Journal of
Emerging Investigators

Evolution of Neuroplastin-65

Collin Cremers and Ian K. Townley

Saint George's School, Spokane, WA

Summary

Variation in the protein neuroplastin-65, encoded by the *NPTN* gene, correlates with intelligence in humans, but how did it evolve? This study investigated how this gene's evolution differed between various organisms. This topic naturally lent itself to evolutionary phylogenetics; a phylogenetic tree of the gene was created to trace the evolution of the gene. We hypothesized that the pattern of mutations in the gene would mostly correlate to the accepted evolution of animals, especially among primates where intelligence is under significant selective pressure (1, 2). The data used published results from forty-eight different organisms' gene sequences. The computer analysis produced a phylogenetic tree that reasonably matched the hypothesis. Neuroplastin mostly evolved along with the individual species, as is evidenced by significant correlations between the neuroplastin tree and other researchers' phylogenetic species trees. Specifically among primates, the data suggest that the gene was under purifying pressure, as hypothesized. However, the impact of this conclusion is limited because the primate subtree branches are poorly supported. Interestingly, among bony fish, there appears to be either positive selection or no selective pressure on this gene. Furthermore, fish neuroplastin genes may be experiencing convergent evolution with mammal neuroplastin genes.

Received: April 28, 2016; **Accepted:** August 6, 2016; **Published:** October 26, 2016

Introduction

Recently, scientists discovered that a specific gene, neuroplastin (*NPTN*), impacts intelligence in humans (3). In order to take that discovery further, this study investigated the gene's function in other organisms and its evolution. Specifically, we focused on how NPTN evolved and how its evolution differed among various organisms.

NPTN belongs to the immunoglobulin superfamily. An immunoglobulin domain is a domain that contains a specific type of β-sheet fold, known as the immunoglobulin fold (4). In all non-human organisms, the acronym for neuroplastin is NPTN. There are two isoforms in humans, neuroplastin-65 (Np-65), expressed in the brain, and neuroplastin-55 (Np-55), expressed throughout the body (5). In this study we are only investigating the Np-65 isoform, which is located on chromosome 15q24 and has 12 exons and four conserved domains (three immunoglobulin domains and one intracellurlar C-terminal domain; 6).The Np-65 gene encodes a 398 amino acid protein.

Data suggests that Np-65 impacts the genetic basis of intelligence in humans, as a study found that the number of minor alleles at rs7171755, near the locus of the Np-65 gene, inversely correlates to average cortical thickness. Specifically, in the left hemisphere, there is an average decrease of .0189 mm (0.7% of the average cortical thickness) per risk allele. The study also found that eight other single nucleotide polymorphisms in the Np-65 gene correlate to smaller cortical thickness in the left side of the brain. On average, individuals with these mutations have reductions in left cortical thickness, which impacts an individual's intelligence (7).

Mutations in the Np-65 gene impact intelligence by causing the protein to be less functional. First, Np-65 normally acts as a synaptic cell adhesion molecule (CAM), so it helps in inter-neural communication (8). As a result, a mutation which limits a CAM's effectiveness would cause significant amounts of synaptic mismatches. Data supports this hypothesis, as researchers found that mice lacking NPTN have an increase in synaptic mismatches and a reduction in the amount of post-synapses in hippocampal regions (9). Therefore, if single nucleotide polymorphisms cause Np-65 to function less efficiently as a cell adhesion molecule, then fewer axons and dendrites would align, and thus fewer neurons would send signals across synapses, reducing an individual's intelligence.

Another crucial function of Np-65 is its activation of other molecules such as kinase p38. The interaction between Np-65 and kinase p38 tunes synaptic strength, and kinase p38 helps stem cells differentiate into neurons (10, 11). Therefore, mutations in Np-65 that damage its ability to activate Kinase p38 would lead to sub-optimal stem cell differentiation. Thus, single nucleotide polymorphisms can impact an individual's intelligence by either harming Np-65's effectiveness as a cell adhesion molecule or by limiting its activation of other molecules. Finally, single nucleotide polymorphisms could change

the amount of the resulting protein, which may impact intelligence in humans.

This study investigated the evolution of NPTN by constructing a phylogenetic tree. Phylogenetic research is currently a very important tool, since looking at phylogeny can aid in understanding a particular protein's function (12). In turn, this can help researchers better understand an organism and the impact of an individual gene on the organism as a whole. Looking at clades and the amount of evolutionary time that passes before a derived characteristic is evolved could help identify specific times that NPTN underwent stronger and weaker selective pressures. Currently, intelligence is one genetic factor that could impact a primate's fitness in its environment, so studying the gene could help us understand some of the factors that impact intelligence (1, 2). Also, due to the ongoing debate about environmental and genetic impacts on intelligence, better understanding of the *NPTN* gene is important. Finally, phylogenetic analysis can serve as a starting point for comparative genetic research or help us better understand the evolution of crucial traits like intelligence (13, 14).

Phylogenetic trees predict the evolution of a gene by looking at changes in amino acid sequences. The hypothesis of this study was that the NPTN phylogenetic tree would correlate to the actual evolution of animals, because more similar organisms tend to have more similar genes. Still, because intelligence isn't strongly selected for in many organisms besides primates, it was reasoned that the gene could mutate somewhat randomly among other organisms and therefore not perfectly reflect the actual evolution of the organisms.

It was also hypothesized that among primates, the gene would be under significant purifying selection. Intelligence can help primates thrive in their environment, allowing them to learn quicker and think faster. Scientists have multiple hypotheses for why intelligence was selected for among primates, ranging from superior food gathering abilities to increased social intelligence (1, 2). Because intelligence would help a primate better fit into its environment, and because NPTN has been directly linked to intelligence, there should be selective pressure for a more effective NPTN gene. If mutations to Np-65 are maladaptive, it seems logical that the gene would be under purifying selection. Currently, the data indicates that the closest evolutionary relatives of humans are bonobos and chimpanzees (15). Therefore, if the hypothesis is correct, they would appear very close to humans in the phylogenetic tree and the branch lengths among the primates would be quite short.

In the NPTN tree, if there are few amino acid changes between two organisms, then this would imply some degree of purifying selection, because purifying selection limits changes in the amino acid sequence. Conversely, significant changes to the amino acid sequence would suggest that the gene underwent positive selection, which would cause the amino acid sequence to mutate in a specific manner. However, major changes could also imply an absence of selective pressure, which would allow the sequence to mutate randomly. It is also important to factor in the amount of time that passes since the divergence of two species. After all, more time would increase the odds that the gene would mutate, whereas recent divergence would offer less opportunity for the gene to mutate.

The results of this study suggest that the NPTN protein generally evolved along with various organisms as a whole. The gene appeared to have undergone either positive selection or no natural selection among fish; also, NPTN in fish may have experienced convergent evolution with primates. Finally, the data suggest that primates underwent purifying selection in NPTN.

Results

The phylogenetic tree of NPTN mostly correlates to the accepted evolution of organisms, as the majority of the organisms can be found near other organisms of the same group (**Figure 1**). For example, in the green subtree of **Figure 1**, containing birds, turtles, and alligators, all of the birds share a closer common ancestor with one another than with either the turtles or the alligators.

The subtree with the birds splits into one subtree that contains the order Passeriformes and another subtree with Pelicaniformes and Falconiformes (**Figure 2A**). Therefore, in this protein, the Pelicaniformes and Falconiformes have a closer common ancestor than they have with the Passeriformes. Since all of the Passeriformes are grouped together, birds of this order have copies of the protein that are similar to one another. Furthermore, among the Passeriformes, the two organisms of the genus *Corvus* appear to have a closer common ancestor than with the genus *Sturnus*. This means that, as the birds adapted, their NPTN adapted in a pattern similar to the complete organisms.

The subtree of bony fish and amphibians further demonstrates that NPTN adapted with the whole organism (**Figure 2B**). Because all of the fish share a more recent common ancestor with one another than with the amphibians, this subtree indicates that NPTN evolved as the organism evolved.. The amphibian breaks off from the bony fish before any of the fish diverged, and among the bony fish, literature suggests that the NPTN tree is reasonably close to actual fish evolution. For example, the NPTN tree shows that the Atlantic herring is the least related to the other fish, and this appears to match other sources (16).

Although the NPTN tree as a whole was relatively

Figure 1: The neuroplastin-65 tree, built using MEGA 6's Maximum Likelihood JTT+G Model. Bootstrapping values are displayed at the nodes. The branch lengths are inconsequential in this tree.

similar to the consensus evolutionary tree of the organisms, there were several major exceptions. First, the rodents were interspersed throughout the tree in several locations (**Figure 1**). Of the four rodents, only two formed a close subtree (the subtree with *Rattus norvegicus* and *Cricetulus griseus*). The other two (*Oryctolagus cuniculus* and *Ictidomys tridecemlineatus*) appeared in their own subtrees.

Another interesting exception occurs among the primates (**Figure 2C**). In some regards, the subtree appears to closely match the accepted literature for the evolution of primates, but in other instances it is vastly different (17). However, It is impossible to draw any conclusions from the shape of this subtree because the bootstrap values are incredibly low (see Methods). The values are likely low because there aren't many differences in the amino acid sequence between the various primates. Therefore, this should be investigated further in order to reach more reliable conclusions.

Also, the location of the fish subtree doesn't match traditional thoughts about the evolution of chordates. According to accepted literature, mammals are more closely related to birds than fish (**Figure 3**; 14). However, this is different in the NPTN tree, which shows mammals are more closely related to fish than birds (**Figure 1**). Additionally, the fish appear to be closely related to amphibians in the NPTN tree, but this doesn't match the consensus among published literature (**Figure 3**).

Several time periods stand out when branch lengths correlate to the number of amino acid changes (**Figure 4**). First, the NPTN gene in arthropods and bacteria has apparently changed significantly, since the branch lengths are so long. This could mean that NPTN evolved in a common ancestor of chordates from a different

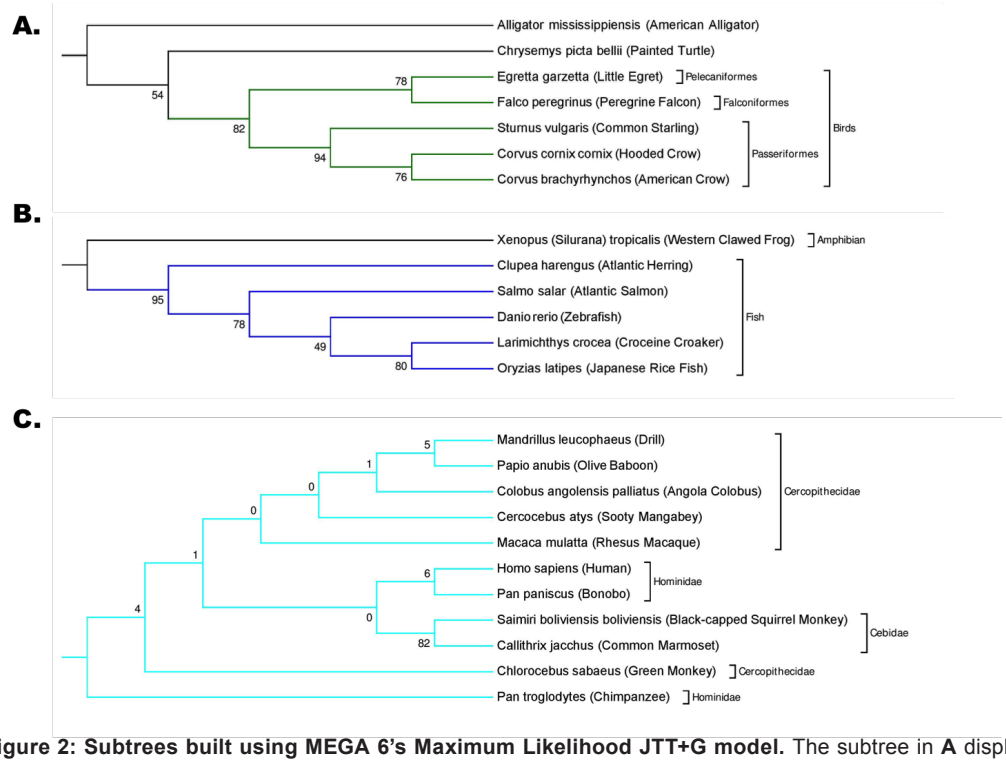

Figure 2: Subtrees built using MEGA 6's Maximum Likelihood JTT+G model. The subtree in **A** displays alligators, turtles, and birds. **B** shows amphibians (pink) and bony fish (teal). **C** displays primates, with each color correlating to a specific family among the primates. For all subtrees, bootstrapping values are displayed at the nodes and branch lengths are inconsequential.

immunoglobulin gene. If this is true, then the NPTN gene in the arthropods and nematodes is not actually a homolog of NPTN, but rather a different immunoglobulin gene. Specifically among the chordates, relatively few amino acid changes have occurred. An interesting exception to this is among the fish, where the gene apparently changed more than in the other chordates. This is evidenced by the long branch lengths in the fish subtree relative to the other chordates. Another interesting fact is that in the primate subtree, almost no change occurs compared to the evolution of the gene as a whole. This is evidenced by the fact that in the NPTN phylogenetic tree, the branches are all very short for the primates.

Discussion

The NPTN phylogenetic tree reveals several important findings. First of all, this protein mostly appears to follow the common evolution of organisms as a whole, which supports the hypothesis that the gene would evolve along with the organisms. Organisms from the same family generally share more recent common ancestors with one another than with other organisms. For example, all of the primates share the most recent common ancestor with each other than with other organisms. The fact that the families cluster suggests that there isn't much selective pressure on the gene. This is because without selective pressure on the gene, close relations between organisms mean there is less time for the amino acid sequence of the gene to mutate. Looking at anomalies in the NPTN tree could identify specific times when the gene was under stronger selective pressure.

Even though the tree's trends mostly match the commonly accepted evolution of organisms, there are several exceptions. One notable exception to this is among the rodents, as the majority of the rodents did not form a rodent subtree (**Figure 1**). It is possible that the

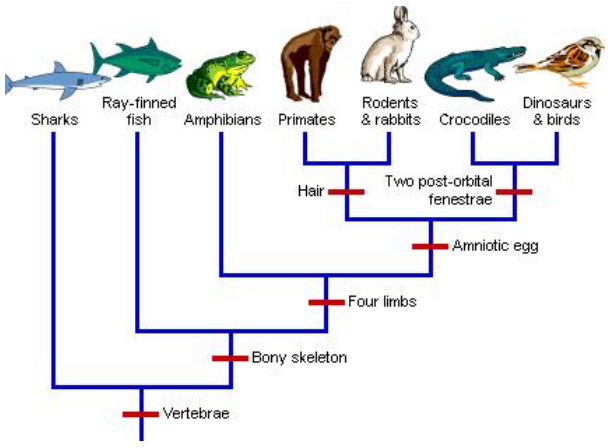

Figure 3: The commonly accepted phylogenetic tree of the evolution of Chordates (14).

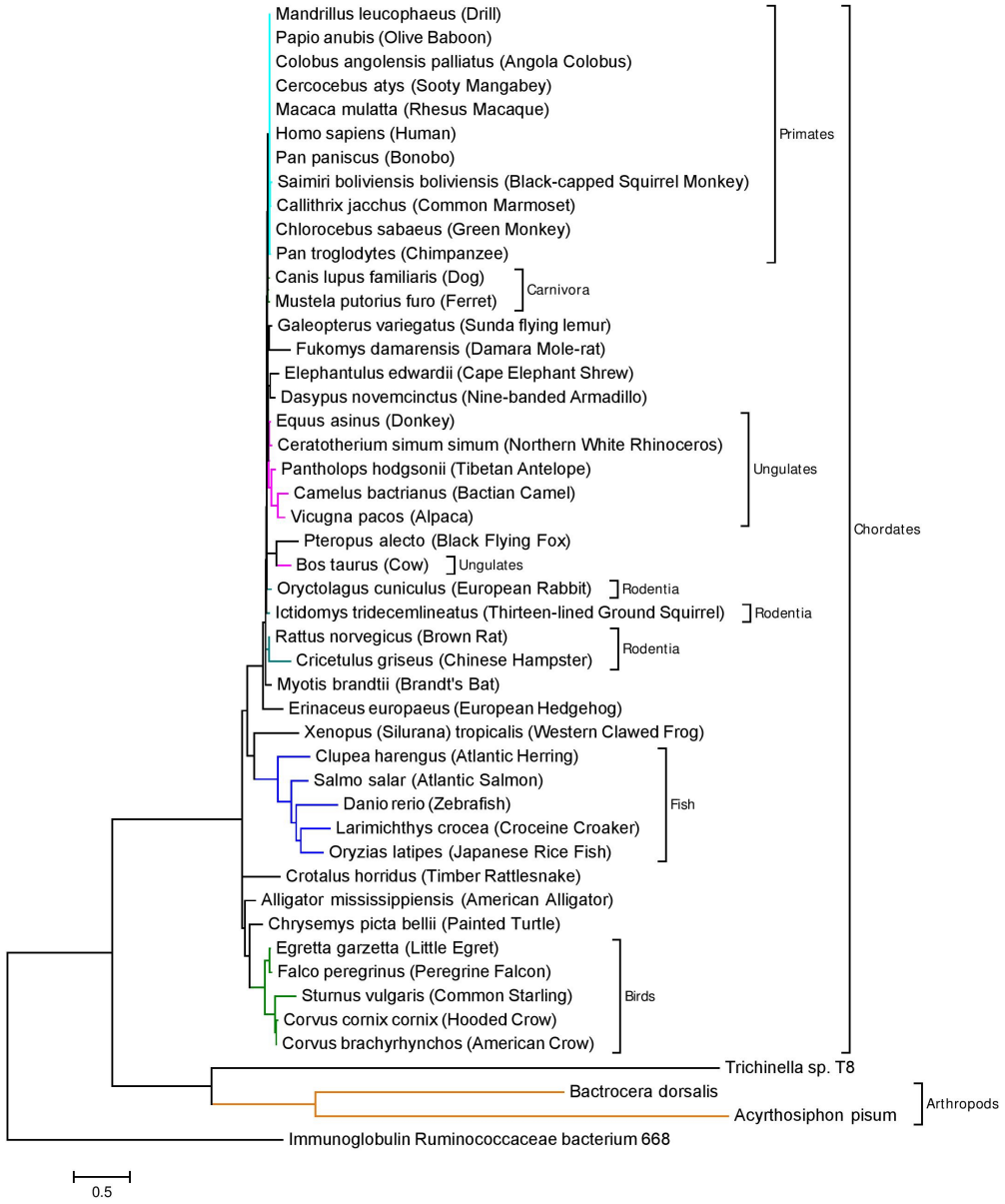

Figure 5: The NPTN tree as a whole, built using MEGA 6's JTT+G Maximum Likelihood model. The branch lengths correlate to the number of amino acid changes. The units of the scale bar are the mean number of substitutions per amino acid site.

NPTN tree doesn't match the actual evolution because the genes of some of the rodents may not be homologs of the *NPTN* gene. Instead, some may be other related immunoglobulin genes. The actual *NPTN* homologs may be within the *Rattus norvegicus* and the *Cricetulus griseus* simply because they form a close subtree, but this hypothesis needs to be investigated further.

Another interesting difference is the position of the fish subtree relative to other families of organisms. Specifically, the NPTN tree indicates that mammals are more closely related to fish than to birds and reptiles (**Figure 1**). This doesn't match the consensus evolution of life on Earth (**Figure 3**; 14). This finding could mean that fish and primates obtained similar amino acid sequences due to convergent evolution. This would cause the amino acid sequences to resemble one another and therefore appear closer together on the NPTN tree.

The amino acid sequence changed minimally among the primates since the branches are so short (**Figure 4**). This could indicate that the gene is under purifying selection, which would limit the number of amino acid sequence changes. Alternatively, it is possible that this simply occurred because the primates diverged very recently relative to life on Earth, so the gene hasn't had

enough time to mutate. The NPTN gene also appeared to undergo significant amino acid changes among the bony fish, which caused the branches to be quite long in that subtree (**Figure 4**). This could mean that there was positive pressure on the gene among the fish, which could cause the amino acid sequence to change over the course of evolutionary time. It is also possible that the gene had little to no selective pressure among the fish, which would allow the gene to mutate freely, resulting in numerous amino acid changes.

As many factors influence intelligence, a phylogenetic tree for a single gene won't perfectly demonstrate the evolution of intelligence. Still, research shows that Np-65 acts as a cell adhesion molecule in humans (9). As a result, it plays a role in building synapses in the brain. If the protein fulfills a similar function in other organisms, then the selective pressures on NPTN would be greater in organisms which benefit from more synapses. This could mean that if additional synapses are associated with a particular manifestation of intelligence (such as bodily-kinesthetic or social intelligence), then we could predict the degree of selective pressure on NPTN in specific organisms. For example, if additional synapses increase an organism's social intelligence, then we would expect that organisms with complex social structures would have stronger selective pressures on the NPTN gene.

There are several potential ways to take this experiment further or improve its reliability. It would be interesting to analyze more organisms since this tree only featured forty-eight organisms. This tree used a smaller subset of organisms in order to simplify the tree and aid in analysis, but it could also be interesting to review the gene's evolution holistically to see if there are larger trends. Adding in more species could also potentially help determine where certain families belong on the evolutionary tree. Additionally, it would be useful to look into improving the bootstrap results, which might be achieved by including more organisms or broadening the scope of the phylogenetic analysis as suggested above. Currently, the bootstrap numbers are quite low, limiting the impact of the conclusions. Generally, bootstrap values of 70% or higher indicate that the data is reliable (see Methods; 18). Finally, it would be interesting to perform a dN/dS test between several species on the phylogenetic tree. This test would help elucidate whether the organisms underwent positive or purifying selection. This could be especially useful for the primates and fish. Currently, it is not possible to reach conclusions about the importance of the anomalies in these subtrees, and the dN/dS test could help resolve the problem.

Methods

This study built the phylogenetic tree with MEGA 6, a program that aligns protein or DNA sequences, calculates phylogenetic trees, and estimates divergence times (19). The amino acid sequences were obtained by searching for the term "Neuroplastin" on the National Center for Biotechnology Information (NCBI), which records all the known amino acid sequences of a particular protein (20). Therefore, every sequence was identified as either neuroplastin or a neuroplastin precursor. This experiment used protein sequences instead of DNA sequences because DNA sequences can mutate more freely than amino acid sequences, which can cause incorrect patterns to arise from the data. The FASTA formats of several organisms from every major family were used. If the gene had multiple isoforms, this experiment used the one that was closest to 398 amino acids, because that is the amount of amino acids in Np-65. This helped increase the odds of obtaining likely precursors to the human Np-65 in the other organisms.

After obtaining the amino acid sequences, the data was aligned using the MUSCLE format (21). Next, MEGA's option to "Find Best DNA/Protein Models (ML)" calculated the best model to use for building the tree. The default settings in the section labeled "Gaps/Missing Data Treatment" were slightly readjusted by changing the setting to "Use All Sites." This section determines how gaps in data are handled; the default option ignores any gaps, but the option to use all sites would treat gaps as mutations. The default settings were changed because according to Dr. Hall in his article, "Building Phylogenetic Trees from Molecular Data with MEGA", ignoring gaps can result in loss of significant amounts of information (22). MEGA 6 recommended using the JTT+G model, so the tree was constructed with this model. The JTT model forms empirical substitution matrices that calculate the odds that a specific amino acid will change over a given period of time (23). Then, the model minimizes the total number of amino acid sequence changes that occur in the tree (24). The NPTN tree was constructed as a Maximum Likelihood Tree (as opposed to a Neighbor-Joining or Minimum Evolution), which aims to maximize the likelihood that the displayed tree is accurate. This model was used because it is considered to be the least susceptible to sampling errors (25).

The phylogenetic tree was rooted on Immunoglobulin Ruminococcaceae Bacterium 668, because this organism diverged from eukaryotes at least two billion years ago (26). After all, the other organisms are eukaryotic, whereas this one is a bacterium (27). Next, the reliability of the NPTN tree was checked by performing a bootstrap analysis (28). In this experiment, 2,000 bootstrapping runs were performed. The bootstrapping method tests the reliability by continuously rebuilding the tree with slight variations and then calculating the likelihood that the branch actually occurred where it

appears on the NPTN tree.

References

1. Jolly, Alison. "The Evolution of Primate Behavior: A Survey of the Primate Order Traces the Progressive Development of Intelligence as a Way of Life." *American Scientist* 73.3 (1985): 230-39. JSTOR. Web. July 2015.
2. Reader, Simon, et al. "The Evolution of Primate General and Cultural Intelligence." *Philosophical Transactions of the Royal Society B* 366.1567 (2011): 1017-27. National Center for Biotechnology Information. Web. July 2015.
3. Knapton, Sarah. "Is Intelligence Written in the Genes?" *The Telegraph*. Telegraph Media Group, 11 Feb. 2014. Web. May 2015.
4. Bork, P., L. Holm, and C. Sander. "The Immunoglobulin Fold Structural Classification, Sequence Patterns and Common Core." *Journal of Molecular Biology* 242.4 (1994): 309-20. Web. Dec. 2015.
5. Smalla, K., et al. "The Synaptic Glycoprotein Neuroplastin Is Involved in Long-term Potentiation at Hippocampal CA1 Synapses." *Proceedings of the National Academy of Sciences* 97.8 (2000): 4327-332. Web. Aug. 2015.
6. Owczarek, S., et al. "Neuroplastin-55 Binds to and Signals through the Fibroblast Growth Factor Receptor." *The FASEB Journal* 24.4 (2009): 1139-150. Web. June 2015.
7. Desrivières, S., et al. "Single Nucleotide Polymorphism in the Neuroplastin Locus Associates with Cortical Thickness and Intellectual Ability in Adolescents." *Molecular Psychiatry* 20.2 (2014): 263-74. Web. June 2015.
8. Thalhammer, Agnes, and Lorenzo A. Cingolani. "Cell Adhesion and Homeostatic Synaptic Plasticity." *Neuropharmacology* 78 (2014): 23-30. Web. June 2015.
9. Beesley, Philip, et al. "The Neuroplastin Adhesion Molecules: Key Regulators of Neuronal Plasticity and Synaptic Function." *Journal of Neurochemistry* 131.3 (2014): 268-83. Web. June 2015.
10. Zarubin, Tyler, and Jiahuai Han. "Activation and Signaling of the P38 MAP Kinase Pathway." *Cell Research* 15.1 (2005): 11-18. Web. June 2015.
11. Empson, Ruth, et al. "The Cell Adhesion Molecule Neuroplastin-65 Inhibits Hippocampal Long-term Potentiation via a Mitogen-activated Protein Kinase P38-dependent Reduction in Surface Expression of GluR1-containing Glutamate Receptors." *Journal of Neurochemistry* 99.3 (2006): 850-60. Web. July 2015.
12. Baxevanis, Andy, and Francis Ouellette. *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. 2nd ed. Hoboken: John Wiley & Sons, 2004. Print.
13. Soltis, Douglas, and Pamela Soltis. "The Role of Phylogenetics in Comparative Genetics." *Plant Physiology* 132.4 (2003): 1790-800. Web. Feb. 2016.
14. "Understanding Evolution." *University of California Museum of Paleontology*, 2016. Web. Feb. 2016.
15. Prüfer, Kay, et al. "The Bonobo Genome Compared with the Chimpanzee and Human Genomes." *Nature* 486.7404 (2012): 527-31. Web. 30 July 2016.
16. Betancur-R., Ricardo, et al. "The Tree of Life and a New Classification of Bony Fishes." *PLoS Currents* (2013). Web. Nov. 2015.
17. Perelman, Polina, et al. "A Molecular Phylogeny of Living Primates." *PLOS Genetics* 7.3 (2011). Web. Nov. 2015.
18. Baldauf, Sandra. "Phylogeny for the Faint of Heart: A Tutorial." *Trends in Genetics.* Department of Biology, University of New York, 2003. Web. 30 Sept. 2016.
19. Koichiro Tamura, et al. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*: 30 2725-2729.
20. The NCBI handbook. *National Library of Medicine (US), National Center for Biotechnology Information*. 2002 Oct. Chapter 18.
21. Edgar, Robert C. (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research* 32(5), 1792-97.
22. Hall, Barry. "Building Phylogenetic Trees from Molecular Data with MEGA." *Molecular Biology and Evolution* 30.5 (2013): 1229-235. Web. Nov. 2015.
23. Abascal, F., et al. "ProtTest: Selection of Best-fit Models of Protein Evolution." *Bioinformatics* 21.9 (2005): 2104-2105. Web. 30 July 2016
24. Vinuesa, Pablo. "Model Fitting in Phylogenetics." *Pablo Vinuesa's Research and Teaching Site*. Center for Genomic Sciences, 1 Feb. 2007. Web. 30 July 2016
25. Kipling, Will. "Principles of Phylogenetics." University of California, Berkely, Spring 2012. Web. Feb. 2016.
26. "When Did Eukaryotic Cells (cells with Nuclei and Other Internal Organelles) First Evolve? What Do We Know about How They Evolved from Earlier Life-forms?" *Scientific American*, 2002. Web. 30 Sept. 2016.
27. Russell, Peter J., et al. *Biology: The Dynamic Science*. 1st ed. Belmont: Thomson Corporation, 2008. Print.
28. Kesar Singh and Minge Xie. Bootstrap: A statistical method, 2008.