



Universidad Nacional Autónoma de México
Facultad de Ciencias
Genómica Computacional
2023-2

Análisis Filogenético de la ropsodina en la familia Octopodidae

Antonio Sebastián Dromundo
Diego Dozal Magnani
Giovani Ramírez Cortés

Introducción

La rodopsina es una proteína fotosensible que desempeña un papel crucial en la visión del humano y muchos otros vertebrados en condiciones de poca luz al convertir la luz en señales eléctricas que el cerebro puede interpretar en imágenes. Se compone principalmente de la opsina y el retinal. La opsina es una proteína que actúa como un receptor de luz, mientras que el retinal es un pigmento sensible a la luz que se une a la opsina. El retinal es esencial para captar la luz (Hubbard, R., & Kropf, A. (1958)).

Los pulpos pueden ajustar la sensibilidad de su visión alterando la cantidad de rodopsina y su estado de absorción de luz, que le permite adaptarse rápidamente a entornos y camuflarse eficazmente. En el caso de los pulpos, la proteína se encuentra en las células de los conos de la retina de los pulpos, que son los responsables de la visión diurna y percepción de colores (Hanke, F. D. & Kelber, A. (2019)).

Pregunta de Investigación

¿Existen diferencias significativas en la proteína rodopsina entre las distintas especies de pulpos ?

Objetivo :

Determinar diferencias entre las especies de pulpos seleccionadas a partir de las secuencias de la rodopsina utilizando métodos estadísticos y análisis filogenético.

Métodos :

- Recopilar archivos FASTA de la proteína rodopsina de 50 especies de pulpos de la base de datos de NCBI : <https://www.ncbi.nlm.nih.gov/>
- Implementar tres métricas de distancia para alineación de secuencias y visualización de árboles filogenéticos y gráficas de proporción de nucleótidos en python.
- Bibliotecas utilizadas en Python : numpy, matplotlib, scipy.

Especies consideradas :

Octopus Sinensis	Abdopus aculeatus	Adelieled one polymorpha	Ameloctopus litoralis	Argonauta nodosus	Bathypolypus sponsalis	Bolitaena pygmaea	Callistoctopus macropus	Callistoctopus ornatus	Cistopus indicus
Enteroctopus dofileini	Haliphron atlanticus	Hapalochlaena maculosa	Japetella diaphana	Macroctopus maorum	Macrotritopus defilippi	Megaleledone setebos	Muusoctopus eureka	Muusoctopus levis	Muusoctopus longibrachus
Muusoctopus oregonensis	Muusoctopus yaquinae	Octopus berrima	Octopus bimaculoides	Octopus californicus	Octopus fitchi	Octopus hummeli ncki	Octopus insularis	Octopus kaurna	Octopus maya

Octopus mimus	Octopus rubescens	Octopus vulgaris	Pareledone aequipapillae	Pareledone albimaculata	Pareledone aurata	Pareledone charcoti	Pareledone cornuta	Pareledone serperastata	Pareledone subtilis
Pareledone turqueti	Paroctopus digueti	Sasakiopus salebrosus	Scaeurgus unicolorrhynchus	Thaumelodone gunteri	Thaumelodone peninsulae	Tremoctopus violaceus	Velodona togata	Vitreledonella richardi	

Tabla 1 : Especies de pulpos

Métricas utilizadas :

Los árboles filogenéticos que se propusieron son conocidos como “árboles basados en distancias”; lo primero que se necesita es comparar la diferencia entre dos cadenas, en nuestro caso, de misma longitud; a continuación se detallaran las métricas que se utilizaron para este experimento:

- La distancia de Jaccard es una medida que calcula la similitud entre dos conjuntos; es decir, para cualquier par de conjuntos A, B podemos definir su distancia entre ambos por la siguiente fórmula:

$$d_J(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

- La distancia de Hamming es una extensión de la distancia discreta; para cualesquiera $A, B \in \Sigma$ definimos su distancia como:

$$d_H(A, B) = \sum_{i=0}^n d_D(A_n, B_n)$$

donde $n = \text{long}(A) = \text{long}(B)$

$$d_D(A_n, B_n) = \begin{cases} 0 & \text{si } A_n = B_n \\ 1 & \text{si } A_n \neq B_n \end{cases}$$

- Para la métrica que se propone debemos definir $\# : \Sigma \rightarrow \mathbf{N}$ dada por :

$$\#(B) = \#_A(B) + \#_T(B) + \#_C(B) + \#_G(B)$$

donde cada función $\#_A, \#_T, \#_C, \#_G : \Sigma \rightarrow \mathbf{N}$ calcula la cantidad de aminoácidos en la secuencia.

Con esto presentado definimos la distancia “Contenido A-T-C-G” entre dos secuencias A y B como:

$$d(A, B) = |\#(A) - \#(B)|$$

Nótese que esta distancia NO es una métrica, ya que no cumple lo siguiente:

$$\text{Si } d(A, B) = 0 \Leftrightarrow A = B$$

Tomemos por ejemplo las cadenas “GATA” y “TAGA”, aunque ambas cadenas tienen el mismo contenido ATCG, estas cadenas son totalmente distintas; entonces, ¿Está mal la propuesta?

Mónica Clapp define una métrica en cualquier conjunto X como una función $d : X \times X \rightarrow \mathbb{R}$ que cumple tres propiedades.

- $d(x, y) = 0 \Leftrightarrow x = y$.
- $d(x, y) = d(y, x)$ para cualesquiera $x, y \in X$.
- $d(x, z) \leq d(x, y) + d(y, z)$ para cualesquiera $x, y, z \in X$.

y mas adelante nos demuestra que cualquier distancia es no negativa.(Mónica Clapp ,2015); en esta demostración se utiliza que $d(x, x) = 0$, aunque esta afirmación suene trivial, es muy importante de resaltar, ya que podemos debilitar el punto (1.) por el siguiente axioma:

1. $d(x, y) = 0$ si $x = y$. (G. Alberto & G. José, 2017, pp. 13)

Por otro lado, Marta Macho debilita aún más esta afirmación diciendo para cada $x \in X$, $d(x, x) = 0$.

Definimos para un conjunto X una nueva función $d: X \times X \rightarrow \mathbb{R}_{\geq 0}$ que cumple las siguientes propiedades.

1. para cada $x \in X$, $d(x, x) = 0$.
2. $d(x, y) = d(y, x)$ para cualesquiera $x, y \in X$.
3. $d(x, z) \leq d(x, y) + d(y, z)$ para cualesquiera $x, y, z \in X$.

estamos contemplando la posibilidad de que existan $x \neq y$ en X con $d(x, y) = 0$. Entonces recibe el nombre de pseudométrica (M. Marta, 2009, pp. 23).

Probemos si nuestra propuesta cumple la definición de pseudométrica:

1. Para cualquier cadena $x \in \Sigma$ tenemos que $d(x, x) = |\#(x) - \#(x)| = |0| = 0$.
2. Dados cualesquiera $x, y \in \Sigma$ se puede notar como

$$d(x, y) = |\#(x) - \#(y)| = |-1| |\#(y) - \#(x)| = d(y, x).$$

3. Finalmente si tenemos $x, y, z \in \Sigma$, por definición de $|\cdot|$ podemos concluir

$$|\#(x) - \#(z)| \leq |\#(x) - \#(y)| + |\#(y) - \#(z)|.$$

Hemos probado que nuestra propuesta es una pseudodistancia. Aunque no es una métrica, esta función mide la diferencia de Contenido A-T-C-G entre dos secuencias de aminoácidos; por lo que sería interesante observar el árbol filogenético que refleja.

Resultados :

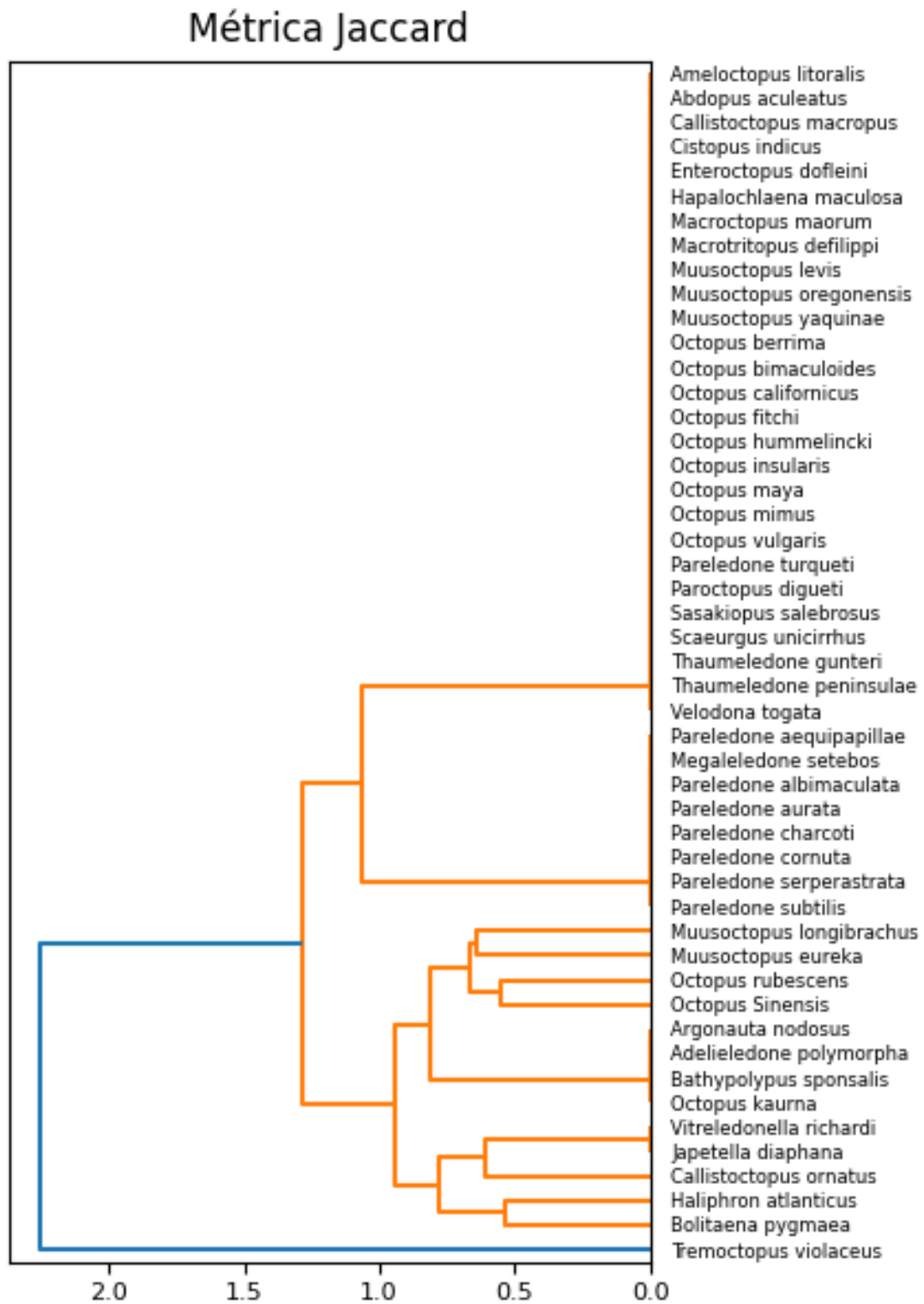


Figura 1 : Árbol Filogenético Jaccard

Arboles Filogenéticos (UPGMA)

Métrica Hamming

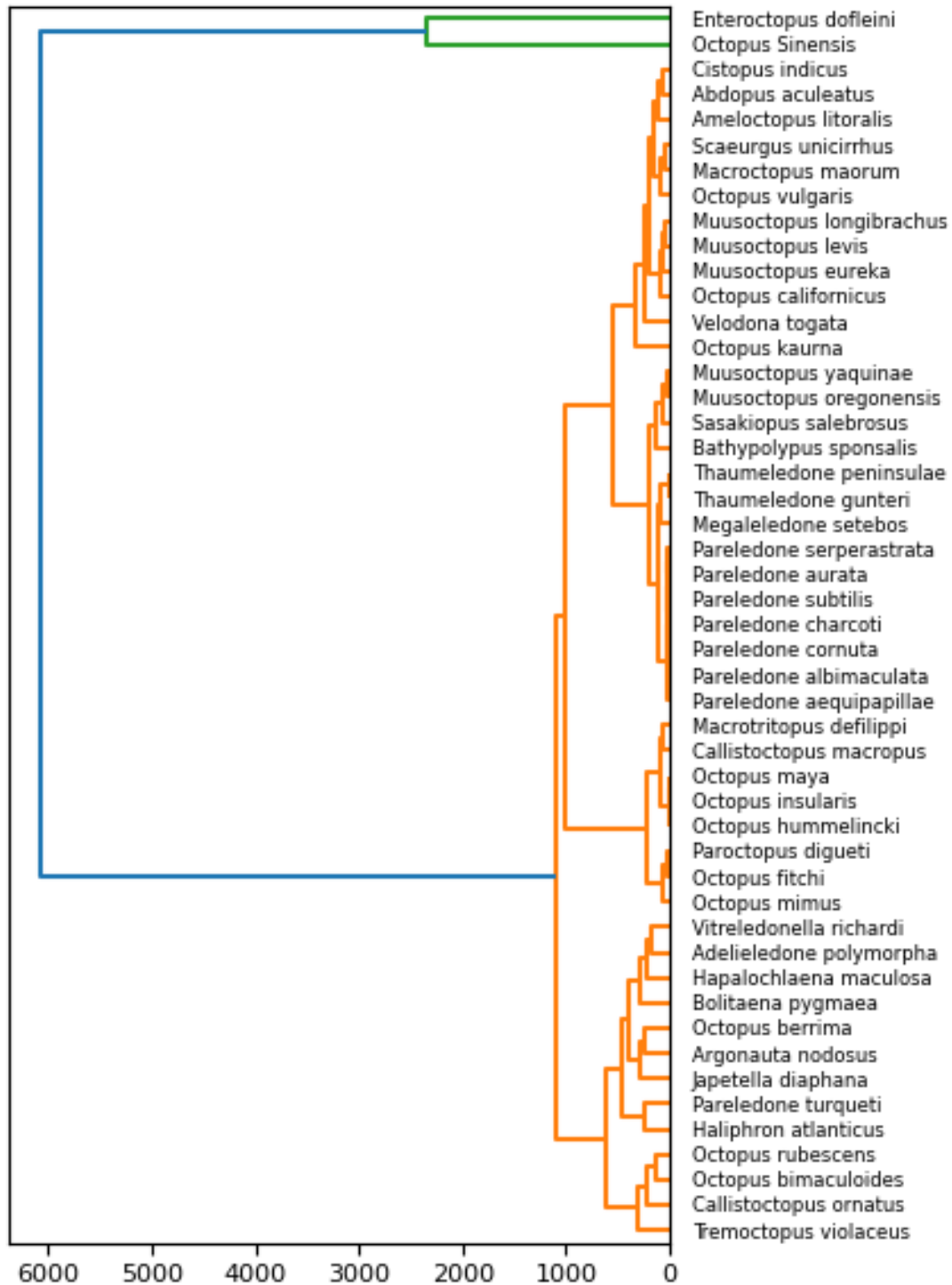


Figura 2 : Árbol Filogenético Hamming

Métrica por Contenido A-T-C-G

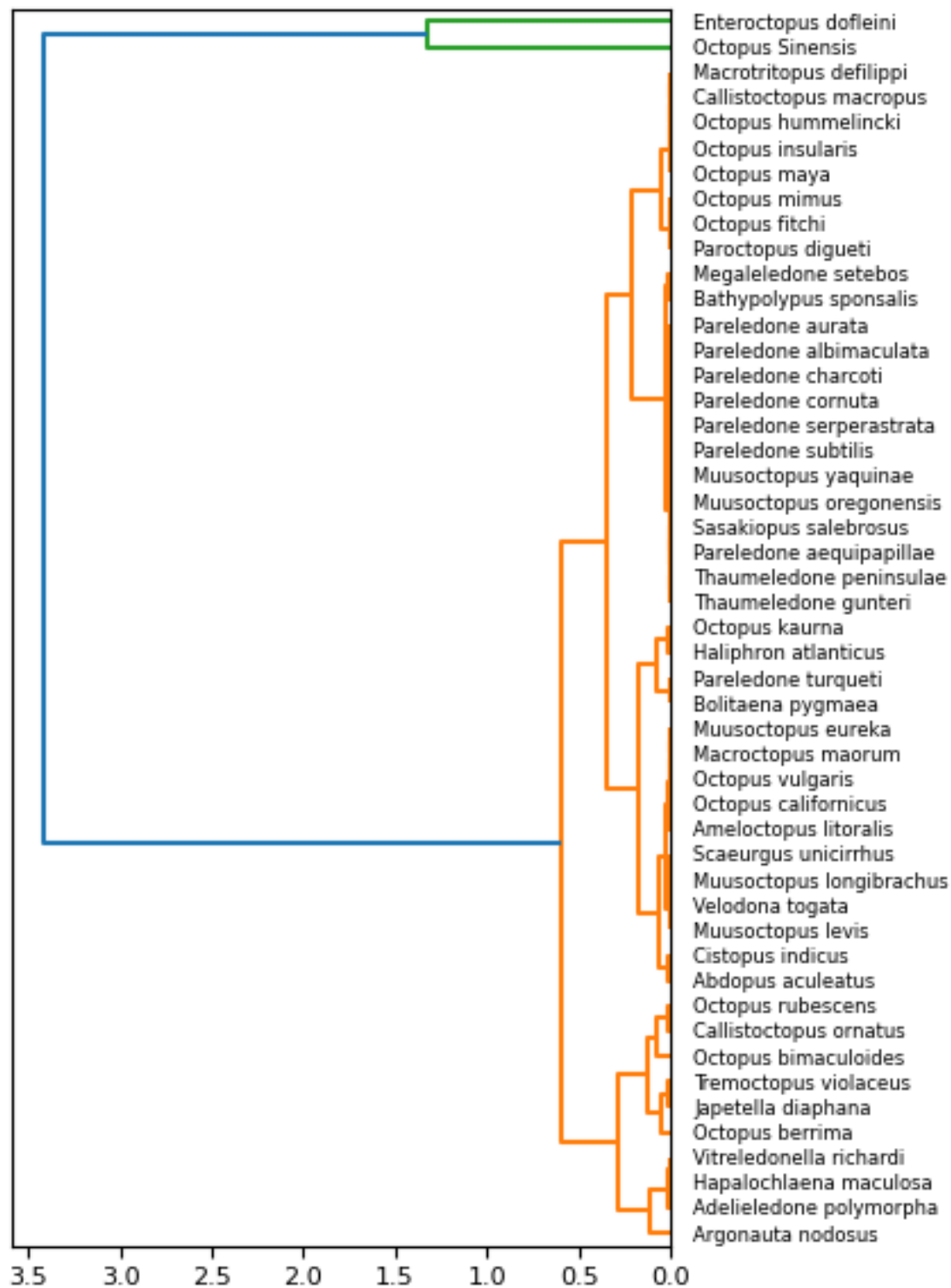


Figura 3 : Árbol Filogenético Contenido

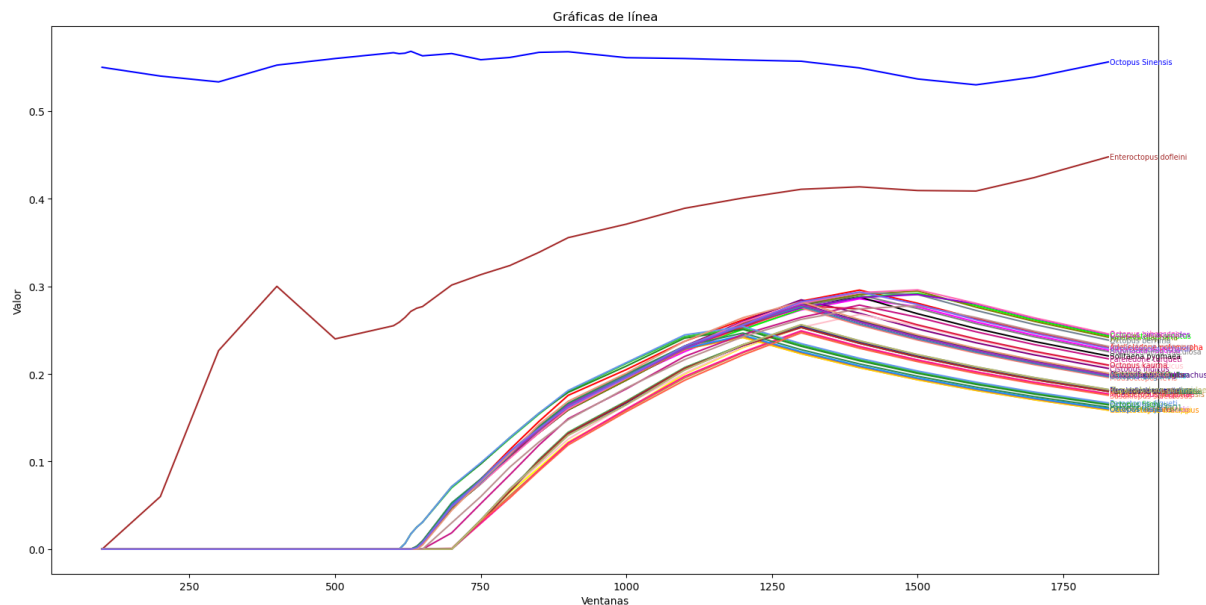


Figura 4 : Proporción de nucleótidos

Discusión :

La proporción de aminoácidos del **Enterocopus Dofleini** es menor a la del **Octopus Sinensis** y mayor que cualquiera de los otros pulpos, por lo que la distancia (en cualquiera de los árboles genéticos presentados) entre estas dos especies es menor que entre cualquier otra especie.

Aunque son cercanos considerando el ancestro común más cercano debido a alguna divergencia en su historial evolutivo (según el alcance presentado), las ramas en los árboles filogenéticos, considerando las métricas de Hamming y por Contenido, están a una distancia considerable entre estas especies.

Algunos datos encontrados en Hanke, F. D., & Kelber, A. (2019) en cuanto a localización geográfica, hábitat son :

Octopus Sinensis

Distribución Geográfica :

Aguas costeras subtropicales y templadas del noroeste del Océano Pacífico, al menos de China,

Hong Kong SAR y la provincia china de Taiwán en el sur, hasta el centro de Japón

Hábitat y Biología :

Frente a la isla de Nanji, China, el desove ocurre cuando la temperatura del agua del mar alcanza los 13°C, y la mayoría de los animales desovan cuando

las temperaturas del agua de mar alcanzan alrededor de 16°C. El número de huevos varía de 22 000 a 170 000 por hembra y el éxito de eclosión disminuye a temperaturas del agua de mar superiores a 28 °C. De la puesta de huevos a la eclosión

El período varía de 20 a 47 días y depende de la temperatura.

Enterocopus Dofleini

(El gigante del pacífico)

Distribución Geográfica :

Norte del Océano Pacífico, desde Japón hasta Baja California, México.

Hábitat y Biología :

La profundidad donde habitan ronda de los 0 a los 1500 metros.

Normalmente caza a sus presas en la noche; su alimentación consiste principalmente de algunos moluscos, crustáceos y en raras ocasiones son capaces de comer aves marinas.

Conclusiones :

Recordemos que los árboles filogenéticos son hipótesis basadas en datos específicos , en nuestro caso una recopilación fue de 50 especies. Por lo que es posible que haya un sesgo de información al resaltar a dos especies de pulpos que parecen estar apartadas del historial de evolución de la mayoría de los pulpos. Sin embargo, la experimentación arroja que esas especies sobresalen por su secuencia de la rodopsina. Investigando un poco más pudimos encontrar la siguiente información :

El **Octopus Sinensis** se distingue por tener una *visión cromática muy desarrollada, *sensibilidad a la polarización de la luz (Hanke, F. D., & Kelber, A. (2019)), la cual sirve para reconocer cambios en las vibraciones de la onda de la luz, y cambios de color y patrones (camuflaje) (*Cephalopods of the world, 2016*) .

El **Enterocopus Dofleini** se distingue por tener también una *visión cromática desarrollada, visión binocular (percepción tridimensional más precisa) y un sentido de detección de movimiento (*Cephalopods of the world, 2016*) .

Algunas características no están directamente relacionadas con la rodopsina, sin embargo consideramos importante mencionarlas para este trabajo.

Bibliografía

- *Árboles filogenéticos (artículo)*. (s/f). Khan Academy. Recuperado el 5 de junio de 2023, de <https://es.khanacademy.org/science/biology/her/tree-of-life/a/phylogenetic-trees>
- *Cómo construir un árbol filogenético*. (s/f). Khan Academy. Recuperado el 5 de junio de 2023, de <https://es.khanacademy.org/science/ap-biology/natural-selection/phylogeny/a/building-an-evolutionary-tree>
- de Grao, M. do T. de F. (s/f). *Grao en Biología*. Udc.es. Recuperado el 5 de junio de 2023, de https://ruc.udc.es/dspace/bitstream/handle/2183/29165/FernandezPineiro_Alberto_TF_G_2021.pdf?sequence=2
- Food and Agriculture Organization of the United Nations. (2016). *Cephalopods of the world: An annotated and illustrated catalogue of cephalopod species known to date: Volume 3: Octopods and vampire squids*. Food & Agriculture Organization of the United Nations (FAO).
- Hanke, F. D., & Kelber, A. (2019). The Eye of the Common Octopus (*Octopus vulgaris*). *Frontiers in Physiology*, 10, 1637. <https://doi.org/10.3389/fphys.2019.01637>
- Hickman, C. P., Keen, S. L., Larson, A. L., Eisenhour, D. J., I'Anson, H., & Roberts, L. S. (2013). *Integrated principles of zoology* (16a ed.). McGraw Hill Higher Education.
- M. Clapp. (2015). *Análisis Matemático* (1a ed.). Colección Papirhos, Serie Textos, Núm. 2, Instituto de Matemáticas de la UNAM

- Hubbard, R., & Kropf, A. (1958). The action of light on rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America*, 44(2), 130–139.
<https://doi.org/10.1073/pnas.44.2.130>
- Rodríguez, S. (s/f). *Análisis Filogenético - Parte 1 | Descripciones generales e importancia*. <https://youtu.be/h1zkTboMF8o>
- The Roslin Institute-Training. (s/f-a). 2. *Phylogenetics & Phylogeography(lecture-part 2)*. <https://youtu.be/2zPXHmqM3rk>
- The Roslin Institute-Training. (s/f-b). *Phylogenetic analysis of pathogens(lecture - part1)*. <https://www.youtube.com/watch?v=t1vAhQvukRY>
- De, E. (s/f). Topologia de espacios metricos. Ehu.eus. Recuperado el 5 de junio de 2023, de https://www.ehu.eus/~mtwmastm/TEM_Managua2009.pdf
- Métricos, E. (s/f). Universidad de Buenos Aires. Uba.ar. Recuperado el 5 de junio de 2023, de https://cms.dm.uba.ar/academico/materias/2docuat2018/calculo_avanzado/2017/apunte.pdf