

# Proyecto Final

## Genómica Computacional

Hernández González Yun

June 6, 2023

### 1 Introducción

Las plantas son uno de los dos mayores grupos de organismos vivos en la tierra, ellas son parte fundamental de la biosfera. Las podemos encontrar en cada rincón de planeta, en toda clase de formas y tamaños. (Fernando, 2012). Desde las expediciones de Carlos Linneo en el siglo XX se ha buscado clasificar las plantas mediante la forma de sus hojas, características físicas, hasta métodos moleculares. Un gen que comparten muchas plantas es ribulosa-1, 5-bisfosfato carboxilasa/oxigenasa (Rubisco), ubicada en los cloroplastos, es una enzima clave para la generación de alimentos de las plantas, esta es la enzima más abundante del planeta. Más del 90% de carbono inorgánico convertido en biomasa es transformado por la enzima RubisCO (Tobias, 2018), mediante Clavin-Benson.(Andersson and Backlund, 2008) Se ha encontrado procesos adaptativos en este gen, sin embargo, aún quedan dudas sobre cual es su significado biológico ante su entorno. Su representación 3D (J. and R., 2021) es:

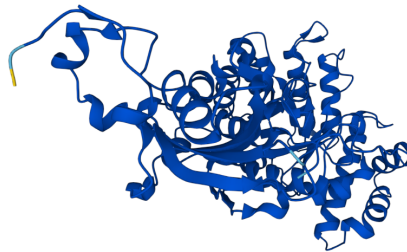


Figure 1: Imagen 3D de rbcl.

### 2 Pregunta de Investigación

¿El gen rbcl es un factor importante en la filogenia?

### 3 Objetivo

El objetivo es encontrar alguna relación en *rbcl* de 15 plantas de 3 familias distintas, tanto en filogenética como ante los factores ambientales. Al mismo tiempo *rbcl* nos podría proporcionar alguna relación sobre sus procesos adaptativos, como la adaptación ante el cambio climático y mayores concentraciones de  $CO_2$  en la atmósfera.

### 4 Métodos.

Para poder responder la pregunta estudiaremos a *rbcl* en 15 diferentes plantas de 3 familias que lo contienen. Las familias que fueron escogidas para el análisis son:

Aizoaceae

- *Sesuvium sesuvioides* - Namibia
- *Mesembryanthemum cordifolium*- Nativa de Provincia del Cabo Oriental de Sudáfrica.
- *Mesembryanthemum crystallinum* - Africa, Sinai y Europa del sur.
- *Tetragonia tetragonioides* - Nueva Zelanda, Australia, Japón, Chile y Argentina.
- *Sesuvium portulacastrum* - África, Asia, Australia, América del Norte y América del Sur.

Cactaceae



- *Lophophora diffusa* - México (Querétaro e Hidalgo).
- *Carnegiea gigantea* - México (Desierto de Sonora).
- *Pereskia aculeata* - América del sur, centro América e islas del Caribe.
- *Tephrocactus geometricus* - Argentina(Catamarca)
- *Stenocereus thurberi* - México( Sonora, Sinaloa, Baja California).

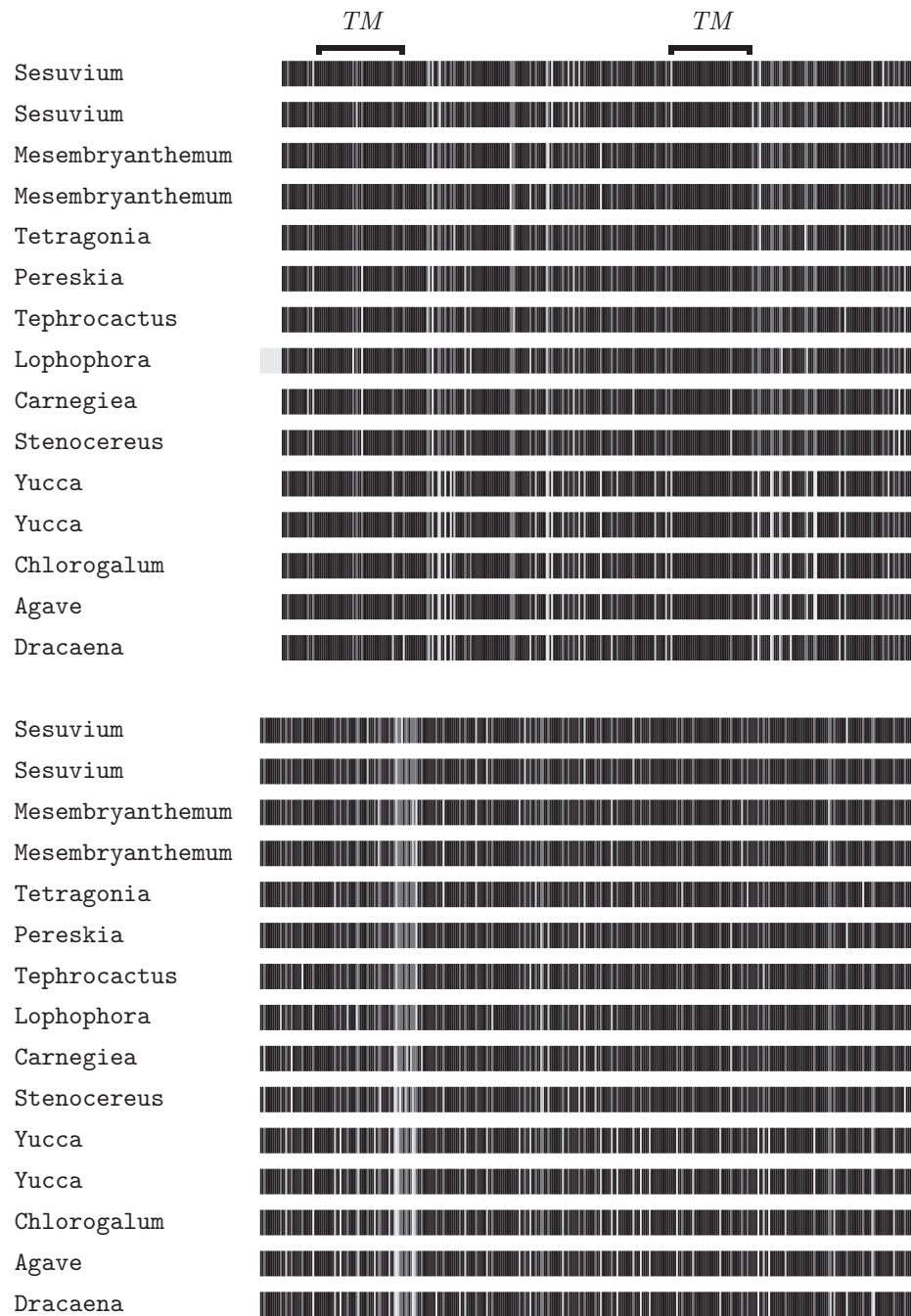
Asparagaceae
















- *Agave angustifolia* - México y centro América.
- *Yucca brevifolia* - Sur oeste de EU(desierto Mojave)
- *Yucca treculeana* - Texas, Nuevo México, Coahuila y Nuevo León.
- *Chlorogalum pomeridianum* - EU(California)
- *Dracaena sanderiana* - Africa (Camerún)
















## 5 Resultados

	CTAAC	TATAAAC	CCAG	GATAC	TGATATCTTGG	
Sesuvium	CTAAC	TATAAAC	CCAG	GATAC	TGATATCTTGG	112
Sesuvium	CTGACTATAAAC	CCAG	GATAC	TGATATCTTGG		112
Mesembryanthemum	CTGACTATCAAC	CCAG	GATAC	TGATATCTTGG		112
Mesembryanthemum	CTGACTATCAAC	CCAG	GATAC	TGATATCTTGG		112
Tetragonia	CTGACTATCAAC	CCAG	GATAC	TGATATCTTGG		112
Pereskia	CTCAATATCAAC	CCAG	GATAC	TGATATCTTGG		112
Tephrocactus	CCGAATATCAAC	CCAG	GATAC	TGATATCTTGG		112
Lophophora	CTGAATATCAAC	CCAG	GATAC	TGATATCTTGG		124
Carnegiea	CTGAATATCAAC	CCAG	GATAC	TGATATCTTGG		112
Stenocereus	CTGAATATCAAC	CCAG	GATAC	TGATATCTTGG		112
Yucca	CTGATTACGAAA	CCAAA	GATAC	TGATATCTTGG		112
Yucca	CTGATTACGAAA	CCAAA	GATAC	TGATATCTTGG		112
Chlorogalum	CTGATTACGAAA	CCAAA	GATAC	TGATATCTTGG		112
Agave	CTGATTACGAAA	CCAAA	GATAC	TGATATCTTGG		112
Dracaena	CTAATTACGAAA	CCAAA	GATAC	TGATATCTTGG		112
consensus	Ct	gA.TA..AA..CC..A..	GATAC..	GATATCTTGG		

 non-conserved  
  $\geq 80\%$  conserved  
 all match



Sesuvium	
Sesuvium	
Mesembryanthemum	
Mesembryanthemum	
Tetragonia	
Pereskia	
Tephrocactus	
Lophophora	
Carnegiea	
Stenocereus	
Yucca	
Yucca	
Chlorogalum	
Agave	
Dracaena	

Sesuvium	
Sesuvium	
Mesembryanthemum	
Mesembryanthemum	
Tetragonia	
Pereskia	
Tephrocactus	
Lophophora	
Carnegiea	
Stenocereus	
Yucca	
Yucca	
Chlorogalum	
Agave	
Dracaena	

Sesuvium	
Sesuvium	
Mesembryanthemum	
Mesembryanthemum	
Tetragonia	
Pereskia	
Tephrocactus	
Lophophora	
Carnegiea	
Stenocereus	
Yucca	
Yucca	
Chlorogalum	
Agave	
Dracaena	

	non-conserved
	similar
	$\geq 50\%$ conserved
	all match

Podemos resaltar que existen secciones que son son altamente conservadas.

Al generar la matriz de distancias, y construir los árboles con NJ, UPGMA y WPGMA es el siguiente:

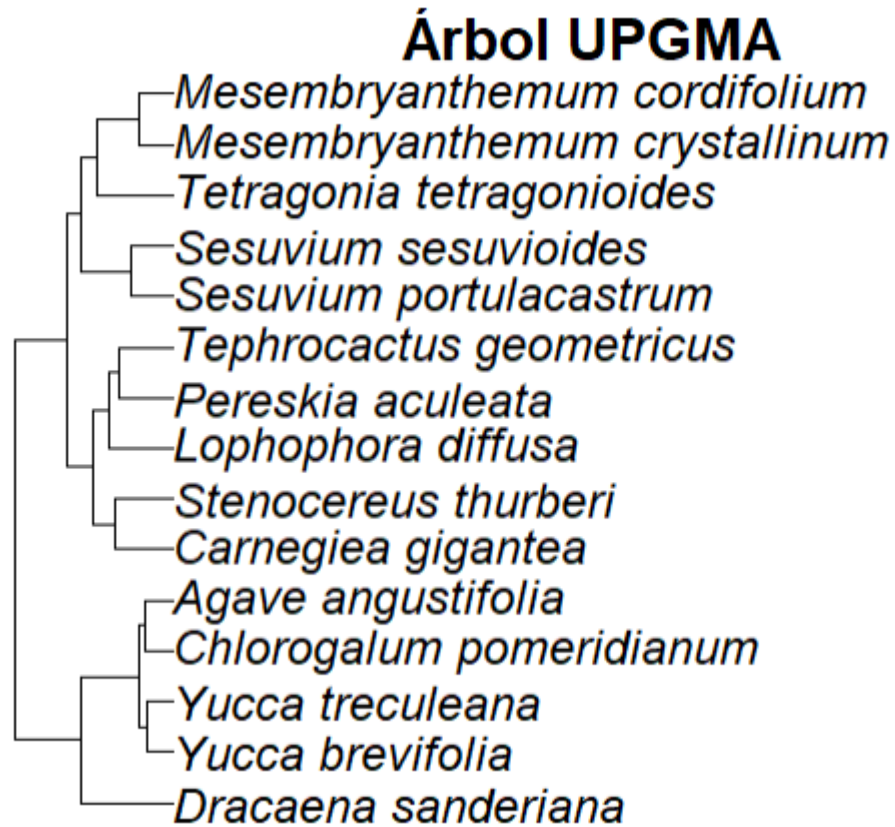


Figure 2: Árbol con UPGMA.

## Árbol WPGMA

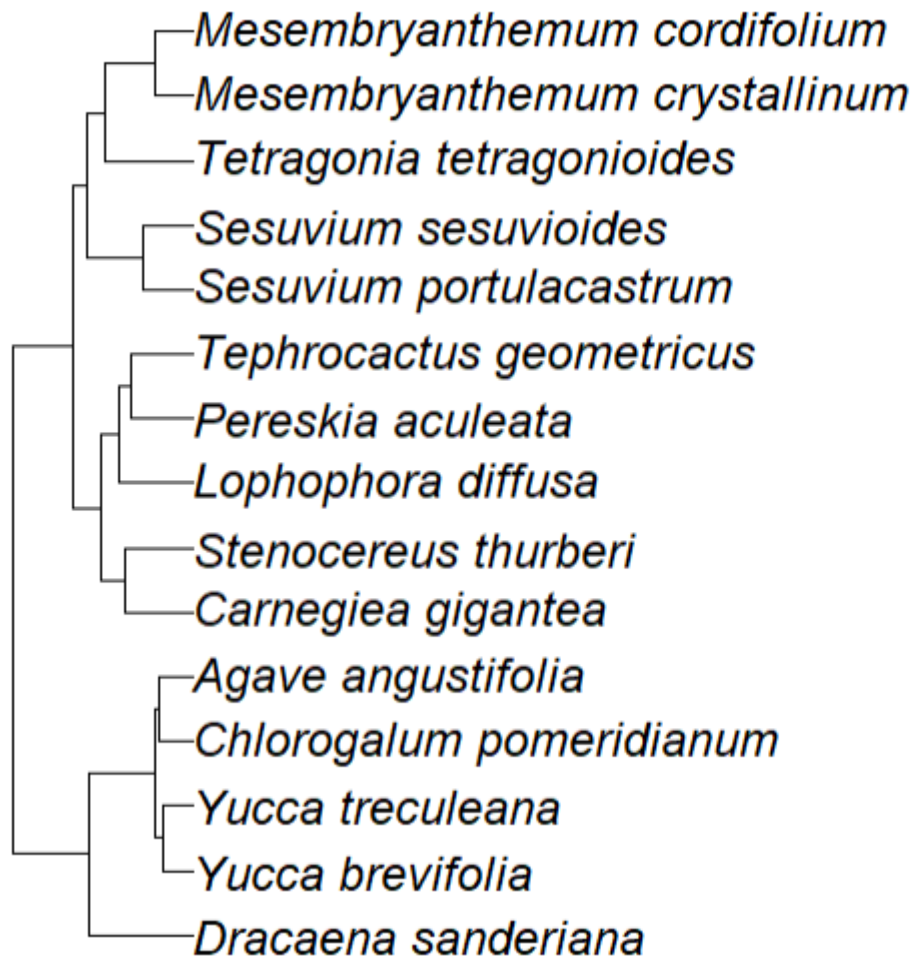


Figure 3: Árbol con WPGMA.



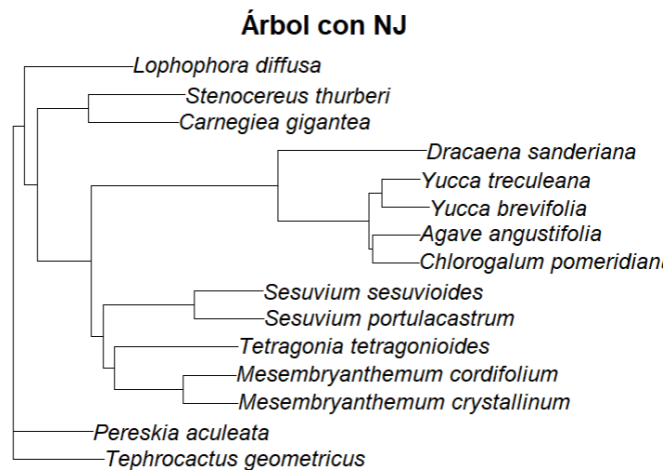


Figure 4: Árbol con NJ

## 6 Discusión

Tras obtener los resultados, es evidente que el gen *rbcl* en las 15 es bastante similar entre sí, esto sugiere que comparten un ancestro en común no muy lejano. Sin embargo, los cambios que este presenta son suficientes para poder clasificar a las plantas de acuerdo a su familia, indicando que este se ha adaptado dependiendo las condiciones ambientales. Esto puede significar que el gen tiende a tener una baja tasa de mutación.

## 7 Conclusión

En conclusión, el estudio de *rbcl* y su papel en la fotosíntesis y el metabolismo de las plantas muy importante en la comprensión de la biología vegetal. De acuerdo con los resultados obtenidos en los árboles filogenéticos, fue posible obtener la filogenia real de las familias estudiadas, mediante técnicas como ClustalW, matrices de distancias y árboles filogenéticos.

## References

- Andersson, I., & Backlund, A. (2008). Structure and function of rubisco [Plant structural biology]. *Plant Physiology and Biochemistry*, 46(3), 275–291. <https://doi.org/https://doi.org/10.1016/j.plaphy.2008.01.001>
- Bodenhofer, U., Bonatesta, E., C., H.-K., & Hochreiter, S. (2015). Msa: An r package for multiple sequence alignment. *Bioinformatics*, 31(24), 3997–3999. <https://doi.org/10.1093/bioinformatics/btv494>

- Fernando, W. G. D. (2012). Plants: An international scientific open access journal to publish all facets of plants, their functions and interactions with the environment and other living organisms. *Plants*, 1(1), 1–5. <https://doi.org/10.3390/plants1010001>
- J., J., & R., E. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, 583–589. <https://doi.org/https://doi.org/10.1038/s41586-021-03819-2>
- Stavrinides, J., & Ochman, H. (2009). Phylogenetic methods en encyclopedia of microbiology. *Editorial Elsevier*, 249–251.
- Tobias, J. (2018). A short history of rubisco: The rise and fall (?) of nature’s predominant co2 fixing enzyme. *Current Opinion in Biotechnology*, 49(Number), 100–107. <https://doi.org/https://doi.org/10.1016/j.copbio.2017.07.017>