**Tracing the Human Lineage Through the GABRA1 Gene**

Alex Rance

**Introduction**

Gamma-aminobutyric acid type A receptor subunit alpha1, or GABRA1, is a gene that provides instruction for making a singular piece, the alpha-1 subunit of the GABAA receptor protein. Gamma-amino butyric acid will attach to GABAA receptors and chlorine ions flow across the cell membrane. The role of gamma-amino butyric acid, GABA, in both children and adults is to prevent the brain from becoming overwhelmed with signals. Mutations within this gene have shown an increase in epilepsy.

Humans have inherited many traits from our more primal, monkey-like ancestors. This paper will attempt to use the idea that humans inherited the GABRA1 gene from our ancestors. Through phylogenetic analysis, the most likely path of evolution of humans is going to be seen. This paper will be using ten different species of primates. These 10 species are all located within the Order of Primates and within the Suborder of Haplorhini. The 10 species can then be lowered into two Parvorders, the Platyrrhini and the Catarrhini. The Platyrrhini is known as the New World Monkeys while the Catarrhini is known as the Old-World Monkeys. The species located within the Platyrrhini Parvorder is the Callithrix jacchus, or the white-tufted-ear marmoset which is in the Callitrichidae Family. The remaining 9 species are in the Parvorder Catarrhini. The Papio anubis, the baboon, the Macaca fascicularis, crab-eating macaque, and the Macaca mulatta, the rhesus macaque, all belonging to family of Cercopithecidae. There are two Families within the other Superfamily that is located within the Parvorder Catarrhini. The Nomascus leucogenys, or the northern white-cheeked gibbon is within the Hylobatidae Family. Homo sapiens, humans, Pan paniscus, bonobo, Pan troglodytes, chimpanzee, Gorilla gorilla, western lowland gorilla, and Pongo abelii, Sumatran orangutan, are all located within the Family Hominidae.

**Methods**

To get all the DNA that will be used in this paper had to be obtained from the sequences of all the species mentioned above. To get the sequences, the human genome coordinates for the GABRA1 gene was converted to the corresponding coordinates within each species using NCBI Genome Browser and the UCSC Genome Browser genome convert tool. After every DNA sample from all 10 species was obtained, they had to be aligned. MEGA was used for all the experiments and tests done for this paper. Due to the different lengths of each species GABRA1 gene, in order for the alignment to work properly, the length of each sequence had to be made equal. Therefore, I deleted every nucleotide after the shortest DNA sequence ended from every other species. MEGA was used to run every phylogenetic test to determine the probable mechanism of evolution present within these ten species.

**Results**

To start the experiments, I took the aligned data and set it to create a tree based on a random tree generated from the data available. An outgroup was specified, which in our case is the white-tufted-ear marmoset, and the tree was rooted in relation to the outgroup. The marmoset was chosen as the outgroup because it is the only taxa located within the Parvorder Platyrrhini. This was computed just to see how the MEGA program would interpret the data available. This tree is shown below in Figure 1.

Text

Description automatically generatedFigure 1. This tree was generated via a random assortment of the data available whilst setting the marmoset as the outgroup.

The next tree that was generated is the tree based on the Maximum Parsimony analysis. I did not apply Bootstrap to create consensus trees. Once again, I used the white-tufted-ear marmoset as the outgroup. The following tree made was created using the Neighbor-Joining method. The evolutionary distances were computed using the number of differences, ie the number of transitions and transversions.

Chart

Description automatically generatedFigure 2. This tree was generated via the Maximum Parsimony method with no bootstrap applied to the tree.

Diagram

Description automatically generatedFigure 3. This tree was generated via the Neighbor-Joining method with no bootstraps applied.

The next generated trees are where the trees start to get more complex. The first of the two trees shown below in Figure 3 is the original tree with the highest log likelihood. The numbers by each associated clusters is the percentage of trees in which those associated taxa are clustered together. Figure 3b shows the Bootstrap consensus Tree that was generated through the Maximum Likelihood method as well as the Jukes-Cantor model and was compiled after 500 bootstrap replicates.

Diagram

Description automatically generatedDiagram

Description automatically generated

Figure 4. The tree on the left, 4a, is showing the tree with the highest log likelihood in accordance with the Maximum Likelihood method. The tree on the right, 4b, is showing the most likely tree after 500 bootstrap replicated were completed with the numbers showing the percent of bootstrap replicates where that particular relationship occurred.

The final series of trees is very similar to the trees shown in Figure 3. However, instead of basing the trees on Maximum Likelihood, these trees used Maximum Parsimony. Just like the trees in Figure 3, Figure 4a shows the evolutionary history that has been inferred using the Maximum Parsimony method. Figure 4b shows the Bootstrap consensus tree after 500 replicates.

A picture containing table

Description automatically generatedDiagram

Description automatically generated

Figure 5. The tree on the left, 5a, is showing the tree with the highest parsimony score in accordance with the Maximum Parsimony method. The tree on the right, 5b, is showing the most likely tree after 500 bootstrap replicates were constructed with the numbers representing the percentage of bootstrap trees where that particular relationship is present.

**Discussion**

Using the 10 taxa for this experiment, I created a tree using the Marmoset as the outgroup and created the most likely order of ancestry to compare the results of the experiments with. The tree below shows the most likely family tree of the taxa included in this study. That tree is based off of the current understanding of how science believes the evolution of Apes came around. Figure 1 does not really show anything of importance as the taxa displayed in that tree were arranged into a random order with the marmoset as the outgroup.



Figure 6. This tree was constructed to represent the most likely representation of the relationship between each of the ten taxa present in the paper.

Figure 2 and Figure 3 are interesting due to the single tree generated by Maximum Parsimony and through the Neighbor-Joining method. I was expecting similar results and I was correct. While the Neighbor-Joining tree is not rooted via the marmoset, when the two trees are compared to each other, there are a series of striking similarities to the other. First of all, humans and rhesus monkeys are said, via this singular tree method, to be very similar to each other via this gene sequence, while bonobos and chimps, which are our closest relatives based on the tree in Figure 5, are more closely related to baboons and the gibbon, respectively.

I could not leave those trees at a single replicate, however. As Figure 1 showed me, there are a variety of ways to arrange the taxa based on the nucleotide sequence. Therefore, the next step was to look at the Bootstrap trees after 500 replicates. For the bootstrap trees, I did not run the Neighbor-Joining method. Instead, I ran the Maximum Likelihood method as well as the Maximum Parsimony method. What makes Figure 3b interesting is that after 500 bootstrap replicates, 50% of those trees have the baboon and the bonobo being sister species in terms of the similarity of the sequences. 30% of the bootstrap trees also indicate a relationship between the crab-eating macaque and the western lowland gorilla. The results of the maximum likelihood bootstrap analysis might indicate the appearance of the gene appearing in species before humans, whether this gene controlled the same function is something that remains to be seen. Within the Maximum Parsimony tree in Figure 4 there are a few more items of interest. 39% of all bootstrap trees have the chimp and the gibbon showing a similar relationship.

In conclusion, the trees generated did not show me what I was expecting. I was expecting a simple relationship that would show the expected evolution of the apes to humans. I was expecting the gene to only create a standard tree. I only did analysis on the tree side of analysis and not on the algorithms that could show the parsimony score. This is because I was working with a gene that had a total length of around 51,000 base pairs. I did not have the time to run an analysis on every region of interest so most, if not all the trees generated were based on the first nucleotide in the overall sequence. If I were to run these experiments again, I would run the Sankoff and the Fitch-Hartigan algorithms to better see what was really going on within a specific region of this gene.