Mitochondrial DNA

Abstract

The experiment examined a 440-nucleotide sequence from a noncoding region of mitochondrial genome. Once the product was amplified using an automated thermal cycler, the high yield of product amplified would be visualized in an agarose gel after gel electrophoresis was performed. The same region of the genome was amplified by each student thus making the gel electrophoresis similar but the amplified student samples differed from each other once they were sequenced. A comparison of control region sequences revealed that most individuals have a special pattern of single nucleotide polymorphisms (SNPs). These sequence differences served as the foundation for multiple investigations on human DNA diversity and the evolution of hominids (i.e. individual comparisons to Neanderthals). It was concluded that the average pairwise sequence divergence among students in the Genetics class for the D-loop sequences was 0.017. The estimated time since we all shared a common mitochondrial haplotype was 236,096.262 years and there was confirmation that the displacement ("Out of Africa") model of human evolution was more consistent with the class data.

Introduction

The purpose of the experiment was to analyze mitochondrial DNA by means of sequenced data in the hopes of estimating species divergence times along with determining geographic origins. In mammals, mitochondrial DNA (mtDNA) is maternally inherited. There is no recombination of alleles as it’s inherited as a single unit, a high display of mutation rates and the ability to trace maternal lineages back in time to a common ancestor. Between 200 and 1700 mitochondria are present in each human somatic cell, hence extraction from the cheek cells (Klug et al., 2012). Mitochondria divide inside cells and are distributed to daughter cells after cell division. They’re passed from the human egg cell to the zygote during fertilization where sperm cells contribute very few mitochondria to the zygote and don’t contribute such organelles to the next generation (Klug et al., 2012). As a consequence, an individual’s cells contain many copies of identical mitochondria derived from the mother.

Mitochondria’s role in cellular energy production is very critical since these organelles are responsible for the production of adenosine triphosphate (ATP). The DNA in mitochondria provides the cell with many of the polypeptides necessary for the oxidative phosphorylation system (Fernandez et al 2003). One part of the mitochondrial genome in particular, named the D-loop (displacement loop), is essential in regulating other familiar cellular processes such as initiating replication and transcription (Fernandez et al 2003). The mitochondrial DNA’s contributions to the overall make-up of nuclear biology have become increasingly evident through research and experimentation, where a deeper understanding of its structural components and processes is vital in assessing its role in the "molecular clock" and mutation rates of cellular DNA as a whole (Fernandez et al 2003). This understanding may reveal pertinent answers to
questions regarding cellular evolutionary history, since the time of divergence between two pieces of DNA may be assumed from knowledge of their respective mutation rates along with the time these mutations generated (Rieux et al 2014). This tactic is referred as the "molecular clock" and it’s applicable in distinguishing the various stages of human evolution (Rieux et al 2014). It may even contribute to the analysis of the millions-year old beginning of humankind itself based on the history of the genome, that consists of the history of the mitochondrial genome itself (Rieux et al 2014). This where the rise of mtDNA in technology comes into play, where sequencing specific regions such as the D-loop in databases such as NCBI’s GenBank contributes towards homology searching. For the purpose of this experiment, homology searching helped verify which models of human evolution was most accurate. The multiregional model believes that humans developed simultaneously from many different archaic populations living in different parts of the world. An assumption of this model is that Neanderthals were modern European’s ancestors, while Java men were modern Asian’s ancestors (Walters-Conte 2015). The most recent common ancestor of modern humans according to this model are on the 1.5-3 million years scale, meaning that some human sequences should be more similar to Neanderthal DNA compared to other humans (Walters-Conte 2015). The other evolutionary model looked at is the Displacement model ("Out of Africa") where humans derived from a single founding population that emerged Africa 200,000-500,000 years ago (Walters-Conte 2015). This group of individuals successively migrated to Europe and Asia and displaced archaic hominids. This model asserts that the divergence time between modern humans and Neanderthals should surpass what’s observed between any two modern human beings (Walters-Conte 2015).

**Methods & Materials**

The experiment began with isolating genomic DNA via extraction. Two 1.5 ml tubes were labeled with permanent marker with a last name. 10 ml of saline solution (0.9% NaCl) was poured into the mouth and vigorously swished around for 30 seconds. The saline solution was then expelled into a paper cup and swirled to mix the cells up. 1.5 mL of the liquid was transferred to a 1.5 mL tube. The sample tube was placed in a balanced microcentrifuge and spun for 5 minutes at maximum speed. The supernatant was carefully removed into a paper cup making sure not to disturb the cell pellet at the bottom of the test tube. A small amount of saline should have remained in the tube. Cells were then resuspended in the remaining saline by pipetting in and out and transferred to a 0.5 mL microcentrifuge tube containing 100 µl of 10% Chelex solution. The tube was labeled with a last name and vortexed to ensure cells were completely suspended in the Chelex solution. The tube containing the cell sample was then heated in a thermocycler for 10 minutes at 100°C. After heating, the tube was shook and placed in a balanced microcentrifuge so that it could be spun for 1 minute. 50 µl of supernatant (containing the DNA) was transferred using a micropipettor to a clean labeled 1.5 mL tube containing the pellet (cell debris and Chelex beads). This sample was used for setting up Polymerase Chain Reactions (PCR). The sample was stored in a freezer until the following week for PCR amplification.

The top and side of a 500 µl tube containing the Read-to-Go PCR be labeled with the last name. A 200 µl micropipettor was used to add entire contents (22.5 µl) of the primer/ddH2O mix
to a PCR tube containing a Ready-To-Go PCR Bead; direct contact with the bead was avoided and the tube was tapped with a finger to dissolve the bead. A 20 µl micropipettor with a fresh tip was used to transfer 2.5 µl of gnomic DNA to a reaction tube and tapped to mix. Reagents were pooled by pulsing in a microcentriguge. The sample was then stored on ice until it was ready to be amplified. The thermal cycler was programmed for 30 cycles, each program linked to a 4°C to hold samples after completing the cycle profile but amplified DNA was also held at room temperature. The cycle profile consisted of: an initial denature at 94°C for 2 minutes followed by a denature at 94°C for 30 seconds. Then annealing occurred at 58°C for 30 seconds and extended at 72°C for 30 seconds. Denaturing at 94°C for 30 seconds occurred 30 times which then lead to a final extension at 72°C for 6 minutes. Lastly there was an indefinite hold at 4°C. Successful amplification of the DNA would be seen the following when a gel electrophoresis was conducted on the agarose gel.

After the experimental procedures were performed, analytical procedures followed. An evolutionary analysis was made by determining the average proportional nucleotide sequence divergence in the mitochondrial DNA control region between Neanderthal sequences and modern humans sequences by using data provided in an Excel file. Thereafter, the average pairwise sequence divergence between humans and chimpanzees was calculated. The average proportional divergences ($p_d$), were also calculated by taking the average of the relevant numbers and converted to the number of substitutions per site, $K_0 (= -\ln [1 - p_d])$. The average proportional nucleotide sequence divergence between modern humans was determined by taking the average of all pairwise sequence comparisons between modern humans and then converting the proportional distance to the average number of substitutions per site, $K_{HH}$. The corrected divergence in the control region sequence provided a measure of the genetic distance between populations. The human-chimp divergence time (5,000,000 years) and the average number of substitutions per site for the human-chimp comparison ($K_{HC}$ from 1b) were used to calculate the rate of nucleotide substitution under the assumption that mitochondrial mutations occurred at a constant rate. Now the divergence time for the most recent common ancestor to all living modern humans ($T_{MRCA}$) could be estimated due to the calculated rate of substitution. From this value one could determine which model of human evolution was more consistent with the estimated modern human divergence, choosing from either the multiregional or displacement models. The divergence time between the Neanderthals and modern humans was calculated using the established date of human divergence of 200,000 years as a calibration point, which provided a more accurate estimate of years per substitution/site. The divergence time, in years, was calculated by dividing $K_{HN}$, the average number of substitutions per site between modern humans and Neanderthals. This same approach was used to determine the lab partners last shared a common ancestor using the K value obtained from $p_d$ between the two sequences in the calculation.

In this last section of the experiment, nucleotide sequences in a database were searched for. An assigned student number associated with personal D-loop DNA sequences. The nucleotide sequence was selected for and copied into a nucleotide BLAST in the GenBank of the NCBI homepage, allowing the D-loop nucleotide sequence to be searched against all other nucleotide sequences in the database. The default settings were changed to “Nucleotide collection (nr/nt)”, Organism: “Homo sapiens”, and Optimize for: “Somewhat similar sequences (blastn)”. After the sequence was entered to search the database for DNA similar to our own. The large yellow box was clicked in under the label “Entry Query Sequence” where our personal DNA sequence was pasted. A blastn search identified the nucleotide sequences (nr) in the
database that were similar to the query sequence entered as input data. The blue “BLAST” was clicked on to submit the search and place the query into the blast Queue where a new page appeared with a request ID.

A “graphic summary” indicating the size of the region and degree of homology shared between the query sequence and database entries after the search had been submitted. A list containing the top 100 DNA sequences, which were similar to the query sequence, was displayed where each entry had an Accession Number linked to a database entry. To the right of each matches’ short descriptions was a “Max Score (S)”. The larger this alignment score, the more significant the match; and the closer the “E value” was to 0, the more significant the match. Generally, E scores greater than 0.05 indicate that the match was due to homology. When the sequence description hypertext was clicked on, it took us to the pairwise alignment between the query and the specific database entry questioned. A series of sequence alignment showing specific nucleotide positions that were shared between the query sequence and the database entry was listed below the list of matches. The two sequences that had the greatest similarity to the query sequence based on their “Max Score” were picked and their Genbank accession numbers and brief description of the database entries were recorded. The nucleotide positions associated with each region of similarity shared between the query and database sequences were also recorded. Then the number of nucleotide positions in the query sequence that were identical to the database entry for each of the two matches was recorded so that the percent identity could be calculated (\(= 100 \times \#\text{identical}/\text{total \#bp in query}\)).

Lastly, the links for several of the sequences listed above were clicked on to view the complete database entry. Each entry contained a lot of information regarding publication reference information, coding sequence position of the genes, predicted polypeptide sequence of the gene product, etc. (Walters-Conte 2015).

**Results**

The extracted mtDNA was entered into the NCBI database and matched with multiple DNA sequences. The top sequence had a 97% identity, a length of 355 bp, and an E-value of \(6 \times 10^{-147}\) corresponding to significance. This data is depicted in Table 1, displaying the accurate similarity of the matching sequence was to the inputted one. Accession numbers for the top two greatest matches were recorded along with the geographic origin of the first top match, which turned out to be Neander Valley, Germany. Table 2 displays the average proportional nucleotide divergence in the mtDNA control region between Neanderthal and modern human sequences; along with the average pairwise sequence divergence between humans and the chimpanzee. The calculated average proportional divergence between the human and the Neanderthal was 0.047 and the average number of substitutions per site between humans and Neanderthals calculated to be 0.048 substitutions per site. The average proportional divergence between humans and chimps calculated out to 0.155 and the average number of substitutions per site was 0.169. Comparison between the \(K\) values of humans and Neanderthals and humans and chimps depicted that the average number of substitutions between humans and chimps was greater. This shows that chimps and modern human species are more diverged and less closely related than humans and Neanderthals. The
average proportion divergence between two modern humans was 0.017 with an average number of substitutions per site of 0.017 substitutions per site. This was the expected small value, verifying a small amount of divergence between the humans, Neanderthals, and chimpanzees. 3.37x10^-8 substitutions per site/year represented how many substitutions per year took place between the divergence time between humans and chimps and their respective $K$ values. This rate was used to calculate the estimated number of years it took for all modern humans divergence from their recent common ancestor which was estimated to be 507461.820 years. This method was also used with different values in order to estimate the amount of years it took for the divergence of modern humans and Neanderthals that turned out to be 558111.086 years. There was a greater divergence between Neanderthals and humans. The divergence in years between the most recent ancestor to all modern humans was smaller indicating a smaller genetic difference between modern humans and Neanderthals. The years of divergence of the last common ancestor between individual students, modern humans, was 236096.262 years, which was the least amount of years calculated, again showing the least genetic difference. All of this data is illustrated in Table III which ultimately supported the displacement model of human evolution.

**Table I: Summary of BLAST search of D-loop sequence**

<table>
<thead>
<tr>
<th>D-loop Sequence</th>
<th>Length of match</th>
<th>Percent Identity of top match</th>
<th>E-score of 1st top match</th>
<th>Accession Number of 1st top match</th>
<th>E-score of 2nd top match</th>
<th>Accession Number of 2nd top match</th>
<th>Geographic Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 3-6 sequence (in blackbord text file)</td>
<td>355 bp</td>
<td>97%</td>
<td>6 * 10^{-147}</td>
<td>KJ446515.1</td>
<td>4 * 10^{-136}</td>
<td>IQ045085.1</td>
<td>Neander Valley, Germany</td>
</tr>
</tbody>
</table>

**Table II: Pairwise Comparisons (chimp and Neanderthal sequences) and with other student sequences**

<table>
<thead>
<tr>
<th>Average proportional divergence (human vs. Neanderthal)</th>
<th>Average number of substitutions per site, $K_{HN}$ (human vs. Neanderthal)</th>
<th>Average proportional divergence (human vs. chimp)</th>
<th>Average number of substitutions per site (human vs. Chimp)</th>
<th>Average proportional divergence (human vs. human)</th>
<th>Average number of substitutions per site (human vs. human)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.047</td>
<td>0.048</td>
<td>0.155</td>
<td>0.169</td>
<td>0.017</td>
<td>0.017</td>
</tr>
</tbody>
</table>
Table III: Computation and Presentation of estimated divergence times

| Rate of nucleotide substitution using the human-chimp divergence time and the average number of substitutions per site for human-chimp comparison, $K_{HC}$ | Estimated divergence time for the most recent common ancestor to all living humans, $T_{MRCA}$ | Divergence time between Neanderthals and modern humans using modern human divergence date, 200,000 years | Divergence time between modern humans and Neanderthals | Divergence years of last shared common ancestor between individuals of students with $p_d=0.020$ and $K=0.020$
|---|---|---|---|---|
| $\frac{K_{HC}}{5,000,000}$ | $T_{MRCA} = \frac{K_{HC}}{3.37 \times 10^{-8}}$ | $\text{average } K_{HN} = \frac{0.017}{200,000}$ | $\frac{K_{HN}}{8.56 \times 10^{-8}} = 0.048$ | $\frac{p_d}{0.025+0.015} = 0.20 = K$
| $= \frac{0.155}{5,000,000}$ | $= \frac{0.017}{3.37 \times 10^{-8}}$ | | $= \frac{8.56 \times 10^{-8}}{0.20}$ | |
| $3.37 \times 10^{-8}$ substitutions per site/year | 507461.820 years | $8.56 \times 10^{-8}$ substitutions per site/year | 558111.086 years | 236096.262 years |

Discussion

Human evolution is very complex as seen by the numerous calculations worked out in this lab. But it's necessary for the advancement of the human race where analysis of genetics sequences plays a crucial part. Base substitutions in mitochondrial genomes help to represent human evolution over a significant time frame within and out of species. This lab had a primary focus on mtDNA comparison between Neanderthals, modern humans, the last common ancestor and chimpanzees. These comparisons helped to emphasize the importance of genetic variance and how divergence of species is an ultimate consequence. Human evolution helps to both distinguish and connect individuals that is greatly supported by the data displayed in Table III. Modern humans can be traced back to a shared common ancestor approximately 236096.262 years. Given the old age of
planet earth, the number is relatively recent. What’s interesting about human evolution is that it’s continuing to evolve as time progresses as seen by the calculated substitutions per year in Table III. Due to frequent mutations causing genetic variance, the human race continues its rate of evolving. This genetic variance is largely due to recombination of DNA.

The experiment helped to verify the displacement model as being the model of human evolution and reject the multiregional model. The multiregional model essentially asserted that humans developed simultaneously from different populations in different areas of the world. Contrary, the displacement model asserted that human evolution derived from one origin, Africa. This human population surfaced 200,000-500,000 years ago and would then migrate to other continents to displace themselves and initiate a new population. Both models of human evolution may differ but share the viewpoint of common ancestor giving rise to modern humans; which subliminally suggests that evolution is heritable seen by small commonalities between species.

Again, the collected data supported displacement model theory. This model suggested that the divergence time between modern humans and Neanderthals would be greater than the time calculated between two modern humans. Table III shows the divergence time between modern humans and Neanderthals as 558111.086 years, which exceeded the divergence time between two modern humans, 236096.262 years. Table II serves as additional support to this theory where the average proportional divergence between modern humans was 0.17 substitutions per site (smaller than the average proportional divergence between modern humans and Neanderthals, 0.48 substitutions per site). The fewer amount of substitutions between modern humans to modern humans showed that there was less genetic difference and that there was a greater similarity to Neanderthals to modern humans.

References

