatagg[A/G]ggcttcattcactgatttcccctattctcaggctacaccc
A7146G	aggctacaccctagaccaaacctacgccaaaatccatttc[A/G]ctatcatattcatcggcgtaaatctaactttcttcccaca
A7184G	tcactatcatattcatcggcgtaaatctaactttcttccc[A/G]caacactttctcggcctatccggaatgccccgacgttact
A7250G	tgccccgacgttactcggactaccccgatgcatacaccac[A/G]tgaaacatcctatcatctgtaggctcattcatttctctaa
A7271G	accccgatgcatacaccacatgaaacatcctatcatctgt[A/G]ggctcattcatttctctaacagcagtaatattaataattt
A735G	atgcaagcatccccgttccagtgagttcaccctctaaatc[A/G]ccacgatcaaaaggaacaagcatcaagcacgcagcaatgc
A7385G	aagtcctaatagtagaagaaccctccataaacctggagtg[A/G]ctatatggatgccccccaccctaccacacattcgaagaac
A7388G	tcctaatagtagaagaaccctccataaacctggagtgact[A/G]tatggatgccccccaccctaccacacattcgaagaacccg
A73G	cgggagctctccatgcatttggtattttcgtctgggggt[A/G]tgcacgcgatagcattgcgagacgctggagccggagcacc
A7403G	aaccctccataaacctggagtgactatatggatgcccccc[A/G]ccctaccacacattcgaagaacccgtatacataaaatcta
A7559G	aggtattagaaaaaccatttcataactttgtcaaagttaa[A/G]ttataggctaaatcctatatatcttaatggcacatgcagc
A7569G	aaaaccatttcataactttgtcaaagttaaattataggct[A/G]aatcctatatatcttaatggcacatgcagcgcaagtaggt

A7768G	caaaactaactaatactaacatctcagacgctcaggaaat[A/G]gaaaccgtctgaactatcctgcccgccatcatcctagtcc
A7771G	aactaactaatactaacatctcagacgctcaggaaataga[A/G]accgtctgaactatcctgcccgccatcatcctagtcctca
A7828G	tgcccgccatcatcctagtcctcatcgccctcccatccct[A/G]cgcatcctttacataacagacgaggtcaacgatccctccc
A8108G	ctcatgagctgtccccacattaggcttaaaaacagatgca[A/G]ttcccggacgtctaaaccaaaccactttcaccgctacacg
A813G	tgcagctcaaaacgcttagcctagccacacccccacggga[A/G]acagcagtgattaacctttagcaataaacgaaagtttaac
A8149G	ttcccggacgtctaaaccaaaccactttcaccgctacacg[A/G]ccgggggtatactacggtcaatgctctgaaatctgtggag
A8158G	gtctaaaccaaaccactttcaccgctacacgaccgggggt[A/G]tactacggtcaatgctctgaaatctgtggagcaaaccaca
A8188G	gaccgggggtatactacggtcaatgctctgaaatctgtgg[A/G]gcaaaccacagtttcatgcccatcgtcctagaattaattc
A8248G	ccatcgtcctagaattaattcccctaaaaatctttgaaat[A/G]gggcccgtatttaccctatagcacccctctacccctct
A8296G	tatttaccctatagcacccctctacccctctagagccc[A/G]ctgtaaagctaacttagcattaaccttttaagttaaagat
A8308G	agcacccctctacccctctagagcccactgtaaagcta[A/G]cttagcattaaccttttaagttaaagattaagagaaccaa
A8389G	cacctctttacagtgaaatgccccaactaaatactaccgt[A/G]tggcccaccataattacccccatactccttacactattcc
A8396G	ttacagtgaaatgccccaactaaatactaccgtatggccc[A/G]ccataattacccccatactccttacactattcctcatcac
A8440G	taattacccccatactccttacactattcctcatcaccca[A/G]ctaaaaatattaaacacaaactaccacctacctccctcac
A8453G	actccttacactattcctcatcacccaactaaaaatatta[A/G]acacaaactaccacctacctccctcaccaaagcccataaa
A8470G	tcatcacccaactaaaaatattaaacacaaactaccacct[A/G]cctccctcaccaaagcccataaaaataaaaaattataaca
A8502G	taccacctacctccctcaccaaagcccataaaaataaaaa[A/G]ttataacaaaccctgagaaccaaaatgaacgaaaatctgt
A8521G	caaagcccataaaaataaaaaattataacaaaccctgaga[A/G]ccaaaatgaacgaaaatctgttcgcttcattcattgcccc
A8566G	aatgaacgaaaatctgttcgcttcattcattgcccccaca[A/G]tcctaggcctacccgccgcagtactgatcattctatttcc
A856G	agcagtgattaacctttagcaataaacgaaagtttaacta[A/G]gctatactaaccccagggttggtcaatttcgtgccagcca
A866G	taacctttagcaataaacgaaagtttaactaagctatacta[A/G]ccccagggttggtcaatttcgtgccagccaccgcggtcac
A8701G	ccaacaatgactaatcaaactaacctcaaaacaaatgata[A/G]ccatacacaacactaaaggacgaacctgatctcttatact
A8710G	actaatcaaactaacctcaaaacaaatgataaccatacac[A/G]acactaaaggacgaacctgatctcttatactagtatcctt
A8805G	ccacaactaacctcctcggactcctgcctcactcatttac[A/G]ccaaccacccaactatctataaacctagccatggccatcc
A8812G	taacctcctcggactcctgcctcactcatttacaccaacc[A/G]cccaactatctataaacctagccatggccatccccttatg
A8836G	ctcatttacaccaaccacccaactatctataaacctagcc[A/G]tggccatccccttatgagcgggcacagtgattataggctt
A8860G	tatctataaacctagccatggccatccccttatgagcgggc[A/G]cagtgattataggctttcgctctaagattaaaaatgccct
A8869G	cctagccatggccatccccttatgagcgggcacagtgatt[A/G]taggctttcgctctaagattaaaaatgccctagcccactt
A8894T	gcgggcacagtgattataggctttcgctctaagattaaaa[A/T]tgccctagcccacttcttaccacaaggcacacctacaccc
A8973G	ccccttatccccatactagttattatcgaaaccatcagcct[A/G]ctcattcaaccaatagccctggccgtacgcctaaccgcta
A9006G	tcagcctactcattcaaccaatagccctggccgtacgcct[A/G]accgctaacattactgcaggccacctactcatgcacctaa
A9007G	tcagcctactcattcaaccaatagccctggccgtacgccta[A/G]ccgctaacattactgcaggccacctactcatgcacctaa
A9033G	tggccgtacgcctaaccgctaacattactgcaggccacct[A/G]ctcatgcacctaattggaagcgccaccctagcaatatcaa
A9052G	taacattactgcaggccacctactcatgcacctaattgga[A/G]gcgccaccctagcaatatcaaccattaaccttccctctac
A9066G	gccacctactcatgcacctaattggaagcgccaccctagc[A/G]atatcaaccattaaccttccctctacacttatcatcttca
A9072G	tactcatgcacctaattggaagcgccaccctagcaatatc[A/G]accattaaccttccctctacacttatcatcttcacaattc
A9109G	atcaaccattaaccttccctctacacttatcatcttcaca[A/G]ttctaattctactgactatcctagaaatcgctgtcgcctt
A9150G	ttctaattctactgactatcctagaaatcgctgtcgcctt[A/G]atccaagcctacgttttcacacttctagtaagcctctacc
A9156G	ttctactgactatcctagaaatcgctgtcgccttaatcca[A/G]gcctacgttttcacacttctagtaagcctctacctgcacg
A9242G	acataatgacccaccaatcacatgcctatcatatagtaaa[A/G]cccagcccatgacccctaacaggggccctctcagccctcc
A9254G	accaatcacatgcctatcatatagtaaaacccagcccatg[A/G]cccctaacaggggccctctcagccctcctaatgacctccg
A9299G	taacaggggccctctcagccctcctaatgacctccggcct[A/G]gccatgtgatttcacttccactccataacgctcctcatac
A9347G	gatttcacttccactccataacgctcctcatactaggcct[A/G]ctaaccaacacactaaccatataccaatgatggcgcgatg
A9355G	ttccactccataacgctcctcatactaggcctactaacca[A/G]cacactaaccatataccaatgatggcgcgatgtaacacga
A93G	ggtattttcgtctggggggtatgcacgcgatagcattgcg[A/G]gacgctggagccggagcaccctatgtcgcagtatctgtct
A9410G	accaatgatggcgcgatgtaacacgagaaagcacatacca[A/G]ggccaccacacaccacctgtccaaaaaggccttcgatacg
A942G	gtcacacgattaacccaagtcaatagaagccggcgtaaag[A/G]gtgttttagatcacccctccccaataaagctaaaactca

A9632G	tcccactcctaaacacatccgtattactcgcatcaggagt[A/G]tcaatcacctgagctcaccatagtctaatagaaaacaacc
A9670G	gtatcaatcacctgagctcaccatagtctaatagaaaaca[A/G]ccgaaaccaaataattcaagcactgcttattacaattta
A9852G	cttccacggacttcacgtcattattggctcaactttcctc[A/G]ctatctgcttcatccgccaactaatatttcactttacatc
C1009T	aaagctaaaactcacctgagttgtaaaaaactccagttga[C/T]acaaaatagactacgaaagtggctttaacatatctgaaca
C10104T	cttcgccttaattttaataatcaacaccctcctagcctta[C/T]tactaataattattacattttgactaccacaactcaacgg
C10142T	tactactaataattattacattttgactaccacaactcaa[C/T]ggctacatagaaaaatccaccccttacgagtgcggcttcg
C10192T	gaaaaatccaccccttacgagtgcggcttcgaccctatat[C/T]ccccgcccgcgtccctttctccataaaattcttcttagta
C10211T	agtgcggcttcgaccctatatcccccgcccgcgtcccttt[C/T]tccataaaattcttcttagtagctattaccttcttattat
C10214T	gcggcttcgaccctatatcccccgcccgcgtccctttctc[C/T]ataaaattcttcttagtagctattaccttcttattatttg
C10394T	ccctaagtctggcctatgagtgactacaaaaaggattaga[C/T]tgaaccgaattggtatatagtttaaacaaaacgaatgatt
C10400T	gtctggcctatgagtgactacaaaaaggattagactgaac[C/T]gaattggtatatagtttaaacaaaacgaatgatttcgact
C1048T	acacaaaatagactacgaaagtggctttaacatatctgaa[C/T]acacaatagctaagacccaaactgggattagataccccac
C10664T	actccctcttagccaatattgtgcctattgccatactagt[C/T]tttgccgcctgcgaagcagcggtgggcctagccctactag
C10670T	tcttagccaatattgtgcctattgccatactagtctttgc[C/T]gcctgcgaagcagcggtgggcctagccctactagtctcaa
C10733T	tagccctactagtctcaatctccaacacatatggcctaga[C/T]tacgtacataacctaaacctactccaatgctaaaactaat
C10822T	caattatattactaccactgacatgactttccaaaaaaca[C/T]ataatttgaatcaacacaaccacccacagcctaattatta
C10920T	ttttttaaccaaatcaacaacaacctatttagctgttccc[C/T]aaccttttcctccgacccccctaacaacccccctcctaata
C10955T	ttccccaaccttttcctccgaccccctaacaaccccctc[C/T]taatactaactacctgactcctacccctcacaatcatggc
C10986A	acccccctcctaatactaactacctgactcctacccctca[C/A]aatcatggcaagccaacgccacttatccagtgaaccacta
C11014T	tcctacccctcacaatcatggcaagccaacgccacttatc[C/T]agtgaaccactatcacgaaaaaaactctacctctctatac
C11047A	acttatccagtgaaccactatcacgaaaaaaactctacct[C/A]tctatactaatctccctacaaatctccttaattataacat
C11061T	ccactatcacgaaaaaaactctacctctctatactaatct[C/T]cctacaaatctccttaattataacattcacagccacagaa
C11071T	gaaaaaaactctacctctctatactaatctccctacaaat[C/T]tccttaattataacattcacagccacagaactaatcatat
C11143T	taatcatattttatatcttcttcgaaaccacacttatccc[C/T]accttggctatcatcacccgatgaggcaaccagccagaac
C11146T	tcatattttatatcttcttcgaaaccacacttatccccac[C/T]ttggctatcatcacccgatgaggcaaccagccagaacgcc
C11177T	tatccccaccttggctatcatcacccgatgaggcaaccag[C/T]cagaacgcctgaacgcaggcacatacttcctattctacac
C11197T	tcacccgatgaggcaaccagccagaacgcctgaacgcagg[C/T]acatacttcctattctacaccctagtaggctcccttcccc
C11242G	acttcctattctacaccctagtaggctcccttcccctact[C/G]atcgcactaatttacactcacaacaccctaggctcactaa
C11242T	acttcctattctacaccctagtaggctcccttcccctact[C/T]atcgcactaatttacactcacaacaccctaggctcactaa
C11332T	tactcactctcactgcccaagaactatcaaactcctgagc[C/T]aacaacttaatatgactagcttacacaatagcttttatag
C11348T	ccaagaactatcaaactcctgagccaataacttaatatga[C/T]tagcttacacaatagcttttatagtaaagatacctcttta
C11549T	caaccccctgacaaaacacatagcctaccccttccttgta[C/T]tatccctatgaggcataattataacaagctccatctgccta
C11623T	tctgcctacgacaaacagacctaaaatcgctcattgcata[C/T]tcttcaatcagccacatagccctcgtagtaacagccattc
C11674T	gccacatagccctcgtagtaacagccattctcatccaaac[C/T]ccctgaagcttcaccggcgcagtcattctcataatcgccc
C11767T	cctcattactattctgcctagcaaactcaaactacgaacg[C/T]actcacagtcgcatcataatcctctctcaaggacttcaaa
C1185T	acgagccacagcttaaaactcaaaggacctggcggtgctt[C/T]atatccctctagaggagcctgttctgtaatcgataaaccc
C11869A	tttgatgacttctagcaagcctcgctaacctcgccttacc[C/A]cccactattaacctactgggagaactctctgtgctagtaa
C11881T	tagcaagcctcgctaacctcgccttaccccccactattaa[C/T]ctactgggagaactctctgtgctagtaaccacgttctcct
C12049T	gctcactcacccaccacattaacaacataaaaccctcatt[C/T]acacgagaaaacaccctcatgttcatacacctatcccca
C12106T	tcatgttcatacacctatcccccattctcctcctatccct[C/T]aaccccgacatcattaccgggttttcctcttgtaaatata
C12153T	gacatcattaccgggttttcctcttgtaaatatagtttaa[C/T]caaaacatcagattgtgaatctgacaacagaggcttacgac
C12403T	cctaaccctgacttccctaattccccccatccttaccacc[C/T]tcgttaaccctaacaaaaaaaactcatacccccattatgt
C12432T	tccttaccaccctcgttaaccctaacaaaaaaaactcata[C/T]ccccattatgtaaaatccattgtcgcatccacctttatta
C12549T	accaagaagttattatctcgaactgacactgagccacaac[C/T]caaacaacccagctctccctaagcttcaaactagactact
C12557T	gttattatctcgaactgacactgagccacaacccaaacaa[C/T]ccagctctccctaagcttcaaactagactacttctccata
C12633A	ccataatattcatccctgtagcattgttcgttacatggtc[C/A]atcatagaattctcactgtgatatataaactcagacccaa
C12669T	ggtccatcatagaattctcactgtgatatataaactcaga[C/T]ccaaacattaatcagttcttcaaatatctactcatcttcc

C12705T	cagacccaaacattaatcagttcttcaaatatctactcat[C/T]ttcctaattaccatactaatcttagttaccgctaacaacc
C12816T	taggaattatatccttcttgctcatcagttgatgatacgc[C/T]cgagcagatgccaacacagcagccattcaagcaatcctat
C12858T	gagcagatgccaacacagcagccattcaagcaatcctata[C/T]aaccgtatcggcgatatcggtttcatcctcgccttagcat
C12876T	cagccattcaagcaatcctatacaaccgtatcggcgatat[C/T]ggtttcatcctcgccttagcatgatttatcctacactcca
C1290T	cagcctatataccgccatcttcagcaaaccctgatgaagg[C/T]tacaaagtaagcgcaagtacccacgtaaagacgttaggtc
C1310T	tcagcaaaccctgatgaaggctacaaagtaagcgcaagta[C/T]ccacgtaaagacgttaggtcaaggtgtagcccatgaggtg
C13110T	ccctactccactcaagcactatagttgtagcaggaatctt[C/T]ttactcatccgcttccacccccctagcagaaaatagcccac
C13116T	tccactcaagcactatagttgtagcaggaatcttcttact[C/T]atccgcttccacccccctagcagaaaatagcccactaatcc
C13129T	ctatagttgtagcaggaatcttcttactcatccgcttccac[C/T]ccctagcagaaaatagcccactaatccaaactctaacact
C13188T	cactaatccaaactctaacactatgcttaggcgctatcac[C/T]actctgttcgcagcagtctgcgcccttacacaaaatgaca
C13293T	cttcaagtcaactaggactcataatagttacaatcggcat[C/T]aaccaaccacacctagcattcctgcacatctgtacccacg
C13383T	aagccatactatttatgtgctccgggtccatcatccacaa[C/T]cttaacaatgaacaagatattcgaaaaataggaggactac
C1342T	cgcaagtacccacgtaaagacgttaggtcaaggtgtagcc[C/T]atgaggtggcaagaaatgggctacattttctaccccagaa
C13506T	gcctagcattagcaggaatacctttcctcacaggtttcta[C/T]tccaaagaccacatcatcgaaaccgcaaacatatcataca
C13749T	cagccggaagcctattcgcaggatttctcattactaacaa[C/T]atttcccccgcatccccccttccaaacaacaatccccctcta
C13758A	gcctattcgcaggatttctcattactaacaacatttcccc[C/A]gcatccccccttccaaacaacaatccccctctacctaaaac
C1375T	tgtagcccatgaggtggcaagaaatgggctacattttcta[C/T]cccagaaaactacgatagcccttatgaaacttaagggtcg
C13880A	ctagacctcaactacctaaccaacaaacttaaaataaaat[C/A]cccactatgcacattttatttctccaacatactcggattc
C13914A	taaaatccccactatgcacattttatttctccaacatact[C/A]ggattctaccctagcatcacacaccgcacaatcccctatc
C13924T	actatgcacattttatttctccaacatactcggattctac[C/T]ctagcatcacacaccgcacaatcccctatctaggccttct
C13934T	ttttatttctccaacatactcggattctaccctagcatca[C/T]acaccgcacaatcccctatctaggccttcttacgagccaa
C13967T	agcatcacacaccgcacaatcccctatctaggccttctta[C/T]gagccaaaacctgcccctactcctcctagacctaacctga
C14097T	ccatcatcacctcaacccaaaaaggcataattaaacttta[C/T]ttcctctctttcttcttcccactcatcctaaccctactcc
C14103A	tcacctcaacccaaaaaggcataattaaactttacttcct[C/T]tctttcttcttcccactcatcctaaccctactcctaatca
C14109T	caacccaaaaaggcataattaaactttacttcctctcttt[C/T]ttcttcccactcatcctaaccctactcctaatcacataac
C14125T	aattaaactttacttcctctctttcttcttcccactcatc[C/T]taaccctactcctaatcacataacctattcccccgagcaa
C14129T	aaactttacttcctctctttcttcttcccactcatcctaa[C/T]cctactcctaatcacataacctattcccccgagcaatctc
C14149T	cttcttcccactcatcctaaccctactcctaatcacataa[C/T]ctattcccccgagcaatctcaattacaatatatacaccaa
C14158T	cactcatcctaaccctactcctaatcacataacctattccc[C/T]cgagcaatctcaattacaatatatacaccaacaaacaatg
C14167T	aaccctactcctaatcacataacctattcccccgagcaat[C/T]tcaattacaatatatacaccaacaaacaatgttcaaccag
C14284T	cccccgcaccaataggatcctcccgaatcaaccctgaccc[C/T]tctccttcataaattattcagcttcctacactattaaagt
C14315T	ccctgacccctctccttcataaattattcagcttcctaca[C/T]tattaaagtttaccacaaccaccaccccatcatactcttt
C14338T	ttattcagcttcctacactattaaagtttaccacaaccac[C/T]accccatcatactctttcacccacaggacgaatcctacct
C14350T	ctacactattaaagtttaccacaaccaccaccccatcata[C/T]tctttcacccacagcaccaatcctacctccatcgctaacc
C14365T	ttaccacaaccaccaccccatcatactctttcacccacag[C/T]accaatcctacctccatcgctaaccccactaaaacactca
C14404T	ggacgaatcctacctccatcgctaaccccactaaaacact[C/T]accaagacctcaacccctgacccccatgcctcaggatact
C14407T	acgaatcctacctccatcgctaaccccactaaaacactcac[C/T]aagacctcaacccctgacccccatgcctcaggatactcct
C14460T	cctgacccccatgcctcaggatactcctcaatagccatcg[C/T]tgtagtatatccaaagacaaccatcattccccctaaataa
C14620T	atactaaacccccataaataggagaaggcttagaagaaaa[C/T]cccacaaaccccattactaaacccacactcaacagaaaca
C14668T	accccattactaaacccacactcaacagaaacaaagcata[C/T]atcattattctcgcacggactacaaccacgaccaatgata
C14766T	atttcaactacaagaacaccaatgaccccaatacgcaaaa[C/T]taacccccctaataaaattaattaaccactcattcatcgac
C14866T	tctccgcatgatgaaacttcggctcactccttggcgcctg[C/T]ctgatcctccaaatcaccacaggactattcctagccatgc
C14867T	tctccgcatgatgaaacttcggctcactccttggcgcctgc[C/T]tgatcctccaaatcaccacaggactattcctagccatgca
C14911T	tcctccaaatcaccacaggactattcctagccatgcacta[C/T]tcaccagacgcctcaaccgccttttcatcaatcgcccaca
C1495T	ttgaacagggccctgaagcgcgtacacaccgcccgtcacc[C/T]tcctcaagtatacttcaaaggacatttaactaaaacccct
C14974T	tttcatcaatcgcccacatcactcgagacgtaaattatgg[C/T]tgaatcatccgctaccttcacgccaatggcgcctcaatat
C14990T	catcactcgagacgtaaattatggctgaatcatccgctac[C/T]ttcacgccaatggcgcctcaatattctttatctgcctctt

C15028A	accttcacgccaatggcgcctcaatattctttatctgcct[C/A]ttcctacacatcgggcgaggcctatattacggatcatttc
C15049T	caatattctttatctgcctcttcctacacatcgggcgagg[C/T]ctatattacggatcatttctctactcagaaacctgaaac
C1508T	tgaagcgcgtacacaccgcccgtcaccctcctcaagtata[C/T]ttcaaaggacatttaactaaaacccctacgcatttatata
C150T	caccctatgtcgcagtatctgtctttgattcctgcctcat[C/T]ctattatttatcgcacctacgttcaatattacaggcgaac
C15136T	ttatcctcctgcttgcaactatagcaacagccttcatagg[C/T]tatgtcctcccgtgaggccaaatatcattctgaggggcca
C15292T	cagtagacagtcccaccctcacacgattctttacctttca[C/T]ttcatcttgcccttcattattgcagccctagcaacactcc
C15331A	acttcatcttgcccttcattattgcagccctagcaacact[C/A]cacctcctattcttgcacgaaacgggatcaaacaacccccc
C15385T	tgcacgaaacgggatcaaacaacccccctaggaatcacctc[C/T]cattccgataaaatcaccttccacccttactacacaatca
C15391T	aaacgggatcaaacaacccccctaggaatcacctcccattc[C/T]gataaaatcaccttccacccttactacacaatcaaagacg
C15452A	ttactacacaatcaaagacgccctcggcttacttctcttc[C/A]ttctctccttaatgacattaacactattctcaccagacc
C15550T	attataccctagccaacccttaaacacccctccccacat[C/T]aagcccgaatgatatttcctattcgcctacacaattctcc
C15664T	ccctattactatccatcctcatcctagcaataatccccat[C/T]ctccatatatccaaacaacaaagcataatatttcgcccac
C15789T	ctaacctgaatcggaggacaaccagtaagctaccctttta[C/T]catcattggacaagtagcatccgtactatacttcacaaca
C15849T	atccgtactatacttcacaacaatcctaatcctaataccaa[C/T]tatctccctaattgaaaacaaaatactcaaatgggcctgt
C15885T	ccaactatctccctaattgaaaacaaaatactcaaatggg[C/T]ctgtccttgtagtataaactaatacaccagtcttgtaaac
C15886T	ccaactatctccctaattgaaaacaaaatactcaaatgggc[C/T]tgtccttgtagtataaactaatacaccagtcttgtaaac
C15904T	aaaacaaaatactcaaatgggcctgtccttgtagtataaa[C/T]taatacaccagtcttgtaaaccggagatgaaaaccttttt
C15946T	aatacaccagtcttgtaaaccggagatgaaaaccttttttc[C/T]aaggacaaatcagagaaaaagtctttaactccaccatta
C15978T	ccttttccaaggacaaatcagagaaaaagtctttaactc[C/T]accattagcacccaaagctaagattctaatttaaactatt
C16108T	tcacccatcaacaaccgctatgtatttcgtacattactgc[C/T]agccaccatgaatattgtacggtaccataaatacttgac
C16147T	gccagccaccatgaatattgtacggtaccataaatacttga[C/T]cacctgtagtacataaaaacccaatccacatcaaaaccc
C16148T	cagccaccatgaatattgtacggtaccataaatacttgac[C/T]acctgtagtacataaaaacccaatccacatcaaaaccccc
C16150T	gccaccatgaatattgtacggtaccataaatacttgacca[C/T]ctgtagtacataaaaacccaatccacatcaaaacccccctc
C16173T	ccataaatacttgaccacctgtagtacataaaaacccaat[C/T]cacatcaaaacccccctccccatgcttacaagcaagtacag
C16174T	ccataaatacttgaccacctgtagtacataaaaacccaatc[C/T]acatcaaaacccccctccccatgcttacaagcaagtacag
C16186T	accacctgtagtacataaaaacccaatccacatcaaaacc[C/T]cctccccatgcttacaagcaagtacagcaatcaaccctca
C16187T	ccacctgtagtacataaaaacccaatccacatcaaaaccc[C/T]ctccccatgcttacaagcaagtacagcaatcaaccctcaa
C16188G	cacctgtagtacataaaaacccaatccacatcaaaacccc[C/G]tccccatgcttacaagcaagtacagcaatcaaccctcaac
C16192T	tgtagtacataaaaacccaatccacatcaaaacccccctcc[C/T]catgcttacaagcaagtacagcaatcaaccctcaactatc
C16218T	atcaaaacccccctccccatgcttacaagcaagtacagcaat[C/T]aaccctcaactatcacacatcaactgcaactccaaagcca
C16223T	acccccctccccatgcttacaagcaagtacagcaatcaacc[C/T]tcaactatcacacatcaactgcaactccaaagccacccct
C16256T	atcaaccctcaactatcacacatcaactgcaactccaaag[C/T]cacccctcacccactaggataccaacaaacctacccaccct
C16259T	caaccctcaactatcacacatcaactgcaactccaaagcca[C/T]ccctcacccactaggataccaacaaacctacccacccctt
C16261T	ccctcaactatcacacatcaactgcaactccaaagccacc[C/T]ctcacccactaggataccaacaaacctacccacccttaac
C16264T	tcaactatcacacatcaactgcaactccaaagccacccct[C/T]acccactaggataccaacaaacctacccacccttaacagt
C16268T	ctatcacacatcaactgcaactccaaagccacccctcacc[C/T]actaggataccaacaaacctacccacccttaacagtacat
C16270T	atcacacatcaactgcaactccaaagccacccctcaccca[C/T]taggataccaacaaacctacccacccttaacagtacatag
C16278T	tcaactgcaactccaaagccacccctcacccactaggata[C/T]caacaaacctacccacccttaacagtacatagtacataaa
C16292T	aaagccacccctcacccactaggataccaacaaacctacc[C/T]acccttaacagtacatagtacataaagccatttaccgtac
C16294T	agccacccctcacccactaggataccaacaaacctaccca[C/T]ccttaacagtacatagtacataaagccatttaccgtacat
C16354T	ataaagccatttaccgtacatagcacattacagtcaaatc[C/T]cttctcgtccccatggatgacccccctcagataggggtcc
C16366T	accgtacatagcacattacagtcaaatcccttctcgtccc[C/T]atggatgacccccctcagataggggtcccttgaccaccatc
C1703T	cgctctgagctaaacctagccccaaacccactccacctta[C/T]taccagacaaccttagccaaaccatttacccaaataaagt
C1706T	tctgagctaaacctagccccaaacccactccaccttacta[C/T]cagacaaccttagccaaaccatttacccaaataaagtata
C1715T	aacctagccccaaacccactccaccttactaccagacaac[C/T]ttagccaaaccatttacccaaataaagtataggcgataga
C1721T	gccccaaacccactccaccttactaccagacaaccttagc[C/T]aaaccatttacccaaataaagtataggcgatagaaattga
C2000T	aaaatagtgggaagatttataggtagaggcgacaaaccta[C/T]cgagcctggtgatagctggttgtccaagatagaatcttag

C2092T	atttgcccacagaaccctctaaatccccttgtaaatttaa[C/T]tgttagtccaaagaggaacagctctttggacactaggaaa
C2217T	taaaagcagccaccaattaagaaagcgttcaagctcaaca[C/T]ccactacctaaaaaatcccaaacatataactgaactcctc
C2218T	aaaagcagccaccaattaagaaagcgttcaagctcaacac[C/T]cactacctaaaaaatcccaaacatataactgaactcctca
C2225A	gccaccaattaagaaagcgttcaagctcaacacccactac[C/A]taaaaaatcccaaacatataactgaactcctcacacccaa
C2259T	cactacctaaaaaatcccaaacatataactgaactcctca[C/T]acccaattggaccaatctatcaccctatagaagaactaat
C2283T	ataactgaactcctcacacccaattggaccaatctatcac[C/T]ctatagaagaactaatgttagtataagtaacatgaaaaca
C2332T	gaactaatgttagtataagtaacatgaaaacattctcctc[C/T]gcataagcctgcgtcagattaaaacactgaactgacaatt
C2380T	cctgcgtcagattaaaacactgaactgacaattaacagcc[C/T]aatatctacaatcaaccaacaagtcattattaccctcact
C2639T	taatcacttgttccttaaatagggacctgtatgaatggct[C/T]cacgagggttcagctgtctcttacttttaaccagtgaaat
C2766T	ccctatggagctttaatttattaatgcaaacagtacctaa[C/T]aaacccacaggtcctaaactaccaaacctgcattaaaaat
C2772T	ggagctttaatttattaatgcaaacagtacctaacaaacc[C/T]acaggtcctaaactaccaaacctgcattaaaaatttcggt
C2789T	atgcaaacagtacctaacaaacccacaggtcctaaactac[C/T]aaacctgcattaaaaatttcggttggggcgacctcggagc
C2835T	tgcattaaaaatttcggttggggcgacctcggagcagaac[C/T]caacctccgagcagtacatgctaagacttcaccagtcaaa
C295T	gcacagccactttccacacagacatcataacaaaaaattt[C/T]caccaaaccccccctcccccgcttctggccacagcactta
C303T	actttccacacagacatcataacaaaaaatttccaccaaa[C/T]cccccctcccccgcttctggccacagcacttaaacacatc
C3116T	accggagtaatccaggtcggtttctatctacnttcaaatt[C/T]ctccctgtacgaaaggacaagagaaataaggcctacttca
C311T	cacagacatcataacaaaaaatttccaccaaacccccccct[C/T]ccccgcttctggccacagcacttaaacacatctctgccaa
C3204T	ccttcccccgtaaatgatatcatctcaacttagtattata[C/T]ccacacccacccaagaacagggtttgttaagatggcagag
C3206T	ttcccccgtaaatgatatcatctcaacttagtattatacc[C/T]acacccacccaagaacagggtttgttaagatggcagagcc
C3212T	cgtaaatgatatcatctcaacttagtattatacccacacc[C/T]acccaagaacagggtttgttaagatggcagagcccggtaa
C340T	aaaccccccctcccccgcttctggccacagcacttaaaca[C/T]atctctgccaaaccccaaaaacaaagaaccctaacaccag
C3450T	aaggccccaacgttgtaggcccctacgggctactacaacc[C/T]ttcgctgacgccataaaactcttcaccaaagagcccctaa
C3495A	ctgacgccataaaactcttcaccaaagagcccctaaaacc[C/A]gccacatctaccatcaccctctacatcaccgccccgacct
C3513T	tcaccaaagagcccctaaaacccgccacatctaccatcac[C/T]ctctacatcaccgccccgaccttagctctcaccatcgctc
C3516A	ccaaagagcccctaaaacccgccacatctaccatcaccct[C/A]tacatcaccgccccgaccttagctctcaccatcgctcttc
C3549T	tcaccctctacatcaccgccccgaccttagctctcaccat[C/T]gctcttctactatgaacccccctccccatacccaaccccc
C3594T	ttctactatgaacccccctccccatacccaaccccctggt[C/T]aacctcaacctaggcctcctatttattctagccacctcta
C3696T	tctgatcagggtgagcatcaaactcaaactacgccctgat[C/T]ggcgcactgcgagcagtagcccaaacaatctcatatgaag
C3738T	gcgcactgcgagcagtagcccaaacaatctcatatgaagt[C/T]accctagccatcattctactatcaacattactaataagtg
C3741T	cactgcgagcagtagcccaaacaatctcatatgaagtcac[C/T]ctagccatcattctactatcaacattactaataagtggct
C3817T	tggctcctttaacctctccacccttatcacaacacaagaa[C/T]acctctgattactcctgccatcatgacccttggccataat
C3975T	gcttcaacatcgaatacgccgcaggccccttcgccctatt[C/T]ttcatagccgaatacacaaacattattataataaacaccc
C3990T	acgccgcaggccccttcgccctattcttcatagccgaata[C/T]acaaacattattataataaacaccctcaccactacaatct
C3992T	gccgcaggccccttcgccctattcttcatagccgaataca[C/T]aaacattattataataaacaccctcaccactacaatcttc
C4025T	gaatacacaaacattattataataaacaccctcaccacta[C/T]aatcttcctaggaacaacatatgacgcactctcccctgaa
C4050T	acaccctcaccactacaatcttcctaggaacaacatatga[C/T]gcactctcccctgaactctacacaacatattttgtcacca
C4140T	ttctaacctccctgttcttatgaattcgaacagcataccc[C/T]cgattccgctacgaccaactcatacacctcctatgaaaaa
C4185T	tccgctacgaccaactcatacacctcctatgaaaaaactt[C/T]ctaccactcaccctagcattacttatatgatatgtctcca
C41T	gatcacaggtctatcaccctattaaccactcacgggagct[C/T]tccatgcatttggtattttcgtctggggggtatgcacgcg
C4312T	tctgataaaagagttactttgatagagtaaataataggag[C/T]ttaaacccccttatttctaggactatgagaatcgaaccca
C4335T	agagtaaataataggagcttaaacccccttatttctagga[C/T]tatgagaatcgaacccatccctgagaatccaaaattctcc
C447G	ttggcggtatgcacttttaacagtcaccccccaactaaca[C/G]attattttcccctcccactcccatactactaatctcatca
C4505T	tactaattaatcccctggcccaacccgtcatctactctac[C/T]atctttgcaggcacactcatcacagcgctaagctcgcact
C4508T	taattaatcccctggcccaacccgtcatctactctaccat[C/T]tttgcaggcacactcatcacagcgctaagctcgcactgat
C4538T	actctaccatctttgcaggcacactcatcacagcgctaag[C/T]tcgcactgatttttacctgagtaggcctagaaataaaca
C462T	tttaacagtcaccccccaactaacacattattttcccctc[C/T]cactcccatactactaatctcatcaatacaacccccgccc
C4670T	aagctgccatcaagtatttcctcacgcaagcaaccgcatc[C/T]ataatccttctaatagctatcctcttcaacaatatactct

C467T	cagtcaccccccaactaacacattattttcccctcccact[C/T]ccatactactaatctcatcaatacaaccccccgcccatcct
C4883T	catccggcctgcttcttctcacatgacaaaaactagcccc[C/T]atctcaatcatataccaaatctctccctcactaaacgtaa
C4960T	gtaagccttctcctcactctctcaatcttatccatcatag[C/T]aggcagttgaggtggattaaaccagacccagctacgcaaa
C497T	ccctcccactcccatactactaatctcatcaatacaaccc[C/T]cgcccatcctacccagcacacacacaccgctgctaacccc
C498T	ccctcccactcccatactactaatctcatcaatacaacccc[C/T]gcccatcctacccagcacacacacaccgctgctaacccca
C5027T	cccagctacgcaaaatcttagcatactcctcaattaccca[C/T]ataggatgaataatagcagttctaccgtacaaccctaaca
C5178A	cagcaccacgaccctactactatctcgcacctgaaacaag[C/A]taacatgactaacacccttaattccatccaccctcctctc
C5178T	cagcaccacgaccctactactatctcgcacctgaaacaag[C/T]taacatgactaacacccttaattccatccaccctcctctc
C5263T	ggaggcctgcccccgctaaccggctttttgcccaaatggg[C/T]cattatcgaagaattcacaaaaaacaatagcctcatcatc
C5300T	gggccattatcgaagaattcacaaaaaacaatagcctcat[C/T]atccccaccatcatagccaccatcaccctccttaacctct
C5331A	tagcctcatcatccccaccatcatagccaccatcaccctc[C/A]ttaacctctacttctacctacgcctaatctactccacctc
C5348T	ccatcatagccaccatcaccctccttaacctctacttcta[C/T]ctacgcctaatctactccacctcaatcacactactcccca
C5360T	ccatcaccctccttaacctctacttctacctacgcctaat[C/T]tactccacctcaatcacactactccccatatctaacaacg
C5375T	acctctacttctacctacgcctaatctactccacctcaat[C/T]acactactccccatatctaacaacgtaaaaataaaatgac
C5478T	cccattcctccccacactcatcgcccttaccacgctactc[C/T]tacctatctccccttttatactaataatcttatagaaatt
C5603T	gcaatacttaatttctgtaacagctaaggactgcaaaacc[C/T]cactctgcatcaactgaacgcaaatcagccactttaatta
C5633T	ctgcaaaaccccactctgcatcaactgaacgcaaatcagc[C/T]actttaattaagctaagcccttactagaccaatgggactt
C5911T	gcttcactcagccattttacctcacccccactgatgttcg[C/T]cgaccgttgactattctctacaaaccacaaagacattgga
C597T	ccccaaagacaccccccacagtttatgtagcttacctcct[C/T]aaagcaatacactgaaaatgtttagacgggctcacatcac
C6045T	cctccttattcgagccgagctgggccagccaggcaacctt[C/T]taggtaacgaccacatctacaacgttatcgtcacagccca
C6077T	gcaaccttctaggtaacgaccacatctacaacgttatcgt[C/T]acagcccatgcatttgtaataatcttcttcatagtaatac
C6104T	acaacgttatcgtcacagcccatgcatttgtaataatctt[C/T]ttcatagtaatacccatcataatcggaggctttggcaact
C61T	attaaccactcacgggagctctccatgcatttggtatttt[C/T]gtctgggggtatgcacgcgatagcattgcgagacgctgg
C6371T	tagacctaaccatcttctccttacacctagcaggtgtctc[C/T]tctatcttaggggccatcaatttcatcacaacaattatca
C6464A	ctgccataacccaataccaaacgcccctcttcgtctgatc[C/A]gtcctaatcacagcagtcctacttctcctatctctcccag
C6473T	cccaataccaaacgcccctcttcgtctgatccgtcctaat[C/T]acagcagtcctacttctcctatctctcccagtcctagctg
C6489A	cctcttcgtctgatccgtcctaatcacagcagtcctactt[C/A]tcctatctctcccagtcctagctgctggcatcactatact
C64T	aaccactcacgggagctctccatgcatttggtattttcgt[C/T]tgggggtatgcacgcgatagcattgcgagacgctggagc
C6518T	cagtcctacttctcctatctctcccagtcctagctgctgg[C/T]atcactatactactaacagaccgcaacctcaacaccacct
C6528T	tctcctatctctcccagtcctagctgctggcatcactata[C/T]tactaacagaccgcaacctcaacaccaccttcttcgaccc
C6531T	cctatctctcccagtcctagctgctggcatcactatacta[C/T]taacagaccgcaacctcaacaccaccttcttcgaccccgc
C6548T	tagctgctggcatcactatactactaacagaccgcaacct[C/T]aacaccaccttcttcgaccccgccggaggaggagacccca
C6554T	ctggcatcactatactactaacagaccgcaacctcaacac[C/T]accttcttcgaccccgccggaggaggagaccccattctat
C6587T	tcaacaccaccttcttcgaccccgccggaggaggagaccc[C/T]attctataccaacacctattctgatttttcggtcaccctg
C6713T	ttgtaacttactactccggaaaaaaagaaccatttggata[C/T]ataggtatggtctgagctatgatatcaattggcttcctag
C6869T	ccataatcatcgctatccccaccggcgtcaaagtatttag[C/T]tgactcgccacactccacggaagcaatatgaaatgatctg
C6960T	ctgagccctaggattcatctttcttttcaccgtaggtggc[C/T]tgactggcattgtattagcaaactcatcactagacatcgt
C7109T	tcataggaggcttcattcactgatttcccctattctcagg[C/T]tacaccctagaccaaacctacgccaaaatccatttcacta
C7112T	taggaggcttcattcactgatttcccctattctcaggcta[C/T]accctagaccaaacctacgccaaaatccatttcactatca
C7274T	ccgatgcatacaccacatgaaacatcctatcatctgtagg[C/T]tcattcatttctctaacagcagtaatattaataattttca
C7476T	aaaatctagacaaaaaaggaaggaatcgaaccccccaaag[C/T]tggtttcaagccaaccccatggcctccatgactttttcaa
C7693T	ttcatgatcacgccctcataatcattttccttatctgctt[C/T]ctagtcctgtatgcccttttcctaacactcacaacaaaac
C7867T	tacgcatcctttacataacagacgaggtcaacgatccctc[C/T]cttaccatcaaatcaattggccaccaatggtactgaacct
C7909T	ttaccatcaaatcaattggccaccaatggtactgaaccta[C/T]gagtacaccgactacggcggactaatcttcaactcctaca
C793T	aagcatcaagcacgcagcaatgcagctcaaaacgcttagc[C/T]tagccacacccccacgggaaacagcagtgattaaccttta
C7948T	acgagtacaccgactacggcggactaatcttcaactccta[C/T]atacttcccccattattcctagaaccaggcgacctgcgac
C8080T	taataattacatcacaagacgtcttgcactcatgagctgt[C/T]cccacattaggcttaaaaacagatgcaattcccggacgtc

C8393T	tctttacagtgaaatgccccaactaaatactaccgtatgg[C/T]ccaccataattacccccatactccttacactattcctcat
C8414T	actaaatactaccgtatggcccaccataattacccccata[C/T]tccttacactattcctcatcacccaactaaaaatattaaa
C8428T	tatggcccaccataattacccccatactccttacactatt[C/T]ctcatcacccaactaaaaatattaaacacaaactaccacc
C8468T	cctcatcacccaactaaaaatattaaacacaaactaccac[C/T]tacctccctcaccaaagcccataaaaataaaaaattataa
C8655T	tgatccccacctccaaatatctcatcaacaaccgactaat[C/T]accacccaacaatgactaatcaaactaacctcaaaacaaa
C8676T	tcatcaacaaccgactaatcaccacccaacaatgactaat[C/T]aaactaacctcaaaacaaatgataaccatacacaacacta
C870T	tttagcaataaacgaaagtttaactaagctatactaaccc[C/T]agggttggtcaatttcgtgccagccaccgcggtcacacga
C8794T	catttttattgccacaactaacctcctcggactcctgcct[C/T]actcatttacaccaaccacccaactatctataaacctagc
C8859T	tatctataaacctagccatggccatccccttatgagcggg[C/T]acagtgattataggctttcgctctaagattaaaaatgccc
C9042T	gcctaaccgctaacattactgcaggccacctactcatgca[C/T]ctaattggaagcgccaccctagcaatatcaaccattaacc
C9201T	acgttttcacacttctagtaagcctctacctgcacgacaa[C/T]acataatgacccaccaatcacatgcctatcatatagtaaa
C9272T	atatagtaaaacccagcccatgacccctaacaggggccct[C/T]tcagccctcctaatgacctccggcctagccatgtgatttc
C9365T	taacgctcctcatactaggcctactaaccaacacactaac[C/T]atataccaatgatggcgcgatgtaacacgagaaagcacat
C9431T	cacgagaaagcacataccaaggccaccacacaccacctgt[C/T]caaaaaggccttcgatacgggataatcctatttattacct
C9536T	ttttctgagccttttaccactccagcctagcccctacccc[C/T]caattaggagggcactggcccccgaacaggcatcaccccgc
C9818T	gcatctacggctcaacatttttgtagccacaggcttcca[C/T]ggacttcacgtcattattggctcaactttcctcactatct
C9822A	ctacggctcaacatttttgtagccacaggcttccacgga[C/A]ttcacgtcattattggctcaactttcctcactatctgctt
C9923T	actttacatccaaacatcactttggcttcgaagccgccgc[C/T]tgatactggcattttgtagatgtggtttgactatttctgt
G10143A	tactactaataattattacatttgactaccacaactcaac[G/A]gctacatagaaaaatccaccccttacgagtgcggcttcga
G10172A	cacaactcaacggctacatagaaaaatccaccccttacga[G/A]tgcggcttcgaccctatatcccccgcccgcgtccctttct
G10427A	gattagactgaaccgaattggtatatagtttaaacaaaac[G/A]aatgatttcgactcattaaattatgataatcatatttacc
G10688A	ctattgccatactagtctttgccgcctgcgaagcagcggt[G/A]ggcctagccctactagtctcaatctccaacacatatggcc
G10914A	ctactatttttaaccaaatcaacaacaacctatttagct[G/A]ttccccaaccttttcctccgaccccctaacaaccccctc
G11149A	atattttatatcttcttcgaaaccacacttatccccacctt[G/A]gctatcatcacccgatgaggcaaccagccagaacgcctga
G11150A	atattttatatcttcttcgaaaccacacttatccccaccttg[G/A]ctatcatcacccgatgaggcaaccagccagaacgcctga
G11440A	tatgactccctaaagcccatgtcgaagcccccatcgctgg[G/A]tcaatagtacttgccgcagtactcttaaaactaggcggct
G11611A	taacaagctccatctgcctacgacaaacagacctaaaatc[G/A]ctcattgcatactcttcaatcagccacatagccctcgtag
G11696A	agccattctcatccaaaccccctgaagcttcaccggcgca[G/A]tcattctcataatcgcccacgggcttacatcctcattact
G12007A	cagccctatactccctctacatatttaccacaacacaatg[G/A]ggctcactcacccaccacattaacaacataaaaccctcat
G1211A	acctggcggtgcttcatatccctctagaggagcctgttct[G/A]taatcgataaaccccgatcaacctcaccacctcttgctca
G12127A	cccattctcctcctatccctcaaccccgacatcattaccgg[G/A]ttttcctcttgtaaatatagtttaaccaaaacatcagatt
G12127A -2	ccattctcctcctatccctcaaccccgacatcattaccgg[G/A]ttttcctcttgtaaatatagtttaaccaaaacatcagatt
G12236A	ccttatttaccgagaaagctcacaagaactgctaactcat[G/A]cccccatgtctaacaacatggctttctcaacttttaaagg
G12372A	aagtaataaccatgcacactactataaccaccctaaccct[G/A]acttccctaattccccccatccttaccaccctcgttaacc
G12501A	ccacctttattatcagtctcttccccacaacaatattcat[G/A]tgcctagaccaagaagttattatctcgaactgacactgag
G12561A	ttatctcgaactgacactgagccacaacccaaacaaccca[G/A]ctctccctaagcttcaaactagactacttctccataatat
G12613A	cttcaaactagactacttctccataatattcatccctgta[G/A]cattgttcgttacatggtccatcatagaattctcactgtg
G12618A	aactagactacttctccataatattcatccctgtagcatt[G/A]ttcgttacatggtccatcatagaattctcactgtgatata
G12795A	tgttcatcggctgagagggcgtaggaattatatccttctt[G/A]ctcatcagttgatgatacgcccgagcagatgccaacacag
G1282A	ctcttgctcagcctatataccgccatcttcagcaaaccct[G/A]atgaaggctacaaagtaagcgcaagtacccacgtaaagac
G13135A	tgtagcaggaatcttcttactcatccgcttccacccccta[G/A]cagaaaatagcccactaatccaaactctaacactatgctt
G13359A	acatctgtacccacgccttcttcaaagccatactatttat[G/A]tgctccgggtccatcatccacaaccttaacaatgaacaag
G13368A	cccacgccttcttcaaagccatactatttatgtgctccgg[G/A]tccatcatccacaaccttaacaatgaacaagatattcgaa
G13477A	tctcacttcaacctccctcaccattggcagcctagcatta[G/A]caggaataccttttcctcacaggtttctactccaaagacca
G13708A	cgaaaataaccccaccctactaaaccccattaaacgcctg[G/A]cagccggaagcctattcgcaggatttctcattactaacaa
G13889A	aactacctaaccaacaaacttaaaataaaatccccactat[G/A]cacattttatttctccaacatactcggattctaccctagc
G1393A	aagaaatgggctacattttctaccccagaaaactacgata[G/A]cccttatgaaacttaagggtcgaaggtggatttagcagta

G14162A	atcctaaccctactcctaatcacataacctattccccga[G/A]caatctcaattacaatatatacaccaacaaacaatggtca
G14323A	cctctccttcataaattattcagcttcctacactattaaa[G/A]tttaccacaaccaccaccccatcatactctttcacccaca
G143A	gccggagcaccctatgtcgcagtatctgtctttgattcct[G/A]cctcatcctattatttatcgcacctacgttcaatattaca
G1442A	aaacttaagggtcgaaggtggatttagcagtaaactaaga[G/A]tagagtgcttagttgaacagggccctgaagcgcgtacaca
G1462A	gatttagcagtaaactaagagtagagtgcttagttgaaca[G/A]ggccctgaagcgcgtacacaccgcccgtcaccctcctcaa
G14831A	ccactcattcatcgacctccccaccccatccaacatctcc[G/A]catgatgaaacttcggctcactccttggcgcctgcctgat
G14861A	caacatctccgcatgatgaaacttcggctcactccttggc[G/A]cctgcctgatcctccaaatcaccacaggactattcctagc
G14905A	gcctgatcctccaaatcaccacaggactattcctagccat[G/A]cactactcaccagacgcctcaaccgccttttcatcaatcg
G1503A	ggccctgaagcgcgtacacaccgcccgtcaccctcctcaa[G/A]tatacttcaaaggacatttaactaaaacccctacgcattt
G15043A	gcgcctcaatattctttatctgcctcttcctacacatcgg[G/A]cgaggcctatattacggatcatttctctactcagaaacct
G15047A	ctcaatattctttatctgcctcttcctacacatcgggcga[G/A]gcctatattacggatcatttctctactcagaaacctgaaa
G15106A	ttctctactcagaaacctgaaacatcggcattatcctcct[G/A]cttgcaactatagcaacagccttcataggctatgtcctcc
G15217A	cagtaattacaaacttactatccgccatcccatacattgg[G/A]acagacctagttcaatgaatctgaggaggctactcagtag
G15221A	ttacaaacttactatccgccatcccatacattgggaca[G/A]acctagttcaatgaatctgaggaggctactcagtagacag
G15257A	gacagacctagttcaatgaatctgaggaggctactcagta[G/A]acagtcccaccctcacacgattctttacctttcacttcat
G15301A	agtcccaccctcacacgattctttacctttcacttcatctt[G/A]cccttcattattgcagccctagcaacactccacctcctat
G15317A	attctttacctttcacttcatcttgcccttcattattgca[G/A]ccctagcaacactccacctcctattcttgcacgaaacggg
G15553A	ataccctagccaaccccttaaacacccctccccacatcaa[G/A]cccgaatgatatttcctattcgcctacacaattctccgat
G15734A	tttcgcccactaagccaatcactttattgactcctagcc[G/A]cagacctcctcattctaacctgaatcggaggacaaccagt
G15773A	cgcagacctcctcattctaacctgaatcggaggacaacca[G/A]taagctaccctttttaccatcattggacaagtagcatccgt
G15883A	taccaactatctccctaattgaaaacaaaatactcaaatg[G/A]gcctgtccttgtagtataaactaatacaccagtcttgtaa
G15884C	accaactatctccctaattgaaaacaaaatactcaaatgg[G/C]cctgtccttgtagtataaactaatacaccagtcttgtaaa
G15927A	tgtccttgtagtataaactaatacaccagtcttgtaaacc[G/A]gagatgaaaacctttttccaaggacaaatcagagaaaaag
G15928A	gtccttgtagtataaactaatacaccagtcttgtaaaccg[G/A]agatgaaaacctttttccaaggacaaatcagagaaaaagt
G1598A	agtcgtaacatggtaagtgtactggaaagtgcacttggac[G/A]aaccagagtgtagcttaacacaaagcacccaacttacact
G16129A	gtatttcgtacattactgccagccaccatgaatattgtac[G/A]gtaccataaatacttgaccacctgtagtacataaaaaccc
G16145A	tgccagccaccatgaatattgtacggtaccataaatactt[G/A]accacctgtagtacataaaaacccaatccacatcaaaacc
G16244A	gcaagtacagcaatcaaccctcaactatcacacatcaact[G/A]caactccaaagccacccctcacccactaggataccaacaa
G16255A	aatcaaccctcaactatcacacatcaactgcaactccaaa[G/A]ccacccctcacccactaggataccaacaaacctacccacc
G16274A	cacatcaactgcaactccaaagccacccctcacccactag[G/A]ataccaacaaacctacccacccttaacagtacatagtaca
G16390A	aatcccttctcgtccccatggatgacccccctcagatagg[G/A]gtcccttgaccaccatcctccgtgaaatcaatatcccgca
G16438A	gaccaccatcctccgtgaaatcaatatcccgcacaagagt[G/A]ctactctcctcgctccgggcccataacacttgggggtagc
G1664A	cacccaacttacacttaggagatttcaacttaacttgacc[G/A]ctctgagctaaacctagccccaaacccactccaccttact
G1709A	gagctaaacctagccccaaacccactccaccttactacca[G/A]acaaccttagccaaaccatttacccaaataaagtataggc
G1719A	tagccccaaacccactccaccttactaccagacaacctta[G/A]ccaaaccatttacccaaataaagtataggcgatagaaatt
G185A	ctcatcctattatttatcgcacctacgttcaatattacag[G/A]cgaacatacttactaaagtgtgttaattaattaatgcttg
G1888A	tctgcataatgaattaactagaaataactttgcaaggaga[G/A]ccaaagctaagacccccgaaaccagacgagctacctaaga
G2056A	ctggttgtccaagatagaatcttagttcaactttaaattt[G/A]cccacagaaccctctaaatccccttgtaaatttaactgtt
G207A	ctacgttcaatattacaggcgaacatacttactaaagtgt[G/A]ttaattaattaatgcttgtaggacataataataacaattg
G2145A	gaggaacagctctttggacactaggaaaaaaccttgtaga[G/A]agagtaaaaaatttaacacccatagtaggcctaaaagcag
G228A	aacatacttactaaagtgtgttaattaattaatgcttgta[G/A]gacataataataacaattgaatgtctgcacagccactttc
G2361A	aacattctcctccgcataagcctgcgtcagattaaaacact[G/A]aactgacaattaacagcccaatatctacaatcaaccaaca
G2758A	cgagaagaccctatggagctttaatttattaatgcaaaca[G/A]tacctaacaaacccacaggtcctaaactaccaaacctgca
G2831A	aacctgcattaaaaatttcggttggggcgacctcggagca[G/A]aacccaacctccgagcagtacatgctaagacttcaccagt
G2849A	cggttggggcgacctcggagcagaacccaacctccgagca[G/A]tacatgctaagacttcaccagtcaaagcgaactactatac
G3010A	caatagggtttacgacctcgatgttggatcaggacatccc[G/A]atggtgcagccgctattaaaggttcgtttgttcaacgatt
G316A	acatcataacaaaaaatttccaccaaaccccccctcccccc[G/A]cttctggccacagcacttaaacacatctctgccaaacccc

G322A	taacaaaaaatttccaccaaaccccccctcccccgcttct[G/A]gccacagcacttaaacacatctctgccaaaccccaaaaac
G3316A	agtcagaggttcaattcctcttcttaacaacatacccatg[G/A]ccaacctcctactcctcattgtacccattctaatcgcaat
G3337A	tcttaacaacatacccatggccaacctcctactcctcatt[G/A]tacccattctaatcgcaatggcattcctaatgcttaccga
G3391A	cgcaatggcattcctaatgcttaccgaacgaaaaattcta[G/A]gctatatacaactacgcaaaggccccaacgttgtaggccc
G3421A	aaaaattctaggctatatacaactacgcaaaggccccaac[G/A]tggtaggcccctacgggctactacaacccttcgctgacgc
G3438A	tacaactacgcaaaggccccaacgttgtaggcccctacgg[G/A]ctactacaacccttcgctgacgccataaaactcttcacca
G3483A	tacaacccttcgctgacgccataaaactcttcaccaaaga[G/A]cccctaaaacccgccacatctaccatcaccctctacatca
G3496T	tgacgccataaaactcttcaccaaagagcccctaaaaccc[G/T]ccacatctaccatcaccctctacatcaccgccccgacctt
G3531A	aacccgccacatctaccatcaccctctacatcaccgcccc[G/A]accttagctctcaccatcgctcttctactatgaaccccc
G3693A	tcctctgatcagggtgagcatcaaactcaaactacgccct[G/A]atcggcgcactgcgagcagtagcccaaacaatctcatatg
G3705A	ggtgagcatcaaactcaaactacgccctgatcggcgcact[G/A]cgagcagtagcccaaacaatctcatatgaagtcaccctag
G3834A	ccacccttatcacaacacaagaacacctctgattactcct[G/A]ccatcatgacccttggccataatatgatttatctccacac
G3849A	cacaagaacacctctgattactcctgccatcatgaccctt[G/A]gccataatatgatttatctccacactagcagagaccaacc
G3882A	gacccttggccataatatgatttatctccacactagcaga[G/A]accaaccgaaccccttcgaccttgccgaaggggagtccg
G3918A	cagagaccaaccgaacccccttcgaccttgccgaagggga[G/A]tccgaactagtctcaggcttcaacatcgaatacgccgcag
G4113A	caacatattttgtcaccaagaccctacttctaacctccct[G/A]ttcttatgaattcgaacagcataccccgattccgctacg
G4491A	ttatacccttcccgtactaattaatcccctggcccaaccc[G/A]tcatctactctaccatctttgcaggcacactcatcacagc
G4580A	cgcactgatttttacctgagtaggcctagaaataaacat[G/A]ctagcttttattccagttctaaccaaaaaaataaaccctcg
G4580A-2	cgcactgatttttacctgagtaggcctagaaataaacat[G/A]ctagcttttattccagttctaaccaaaaaaataaaccctc
G4924A	atctcaatcatataccaaatctctccctcactaaacgtaa[G/A]ccttctcctcactctctcaatcttatccatcatagcaggc
G4991A	ccatcatagcaggcagttgaggtggattaaaccagaccca[G/A]ctacgcaaaatcttagcatactcctcaattacccacatag
G5046A	agcatactcctcaattacccacataggatgaataatagca[G/A]ttctaccgtacaaccctaacataaccattcttaatttaac
G5147A	ctaccgcattcctactactcaacttaaactccagcaccac[G/A]accctactactatctcgcacctgaaacaagctaacatgac
G5177A	ccagcaccacgaccctactactatctcgcacctgaaacaa[G/A]ctaacatgactaacacccttaattccatccaccctcctct
G5252A	tcctctccctaggaggcctgcccccgctaaccggcttttt[G/A]cccaaatgggccattatcgaagaattcacaaaaaacaata
G5261A	taggaggcctgcccccgctaaccggctttttgcccaaatg[G/A]gccattatcgaagaattcacaaaaaacaatagcctcatca
G5262A	aggaggcctgcccccgctaaccggctttttgcccaaatgg[G/A]ccattatcgaagaattcacaaaaaacaatagcctcatcat
G5460A	tgaacatacaaaacccaccccattcctccccacactcatc[G/A]cccttaccacgctactcctacctatctccccttttatact
G5563A	ttaaatacagaccaagagccttcaaagccctcagtaagtt[G/A]caatacttaatttctgtaacagctaaggactgcaaaaccc
G5585A	caaagccctcagtaagttgcaatacttaatttctgtaaca[G/A]ctaaggactgcaaaaccccactctgcatcaactgaacgca
G5773A	ctcccgccgccgggaaaaaaggcgggagaagccccggcag[G/A]tttgaagctgcttcttcgaatttgcaattcaatatgaaaa
G5824A	cttcttcgaatttgcaattcaatatgaaaatcacctcgga[G/A]ctggtaaaaagaggcctaacccctgtctttagatttacag
G6023A	gagtcctaggcacagctctaagcctccttattcgagccga[G/A]ctgggccagccaggcaaccttctaggtaacgaccacatct
G6026A	cctaggcacagctctaagcctccttattcgagccgagct[G/A]ggccagccaggcaaccttctaggtaacgaccacatctaca
G6150A	agtaatacccatcataatcggaggctttggcaactgacta[G/A]ttcccctaataatcggtgcccccgatatggcgtttccccg
G6249A	ctgactcttacctccctctctcctactcctgctcgcatct[G/A]ctatagtggaggccggagcaggaacaggttgaacagtcta
G6261A	tccctctctcctactcctgctcgcatctgctatagtggag[G/A]ccggagcaggaacaggttgaacagtctaccctcccttagc
G6446A	ttatcaatataaaacccccctgccataacccaataccaaac[G/A]cccctcttcgtctgatccgtcctaatcacagcagtcctac
G6480A	ccaaacgcccctcttcgtctgatccgtcctaatcacagca[G/A]tcctacttctcctatctctcccagtcctagctgctggcat
G6734A	aaaaagaaccatttggatacataggtatggtctgagctat[G/A]atatcaattggcttcctagggtttatcgtgtgagcacacc
G6755A	taggtatggtctgagctatgatatcaattggcttcctagg[G/A]tttatcgtgtgagcacaccatatatttacagtaggaatag
G6899A	aagtatttagctgactcgccacactccacggaagcaatat[G/A]aaatgatctgctgcagtgctctgagccctaggattcatct
G68A	actcacgggagctctccatgcatttggtattttcgtctgg[G/A]gggtatgcacgcgatagcattgcgagacgctggagccgga
G6917A	ccacactccacggaagcaatatgaaatgatctgctgcagt[G/A]ctctgagccctaggattcatctttcttttcaccgtaggtg
G7013A	tattagcaaactcatcactagacatcgtactacacgacac[G/A]tactacgttgtagcccacttccactatgtcctatcaatag
G709A	tctattagctcttagtaagattacacatgcaagcatcccc[G/A]ttccagtgagttcaccctctaaatcaccacgatcaaaagg
G719A	cttagtaagattacacatgcaagcatccccgttccagtga[G/A]ttcaccctctaaatcaccacgatcaaaaggaacaagcatca

G7226A	aacactttctcggcctatccggaatgccccgacgttactc[G/A]gactaccccgatgcatacaccacatgaaacatcctatcat
G7340A	tattaataattttcatgatttgagaagccttcgcttcgaa[G/A]cgaaaagtcctaatagtagaagaaccctccataaacctgg
G7444A	ccctaccacacattcgaagaacccgtatacataaaatcta[G/A]acaaaaaaggaaggaatcgaaccccccaaagctggtttca
G7598A	aattataggctaaatcctatatatcttaatggcacatgca[G/A]cgcaagtaggtctacaagacgctacttcccctatcataga
G75A	ggagctctccatgcatttggtattttcgtctgggggggtat[G/A]cacgcgatagcattgcgagacgctggagccggagcaccct
G7604A	aggctaaatcctatatatcttaatggcacatgcagcgcaa[G/A]taggtctacaagacgctacttcccctatcatagaagagct
G7642A	aagtaggtctacaagacgctacttcccctatcatagaaga[G/A]cttatcacctttcatgatcacgccctcataatcattttcc
G7664A	ttcccctatcatagaagagcttatcacctttcatgatcac[G/A]ccctcataatcattttccttatctgcttcctagtcctgta
G7697A	tgatcacgccctcataatcattttccttatctgcttccta[G/A]tcctgtatgcccttttcctaacactcacaacaaaactaac
G7706A	cctcataatcattttccttatctgcttcctagtcctgtat[G/A]cccttttcctaacactcacaacaaaactaactaatactaa
G7762A	tcacaacaaaactaactaatactaacatctcagacgctca[G/A]gaaatagaaaccgtctgaactatcctgcccgccatcatcc
G7805A	aatagaaaccgtctgaactatcctgcccgccatcatccta[G/A]tcctcatcgccctcccatccctacgcatcctttacataac
G7830A	cccgccatcatcctagtcctcatcgccctcccatccctac[G/A]catcctttacataacagacgaggtcaacgatccctccctt
G7912A	ccatcaaatcaattggccaccaatggtactgaacctacga[G/A]tacaccgactacggcggactaatcttcaactcctacatac
G8155A	gacgtctaaaccaaaccactttcaccgctacacgaccggg[G/A]gtatactacggtcaatgctctgaaatctgtggagcaaacc
G8251A	tcgtcctagaattaattcccctaaaaatctttgaaatagg[G/A]cccgtatttaccctatagcacccctctacccctctaga
G8292A	cccgtatttaccctatagcacccctctacccctctaga[G/A]cccactgtaaagctaacttagcattaaccttttaagttaa
G8387A	aacacctctttacagtgaaatgccccaactaaatactacc[G/A]tatggcccaccataattacccccatactccttacactatt
G8572A	cgaaaatctgttcgcttcattcattgcccccacaatccta[G/A]gcctacccgccgcagtactgatcattctatttccccctct
G8592A	tcattgcccccacaatcctaggcctacccgccgcagtact[G/A]atcattctatttccccctctattgatccccacctccaaat
G8616A	tacccgccgcagtactgatcattctatttccccctctatt[G/A]atccccacctccaaatatctcatcaacaaccgactaatca
G8616T	tacccgccgcagtactgatcattctatttccccctctatt[G/T]atccccacctccaaatatctcatcaacaaccgactaatca
G8697A	ccacccaacaatgactaatcaaactaacctcaaaacaaat[G/A]ataaccatacacaacactaaaggacgaacctgatctctta
G8764A	aacctgatctcttatactagtatccttaatcatttttatt[G/A]ccacaactaacctcctcggactcctgcctcactcatttac
G8838A	catttacaccaaccacccaactatctataaacctagccat[G/A]gccatccccttatgagcgggcacagtgattataggctttc
G8839A	catttacaccaaccacccaactatctataaacctagccat[G/A]ccatccccttatgagcgggcacagtgattataggctttcg
G8839A-2	atttacaccaaccacccaactatctataaacctagccatg[G/A]ccatccccttatgagcgggcacagtgattataggctttcg
G8994A	ttatcgaaaccatcagcctactcattcaaccaatagccct[G/A]gccgtacgcctaaccgctaacattactgcaggccacctac
G9053A	taacattactgcaggccacctactcatgcacctaattggaa[G/A]cgccaccctagcaatatcaaccattaaccttccctctac
G9055A	cattactgcaggccacctactcatgcacctaattggaagc[G/A]ccaccctagcaatatcaaccattaaccttccctctacact
G9123A	ttccctctacacttatcatcttcacaattctaattctact[G/A]actatcctagaaatcgctgtcgccttaatccaagcctacg
G9139A	catcttcacaattctaattctactgactatcctagaaatc[G/A]ctgtcgccttaatccaagcctacgttttcacacttctagt
G9182A	gtcgccttaatccaagcctacgttttcacacttctagtaa[G/A]cctctacctgcacgacaacacataatgacccaccaatcac
G9266A	cctatcatatagtaaaacccagcccatgacccctaacagg[G/A]gccctctcagccctcctaatgacctccggcctagccatgt
G92A	tggtattttcgtctgggggggtatgcacgcgatagcattgc[G/A]agacgctggagccggagcaccctatgtcgcagtatctgtc
G9300A	aacaggggccctctcagccctcctaatgacctccggccta[G/A]ccatgtgatttcacttccactccataacgctcctcatact
G9305A	gggccctctcagccctcctaatgacctccggcctagccat[G/A]tgatttcacttccactccataacgctcctcatactaggcc
G930A	ccagccaccgcggtcacacgattaacccaagtcaatagaa[G/A]ccggcgtaaagagtgttttagatcacccccctccccaataa
G9438A	aagcacataccaaggccaccacacaccacctgtccaaaaa[G/A]gccttcgatacgggataatcctatttattacctcagaagt
G9477A	aggccttcgatacgggataatcctatttattacctcagaa[G/A]ttttttcttcgcaggatttttctgagccttttaccactc
G94A	gtattttcgtctgggggggtatgcacgcgatagcattgcga[G/A]acgctggagccggagcaccctatgtcgcagtatctgtctt
G951A	ttaacccaagtcaatagaagccggcgtaaagagtgtttta[G/A]atcacccccctccccaataaagctaaaactcacctgagttg
G9548A	tttaccactccagcctagcccctacccccaattaggagg[G/A]cactggccccgaacaggcatcaccccgctaaatcccctag
G9548A-2	tttaccactccagcctagcccctacccccaattaggagg[G/A]cactggcccccaacaggcatcaccccgctaaatcccctaga
G9554A	actccagcctagcccctacccccaattaggagggcactg[G/A]ccccgaacaggcatcaccccgctaaatcccctagaagtcc
G9612A	cccgctaaatcccctagaagtcccactcctaaacacatcc[G/A]tattactcgcatcaggagtatcaatcacctgagctcacca

G9804A	caccatttccgacggcatctacggctcaacatttttgta[G/A]ccacaggcttccacggacttcacgtcattattggctcaac
G988A	ttagatcacccccctccccaataaagctaaaactcacctga[G/A]ttgtaaaaaactccagttgacacaaaatagactacgaaag
G9921A	tcactttacatccaaacatcactttggcttcgaagccgcc[G/A]cctgatactggcattttgtagatgtggtttgactatttct
G9947A	gcttcgaagccgccgcctgatactggcattttgtagatgt[G/A]gtttgactatttctgtatgtctccatctattgatgagggt
G9948A	cttcgaagccgccgcctgatactggcattttgtagatgtg[G/A]tttgactatttctgtatgtctccatctattgatgagggtc
G9962A	cctgatactggcattttgtagatgtggtttgactatttct[G/A]tatgtctccatctattgatgagggtcttactcttttagta
G9986A	tggtttgactatttctgtatgtctccatctattgatgagg[G/A]tcttactcttttagtataaatagtaccgttaacttccaat
T10007C	tctccatctattgatgagggtcttactcttttagtataaa[T/C]agtaccgttaacttccaattaactagttttgacaacattc
T10031C	actcttttagtataaatagtaccgttaacttccaattaac[T/C]agttttgacaacattcaaaaaagagtaataaacttcgcct
T10034C	cttttagtataaatagtaccgttaacttccaattaactag[T/C]tttgacaacattcaaaaaagagtaataaacttcgccttaa
T1005C	caataaagctaaaactcacctgagttgtaaaaaactccag[T/C]tgacacaaaatagactacgaaagtggctttaacatatctg
T10084C	attcaaaaaagagtaataaacttcgccttaattttaataa[T/C]caacaccctcctagccttactactaataattattacattt
T10101C	aaacttcgccttaattttaataatcaacaccctcctagcc[T/C]tactactaataattattacatttgactaccacaactcaa
T10118C	taataatcaacaccctcctagccttactactaataattat[T/C]acatttgactaccacaactcaacggctacatagaaaaat
T10166C	gactaccacaactcaacggctacatagaaaaatccacccc[T/C]tacgagtgcggcttcgaccctatatcccccgcccgcgtcc
T10187C	acatagaaaaatccaccccttacgagtgcggcttcgaccc[T/C]atatcccccgcccgcgtccctttctccataaaattcttct
T10238C	cccgcgtccctttctccataaaattcttcttagtagctat[T/C]accttcttattatttgatctagaaattgccctccttttac
T10245C	ccctttctccataaaattcttcttagtagctattaccttc[T/C]tattatttgatctagaaattgccctccttttacccctacc
T10256C	taaaattcttcttagtagctattaccttcttattatttga[T/C]ctagaaattgccctccttttacccctaccatgagccctac
T10321C	ctaccatgagccctacaaacaactaacctgccactaatag[T/C]tatgtcatccctcttattaatcatcatcctagccctaagt
T10370C	ccctcttattaatcatcatcctagccctaagtctggccta[T/C]gagtgactacaaaaaggattagactgaaccgaattggtat
T10389C	cctagccctaagtctggcctatgagtgactacaaaaagga[T/C]tagactgaaccgaattggtatatagtttaaacaaaacgaa
T10463C	aaacgaatgatttcgactcattaaattatgataatcatat[T/C]taccaaatgcccctcatttacataaatattatactagcat
T10493C	ataatcatatttaccaaatgcccctcatttacataaatat[T/C]atactagcatttaccatctcacttctaggaatactagtat
T10535C	tactagcatttaccatctcacttctaggaatactagtata[T/C]cgctcacacctcatatcctccctactatgcctagaaggaa
T10667C	ccctcttagccaatattgtgcctattgccatactagtctt[T/C]gccgcctgcgaagcagcggtgggcctagccctactagtct
T10724C	gcggtgggcctagccctactagtctcaatctccaacacata[T/C]ggcctagactacgtacataacctaaacctactccaatgct
T10786C	ctaaacctactccaatgctaaaactaatcgtcccaacaat[T/C]atattactaccactgacatgactttccaaaaaacacataa
T10810C	taatcgtcccaacaattatattactaccactgacatgact[T/C]tccaaaaaacacataatttgaatcaacacaaccacccaca
T10873C	tcaacacaaccacccacagcctaattattagcatcatccc[T/C]ctactatttttaaccaaatcaacaacaacctatttagct
T10907C	catccctctactatttttaaccaaatcaacaacaacccta[T/C]ttagctgttccccaaccttttcctccgacccccctaacaac
T10915C	tactatttttaaccaaatcaacaacaacctatttagctg[T/C]tccccaaccttttcctccgacccccctaacaaccccccctcc
T10927C	accaaatcaacaacaacctatttagctgttccccaacctt[T/C]tcctccgacccccctaacaacccccccctcctaatactaacta
T1095C	tagctaagacccaaactgggattagataccccactatgct[T/C]agccctaaacctcaacagttaaatcaacaaaactgctcgc
T11009C	ctgactcctacccctcacaatcatggcaagccaacgccac[T/C]tatccagtgaaccactatcacgaaaaaaactctacctctc
T11025C	acaatcatggcaagccaacgccacttatccagtgaaccac[T/C]atcacgaaaaaaactctacctctctatactaatctccta
T1107C	aaactgggattagataccccactatgcttagccctaaacc[T/C]caacagttaaatcaacaaaactgctcgccagaacactacg
T11147C	tcatattttatatcttcttcgaaaccacacttatccccacc[T/C]tggctatcatcacccgatgaggcaaccagccagaacgcc
T1119C	gataccccactatgcttagccctaaacctcaacagttaaa[T/C]caacaaaactgctcgccagaacactacgagccacagctta
T11204C	atgaggcaaccagccagaacgcctgaacgcaggcacatac[T/C]tcctattctacaccctagtaggctcccttcccctactcat
T11253C	tacaccctagtaggctcccttcccctactcatcgcactaa[T/C]ttacactcacaacaccctaggctcactaaacattctacta
T11255C	caccctagtaggctcccttcccctactcatcgcactaatt[T/C]acactcacaacaccctaggctcactaaacattctactact
T11260C	tagtaggctcccttcccctactcatcgcactaatttacac[T/C]cacaacaccctaggctcactaaacattctactactcactc
T11287C	cactaatttacactcacaacaccctaggctcactaaacat[T/C]ctactactcactctcactgcccaagaactatcaaactcct
T11299C	ctcacaacaccctaggctcactaaacattctactactcac[T/C]ctcactgcccaagaactatcaaactcctgagccaacaact
T11339C	tctcactgcccaagaactatcaaactcctgagccaataac[T/C]taatatgactagcttacacaatagcttttatagtaaagat
T11353C	aactatcaaactcctgagccaataacttaatatgactagc[T/C]tacacaatagcttttatagtaaagatacctctttacggac

T11368C	gagccaataacttaatatgactagcttacacaatagcttt[T/C]atagtaaagatacctctttacggactccacttatgactcc
T11383C	tatgactagcttacacaatagcttttatagtaaagatacc[T/C]ctttacggactccacttatgactccctaaagcccatgtcg
T11410C	tagtaaagatacctctttacggactccacttatgactccc[T/C]aaagcccatgtcgaagcccccatcgctgggtcaatagtac
T11701C	ttctcatccaaacccccctgaagcttcaccggcgcagtcat[T/C]ctcataatcgcccacgggcttacatcctcattactattct
T11722C	gcttcaccggcgcagtcattctcataatcgcccacgggct[T/C]acatcctcattactattctgcctagcaaactcaaactacg
T11732C	cgcagtcattctcataatcgcccacgggcttacatcctca[T/C]tactattctgcctagcaaactcaaactacgaacgcactca
T11854C	tcccactaatagctttttgatgacttctagcaagcctcgc[T/C]aacctcgccttaccccccactattaacctactgggagaac
T11864C	agctttttgatgacttctagcaagcctcgctaacctcgcc[T/C]taccccccactattaacctactgggagaactctctgtgct
T1189C	gccacagcttaaaactcaaaggacctggcggtgcttcata[T/C]ccctctagaggagcctgttctgtaatcgataaaccccgat
T11935C	tctctgtgctagtaaccacgttctcctgatcaaatatcac[T/C]ctcctacttacaggactcaacatactagtcacagccctat
T1193C	cagcttaaaactcaaaggacctggcggtgcttcatatccc[T/C]ctagaggagcctgttctgtaatcgataaaccccgatcaac
T12083G	ctcattcacacgagaaaacaccctcatgttcatacaccta[T/G]cccccattctcctcctatccctcaaccccgacatcattac
T12136C	tcctatccctcaaccccgacatcattaccgggttttcctc[T/C]tgtaaatatagtttaaccaaaacatcagattgtgaatctg
T12189C	ttaaccaaaacatcagattgtgaatctgacaacagaggct[T/C]acgaccccttatttaccgagaaagctcacaagaactgcta
T12285C	tctaacaacatggctttctcaacttttaaaggataacagc[T/C]atccattggtcttaggccccaaaaattttggtgcaactcc
T12338C	taggccccaaaaattttggtgcaactccaaataaaagtaa[T/C]aaccatgcacactactataaccaccctaaccctgacttcc
T12354C	tggtgcaactccaaataaaagtaataaccatgcacactac[T/C]ataaccaccctaaccctgacttccctaattccccccatcct
T12396C	taaccaccctaaccctgacttccctaattccccccatcct[T/C]accaccctcgttaaccctaacaaaaaaaactcataccccc
T12408C	ccctgacttccctaattccccccatccttaccaccctcgt[T/C]aaccctaacaaaaaaaactcataccccccattatgtaaaat
T12414C	cttccctaattccccccatccttaccaccctcgttaaccc[T/C]aacaaaaaaaactcataccccccattatgtaaaatccattg
T12438C	ccaccctcgttaaccctaacaaaaaaaactcatacccccca[T/C]tatgtaaaatccattgtcgcatccacctttattatcagtc
T1243C	cctgttctgtaatcgataaaccccgatcaacctcaccacc[T/C]cttgctcagcctatataccgccatcttcagcaaaccctga
T12441C	ccctcgttaaccctaacaaaaaaaactcataccccccatta[T/C]gtaaaatccattgtcgcatccacctttattatcagtctct
T12468C	catacccccattatgtaaaatccattgtcgcatccacctt[T/C]attatcagtctcttccccacaacaatattcatgtgcctag
T12477C	attatgtaaaatccattgtcgcatccacctttattatcag[T/C]ctcttccccacaacaatattcatgtgcctagaccaagaagt
T12519C	tcttccccacaacaatattcatgtgcctagaccaagaagt[T/C]attatctcgaactgacactgagccacaacccaaacaaccc
T12609C	taagcttcaaactagactacttctccataatattcatccc[T/C]gtagcattgttcgttacatggtccatcatagaattctcac
T12696C	atataaactcagacccaaacattaatcagttcttcaaata[T/C]ctactcatcttcctaattaccatactaatcttagttaccg
T12714C	acattaatcagttcttcaaatatctactcatcttcctaat[T/C]accatactaatcttagttaccgctaacaacctattccaac
T12732C	aatatctactcatcttcctaattaccatactaatcttagt[T/C]accgctaacaacctattccaactgttcatcggctgagagg
T12793C	actgttcatcggctgagagggcgtaggaattatatccttc[T/C]tgctcatcagttgatgatacgcccgagcagatgccaacac
T12954C	ccaactcatgagacccacaacaaatagcccttctaaacgc[T/C]aatccaagcctcaccccactactaggcctcctcctagcag
T13020C	gcctcctcctagcagcagcaggcaaatcagcccaattagg[T/C]ctccacccctgactcccctcagccatagaaggccccaccc
T13111C	ccctactccactcaagcactatagttgtagcaggaatcttc[T/C]tactcatccgcttccacccccctagcagaaaatagcccac
T13143C	gaatcttcttactcatccgcttccacccccctagcagaaaa[T/C]agcccactaatccaaactctaacactatgcttaggcgcta
T13215C	taggcgctatcaccactctgttcgcagcagtctgcgccct[T/C]acacaaaatgacatcaaaaaaatcgtagccttctccactt
T13281C	tagccttctccacttcaagtcaactaggactcataatagt[T/C]acaatcggcatcaaccaaccacacctagcattcctgcaca
T13326C	tcggcatcaaccaaccacacctagcattcctgcacatctg[T/C]acccacgccttcttcaaagccatactatttatgtgctccg
T13404C	ccgggtccatcatccacaaccttaacaatgaacaagatat[T/C]cgaaaaataggaggactactcaaaaccatacctctcactt
T13437C	aagatattcgaaaaataggaggactactcaaaaccatacc[T/C]ctcacttcaacctccctcaccattggcagcctagcattag
T13461C	tactcaaaaccatacctctcacttcaacctccctcaccat[T/C]ggcagcctagcattagcaggaataccttttcctcacaggtt
T13602C	ctatctattactctcatcgctacctccctgacaagcgccta[T/C]agcactcgaataattcttctcaccctaacaggtcaacctc
T13617C	tcgctacctccctgacaagcgcctatagcactcgaataat[T/C]cttctcaccctaacaggtcaacctcgcttccccaccctta
T13635C	gcgcctatagcactcgaataattcttctcaccctaacagg[T/C]caacctcgcttccccacccttactaacattaacgaaaata
T13641C	atagcactcgaataattcttctcaccctaacaggtcaacc[T/C]cgcttccccacccttactaacattaacgaaaataacccca
T13656C	ttcttctcaccctaacaggtcaacctcgcttccccaccct[T/C]actaacattaacgaaaataaccccaccctactaaacccca
T13734C	ccattaaacgcctggcagccggaagcctattcgcaggatt[T/C]ctcattactaacaacatttcccccgcatccccccttccaaa

T13743C	gcctggcagccggaagcctattcgcaggatttctcattac[T/C]aacaacatttcccccgcatccccсttccaaacaacaatccc
T13899C	ccaacaaacttaaaataaaatccccactatgcacatttta[T/C]ttctccaacatactcggattctaccctagcatcacacacc
T1391C	gcaagaaatgggctacattttctaccccagaaaactacga[T/C]agcccttatgaaacttaagggtcgaaggtggatttagcag
T13965C	ctagcatcacacaccgcacaatcccctatctaggccttct[T/C]acgagccaaaacctgcccctactcctcctagacctaacct
T14000A	cttcttacgagccaaaacctgcccctactcctcctagacc[T/A]aacctgactagaaaagctattacctaaaacaatttcacag
T14020C	gcccctactcctcctagacctaacctgactagaaaagcta[T/C]tacctaaaacaatttcacagcaccaaatctccacctccat
T1406C	cattttctaccccagaaaactacgatagcccttatgaaac[T/C]taagggtcgaaggtggatttagcagtaaactaagagtaga
T1407A	attttctaccccagaaaactacgatagcccttatgaaact[T/A]aagggtcgaaggtggatttagcagtaaactaagagtagag
T14094C	cctccatcatcacctcaacccaaaaaggcataattaaact[T/C]tacttcctctctttcttcttcccactcatcctaaccctac
T14179C	aatcacataacctattcccccgagcaatctcaattacaat[T/C]tatacaccaacaaacaatgttcaaccagtaactactacta
T14182C	cacataacctattcccccgagcaatctcaattacaatata[T/C]acaccaacaaacaatgttcaaccagtaactactactaatc
T1420C	gaaaactacgatagcccttatgaaacttaagggtcgaagg[T/C]ggatttagcagtaaactaagagtagagtgcttagttgaac
T14325C	tctccttcataaattattcagcttcctacactattaaagt[T/C]taccacaaccaccaccccatcatactctttcacccacagg
T14371C	acaaccaccaccccatcatactctttcacccacaggacga[T/C]cctacctccatcgctaaccccactaaaacactcaccaag
T14374C	ccaccaccccatcatactctttcacccacaggacgaatcc[T/C]acctccatcgctaaccccactaaaacactcaccaagacct
T14470C	atgcctcaggatactcctcaatagccatcgctgtagtata[T/C]ccaaagacaaccatcattcccсctaaataaattaaaaaa
T14484C	tcctcaatagccatcgctgtagtatatccaaagacaacca[T/C]cattcccсctaaataaattaaaaaaactattaaacccata
T14488C	caatagccatcgctgtagtatatccaaagacaaccatcat[T/C]cccсctaaataaattaaaaaaactattaaacccatataac
T14550C	aactattaaacccatataacctcccccaaaattcagaataa[T/C]aacacacccgaccacaccgctaacaatcaatactaaaccc
T14577C	aaattcagaataataacacacccgaccacaccgctaacaa[T/C]caatactaaacccccataaataggagaaggcttagaagaa
T14581C	tcagaataataacacacccgaccacaccgctaacaatcaa[T/C]actaaacccccataaataggagaaggcttagaagaaaacc
T146C	ggagcaccctatgtcgcagtatctgtctttgattcctgcc[T/C]catcctattatttatcgcacctacgttcaatattacaggc
T14766C	atttcaactacaagaacaccaatgaccccaatacgcaaaa[T/C]taacccсctaataaaattaattaaccactcattcatcgac
T14783C	accaatgaccccaatacgcaaaactaaccccctaataaaa[T/C]taattaaccactcattcatcgacctccccaccccatccaa
T14971C	ccttttcatcaatcgcccacatcactcgagacgtaaatta[T/C]ggctgaatcatccgctaccttcacgccaatggcgcctcaa
T15071C	cctacacatcgggcgaggcctatattacggatcatttctc[T/C]actcagaaacctgaaacatcggcattatcctcctgcttgc
T15099C	ggatcatttctctactcagaaacctgaaacatcggcatta[T/C]cctcctgcttgcaactatagcaacagccttcataggctat
T15184C	tcccgtgaggccaaatatcattctgaggggccacagtaat[T/C]acaaacttactatccgccatcccatacattgggacagacc
T15191C	aggccaaatatcattctgaggggccacagtaattacaaac[T/C]tactatccgccatcccatacattgggacagacctagttca
T1520C	acaccgcccgtcaccctcctcaagtatacttcaaaggaca[T/C]ttaactaaaacccctacgcatttatatagaggagacaagt
T15229C	acttactatccgccatcccatacattgggacagacctagt[T/C]caatgaatctgaggaggctactcagtagacagtcccaccct
T15299C	acagtcccaccctcacacgattctttacctttcacttcatc[T/C]tgcccttcattattgcagccctagcaacactccacctcct
T152C	ccctatgtcgcagtatctgtctttgattcctgcctcatcc[T/C]attatttatcgcacctacgttcaatattacaggcgaacat
T15412C	taggaatcacctcccattccgataaaatcaccttccaccc[T/C]tactacacaatcaaagacgccctcggcttacttctcttcc
T15479C	cttacttctcttccttctctccttaatgacattaacacta[T/C]tctcaccagacctcctaggcgacccagacaattataccct
T15511C	taacactattctcaccagacctcctaggcgacccagacaa[T/C]tataccctagccaaccccttaaacacccctccccacatca
T15514C	cactattctcaccagacctcctaggcgacccagacaatta[T/C]accctagccaaccccttaaacacccctccccacatcaagc
T15519C	attctcaccagacctcctaggcgacccagacaattataccc[T/C]agccaaccccttaaacacccctccccacatcaagcccgaat
T15530C	cctcctaggcgacccagacaattataccctagccaacccc[T/C]taaacacccctccccacatcaagcccgaatgatatttcct
T15565C	accccttaaacacccctccccacatcaagcccgaatgata[T/C]ttcctattcgcctacacaattctccgatccgtccctaaca
T15622C	caattctccgatccgtccctaacaaactaggaggcgtcct[T/C]gccctattactatccatcctcatcctagcaataatcccca
T15670C	tactatccatcctcatcctagcaataatccccatcctcca[T/C]atatccaaacaacaaagcataatatttcgcccactaagcc
T15693C	ataatccccatcctccatatatccaaacaacaaagcataa[T/C]atttcgcccactaagccaatcactttattgactcctagcc
T15697C	tccccatcctccatatatccaaacaacaaagcataatatt[T/C]cgcccactaagccaatcactttattgactcctagccgcaga
T15784C	tcattctaacctgaatcggaggacaaccagtaagctaccc[T/C]tttaccatcattggacaagtagcatccgtactatacttca
T15852C	gtactatacttcacaacaatcctaatcctaataccaacta[T/C]ctccctaattgaaaacaaaatactcaaatgggcctgtcct
T15900C	attgaaaacaaaatactcaaatgggcctgtccttgtagta[T/C]aaactaatacaccagtcttgtaaaccggagatgaaaacct

T15942C	aactaatacaccagtcttgtaaaccggagatgaaaacctt[T/C]ttccaaggacaatcagagaaaaagtctttaactccacca
T15968C	gagatgaaaaccttttccaaggacaaatcagagaaaaag[T/C]ctttaactccaccattagcacccaaagctaagattctaat
T16075C	ggaagcagatttgggtaccacccaagtattgactcaccca[T/C]caacaaccgctatgtatttcgtacattactgccagccacc
T16093C	cacccaagtattgactcacccatcaacaaccgctatgtat[T/C]tcgtacattactgccagccaccatgaatattgtacggtac
T16124C	gctatgtatttcgtacattactgccagccaccatgaatat[T/C]gtacggtaccataaatacttgaccacctgtagtacataa
T16124C -2	gctatgtatttcgtacattactgccagccaccatgaatat[T/C]gtacggtaccataaatacttgaccacctgtagtacataaa
T16126C	tatgtatttcgtacattactgccagccaccatgaatattg[T/C]acggtaccataaatacttgaccacctgtagtacataaaaa
T16154C	ccatgaatattgtacggtaccataaatacttgaccacctg[T/C]agtacataaaaacccaatccacatcaaaacccccctcccca
T16172C	accataaatacttgaccacctgtagtacataaaaacccaa[T/C]ccacatcaaaacccccctccccatgcttacaagcaagtaca
T16189C	acctgtagtacataaaaacccaatccacatcaaaaccccc[T/C]ccccatgcttacaagcaagtacagcaatcaaccctcaact
T16224C	ccccctccccatgcttacaagcaagtacagcaatcaaccc[T/C]caactatcacacatcaactgcaactccaaagccacccctc
T16271C	tcacacatcaactgcaactccaaagccacccctcacccac[T/C]aggataccaacaaacctacccacccttaacagtacatagt
T16298C	acccctcacccactaggataccaacaaacctacccaccct[T/C]aacagtacatagtacataaagccatttaccgtacatagca
T16304C	cacccactaggataccaacaaacctacccacccttaacag[T/C]acatagtacataaagccatttaccgtacatagcacattac
T16311C	taggataccaacaaacctacccacccttaacagtacatag[T/C]acataaagccatttaccgtacatagcacattacagtcaaa
T16342C	agtacatagtacataaagccatttaccgtacatagcacat[T/C]acagtcaaatcccttctcgtccccatggatgaccccc
T16362C	atttaccgtacatagcacattacagtcaaatcccttctcg[T/C]ccccatggatgacccccctcagataggggtcccttgacca
T16381C	ttacagtcaaatcccttctcgtccccatggatgaccccc[T/C]cagataggggtcccttgaccaccatcctccgtgaaatcaa
T16505C	cacttgggggtagctaaagtgaactgtatccgacatctgg[T/C]tcctacttcagggtcataaagcctaaatagcccacacgtt
T16519C	taaagtgaactgtatccgacatctggttcctacttcaggg[T/C]cataaagcctaaatagcccacacgttccccttaaataaga
T1700C	gaccgctctgagctaaacctagccccaaacccactccacc[T/C]tactaccagacaaccttagccaaaccatttacccaaataa
T1738C	ccttactaccagacaaccttagccaaaccatttacccaaa[T/C]aaagtataggcgatagaaattgaaacctggcgcaatagat
T1766C	catttacccaaataaagtataggcgatagaaattgaaacc[T/C]ggcgcaatagatatagtaccgcaagggaaagatgaaaaat
T1780C	aagtataggcgatagaaattgaaacctggcgcaatagata[T/C]agtaccgcaagggaaagatgaaaaattataaccaagcata
T1819C	atagtaccgcaagggaaagatgaaaaattataaccaagca[T/C]aatatagcaaggactaacccctataccttctgcataatga
T1822C	gtaccgcaagggaaagatgaaaaattataaccaagcataa[T/C]atagcaaggactaacccctataccttctgcataatgaatt
T1824C	accgcaagggaaagatgaaaaattataaccaagcataata[T/C]agcaaggactaacccctataccttctgcataatgaattaa
T1834C	aaagatgaaaaattataaccaagcataatatagcaaggac[T/C]aacccctataccttctgcataatgaattaactagaaataa
T1850C	aaccaagcataatatagcaaggactaacccctataccttc[T/C]gcataatgaattaactagaaataactttgcaaggagagcc
T195C	tatttatcgcacctacgttcaatattacaggcgaacatac[T/C]tactaaagtgtgttaattaattaatgcttgtaggacataa
T1977C	aagagcacacccgtctatgtagcaaaatagtgggaagatt[T/C]ataggtagaggcgacaaacctaccgagcctggtgatagct
T199C	tatcgcacctacgttcaatattacaggcgaacatacttac[T/C]aaagtgtgttaattaattaatgcttgtaggacataataat
T2010C	gaagatttataggtagaggcgacaaacctaccgagcctgg[T/C]gatagctggttgtccaagatagaatcttagttcaacttta
T2083C	caactttaaatttgcccacagaaccctctaaatccccttg[T/C]aaatttaactgttagtccaaagaggaacagctctttggac
T2141C	caaagaggaacagctctttggacactaggaaaaaaccttg[T/C]agagagagtaaaaaatttaacacccatagtaggcctaaaa
T2158C	ttggacactaggaaaaaaccttgtagagagagtaaaaaat[T/C]taacacccatagtaggcctaaaagcagccaccaattaaga
T2159C	tggacactaggaaaaaaccttgtagagagagtaaaaaatt[T/C]aacacccatagtaggcctaaaagcagccaccaattaagaa
T217C	tattacaggcgaacatacttactaaagtgtgttaattaat[T/C]aatgcttgtaggacataataacaattgaatgtctgcac
T2330C	aagaactaatgttagtataagtaacatgaaaacattctcc[T/C]ccgcataagcctgcgtcagattaaaacactgaactgacaa
T2352C	aacatgaaaacattctcctccgcataagcctgcgtcagat[T/C]aaaacactgaactgacaattaacagcccaatatctacaat
T236C	actaaagtgtgttaattaattaatgcttgtaggacataa[T/C]aataacaattgaatgtctgcacagccactttccacacaga
T2387C	cagattaaaacactgaactgacaattaacagcccaatatc[T/C]acaatcaaccaacaagtcattattaccctcactgtcaacc
T2404C	ctgacaattaacagcccaatatctacaatcaaccaacaag[T/C]cattattaccctcactgtcaacccaacacaggcatgctca
T2416C	agcccaatatctacaatcaaccaacaagtcattattaccc[T/C]cactgtcaacccaacacaggcatgctcataaggaaaggtt
T2442C	agtcattattaccctcactgtcaacccaacacaggcatgc[T/C]cataaggaaaggttaaaaaaagtaaaaggaactcggcaaa
T2445C	cattattaccctcactgtcaacccaacacaggcatgctca[T/C]aaggaaaggttaaaaaaagtaaaaggaactcggcaaatct
T2483C	cataaggaaaggttaaaaaaagtaaaaggaactcggcaaa[T/C]cttaccccgcctgtttaccaaaaacatcacctctagcatc

T2626C	tgcaaaggtagcataatcacttgttccttaaatagggacc[T/C]gtatgaatggctccacgagggttcagctgtctcttacttt
T2850C	ggttggggcgacctcggagcagaacccaacctccgagcag[T/C]acatgctaagacttcaccagtcaaagcgaactactatact
T2863C	tcggagcagaacccaacctccgagcagtacatgctaagac[T/C]tcaccagtcaaagcgaactactatactcaattgatccaat
T2885C	agcagtacatgctaagacttcaccagtcaaagcgaactac[T/C]atactcaattgatccaataacttgaccaacggaacaagtt
T3083C	caacgattaaagtcctacgtgatctgagttcagaccggag[T/C]aatccaggtcggtttctatctaccttcaaattcctccctg
T318C	atcataacaaaaaatttccaccaaaccccccctcccccgc[T/C]tctggccacagcacttaaacacatctctgccaaaccccaa
T3197C	caaagcgccttcccccgtaaatgatatcatctcaacttag[T/C]attatacccacacccacccaagaacagggtttgttaagat
T3278C	ggcagagcccggtaatcgcataaaacttaaaactttacag[T/C]cagaggttcaattcctcttcttaacaacatacccatggcc
T3290C	taatcgcataaaacttaaaactttacagtcagaggttcaa[T/C]tcctcttcttaacaacatacccatggccaacctcctactc
T3336C	ttcttaacaacatacccatggccaacctcctactcctcat[T/C]gtacccattctaatcgcaatggcattcctaatgcttaccg
T3338C	cttaacaacatacccatggccaacctcctactcctcattg[T/C]acccattctaatcgcaatggcattcctaatgcttaccgaa
T3372C	tcattgtacccattctaatcgcaatggcattcctaatgct[T/C]accgaacgaaaaattctaggctatatacaactacgcaaag
T3394C	aatggcattcctaatgcttaccgaacgaaaaattctaggc[T/C]atatacaactacgcaaaggccccaacgttgtaggcccccta
T3396C	tggcattcctaatgcttaccgaacgaaaaattctaggcta[T/C]atacaactacgcaaaggccccaacgtggtaggcccctacg
T3398C	gcattcctaatgcttaccgaacgaaaaattctaggctata[T/C]acaactacgcaaaggccccaacgtggtaggcccctacggc
T3423C	aaattctaggctatatacaactacgcaaaggccccaacgt[T/C]gtaggcccctacgggctactacaacccttcgctgacgcca
T3618C	tacccaacccccctggtcaacctcaacctaggcctcctatt[T/C]attctagccacctctagcctagccgtttactcaatcctct
T3645C	taggcctcctatttattctagccacctctagcctagccgt[T/C]tactcaatcctctgatcagggtgagcatcaaactcaaact
T3777C	tcaccctagccatcattctactatcaacattactaataag[T/C]ggctcctttaacctctccacccttatcacaacacaagaac
T3826C	taacctctccacccttatcacaacacaagaacacctctga[T/C]tactcctgccatcatgacccttggccataatatgatttat
T3847C	aacacaagaacacctctgattactcctgccatcatgaccc[T/C]tggccataatatgatttatctccacactagcagagaccaa
T3866C	ttactcctgccatcatgacccttggccataatatgattta[T/C]ctccacactagcagagaccaaccgaaccccttcgacctt
T3999C	gccccttcgccctattcttcatagccgaatacacaaacat[T/C]attataataaacaccctcaccactacaatcttcctaggaa
T4047C	aaacaccctcaccactacaatcttcctaggaacaacata[T/C]gacgcactctcccctgaactctacacaacatattttgtca
T4062C	ctacaatcttcctaggaacaacatatgacgcactctcccc[T/C]gaactctacacaacatattttgtcaccaagaccctacttc
T4080C	caacatatgacgcactctcccctgaactctacacaacata[T/C]tttgtcaccaagaccctacttctaacctccctgttcttat
T408C	accctaacaccagcctaaccagatttcaaattttatcttt[T/C]ggcggtatgcacttttaacagtcaccccccaactaacaca
T4209C	tcctatgaaaaaacttcctaccactcaccctagcattact[T/C]atatgatatgtctccatacccattacaatctccagcattc
T4216C	aaaaaacttcctaccactcaccctagcattacttatatga[T/C]atgtctccatacccattacaatctccagcattccccctca
T4218C	aaaacttcctaccactcaccctagcattacttatatgata[T/C]gtctccatacccattacaatctccagcattccccctcaaa
T4248C	ttatatgatatgtctccatacccattacaatctccagcat[T/C]ccccctcaaacctaagaaatatgtctgataaaagagttac
T4370C	taggactatgagaatcgaacccatccctgagaatccaaaa[T/C]tctccgtgccacctatcacaccccatcctaaagtaaggtc
T4452C	gctaaataagctatcgggcccataccccgaaaatgttggt[T/C]ataccttcccgtactaattaatcccctggcccaacccgt
T4454C	taaataagctatcgggcccataccccgaaaatgttggtta[T/C]acccttcccgtactaattaatcccctggcccaacccgtca
T4553C	caggcacactcatcacagcgctaagctcgcactgatttttt[T/C]acctgagtaggcctagaaataaacatgctagcttttattc
T4579C	tcgcactgattttttacctgagtaggcctagaaataaaca[T/C]gctagcttttattccagttctaaccaaaaaaataaaccct
T4586C	gatttttacctgagtaggcctagaaataaacatgctagc[T/C]tttattccagttctaaccaaaaaaataaaccctcgttcca
T4639C	ctaaccaaaaaaataaaccctcgttccacagaagctgcca[T/C]caagtatttcctcacgcaagcaaccgcatccataatcctt
T4646C	aaaaaataaaccctcgttccacagaagctgccatcaagta[T/C]ttcctcacgcaagcaaccgcatccataatccttctaatag
T4695C	gcaagcaaccgcatccataatccttctaatagctatcctc[T/C]tcaacaatatactctccggacaatgaaccataaccaatac
T4703C	ccgcatccataatccttctaatagctatcctcttcaacaa[T/C]atactctccggacaatgaaccataaccaatactaccaatc
T4733C	tcttcaacaatatactctccggacaatgaaccataaccaa[T/C]actaccaatcaatactcatcattaataatcataatagcta
T4736C	tcaacaatatactctccggacaatgaaccataaccaatac[T/C]accaatcaatactcatcattaataatcataatagctatag
T4742C	atatactctccggacaatgaaccataaccaatactaccaa[T/C]caatactcatcattaataatcataatagctatagcaataa
T4772C	atactaccaatcaatactcatcattaataatcataatagc[T/C]atagcaataaaactaggaatagccccctttcacttctgag
T482C	taacacattattttcccctcccactcccatactactaatc[T/C]catcaatacaacccccgcccatcctacccagcacacacac
T485C	cacattattttcccctcccactcccatactactaatctca[T/C]caatacaacccccgcccatcctacccagcacacacacacc

T489C	ttattttcccctcccactcccatactactaatctcatcaa[T/C]acaacccccgcccatcctacccagcacacacacaccgctg
T4928C	caatcatataccaaatctctccctcactaaacgtaagcct[T/C]ctcctcactctctcaatcttatccatcatagcaggcagttg
T4977C	tctctcaatcttatccatcatagcaggcagttgaggtgga[T/C]taaaccaaacccagctacgcaaaatcttagcatactcctc
T5004C	cagttgaggtggattaaaccaaacccagctacgcaaaatc[T/C]tagcatactcctcaattacccacataggatgaataatagc
T5048C	catactcctcaattacccacataggatgaataatagcagt[T/C]ctaccgtacaaccctaacataaccattcttaatttaacta
T5081C	tagcagttctaccgtacaaccctaacataaccattcttaa[T/C]ttaactatttatattatcctaactactaccgcattcctac
T5093C	cgtacaaccctaacataaccattcttaatttaactattta[T/C]attatcctaactactaccgcattcctactactcaacttaa
T5096C	acaaccctaacataaccattcttaatttaactatttatat[T/C]atcctaactactaccgcattcctactactcaacttaaact
T5201C	tcgcacctgaaacaagctaacatgactaacacccttaat[T/C]ccatccaccctcctctccctaggaggcctgcccccgctaa
T5277C	gctaaccggctttttgcccaaatgggccattatcgaagaa[T/C]tcacaaaaaacaatagcctcatcatccccaccatcatagc
T5393C	gcctaatctactccacctcaatcacactactccccatatc[T/C]aacaacgtaaaaataaaatgacagtttgaacatacaaaac
T5426C	ccatatctaacaacgtaaaaataaaatgacagtttgaaca[T/C]acaaaacccaccccattcctccccacactcatcgccctta
T5492C	cactcatcgcccttaccacgctactcctacctatctcccc[T/C]tttatactaataatcttatagaaatttaggttaaatacag
T5495C	tcatcgcccttaccacgctactcctacctatctcccctttt[T/C]atactaataatcttatagaaatttaggttaaatacagacc
T5580C	gccttcaaagccctcagtaagttgcaatacttaatttctg[T/C]aacagctaaggactgcaaaaccccactctgcatcaactga
T55A	caccctattaaccactcacgggagctctccatgcatttgg[T/A]attttcgtctgggggggtatgcacgcgatagcattgcgaga
T5628C	aaggactgcaaaaccccactctgcatcaactgaacgcaaa[T/C]cagccactttaattaagctaagcccttactagaccaatgg
T5655C	aactgaacgcaaatcagccactttaattaagctaagccct[T/C]actagaccaatgggacttaaacccacaaacacttagttaa
T5788C	aaaaaggcgggagaagccccggcaggtttgaagctgcttc[T/C]tcgaatttgcaattcaatatgaaaatcacctcggagctgg
T5814C	tttgaagctgcttcttcgaatttgcaattcaatatgaaaa[T/C]cacctcggagctggtaaaaagaggcctaacccctgtcttt
T5826C	tcttcgaatttgcaattcaatatgaaaatcacctcggagc[T/C]ggtaaaaagaggcctaacccctgtctttagatttacagtc
T5892C	tttagatttacagtccaatgcttcactcagccattttacc[T/C]cacccccactgatgttcgccgaccgttgactattctctac
T593C	caaaccccaaagacacccccacagtttatgtagcttacc[T/C]cctcaaagcaatacactgaaaatgtttagacgggctcaca
T59C	ctattaaccactcacgggagctctccatgcatttggtatt[T/C]tcgtctgggggggtatgcacgcgatagcattgcgagacgct
T6071C	agccaggcaaccttctaggtaacgaccacatctacaacgt[T/C]atcgtcacagcccatgcatttgtaataatcttcttcatag
T6152C	taatacccatcataatcggaggctttggcaactgactagt[T/C]cccctaataatcggtgcccccgatatggcgtttccccgca
T6185C	gactagttcccctaataatcggtgcccccgatatggcgtt[T/C]ccccgcataaacaacataagcttctgactcttacctccct
T6216C	tatggcgtttccccgcataaacaacataagcttctgactc[T/C]tacctccctctctcctactcctgctcgcatctgctatagt
T6221C	cgtttccccgcataaacaacataagcttctgactcttacc[T/C]ccctctctcctactcctgctcgcatctgctatagtggagg
T6293C	tagtggaggccggagcaggaacaggttgaacagtctaccc[T/C]cccttagcagggaactactcccaccctggagcctccgtag
T629C	tacctcctcaaagcaatacactgaaaatgtttagacgggc[T/C]cacatcaccccataaacaaataggtttggtcctagccttt
T6365C	cctccgtagacctaaccatcttctccttacacctagcagg[T/C]gtctcctctatcttaggggccatcaatttcatcacaacaa
T6374C	acctaaccatcttctccttacacctagcaggtgtctcctc[T/C]atcttaggggccatcaatttcatcacaacaattatcaata
T63C	taaccactcacgggagctctccatgcatttggtattttcg[T/C]ctgggggggtatgcacgcgatagcattgcgagacgctggagc
T6413C	atcttaggggccatcaatttcatcacaacaattatcaa[T/C]ataaaacccccctgccataacccaataccaaacgcccctct
T6515C	cagcagtcctacttctcctatctctcccagtcctagctgc[T/C]ggcatcactatactactaacagaccgcaacctcaacacca
T656C	tgtttagacgggctcacatcaccccataaacaaataggtt[T/C]ggtcctagcctttctattagctcttagtaagattacacat
T6719C	cttactactccggaaaaaaagaaccatttggatacatagg[T/C]atggtctgagctatgatatcaattggcttcctagggttta
T6776C	tatcaattggcttcctagggtttatcgtgtgagcacacca[T/C]atatttacagtaggaatagacgtagacacacgagcatatt
T680C	cataaacaaataggtttggtcctagcctttctattagctc[T/C]tagtaagattacacatgcaagcatccccgttccagtgagt
T681C	ataaacaaataggtttggtcctagcctttctattagctct[T/C]agtaagattacacatgcaagcatccccgttccagtgagtt
T6956C	tgctctgagccctaggattcatctttcttttcaccgtagg[T/C]ggcctgactggcattgtattagcaaactcatcactagaca
T7022C	actcatcactagacatcgtactacacgacacgtactacgt[T/C]gtagcccacttccactatgtcctatcaataggagctgtat
T710C	ctattagctcttagtaagattacacatgcaagcatccccg[T/C]tccagtgagttcaccctctaaatcaccacgatcaaaagga
T7175C	aaatccatttcactatcatattcatcggcgtaaatctaac[T/C]ttcttcccacaacactttctcggcctatccggaatgcccc
T721C	tagtaagattacacatgcaagcatccccgttccagtgagt[T/C]caccctctaaatcaccacgatcaaaaggaacaagcatcaa
T7258C	cgttactcggactaccccgatgcatacaccacatgaaaca[T/C]cctatcatctgtaggctcattcatttctctaacagcagta

T7270C	taccccgatgcatacaccacatgaaacatcctatcatctg[T/C]aggctcattcatttctctaacagcagtaatattaataatt
T72C	acgggagctctccatgcatttggtattttcgtctgggggg[T/C]atgcacgcgatagcattgcgagacgctggagccggagcac
T7309C	gtaggctcattcatttctctaacagcagtaatattaataa[T/C]tttcatgatttgagaagccttcgcttcgaagcgaaaagtc
T7621C	tcttaatggcacatgcagcgcaagtaggtctacaagacgc[T/C]acttcccctatcatagaagagcttatcacctttcatgatc
T7624A	taatggcacatgcagcgcaagtaggtctacaagacgctac[T/A]tcccctatcatagaagagcttatcacctttcatgatcacg
T7645C	taggtctacaagacgctacttcccctatcatagaagagct[T/C]atcacctttcatgatcacgccctcataatcattttcctta
T7660C	ctacttcccctatcatagaagagcttatcacctttcatga[T/C]cacgccctcataatcattttccttatctgcttcctagtcc
T7705C	ccctcataatcattttccttatctgcttcctagtcctgta[T/C]gcccttttcctaacactcacaacaaaactaactaatacta
T7711C	taatcattttccttatctgcttcctagtcctgtatgccct[T/C]ttcctaacactcacaacaaaactaactaatactaacatct
T7759C	cactcacaacaaaactaactaatactaacatctcagacgc[T/C]caggaaatagaaaccgtctgaactatcctgcccgccatca
T7961C	ctacggcggactaatcttcaactcctacatacttccccca[T/C]tattcctagaaccaggcgacctgcgactccttgacgttga
T7999C	cattattcctagaaccaggcgacctgcgactccttgacgt[T/C]gacaatcgagtagtactcccgattgaagcccccattcg
T8227C	gagcaaaccacagtttcatgcccatcgtcctagaattaat[T/C]cccctaaaaatctttgaaatagggcccgtatttaccctat
T825A	cgcttagcctagccacacccccacgggaaacagcagtgat[T/A]aacctttagcaataaacgaaagtttaactaagctatacta
T8383C	aaccaacacctctttacagtgaaatgccccaactaaatac[T/C]accgtatggcccaccataattacccccatactccttacac
T8419C	atactaccgtatggcccaccataattacccccatactcct[T/C]acactattcctcatcacccaactaaaaatattaaacacaa
T8448C	cccatactccttacactattcctcatcacccaactaaaaa[T/C]attaaacacaaactaccacctacctccctcaccaaagccc
T8450C	catactccttacactattcctcatcacccaactaaaaata[T/C]taaacacaaactaccacctacctccctcaccaaagcccat
T8503C	accacctacctccctcaccaaagcccataaaaataaaaaa[T/C]tataacaaaccctgagaaccaaaatgaacgaaaatctgt
T8602C	cacaatcctaggcctacccgccgcagtactgatcattcta[T/C]ttcccctctattgatccccacctccaaatatctcatcaa
T8618C	tacccgccgcagtactgatcattctatttcccctctattga[T/C]ccccacctccaaatatctcatcaacaaccgactaatca
T8705C	caatgactaatcaaactaacctcaaaacaaatgataacca[T/C]acacaacactaaaggacgaacctgatctcttatactagta
T8843C	acaccaaccacccaactatctataaacctagccatggcca[T/C]cccccttatgagcgggcacagtgattataggctttcgctct
T8877C	tggccatccccttatgagcgggcacagtgattataggctt[T/C]cgctctaagattaaaaatgccctagcccacttcttaccac
T8928C	ttaaaaatgccctagcccacttcttaccacaaggcacacc[T/C]acaccccttatccccatactagttattatcgaaaccatca
T8937C	ccctagcccacttcttaccacaaggcacacctacacccct[T/C]atccccatactagttattatcgaaaccatcagcctactca
T89C	atttggtattttcgtctgggggggtatgcacgcgatagcat[T/C]gcgagacgctggagccggagcaccctatgtcgcagtatct
T9070G	cctactcatgcacctaattggaagcgccaccctagcaata[T/G]caaccattaaccttccctctacacttatcatcttcacaat
T9077C	atgcacctaattggaagcgccaccctagcaatatcaacca[T/C]taaccttccctctacacttatcatcttcacaattctaatt
T9083C	ctaattggaagcgccaccctagcaatatcaaccattaacc[T/C]tccctctacacttatcatcttcacaattctaattctactg
T9090C	gaagcgccaccctagcaatatcaaccattaaccttccctc[T/C]acacttatcatcttcacaattctaattctactgactatcc
T9101C	ctagcaatatcaaccattaaccttccctctacacttatca[T/C]cttcacaattctaattctactgactatcctagaaatcgct
T9103C	agcaatatcaaccattaaccttccctctacacttatcatc[T/C]tcacaattctaattctactgactatcctagaaatcgctgt
T9117C	ttaaccttccctctacacttatcatcttcacaattctaat[T/C]ctactgactatcctagaaatcgctgtcgccttaatccaag
T9128C	tctacacttatcatcttcacaattctaattctactgacta[T/C]cctagaaatcgctgtcgccttaatccaagcctacgttttc
T9165C	ctatcctagaaatcgctgtcgccttaatccaagcctacgt[T/C]ttcacacttctagtaagcctctacctgcacgacaacacat
T9311C	tctcagccctcctaatgacctccggcctagccatgtgatt[T/C]cacttccactccataacgctcctcatactaggcctactaa
T9386C	tactaaccaacacactaaccatataccaatgatggcgcga[T/C]gtaacacgagaaagcacataccaaggccaccacacaccac
T9509C	cctcagaagtttttttcttcgcaggatttttctgagcctt[T/C]taccactccagcctagcccctacccccaattaggagggc
T9540C	ctgagccttttaccactccagcctagcccctacccccaa[T/C]taggagggcactggcccccaacaggcatcaccccgctaaa
T9581C	taggagggcactggccccgaacaggcatcaccccgctaaa[T/C]cccctagaagtcccactcctaaacacatccgtattactcg
T9647C	acatccgtattactcgcatcaggagtatcaatcacctgagc[T/C]caccatagtctaatagaaaacaaccgaaaccaaataattc
T9656C	tactcgcatcaggagtatcaatcacctgagctcaccatag[T/C]ctaatagaaaacaaccgaaaccaaataattcaagcactgc
T9682C	tgagctcaccatagtctaatagaaaacaaccgaaaccaaa[T/C]aattcaagcactgcttattacaattttactgggtctctat
T9698C	taatagaaaacaaccgaaaccaaataattcaagcactgct[T/C]attacaattttactgggtctctattttaccctcctacaag
T9716C	accaaataattcaagcactgcttattacaattttactggg[T/C]ctctattttaccctcctacaagcctcagagtacttcgagt
T9758C	tctattttaccctcctacaagcctcagagtacttcgagtc[T/C]cccttcaccatttccgacggcatctacggctcaacatttt

T980C	agagtgttttagatcaccccctccccaataaagctaaaac[T/C]cacctgagttgtaaaaaactccagttgacacaaaatagac
T9833C	catttttgtagccacaggcttccacggacttcacgtcat[T/C]attggctcaactttcctcactatctgcttcatccgccaac
T9836C	tttttgtagccacaggcttccacggacttcacgtcattat[T/C]ggctcaactttcctcactatctgcttcatccgccaactaa
T9861C	acttcacgtcattattggctcaactttcctcactatctgc[T/C]tcatccgccaactaatatttcactttacatccaaacatca
T9899C	gcttcatccgccaactaatatttcactttacatccaaaca[T/C]cactttggcttcgaagccgccgcctgatactggcattttgt
T9938C	atcactttggcttcgaagccgccgcctgatactggcattt[T/C]gtagatgtggtttgactatttctgtatgtctccatctatt
T9944C	ttggcttcgaagccgccgcctgatactggcattttgtaga[T/C]gtggtttgactatttctgtatgtctccatctattgatgag
T9977C	ttgtagatgtggtttgactatttctgtatgtctccatcta[T/C]tgatgagggtcttactcttttagtataaatagtaccgtta

Table 2: SNP sequences were chosen based on published data [18, 29-31, 33]. The SNP ID refers to the base position within the rCRS. The base to the left of the ID number is the reference base, and the base on the right of the ID number is the mutation. The primer sequences include 40 bases up and downstream of the SNP.

A SNP Score File was generated by Illumina for each SNP designed for the array. This file includes: a SNP Score, Designability Rank, Failure Codes, and Validation Class [54]. The SNP Score reflects the design of each polymorphism and its predicted performance within an assay based on a scale of 0.000-1.000, with 1.000 being the most successful. GoldenGate®-validated SNPs with original assay oligos on file with Illumina will have SNP scores of 1.1. The Designability Rank estimates the success rate of each SNP and its risk to the oligonucleotide pool (OPA) based on a scale of 0 to 1. A value less than 0.4 has a low predicted success rate and is a high risk to OPA, a value of 0.5 has a moderate predicted success rate and moderate risk to OPA, and a value greater than 0.6 has a high predicted success rate and a low risk to OPA. Every SNP designed for my HHMI array was assigned a SNP Score and Designability Rank of 0. There are seven failure codes Illumina can also assign to a SNP. Each SNP within my array contained failure codes 2 and 5. Failure code 2 appears when a SNP is in a duplicated or repetitive region, and failure code 5 appears when there are concerns about multiple contributing issues. Other failure codes include: a melting temperature outside the assay limits, tri- or quad-allelic SNPs, and SNPs within 60 nucleotides of each other. Validation class represents the status of each SNP based on previously collected data; a SNP is

GoldenGate™-validated if it successfully generates polymorphic results on an Illumina platform. All SNPs generated for my array were non-validated, meaning they had not been previously assayed by Illumina and are potentially monomorphic. These mediocre predictions for array performance are mainly due to the novel use of mitochondrial DNA for a GoldenGate™ assay since the Illumina platform is designed for the diploid human genome.

3.2 Estimated Maternal Background

Participant ID	Suspected Maternal Background
TB	Poland
CL	England and Ireland
RDW	Northern Norway, Jewish (Poland, German), Swedish/Danish
HH	Northwestern Germany, Netherlands
RBH	Southeast China, Hong Kong
MA	Poland, Eastern European
DR	Central Italy
JT	Western European: England, Ireland, Germany
SMW	Eastern European: Poland, Ukraine
KB	Canada, Scotland
AF	Russia
CC	Southeast China, Hong Kong
DRW	Eastern Europe

Table 3: Estimated maternal background of each participant was gathered for the most distant female ancestor known.

Based on the estimated maternal background of the participants, predictions can be made concerning their potential haplogroups [17]. Following analysis with the Illumina array, comparisons can be made between the experimental results and these predictions.

3.3 Identification of mtDNA

Mitochondrial DNA was identified by PCR and gel electrophoresis. Two primer pairs were selected to amplify 800-900 bp regions of the mitochondrial genome [52]. The sequences of these primers, their length and position according to the rCRS are presented below in Table 4.

Table 4 – Primer Pair Selection for PCR

Primer Name	Primer Sequence (5'-3')	3' position	Length
4F	AAATCTTACCCCGCCTGTTT	2499	887
4R	AGGAATGCCATTGCGATTAG	3346	
17F	TCACTCTCACTGCCCAAGAA	11314	802
17R	GGAGAATGGGGGATAGGTGT	10276	

Table 4: Names and primer sequences of the two sets of primer pairs used to specifically amplify the mtDNA genome are represented as well as the position within the genome relative to the rCRS and the length of the PCR amplicon [52].

Two different samples for each participant were examined: the purified DNA following the DNA Genotek protocol (un-amplified) and the amplified DNA following the REPLI-g Amplification (See Sections 2.3 and 2.4).

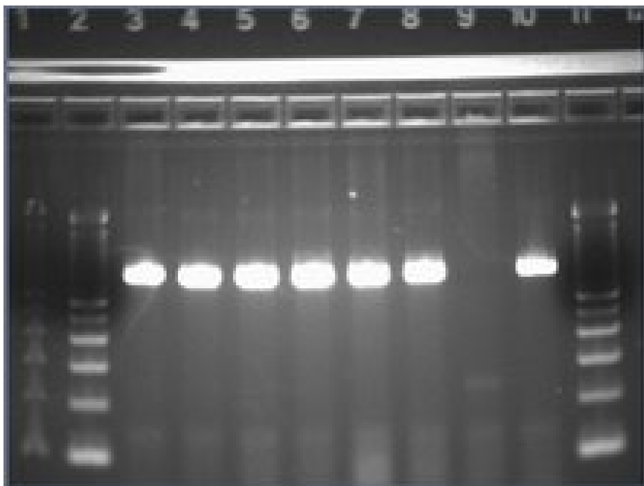

Figure 3a – Un-amplified mtDNA with Primer Pair 17, Samples 1-8

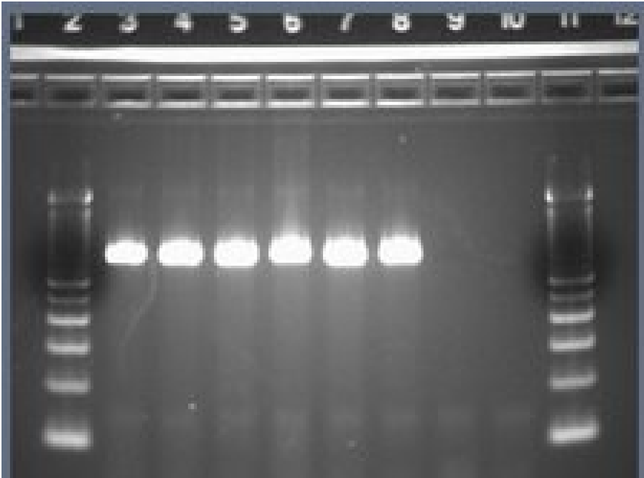

Figure 3b – Un-amplified mtDNA with Primer Pair 17, Samples 9-14
Figure 3: Fourteen saliva samples were collected and the DNA purified following the DNA Genotek Oragene™ Self-Collection Kit protocol. The whole-genome purified DNA was run through PCR using mtDNA-specific primer pair 17 [52] to generate an 800 base-pair amplicon. Five µL aliquots of PCR product were loaded in a 2% agarose gel run for 30 min at 60V. The Invitrogen 100 bp TrackIt™ ladder was used to identify amplicon length. Lanes 3-10 in Figure 2a, and 3-8 in 2b contain the 14 saliva samples. Lanes 9 and 10 on Figure 2b were run as negative controls with buffer only.

Thirteen of the fourteen samples have intense bands in the position corresponding to a size of 800 base pairs indicating the presence of mtDNA in these samples. Sample RDW was the exception and did not present a band in the gel.

3.4 mtDNA Amplification

The Oragene™ Self-Collection Protocol isolates both genomic and mitochondrial DNA. Although there are hundreds more copies of the mitochondrial genome present per cell, the size of the nuclear DNA (three million kilobases compared to sixteen kilobases) can overwhelm, and therefore interfere with, the downstream analysis of mitochondrial DNA. Typical DNA preparations contain around 0.1% mitochondrial DNA, which would lead to an estimate of 0.11 µg in our 110 µg DNA preparation. In order to increase the concentration of the mitochondrial DNA in the sample, the Qiagen REPLI-g

Mitochondrial DNA Kit was used. This kit is designed to specifically amplify mtDNA present in whole DNA samples using Multiple Displacement Amplification (MDA), which is much less cumbersome than performing a mtDNA purification. The typical yield from this amplification is 3-5 μg per 50 μL reaction or an estimated 2×10^9 mitochondrial genome copies per μL – a significantly large amount for sensitive downstream applications – with an average product length greater than ten kilobases. This preparation takes advantage of a uniquely-processive DNA polymerase which carries out isothermal gene amplification, including a 3'→ 5' exonuclease proofreading activity in the presence of exonuclease-resistant primers, to obtain a large DNA yield.

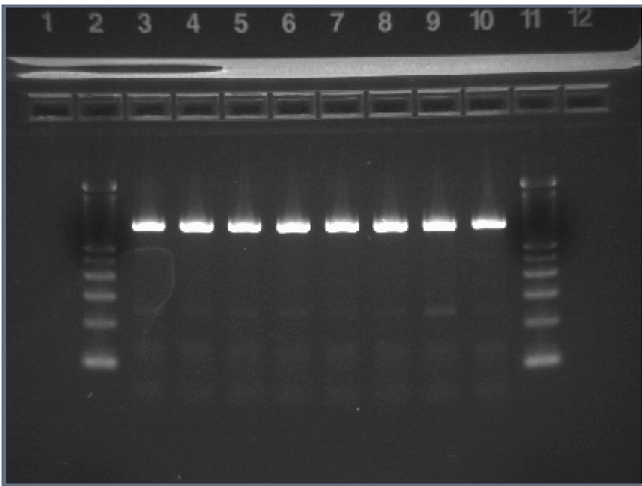

Figure 4a – Amplified mtDNA with Primer Pair 17, Samples 1-8

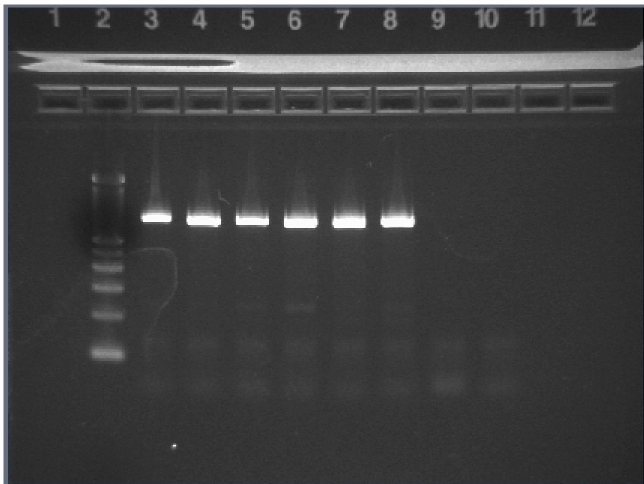

Figure 4b – Amplified mtDNA with Primer Pair 17, Samples 9-14
Figure 4: Fourteen saliva samples were collected and the DNA purified following the DNA Genotek Oragene™ Self-Collection Kit protocol. The whole-genome purified DNA was run through the Qiagen REPLI-g mtDNA Amplification kit to selectively amplify mtDNA. This amplified sample was run through PCR using mtDNA-specific primer pair 17 [52] to generate an 800 base-pair amplicon. Five µL aliquots of PCR product were loaded in a 2% agarose gel run for 30 min at 60V. The Invitrogen 100 bp TrackIt™ ladder was used to identify amplicon length. Lanes 3-10 in Figure 2a, and 3-8 in 2b contain the 14 saliva samples. Lanes 9 and 10 on Figure 2b were run as negative controls with buffer only.

All 14 samples have bands representing 800 base pairs in length, which can be associated with the presence of mtDNA. These samples were diluted 1:1000 prior to PCR since the high yield of mtDNA from the amplification protocol caused PCR inhibition.

Using these results, the presence of mtDNA within each saliva sample was confirmed.

3.5 SNP Calls

The SNP calls made by Illumina's software BeadStudio are assigned based on the comparison of the intensity values from the Cy3 and Cy5 fluorescent dyes. Each of these dyes represents either the natural base or the mutation, and by comparison of the intensity

values the nucleotide base present at a given locus can be determined. BeadStudio, however, is designed for diploid organisms such as the human genome, and because of the haploid nature of the mitochondrial DNA, the data output from the BeadChips was manually clustered using Microsoft Excel. After evaluating the theta values at 461 loci to determine the mutations that each sample contained, numerous false haplogroup designations for each individual were seen compared to the expected results (Appendix, Table 1).

Upon evaluation of the raw intensity values, it was determined that, most likely the assay had failed. The inherent controls present with the Illumina array showed that although the first and second hybridization steps appeared to work, the extension, ligation, and PCR steps did not. Although clustering of the data is possible based on the ratios generated by the fluorescent signals, the intensity of the signals was so far below the threshold that the data generated is insignificant and can be considered background noise.

Table 5: Illumina Control Dashboard

Category	Sample_Name	Theta	R
Allele Specific Extension	1763536022_A1	0.528	47
Allele Specific Extension	1763536022_A2	0.521	42
Allele Specific Extension	1763536022_B1	0.52	45
Allele Specific Extension	1763536022_B2	0.52	45
Allele Specific Extension	1763536022_C1	0.552	43
Allele Specific Extension	1763536022_C2	0.614	47
Allele Specific Extension	1763536022_D1	0.585	48
Allele Specific Extension	1763536022_D2	0.5	44
Allele Specific Extension	1763536022_E1	0.553	51
Allele Specific Extension	1763536022_E2	0.554	42
Allele Specific Extension	1763536022_F1	0.622	52
Allele Specific Extension	1763536022_F2	0.52	45
Allele Specific Extension	1763536022_G1	0.56	45
Allele Specific Extension	1763536022_G2	0.521	42
Allele Specific Extension	1763536022_H1	0.556	48
Allele Specific Extension	1763536022_H2	0.607	47

Allele Specific Extension	1763536022_A1	0.532	42
Allele Specific Extension	1763536022_A2	0.511	42
Allele Specific Extension	1763536022_B1	0.552	43
Allele Specific Extension	1763536022_B2	0.549	46
Allele Specific Extension	1763536022_C1	0.584	43
Allele Specific Extension	1763536022_C2	0.56	45
Allele Specific Extension	1763536022_D1	0.57	45
Allele Specific Extension	1763536022_D2	0.561	44
Allele Specific Extension	1763536022_E1	0.611	49
Allele Specific Extension	1763536022_E2	0.566	41
Allele Specific Extension	1763536022_F1	0.585	48
Allele Specific Extension	1763536022_F2	0.512	39
Allele Specific Extension	1763536022_G1	0.558	47
Allele Specific Extension	1763536022_G2	0.607	47
Allele Specific Extension	1763536022_H1	0.573	43
Allele Specific Extension	1763536022_H2	0.521	42
Allele Specific Extension	1763536022_A1	0.614	47
Allele Specific Extension	1763536022_A2	0.5	44
Allele Specific Extension	1763536022_B1	0.582	44
Allele Specific Extension	1763536022_B2	0.588	46
Allele Specific Extension	1763536022_C1	0.579	46
Allele Specific Extension	1763536022_C2	0.582	44
Allele Specific Extension	1763536022_D1	0.593	44
Allele Specific Extension	1763536022_D2	0.49	45
Allele Specific Extension	1763536022_E1	0.641	45
Allele Specific Extension	1763536022_E2	0.605	43
Allele Specific Extension	1763536022_F1	0.575	48
Allele Specific Extension	1763536022_F2	0.57	45
Allele Specific Extension	1763536022_G1	0.542	43
Allele Specific Extension	1763536022_G2	0.575	42
Allele Specific Extension	1763536022_H1	0.56	45
Allele Specific Extension	1763536022_H2	0.556	40
Allele Specific Extension	1763536022_A1	0.607	47
Allele Specific Extension	1763536022_A2	0.455	40
Allele Specific Extension	1763536022_B1	0.539	47
Allele Specific Extension	1763536022_B2	0.546	49
Allele Specific Extension	1763536022_C1	0.55	45
Allele Specific Extension	1763536022_C2	0.545	50
Allele Specific Extension	1763536022_D1	0.556	40
Allele Specific Extension	1763536022_D2	0.5	42
Allele Specific Extension	1763536022_E1	0.611	49
Allele Specific Extension	1763536022_E2	0.573	43
Allele Specific Extension	1763536022_F1	0.575	48
Allele Specific Extension	1763536022_F2	0.57	45
Allele Specific Extension	1763536022_G1	0.51	46
Allele Specific Extension	1763536022_G2	0.53	45
Allele Specific Extension	1763536022_H1	0.55	45

Allele Specific Extension	1763536022_H2	0.51	43
PCR Uniformity	1763536022_A1	0.56	38
PCR Uniformity	1763536022_A2	0.556	40
PCR Uniformity	1763536022_B1	0.552	43
PCR Uniformity	1763536022_B2	0.542	43
PCR Uniformity	1763536022_C1	0.607	47
PCR Uniformity	1763536022_C2	0.593	44
PCR Uniformity	1763536022_D1	0.59	40
PCR Uniformity	1763536022_D2	0.622	45
PCR Uniformity	1763536022_E1	0.663	48
PCR Uniformity	1763536022_E2	0.622	45
PCR Uniformity	1763536022_F1	0.601	49
PCR Uniformity	1763536022_F2	0.57	45
PCR Uniformity	1763536022_G1	0.542	43
PCR Uniformity	1763536022_G2	0.61	45
PCR Uniformity	1763536022_H1	0.554	42
PCR Uniformity	1763536022_H2	0.59	45
PCR Uniformity	1763536022_A1	0.549	46
PCR Uniformity	1763536022_A2	0.554	42
PCR Uniformity	1763536022_B1	0.52	45
PCR Uniformity	1763536022_B2	0.5	40
PCR Uniformity	1763536022_C1	0.541	44
PCR Uniformity	1763536022_C2	0.614	44
PCR Uniformity	1763536022_D1	0.5	42
PCR Uniformity	1763536022_D2	0.561	44
PCR Uniformity	1763536022_E1	0.556	40
PCR Uniformity	1763536022_E2	0.531	43
PCR Uniformity	1763536022_F1	0.529	46
PCR Uniformity	1763536022_F2	0.476	38
PCR Uniformity	1763536022_G1	0.539	47
PCR Uniformity	1763536022_G2	0.542	43
PCR Uniformity	1763536022_H1	0.573	43
PCR Uniformity	1763536022_H2	0.529	46
Gender	1763536022_A1	0.602	44
Gender	1763536022_A2	0.512	39
Gender	1763536022_B1	0.521	44
Gender	1763536022_B2	0.582	44
Gender	1763536022_C1	0.614	47
Gender	1763536022_C2	0.618	46
Gender	1763536022_D1	0.5	38
Gender	1763536022_D2	0.566	41
Gender	1763536022_E1	0.609	42
Gender	1763536022_E2	0.61	45
Gender	1763536022_F1	0.633	51
Gender	1763536022_F2	0.561	44
Gender	1763536022_G1	0.49	43
Gender	1763536022_G2	0.561	44

Gender	1763536022_H1	0.602	44
Gender	1763536022_H2	0.582	44
Gender	1763536022_A1	0.46	45
Gender	1763536022_A2	0.469	43
Gender	1763536022_B1	0.474	52
Gender	1763536022_B2	0.471	46
Gender	1763536022_C1	0.465	51
Gender	1763536022_C2	0.575	48
Gender	1763536022_D1	0.49	46
Gender	1763536022_D2	0.481	48
Gender	1763536022_E1	0.541	55
Gender	1763536022_E2	0.544	52
Gender	1763536022_F1	0.454	58
Gender	1763536022_F2	0.491	50
Gender	1763536022_G1	0.5	44
Gender	1763536022_G2	0.519	48
Gender	1763536022_H1	0.481	48
Gender	1763536022_H2	0.444	48
Extension Gap	1763536022_A1	0.528	47
Extension Gap	1763536022_A2	0.511	42
Extension Gap	1763536022_B1	0.527	50
Extension Gap	1763536022_B2	0.566	48
Extension Gap	1763536022_C1	0.552	43
Extension Gap	1763536022_C2	0.545	50
Extension Gap	1763536022_D1	0.549	46
Extension Gap	1763536022_D2	0.554	50
Extension Gap	1763536022_E1	0.609	54
Extension Gap	1763536022_E2	0.51	43
Extension Gap	1763536022_F1	0.509	50
Extension Gap	1763536022_F2	0.49	46
Extension Gap	1763536022_G1	0.522	41
Extension Gap	1763536022_G2	0.521	44
Extension Gap	1763536022_H1	0.5	49
Extension Gap	1763536022_H2	0.469	43
First Hybridization	1763536022_A1	0.57	45
First Hybridization	1763536022_A2	0.51	43
First Hybridization	1763536022_B1	0.55	45
First Hybridization	1763536022_B2	0.593	49
First Hybridization	1763536022_C1	0.647	49
First Hybridization	1763536022_C2	0.59	45
First Hybridization	1763536022_D1	0.556	40
First Hybridization	1763536022_D2	0.533	40
First Hybridization	1763536022_E1	0.602	44
First Hybridization	1763536022_E2	0.554	42
First Hybridization	1763536022_F1	0.585	48
First Hybridization	1763536022_F2	0.489	42
First Hybridization	1763536022_G1	0.524	38

First Hybridization	1763536022_G2	0.584	43
First Hybridization	1763536022_H1	0.607	47
First Hybridization	1763536022_H2	0.566	48
First Hybridization	1763536022_A1	0.579	46
First Hybridization	1763536022_A2	0.587	41
First Hybridization	1763536022_B1	0.622	48
First Hybridization	1763536022_B2	0.596	42
First Hybridization	1763536022_C1	0.634	44
First Hybridization	1763536022_C2	0.584	43
First Hybridization	1763536022_D1	0.617	42
First Hybridization	1763536022_D2	0.575	48
First Hybridization	1763536022_E1	0.641	45
First Hybridization	1763536022_E2	0.646	44
First Hybridization	1763536022_F1	0.59	45
First Hybridization	1763536022_F2	0.542	43
First Hybridization	1763536022_G1	0.596	42
First Hybridization	1763536022_G2	0.579	46
First Hybridization	1763536022_H1	0.587	41
First Hybridization	1763536022_H2	0.552	43
Second Hybridization	1763536022_A1	0.001	21160
Second Hybridization	1763536022_A2	0.002	9388
Second Hybridization	1763536022_B1	0.001	23616
Second Hybridization	1763536022_B2	0.001	20961
Second Hybridization	1763536022_C1	0.001	23694
Second Hybridization	1763536022_C2	0.02	1031
Second Hybridization	1763536022_D1	0.001	23680
Second Hybridization	1763536022_D2	0.001	23803
Second Hybridization	1763536022_E1	0.001	24575
Second Hybridization	1763536022_E2	0.001	23016
Second Hybridization	1763536022_F1	0.001	23432
Second Hybridization	1763536022_F2	0.001	15973
Second Hybridization	1763536022_G1	0.001	19284
Second Hybridization	1763536022_G2	0.002	12121
Second Hybridization	1763536022_H1	0.002	11370
Second Hybridization	1763536022_H2	0.001	22928
Second Hybridization	1763536022_A1	0.001	23922
Second Hybridization	1763536022_A2	0.001	15894
Second Hybridization	1763536022_B1	0.001	28676
Second Hybridization	1763536022_B2	0.001	26550
Second Hybridization	1763536022_C1	0.001	27021
Second Hybridization	1763536022_C2	0.008	3054
Second Hybridization	1763536022_D1	0.001	27639
Second Hybridization	1763536022_D2	0.001	27228
Second Hybridization	1763536022_E1	0.001	28828
Second Hybridization	1763536022_E2	0.001	26563
Second Hybridization	1763536022_F1	0.001	26922
Second Hybridization	1763536022_F2	0.001	23043

Second Hybridization	1763536022_G1	0.001	27009
Second Hybridization	1763536022_G2	0.001	17335
Second Hybridization	1763536022_H1	0.001	15350
Second Hybridization	1763536022_H2	0.001	27112
Second Hybridization	1763536022_A1	0.002	14294
Second Hybridization	1763536022_A2	0.002	12237
Second Hybridization	1763536022_B1	0.002	13737
Second Hybridization	1763536022_B2	0.001	14854
Second Hybridization	1763536022_C1	0.002	14739
Second Hybridization	1763536022_C2	0.004	5987
Second Hybridization	1763536022_D1	0.001	15088
Second Hybridization	1763536022_D2	0.001	15335
Second Hybridization	1763536022_E1	0.002	15589
Second Hybridization	1763536022_E2	0.001	15537
Second Hybridization	1763536022_F1	0.001	15472
Second Hybridization	1763536022_F2	0.001	14473
Second Hybridization	1763536022_G1	0.001	14933
Second Hybridization	1763536022_G2	0.002	15160
Second Hybridization	1763536022_H1	0.002	11442
Second Hybridization	1763536022_H2	0.001	16496
Second Hybridization	1763536022_A1	0.002	12022
Second Hybridization	1763536022_A2	0.002	11017
Second Hybridization	1763536022_B1	0.002	11560
Second Hybridization	1763536022_B2	0.002	13088
Second Hybridization	1763536022_C1	0.002	13044
Second Hybridization	1763536022_C2	0.006	3860
Second Hybridization	1763536022_D1	0.002	12650
Second Hybridization	1763536022_D2	0.002	12988
Second Hybridization	1763536022_E1	0.002	13140
Second Hybridization	1763536022_E2	0.002	13252
Second Hybridization	1763536022_F1	0.002	12927
Second Hybridization	1763536022_F2	0.001	12737
Second Hybridization	1763536022_G1	0.002	13074
Second Hybridization	1763536022_G2	0.002	11203
Second Hybridization	1763536022_H1	0.002	10172
Second Hybridization	1763536022_H2	0.001	13239
Second Hybridization	1763536022_A1	0.593	44
Second Hybridization	1763536022_A2	0.51	43
Second Hybridization	1763536022_B1	0.56	45
Second Hybridization	1763536022_B2	0.556	40
Second Hybridization	1763536022_C1	0.579	46
Second Hybridization	1763536022_C2	0.541	44
Second Hybridization	1763536022_D1	0.521	44
Second Hybridization	1763536022_D2	0.56	45
Second Hybridization	1763536022_E1	0.586	53
Second Hybridization	1763536022_E2	0.531	43
Second Hybridization	1763536022_F1	0.561	44

Second Hybridization	1763536022_F2	0.51	43
Second Hybridization	1763536022_G1	0.619	50
Second Hybridization	1763536022_G2	0.528	47
Second Hybridization	1763536022_H1	0.549	46
Second Hybridization	1763536022_H2	0.587	41
Second Hybridization	1763536022_A1	0.53	45
Second Hybridization	1763536022_A2	0.459	44
Second Hybridization	1763536022_B1	0.5	45
Second Hybridization	1763536022_B2	0.517	52
Second Hybridization	1763536022_C1	0.549	46
Second Hybridization	1763536022_C2	0.549	46
Second Hybridization	1763536022_D1	0.521	44
Second Hybridization	1763536022_D2	0.48	45
Second Hybridization	1763536022_E1	0.577	47
Second Hybridization	1763536022_E2	0.556	48
Second Hybridization	1763536022_F1	0.554	50
Second Hybridization	1763536022_F2	0.526	52
Second Hybridization	1763536022_G1	0.572	50
Second Hybridization	1763536022_G2	0.458	43
Second Hybridization	1763536022_H1	0.558	47
Second Hybridization	1763536022_H2	0.541	44
PCR Contamination	1763536022_A1	0.63	42
PCR Contamination	1763536022_A2	0.621	37
PCR Contamination	1763536022_B1	0.596	42
PCR Contamination	1763536022_B2	0.622	45
PCR Contamination	1763536022_C1	0.68	46
PCR Contamination	1763536022_C2	0.614	44
PCR Contamination	1763536022_D1	0.671	43
PCR Contamination	1763536022_D2	0.603	39
PCR Contamination	1763536022_E1	0.614	47
PCR Contamination	1763536022_E2	0.605	43
PCR Contamination	1763536022_F1	0.618	46
PCR Contamination	1763536022_F2	0.532	42
PCR Contamination	1763536022_G1	0.59	40
PCR Contamination	1763536022_G2	0.612	40
PCR Contamination	1763536022_H1	0.617	42
PCR Contamination	1763536022_H2	0.609	42
PCR Contamination	1763536022_A1	0.581	39
PCR Contamination	1763536022_A2	0.549	37
PCR Contamination	1763536022_B1	0.654	44
PCR Contamination	1763536022_B2	0.638	43
PCR Contamination	1763536022_C1	0.643	41
PCR Contamination	1763536022_C2	0.671	43
PCR Contamination	1763536022_D1	0.646	44
PCR Contamination	1763536022_D2	0.612	40
PCR Contamination	1763536022_E1	0.646	44
PCR Contamination	1763536022_E2	0.693	47

PCR Contamination	1763536022_F1	0.634	44
PCR Contamination	1763536022_F2	0.551	35
PCR Contamination	1763536022_G1	0.612	40
PCR Contamination	1763536022_G2	0.661	39
PCR Contamination	1763536022_H1	0.658	43
PCR Contamination	1763536022_H2	0.626	43
PCR Contamination	1763536022_A1	0.658	43
PCR Contamination	1763536022_A2	0.536	38
PCR Contamination	1763536022_B1	0.596	42
PCR Contamination	1763536022_B2	0.603	39
PCR Contamination	1763536022_C1	0.626	43
PCR Contamination	1763536022_C2	0.634	44
PCR Contamination	1763536022_D1	0.56	38
PCR Contamination	1763536022_D2	0.573	43
PCR Contamination	1763536022_E1	0.661	45
PCR Contamination	1763536022_E2	0.646	44
PCR Contamination	1763536022_F1	0.561	44
PCR Contamination	1763536022_F2	0.542	43
PCR Contamination	1763536022_G1	0.596	42
PCR Contamination	1763536022_G2	0.63	42
PCR Contamination	1763536022_H1	0.617	42
PCR Contamination	1763536022_H2	0.556	40
PCR Contamination	1763536022_A1	0.519	47
PCR Contamination	1763536022_A2	0.49	45
PCR Contamination	1763536022_B1	0.518	50
PCR Contamination	1763536022_B2	0.556	48
PCR Contamination	1763536022_C1	0.529	46
PCR Contamination	1763536022_C2	0.554	50
PCR Contamination	1763536022_D1	0.529	46
PCR Contamination	1763536022_D2	0.5	44
PCR Contamination	1763536022_E1	0.52	45
PCR Contamination	1763536022_E2	0.558	47
PCR Contamination	1763536022_F1	0.535	51
PCR Contamination	1763536022_F2	0.459	44
PCR Contamination	1763536022_G1	0.49	43
PCR Contamination	1763536022_G2	0.471	46
PCR Contamination	1763536022_H1	0.5	47
PCR Contamination	1763536022_H2	0.536	50

Table 5: The control values from BeadStudio for the HHMI Illumina array. These controls determine which step(s) failed; these categories are listed in column 1 for each sample. The samples (A1-G2) are represented in column 2. The θ and R values are the polar coordinates of the comparison of raw intensity values.

For each of the control steps in this assay, the only step with significant signal intensities was in the second hybridization step. A significant signal intensity should

have an R value greater than 1,000, and ideally greater than 10,000. The other control steps for my samples have R values less than 100.

3.6 FamilyTree® DNA Results

As a positive control for this experiment, two study participants (JT and SMW) sent buccal swab samples to be analyzed by FamilyTree® DNA. The HVR I and 2 regions were sequenced and the differences were compared to the Cambridge Reference Sequence (CRS)(Table 6, 7). Participants SMW and JT were determined to be from Haplogroups K and W respectively.

Table 6: SNPs for Sample SMW

HVR1 Haplogroup	K
HVR1 differences from CRS	16223T
	16224C
	16234T
	16311C
	16519C
HVR2 differences from CRS	73G
	114T
	263G
	315.1C
	497T

Table 6: FamilyTree® DNA Results for Sample SMW showing haplogroup designation K and the differences within the hypervariable regions (HVR) I and II from the CRS.

Table 7: SNPs for Sample JT

HVR1 Haplogroup	W
HVR1 differences from CRS	16223T
	16292T
	16311C
	16320T
	16519C
HVR2 differences from CRS	73G

	189G
	195C
	204C
	207A
	263G
	309.1C
	315.1C
	554G
	569A

Table 7: FamilyTree® DNA Results for Sample JT showing haplogroup designation W and the differences within the hypervariable regions (HVR) I and II from the CRS.

Haplogroup K is part of the super-haplogroup U found in Europe and dates to approximately 16,000 years ago. It has been suggested that individuals from this haplogroup participated in the pre-Neolithic expansion following the Last Glacial Maximum [17]. Haplogroup K contains multiple closely related lineages and shows a high concentration in Eastern Europe. Haplogroup W is mainly distributed through Western Europe and dates to approximately 25,000 years ago. Similar to haplogroup K, it has been suggested that individuals bearing haplogroup W participated in the population of Europe following the Last Glacial Maximum [17].

Chapter 4: FUTURE WORK

The immediate next step taken on this project will be to identify the problematic portions of the GoldenGate^{TM} array protocol that contributed to the assay's failure. Initial investigation suggests the enzyme necessary to carry out the extension-ligation reaction did not function properly. Collaboration is ongoing with Illumina to assess different possibilities such as incorrect buffer compositions, incorrect incubation times, and insufficient enzyme concentrations. The following assay should be carried out using the Illumina-recommended Titanium Taq Polymerase (Cat. No. 639220, Clontech; Mountain View, CA).

Following successful re-analysis of the samples with the GoldenGate^{TM} assay to determine appropriate SNP identification and haplogroup designations, it would be very helpful to assay more individuals with known maternal backgrounds. In addition, this experiment will help determine which SNP loci provide the most accurate and useful ancestral information. Eventually, in order to transition this project to an undergraduate laboratory, it would be necessary to down-scale the number of SNPs investigated. As a positive control for the GoldenGate^{TM} assay, several mtDNA samples could be fully sequenced to determine which bases are present at each position. Sequencing would assist in determining the accuracy of the SNP calls from the Illumina assay. Also, it would be useful to collaborate with undergraduate students in the Computer Science Department to generate a program to cluster the data and assign haplogroup designations based on the intensity outputs and theta values.

Since the eventual goal of the HHMI Project is to integrate the advanced technology of the Illumina GoldenGate™ microarray into an undergraduate laboratory, it is essential to consider possible modifications. Ideally, a platform will be developed using a small number of SNPs (20-50) with hybridizations visualized by colorimetric or fluorescent signals in order to identify an individual's haplogroup. This approach is being pursued currently in the laboratory.

APPENDIX

Table 1: SNPs for Samples

SNP	A1	A2	B1	B2	C1	C2	D1	D2	E1	E2	F1	F2	G1	G2	H1	H2
T 6185 C			L0	L0	L0		L0	L0			L0			L0	L0	L0
C 3516 A	L0															
A 9347 G							L0		L0							
C 10664 T														L0		
T 10915 C			L0	L0			L0									
A 13276 G	L0	L0	L0	L0			L0	L0			L0				L0	L0
G 1719 A	L0f															
G 1719 A	L3h															
G 1719 A	N1,X															
G 10143 A			L0f	L0f				L0f	L0f		L0f			L0f	L0f	L0f
T 11287 C							L0f				L0f				L0f	
T 15852 C		L0f														
A 12720 G				L0a,d					L0a,d		L0a,d				L0a,d	
A 8566 G	L0a			L0a	L0a		L0a				L0a			L0a		
G 8994 A		L0k, W														
C 10920 T		L0k	L0k												L0k	
T 13020 C			L0k	L0k	L0k		L0k		L0k		L0k			L0k	L0k	
T 13020 C			U2e	U2e	U2e		U2e		U2e		U2e			U2e	U2e	
A 10499 G	L0k		L0k		L0k		L0k				L0k			L0k	L0k	
A 10876 G							L0k				L0k					
A 10876 G							U2e				U2e					
T 11299 C									L0k, K		L0k, K					
A 11653 G	L0k	L0k		L0k	L0k		L0k	L0k	L0k		L0k				L0k	
T 14020 C			L0k	L0k	L0k	L0k	L0k				L0k			L0k	L0k	L0k
T 14182 C			U5b				U5b		U5b		U5b					
T 14182 C			L0k				L0k		L0k		L0k					
T 14371 C	L0k	L0k	L0k	L0k	L0k	L0k	L0k	L0k			L0k			L0k		L0k
T 14374 C	L0k		L0k	L0k			L0k	L0k							L0k	
G 8251 A		L0d														
G 8251 A		M12														
G 8251 A		W														
C 498 T		L0d	L0d	L0d										L0d	L0d	
T 710 C	L1b	L1b	L1b	L1b	L1b	L1b	L1b	L1b	L1b		L1b				L1b	L1b
T 2352 C			L1b				L1b		L1b							
T 2352 C			L3e				L3e		L3e							
G 709 A							B5b		B5b							
G 709 A							L1b		L1b							
G 709 A							R11		R11							
G 709 A							T		T							
T 1738 C	L1b		L1b		L1b		L1b				L1b					
A 5036 G	L1b		L1b	L1b			L1b								L1b	
G 5046 A								L1b								
	A1	**A2**	**B1**	**B2**	**C1**	**C2**	**D1**	**D2**	**E1**	**E2**	**F1**	**F2**	**G1**	**G2**	**H1**	**H2**

T 5655 C	L1b	L1b	L1b	L1b			L1b	L1b	L1b					L1b		
T 12519 C			L1b	L1b			L1b							L1b		
A 14203 G	L1b		L1b	L1b			L1b		L1b		L1b				L1b	
T 6071 C									L1c							
A 12810 G	L1c	L1c	L1c				L1c	L1c	L1c		L1c			L1c	L1c	
A 9072 G				L1c				L1c	L1c		L1c					
A 13485 G	L1c		L1c	L1c	L1c	L1c	L1c	L1c	L1c		L1c				L1c	
C 14911 T											L1c				L1c	
T 3423 C			L5				L5									
C 12432 T						L5										
T 2416 C	L2		L2													
T 7175 C	L2a	L2a	L2a		L2a	L2a	L2a		L2a					L2a		L2a
C 7274 T																L2a
A 7771 G	L2a	L2a	L2a	L2a	L2a		L2a	L2a	L2a							L2a
A 13803 G			L2a	L2a			L2a	L2a			L2a			L2a		
T 15784 C			Z		Z	Z	Z	Z	Z		Z				Z	Z
T 15784 C			L2a1		L2a1	L2a1	L2a1	L2a1	L2a1		L2a1				L2a1	L2a1
G 709 A							W		W							
G 709 A							L2a2		L2a2							
G 15217 A									L2b,c					L2b,c		L2b,c
C 1706 T		L2b														
A 2358 G			L2b			L2b	L2b	L2b	L2b		L2b				L2b	
T 4370 C	L2b		L2b	L2b		L2b		L2b	L2b		L2b			L2b	L2b	L2b
C 5027 T		L2b														
T 5814 C		L2b									L2b					
C 6713 T			L2b				L2b		L2b							L2b
G 8387 A																L2b
A 14059 G				L2b		L2b			L2b							L2b
T 680 C			L2c				L2c				L2c					
C 15849 T														L2c		
C 3594 T						L0/L1/L2										
T 12609 C			L4		L4						L4			L4		
A 13470 G			L4						L4		L4					
A 8701 G	L/M		L/M	L/M		L/M	L/M		L/M		L/M			L/M		L/M
A 10398 G			L/M		L/M		L/M		L/M							
A 10398 G			B5b		B5b		B5b		B5b							
A 10398 G			R11		R11		R11		R11							
A 10398 G			J		J		J		J							
	A1	**A2**	**B1**	**B2**	**C1**	**C2**	**D1**	**D2**	**E1**	**E2**	**F1**	**F2**	**G1**	**G2**	**H1**	**H2**
T 9540 C	L/M	L/M	L/M	L/M			L/M		L/M		L/M				L/M	
G 15301 A			L/M													
G 5773 A		L3b				L3b										
A 10086 G			L3b			L3b	L3b		L3b							
A 15311 G	L3b		L3b	L3b	L3b		L3b	L3b	L3b		L3b			L3b	L3b	L3b
A 15824 G	L3b			L3b			L3b									
G 5147 A		L3d														
T 8618 C	L3d	L3d	L3d	L3d	L3d	L3d	L3d	L3d	L3d		L3d			L3d	L3d	L3d

T 1822 C	L3f						L3f				L3f					
T 3396 C				L3f			L3f									
T 15514 C	L3f		L3f	L3f	L3f			L3f	L3f		L3f			L3f	L3f	
T 9509 C	L3h		L3h	L3h	L3h	L3h	L3h	L3h			L3h				L3h	
A 11590 G	L3h						L3h				L3h			L3h		L3h
T 7645 C	L3i		L3i	L3i	L3i		L3i	L3i	L3i		L3i				L3i	L3i
T 11260 C		L3i		L3i	L3i				L3i		L3i					L3i
C 13967 T								L3i								
A 15758 G	L3i		L3i				L3i	L3i								
T 489 C	M, J		M, J	M, J			M, J				M, J				M, J	
T 14783 C	M		M				M	M	M		M			M		
G 15043 A	M															
C 447 G	M2a						M2a									M2a
T 1780 C							M2a				M2a					
A 8396 G											M2a					
G 4850 A (2)		M3		M3		M3										
C 5178 A		D														D
C 8414 T						D4										
C 2092 T		D1	D1													
T 1107 C			D5	D5					D5		D5					
A 13263 G		C	C				C	C			C				C	
T 9090 C	Z										Z					
T 14470 C				M8a, X			M8a, X				M8a, X					
T 2626 C											M7a					
A 2768 G	Q2						Q2		Q2		Q2				Q2	
	A1	**A2**	**B1**	**B2**	**C1**	**C2**	**D1**	**D2**	**E1**	**E2**	**F1**	**F2**	**G1**	**G2**	**H1**	**H2**
C 4335 T		Q2														
A 663 G			A											A	A	A
A 1736 G	A	A	A	A			A	A	A							A
T 4248 C		A	A	A				A							A	
A 3505 G	W			W	W	W	W	W	W		W				W	W
T 5048 C							W								W	
C 11674 T											W					
A 11947 G			W				W		W		W				W	
T 12414 C			W				W		W		W					W
T 10238 C	N1	N1	N1	N1	N1		N1	N1	N1		N1				N1	
T 10238 C	B4a	B4a	B4a	B4a	B4a		B4a	B4a	B4a		B4a				B4a	
A 13966 G	X		X	X				X	X						X	
G 15927 A						B5b		B5b							B5b	
T 10031 C							R11									
C 11061 T		R11						R11								
A 16381 G				R11					R11							
T 10118 C											P					
A 3203 G														P1a	P1a	
A 4122 G			P1a	P1a	P1a		P1a		P1a		P1a					
C 8859 T					P1a											

A 14890 G			P1a				P1a	P1a								
A 12308 G			U		U	U	U	U	U		U				U	U
C 2218 T								U1a								
A 1811 G	U						U									
A 3720 G							U2e				U2e			U2e		
A 5390 G			U2e	U2e			U2e				U2e					
T 5426 C		U2e									U2e					
C 11197 T			U2e					U2e								
T 11732 C			U2e													
T 13734 C		U2e														
A 15907 G							U2e		U2e							
T 4703 C	U3a	U3a	U3a	U3a				U3a			U3a			U3a		
G 9266 A					U3a											U3a
A 10506 G									U3a							
T 9656 C	U3b			U3b				U3b	U3b		U3b			U3b		U3b
	A1	**A2**	**B1**	**B2**	**C1**	**C2**	**D1**	**D2**	**E1**	**E2**	**F1**	**F2**	**G1**	**G2**	**H1**	**H2**
T 13743 C	U3b		U3b	U3b	U3b		U3b		U3b					U3b	U3b	U3b
T 4646 C						U4					U4					
A 6047 G	U4		U4	U4	U4		U4				U4			U4		U4
T 15693 C							U4		U4		U4					
A 3348 G			U6		U6	U6	U6	U6	U6		U6			U6	U6	U6
T 980 C			U7				U7							U7	U7	
T 10084 C	U7	U7	U7	U7	U7	U7	U7	U7	U7		U7			U7		U7
C 10142 T											U7					
T 9698 C			U8/K	U8/K				U8/K	U8/K							
A 3480 G							K		K							
T 3197 C			U5													
G 9477 A														U5		
T 13617 C											U5					
A 14793 G	U5a	U5a	U5a	U5a	U5a	U5a	U5a	U5a			U5a			U5a	U5a	U5a
A 7768 G				U5b							U5b					
T 2442 C	pHV1		pHV1		pHV1		pHV1	pHV1	pHV1		pHV1				pHV1	pHV1
T 3847 C			pHV1				pHV1		pHV1		pHV1			pHV1	pHV1	pHV1
C 13188 T							pHV1								pHV1	
7028 C	H			H	H		H	H	H		H			H	H	H
C 15904 T					V											
1438 A		H2				H2										
T 4216 C		JT	JT					JT	JT		JT					
A 11251 G		JT	JT	JT	JT		JT	JT	JT					JT	JT	JT
A 12612 G											J					
C 7476 T																J2
G 8697 A	T															
T 10463 C	T		T	T							T			T	T	T
G 14905 A											T					
A 14233 G			T2													
T 16519 C		B	B		B		B		B						B	

Table 1: Manual clustering and analysis of the theta values showed 182 SNPs present in total for the samples assayed. The SNP locus is represented in column 1. The haplogroup designations are shown if the SNP was present in a certain sample. Clearly, there are multiple haplogroup designations per sample, indicating a failed reaction.

References:

1. Collins, F. and V. McKusick, *Implications of the Human Genome Project for Medical Science.* Journal of the American Medical Association, 2001. **285**(5): p. 540-544.
2. Olden, K. and S. WIlson, *Environmental health and genomics: visions and implications.* Nature Reviews Genetics, 2000. **1**: p. 149-154.
3. Collins, F., *Medical and societal consequences of the human genome project.* New England Journal of Medicine, 1999. **341**(1): p. 28-37.
4. Hood, L. and D. Galas, *The digital code of DNA.* Nature, 2003. **421**: p. 444-449.
5. Budowle, B., et al., *Forensics and mitochondrial DNA: Applications, debates and foundations.* Annual Reviews Genomics Human Genetics, 2003. **4**: p. 119-141.
6. *The human genome project*, D.o. Energy, Editor. 2003, The Human Genome Program.
7. Wakeley, J., et al., *The discovery of single-nucleotide polymorphisms -- and inferences about human demographic history.* American Journal of Human Genetics, 2001. **69**: p. 1332-1347.
8. Brown, W.M., *Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis.* Proc. Natl. Acad. Sci. USA, 1980. **77**: p. 3605-3609.
9. Torroni, A., et al., *Harvesting the fruit of the human mtDNA tree.* Trends in Genetics, 2006. **22**(6): p. 339-345.
10. Pakendorf, B. and M. Stoneking, *Mitochondrial DNA and human evolution.* Annual Reviews Genomics Human Genetics, 2005. **6**: p. 165-183.
11. Ingman, M. and U. Gyllensten, *Analysis of the complete human mtDNA genome: methodology and inferences for human evolution.* Journal of Heredity, 2001. **92**(6): p. 454-461.
12. Brown, W.M., et al., *Mitochondrial DNA sequences of primates: tempo and mode of evolution.* Journal of Molecular Evolution, 1982. **18**: p. 225-239.
13. Cann, R., M. Stoneking, and A. Wilson, *Mitochondrial DNA and human evolution.* Nature 1987. **325**: p. 31-36.
14. Awadalla, P., A. Eyre-Walker, and J.M. Smith, *Linkage disequilibrium and recombination in hominid mitochondrial DNA.* Science, 1999. **286**: p. 2524-2525.
15. Kumar, S., et al., *Questioning evidence for recombination in human mitochondrial DNA.* Science, 2000. **288**: p. 1931.
16. Vigilant, L., et al., *African populations and the evolution of human mitochondrial DNA.* Science, 1991. **253**: p. 1503-1507.
17. *The National Genographic Project*. 2007, National Geographic and IBM: Washington D.C.
18. Herrnstadt, C., et al., *Reduced-Median-Network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian and European haplogroups.* American Journal of Human Genetics, 2002. **70**: p. 1152-1171.
19. Huson, D. and D. Bryant, *Application of phylogenetic networks in evolutionary studies.* Molecular Biology and Evolution, 2006. **23**(2): p. 254-267.

20. Nakhleh, L. and L.-S. Wang. *Phylogenetic networks, trees and clusters*. [cited.
21. Maca-Meyer, N., et al., *Major genomic mitochondrial lineages delineate early human expansions.* BMC Genetics, 2001. **2**: p. 13-20.
22. Torroni, A., et al., *Classification of european mtDNAs from an analysis of three european populations.* Genetics, 1996. **144**: p. 1835-1850.
23. Finnila, S., M. Lehtonen, and K. Majamaa, *Phylogenetic network for european mtDNA.* American Journal of Human Genetics, 2001. **68**: p. 1475-1484.
24. Kong, Q.-P., et al., *Phylogeny of east asian mitochondrial DNA lineages inferred from complete sequences.* American Journal of Human Genetics, 2003. **73**: p. 671-676.
25. Ingman, M. and U. Gyllensten, *Mitochondrial genome variation and evolutionary history of Australian and New Guinean aborigines.* Genome Research, 2003. **13**: p. 1600-1606.
26. Ingman, M., et al., *Mitochondrial genome variation and the origin of modern humans.* Nature, 2000. **408**: p. 708-713.
27. Tanaka, M., et al., *Mitochondrial genome variation in eastern Asia and the peopling of Japan.* Genome Research, 2004. **14**: p. 1832-1850.
28. Richards, M., et al., *Phylogeography of mitochondrial DNA in western Europe.* Annals of Human Genetics, 1998. **62**: p. 241-260.
29. Ingman, M. and U. Gyllensten, *mtDB: Human mitochondrial genome database, a resource for population genetics and medical sciences.* Nucleic Acids Research, 2005. **34**: p. D749-D751.
30. Attimonelli, M., et al., *HmtDNA, a human mitochondrial genomic resource based on variability studies supporting population genetics and biomedical research.* BMC Bioinformatics, 2005. **6**(S4): p. 1-9.
31. Ruiz-Pesini, E., et al., *An enhanced MITOMAP with a global mtDNA mutational phylogeny.* Nucleic Acids Research, 2007. **35**: p. D823-D828.
32. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome.* Nature, 1981. **290**: p. 457-465.
33. Andrews, R., et al., *Reanalysis and revision of the cambridge reference sequence for human mitochondrial DNA.* Nature Genetics, 1999. **23**: p. 147.
34. Wallace, D., *Mitochondrial DNA variation in human evolution, degenerative disease, and aging.* American Journal of Human Genetics, 1995. **57**: p. 201-223.
35. Coble, M., et al., *Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians.* International Journal of Legal Medicine, 2004. **118**: p. 137-146.
36. Shriver, M. and R. Kittles, *Genetic ancestry and the search for personalized genetic histories.* Nature Reviews Genetics, 2004. **5**: p. 611-618.
37. *Howard Hughes*. http://en.wikipedia.org/wiki/Howard_hughes 2007 [cited.
38. Brown, P. and P. Broeske, *Howard Hughes: The Untold Story*. 1996, New York: Penguin Books.
39. Belosludtsev, Y., B. Iverson, and S. Lemeshki, *DNA microarrays based on noncovalent oligonucleotide attachment and hybridization in two dimensions.* Analytical Biochemistry, 2001. **292**(2): p. 250-256.
40. Clerici, L., F. Campagnari, and J.F.M. De Rooji, *Preparation of polydeoxynucleotides linked to solid support by coupling CNBr-activated*

cellulose with 5'-NH2-terminated oligo and poly(pdT)'s. Nucleic Acids Research, 1979. **6**(1): p. 247-258.

41. de Rooji, J.F.M., et al., *Synthesis of functionalized DNA fragments suitable for reversible attachment to activated cellulose.* Tetrahedron, 1979. **35**: p. 2913-2926.
42. Watkins, S., *Use and development of microarrays as an education tool and for bacterial identification*, in *Chemistry*. 2007, Tufts University: Medford. p. 1-76.
43. Schena, M., et al., *Quanitative monitoring of gene expression patterns with a complementary DNA microarray.* Science, 1995. **270**(5235): p. 467-470.
44. Lander, E.S., *Array of hope.* Nature Genetics, 1999. **21**(1, Suppl.): p. 3-4.
45. *DNA Microarray Factsheet*, N.C.B. I., Editor. 2004, http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html.
46. Rubenstein, K., *Microarrays: technologies, applications, markets*, in *Life Sciences Reports*, C.H. Associates, Editor. 2004. p. 1-9.
47. Shen, R., et al., *High-throughput SNP genotyping on universal bead arrays.* Mutation Research, 2004. **573**: p. 70-82.
48. Gunderson, K., et al., *A genome-wide scalable SNP genotyping assay using microarray technology.* Nature Genetics, 2005. **37**(5): p. 549-554.
49. Michael, K.L., et al., *Randomly ordered addressable high-density optical sensor arrays.* Analytical Chemistry, 1998. **70**: p. 1242-1248.
50. Gunderson, K., et al., *Decoding Randomly Ordered DNA Arrays.* Genome Research, 2004. **14**: p. 870-877.
51. Cormier, T., et al., *SNP genotyping of saliva DNA using Affymetrix GeneChip targeted genotyping system*, Affymetrix. p. http://www.dnagenotek.com/pdf_files/Affymetrix_SNP_Genotyping.pdf.
52. Rieder, M., et al., *Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome.* Nucleic Acids Research, 1998. **26**(4): p. 967-973.
53. Illumina, *Experienced user cards for GoldenGate protocol.* 2007: p. 1-20.
54. Acierno, J. (2007) *Custom OPA design using the Illumina assay design tool.* Illumina, Inc. **Volume**, 1-35