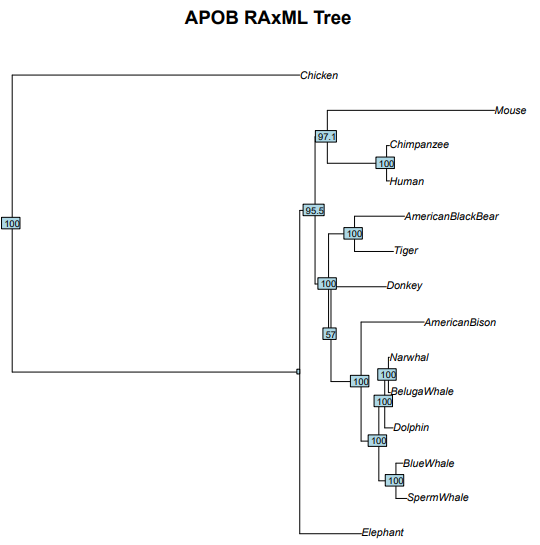
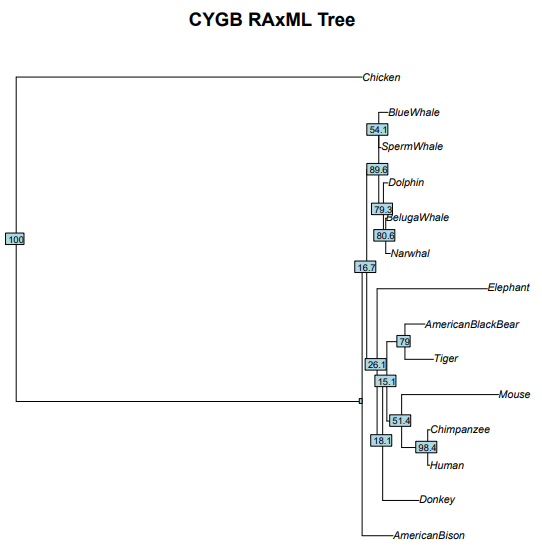
**Investigating the evolutionary history of oxygen-interacting proteins in deep-diving aquatic mammals and terrestrial mammals through phylogenetic computations**

**Introduction**

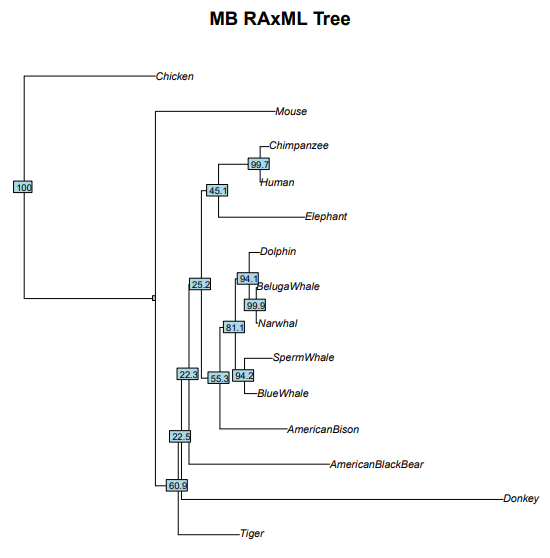
Many aquatic mammals have the ability to utilize deep-diving capabilities with their developed molecular and evolutionary physiology (Berenbrink, 2020). Aside from some physical and habitation differences, land and aquatic mammals vastly differ at the molecular level through their evolutionary history. Evidence from previous studies suggest further evolution in marine mammals to prevent hypoxia, or blood oxygen deprivation, in the brain and muscles. Their molecular characteristics allowed for a crucial utilization of proteins that bind with or transfer oxygen throughout the body like cytoglobin and myoglobin. Terrestrial or land mammals also share these globin proteins and their functions, but express lower amounts of these proteins in their muscle tissues (ScienceDaily, 2015; Williams et al., 2008).  
 Two of the sequenced proteins of this study, cytoglobin and myoglobin are two integral proteins for vertebrate animals. They both function in cellular oxygen homeostasis and binding to oxygen reversibly, but they occupy different spaces in the body. Cytoglobin’s role is paired with another protein, neuroglobin, in exchanging oxygen to almost every tissue type, but mostly to mammalian brain tissues. Comparatively, myoglobin is predominately located in cardiac or striated muscle, contributing a red color to the tissues. Myoglobin can store oxygen in these tissues and facilitate O2 diffusion whilst also removing toxic nitric oxide (Pesce et al., 2002). A plethora of studies have related both of these proteins to hypoxia-responsive adaptations in vertebrates, contributing to aquatic mammals’ ability to not black out when deep-diving (Bryner, 2007; ScienceDaily, 2015; Williams et al., 2007).  
 Though, from a multitude of experiments, scientists mainly analyzed the biochemistry properties of these proteins in mammals. There was only a single experiment that constructed evolutionary trees solely for myoglobin (Isogai et al., 2021). A computational investigation of the evolutionary history via phylogenetic tree with ancestral state reconstructions has also not yet been conducted to explain this change. Therefore, the purpose of this study was to search for evolutionary evidence and differences of oxygen-related proteins in aquatic versus land mammals compared with a control gene. The control gene, apolipoprotein B, that was proposed is exclusive to mammals and poses no difference in aquatic or terrestrial mammal. This gene is not related to oxygen binding and is produced in the intestine to carry fats in the bloodstream (U.S. National Library of Medicine, 2021).   
**Methods** To test for this, the genes’ sequences were to be analyzed through phylogenetic tree construction. Genes with 1-to-1 orthologs for the chosen sample of mammals were chosen and downloaded as .fa files from the genome browser Ensembl. The chosen land mammals included the American bison, American black bear, chimpanzee, donkey, elephant, human, mouse, and tiger; the marine mammals were the dolphin, narwhal, sperm whale, beluga whale, and blue whale; and the chosen outgroup was the chicken from the birds and reptiles category. The unaligned coding sequences (CDS) of the human orthologues for myoglobin (MB), cytoglobin (CYGB), and apolipoprotein B (APOB) were read and aligned in RStudio using Clustal Omega Once the Phylip output files were formed, they were uploaded to UNC Charlotte’s Centaurus cluster. There, RAxML was run with 1000 bootstrap replicates under the GTRGAMMA model to create the maximum likelihood tree in .bestTree file format. Then, tree distances were calculated from each combinational pair of trees. This was to quantify the comparison of the control gene to the experimental genes. Finally, an overall ancestral state reconstruction was performed with maximum parsimony on the three trees to see if there were any model character evolutionary changes between terrestrial and aquatic mammals.

**Results**

**Figure 1.** Maximum likelihood tree of the control APOB gene constructed through RAxML with 1000 bootstrap replicates.



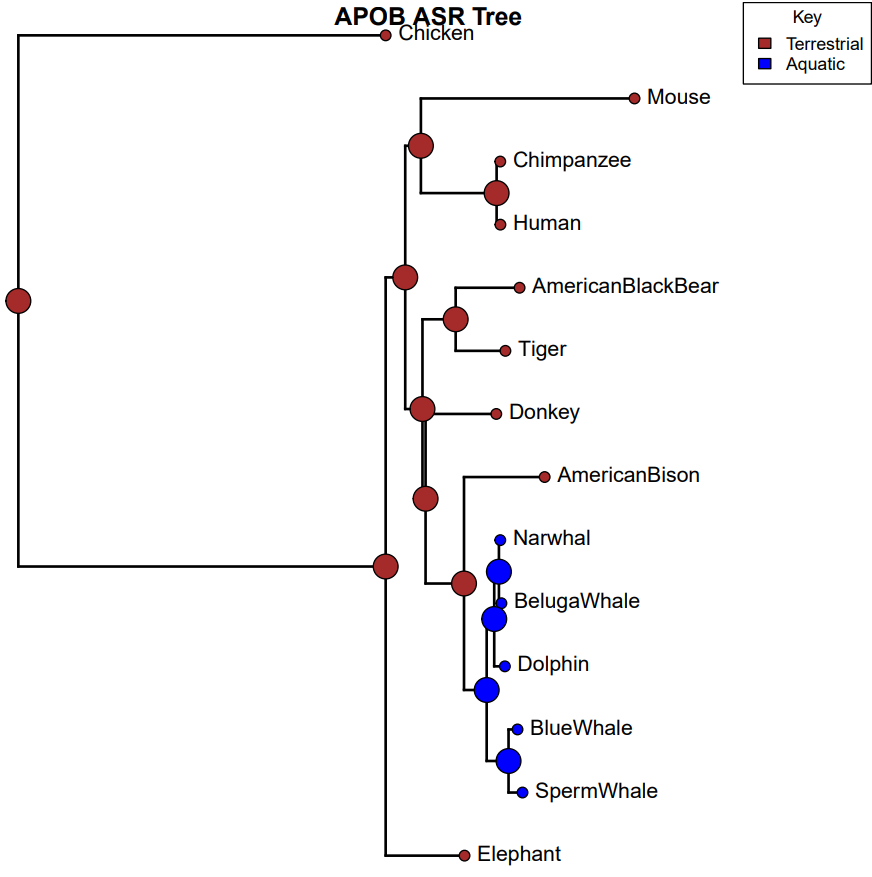
**Figure 2.** Maximum likelihood tree of the experimental CYGB gene constructed through RAxML with 1000 bootstrap replicates.

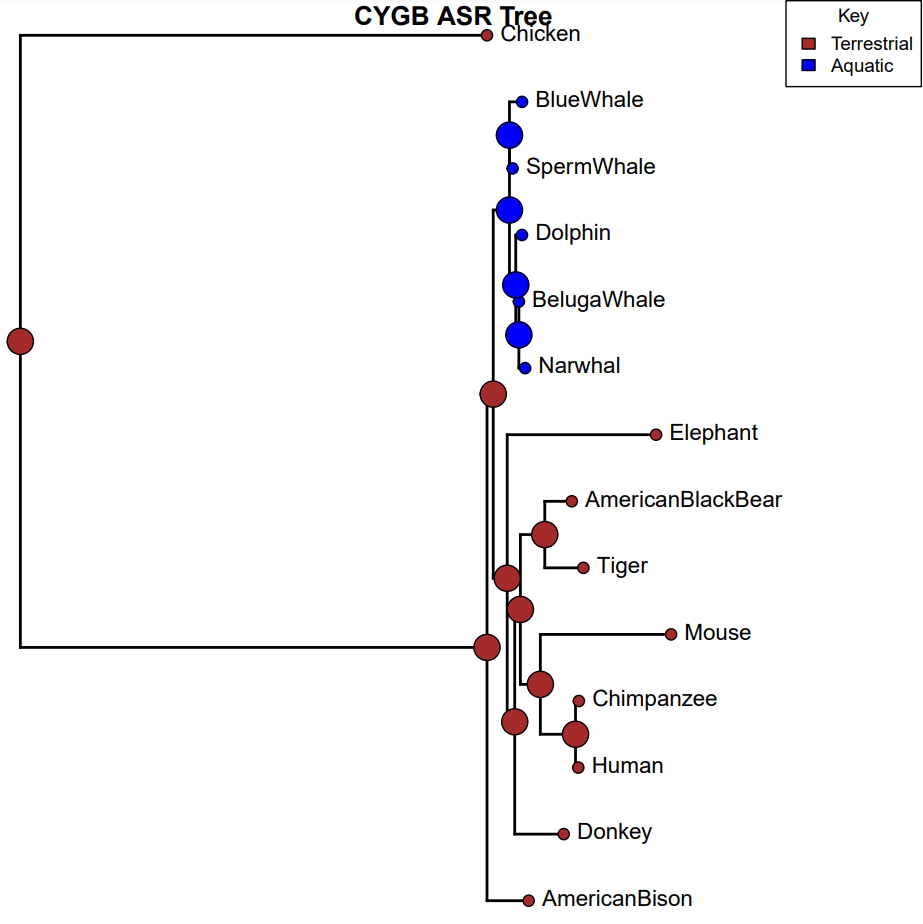


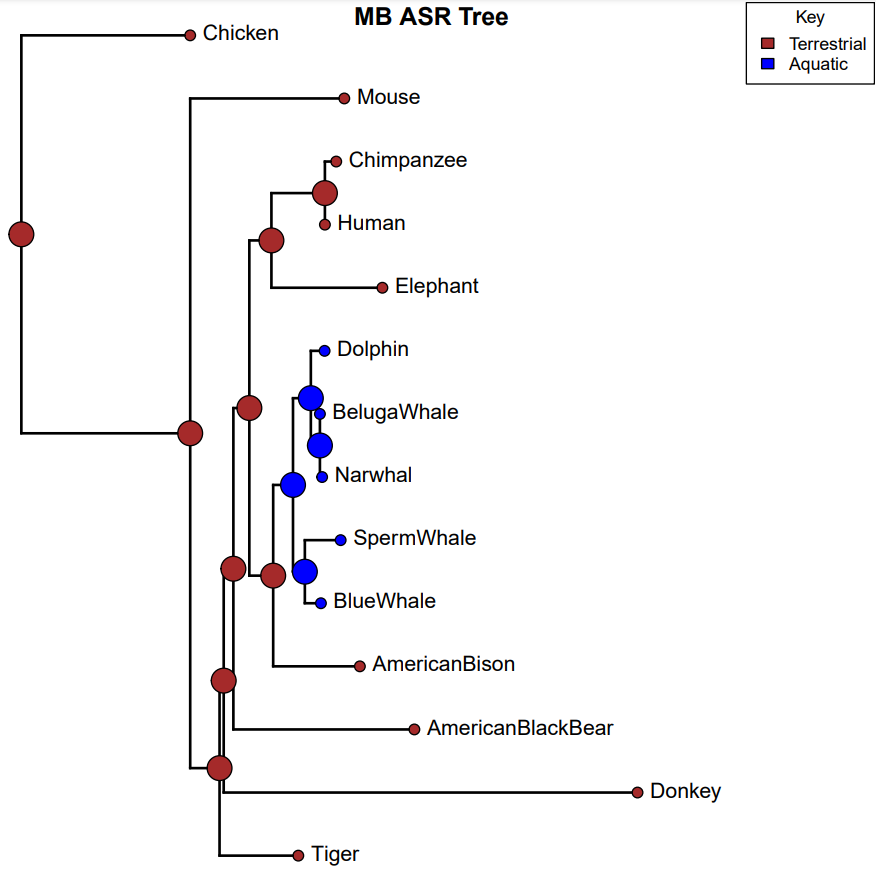
**Figure 3.** Maximum likelihood tree of the experimental MB gene constructed through RAxML with 1000 bootstrap replicates.

**Table 1.** Quantitative comparison of tree distance of two trees from the R package phangorn.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Trees Compared | Symmetric Difference | Branch Score Difference | Path Difference | Quadratic Path Difference |
| APOB & CYGB | 8.0000000 | 0.2602764 | 14.1421356 | 1.4866391 |
| APOB & MB | 10.0000000 | 0.5584069 | 16.0000000 | 1.8905536 |
| MB & CYGB | 12.0000000 | 0.5007127 | 19.0787840 | 2.0334587 |

 **Figure 4.** Ancestral state reconstruction with maximum parsimony of the APOB control gene.

   
**Figure 5.** Ancestral state reconstruction with maximum parsimony of the CYGB experimental gene.



**Figure 6.** Ancestral state reconstruction with maximum parsimony of the MB experimental gene.

**Discussion**

For **Figures 1 to 3**, the aquatic mammals were all grouped together and diverge collectively from the other terrestrial mammals with relatively low bootstrap support values. The marine mammals’ clade diverged latest in the myoglobin and apolipoprotein B trees, but not for cytoglobin. This also shared different nodes across all trees. **Figure 2** for the CYGB gene has the lowest support value of 16.7 for the aquatic mammals’ clade across all three trees, possibly due to the replicates having poor agreement. From **Table 1**, the juxtaposed quantitative comparisons between two respective trees were performed to give the symmetric, branch score, path, and quadratic differences. The biggest number that implies the biggest difference was the path difference in all 3 comparisons, which makes sense as all of the branches are in different locations throughout the trees. The lowest scores were the branch scores, which also makes sense as the branch lengths of all the sequences of the same species are relatively similar among all the trees. The other differences were not as significant in part of being the greatest or lowest value, being intermediate values. In **Figures 4 to 6**, the ancestral state reconstruction posed that there were no outstanding character state changes aside from when the aquatic or blue branches separate. It only changed the state from terrestrial to aquatic mammals once towards the end of the trees.

The APOB gene does not provide many differences other than node locations for the marine clade specifically apart from the oxygen-interacting genes of MB and CYGB. There is not enough evidence to conclude that proteins that interact with oxygen like MB and CYGB are vastly different from a control gene that does not interact with oxygen like APOB. MB and CYGB may not have diverged specifically for aquatic mammals from land mammals according to **Figures 1, 2, and 3,** as a similar divergence occurred in the APOB tree as well. This may suggest evolutionary divergence for the aquatic mammals from land mammals as a whole rather than selecting for more utilized genes. It makes sense that previous studies have not concluded a sequential evolutionary distinction, but instead found the evidence through the biochemistry portions of these proteins (Bryner, 2007; ScienceDaily, 2015; Williams et al., 2007). The way that marine mammals can utilize most of their mammalian genes could have stemmed from the adaptability and variation in different habitats.

In future directions or studies, other genes could be tested for, even including 1-to-many or many-to-many orthologues of mammalian genes, or other genes. Since the chosen sample of mammals could not be entirely representative of an aquatic mammal deviation, more mammals could be selected and/or possibly include other classifications from the animal kingdom. For the trees constructed in this study, the rates of evolutionary state changes were assumed to be equal rates in both directions, which could not be the realistic historical case. More research is needed to see if evolving from an aquatic habitat to a terrestrial habitat has the same evolutionary cost or vice versa. Lastly, the model of rate changes could also be chosen more specifically to fit historical transition events like the Devonian tetrapod land invasion (Berenbrink, 2020).

**References**

Berenbrink, M. (2020, November 10). *The role of myoglobin in the evolution of mammalian diving capacity – the August Krogh principle applied in molecular and Evolutionary Physiology*. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology. Retrieved December 10, 2021, from https://www.sciencedirect.com/science/article/pii/S1095643320301963.

Bryner, J. (2007, December 18). *Why deep-diving mammals don't black out*. LiveScience. Retrieved December 10, 2021, from <https://www.livescience.com/10549-deep-diving-mammals-black.html>.

Isogai, Y., Imamura, H., Nakae, S., Sumi, T., Takahashi, K.-ichi, & Shirai, T. (2021, July 30). *Common and unique strategies of myoglobin evolution for deep-sea adaptation of diving mammals*. iScience. Retrieved December 10, 2021, from https://www.sciencedirect.com/science/article/pii/S2589004221008889.

Pesce, A., Bolognesi, M., Bocedi, A., Ascenzi, P., Dewilde, S., Moens, L., Hankeln, T., & Burmester, T. (2002, December). *Neuroglobin and cytoglobin. fresh blood for the vertebrate globin family*. EMBO reports. Retrieved December 10, 2021, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1308314/.

ScienceDaily. (2015, September 25). *Deep-diving whales could hold answer for synthetic blood*. ScienceDaily. Retrieved December 10, 2021, from https://www.sciencedaily.com/releases/2015/09/150925131427.htm#:~:text=%22Whales%20and%20other%20deep%2Ddiving,and%20Dorothy%20Looney%20Professor%20of.

U.S. National Library of Medicine. (2021, May 18). *APOB gene: Medlineplus Genetics*. MedlinePlus. Retrieved December 10, 2021, from https://medlineplus.gov/genetics/gene/apob/.

Williams, T. M., Zavanelli, M., Miller, M. A., Goldbeck, R. A., Morledge, M., Casper, D., Pabst, D. A., McLellan, W., Cantin, L. P., & Kliger, D. S. (2008, April 7). *Running, swimming and diving modifies neuroprotecting globins in the mammalian brain*. Proceedings. Biological sciences. Retrieved December 10, 2021, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2596902/.