Mitochondrial DNA Sequencing Report from Argus Biosciences

Results for order 12482

Your mitochondrial DNA has the following polymorphisms: A**73**G T**195**C A**263**G **315**insC C**497**T **524**insAC A**750**G T**1189**C A**1438**G A**1811**G A**2706**G A**3480**G A**4769**G G**5460**A G**6182**A C**7028**T A**7245**G A**8860**G G**9055**A T**9148**C T**9698**C A**10398**G A**10550**G T**11299**C A**11467**G G**11719**A T**12235**C A**12308**G G**12372**A C**14167**T C**14766**T T**14798**C A**15326**G G**15930**A T**16093**C C**16179**T T**16224**C T**16311**C T**16519**C. Based on these polymorphisms, you and your maternal ancestors belong to haplogroup K1a.

Methods

Mitochondrial DNA comprising the coding region and the control region was amplified by polymerase chain reaction and the fragments were purified for DNA sequencing using an Applied Biosystems Genetic Analyzer. The DNA sequence was compared to the revised Cambridge Reference Sequence (rCRS) using the program GEN-SNiP. GEN-SNiP identifies bases where your mtDNA differs from the reference sequence and prints these differences (polymorphisms). The list of polymorphisms found in your DNA was compared to a database of polymorphisms that are diagnostic for various haplogroups. You share with other members of your haplogroup a common maternal lineage. Haplogroups are often associated with diverse geographical regions, reflecting the migration patterns of our ancient ancestors.

Genomic map of human mitochondrial DNA

The genome is divided into the control region and the coding region. The control region regulates transcription (DNA to RNA) and DNA replication (making DNA copies). The coding region contains DNA sequence used to make proteins or RNA. There are 13 protein coding genes (Cytb, ND1 to 8, Cox1 to 3, ATP6 and 8), two ribosomal RNAs (12S and 16S), and 22 tRNAs (single capital letters). Numbering is counter-clockwise from base number 1, in the middle of the control region, to base number 16,569.

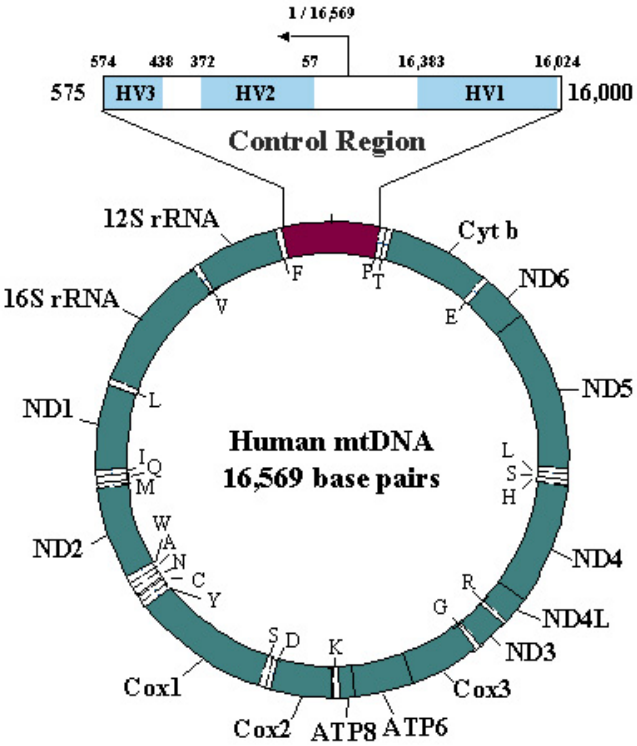


Table of Polymorphisms

Polymorphisms are specific spots in your DNA that differ from a standard reference sequence. A263G indicates a change from A to G at position 263, for example. The list of polymorphisms captures all of the relevant information in the DNA sequence, and has the advantage that it is much easier to work with than a long string of nucleotides (A, T, G & C). The presence of a specific polymorphism or set of polymorphisms determines which haplogroup your mtDNA belongs in.

DNA codes for proteins. Three DNA letters (ACG or T) code for a single amino acid. Sometimes a change in the DNA is "silent", meaning that the protein sequence is identical in both forms of the gene. For example, both CTA and CTG (a single A to G substitution polymorphism) instruct the cellular protein synthesis machinery to add the amino acid leucine. Thus, the DNA change - CTA to CTG - is "silent" at the protein level. "Non-silent" changes in the DNA result in altered protein sequence: AAA codes for the amino acid lysine, for example, while GAA (another A to G substitution polymorphism) codes for glutamic acid. This change in one amino acid may - or may not - have a dramatic effect on the function of the protein.

A change of a base from an A to a G, or C to T, etc, is called a substitution: one base is substituted with another. In addition to substitutions, your DNA may have short insertions and deletions. Insertions are indicted using the notation “ins”, deletions as “del”.

The table lists polymorphisms found in your mtDNA, their diagnostic use in mtDNA phylogeny, their prevalence, and the region of the genome in which they occur. For full length sequence, the table includes changes in amino acid sequence.

- The first column lists the polymorphisms found in your mtDNA.

- The next column shows the frequency with which your polymorphism is found in over 2400 published mtDNA genomes (www.genpat.uu.se/mtDB/). The percentage is based on a non-representative population of mtDNAs, so it should be used only as a rough guide for the frequency of a given polymorphism. Polymorphisms that appear to be very common, such as A263G, found in over 99% of mtDNAs, actually represent a rare polymorphism in the reference sequence.

- The “Location” column identifies the region of the mitochondrial genome that the polymorphism maps to. If you ordered hypervariable region sequencing, the polymorphisms will map to the D-Loop, which resides in the control region. The location of mitochondrial genes on the circular genome can be found on the genomic map on the first page of this report.

- The “Codon” column identifies which codon is altered by polymorphisms in coding regions. Codons are three letter “words” that direct the addition of a

specific amino acid to the growing protein polypeptide. If, for example, the polymorphism occurs at nucleotide number 12 in the protein coding region, it occurs in codon number 4 (codon 4 covers nucleotides 10, 11 and 12). Most polymorphisms are "silent", i.e. they do not change the type of amino acid that is incorporated into the protein.

- The "Amino Change" column lists the effect of the polymorphism on the protein coding sequence.

- The "Haplogroup Assignment" column shows the role of each polymorphism in determining your haplogroup. The more similar your mtDNA is to the reference sequence - which is in haplogroup H - the fewer the polymorphisms. Many haplogroup H members will thus not have diagnostic polymorphisms.

- The academic publications used to determine haplogroups are noted in this column (Kiv_06, etc) and are listed in the appendix and on our website at http://argusbio.com/papers.html.

Polymorphism	Prevalence %	Location	Codon	Amino Change	Haplogroup Assignment
A73G	84.23	D-loop			
T195C	12.58	D-loop			K1b2 (Beh_06); K1a9 (Beh_06)
A263G	99.00	D-loop			
315insC	85.00	D-loop			
C497T	1.31	D-loop			K1a (Pal_04, Ach_05, Beh_06)
524 insAC	4.00	D-loop			
A750G	99.11	12S rRNA			
T1189C	3.16	12S rRNA			K1 (Pal_04, Ach_05, Beh_06, Kiv_06)
A1438G	96.52	12S rRNA			
A1811G	7.05	16S rRNA			
A2706G	80.72	16S rRNA			
A3480G	4.09	ND1	58	Lys -> Lys	K (Ach_05, Kiv_06, Beh_06, Pal_04)
A4769G	98.91	ND2	100	Met -> Met	
G5460A	6.72	ND2	331	Ala -> Thr	
G6182A	0.53	Cox1	93	Ala -> Ala	
C7028T	81.49	Cox1	375	Ala -> Ala	
A7245G	0.12	Cox1	448	Thr -> Ala	
A8860G	99.76	ATPase6	112	Thr -> Ala	
G9055A	4.21	ATPase6	177	Ala -> Thr	K (Kiv_06, Pal_04); K/U8b (Ach_05)
T9148C	0.12	ATPase6	208	Leu -> Leu	
T9698C	4.37	Cox3	164	Leu -> Leu	K (Beh_06)
A10398G	46.21	ND3	114	Thr -> Ala	K1 (Pal_04, Ach_05, Beh_06)
A10550G	3.65	ND4L	27	Met -> Met	K (Ach_05, Kiv_06; Beh_06)
T11299C	4.33	ND4	180	Thr -> Thr	K (Pal_04, Ach_05, Beh_06, Kiv_06)

Polymorphism	Prevalence %	Location	Codon	Amino Change	Haplogroup Assignment
A11467G	11.02	ND4	236	Leu -> Leu	
G11719A	77.76	ND4	320	Gly -> Gly	
T12235C	0.12	tRNA Ser(2)			
A12308G	11.02	tRNA Leu			
G12372A	12.76	ND5	12	Leu -> Leu	
C14167T	4.17	ND6	169	Glu -> Glu	K (Kiv_06, Beh_06, Pal_04)
T14766C	22.52	Cytb	7	Ile -> Thr	
T14798C	7.49	Cytb	18	Phe -> Leu	K (Kiv_06); K (Ach_05, Pal_04, Beh_06)
A15326G	99.31	Cytb	194	Thr -> Ala	
G15930A	1.50	tRNA Thr			
T16093C	4.54	D-loop			K1a1a (Beh_06); K1a5 (Beh_06); K1a3a (Beh_06)
C16179T	0.31	D-loop			
T16224C	4.11	D-loop			K (Ach_05, Ric_98, Beh_06, Yao_04, Pal_04)
T16311C	14.71	D-loop			K (Ric_98, Yao_04, Beh_06, Pal_04)
T16519C	57.41	D-loop			

Table 1. Your Polymorphisms vs rCRS

Haplogroup Assignment and Your Polymorphisms

Your mtDNA has both common and rare polymorphisms. Based on these polymorphisms, you are a member of mitochondrial haplogroup K1a.

- The polymorphism A73G is very common, found in over 80% of public mtDNAs. It is used for classification of several sub-haplogroups within HV, such as H1a (Loo_04).

- The polymorphism T195C occurs in roughly 1 in 8 mtDNAs. It is not strongly associated with a single major haplogroup, but is used occasionally to distinguish minor haplogroups (K1b2 vs K1b1, for example).

- A263G is a substitution mutation: the "A" at position 263 in the reference sequence has been substituted with "G" in your mtDNA. This polymorphism occurs in the vast majority (>99%) of mtDNAs. The rCRS sequence has a rare mutation (A) at this spot.

- 315insC indicates an insertion of a C base at position 315. This occurs in over 80% of public mtDNAs. The polyC tracts at positions 309 and 315 are hypermutable, with length changes caused by extra Cs being quite common.

Because of the rapid rate of mutation this region is not usually helpful for determining haplogroup.

- The polymorphism C497T occurs in 1 in 50 mtDNA samples. It is diagnostic for sub-haplogroup K1a.

- 524insA, 524insC indicates insertion of the bases A and C at position 524 in the third hypervariable region. This insertion, found in roughly 1 in 25 public sequences, occurs at the end of a stretch of five CA repeats. The alignment of your DNA with the reference sequence is shown below. The hyphens in the rCRS sequence indicate the site of the insertions. Alternative notations for this insertion are: 524.1C 524.2A; 524insAC.

- The polymorphism A750G is very common, occurring in 99% of public mtDNA molecules. It is a rare polymorphism that happens to occur in the reference DNA, so it is reported in most mtDNA analyses. A750G is used to differentiate L3e3 and L3e4 (Kiv_06).

- T1189C is found in about 3% of public mtDNAs. It is diagnostic for haplogroup K1.

- The polymorphism A1438G is common, occurring in over 95% of public mtDNAs.

- A1811G is a 16S rRNA polymorphism that occurs in about 1 in 14 mtDNA genomes. It is diagnostic for haplogroup U.

- A2706G is a common polymorphism, occurring in 80% or so of public mtDNAs. In conjunction with other, more restricted, polymorphisms, it is use to assign genomes to U2b, H and L0d1 haplogroups (Kiv_06).

- A3480G, found in 1 in 25 mtDNAs, is diagnostic for haplogroup K.

- The polymorphism A4769G occurs in the ND2 protein, but does not alter the amino acid sequence. It is very common, being found in about 99% of human mtDNA molecules.

- This is a recurrent polymorphism with a frequency of about 7%. G5460A changes the ND2 protein at amino acid 331 from alanine to threonine.

- G5460A occurs with a population frequency of about 1 in 200. It is a silent polymorphism in the Cox1 gene.

- C7028T is a very common polymorphism found in roughly 4 of 5 public mtDNA genomes. It occurs in the 375th codon of subunit I of cytochrome c oxidase, but does not change the amino acid sequence.

- A7245G is a non-silent polymorphism that changes the threonine to alanine at position 448 of Cox1 gene.

- The polymorphism A8860G is found in nearly all (>99%) mtDNAs. It occurs in the ATPase6 protein, changing the amino acid sequence at position 112 from threonine, in the rCRS genome, to alanine.

- G9055A has a frequency of roughly 4% and is associated with haplogroup K. It is a non-silent polymorphism I the ATPase6 gene.

- T9148C is a silent polymorphism in the ATPase6 gene.

- T9698C is a silent polymorphism in the Cox3 gene with a frequency of about 4%.

- A10398G is a common polymorphism that alters codon 114 of the ND3 gene. Article: Mitochondrial DNA G10398A Polymorphism and Invasive Breast Cancer in African-American Women

- A10550G is diagnostic for haplogroup K. It is found in about 1 in 25 mtDNAs.

- T11299C is a silent polymorphism in the ND4 gene associated with haplogroup K.

- The haplogroup U-specific polymorphism A11467G occurs in roughly 1 in 10 mtDNAs in the public databases. It is located on the mitochondrial genome within the ND4 gene. It is a silent polymorphism.

- G11719A is found in 3 of 4 mtDNAs. It is located within the ND4 gene and is silent - that is., it does not affect protein sequence.

- T12235C is a rare polymorphism that maps to the tRNA Ser gene.

- The polymorphism A12308G is diagnostic for haplogroup U. Within the mitochondrial genome, it is located in the tRNA that delivers the amino acid leucine to mitochondrial proteins. This polymorphism is present in about 1 in 9 mtDNAs.

- G12372A is associated with haplogroup U. Roughly speaking, it is found in 1 in 8 mtDNA genomes. The nucleotide change, from a G to an A, is located on the mitochondrial genome in the 12th codon of the ND5 gene; the amino acid sequence is not altered by this change in the DNA sequence.

- C14167T is diagnostic for haplogroup K. It is a silent polymorphism in the ND6 gene.

- T14766C is found in 1 in 5 mitochondrial DNA genomes. It changes the "code" of codon 7, in the cytochrome b gene, from isoleucine to threonine.

- T14798C is a non-silent (Phe -> Leu) polymorphism in the Cytb gene. It is diagnostic for haplogroup K.

- The polymorphism A15326G is found in the large majority (>99%) of mtDNAs. It occurs in the Cytb protein, changing the amino acid sequence at position 194 from threonine, in the rCRS genome, to alanine.

- G15930A is a tRNA threonine polymorphism that occurs in about 1.5% of mtDNAs.

- T16093C is associated with several sub-branches within the K1 family. More DNA sequence information is needed to resolve which of these sub-groups the sample belongs in.

- C16179T is a relatively rare polymorphism (0.3% of public mtDNAs) in the D-Loop region.

- T16224C is diagnostic for haplogroup K. It found in about 4% of public mtDNAs.

- T16311C occurs in roughly 1 in 7 mtDNAs. It is associated with several haplogroups and sub-haplogroups.

- The polymorphism T16519C arises from a substitution mutation: the "T" at position 16,519 of the reference sequence is replaced by a "C" in the test sequence. This site is polymorphic in many (57%) mtDNAs, of various haplogroups. It is occasionally helpful in dissecting sub-haplogroups, when used in conjunction with other more reliable polymorphisms.

The haplogroup K1a is divided into several sub-haplogroups: K1a1 to K1a9. Your mtDNA does not have the diagnostic markers for any of these sub-groups and so is classed simply as K1a. The table below shows the diagnostic markers that are missing in your mtDNA and that are diagnostic for K1a sub-haplogroups.

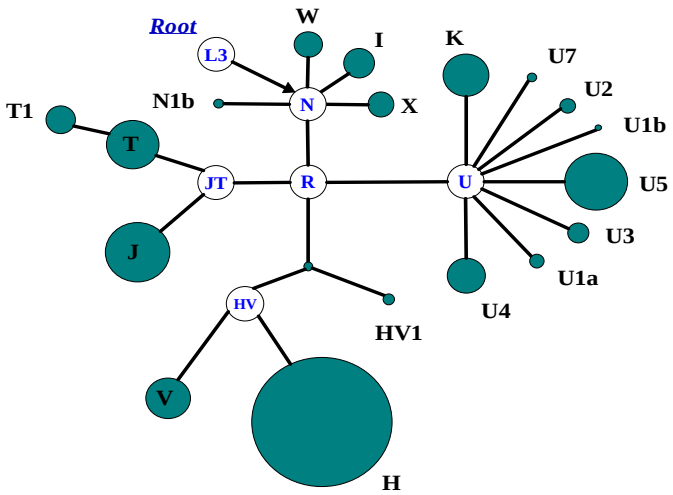
Polymorphism	Haplogroup Assignment
G11914A	K1a1 (Beh_06, Pal_04)
T11025C	K1a2 (Beh_06, Pal_04)
A13117G	K1a3 (Beh_06)
T11485C	K1a4 (Beh_06)
C4640A	K1a5 (Beh_06)
T9647C	K1a5 (Beh_06)
C8703	K1a6 (Beh_06)
C16527T	K1a6 (Beh_06)
C431T	K1a7 (Beh_06)
A4310G	K1a7 (Beh_06)
C295T	K1a8 (Beh_06)
C7927G	K1a8 (Beh_06)
A16524G	K1a9 (Beh_06)

Table 2. Absent diagnostic polymorphisms are useful for defining your mtDNA group.

There is a wealth of information available on the web and in academic publications dealing with mitochondrial DNA in general and your haplogroup in particular. This link will search the web for sites dealing with. Haplogroup K1.

Genogram

This figure shows the relative size of mitochondrial haplogroups found in modern Europe. The size of the circles reflects the prevalence of the haplogroup. Haplogroup H is the major type of mitochondrial DNA in Europe, followed by J, T, U5, U4, K, V, W, I and X in rough order of frequency. Note that the haplogroup H can be divided into multiple sub-groups (not shown). Also, note that haplogroup K is part of the U super-family.



External links

Your Haplogroup

- Spread of Your Haplogroup, from *National Geographic* (Click on "Genetic Markers" and select your haplogroup).

- List of Haplogroup members and countries of origin, from Mitosearch.

- Frequency of various haplogroups in Europe, Helgason, et al. (see Table 2).

General Interest

- mtDNA, Argus Biosciences

- Mitochondrial DNA and Human History, Wellcome Trust, UK

- Mitosearch, Search for people who share your polymorphisms

- The International Society of Genetic Genealogy, haplogroups of famous people and other features of interest.

Selected Articles:

1. Saami and Berbers—An Unexpected Mitochondrial DNA Link, Achilli et al, 2005

2. Phylogeography of Mitochondrial DNA in Western Europe, Richards, et al, 1998

3. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Pereira, 2005

4. mtDNA and the Islands of the North Atlantic: Estimating the Proportions of Norse and Gaelic Ancestry , Helgason et al., 2001

5. The molecular dissection of mtDNA Haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Achilli, et al., 2004.

6. Phylogeny of Mitochondrial DNA Macrohaplogroup N in India, Based on Complete Sequencing: Implications for the Peopling of South Asia, Palanichamy, et al., 2004
7. The Making of the African mtDNA Landscape, Salas, 2002

8. Whole-mtDNA Genome Sequence Analysis of Ancient African Lineages, Gonder, et al., 2007

9. The role of selection in the evolution of human mitochondrial genomes, Kivisild, et al. 2006

10. Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia, Loogvali, 2004

General Background on Mitochondrial DNA (mtDNA)

What is DNA?

DNA is like a recipe - it is a set of instructions to make something. The DNA you were born with is a recipe to make You. Your eye color, your height, your intelligence, everything about who you are as a human being is coded for in the genes that you were born with.

What are mitochondria?

Mitochondria are the "powerhouses of the cell". Their function is to break down sugars and release energy for use by the cell. Cells that are energy-intensive, such as muscle cells, have more mitochondria than cells with low energy needs. In the diagram below, the mitochondria are the purple compartments with the thread-like membranes inside.

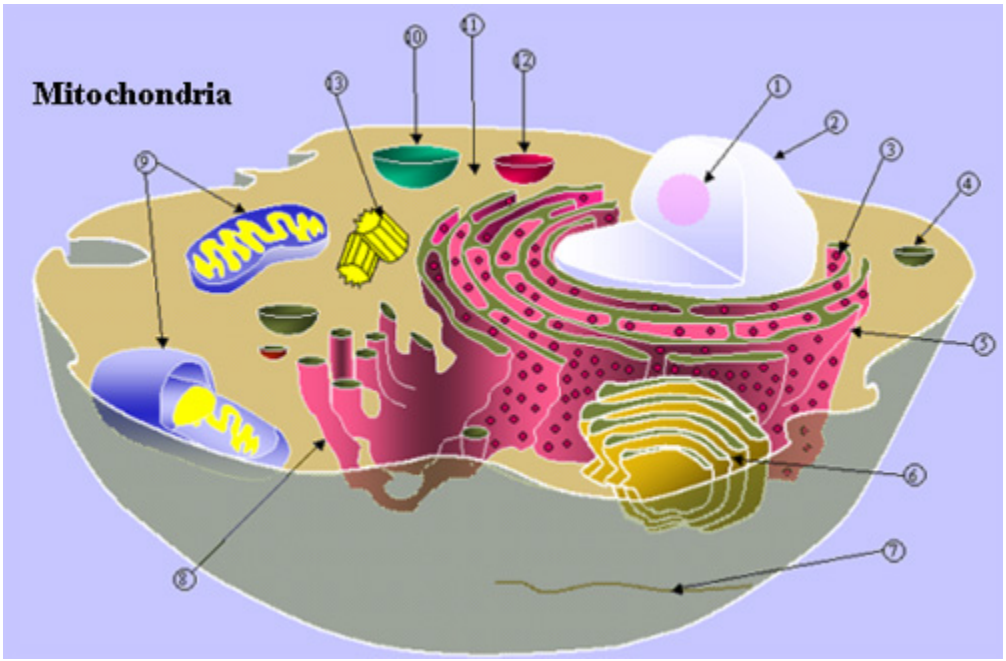
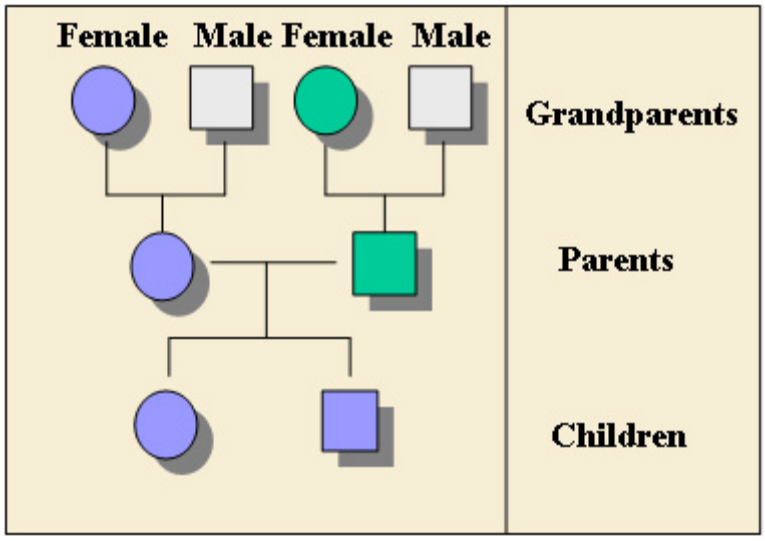


Diagram of a typical animal cell. Organelles are labeled as follows: 1) Nucleolus, 2) Nucleus 3) Ribosome 4) Vesicle 5) Rough endoplasmic reticulum 6)Golgi apparatus 7) Cytoskeleton 8) Smooth endoplasmic reticulum 9) Mitochondrion 10) Vacuole 11)

Cytoplasm 12) Lysosome 13) Centriole. Image by Magnus Manske (from Nupedia, reproduced with permission).

Mitochondrial DNA

Each mitochondrion has its own DNA, or genome, separate from the DNA in the nucleus. The mitochondrial genome is a circular molecule of double-stranded DNA, 16,569 base pairs long. A base is a specific component of the DNA and is made of adenine, thymine, guanine or cytosine (A, T, G, C). Within the genome, there is an approximately 1100 base long regulatory region, called the D-loop. Because this region accumulates genetic changes faster than the rest of the genome, it is also referred to as the hypervariable region. The remainder of the mitochondrial genome is coding DNA - it is copied into RNA molecules that perform downstream functions within the cell. The mitochondrial genome codes for 13 proteins (used in energy production by the mitochondria) two ribosomal RNAs (used for protein synthesis) and 22 transfer RNAs (also used for protein synthesis).



How mtDNA is inherited. Mitochondrial DNA is inherited only from the mother: the fertilized egg destroys the mitochondria of the sperm. Because of this selective matrilineal transmission, mitochondrial DNA sequences can be used to by population geneticists and evolutionary biologists to shed light on the unbroken genetic line connecting us to our maternal ancestors.

Note that the children inherit their maternal grandmother's mitochondrial DNA (in purple, left side of diagram) without contribution from either grandfather, or the father, or the paternal grandmother.

Polymorphisms

Mutations, when they occur, can be passed down to the children. These mutations, or polymorphisms, tell a story about your past. Part of that story is told simply by the number of polymorphisms identified in your mtDNA. Because the genome accumulates mutations at a linear rate over time, the polymorphisms represent a sort of molecular clock: the more polymorphisms that differ between two people's mtDNA, the longer ago in the past they shared a common ancestor. For example, while an African-American and

a European might have 75 variable polymorphisms, two people of European descent might have only 25 variable polymorphisms, reflecting a more recent common ancestor.

Your polymorphisms also tell a story about place, about where your ancestors came from. Imagine a small group of people, migrating out of the Middle East and into a locale somewhere in Western Europe. If they succeed in colonizing the region, they will pass on their particular mtDNA onto their descendants. Skipping forward to today, these polymorphisms can now be associated with particular geographic areas and populations. In conjunction with linguistic and anthropological studies, researchers have constructed ancient migration patterns based on the presence of these polymorphisms in human populations.

Some polymorphisms are quite common, represented in over 50% of a given population. This may be due to a founder effect, as mentioned above. There is also some evidence that certain polymorphisms may have rendered their carriers resistant to certain diseases, giving them a selective advantage over non-carriers. A recent article published in Lancet, for example, claims that mtDNA haplogroup H is a strong independent predictor of increased chance of survival after sepsis, and goes on to suggest that this resistance may have contributed to making this haplogroup the most common on Europe. There are also studies of which polymorphisms are more likely to be found in centenarians.

Other polymorphisms are quite rare, occurring only once in a thousand or more mitochondrial genomes. There are still relatively few full length genomes available for comparison, however; the actual frequency of a given polymorphism will be better known as more and more genomic sequences become available.

The revised Cambridge Reference Sequence

After completion of your sequencing project, the sequence of your mtDNA is compared to a standard sequence, called the revised Cambridge Reference Sequence, or rCRS. (The original had several mistakes corrected in the revised version). The rCRS sets the numbering for each base, so that any two mtDNAs can be compared. This is important because, due to small deletions and insertions of DNA in many genomes, any two genomes would quickly become out of register. To get the numbering right, each genome is first aligned to the standard rCRS, introducing gaps or insertions as needed, and then each of the 16,569 or so paired bases is numbered relative to the standard genome. Polymorphisms are generally written like this: "A750G", which means that the A at position 750 in the rCRS is changed to a G in the equivalent spot on the sample genome.

Mutation vs Polymorphism

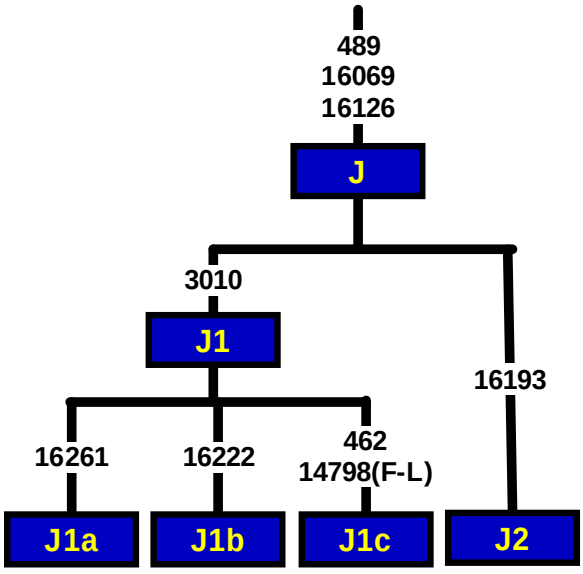
The words polymorphism and mutation are often used interchangeably in talking about mt DNA. One way to distinguish them is that a mutation becomes a polymorphism as it gains a foothold within a population. If you inherit a change in your mitochondrial DNA that originated in your mother's egg, that is a mutation, defined as a genetic alteration that

occurs during transmission of a gene from one generation to the next. If after some number of generations your descendants still carry this alteration and they come to represent a significant proportion of the population, say over 1%, the alteration can be called a polymorphism, in the sense that it is one of the variant types found in the pool of mitochondrial genomes. Mitochondria have been around for over a billion years. They are clearly very well-adapted and most mutations will be lost after a few generations. If, on the other hand, the mutation confers some sort of survival benefit in a given environment, as in the sepsis story referred to above, they are more likely to make the leap from a mutation in one person to a polymorphism within the wider population.

Haplogroups

Your haplogroup identifies your ethnic and geographic origins on your maternal line. Members of a haplogroup are related to each other by common descent. Mitochondrial haplogroups are sometimes referred to as maternal clans, since members share a common maternal ancestor. There are nine main haplogroups in Europe, and about 30 worldwide

Haplogroups are defined by polymorphisms. For example, if upon comparing your mitochondrial to the standard sequence, polymorphisms are identified at positions 489, 16,069 and 16,126, you will be classed in haplogroup J. If you also have polymorphisms at positions 3010 and 16,261, you can be more precisely placed within the sub-haplogroup J1a.



A portion of the phylogenetic tree for haplogroup J. Haplogroups are defined by polymorphisms

Phylogenetic Trees

A phylogenetic tree is a graphical representation of the evolutionary relationship between groups. Like a family tree, it traces diverging lines of descent from a common ancestor. The phylogenetic tree of human mitochondrial DNA depicts the relationship of global mitochondrial haplogroups. The first mitochondrial tree was based on the pioneering research of Allan Wilson in the mid-1980s (Cann, et al).

Haplotypes and Haplogroups

In the roughly 8,000 generations that separate us from our common African ancestors, our mtDNA has diverged. Though the difference between any two people is less than 1%., there are enough differences to see patterns in the DNA sequences. The set of polymorphisms for an individual is called a haplotype. Similar haplotypes can be grouped together into haplogroups.

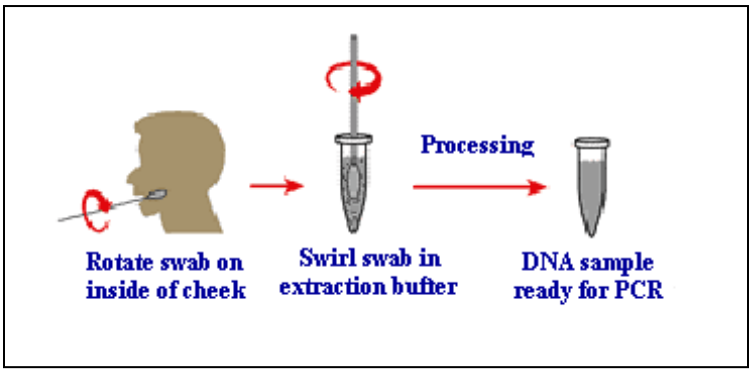
Mitochondrial Eve

The phylogenetic tree of mtDNA has a single source, a single mitochondrial genome at the root of the tree. Humans did not arise separately in China and Australia and Europe - those populations are derived from a common ancestral population. The root of the tree is in Africa. mtDNA from Africa has greater genetic diversity than mtDNA from other regions, the result a more ancient lineage. There are a couple of things we can say about the woman who has the distinction of having copies of her mitochondrial genome present is every person living today. She lived in Africa - so on some level we are all Africans. She had at least two daughters: if she just had one, then that daughter would be our most recent common mitochondrial ancestor, not her mother. And in all probability there was nothing special about her - she was a member of a clan that included women much like her. Her founder status is a matter of chance, it could just as easily been another woman.

Scientific Procedures

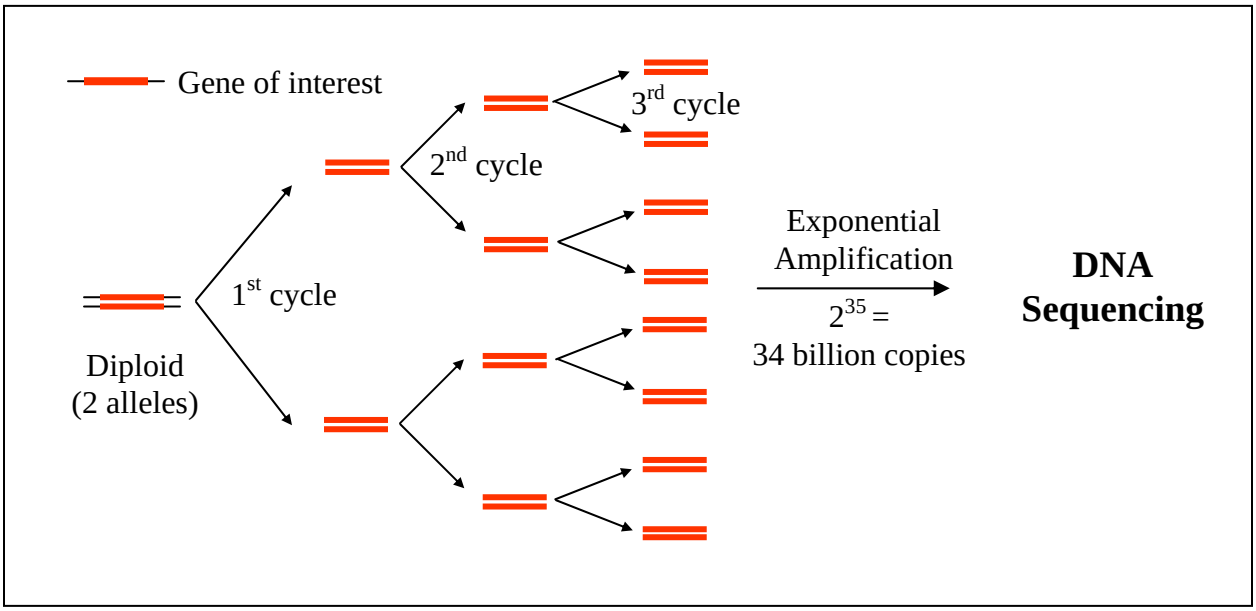
DNA Preparation

The process starts with collecting buccal (cheek) cells from inside your mouth. These cells are broken open in a special extraction solution and the DNA is prepared for amplification by polymerase chain reaction (PCR).



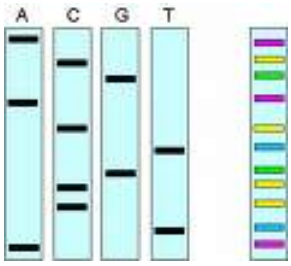
PCR

The DNA has to be amplified to get enough of it for sequencing. This is done using the polymerase chain reaction (PCR), which makes many copies of the regions we want to sequence. In the example, two copies of a gene to be sequenced are amplified to 34 billion copies in 35 cycles of gene doubling. The PCR products are then used for DNA sequencing.



Gene amplification by polymerase chain reaction. The number of genes is doubled with each cycle. The entire PCR amplification is complete in about two hours.

DNA Sequencing



The PCR products are used as templates in a biochemical reaction that generates single-stranded pieces of DNA, one type each for each base (A, T, G or C). The mixture of single-stranded DNA molecules is run through a matrix that separates the strands based on their size.

Finding Polymorphisms

Once we have the DNA for the "test" sequence, we compare it base-for-base with a reference DNA to identify polymorphisms. The two DNAs are aligned and differences are annotated using the program GEN-SNiP, developed at Argus Biosciences in collaboration with SooryaKiran Bioinformatics of Kerala, India. In the example box, the C at position 2 in the reference sequence is changed to a T in the test sequence.

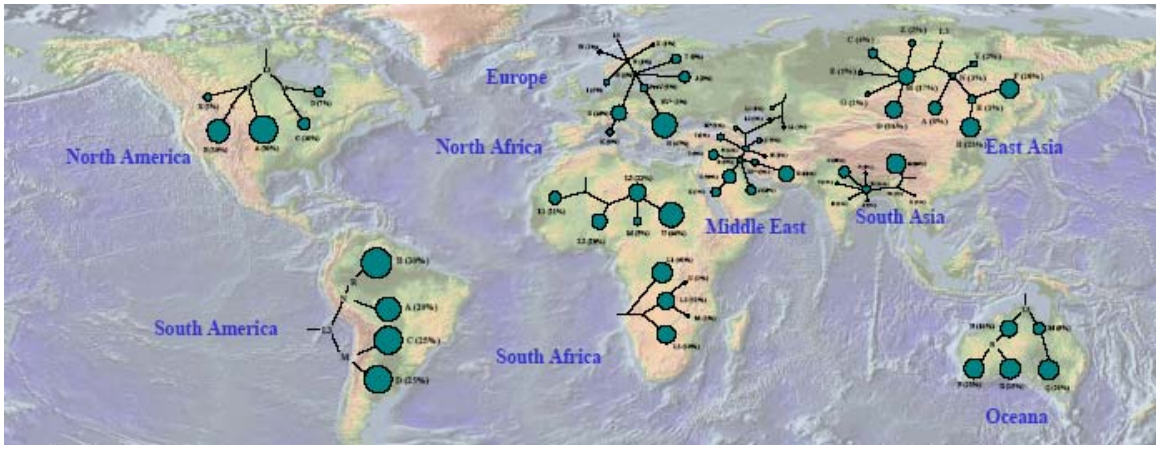
```
Test sequence: ATCGTGCTA
    Reference: ACCGTGCTA
     Position: 123456789
 Polymorphism: C2T
```

Assigning your Haplogroup

Your haplogroup is determined by the presence of certain diagnostic polymorphisms. Once the DNA sequence is obtained, it is compared to a reference sequence to determine the position and nature of each polymorphism. The list of polymorphisms contained in your mtDNA is then checked against a comprehensive spreadsheet of polymorphisms - and associated haplogroup assignments - extracted from the scientific literature. The results of this analysis are tabulated in the table of polymorphisms. Articles we use to associate specific polymorphisms with mitochondrial haplogroups are available on our website at http://argusbio.com/papers.html.

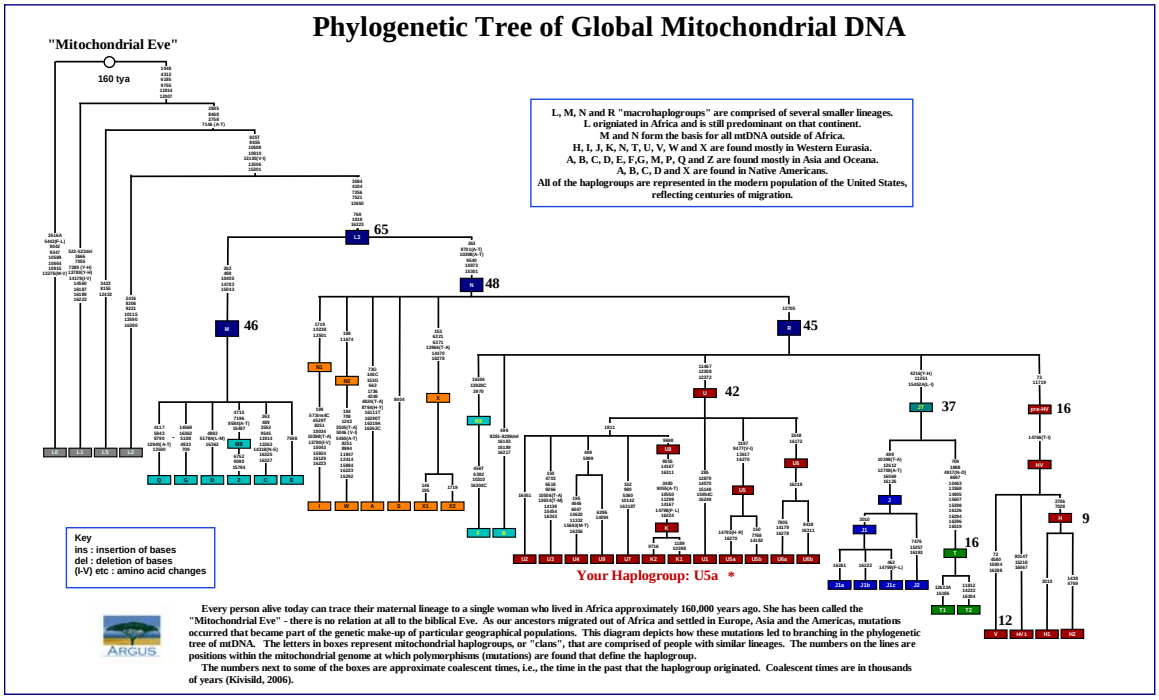
Genograms

Genograms depict the relative size of mitochondrial haplogroup populations, as well as how they are related to each other by descent. Genogram may reflect mtDNA populations in Europe, Asia, N. and S. America, Oceana or Africa. Your package of results also includes a map of the world with mtDNA genograms overlaid on the continents (thumbnail below).



Phylogenetic Trees

A phylogenetic tree is a “family tree” showing how all of the global haplogroups are related to each other. A personalized phylogenetic tree showing how your haplogroup is related to other haplogroups is included in your package. We have a number of tree types, typically sending the one that best illustrates the relationship of your haplogroup to others.



Example of the personalized phylogenetic tree. The tree in the final report is of higher resolution.

glossary

A	
Allele	One of several alternative forms of a gene or DNA sequence at a specific chromosomal location (locus). At each autosomal locus an individual possesses two alleles, one from each parent.
Amplification	The process of making identical genetic copies of a specific region of DNA. PCR is a powerful technique for amplifying specific regions of DNA.
Ancestral Clan Mother	A woman who is considered to be the ancient maternal ancestor from whom all people in a particular haplogroup (clan) are descended.
Ancestry	A person's line of descent.
Anthropology	The study of humankind, including the comparative study of societies and cultures, and the science of human zoology and evolution.
Atlantic Modal Haplotype or AMH	Most common haplotype found in Europe.
Autosomes	The non-sex chromosomes. Humans have 23 pairs of chromosomes within the nucleus. Chromosomes 1 through 22 are autosomal. The other pair has the special property of determining one's gender: XX in females and XY in males.
B	
Base	Adenine (A), cytosine (C), guanine, (G) or thymine (T) are the four bases in the DNA. The chemical building blocks of DNA. These bases pair up to form the "rungs" of the DNA double helix.
Base Pair	The DNA bases are always held together in pairs by weak hydrogen bonds attaching to one of the strands in the DNA double helix. Adenine always pairs with thymine, and guanine always pairs with cytosine.
BLAST	A family of programs that search sequence databases for matches to a query sequence.
C	

Cambridge Reference Sequence (CRS)	The first complete sequence of human mtDNA, published in 1981. Recently revised to correct minor errors in the initial sequence (revised CRS or rCRS). Each mtDNA haplotype is described by the differences it shows with the rCRS. The nucleotides of this standard molecule are numbered from 1 to 16569.The history of the sequence changes is described in the Mitomap website.
Chromosome	Long strands of DNA on which genes are found. Each human cell has 46 chromosomes in 23 pairs. One member of each pair is inherited from the mother, the other from the father. Mitochondrial DNA is inherited only from the mother.
Clade	A group comprising all the evolutionary descendants of a common ancestor. Also called "clan" or haplogroup.
Coalescence time	Time in the past at which two or more lines of descent split from a common ancestor.
Coding DNA	DNA that encodes the amino acid sequence of a protein, or for a functional mature RNA (transfer RNA or ribosomal RNA).
Codon	A nucleotide triplet that specifies an amino acid or a translation stop signal.
Complementary strands	Two nucleic acid strands are complementary in sequence if they can form a stable double-stranded structure. Base pairs are formed between adenine and thymine (AT) and guanine and cytosine (GC).The GC pairing has three hydrogen bonds and is thus stronger (i.e., requires more energy to melt) than the AT pairing, which has two hydrogen bonds.
CRS	See Cambridge Reference Sequence.
Cytosine	The "C" in ATGC, the four bases found in DNA."C" is short for cytosine, a base that bonds with Guanine (G) in double stranded DNA.
D	
Displacement (D) loop region	In mitochondrial DNA, the D-loop is a short triple-stranded region that contains regulatory sequences. The D-loop contains several hypervariable regions that have relatively higher rate of mutation compared to the coding region of mtDNA. Transcription (DNA to RNA) originates from two closely spaced promoters located in the D loop region. The replication (i.e., duplication of DNA) of both strands is unidirectional and starts at specific "origins of replication" in the D loop.
DNA (deoxyribonucleic acid)	The double helix-shaped molecule that holds an organism's genetic information. The DNA in each cell contains over 3 billion base pairs coding the approximately 25,000 genes that make up the human genome. DNA is composed of sugars, phosphates, and four nucleotide bases: adenine, guanine, cytosine, and thymine (A, G, C, T). The bases bind together in specific pairs - A:T and G:C.
DNA Letters	The DNA molecule is composed of a string of four chemicals called adenine, cytosine, guanine and thymine, normally abbreviated to A, C, G and T, respectively.
DNA Sequence	The order or arrangement of the DNA letters (A, T, C & G) making up the DNA molecule.
DNA sequencing	Laboratory procedure for determining the exact order of bases (ATCG) in a strand of DNA. In "bi-directional sequencing" the sequence of both complementary strands is obtained.
DNA	Short for DeoxyriboNucleic Acid. The genetic material carried by all animals and plants that allows transmission of characteristics from one generation to the next.
E	
Electropherogram	The display of DNA sequence information from a capillary-based genetic analyzer.
Enzyme	A protein that catalyzes a specific chemical reaction.
G	

Gene	The basic unit of inheritance. A gene is a sub-unit of DNA in a particular position on a particular chromosome that contains the genetic code to make a particular protein.
Genealogy	The study of lines of descent. Alternatively, a line of descent traced continuously from an ancestor.
Generation time	The number of years between the birth of parents and the birth of their children.
Genetic Ancestry	Line of descent supported by genetic evidence.
Genetic Marker	Any part of the DNA molecule that expresses variability within a population and that can be used for analysis of that population.
Genogram	Diagram that depicts the phylogenetic relationship of mitochondrial haplogroups (the line connecting the circles) and the relative size of the populations (the area of the labeled circles).
Genome	The entire complement of genetic material in a nucleus (23 pairs of chromosomes) or an organelle (mitochondrial DNA). The mitochondrial genome is 16,569 bases long, circular and resides within the mitochondrion.
Genotype	The set of genes of an individual.
Guanine	The "G" in A, T, G & C, the four bases found in DNA."G" is short for guanine, a base that bonds with Cytosine (C) in double stranded DNA.
H	
Haplogroup	A group of haplotypes that share common ancestry defined by shared sequence. Haplogroups are the main branches of the human genealogical tree and reflect early human migrations. A haplogroup contains all the direct descendants of a single person (man or woman) who passed on a specific genetic marker or mutation.
Haplotype	Different combinations of polymorphisms are known as haplotypes. A set of closely linked alleles (genes or DNA polymorphisms) inherited as a unit. A contraction of the phrase "haploid genotype."
Heteroplasmy	The presence of two or more mtDNA genotypes in a single DNA sample.
Homoplasmy	Having all copies of mtDNA the same within a cell or organism (compare heteroplasmy).Not to be confused with homoplasy, which means something entirely different!
Homoplasy	In mitochondrial or chromosomal DNA, a situation in which the same polymorphism arises independently in two or more haplogroups. Homoplastic polymorphisms are more likely in regions of high mutation rate and limit the "informativeness" of the polymorphism. In a broader biological context, homoplasy refers to similar features that are NOT derived form a common ancestral feature, and have thus arisen from parallelism, convergence, or chance.
Human Migration	The movement of historic populations of Man out of Africa and across the continents of the world.
Hypervariable region	See D-loop, above.
I	
Identity by descent (IBD)	Alleles that are identical because they have both been inherited from a common ancestor, as opposed to identity by state (IBS).
Identity by state (IBS)	Coincidental possession of alleles that appear to be identical but have not been proved to be of common descent.
L	
Locus	The position of a particular gene on a chromosome.
M	

Marker	A physical location (locus) on the chromosome.
Matrilineal	Passed down exclusively from the mother.
Meiosis	A special type of cell division that occurs when gametes (spermatozoa and eggs) are formed. Each gamete contains just one version of each gene.
Metabolism	The chemical processes occurring within an organism that are necessary for the maintenance of life.
Migration Routes	The lines of travel taken by our ancestors as they migrated out of Africa and colonized the other continents.
Mitochondria	Plural of mitochondrion. Mitochondria are the "powerhouses of the cell". Their function is to break down sugars and release energy for use by the cell. They have their own circular genome.
Mitochondrial DNA or mtDNA	The circular, 16,569 base pair long genome of the mitochondrion. It encodes 13 protein coding genes, two ribosomal RNAs and 22 transfer RNAs. Because the sperm cell does not contribute any mitochondria to the fertilized egg, all of the mitochondria in both male and female offspring are derived from the mother. It is ideal for tracing recent human history, such as the emergence of ethnically distinct lineages or haplogroups, because of its relatively high mutation rate and lack of recombination.
Mitochondrial Eve	Also known as African Eve. The human female who lived approximately 150,000 to 200,000 years ago and from whom everyone on this planet is descended through the maternal line.
MitoMap	An online human mitochondrial genome database (www.mitomap.org)
Molecular clock	The clock-like regularity of the change of a gene over geological time. Different genetic regions may have quite different rates of change.
MRCA	Most recent common ancestor.
mtDNA	See Mitochondrial DNA.
Mutation	An inheritable alteration in the genetic material. At the level of the whole organism, mutations can be divided into germ line and somatic types. Those that occur in the cells that make spermatozoa or eggs (germ cells) can be passed on to the next generation. Mutations that occur in somatic (non-germ) cells and are not transmitted to progeny. Somatic mutations are common in cancer tissue. A non-synonymous or missense mutation results in an amino acid change in a protein. A synonymous or silent mutation replace one codon with another that encodes the same amino acid.
Mutation rate	The rate at which changes occur in DNA sequence. Regions coding for proteins have a lower mutation rate than non-coding DNA since they may alter protein structure and are thus selected against. As an example, say the mutation rate of the coding region of mtDNA is approximately 1.7% per million years. Thus if a population of mtDNAs has an average difference of 0.3% within the coding region, one can estimate that they began to diverge from a common ancestor around 175,000 years ago (i.e., 0.3/1.7 million years ago).Because the mutation rate of the D-loop region in mtDNA is faster than the rate for the coding region, it will take less time to reach the same level of population diversity for this region. Because of this high mutation rate, reliance on the hypervariable D-loop region to establish phylogenetic networks is limited by the effects of saturation and homoplasy.
N	
Non-synonymous mutation	A mutation in DNA that alters the amino acid composition of a protein.
Nucleotide	A nucleotide is formed by adding a sugar unit and a phosphate to a base. Nucleotide trisphosphate are used by the cell to form the nucleic acid polymers RNA and DNA. When referring to A, G, C or T, “nucleotides” is commonly used interchangeably with “bases”.

Nucleus	The membrane bound organelle containing the chromosomes.
O	
OMIM - Online Mendelian Inheritance in Man	A central database of human genes involved in disease.
Organelle	A structure within a cell, such as a mitochondrion, that performs a specific function. Organelle = "little organ".
Organism	An individual animal, plant or single-celled life form.
P	
Parallelism	An evolutionary event where two identical changes occur independently.
PCR - Polymerase Chain Reaction	A biochemical technique that allows the specific amplification (production of multiple copies) of extremely small amounts of particular DNA fragments using DNA polymerase and specific primers.
Phylogeny	The inferred lines of descent from a common ancestor. The etymology of the word is interesting: phylo- means "tribe" or "clan"; -geny means "origins".
Polymorphism	The simultaneous occurrence of two or more versions of a gene in a population. In mitochondrial DNA, it usually refers to different bases at a particular position, such as A750G.The frequency of the rarest form of the polymorphism is higher than can be maintained by recurrent mutation.
Primer	A short DNA molecule used for PCR and for sequencing.
Protein	Proteins are made up of amino acids - they are the main building blocks of our cells.
Purine	Adenine and Guanine are purine bases; thymine and cytosine are pyrimidine bases. urines consist of a six-membered and a five-membered nitrogen-containing ring, fused together. Pyridmidines have only a six-membered nitrogen-containing ring.
Pyrimidine	Thymine and cytosine are pyrimidine bases.
R	
Recombination	The reshuffling of the genes in a fertilized egg as a result of crossing-over and re-assortment of the chromosomes during meiosis (i.e., during the formation of the sperm and egg).Mitochondrial DNA does not recombine.
Replication	The process by which the DNA double helix makes an exact copy of itself or of a fragment. It uses the DNA as a template for the synthesis of new DNA strands.
Root Node	The original sequence of any specific clan mother or father.
S	
Sequence	See DNA Sequence.
Sequencing	The determination of the order of the four DNA letters within the DNA molecule. Once the DNA is extracted and purified from a cell sample, it is amplified and the sequence determined using a DNA sequencer.
Sex chromosomes	The X and Y chromosomes. Normally males have one X and one Y and females have two X's. Sex chromosomes make up the 23rd pair, different from the other 22 pairs that are the autosomal DNA.
Silent Mutation	See synonymous mutation
SNP - Single Nucleotide Polymorphism	Changes in the DNA that happen when a single nucleotide (A, T, G, or C) in the genome sequence is altered. A person has many SNP's that together create a unique DNA pattern for that individual.
Substitution mutation	Mutation in which one base is replaced by another base.

Synonymous mutation	A change in DNA sequence that does not result in a change in protein sequence. Also called silent mutations, these are usually "invisible" to selective pressure since they don't alter protein sequence.
T	
Thymine	The "T" in ATGC, the four bases found in DNA."T" is short for thymine, a base that bonds with adenine (A) in double stranded DNA.
TMRCA	Time to the Most Recent Common Ancestor.
Trace	See electropherogram.
Transition Mutation	A type of base substitution in which a pyrimidine (C, T) is replaced by another pyrimidine (T, C), or a purine (A,G) is replaced by another purine (G, A). Transitions are much more common than transversions. C to T, T to C, A to G, G to A are transitions.
Transversion Mutation	A type of base substitution in which a pyrimidine is replaced by a purine, or vice versa. C to A, C to G, T to A, T to G, A to C, A to T, G to C and G to T are transversions.
Transmission event	The passage of genes from one generation to the next.
Tribes	Traditionally used to describe a large number of people with the same culture and dialect.
X	
X chromosome	One of the two sex chromosomes, X and Y. X is the sex chromosome that is present in both sexes: singly in males and doubly in females.
Z	
Zygote	A fertilized egg.

Appendix

References used to assign haplogroups

Ref	Author	Year	Title
Ach_04	Achilli	2004	The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool.
Ach_05	Achilli	2005	Saami and Berbers--an unexpected mitochondrial DNA link. (Hg U)
Beh_06	Behar	2006	The Matrilineal Ancestry of Ashkenazi Jewry: Portrait of a Recent Founder Event
Bra_06	Bradstatter	2006	Dissection of mitochondrial superhaplogroup H using coding region SNPs"
Kiv_99	Kivisild	1999	The Place of the Indian mtDNA Variants in the Global Network of Maternal Lineages and the Peopling of the Old World
Kiv_02	Kivisild	2002	The Emerging Limbs and Twigs of the East Asian mtDNA Tree
Kiv_06	Kivisild	2006	The role of selection in the evolution of human mitochondrial genomes.
Kon_03	Kong	2003	Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences.
Loo_04	Loogvali	2004	Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia.
MM_01	Maca-Meyer	2001	Major genomic mitochondrial lineages delineate early human expansions
MM_03	Maca-Meyer	2003	Mitochondrial DNA transit between West Asia and North Africa inferred from U6 phylogeography
Pal_04	Palanichamy	2004	Phylogeny of Mitochondrial DNA Macrohaplogroup N in India, Based on Complete Sequencing: Implications for the Peopling of South Asia
Pal_06	Palanichamy	2006	Comment on ‘‘Reconstructing the Origin of Andaman Islanders’’
R03	Reidla	2003	Origin and Diffusion of Haplogroup X
Ric_98	Richards	1998	Phylogeography of mtDNA in Western Europe
Ric_00	Richards	2000	Tracing European Founder Lineages in the Near Eastern mtDNA Pool
Sal_04	Salas	2004	The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade
Sal_02	Salas	2002	The Making of the African mtDNA Landscape
Tam_04	Tambets	2004	The western and eastern roots of the Saami--the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes.
Tan_04	Tanaka	2004	Mitochondrial genome variation in eastern Asia and the peopling of Japan.
Tor_92	Torroni	1992	Native American Mitochondrial DNA Analysis Indicates That the Amerind and the Nadene Populations Were Founded by Two Independent Migrations

Ref	Author	Year	Title
Tro_03	Trovoada	2003	Pattern of mtDNA Variation in Three Populations from Sao Tome e Prıncipe
Yao_04	Yao	2004	Different Matrilineal Contributions to Genetic Structure of Ethnic Groups in the Silk Road Region of China

Your mtDNA Sequence

The DNA sequence for your complete mitochondrial genome, starting at nucleotide position 1.

```
>12482
GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTTCGTCTGGGGGGTGTGCACGC
GATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATTTA
TCGCACCTACGTTCAATATTACAGGCGAACATACCTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA
ACAATTGAATGTCTGCACAGCCGCTTTCCACACAGACATCATAACAAAAAATTTCCACCAAACCCCCCCTCCCCCCGCTT
CTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAAACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATT
TTATCTTTTGGCGGTATGCACTTTTAACAGTCACCCCCCAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAA
TCTCATCAATACAACCCTCGCCCATCCTACCCAGCACACACACACACCGCTGCTAACCCCATACCCCGAACCAACCAAAC
CCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAAGCAATACACTGAAAATGTTTAGACGGGCTCACATCAC
CCCATAAACAAATAGGTTTGGTCCTAGCCTTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCATCCCCGTTCCAGTG
AGTTCACCCTCTAAATCACCACGATCAAAAGGGACAAGCATCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGC
CACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCCAGGGTTG
GTCAATTTCGTGCCAGCCACCGCGGTCACACGATTAACCCAAGTCAATAGAAGCCGGCGTAAAGAGTGTTTTAGATCACC
CCCTCCCCAATAAAGCTAAAACTCACCTGAGTTGTAAAAAACTCCAGTTGACACAAAATAGACTACGAAAGTGGCTTTAA
CATATCTGAACACACAATAGCTAAGACCCAAACTGGGATTAGATACCCCACTATGCTTAGCCCTAAACCTCAACAGTTAA
ATCAACAAAACTGCTCGCCAGAACACTACGAGCCACAGCTTAAAACTCAAAGGACCTGGCGGTGCTTCATACCCCTCTAG
AGGAGCCTGTTCTGTAATCGATAAACCCCGATCAACCTCACCACCTCTTGCTCAGCCTATATACCGCCATCTTCAGCAAA
CCCTGATGAAGGCTACAAAGTAAGCGCAAGTACCCACGTAAAGACGTTAGGTCAAGGTGTAGCCCATGAGGTGGCAAGAA
ATGGGCTACATTTTCTACCCCAGAAAACTACGATAGCCCTTATGAAACTTAAGGGTCGAAGGTGGATTTAGCAGTAAACT
GAGAGTAGAGTGCTTAGTTGAACAGGGCCCTGAAGCGCGTACACACCGCCCGTCACCCTCCTCAAGTATACTTCAAAGGA
CATTTAACTAAAACCCCTACGCATTTATATAGAGGAGACAAGTCGTAACATGGTAAGTGTACTGGAAAGTGCACTTGGAC
GAACCAGAGTGTAGCTTAACACAAAGCACCCAACTTACACTTAGGAGATTTCAACTTAACTTGACCGCTCTGAGCTAAAC
CTAGCCCCAAACCCACTCCACCTTACTACCAGACAACCTTAGCCAAACCATTTACCCAAATAAAGTATAGGCGATAGAAA
TTGAAACCTGGCGCAATAGATATAGTACCGCAAGGGAAAGATGAAAAATTATAGCCAAGCATAATATAGCAAGGACTAAC
CCCTATACCTTCTGCATAATGAATTAACTAGAAATAACTTTGCAAGGAGAGCCAAAGCTAAGACCCCCGAAACCAGACGA
GCTACCTAAGAACAGCTAAAAGAGCACACCCGTCTATGTAGCAAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACC
TACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAATCTTAGTTCAACTTTAAATTTGCCCACAGAACCCTCTAAATCC
CCTTGTAAATTTAACTGTTAGTCCAAAGAGGAACAGCTCTTTGGACACTAGGAAAAAACCTTGTAGAGAGAGTAAAAAAT
TTAACACCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACACCCACTACCTAAAAAATCCCA
AACATATAACTGAACTCCTCACACCCAATTGGACCAATCTATCACCCTATAGAAGAACTAATGTTAGTATAAGTAACATG
AAAACATTCTCCTCCGCATAAGCCTGCGTCAGATTAAAACACTGAACTGACAATTAACAGCCCAATATCTACAATCAACC
AACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAGGAAAGGTTAAAAAAAGTAAAAGGAACTCG
GCAAATCTTACCCCGCCTGTTTACCAAAAACATCACCTCTAGCATCACCAGTATTAGAGGCACCGCCTGCCCAGTGACAC
ATGTTTAACGGCCGCGGTACCCTAACCGTGCAAAGGTAGCATAATCACTTGTTCCTTAAATAGGGACCTGTATGAATGGC
TCCACGAGGGTTCAGCTGTCTCTTACTTTTAACCAGTGAAATTGACCTGCCCGTGAAGAGGCGGGCATGACACAGCAAGA
CGAGAAGACCCTATGGAGCTTTAATTTATTAATGCAAACAGTACCTAACAAACCCACAGGTCCTAAACTACCAAACCTGC
ATTAAAAATTTCGGTTGGGGCGACCTCGGAGCAGAACCCAACCTCCGAGCAGTACATGCTAAGACTTCACCAGTCAAAGC
GAACTACTATACTCAATTGATCCAATAACTTGACCAACGGAACAAGTTACCCTAGGGATAACAGCGCAATCCTATTCTAG
AGTCCATATCAACAATAGGGTTTACGACCTCGATGTTGGATCAGGACATCCCGATGGTGCAGCCGCTATTAAAGGTTCGT
TTGTTCAACGATTAAAGTCCTACGTGATCTGAGTTCAGACCGGAGTAATCCAGGTCGGTTTCTATCTACTTCAAATTCCT
CCCTGTACGAAAGGACAAGAGAAATAAGGCCTACTTCACAAAGCGCCTTCCCCCGTAAATGATATCATCTCAACTTAGTA
TTATACCCACACCCACCCAAGAACAGGGTTTGTTAAGATGGCAGAGCCCGGTAATCGCATAAAACTTAAAACTTTACAGT
CAGAGGTTCAATTCCTCTTCTTAACAACATACCCATGGCCAACCTCCTACTCCTCATTGTACCCATTCTAATCGCAATGG
CATTCCTAATGCTTACCGAACGAAAAATTCTAGGCTATATACAACTACGCAAAGGCCCCAACGTTGTAGGCCCCTACGGG
CTACTACAACCCTTCGCTGACGCCATAAAACTCTTCACCAAGGAGCCCCTAAAACCCGCCACATCTACCATCACCCTCTA
CATCACCGCCCCGACCTTAGCTCTCACCATCGCTCTTCTACTATGAACCCCCCTCCCCATACCCAACCCCCTGGTCAACC
TCAACCTAGGCCTCCTATTTATTCTAGCCACCTCTAGCCTAGCCGTTTACTCAATCCTCTGATCAGGGTGAGCATCAAAC
TCAAACTACGCCCTGATCGGCGCACTGCGAGCAGTAGCCCAAACAATCTCATATGAAGTCACCCTAGCCATCATTCTACT
ATCAACATTACTAATAAGTGGCTCCTTTAACCTCTCCACCCTTATCACAACACAAGAACACCTCTGATTACTCCTGCCAT
CATGACCCTTGGCCATAATATGATTTATCTCCACACTAGCAGAGACCAACCGAACCCCCTTCGACCTTGCCGAAGGGGAG
TCCGAACTAGTCTCAGGCTTCAACATCGAATACGCCGCAGGCCCCTTCGCCCTATTCTTCATAGCCGAATACACAAACAT
TATTATAATAAACACCCTCACCACTACAATCTTCCTAGGAACAACATATGACGCACTCTCCCCTGAACTCTACACAACAT
ATTTTGTCACCAAGACCCTACTTCTAACCTCCCTGTTCTTATGAATTCGAACAGCATACCCCCGATTCCGCTACGACCAA
CTCATACACCTCCTATGAAAAAACTTCCTACCACTCACCCTAGCATTACTTATATGATATGTCTCCATACCCATTACAAT
CTCCAGCATTCCCCCTCAAACCTAAGAAATATGTCTGATAAAAGAGTTACTTTGATAGAGTAAATAATAGGAGCTTAAAC
```

CCCCTTATTTCTAGGACTATGAGAATCGAACCCATCCCTGAGAATCCAAAATTCTCCGTGCCACCTATCACACCCCATCC
TAAAGTAAGGTCAGCTAAATAAGCTATCGGGCCCATACCCCGAAAATGTTGGTTATACCCTTCCCGTACTAATTAATCCC
CTGGCCCAACCCGTCATCTACTCTACCATCTTTGCAGGCACACTCATCACAGCGCTAAGCTCGCACTGATTTTTTACCTG
AGTAGGCCTAGAAATAAACATGCTAGCTTTTATTCCAGTTCTAACCAAAAAAATAAACCCTCGTTCCACAGAAGCTGCCA
TCAAGTATTTCCTCACGCAAGCAACCGCATCCATAATCCTTCTAATAGCTATCCTCTTCAACAATATACTCTCCGGACAA
TGAACCATAACCAATACTACCAATCAATACTCATCATTAATAATCATAATGGCTATAGCAATAAAACTAGGAATAGCCCC
CTTTCACTTCTGAGTCCCAGAGGTTACCCAAGGCACCCCTCTGACATCCGGCCTGCTTCTTCTCACATGACAAAAACTAG
CCCCCATCTCAATCATATACCAAATCTCTCCCTCACTAAACGTAAGCCTTCTCCTCACTCTCTCAATCTTATCCATCATA
GCAGGCAGTTGAGGTGGATTAAACCAAACCCAGCTACGCAAAATCTTAGCATACTCCTCAATTACCCACATAGGATGAAT
AATAGCAGTTCTACCGTACAACCCTAACATAACCATTCTTAATTTAACTATTTATATTATCCTAACTACTACCGCATTCC
TACTACTCAACTTAAACTCCAGCACCACGACCCTACTACTATCTCGCACCTGAAACAAGCTAACATGACTAACACCCTTA
ATTCCATCCACCCTCCTCTCCCTAGGAGGCCTGCCCCCGCTAACCGGCTTTTTGCCCAAATGGGCCATTATCGAAGAATT
CACAAAAAACAATAGCCTCATCATCCCCACCATCATAGCCACCATCACCCTCCTTAACCTCTACTTCTACCTACGCCTAA
TCTACTCCACCTCAATCACACTACTCCCCATATCTAACAACGTAAAAATAAAATGACAGTTTGAACATACAAAACCCACC
CCATTCCTCCCCACACTCATCACCCTTACCACGCTACTCCTACCTATCTCCCCTTTTATACTAATAATCTTATAGAAATT
TAGGTTAAATACAGACCAAGAGCCTTCAAAGCCCTCAGTAAGTTGCAATACTTAATTTCTGTAACAGCTAAGGACTGCAA
AACCCCACTCTGCATCAACTGAACGCAAATCAGCCACTTTAATTAAGCTAAGCCCTTACTAGACCAATGGGACTTAAACC
CACAAACACTTAGTTAACAGCTAAGCACCCTAATCAACTGGCTTCAATCTACTTCTCCCGCCGCCGGGAAAAAAGGCGGG
AGAAGCCCCGGCAGGTTTGAAGCTGCTTCTTCGAATTTGCAATTCAATATGAAAATCACCTCGGAGCTGGTAAAAAGAGG
CCTAACCCCTGTCTTTAGATTTACAGTCCAATGCTTCACTCAGCCATTTTACCTCACCCCCACTGATGTTCGCCGACCGT
TGACTATTCTCTACAAACCACAAAGACATTGGAACACTATACCTATTATTCGGCGCATGAGCTGGAGTCCTAGGCACAGC
TCTAAGCCTCCTTATTCGAGCCGAGCTGGGCCAGCCAGGCAACCTTCTAGGTAACGACCACATCTACAACGTTATCGTCA
CAGCCCATGCATTTGTAATAATCTTCTTCATAGTAATACCCATCATAATCGGAGGCTTTGGCAACTGACTAGTTCCCCTA
ATAATCGGTGCCCCCGATATGGCATTTCCCCGCATAAACAACATAAGCTTCTGACTCTTACCTCCCTCTCTCCTACTCCT
GCTCGCATCTGCTATAGTGGAGGCCGGAGCAGGAACAGGTTGAACAGTCTACCCTCCCTTAGCAGGGAACTACTCCCACC
CTGGAGCCTCCGTAGACCTAACCATCTTCTCCTTACACCTAGCAGGTGTCTCCTCTATCTTAGGGGCCATCAATTTCATC
ACAACAATTATCAATATAAAACCCCCTGCCATAACCCAATACCAAACGCCCCTCTTCGTCTGATCCGTCCTAATCACAGC
AGTCCTACTTCTCCTATCTCTCCCAGTCCTAGCTGCTGGCATCACTATACTACTAACAGACCGCAACCTCAACACCACCT
TCTTCGACCCCGCCGGAGGAGGAGACCCCATTCTATACCAACACCTATTCTGATTTTTCGGTCACCCTGAAGTTTATATT
CTTATCCTACCAGGCTTCGGAATAATCTCCCATATTGTAACTTACTACTCCGGAAAAAAAGAACCATTTGGATACATAGG
TATGGTCTGAGCTATGATATCAATTGGCTTCCTAGGGTTTATCGTGTGAGCACACCATATATTTACAGTAGGAATAGACG
TAGACACACGAGCATATTTCACCTCCGCTACCATAATCATCGCTATCCCCACCGGCGTCAAAGTATTTAGCTGACTCGCC
ACACTCCACGGAAGCAATATGAAATGATCTGCTGCAGTGCTCTGAGCCCTAGGATTCATCTTTCTTTTCACCGTAGGTGG
CCTGACTGGCATTGTATTAGCAAACTCATCACTAGACATCGTACTACACGACACGTACTACGTTGTAGCTCACTTCCACT
ATGTCCTATCAATAGGAGCTGTATTTGCCATCATAGGAGGCTTCATTCACTGATTTCCCCTATTCTCAGGCTACACCCTA
GACCAAACCTACGCCAAAATCCATTTCACTATCATATTCATCGGCGTAAATCTAACTTTCTTCCCACAACACTTTCTCGG
CCTATCCGGAATGCCCCGACGTTACTCGGACTACCCCGATGCATACGCCACATGAAACATCCTATCATCTGTAGGCTCAT
TCATTTCTCTAACAGCAGTAATATTAATAATTTTCATGATTTGAGAAGCCTTCGCTTCGAAGCGAAAAGTCCTAATAGTA
GAAGAACCCTCCATAAACCTGGAGTGACTATATGGATGCCCCCCACCCTACCACACATTCGAAGAACCCGTATACATAAA
ATCTAGACAAAAAAGGAAGGAATCGAACCCCCCAAAGCTGGTTTCAAGCCAACCCCATGGCCTCCATGACTTTTTCAAAA
AGGTATTAGAAAAACCATTTCATAACTTTGTCAAAGTTAAATTATAGGCTAAATCCTATATATCTTAATGGCACATGCAG
CGCAAGTAGGTCTACAAGACGCTACTTCCCCTATCATAGAAGAGCTTATCACCTTTCATGATCACGCCCTCATAATCATT
TTCCTTATCTGCTTCCTAGTCCTGTATGCCCTTTTCCTAACACTCACAACAAAACTAACTAATACTAACATCTCAGACGC
TCAGGAAATAGAAACCGTCTGAACTATCCTGCCCGCCATCATCCTAGTCCTCATCGCCCTCCCATCCCTACGCATCCTTT
ACATAACAGACGAGGTCAACGATCCCTCCCTTACCATCAAATCAATTGGCCACCAATGGTACTGAACCTACGAGTACACC
GACTACGGCGGACTAATCTTCAACTCCTACATACTTCCCCCATTATTCCTAGAACCAGGCGACCTGCGACTCCTTGACGT
TGACAATCGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACGTCTTGCACTCATGAGCTG
TCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACGTCTAAACCAAACCACTTTCACCGCTACACGACCGGGGGTA
TACTACGGTCAATGCTCTGAAATCTGTGGAGCAAACCACAGTTTCATGCCCATCGTCCTAGAATTAATTCCCCTAAAAAT
CTTTGAAATAGGGCCCGTATTTACCCTATAGCACCCCCTCTACCCCCTCTAGAGCCCACTGTAAAGCTAACTTAGCATTA
ACCTTTTAAGTTAAAGATTAAGAGAACCAACACCTCTTTACAGTGAAATGCCCCAACTAAATACTACCGTATGGCCCACC
ATAATTACCCCCATACTCCTTACACTATTCCTCATCACCCAACTAAAAATATTAAACACAAACTACCACCTACCTCCCTC
ACCAAAGCCCATAAAAATAAAAAATTATAACAAACCCTGAGAACCAAAATGAACGAAAATCTGTTCGCTTCATTCATTGC
CCCCACAATCCTAGGCCTACCCGCCGCAGTACTGATCATTCTATTTCCCCCTCTATTGATCCCCACCTCCAAATATCTCA
TCAACAACCGACTAATCACCACCCAACAATGACTAATCAAACTAACCTCAAAACAAATGATAACCATACACAACACTAAA
GGACGAACCTGATCTCTTATACTAGTATCCTTAATCATTTTTATTGCCACAACTAACCTCCTCGGACTCCTGCCTCACTC
ATTTACACCAACCACCCAACTATCTATAAACCTAGCCATGGCCATCCCCTTATGAGCGGGCGCAGTGATTATAGGCTTTC
GCTCTAAGATTAAAAATGCCCTAGCCCACTTCTTACCACAAGGCACACCTACACCCCTTATCCCCATACTAGTTATTATC
GAAACCATCAGCCTACTCATTCAACCAATAGCCCTGGCCGTACGCCTAACCGCTAACATTACTGCAGGCCACCTACTCAT
GCACCTAATTGGAAGCACCACCCTAGCAATATCAACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTC
TACTGACTATCCTAGAAATCGCTGTCGCCCTAATCCAAGCCTACGTTTTCACACTTCTAGTAAGCCTCTACCTGCACGAC
AACACATAATGACCCACCAATCACATGCCTATCATATAGTAAAACCCAGCCCATGACCCCTAACAGGGGCCCTCTCAGCC
CTCCTAATGACCTCCGGCCTAGCCATGTGATTTCACTTCCACTCCATAACGCTCCTCATACTAGGCCTACTAACCAACAC
ACTAACCATATACCAATGATGGCGCGATGTAACACGAGAAAGCACATACCAAGGCCACCACACACCACCTGTCCAAAAAG
GCCTTCGATACGGGATAATCCTATTTATTACCTCAGAAGTTTTTTTCTTCGCAGGATTTTTCTGAGCCTTTTACCACTCC
AGCCTAGCCCCTACCCCCCAATTAGGAGGGCACTGGCCCCCAACAGGCATCACCCCGCTAAATCCCCTAGAAGTCCCACT
CCTAAACACATCCGTATTACTCGCATCAGGAGTATCAATCACCTGAGCTCACCATAGTCTAATAGAAAACAACCGAAACC
AAATAATTCAAGCACTGCTCATTACAATTTTACTGGGTCTCTATTTTACCCTCCTACAAGCCTCAGAGTACTTCGAGTCT
CCCTTCACCATTTCCGACGGCATCTACGGCTCAACATTTTTTGTAGCCACAGGCTTCCACGGACTTCACGTCATTATTGG
CTCAACTTTCCTCACTATCTGCTTCATCCGCCAACTAATATTTCACTTTACATCCAAACATCACTTTGGCTTCGAAGCCG
CCGCCTGATACTGGCATTTTGTAGATGTGGTTTGACTATTTCTGTATGTCTCCATCTATTGATGAGGGTCTTACTCTTTT

AGTATAAATAGTACCGTTAACTTCCAATTAACTAGTTTTGACAACATTCAAAAAAGAGTAATAAACTTCGCCTTAATTTT
AATAATCAACACCCTCCTAGCCTTACTACTAATAATTATTACATTTTGACTACCACAACTCAACGGCTACATAGAAAAAT
CCACCCCTTACGAGTGCGGCTTCGACCCTATATCCCCCGCCCGCGTCCCTTTCTCCATAAAATTCTTCTTAGTAGCTATT
ACCTTCTTATTATTTGATCTAGAAATTGCCCTCCTTTTACCCCTACCATGAGCCCTACAAACAACTAACCTGCCACTAAT
AGTTATGTCATCCCTCTTATTAATCATCATCCTAGCCCTAAGTCTGGCCTATGAGTGACTACAAAAAGGATTAGACTGAG
CCGAATTGGTATATAGTTTAAACAAAACGAATGATTTCGACTCATTAAATTATGATAATCATATTTACCAAATGCCCCTC
ATTTACATAAATATTATACTAGCATTTACCATCTCACTTCTAGGAATACTAGTATATCGCTCACACCTCATGTCCTCCCT
ACTATGCCTAGAAGGAATAATACTATCGCTGTTCATTATAGCTACTCTCATAACCCTCAACACCCACTCCCTCTTAGCCA
ATATTGTGCCTATTGCCATACTAGTCTTTGCCGCCTGCGAAGCAGCGGTGGGCCTAGCCCTACTAGTCTCAATCTCCAAC
ACATATGGCCTAGACTACGTACATAACCTAAACCTACTCCAATGCTAAAACTAATCGTCCCAACAATTATATTACTACCA
CTGACATGACTTTCCAAAAAACACATAATTTGAATCAACACAACCACCCACAGCCTAATTATTAGCATCATCCCTCTACT
ATTTTTTAACCAAATCAACAACAACCTATTTAGCTGTTCCCCAACCTTTTCCTCCGACCCCCTAACAACCCCCCTCCTAA
TACTAACTACCTGACTCCTACCCCTCACAATCATGGCAAGCCAACGCCACTTATCCAGTGAACCACTATCACGAAAAAAA
CTCTACCTCTCTATACTAATCTCCCTACAAATCTCCTTAATTATAACATTCACAGCCACAGAACTAATCATATTTTATAT
CTTCTTCGAAACCACACTTATCCCCACCTTGGCTATCATCACCCGATGAGGCAACCAGCCAGAACGCCTGAACGCAGGCA
CATACTTCCTATTCTACACCCTAGTAGGCTCCCTTCCCCTACTCATCGCACTAATTTACACTCACAACACCCTAGGCTCA
CTAAACATTCTACTACTCACCCTCACTGCCCAAGAACTATCAAACTCCTGAGCCAACAACTTAATATGACTAGCTTACAC
AATAGCTTTTATAGTAAAGATACCTCTTTACGGACTCCACTTATGACTCCCTAAAGCCCATGTCGAAGCCCCCATCGCTG
GGTCAATAGTACTTGCCGCAGTACTCTTGAAACTAGGCGGCTATGGTATAATACGCCTCACACTCATTCTCAACCCCCTG
ACAAAACACATAGCCTACCCCTTCCTTGTACTATCCCTATGAGGCATAATTATAACAAGCTCCATCTGCCTACGACAAAC
AGACCTAAAATCGCTCATTGCATACTCTTCAATCAGCCACATAGCCCTCGTAGTAACAGCCATTCTCATCCAAACCCCCT
GAAGCTTCACCGGCGCAGTCATTCTCATAATCGCCCACGGACTTACATCCTCATTACTATTCTGCCTAGCAAACTCAAAC
TACGAACGCACTCACAGTCGCATCATAATCCTCTCTCAAGGACTTCAAACTCTACTCCCACTAATAGCTTTTTGATGACT
TCTAGCAAGCCTCGCTAACCTCGCCTTACCCCCCACTATTAACCTACTGGGAGAACTCTCTGTGCTAGTAACCACGTTCT
CCTGATCAAATATCACTCTCCTACTTACAGGACTCAACATACTAGTCACAGCCCTATACTCCCTCTACATATTTACCACA
ACACAATGGGGCTCACTCACCCACCACATTAACAACATAAAACCCTCATTCACACGAGAAAACACCCTCATGTTCATACA
CCTATCCCCCATTCTCCTCCTATCCCTCAACCCCGACATCATTACCGGGTTTTCCTCTTGTAAATATAGTTTAACCAAAA
CATCAGATTGTGAATCTGACAACAGAGGCTTACGACCCCTTATTTACCGAGAAAGCTCACAAGAACTGCTAACTCACGCC
CCCATGTCTAACAACATGGCTTTCTCAACTTTTAAAGGATAACAGCTATCCATTGGTCTTAGGCCCCAAGAATTTTGGTG
CAACTCCAAATAAAAGTAATAACCATGCACACTACTATAACCACCCTAACCCTAACTTCCCTAATTCCCCCCATCCTTAC
CACCCTCGTTAACCCTAACAAAAAAAACTCATACCCCCATTATGTAAAATCCATTGTCGCATCCACCTTTATTATCAGTC
TCTTCCCCACAACAATATTCATGTGCCTAGACCAAGAAGTTATTATCTCGAACTGACACTGAGCCACAACCCAAACAACC
CAGCTCTCCCTAAGCTTCAAACTAGACTACTTCTCCATAATATTCATCCCTGTAGCATTGTTCGTTACATGGTCCATCAT
AGAATTCTCACTGTGATATATAAACTCAGACCCAAACATTAATCAGTTCTTCAAATATCTACTCATCTTCCTAATTACCA
TACTAATCTTAGTTACCGCTAACAACCTATTCCAACTGTTCATCGGCTGAGAGGGCGTAGGAATTATATCCTTCTTGCTC
ATCAGTTGATGATACGCCCGAGCAGATGCCAACACAGCAGCCATTCAAGCAATCCTATACAACCGTATCGGCGATATCGG
TTTCATCCTCGCCTTAGCATGATTTATCCTACACTCCAACTCATGAGACCCACAACAAATAGCCCTTCTAAACGCTAATC
CAAGCCTCACCCCACTACTAGGCCTCCTCCTAGCAGCAGCAGGCAAATCAGCCCAATTAGGTCTCCACCCCTGACTCCCC
TCAGCCATAGAAGGCCCCACCCCAGTCTCAGCCCTACTCCACTCAAGCACTATAGTTGTAGCAGGAATCTTCTTACTCAT
CCGCTTCCACCCCCTAGCAGAAAATAGCCCACTAATCCAAACTCTAACACTATGCTTAGGCGCTATCACCACTCTGTTCG
CAGCAGTCTGCGCCCTTACACAAAATGACATCAAAAAAATCGTAGCCTTCTCCACTTCAAGTCAACTAGGACTCATAATA
GTTACAATCGGCATCAACCAACCACACCTAGCATTCCTGCACATCTGTACCCACGCCTTCTTCAAAGCCATACTATTTAT
GTGCTCCGGGTCCATCATCCACAACCTTAACAATGAACAAGATATTCGAAAAATAGGAGGACTACTCAAAACCATACCTC
TCACTTCAACCTCCCTCACCATTGGCAGCCTAGCATTAGCAGGAATACCTTTCCTCACAGGTTTCTACTCCAAAGACCAC
ATCATCGAAACCGCAAACATATCATACACAAACGCCTGAGCCCTATCTATTACTCTCATCGCTACCTCCCTGACAAGCGC
CTATAGCACTCGAATAATTCTTCTCACCCTAACAGGTCAACCTCGCTTCCCCACCCTTACTAACATTAACGAAAATAACC
CCACCCTACTAAACCCCATTAAACGCCTGGCAGCCGGAAGCCTATTCGCAGGATTTCTCATTACTAACAACATTTCCCCC
GCATCCCCCTTCCAAACAACAATCCCCCTCTACCTAAAACTCACAGCCCTCGCTGTCACTTTCCTAGGACTTCTAACAGC
CCTAGACCTCAACTACCTAACCAACAAACTTAAAATAAAATCCCCACTATGCACATTTTATTTCTCCAACATACTCGGAT
TCTACCCTAGCATCACACACCGCACAATCCCCTATCTAGGCCTTCTTACGAGCCAAAACCTGCCCCTACTCCTCCTAGAC
CTAACCTGACTAGAAAAGCTATTACCTAAAACAATTTCACAGCACCAAATCTCCACCTCCATCATCACCTCAACCCAAAA
AGGCATAATTAAACTTTACTTCCTCTCTTTCTTCTCCCACTCATCCTAACCCTACTCCTAATCACATAACCTATTCCCC
CGAGCAATTTCAATTACAATATATACACCAACAAACAATGTTCAACCAGTAACTACTACTAATCAACGCCCATAATCATA
CAAAGCCCCCGCACCAATAGGATCCTCCCGAATCAACCCTGACCCCTCTCCTTCATAAATTATTCAGCTTCCTACACTAT
TAAAGTTTACCACAACCACCACCCCATCATACTCTTTCACCCACAGCACCAATCCTACCTCCATCGCTAACCCCACTAAA
ACACTCACCAAGACCTCAACCCCTGACCCCCATGCCTCAGGATACTCCTCAATAGCCATCGCTGTAGTATATCCAAAGAC
AACCATCATTCCCCCTAAATAAATTAAAAAAACTATTAAACCCATATAACCTCCCCCAAAATTCAGAATAATAACACACC
CGACCACACCGCTAACAATCAATACTAAACCCCCATAAATAGGAGAAGGCTTAGAAGAAAACCCCACAAACCCCATTACT
AAACCCACACTCAACAGAAACAAAGCATACATCATTATTCTCGCACGGACTACAACCACGACCAATGATATGAAAAACCA
TCGTTGTATTTCAACTACAAGAACACCAATGACCCCAATACGCAAAATTAACCCCCTAATAAAATTAATTAACCACTCAC
TCATCGACCTCCCCACCCCATCCAACATCTCCGCATGATGAAACTTCGGCTCACTCCTTGGCGCCTGCCTGATCCTCCAA
ATCACCACAGGACTATTCCTAGCCATGCACTACTCACCAGACGCCTCAACCGCCTTTTCATCAATCGCCCACATCACTCG
AGACGTAAATTATGGCTGAATCATCCGCTACCTTCACGCCAATGGCGCCTCAATATTCTTTATCTGCCTCTTCCTACACA
TCGGGCGAGGCCTATATTACGGATCATTTCTCTACTCAGAAACCTGAAACATCGGCATTATCCTCCTGCTTGCAACTATA
GCAACAGCCTTCATAGGCTATGTCCTCCCGTGAGGCCAAATATCATTCTGAGGGGCCACAGTAATTACAAACTTACTATC
CGCCATCCCATACATTGGGACAGACCTAGTTCAATGAATCTGAGGAGGCTACTCAGTAGACAGTCCCACCCTCACACGAT
TCTTTACCTTTCACTTCATCTTGCCCTTCATTATTGCAGCCCTAGCAGCACTCCACCTCCTATTCTTGCACGAAACGGGA
TCAAACAACCCCCTAGGAATCACCTCCCATTCCGATAAAATCACCTTCCACCCTTACTACACAATCAAAGACGCCCTCGG
CTTACTTCTCTTCCTTCTCTCCTTAATGACATTAACACTATTCTCACCAGACCTCCTAGGCGACCCAGACAATTATACCC
TAGCCAACCCCTTAAACACCCCTCCCCACATCAAGCCCGAATGATATTTCCTATTCGCCTACACAATTCTCCGATCCGTC
CCTAACAAACTAGGAGGCGTCCTTGCCCTATTACTATCCATCCTCATCCTAGCAATAATCCCCATCCTCCATATATCCAA

ACAACAAAGCATAATATTTCGCCCACTAAGCCAATCACTTTATTGACTCCTAGCCGCAGACCTCCTCATTCTAACCTGAA
TCGGAGGACAACCAGTAAGCTACCCTTTTACCATCATTGGACAAGTAGCATCCGTACTATACTTCACAACAATCCTAATC
CTAATACCAACTATCTCCCTAATTGAAAACAAAATACTCAAATGGGCCTGTCCTTGTAGTATAAACTAATACACCAGTCT
TGTAAACCGGAAATGAAAACCTTTTTCCAAGGACAAATCAGAGAAAAAGTCTTTAACTCCACCATTAGCACCCAAAGCTA
AGATTCTAATTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCACCCATCAA
CAACCGCTATGTATCTCGTACATTACTGCCAGCCACCATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTA
CATAAAAACCCAATCCACATTAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCCCCAACTATCACACAT
CAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATACCAACAAACCTACCCACCCTTAACAGTACATAGCACATAAA
GCCATTTACCGTACATAGCACATTACAGTCAAATCCCTTCTCGTCCCCATGGATGACCCCCCTCAGATAGGGGTCCCTTG
ACCACCATCCTCCGTGAAATCAATATCCCGCACAAGAGTGCTACTCTCCTCGCTCCGGGCCCATAACACTTGGGGGTAGC
TAAAGTGAACTGTATCCGACATCTGGTTCCTACTTCAGGGCCATAAAGCCTAAATAGCCCACACGTTCCCCTTAAATAAG
ACATCACGATG