

Analysis of Open Chromatin Regions Linked to Hypselodontology

Shashank Katiyar and Jonathan Zhu

Abstract

Teeth are a common feature across a wide variety of mammals, giving many of them an evolutionary advantage to consume varying forms of prey. Of particular interest is hypselodontology; this refers to an organism's capability to infinitely grow their teeth in response to the teeth wearing down over time. While most common in rodents, this trait is noteworthy in that it also occurs in a variety of other unrelated mammalian species that vary in diet and brain composition. Additionally, substantial evidence suggests that the development and maintenance of teeth is highly linked to an animal's nervous system; in the case of hypselodontology, the stem cells that allow for infinite tooth growth are linked to a variety of neural processes.

Our analysis of open chromatin regions (OCRs) between hypselodonts and non-hypselodonts focused on oligodendrocyte precursor cells (OPCs) due to their similarity to neural and dental stem cells, as well as on oligodendrocytes due to their similarity to Schwann cells. We tested three different methods for determining significantly different OCRs: t-tests, phylolm, and phyloglm; we found that phyloglm has the best performance. Our results on OPCs yielded OCRs with relations to neural and general stem cell proliferation, and our results on oligodendrocytes found OCrs with relations to nervous system development. Such results have potential connections to hypselodontology but may need future experimental validation for confirmation.

Our results provide starting points for much future research into hypselodontology as a whole; due to its relation with the nervous system and development, this also allows future work to learn more about the development of the dental system in tandem with the nervous system. Furthermore, our results can form as a basis for future tooth regeneration processes in humans.

Background

Hypselodontology is defined as the continuous growth of teeth in animals, particularly mammals; it is not to be confused with hypsodonty, which refers to a specific morphology of the teeth in the animal, not its infinite growth. In loose terms, hypselodontology relies on the presence of dental stem cells, particularly mesenchymal (internal) and epithelial (external) stem cells to regenerate the enamel that forms a tooth as well as the nerves and blood vessels within the tooth (Tummers and Thesleff, 2003). Hypselodontology specifically counteracts the general wear on teeth over time; as time passes and an animal eats more, the food it eats will wear down its teeth, and the ability of these teeth to grow infinitely can provide an evolutionary advantage.

Hypselodontology has been particularly notable in rodent incisors (Klein et al., 2008), where the infinite ability to grow originates from stem cells in the mesenchyme and epithelial cells of their incisors and is counterbalanced with abrasion, particularly from enamel not being present on some surfaces of the tooth. Various genetic studies have identified certain genes as contributing to this hypselodontology in mice; Klein et al. (2008) particularly notes the *Spry* family of genes, and Wang et al. (2007) describes a large genetic regulatory network behind the phenomenon, noting the genes *Follistatin*, various FGFs, and *Bmp*.

However, this phenomenon is not just present in rodents. Renvoisé and Michon (2014) present a phylogenetic tree labelling various mammalian orders that exhibit hypselodontology,

including Lagomorpha (includes many rabbits), Sirenia (includes manatees and dugongs), Proboscidea (includes elephants) and Hyracoidea (hyraxes) as other mammalian orders that share incisor hypselodonty with rodents. However, other mammals have shown hypselodonty in all other tooth types (canines, premolars and molars). In particular, Diprotodontia, an order of marsupials including koalas, also have members that exhibit hypselodonty, thus showing convergent evolution of this trait.

While this trait may seem to be far from a neural trait in the sense that it deals primarily with teeth, it is important to note that teeth and the nervous system have a fairly nuanced relationship. A review by Duan et al. (2022) states that teeth and the nervous system share developmental origins; in particular, dental mesenchyme is derived from cranial neural crest (CNC) cells. Furthermore, this same review examines literature discussing the relationship between teeth and various neural processes, including the relationship between dental disease and neurophysiological changes. However, key to the examination of hypselodonty is the relationship of the nervous system to dental stem cells. The same review by Duan et al. (2022) summarizes many key pieces of literature linking neural processes to dental stem cell regulation; some of these linkages include inferior alveolar nerve (IAN) secretion influencing mesenchymal stem cell proliferation via Wnt signaling (Zhao et al., 2014).

This trait has already been annotated across the species in Zoonomia, and our next section elaborates on the ones we specifically chose.

Species

We plan to use species from different Mammalian orders which have the hypselodonty trait. Out of the Zoonomia species, 63 species have this trait. We want to include species from many different mammalian orders, hence we tried to find among them, species from as many different mammalian orders as possible. These were the results:

- 4 Artiodactyla (includes cows, giraffes): *Catagonus wagneri*, *Elaphurus davidianus*, *Sus scrofa*, *Tragulus javanicus*
- 2 Cetacea (includes whales, dolphins): *Mesoplodon bidens*, *Monodon monoceros*
- 3 Lagomorpha (includes many rabbits): *Lepus americanus*, *Oryctolagus cuniculus*, *Ochotona princeps*
- 1 Primates (includes humans, chimps): *Daubentonia madagascariensis*
- 10 Rodentia (includes mouse, rat): *Mus musculus*, *Rattus norvegicus*, *Cavia porcellus*, *Chinchilla lanigera*, *Hydrochoerus hydrochaeris*, *Acomys cahirinus*, *Peromyscus maniculatus*, *Dipodomys ordii*, *Cricetulus griseus*, *Myocastor coypus*. There were more rodentia species, but we plan to focus on these 10 first, as they are the most popular (estimated according to the number of papers with the species name on PubMed) and hence will likely have the best quality data from them.

Among the chosen species, as explained above, for the Rodentia, hypselodonty in their incisors is an identifying feature. It helps them to gnaw and chew on tough materials which is helpful for survival in different environments. Lagomorpha species show hypselodonty in both, their molars and incisors, making them an interesting parallel to compare to Rodentia. The primate *Daubentonia madagascariensis*, which is a lemur species from Madagascar, is especially interesting as it is an exception among primates to exhibit hypselodonty (Fleagle, 1999). Since

this species is highly different from others but still has the trait, comparing it with the others will help us find relevant genomic regions.

Cell Types

For this project, we chose to analyze two cell types: Oligodendrocyte precursor cells (OPCs) and oligodendrocytes. OPCs, as the name suggests, give rise to oligodendrocytes during an organism's developmental process. Found throughout the brain, these cells are typically distinguished by their genetic expression; in particular, these cells express the NG2 and PDGF receptors (Nishiyama et al., 1996). These cells have been widely studied and linked to a variety of neurological processes.

Some linkage also may exist between these cells and hypselodonty, or at least to dental cells. For example, Degistirici et al. (2008) found that the PDGF receptors (among other markers) were critical in determining the migration of a cell from the neural tube to develop into a dental precursor cell. Furthermore, a study by Askari et al. (2015) demonstrated that dental stem cells differentiated into oligodendrocyte precursor cells upon the overexpression of the Olig2 transcription factor. While both of these studies performed some experiments on mice, a known hypselodont, mice were not specifically analyzed in either of these studies with respect to their teeth; in fact, the dental stem cells were extracted from human patients. However, we hypothesize that a differential pattern of open chromatin may be observed in the oligodendrocyte precursor cells in hypselodonts than non-hypselodonts due to their need for continuously growing teeth and a constant supply of dental stem cells.

The second cell type we chose to analyze are oligodendrocytes. Oligodendrocytes are glial cells in the CNS (central nervous system). They are similar to astrocytes, but have fewer protuberances. They are responsible for providing support to neurons in the CNS and creating myelin sheath (insulation) around axons of neurons in the CNS. One Oligodendrocyte can provide insulation to up to 50 axons. Schwann cells (below) on the other hand, can cover only one axon per cell.

It is known that dental Mesenchymal Stem Cells produce odontoblasts and pulp cells during the process of development, growth and most importantly, regeneration of mouse incisors, and the fact that Schwann cells and Schwann cell progenitors directly give rise to dental Mesenchymal Stem Cells (Duan et al., 2022). We are selecting Oligodendrocytes because from the given list of striatal cell types, Oligodendrocytes are most closely related to Schwann Cells. Schwann cells are just the peripheral nervous system counterparts of Oligodendrocytes. Our expectation is that because of their similarity to Schwann cells, and the known role of Schwann cells in regeneration of mouse incisors, there may be some linkage between hypselodonty trait and Oligodendrocytes.

Visualization

The full Zoonomia tree whose species are annotated for the presence or absence of hypselodonty is given below:

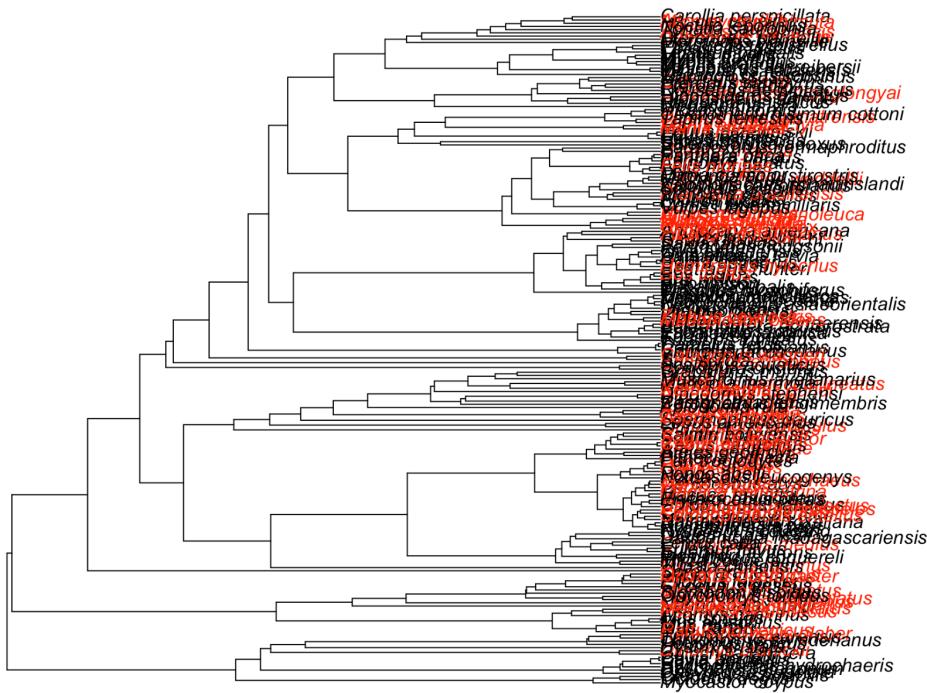


Figure 1: Full Zoonomia tree, with the exception of species not annotated for hypselodonty. Red: Species exhibiting hypselodonty; Black: Species not exhibiting hypselodonty

Here, organism names in red are ones that exhibit hypselodonty, while organisms that do not have their names in black, while the species without the hypselodonty annotation are removed. The tree was constructed based on molecular distance by Cactus.

We have also created a condensed tree of only species labelled as hypselodonts in the Zoonomia dataset, given below. Coloring of the names has been done by clade; blue names represent Laurasiatherians, black names represent Euarchontans, and red names represent Gilres.



Figure 2: Condensed tree of only species labelled as hypselodonts in the Zoonomia dataset. Coloring of the names has been done by clade; Blue: Laurasiatherians; Black: Euarchontans; Red: Gilres.

Analysis

For our analysis, we look at predicted open chromatin regions for oligodendrocytes and oligodendrocyte precursor cells (OPCs) between two groups of hypselodont mammals and non-hypselodont mammals. For both cell types, we took a subset of 10,000 peaks from a 2 million+ peak dataset across 240 Zoonomia mammals. We then used the phylolm package and functions phylolm and phyloglm to analyze correlations between hypselodonty and all predicted open chromatin peak regions. Furthermore, we also used t-tests to look further into differences in predicted peak values between the two groups.

Once correlations were obtained, we looked at some of the top negatively and positively correlated regions of predicted open chromatin. For the top positively and negatively correlated region, we used the UCSC Genome Browser to find the closest gene and report on these close genes. Afterwards, we subsetted the predicted peaks into those confined to a single genome so as to run a GREAT analysis and find enriched regions of the genome; the enriched portions were those that formed the top 200 positive or negative hits. Finally, we looked at existing databases (Allen Brain Map and CATLAS) to examine single-cell results for the genes of

interest. For both GREAT and single-cell analyses, these existing resources allow for the comparison of human (non-hypsodont) and mouse (hypsdont) genomes and results. This unique comparison allows us to examine not only the prevalence of genes throughout various cells in the brain and general enriched regions of the genome, but also to compare the results for human and mouse to potentially find more links to hypsdonty.

1. Comparison of Methods used

In previous work analyzing open chromatin differences, we have used various statistical techniques to analyze the data, make discoveries, and identify large, most significant differences. These tests include the simple t-test, which looks at raw numerical value without any correction, as well as specialized methods provided in the phylolm package, which perform comparisons while correcting for evolutionary relationships given by a tree. To decide the method which guides our primary results, we compared all three methods on OPC peak data.

The simple t-test is one of the most commonly used statistical tests. While it forms a very robust test, it does rely on strong assumptions; specifically, it requires that the tested groups follow a normal distribution. At a high level, the t-test for two sample differences examines pooled variance to determine the probability of the two sample groups coming from the same distribution; given a low probability, we can reject the null hypothesis of these two sample groups coming from the same distribution and therefore can reject the null hypothesis of the difference of means between the two groups being 0.

In our analysis, we found that t-tests produce a fairly uniform distribution of p-values across the tested values (Figure 3). We estimated the effect size of the test by looking at the difference of means between the sample groups, and these formed a histogram centered around 0 (Figure 3). While the effect size is to be expected as most regions would not have large mean differences, the uniform p-value distribution indicates that any significant p-values could be false positives.

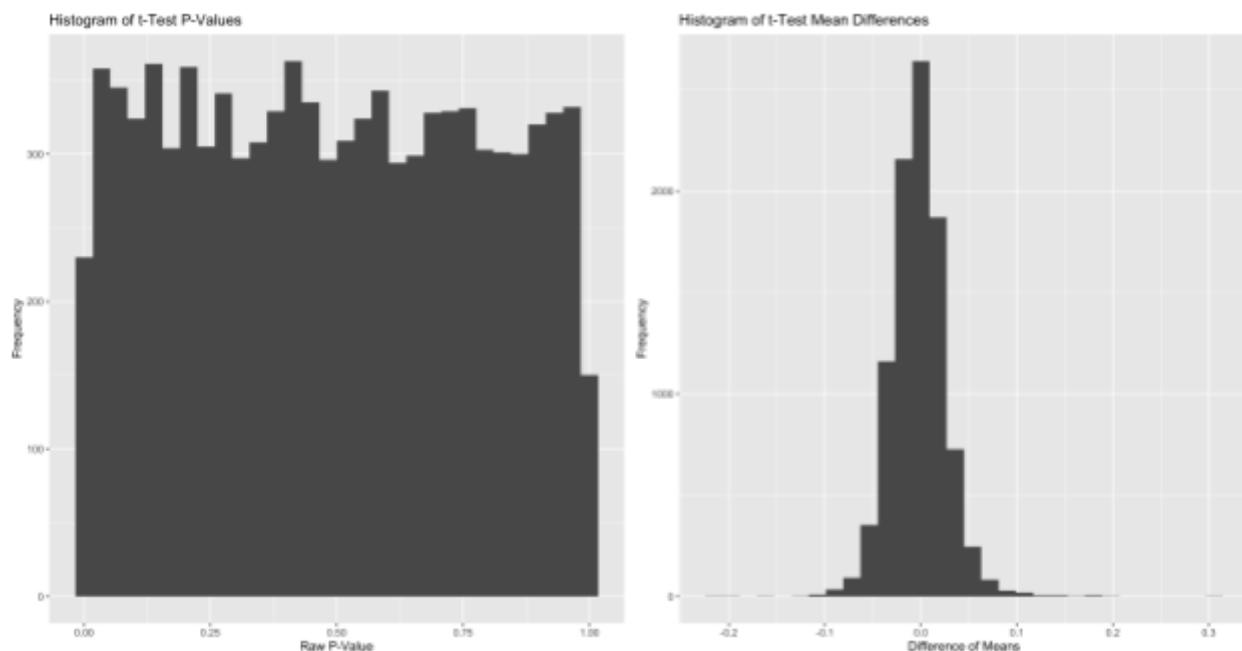


Figure 3: Results of t-tests applied to open chromatin differences between hypselodonts and non-hypselodonts for 10,000 peaks. (Left) p-value distribution; (right) mean difference distribution.

Furthermore, when looking at the top positively and negatively associated regions found with t-tests (i.e., the most significant test results with a positive and negative mean difference, which correspond to the regions more highly open in hypselodonts vs. non-hypselodonts) we find that the t-test prioritizes some of the largest significance due to a low variance caused with smaller group sizes (Figure 4). This, along with the distribution of p-values, indicates that the t-test is least likely to give insightful results.

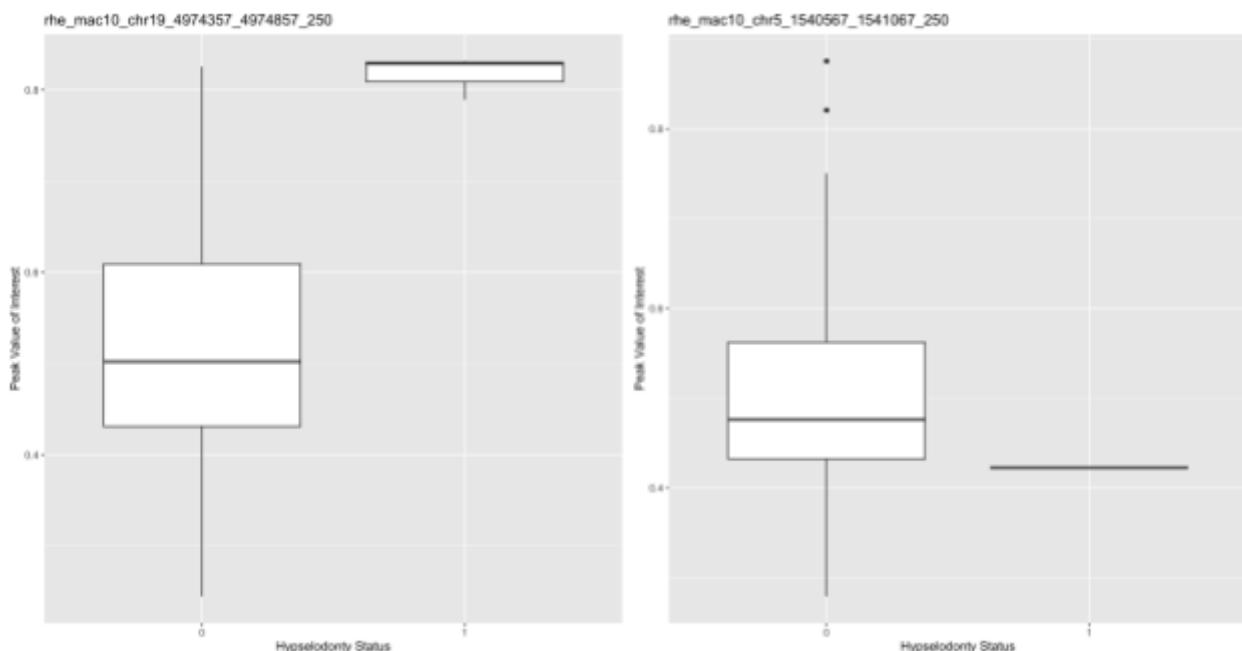


Figure 4: Barplot distributions of values from the most significant t-test results. (Left) Most positively associated region, with larger values in hypselodonts. (Right) Most negatively associated region.

A common alternative to the t-test is a method that accounts for evolutionary relationships that may otherwise confound analyses. The principle behind this is that closely related organisms are more likely to have similar values for the tested peak, so phyloglm's methods seek to correct for these to reduce false positives that may otherwise yield a significant result with the t-test. After correction, the test is a simple one-variable linear regression, with significance tested on the slope by standard methodologies, whose exact procedure is beyond the scope of this report. On a high level, the test assumes the null hypothesis that the slope of the regression is 0; a low p-value indicates a high likelihood that there is a relationship between the independent variable and dependent variable. In our case, hypsodonty status is converted to a binary variable (0 for non-hypselodonts, 1 for hypselodonts) for the regression and is treated as the dependent variable; while it can be treated as the independent variable, this is for consistency with phyloglm, explained later in this report.

PhyloLM's standard linear regression method, often just called phyloLM, produced a right-skewed histogram of p-values with 30 being significant after adjustment, and an expected distribution of effect sizes (estimated here as slope) centered around 0 (Figure 5). However, upon inspection of the most positively and negatively associated regions, phyloLM failed to capture large differences between the tested groups (Figure 5).

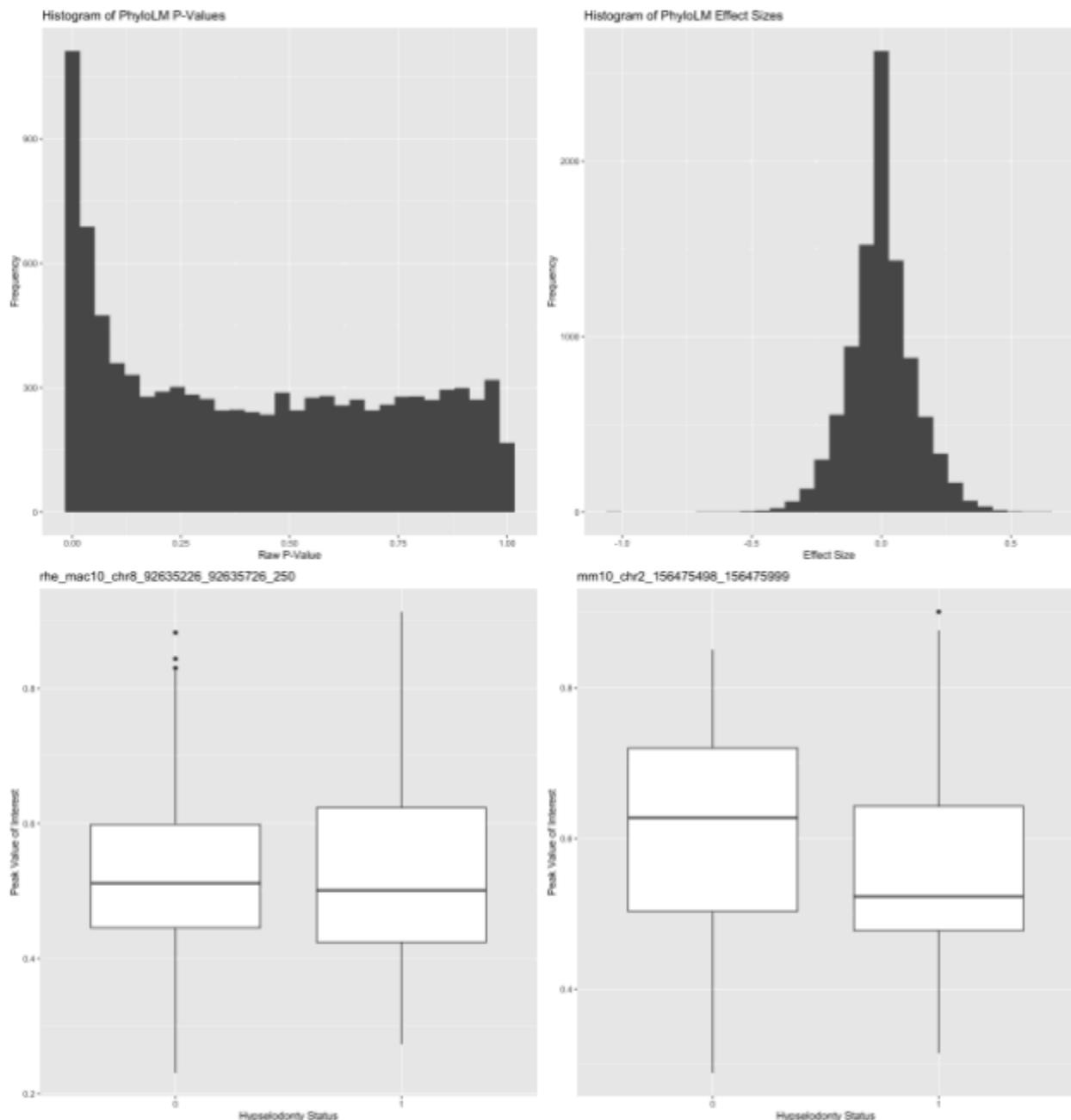


Figure 5: Results of phyloLM applied to open chromatin differences between hypselodonts and non-hypsodonts for 10,000 peaks. (Top-left) p-value distribution; (top-right) slope distribution; (bottom) barplot distributions of values from the most significant t-test results; (bottom-left) most

positively associated region, with larger values in hypselodonts; (bottom-right) most negatively associated region.

An alternative to phylolm is a similar method, phyloglm, with better support for binary traits such as ours. In short, phyloglm applies generalized linear models while accounting for correction with a given evolutionary tree; the model applied in our case is a simple logistic regression model operating on one variable that learns a specific threshold α_i for each region i with peak value x_i and a weight w_i . Then, if $w_i x_i < \alpha_i$ for a single instance, the model predicts that instance to be 0 (non-hypselodont) or 1 (hypselodont). While significantly more difficult to learn with much higher runtime than phylolm (due to the lack of a closed-form solution for the parameters), this model is designed for binary traits and thus captures group differences most effectively in theory.

In practice, our results of phyloglm yielded another right-skewed histogram, but no p-values were significant after adjustment; we also observed the effect sizes distributed around 0 as expected (Figure 6). However, inspection of the values for the most positively and negatively associated regions yielded neatly separated results for the majority of the items in each group (Figure 6). Because of this, we opted for the results of phyloglm to inform the results on which we elaborate in later sections.

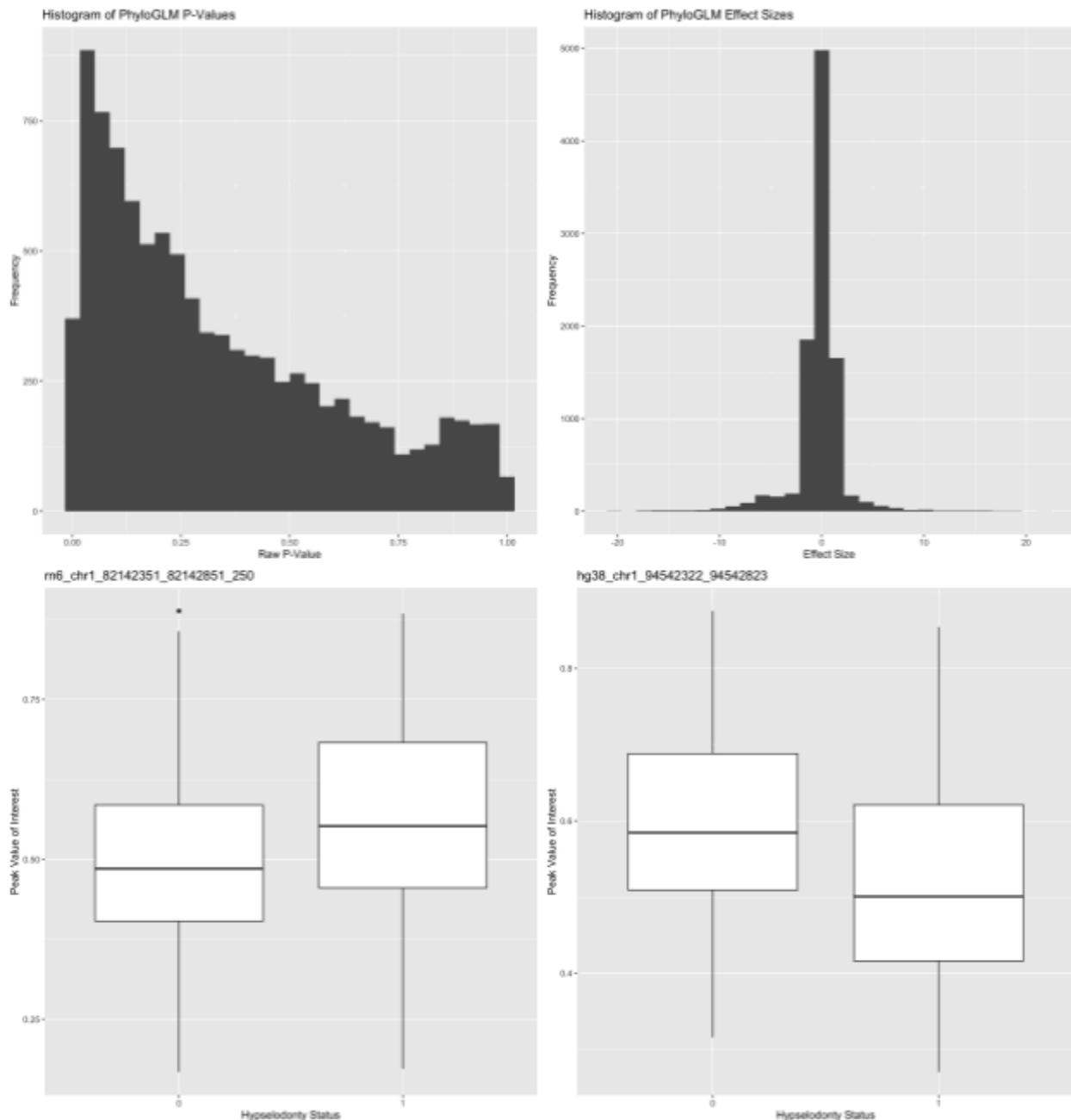


Figure 6: Results of phyloglm applied to open chromatin differences between hypselodonts and non-hypsodonts for 10,000 peaks. (Top-left) p-value distribution; (top-right) slope distribution; (bottom) barplot distributions of values from the most significant t-test results; (bottom-left) most positively associated region, with larger values in hypselodonts; (bottom-right) most negatively associated region.

2. Analysis of Results: OPCs

After the comparison of methods on OPCs in the previous section, we expanded on the results found from phyloglm. Some of the other results we found were interesting, however, and notable ones are put into supplemental appendices at the end of this report.

For the most positively associated phyloGLM region (rn6, chr1, 82142351-82142851), we found the closest gene to be *Rattus norvegicus* capicua transcriptional repressor (CIC) (Figure 7). Originally found in *Drosophila*, the CIC repressor has been linked to lung and immune system development, but also many aspects of brain development (Lee, 2020). Many studies examine CIC's role in cell proliferation, linking it to a variety of neural tumors (Kilian et al., 2020), but it has also been linked to neural stem cells, indicating a potential link to hypselodontology. Ahmad et al. (2021) found that deletion of CIC in the forebrain led to increased proliferation of neural stem cells. Interestingly, we expect the opposite effect as this region is closer to a region of open chromatin with higher peaks for hypselodonts that would have more such proliferation. The same study by Ahmad et al. (2021) confirmed that the loss of CIC led to cells becoming closer in identity to OPCs, while Yang et al. (2017) showed that the loss of CIC in neural stem cells halted development at OPCs. This arguably makes the results even more confusing, as we would expect to see more of this repressor active in non-hypselodonts. This discrepancy between observed and expected results might indicate that this region of open chromatin is not directly related to CIC, or could be used to hold a repressor that lowers CIC expression. Additionally, it also suggests that further study into the cells common in the forebrain such as neurons and astrocytes may produce further links to hypselodontology. It could also suggest that hypselodonts need much more control with regards to tumor suppression due to their constantly dividing tooth cells, but higher expression of a proliferation repressor would also lower tooth regeneration ability. Overall, more examination is necessary to relate this region to hypselodontology, possibly with experimental evidence.

For the most negatively associated phyloGLM region (hg38, chr1, 94542322-94542823), we found the closest gene to be Coagulation Factor III (F3) (Figure 7). F3 plays a significant role in the brain, with it being linked to memory and synaptic plasticity (Gulisano et al., 2016), but also neurogenesis. Critically, Bizzoca et al. (2012) found an inverse relationship between F3 and neurogenesis. Xenaki et al. (2011) find a similar result, where F3 expression suppresses the activity of Shh, which normally promotes neurogenesis and is responsible for many aspects of tooth development (Hardcastle et al., 1998). This indicates a likely relation between this gene and hypselodontology. Since hypselodonts need constant cell proliferation to regenerate their teeth, it makes sense that F3 open chromatin would be lowered in the hypselodonts, especially in a cell that is similar to neural and dental stem cells. While some research has been performed on F3 in mice (a common hypselodont), the researchers do not specifically focus on the mice teeth, so experimental validation of this link is still necessary.

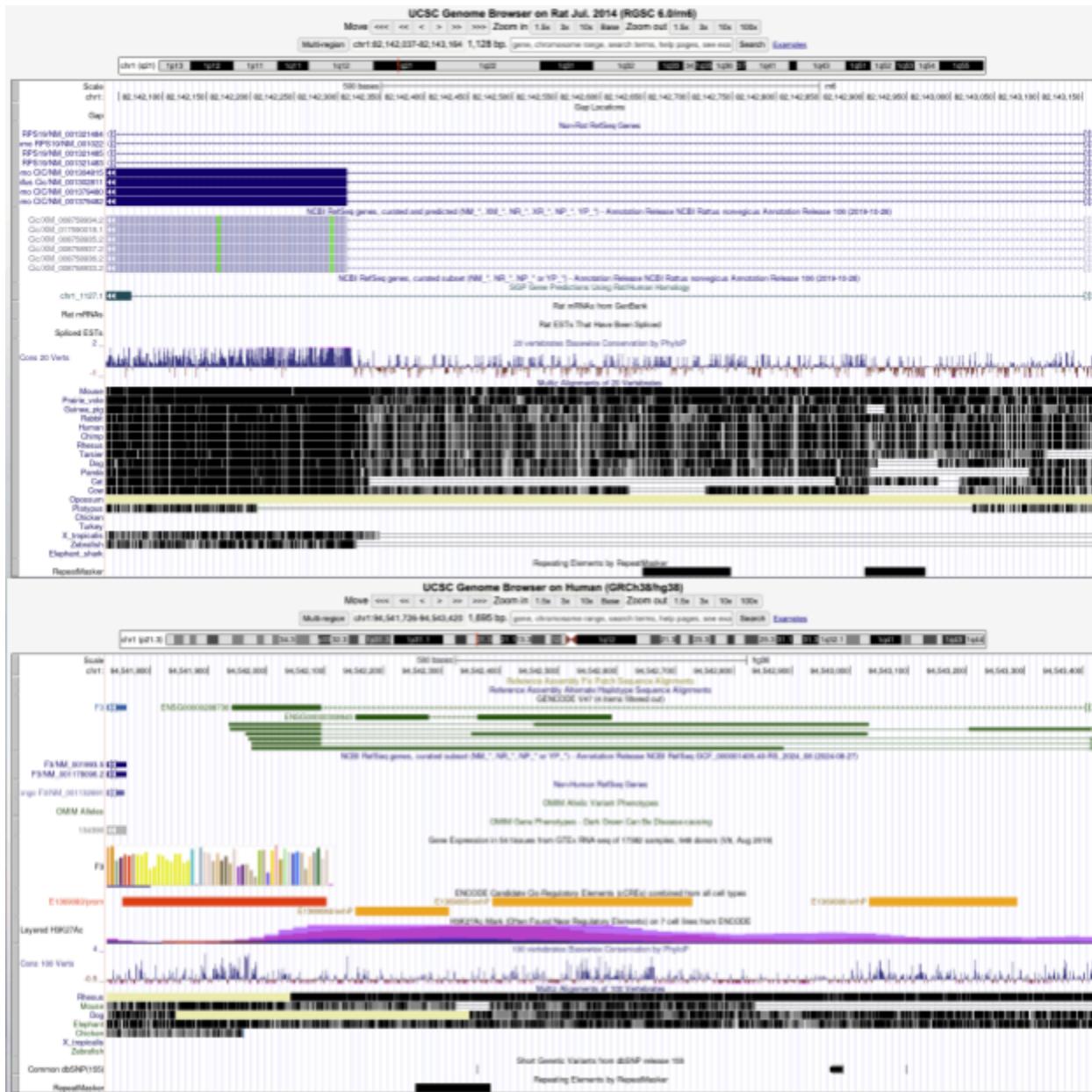


Figure 7: UCSC Genome Browser tracks for the most positively (top) and negatively (bottom) associated regions from our analysis of OPCs. Browser windows have been extended to view the nearest gene.

After subsetting the predictions by genome and retrieving the top positive and negative hits for both human and mouse genomes, a GREAT gene ontology analysis yielded results with some potential linkage to hypselodonty. For our comparison to the human genome, we found one enriched biological process, skeletal muscle cell differentiation, for our positively associated peaks. On a high level, some link could exist between hypselodonty and this process; namely, mesenchymal stem cells differentiate into both skeletal muscle cells and dental cells, so this could form a potential relationship. Some studies confirm a link between some of the genes

involved in muscle cell differentiation (myogenesis) and dental health; Hong et al. (2024) found lower levels of MYOG, a factor involved in myogenesis, to be associated with worse dental health. Furthermore, Shh is included among these genes, and its link with tooth development has been confirmed in various studies (Hardcastle et al., 1998, Hosoya et al., 2020). Further experimental evidence would be needed to confirm a link between this process and hypselodonty, however.

For our comparison to the mouse genome, we found one enriched biological process and two enriched molecular functions for our positively associated regions, all of which have to do with RNA Polymerase II (RNAPolII) transcription. RNAPolII is largely responsible for the transcription of mRNA, meaning that hypselodonts in general transcribe more genes. It is known that cells that constantly divide (i.e., have faster growth rates) produce more mRNA (García-Martínez et al., 2016), so this could have some relationship to hypselodonts' constantly dividing dental cells. While some studies examine mRNA levels of dental cells for specific genes, we were unable to find studies that examined total mRNA count between hypselodont dental cells and non-hypselodont dental cells. As such, experimental validation is necessary to confirm these processes as linking to hypselodonty.

Our single-cell analysis with the Allen Brain Map and CATLAS also allowed us to perform comparisons between human (non-hypselodont) and mouse (hypselodont), allowing us to potentially draw meaningful conclusions from our comparisons. First, we performed a visualization of the genes we analyzed earlier in the Allen Brain Map, visualized in a UMAP plot in Figure 8. In general, the results of this are not particularly insightful; CIC is expressed more in mouse brain than human brain as we might have expected from our earlier results, but these are expressed throughout the various cell types. F3, on the other hand, is concentrated in astrocytes, with it being more highly expressed in mouse astrocytes than human astrocytes.



Figure 8: Allen Brain Map UMAP visualizations of gene expression in mouse brain. (Left) CIC, (Right) F3, concentrated in astrocytes. Compared to human brain, CIC has higher expression in mouse, and F3 has lower expression in human astrocytes.

Results of visualization on CATLAS follow similar patterns; CIC is expressed more in mouse versus human, but there is some cell-type specificity. While OPCs do show more CIC in mouse than human, the same also holds for Oligodendrocytes, and these have higher expression of CIC than OPCs. We similarly found that F3 was largely expressed in humans and much less so in mouse, though a general concentration of F3 in astrocytes was still present. These results indicate that the processes we discovered through OPC may not be cell-type specific and may apply to a broader set of cells, and they also indicate that future work may find astrocytes as a particularly useful cell type to study in relation to hypselodonty.

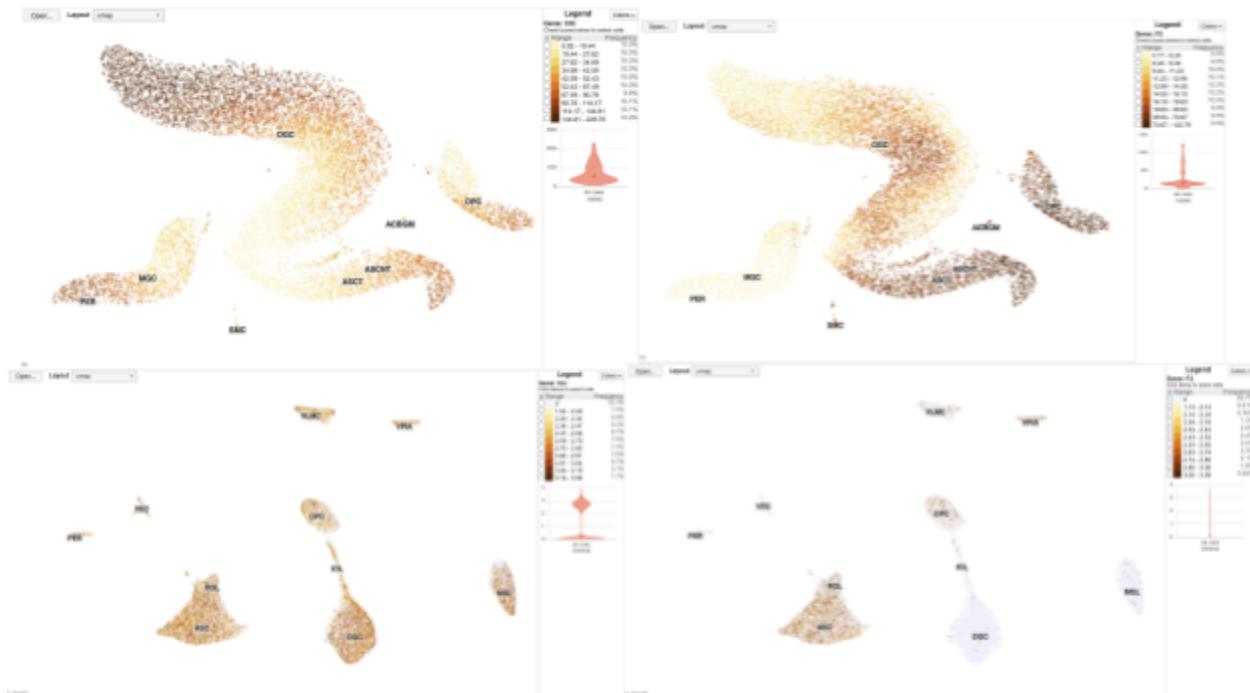


Figure 9: CATLAS Open Chromatin Visualization of the genes identified to be differential for OPCs, compared in both human and mouse brains. (Top) Human, (Bottom) Mouse, (Left) CIC, (Right) F3.

3. Analysis of Results: Oligodendrocytes

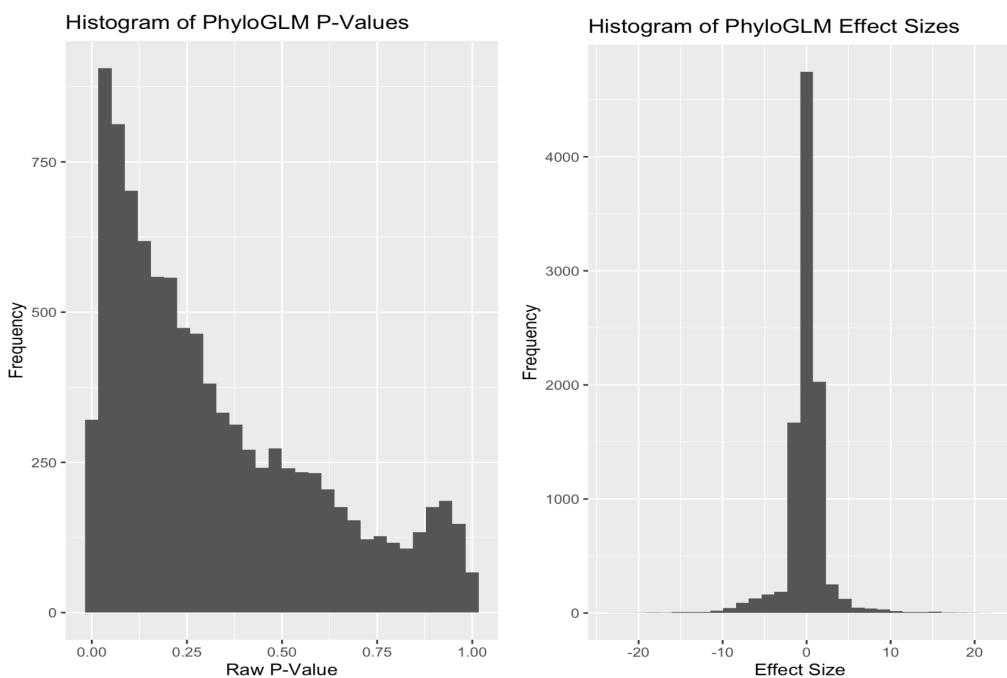
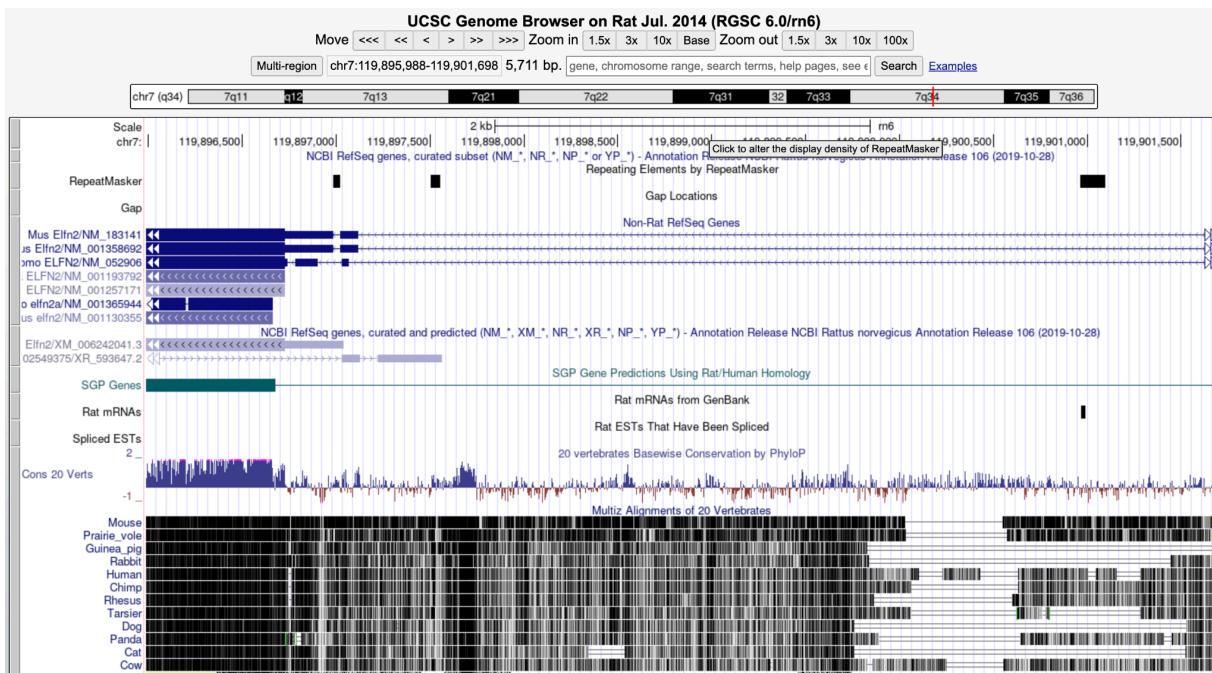


Figure 10: The distribution of p-values and effect sizes for the association testing using phyloglm for Oligodendrocytes.

In the Oligodendrocytes analysis, we found 2308 peaks which were significantly associated with the hypselodonty trait at the significance level of 0.1 (adjusted p-values). The most positively associated peak we found in hypselodonts vs non-hypselodonts was: rn6, chr7, 119898593-119899093. The closest gene to this peak region is ELFN2 (extracellular leucine-rich repeat and fibronectin type III). The main GO annotations related to this gene include phosphatase binding and protein phosphatase inhibitor activity. It is predicted to be involved in synaptic membrane adhesion (“ELFN2 - Gene - NCBI,” 2024). Since protein phosphatases like PP1, PP2A, and PP2B are known to regulate phosphorylation states of synaptic proteins, and hence modulate processes like synapse formation and receptor

trafficking, the effect on synaptic membrane adhesion is expected and can have significant downstream effects on neural connections which could be linked to hypselodontology. Martemyanov et. al. (2019) also found the significance of ELFN2 protein as a postsynaptic cell adhesion molecule in the brain and its function in controlling group III mGluRs (metabotropic glutamate receptors). The mGluRs have essential functions in neuron excitability, learning, and memory, but there isn't any solid evidence of a link with hypselodontology yet. However, given that all these molecules are so intricately linked to brain function, there may very well be mechanisms through which ELFN2 is linked to hypselodontology.

The most negatively associated peak we found in hypselodonts vs non-hypselodonts was: hg38, chr2, 211745236-211745737. The closest gene to this peak region is ERBB4 (erb-b2 receptor tyrosine kinase 4). The corresponding protein is known to regulate cell differentiation and proliferation ("ERBB4 - Gene - NCBI," 2024). It plays an essential role in the development of Central Nervous System through its interaction with neuregulins (neuron supporting celltypes, including oligodendrocytes, astrocytes, microglia etc.). Upon doing a literature review, we found a paper linking this gene directly to tooth development in rodents (Lillesaar et. al. (2002)). They found that NRG-1 (neuregulin-1) and receptor ERBB4 (and ERBB3) are expressed locally during tooth development in rodents. One possible explanation for this could be that the neuregulin/ERBB interaction in mesenchymal cells could trigger the proliferation of teeth in rodents. This explanation is plausible because the neuregulin/ERBB interaction is known to cause proliferation of cells in different parts of the body (Mei, L., & Nave, K. A. (2014)).



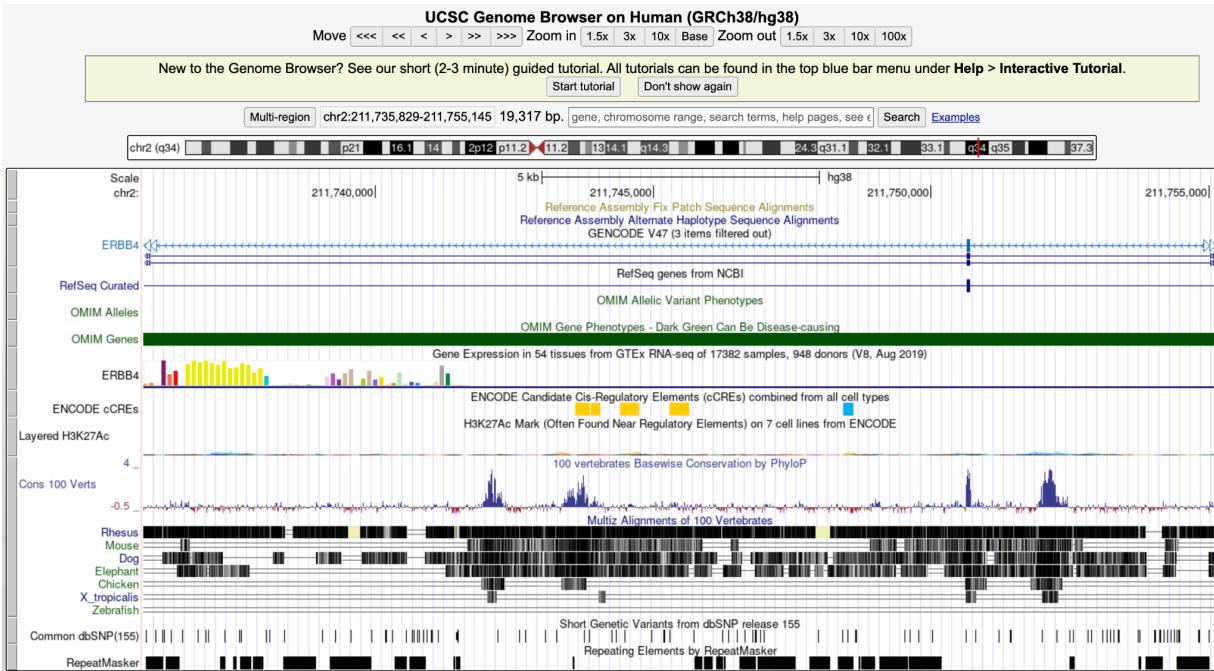


Figure 11: UCSC Genome Browser tracks for the most positively (top) and negatively (bottom) associated regions from our analysis of Oligodendrocytes. Browser windows have been extended to view the nearest gene.

Then, we submitted our top 200 positively and negatively associated peaks to the GREAT gene Ontology analysis tool, for both the human and mouse background genomes. We found significant terms only in one case: when comparing the top 200 negatively associated peaks with the mouse genome as background. The found terms included 1 Biological term: axon ensheathment and 6 Mouse Phenotype terms, which did not have any direct link with the brain or hypselodonty. The axon ensheathment term is related to the myelination of axons of neurons, indicating that the identified negatively associated peaks have some neural link, but upon literature review, we couldn't find any direct link between this term and hypselodonty.

Then, looked into the data for our identified genes in the Allen Brain Map. For the ELF2 gene, we found that there is some expression of it in VIP (inhibitory) neurons and less in other cell types. For the ERBB4 gene, we found distinctly high expression levels in most inhibitory neuron types and very low expression in other celltypes. One validation for our findings for the ELF2 gene was that there was low expression of this gene in the Oligodendrocytes from the Human Allen Brain Map, but there was high expression of it in the Mouse Allen Brain Map. Since mice are hypselodont but humans are not, the differential high expression of the nearest gene to our most positively associated peak provided indication that that peak might be playing a role in the expression of the ELF2 gene.

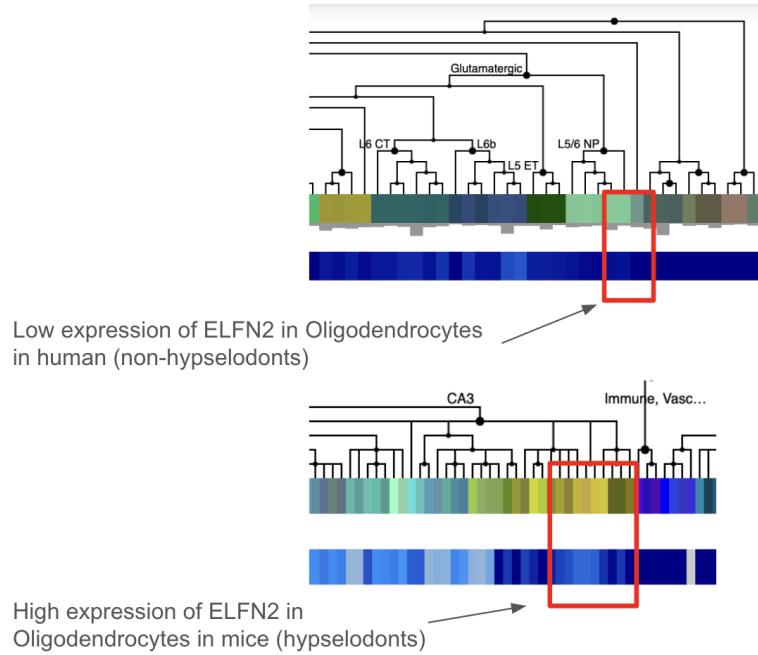


Figure 12: Heatmaps showing expression of *ELFN2* in Human (top) and Mouse (bottom) Allen Brain Maps. Highlighted areas represent Oligodendrocytes.

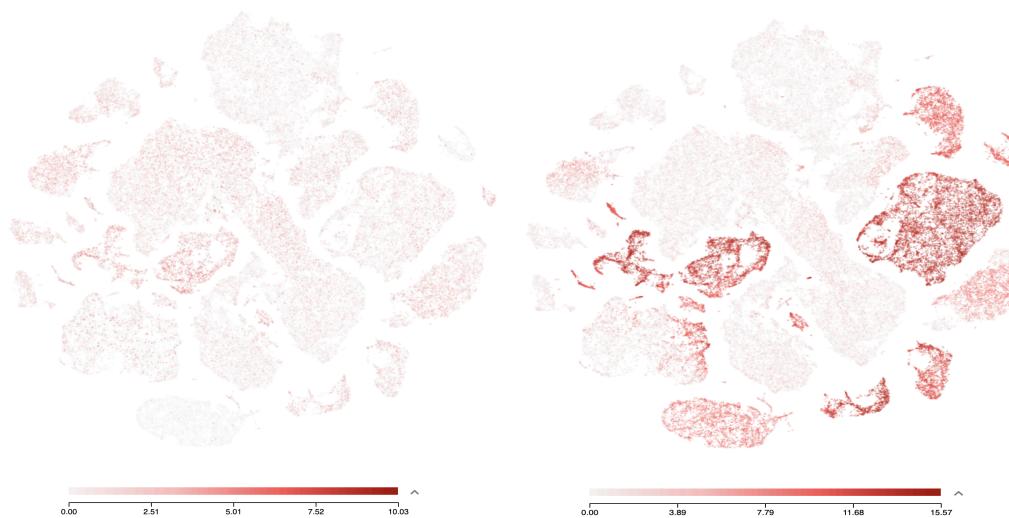


Figure 13: Allen Brain Map UMAP visualizations of gene expression in mouse brain. (Left) *ELFN2*, (Right) *ERBB4*, concentrated in inhibitory neuron types.

Upon analyzing the Catlas resource, we found that the open chromatin peaks in different cell types corresponding to the two genes matched the found expression patterns of the genes in the different brain cell types from the Allen Brain Map. This further strengthens the evidence of the link of these genes with the brain and the possibility of influencing the hypsodonty trait in species.

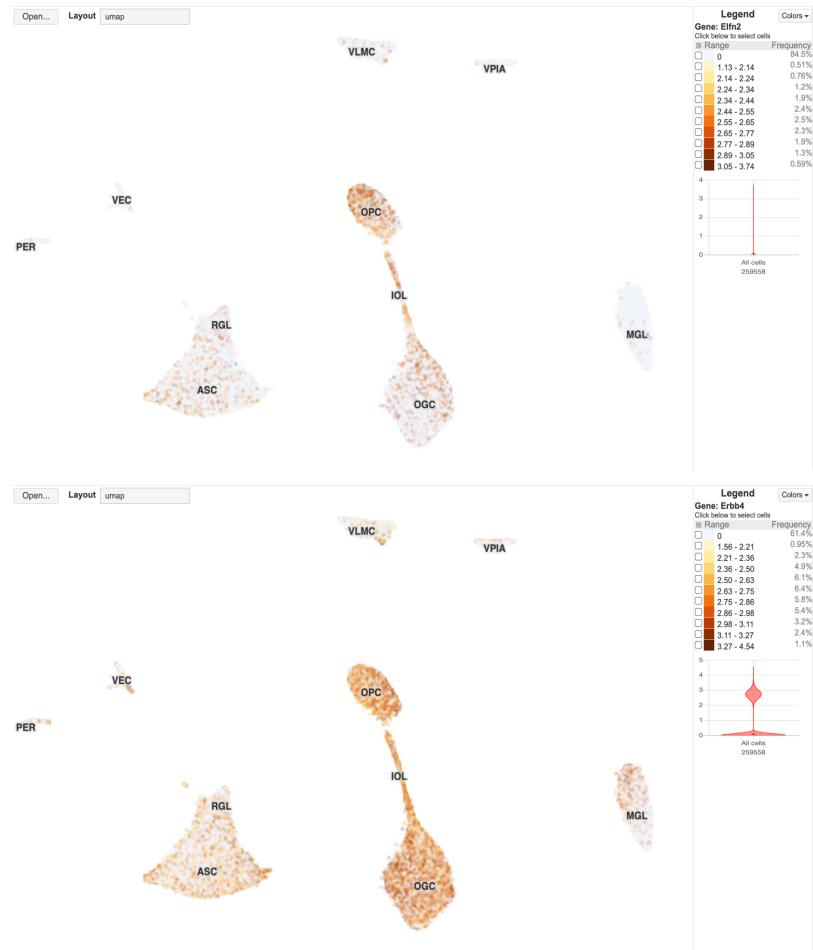


Figure 14: CATLAS Open Chromatin Visualization of the genes identified to be differential for Oligodendrocytes, compared in mouse brain. (Top) *ELFN2*, (Bottom) *ERBB4*.

4. Comparison of Cell Types

For both cell types, we used phyloglm to test for significantly different open chromatin regions. After adjustment, none of these tests provided a significant p-value. There may be several causes for this; one is that the default adjustment method for p-values in R is Bonferroni correction, which is good at reducing Type I errors (false positives) but also brings an increased rate of Type II errors (false negatives). Given that phyloglm was able to detect open chromatin differences better and more confidently than other methods, the possibility of false negatives with this adjustment is likely.

Despite this, our results yielded interesting insights on the relationship between hypselodonty and OPCs and oligodendrocytes and we were able to find peaks that had some potential association with the trait. For OPCs, we found genes that were related to cell division as expected, and our analysis of oligodendrocytes yielded genes that related to nervous system control and development. Our results for GREAT gene ontology revealed the peaks associated with OPCs had some relation to hypselodonty in the form of increased transcription and skeletal muscle cell differentiation, though the result obtained for oligodendrocytes (axon ensheathment)

has no known direct links with hypselodontology yet. Single-cell analysis showed that while oligodendrocyte peaks followed expected patterns based on our results, the same could not be said for the OPC peaks.

For both cell types, we do not have enough evidence to link one of these more strongly to hypselodontology, given that both cell types show few (if any) significantly different peaks and have varying relations to the genes with which the peaks are associated, as well as varying GREAT and single-cell results. More research is needed to assess the different impacts these cell types have on hypselodontology.

Conclusions and Future Work

Overall, our results showed interesting insights into the relationship between hypselodontology and OPCs and oligodendrocytes, with some genes being particularly insightful into the understanding of hypselodontology's relationship to neural systems. Such insights could be used to further study the development of dental systems in hypselodonts and non-hypselodonts, the development of the dental system in tandem with the nervous system, and could even be applied to dental regeneration studies in humans.

We highlight three areas of further research that could be conducted. Firstly, we hypothesize that more in-depth statistical analysis may be necessary. In our comparison of the peaks significantly associated with hypselodontology for OPCs and oligodendrocytes, we discuss how the Bonferroni test may be too corrective and lead to Type II (false negative) errors. Furthermore, we hypothesize that studies into expression quantitative trait loci (eQTLs) (i.e., SNP-gene pairs, or even OCR-gene pairs) could lead to even more insightful discoveries about genes and processes that relate to hypselodontology. Specifically, while looking at nearby genes can often yield noteworthy insights, it is not necessarily the case that the open chromatin region is related to the gene to which it is closest. Further statistical analyses could include the analysis of combinations of OCRs, or the inclusion of more peaks; due to the limitations of time and computational power, we looked at a random sample of 10,000 out of millions, so including more peaks could yield even more insights into hypselodontology.

Secondly, we hypothesize that the analysis of other cell types may be noteworthy. In our single-cell results on OPCs examined in the Allen Brain Map, we found that the negatively associated gene of interest, F3, was highly expressed in mouse astrocytes (Figure 8). Furthermore, teeth contain a complex system of nerves, so there may even be some merit in studying neurons of various types.

Finally, we propose that further experimental work is needed not only to validate the results found in our analyses, but also to provide future directions of work. Namely, while cells in the brain itself may have some relation to hypselodontology (per our analyses), the trait as a whole could be more easily understood with similar single-cell analysis on cells in the dental pulp or mesenchyme. Furthermore, a comparison of different strains of mice can be performed to analyze each strain's capability to continuously grow teeth, which could provide further nuance into the genes involved in hypselodontology.

Appendix A: Other Noteworthy Results and Analyses for OPCs

In our comparison of methods on OPCs, we found some noteworthy results for our usage of other methodologies. While we did not expand on them very in-depth for the purposes of this report (i.e., we did not conduct an advanced literature review, supply associated UCSC browser track, or examine single-cell results), we believed these results to be interesting enough to mention and briefly discuss for potential future research.

1. Results of PhyloLM

The top positive hit of phyloLM was found to be closest to *Macaca mulatta* RUNX1 translocation partner 1 (RUNX1T1), while the top negative hit was found to be closest to *Mus musculus* erythrocyte membrane protein band 4.1 like 1 (EPB41I1). RUNX1T1 is commonly associated with a variety of cancers, especially leukemia. Some evidence suggests that it may also be associated with the nervous system, but we were unable to find strong evidence linking it to neural or stem cell differentiation. EPB41I1 is a known tumor suppressor gene, and Han et al. (2019) found that its decreased expression led to the increase in glioblastoma proliferation, so there may be a link here with hypselodonty as well, where the decrease in this gene allows hypselodonts to have continuously-growing teeth.

2. Results of T-Tests

The top positive t-test hit was found to be closest to *Macaca mulatta* lysine demethylase 4B (KDM4B), while the top negative hit was found to be closest to *Macaca mulatta* UV stimulated scaffold protein A (UVSSA). KDM4B is highly expressed in many cancers, and its down-regulation led to abnormal development in mice (Wilson and Krieg, 2019). While we were unable to find studies specifically linking it to neural cell proliferation or the dental system, this can form a potential link for future research. UVSSA is known in humans to repair DNA damage by UV light; despite this, we were unable to find studies showcasing its involvement in neural or dental processes, likely because these systems are not often exposed to UV light.

3. GREAT Gene Ontology

For our negatively associated phyloLM peaks compared to the mouse genome, we found enrichment in cortical cytoskeleton and cortical actin cytoskeleton. This could relate to the migration of cells, but it would not necessarily make sense as to why cell migration would be down-regulated in hypselodonts as the constant division of cells in teeth would imply higher levels of migration. As such, a further link to hypselodonty requires more research and/or experimental results.

References

- Duan, Y., Liang, Y., Yang, F., Ma, Y.. Neural Regulations in Tooth Development and Tooth-Periodontium Complex Homeostasis: A Literature Review. *Int J Mol Sci.* 2022 Nov 16;23(22):14150. doi: 10.3390/ijms232214150. PMID: 36430624; PMCID: PMC9698398.

Klein, O. D., Lyons, D. B., Balooch, G., Marshall, G. W., Basson, M. A., Peterka, M., ... & Martin, G. R. (2008). An FGF signaling loop sustains the generation of differentiated progeny from stem cells in mouse incisors.

Renvoisé, E., & Michon, F. (2014). An Evo-Devo perspective on ever-growing teeth in mammals and dental stem cell maintenance. *Frontiers in Physiology*, 5, 324.

Tummers, M., Thesleff, I. Root or crown: a developmental choice orchestrated by the differential regulation of the epithelial stem cell niche in the tooth of two rodent species. *Development* 15 March 2003; 130 (6): 1049–1057. doi: <https://doi.org/10.1242/dev.00332>

Wang X-P, Suomalainen M, Felszeghy S, Zelarayan LC, Alonso MT, Plikus MV, et al. (2007) An Integrated Gene Regulatory Network Controls Stem Cell Proliferation in Teeth. *PLoS Biol* 5(6): e159. <https://doi.org/10.1371/journal.pbio.0050159>

Zhao, H., Feng, J., Seidel, K., Shi, S., Klein, O., Sharpe, P., & Chai, Y. (2014). Secretion of shh by a neurovascular bundle niche supports mesenchymal stem cell homeostasis in the adult mouse incisor. *Cell Stem Cell*, 14(2), 160-173.

Fleagle, J. G. (1999). Neotropical Paleontology. Evolutionary Anthropology Issues News and Reviews, 8(3), 77–78.

[https://doi.org/10.1002/\(sici\)1520-6505\(1999\)8:3%3C77::aid-avan1%3E3.0.co;2-7](https://doi.org/10.1002/(sici)1520-6505(1999)8:3%3C77::aid-avan1%3E3.0.co;2-7)

Askari, N., Yaghoobi, M. M., Shamsara, M., & Esmaeili-Mahani, S. (2015). Human dental pulp stem cells differentiate into oligodendrocyte progenitors using the expression of Olig2 transcription factor. *Cells Tissues Organs*, 200(2), 93-103.

Degistirici, Ö., Jaquiere, C., Schönebeck, B., Siemonsmeier, J., Götz, W., Martin, I., & Thie, M. (2008). Defining properties of neural crest-derived progenitor cells from the apex of human developing tooth. *Tissue Engineering Part A*, 14(2), 317-330.

Nishiyama, A. L. X. G. H. H., Lin, X. H., Giese, N., Heldin, C. H., & Stallcup, W. B. (1996). Co-localization of NG2 proteoglycan and PDGF α -receptor on O2A progenitor cells in the developing rat brain. *Journal of neuroscience research*, 43(3), 299-314.

Ahmad ST, Rogers AD, Chen MJ, Dixit R, Adnani L, Frankiw LS, Lawn SO, Blough MD, Alshehri M, Wu W, Marra MA, Robbins SM, Cairncross JG, Schuurmans C, Chan JA. Capicua regulates neural stem cell proliferation and lineage specification through control of Ets factors. *Nat Commun*. 2019 May 1;10(1):2000. doi: 10.1038/s41467-019-09949-6. PMID: 31043608; PMCID: PMC6494820.

Bizzoca A, Corsi P, Polizzi A, Pinto MF, Xenaki D, Furley AJ, Gennarini G. F3/Contactin acts as a modulator of neurogenesis during cerebral cortex development. *Dev Biol*. 2012 May 1;365(1):133-51. doi: 10.1016/j.ydbio.2012.02.011. Epub 2012 Feb 21. PMID: 22360968.

Gulisano W, Bizzoca A, Gennarini G, Palmeri A, Puzzo D. Role of the adhesion molecule F3/Contactin in synaptic plasticity and memory. *Mol Cell Neurosci*. 2017 Jun;81:64-71. doi: 10.1016/j.mcn.2016.12.003. Epub 2016 Dec 28. PMID: 28038945.

Kilian M, Friedrich M, Sanghvi K, Green E, Pusch S, Kawauchi D, Löwer M, Sonner JK, Krämer C, Zaman J, Jung S, Breckwoldt MO, Willimsky G, Eichmüller SB, von Deimling A, Wick W, Sahm F, Platten M, Bunse L. T-cell Receptor Therapy Targeting Mutant Capicua Transcriptional Repressor in Experimental Gliomas. *Clin Cancer Res*. 2022 Jan 15;28(2):378-389. doi: 10.1158/1078-0432.CCR-21-1881. Epub 2021 Nov 15. PMID: 34782365; PMCID: PMC9401455.

Lee Y. Regulation and function of capicua in mammals. *Exp Mol Med*. 2020 Apr;52(4):531-537. doi: 10.1038/s12276-020-0411-3. Epub 2020 Apr 1. PMID: 32238859; PMCID: PMC7210929.

Xenaki D, Martin IB, Yoshida L, Ohyama K, Gennarini G, Grumet M, Sakurai T, Furley AJ. F3/contactin and TAG1 play antagonistic roles in the regulation of sonic hedgehog-induced cerebellar granule neuron progenitor proliferation. *Development*. 2011 Feb;138(3):519-29. doi: 10.1242/dev.051912. PMID: 21205796; PMCID: PMC3014637.

Yang R, Chen LH, Hansen LJ, Carpenter AB, Moure CJ, Liu H, Pirozzi CJ, Diplas BH, Waitkus MS, Greer PK, Zhu H, McLendon RE, Bigner DD, He Y, Yan H. Cic Loss Promotes Gliomagenesis via Aberrant Neural Stem Cell Proliferation and Differentiation. *Cancer Res*. 2017 Nov 15;77(22):6097-6108. doi: 10.1158/0008-5472.CAN-17-1018. Epub 2017 Sep 22. PMID: 28939681; PMCID: PMC5690824.

García-Martínez J, Delgado-Ramos L, Ayala G, Pelechano V, Medina DA, Carrasco F, González R, Andrés-León E, Steinmetz L, Warringer J, Chávez S, Pérez-Ortín JE. The cellular growth rate controls overall mRNA turnover, and modulates either transcription or degradation rates of particular gene regulons. *Nucleic Acids Res*. 2016 May 5;44(8):3643-58. doi: 10.1093/nar/gkv1512. Epub 2015 Dec 29. PMID: 26717982; PMCID: PMC4856968.

Han X, Wang X, Li H, Zhang H. Mechanism of microRNA-431-5p-EPB41L1 interaction in glioblastoma multiforme cells. *Arch Med Sci*. 2019 Oct;15(6):1555-1564. doi: 10.5114/aoms.2019.88274. Epub 2019 Sep 26. PMID: 31749885; PMCID: PMC6855151.

Wilson C, Krieg AJ. KDM4B: A Nail for Every Hammer? *Genes (Basel)*. 2019 Feb 12;10(2):134. doi: 10.3390/genes10020134. PMID: 30759871; PMCID: PMC6410163.

Hardcastle, Z., Mo, R., Hui, C. C., & Sharpe, P. T. (1998). The Shh signalling pathway in tooth development: defects in Gli2 and Gli3 mutants. *Development*, 125(15), 2803-2811.

Hong, S. W., Baek, J. H., Kim, K., & Kang, J. H. (2024). Complex interplay of oral health, muscle and bone metabolism, and frailty in older individuals. *Clinical oral investigations*, 28(1), 116.

Hosoya, A., Shalehin, N., Takebe, H., Shimo, T., & Irie, K. (2020). Sonic hedgehog signaling and tooth development. International Journal of Molecular Sciences, 21(5), 1587.

Fried K, Risling M, Tidcombe H, Gassmann M, Lillesaar C. Expression of ErbB3, ErbB4, and neuregulin-1 mRNA during tooth development. *Dev Dyn*. 2002 Jul;224(3):356-60. doi: 10.1002/dvdy.10114. PMID: 12112465.

Mei L, Nave KA. Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases. *Neuron*. 2014 Jul 2;83(1):27-49. doi: 10.1016/j.neuron.2014.06.007. PMID: 24991953; PMCID: PMC4189115.