

**Centro de Investigación Científica y de Educación
Superior de Ensenada, Baja California**



**Maestría en Ciencias
Ciencias de la Vida
con orientación en Biología Ambiental**

**Evolución molecular del sueño unihemisférico en Cetáceos y
Pinnípedos: una aproximación genómica comparativa**

Tesis para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

Sebastián Álvarez Costes

Ensenada, Baja California, México
2021

Tesis defendida por
Sebastián Álvarez Costes

y aprobada por el siguiente Comité

Dra. María Clara Arteaga Uribe
Directora de tesis

Miembros del comité

Dra. Fadia Sara Ceccarelli

Dr. Axayácatl Rocha Olivares

Dra. Maribel Hernández Rosales



Dra. Patricia Juárez Camacho
Coordinadora del Posgrado en Ciencias de la Vida

Dr. Pedro Negrete Regagnon
Director de Estudios de Posgrado

Resumen de la tesis que presenta **Sebastián Álvarez Costes** como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Vida con orientación en Biología Ambiental

Evolución molecular del sueño unihemisférico en Cetáceos y Pinnípedos: una aproximación genómica comparativa

Resumen aprobado por:

Dra. María Clara Arteaga Uribe
Directora de tesis

El sueño unihemisférico de onda corta es muy inusual en vertebrados, y se caracteriza por la presencia de ciclos NREM y ausencia de REM. Este sueño solo está presente de manera obligada en Cetáceos y de manera facultativa en algunos Pinnípedos y otros grupos de organismos. Sobre el sueño unihemisférico se conoce el funcionamiento fisiológico y las diferencias con el sueño bihemisférico, sin embargo, existen interrogantes acerca de su origen, evolución y base genética. El objetivo de este trabajo fue identificar los genes reguladores del sueño unihemisférico en Cetáceos y Pinnípedos mediante genómica comparativa. Para ello se identificaron los genes ortólogos relacionados con procesos del sueño compartidos entre Cetáceos y Pinnípedos, se llevó a cabo una reconciliación del árbol de genes relacionados con el sueño y el árbol filogenéticos de mamíferos lo que permitió identificar los genes relacionados al sueño unihemisférico. Además, se evaluó la tasa de selección de esos genes en ramas clave del árbol evolutivo y se infirió su estado ancestral distintos puntos de la historia evolutiva. La reconciliación de árboles de genes con el árbol de especies demostró que entre los posibles genes reguladores del sueño unihemisférico están CASP1 en Cetáceos y PER3 en Cetáceos y Otáridos. CASP1 es importante por su función en la ruta metabólica de los ciclos NREM de Cetáceos, que requieren de una mayor influencia de este gen para regular sus ciclos. Para PER3 se demostró que existe un patrón de selección que beneficia a los homocigotos en Pinnípedos, es posible que la selección que actúa sobre este gen les haya conferido la adaptación a los primeros Pinnípedos y a los Otáridos para poder mantenerse en sueño NREM. Las reconstrucciones del estado ancestral de estos genes demostraron que existen sustituciones de aminoácidos particulares en CASP1 de Cetáceos desde su último ancestro común y sustituciones de aminoácidos particulares en PER3 de Pinnípedos y una pérdida de 7 aminoácidos en Phócidos. Estos resultados sugieren que PER3 es un gen que influencia el sueño unihemisférico en Cetáceos y en Otáridos y que CASP1 es el posible responsable de que los Cetáceos presenten el sueño unihemisférico de manera obligada. Futuros estudios enfocados en estas proteínas permitirán evidenciar las características estructurales que les confieren estas similitudes y diferencias en el sueño unihemisférico, así como las diferentes ventajas adaptativas que estas proteínas les pudieran dar y si existe una convergencia evolutiva molecular.

Palabras clave: Sueño unihemisférico, sueño NREM, genómica comparativa, Cetáceos, Pinnípedos, CASP1, PER3.

Abstract of the thesis presented by **Sebastián Álvarez Costes** as a partial requirement to obtain the Master of Science degree in Life Sciences with orientation in Environmental Biology

Molecular evolution of unihemispheric sleep in Cetaceans and Pinnipeds: a comparative genomics approach

Abstract approved by:

Dra. María Clara Arteaga Uribe
Thesis director

Short-wave unihemispheric sleep is very unusual in vertebrates, characterized by the presence of NREM cycles and the absence of REM. This type of sleep is obligatory in Cetaceans and facultative in some Pinnipeds and other groups of organisms. Regarding unihemispheric sleep in these species, the physiological functioning and differences with bihemispheric sleep are known, however, there are questions about its origin, evolution and genetic basis. The objective of this study was to identify unihemispheric sleep regulatory genes in Cetaceans and Pinnipeds using comparative genomics. The orthologous genes related to sleep processes and shared between these two groups of organisms were identified, a reconciliation of the tree of genes related to sleep and species tree of Cetaceans and Pinnipeds was carried out to identify those related to unihemispheric sleep. The selection rate of these genes in key branches of the evolutionary tree was evaluated and their ancestral sequence was inferred at different points in evolutionary history. The reconciliation of gene trees with the species tree showed that the possible unihemispheric sleep regulatory genes are CASP1 in Cetaceans and PER3 in Cetaceans and Otharids. CASP1 was only observed in Cetaceans and due to its function in the metabolic pathway of the NREM cycles, these species require a greater influence of this gene to regulate these cycles. For PER3 there is a selection pattern that benefits homozygosity in Pinnipeds. Therefore, it is possible that the selection that acts on this gene has conferred adaptation to the first Pinnipeds and Otharids to be able to stay in NREM sleep. The ancestral gene reconstruction of CASP1 revealed derived amino acid substitutions in Cetaceans inherited from their last common ancestor. Likewise, derived amino acid substitutions were identified in Pinniped PER3 and 7 aminoacid deletions in Phocids. These results suggest that PER3 is a gene that influences unihemispheric sleep in Cetaceans and Otharids and that CASP1 is possibly responsible for the fact that Cetaceans present obligatory unihemispheric sleep. Future studies focused on these proteins could allow to demonstrate the structural characteristics that these similarities and differences confer on them in unihemispheric sleep, as well as the different adaptive advantages that these proteins could give them and if there is a molecular evolutionary convergence.

Keywords: Unihemispheric sleep, NREM sleep, comparative genomics, Cetaceans, Pinnipeds, CASP1, PER3.

Agradecimientos

La realización de esta tesis comprende un largo proceso de aprendizaje académico, profesional y personal que me ha permitido ampliar mi panorama y confirmar mis intereses y planes futuros; es un trabajo que involucra a diferentes instituciones, así como a diferentes personas con las cuales estoy infinitamente agradecido.

Empezando por mis padres Juan Ramón Álvarez Cuspinera y Denise Costes Intriago, quienes con gran esfuerzo me han apoyado moral y económicamente en este largo, pero placentero trayecto, porque creyeron y depositaron su confianza en mí para perseguir mi sueño, y poder lograr mis objetivos. A ellos y al resto de mi familia, dedico este trabajo, resultado de mi esfuerzo y de su incondicional apoyo.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el apoyo brindado para la realización de este posgrado por medio de la beca al CVU número 994882.

Al Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), al Posgrado en Ciencias de la Vida y al Departamento de Biología de la Conservación por el apoyo en mi formación como persona e investigador. Incluidos todos los profesores que me impartieron clases, por su apoyo fundamental en ampliar mi conocimiento sobre distintos temas.

A mi directora de tesis, la Dra. María Clara Arteaga Uribe, por la confianza proporcionada desde el comienzo, por motivarme a continuar creciendo profesional y personalmente. Gracias María por aceptarme como alumno y por aceptar colaborar en este proyecto, en el cual hemos aprendido mucho los dos.

A los miembros de mi comité de tesis, la Dra. Fadia Sara Ceccarelli, el Dr. Axayácatl Rocha Olivares y la Dra. Maribel Hernández Rosales, porque gracias a su experiencia en el área, y a sus valiosas aportaciones y comentarios, han ayudado a lograr este trabajo que tiene gran potencial para la investigación.

Agradezco particularmente a la Dra. Maribel Hernández Rosales y a su alumno José Antonio Ramírez Rafael, miembros del Laboratorio de Bioinformática y Redes Complejas del CINVESTAV. Gracias por sus valiosas enseñanzas en técnicas bioinformáticas, por su apoyo en la realización de los análisis y por sus aportaciones y recomendaciones para la interpretación de resultados.

Tabla de contenido

Resumen en español.....	ii
Resumen en inglés.....	iii
Agradecimientos.....	iv
Lista de figuras.....	vii
Lista de tablas.....	viii
Glosario.....	ix
Capítulo 1. Introducción.....	1
1.2 Justificación.....	6
1.3 Hipótesis.....	6
1.4 Objetivos.....	7
1.4.1 Objetivo General.....	7
1.4.2 Objetivos Específicos.....	7
Capítulo 2. Metodología.....	8
2.1 Obtención de datos genómicos.....	8
2.2 Identificación de ortólogos basada en inferencia por grafos.....	9
2.3 Filtrado e identificación de genes relacionados al sueño.....	10
2.4 Reconciliación del árbol de genes con el árbol de especies.....	10
2.4.1 Árbol filogenético.....	10
2.4.2 Reconciliación.....	11
2.4.3 Árboles de genes individuales.....	12
2.5 Análisis filogenéticos y selección dN/dS.....	12
2.6 Inferencia del estado ancestral.....	14
Capítulo 3. Resultados.....	16
3.1 Obtención de datos genómicos.....	16
3.2 Identificación de genes ortólogos.....	17
3.3 Filtrado e identificación de genes relacionados al sueño.....	17

3.4 Reconciliación del árbol de genes con el árbol de especies	19
3.4.1 Árbol filogenético base	19
3.4.2 Árboles de genes individuales.....	22
3.5 Análisis de selección dN/dS para CASP1 y PER3	27
3.6 Inferencia del estado ancestral para los genes candidatos CASP1 y PER3.	28
Capítulo 4. Discusión.....	34
4.1 Diferencias en número de genes relacionados al sueño entre Cetáceos y Pinnípedos.....	34
4.2 Genes relacionados al sueño unihemisférico	37
4.3 Genes relacionados al sueño bihemisférico	40
Capítulo 5. Conclusiones y perspectivas	42
Literatura citada.....	43

Lista de figuras

Figura 1. Árbol filogenético por máxima verosimilitud de la filogenia de mamíferos utilizando 13 genes (9 nucleares y 5 mitocondriales) y calibrado con el registro fósil de Cetáceos y Pinnípedos.	20
Figura 2. Reconciliación del árbol de genes relacionados al sueño con el árbol de especies.	21
Figura 3. Agrupamiento de genes entre las distintas especies, donde cada columna representa un gen y cada círculo negro representa en qué especies se encuentra ese gen.	22
Figura 4. a) Árbol del gen ADORA2A. b) Reconciliación del árbol de la familia del gen ADORA2A con el árbol de especies.	23
Figura 5. a) Reconciliación del árbol de genes de la familia PTGER4 con el árbol de especies. b) Reconciliación del árbol de genes de la familia CST3 con el árbol de especies.	24
Figura 6. a) Árbol del gen CASP1. b) Reconciliación del árbol de la familia de genes relacionados a CASP1 con el árbol de especies.	25
Figura 7. a) Árbol de reconciliación del gen PER3. b) Reconciliación del árbol de genes de la familia PER3 con el árbol de especies.	26
Figura 8. a) Árbol del gen PER3. b) Red de ortólogos relacionados a PER3.	27
Figura 9. a) Árbol filogenético por máxima verosimilitud de las secuencias de aminoácidos del gen CASP1. b) Posición de las sustituciones y eliminaciones de aminoácidos dentro de la secuencia de la proteína CASP1 de las especies actuales y de los ancestros comunes.	30
Figura 10. a) Árbol filogenético por máxima verosimilitud de las secuencias de aminoácidos del gen PER3. b) Posición de las sustituciones y eliminaciones de aminoácidos dentro de la secuencia de la proteína PER3 de las especies actuales y de los ancestros comunes.	33

Lista de tablas

Tabla 1. Lista de especies con los respectivos números de acceso a sus genomas en el NCBI.	8
Tabla 2. Número de secuencias utilizadas por archivo fasta para el análisis de Proteinortho.	16
Tabla 3. Estadísticas del análisis de genes ortólogos utilizando la herramienta revolutionhl.	17
Tabla 4. Genes relacionados al sueño con relaciones de homología entre las 18 especies y su función....	18
Tabla 5. Análisis de selección por grupo para los genes CASP1 y PER3.	28
Tabla 6. Modelos de sustitución nucleotídica y de aminoácidos utilizados en el análisis de máxima verosimilitud para la reconstrucción ancestral de las secuencias de nucleótidos de los genes CASP1 y PER3.	28
Tabla 7. Estadísticas de la reconstrucción del estado ancestral del gen CASP1.	29
Tabla 8. Estadísticas de la reconstrucción del estado ancestral del gen PER3.	31

Glosario

Convergencia evolutiva molecular: Cuando un gen que determina un rasgo o característica, evoluciona de manera independiente en dos o más linajes (Zakon 2002).

Evolución convergente: Desarrollo independiente de un rasgo similar en taxones filogenéticamente distantes (Doolittle 1994; Pickersgill 2018).

Evolución paralela: Desarrollo independiente de un rasgo similar en taxones con una ascendencia común relativamente reciente (Pickersgill 2018).

Ganancia de genes: Proceso mediante el cual, por diferentes causas, una especie incorpora genes nuevos en su genoma (Lynch y Conery 2003).

Genes homólogos: Genes con una ascendencia evolutiva común (Kawashima 2019).

Genes ortólogos: Genes homólogos que divergieron como resultado de un evento de especiación (Kawashima 2019).

Genes parálogos: Genes homólogos que divergieron como resultado de un evento de duplicación de genes, usualmente con una subsecuente divergencia en función (Kawashima 2019).

Genes out-parálogos: Genes parálogos que fueron duplicados antes del evento de especiación. (Kawashima 2019)

Reconstrucción de genes ancestrales: Secuencia de un gen en un ancestro común inferida a partir de información de ese mismo gen en especies actuales (Merkl y Sterner 2016).

Ontología: Descripción de un gen y los atributos del producto génico en cualquier organismo (Ashburner *et al.* 2000).

Pérdida de genes: Proceso mediante el cual, por diferentes causas, una especie pierde genes en su genoma (Lynch y Conery 2003).

Selección purificadora: Tipo de selección natural que ocurre cuando hay presiones selectivas que actúan contra dos extremos de un rasgo y, por lo tanto, el rasgo intermedio se selecciona beneficiando a los heterocigotos (Futuyma 2005).

Selección diversificadora: Tipo de selección natural que ocurre cuando las presiones selectivas actúan a favor de los dos extremos, beneficiando a los homocigotos y en contra del rasgo intermedio (Futuyma 2005).

Selección direccional: Tipo de selección natural que ocurre cuando las presiones selectivas actúan a favor de un extremo de un rasgo (Futuyma 2005).

Capítulo 1. Introducción

El sueño es un estado de quiescencia obligatorio para la restauración y mantenimiento del equilibrio homeostático del cuerpo, que se caracteriza por una reducida respuesta a estímulos ambientales. La organización del sueño está regulada principalmente por los ciclos circadianos y por la presión homeostática del sueño. Los ciclos circadianos sincronizan y determinan el momento del día en el cuál llevar a cabo el proceso de sueño, mientras que la presión homeostática del sueño es responsable de mantener el balance de sueño diario (Eban-Rothschild *et al.* 2017).

En términos generales, el sueño está dividido en dos etapas en las que se llevan a cabo procesos cognitivos y fisiológicos tales como el procesamiento de la memoria, la termorregulación y la asignación de energía: la fase de movimiento no rápido de ojos NREM (del inglés *Non-Rapid Eye Movement*) y la fase de movimiento rápido de ojos REM (del inglés *Rapid Eye Movement*). A pesar de que el sueño y sus etapas son procesos conservados en vertebrados, existen variaciones en el tipo, duración y frecuencia de las fases (Schmidt 2014; Aulsebrook *et al.* 2016).

El sueño más común en los vertebrados es el sueño bihemisférico, que está regulado por ciclos circadianos y se caracteriza por hipotonía muscular, presencia de ciclos REM y el cierre de ambos ojos, reflejo del estado de sueño en ambos hemisferios (Mascetti 2016). En este tipo de sueño, hay cambios periódicos entre los dos distintos estados antes mencionados: REM y NREM, los cuales se diferencian por los distintos tipos de actividad cerebral que presentan. En la fase NREM hay oscilaciones globales de actividad cerebral de onda corta, evidencia de que es la etapa de sueño más profunda, cuya principal función es la consolidación de la memoria y el descanso. La fase REM, que inicia al terminar la NREM, se caracteriza por ser de sueño ligero, movimientos rápidos de ojos y ondas cerebrales más largas. Estos episodios son en general más cortos que los NREM, y se ha hipotetizado que su función es compensar por los procesos que se están llevando a cabo en el NREM. Durante el NREM el organismo entra en un estado de nula actividad, entonces, el REM compensa ese estado aumentando la actividad cerebral para mantener las funciones del cuerpo como la termorregulación y mantener al organismo alerta a cualquier estímulo del ambiente durante los ciclos de sueño. Además, también se ha hipotetizado que estos ciclos NREM-REM son fundamentales para la plasticidad y excitabilidad neuronal y procesamiento de información (Vyazovskiy y Delogu 2014; Aulsebrook *et al.* 2016).

El sueño unihemisférico de onda corta, es un tipo de sueño inusual en vertebrados, en el que solo se induce el sueño en un hemisferio a la vez, mientras el otro hemisferio permanece despierto y alerta, alternando

entre un hemisferio y otro en los diferentes ciclos de sueño (Mascetti 2016). Este tipo de sueño se caracteriza por la ausencia de la etapa REM, habiendo la posibilidad de que haya desaparecido completamente o que exista en una forma diferente, aun no identificada, a la conocida para el sueño bihemisférico (Mascetti 2016). Este sueño se presenta en mamíferos acuáticos tales como los miembros de la familia Otariidae del orden Pinnipedia, las especies del orden Cetacea y también en algunas aves (Rattenborg *et al.* 2016). Entre los organismos en los que se presenta este tipo de sueño, los Cetáceos son los únicos en poseerlo exclusivamente; mientras que en los Otáridos es un tipo de sueño facultativo, ya que alternan entre sueño bihemisférico y unihemisférico dependiendo de si duermen en tierra o agua (Lyamin *et al.* 2002, 2018; Kedziora *et al.* 2012; Mascetti 2016; Kendall-Bar *et al.* 2019). Cabe destacar que, dentro del grupo de los Pinnípedos, los miembros de la familia Phocidae, solo presentan sueño bihemisférico mientras que en los miembros de la familia Odobenidae aún no se ha identificado su tipo de sueño (Lyamin *et al.* 2008, 2018). En cuanto a los Sirenios, mamíferos que también son completamente acuáticos, se ha encontrado evidencia de que estas especies presentan sueño de onda corta y de onda larga, con ciertas similitudes con el sueño de Cetáceos y Pinnípedos, pero con presencia de ondas largas que no se observan en otros mamíferos acuáticos, por lo que se concluye que los Sirenios deben de tener una estrategia diferente a la de Cetáceos y Otáridos para llevar a cabo esta función fisiológica en el ambiente acuático (Mukhametov *et al.* 1992).

En mamíferos acuáticos (Cetáceos, Otáridos) el sueño unihemisférico les permite, además de llevar a cabo el proceso de descanso y recuperación, mantener activo el cuerpo para poder mantener la termorregulación en el agua, la respiración y el estado de vigilancia (Mascetti 2016). Además, el sueño unihemisférico en comparación con el bihemisférico, tiene como consecuencia una reducción en el tiempo de descanso y de recuperación, pero a pesar de esto, no existen consecuencias de esta reducción en el comportamiento o salud de las especies que lo presentan (Mascetti 2016).

Durante la historia evolutiva de los organismos que tienen este tipo de sueño, distintas presiones selectivas los han llevado a presentarlo de manera obligada o facultativa. Los Cetáceos, único grupo con sueño unihemisférico obligado, no presentan sueño tipo REM y durante esta etapa del sueño, la termorregulación disminuye drásticamente. Se hipotetiza que el sueño unihemisférico es una adaptación a la termorregulación en el ambiente acuático, provocando que los Cetáceos lo presenten de manera obligada y que los Otáridos lo presentan únicamente al dormir en el agua. Tener un hemisferio despierto manteniendo las funciones normales del cuerpo, permite a los organismos mantener la termogénesis muscular constante (Lyamin *et al.* 2008; Aulsebrook *et al.* 2016; Mascetti 2016). Por otra parte, los Cetáceos, al ser completamente acuáticos, están obligados a mantener un ritmo respiratorio y de nado

constante, manteniendo la posición en la columna de agua, con el objetivo de mantenerse cerca de la superficie y no derivar durante un sueño profundo. Además, al ser organismos muy sociales, este tipo de sueño les permite mantener la estructura del grupo en situaciones donde no hay contacto auditivo y estar alerta ante posibles amenazas (Lyamin *et al.* 2008; Mascetti 2016).

Distintos estudios donde se han llevado a cabo observaciones y comparaciones de la fenomenología entre el sueño bihemisférico de mamíferos y el unihemisférico de Cetáceos, han llegado a tres posibles conclusiones acerca de la evolución de este último tipo de sueño en este grupo: La primera hipótesis explica que este tipo de sueño presente en las especies actuales evolucionó como una característica nueva a partir de su último ancestro común. La segunda, que alguno de los antepasados del último ancestro común de los Cetáceos pudo haber presentado algún tipo de sueño que dio lugar a lo observado en las especies actuales. Por último, la tercera hipótesis plantea que este tipo de sueño y su fenomenología evolucionó convergentemente en los distintos linajes de Cetáceos (Lyamin *et al.* 2008).

Mediante análisis anatómicos y fisiológicos, se ha determinado la presión de selección del ambiente que ha podido modificar los procesos termorregulatorios de Cetáceos y Otáridos en el ambiente acuático. Esta presión de selección se refiere al incremento en la pérdida de calor relacionado con la invasión desde el medio terrestre al medio acuático, que conllevó a presentar sueño unihemisférico de onda corta con ausencia de etapa REM. El resultado fisiológico de adquirir este tipo de sueño es el incremento de la actividad muscular, el tono muscular por movimiento constante y la producción de noradrenalina. Estos tres cambios fisiológicos dan como resultado producción de calor, fundamental para el mantenimiento de la temperatura del cuerpo en este ambiente (Lyamin *et al.* 2008).

Es importante determinar qué tipo de selección y cómo ha afectado a los genes relacionados con el sueño unihemisférico en ramas clave de los árboles filogenéticos de Cetáceos y Pinnípedos, influenciando la diversidad de estos genes en estos grupos. Estas ramas clave están definidas como eventos evolutivos en los cuáles hubo un proceso de divergencia y en el que se produjeron nuevas variantes. En el caso de los Cetáceos, algunos de los eventos clave son: Invasión de tierra a mar por Pakicetus hace 55 ma, evolución de alimentación por filtración (Mysticeti) hace 24 ma, evolución de la ecolocalización (Odontoceti) hace 34 ma, evolución del comportamiento de buceo extremo (Physeteridae) hace 30 ma y la evolución de grupos sociales adhesivos (Delphinidae) hace 10 ma (Zurano *et al.* 2019; McGowen *et al.* 2020). Para los Pinnípedos, existen menos eventos clave por ser un grupo de organismos relativamente nuevo, siendo la más importante la transición del ambiente terrestre al acuático por Puijila hace aproximadamente 20 ma (Rybczynski *et al.* 2009). Analizar el tipo de selección bajo el que se encuentran los genes es importante

para entender cómo han evolucionado. La selección natural en los genes puede promover que existan nuevas variedades y fenotipos que puedan responder mejor ante la presión ambiental (Vallender y Lahn 2004).

El sueño unihemisférico presente en Cetáceos y Pinnípedos tiene diferencias considerables, siendo obligado en Cetáceos y facultativo en Pinnípedos, lo que puede ser resultado de la actividad de distintos genes y por lo tanto de distintos procesos moleculares en la regulación de esta característica. Los Pinnípedos y Cetáceos pertenecen a la clase Mammalia, sin embargo, son muy divergentes y han tenido procesos evolutivos muy distantes. Los Pinnípedos pertenecen al grupo de los Carnívoros, originado hace aproximadamente 65 ma, mientras los Pinnípedos divergieron hace aproximadamente 25 ma (Nyakatura y Bininda-Emonds 2012). Dentro del grupo de los Pinnípedos, los Otáridos divergieron hace aproximadamente 20 ma y los Phócidos hace aproximadamente 11 ma (Nyakatura y Bininda-Emonds 2012). Por su parte, los Cetáceos pertenecen al grupo Cetartiodactyla, que se originó hace aproximadamente 70 ma y divergieron de los rumiantes hace aproximadamente 55 ma (Zurano *et al.* 2019). Con base en la lejanía del último ancestro común entre Cetáceos y Pinnípedos, su relación filogenética tan distante, y la diferencia de que en Cetáceos esta característica es obligada y en Pinnípedos facultativa, es plausible que la regulación del sueño unihemisférico esté dada por genes diferentes.

Para identificar cuáles son los posibles genes relacionados al sueño unihemisférico en Cetáceos y Pinnípedos, y analizar las similitudes y diferencias entre estos genes, es posible utilizar la genómica comparativa. Mediante la genómica comparativa se pueden identificar los genes homólogos relacionados con el sueño unihemisférico en distintas especies, con el fin de encontrar los patrones de homología en ambos grupos, identificar los genes candidatos a ser reguladores de este tipo de sueño en estas especies, y, en última instancia, identificar si podría existir una convergencia evolutiva molecular de estos genes entre ambos grupos (Kawashima 2019). Existen dos principios básicos de la genómica comparativa: las secuencias de ADN de las especies son conservadas, es decir, que los genes que codifican para proteínas responsables de alguna función serán conservados en especies emparentadas con presiones selectivas similares; y que existen secuencias que codifican proteínas que diferencian a las especies haciéndolas divergentes (Wei *et al.* 2002; Hardison 2003). Con base en lo anterior, es posible enfocarse en las secuencias conservadas de los genes homólogos relacionados con ciclos circadianos y presión homeostática del sueño en Cetáceos y Pinnípedos con el fin de explorar las similitudes y diferencias en los genes que regulan el sueño unihemisférico en estas especies (Koonin 2005; Gabaldón y Koonin 2013; Kawashima 2019).

No existen muchos estudios enfocados en identificar genes relacionados con alguna característica o en elucidar convergencias evolutivas moleculares en mamíferos y Cetáceos utilizando genómica comparativa. La evolución molecular de la ecolocalización en Cetáceos se analizó haciendo búsqueda de genes homólogos y utilizando reconstrucción de secuencias ancestrales de distintos ancestros en este grupo. Las comparaciones se realizaron con las secuencias de murciélagos existentes, que son el único otro grupo de mamíferos que presentan ecolocalización y que, además, son filogenéticamente muy distantes de los Cetáceos (Liu *et al.* 2018). Se identificó como gen responsable de la ecolocalización en estos grupos a la Prestina, habiendo evidencia de que las Prestinas son responsables del cambio en el sistema auditivo de estos organismos a partir del último ancestro común de Odontocetos, y que estos cambios tienen convergencia funcional con la Prestina que presentan los murciélagos. Esta convergencia funcional es producto de sustituciones de aminoácidos que fueron adaptativas y se dieron tanto en murciélagos como en Cetáceos (Liu *et al.* 2018).

En cuanto a otros mamíferos, se han identificado los genes responsables del desarrollo de venenos órales en algunas especies de musarañas y el solenodón, analizando los genes homólogos y la evolución convergente molecular de estos venenos en estas especies (Casewell *et al.* 2019). El sistema de entrega del veneno varía considerablemente entre estas especies en morfología, sin embargo, no era claro si la característica de ser venenoso era el estado ancestral en todo el grupo o si fue el resultado de orígenes independientes en musarañas y el solenodón. Análisis genómico-comparativos identificaron a la serina proteasa kallikreina 1 (KLK1) como el gen responsable de esta característica, y, además, que evolucionó independientemente en las musarañas y en el solenodón. Esta proteína sufrió adaptaciones y cambios para producir veneno, de manera independiente y convergente en musarañas y el solenodón, después de la divergencia de estos dos grupos en el Cretácico tardío y es un ejemplo de cómo una proteína puede presentar cambios adaptativos de manera independiente en dos grupos, por presión selectiva (Casewell *et al.* 2019). Las metodologías para poder identificar genes reguladores de alguna característica se basan principalmente en la búsqueda de genes homólogos entre las especies que presentan la característica de interés, en la identificación de los genes candidatos, en la reconstrucción de genes ancestrales y por último en la evaluación y comparación de las características de la proteína en los distintos linajes para comprobar si existe convergencia funcional. En el presente trabajo se identificarán los genes candidatos a regular el sueño unihemisférico en Cetáceos y Pinnípedos, evaluando si esta característica está regulada por los mismos genes en ambos grupos, y analizando las características, similitudes y diferencias en los genes encontrados. Estos análisis permitirán analizar si existen diferencias en la regulación del sueño unihemisférico de estas especies, producto de las diferencias observadas en si este rasgo es obligado o facultativo.

1.2 Justificación

El origen, evolución y base genética del sueño unihemisférico en Cetáceos y Pinnípedos son interrogantes que aún persisten. Se han llevado a cabo muchos estudios en los cuales se evidencia el funcionamiento estructural, la base fisiológica y las diferencias con los tipos de sueño bihemisférico NREM y REM modelando distintas condiciones para saber la estructura y funcionamiento del cerebro durante este tipo de sueño. Sin embargo, aún se desconoce la base molecular, es decir, qué genes están involucrados en la regulación de este tipo de sueño, así como su evolución en los distintos linajes que lo poseen en respuesta a la presión selectiva de su ambiente. Los Cetáceos y Pinnípedos presentan esta característica de manera diferente, siendo obligada en Cetáceos y facultativa en Pinnípedos, por lo que es importante analizar si esta característica está regulada por los mismos genes o por genes diferentes en estas especies. Es importante conocer cuáles son los genes y la base molecular de este tipo de sueño en los organismos que lo presentan para poder inferir su origen evolutivo, el funcionamiento de las proteínas que lo regulan, las ventajas adaptativas en los distintos ambientes y las diferencias que existen entre el sueño unihemisférico obligado y facultativo.

1.3 Hipótesis

Debido a que los grupos de organismos analizados presentan características fisiológicas, etológicas y morfológicas similares en su tipo de sueño, pero a su vez difieren en si esta característica es obligada o facultativa:

- Los genes identificados como reguladores del sueño unihemisférico serán diferentes en Cetáceos y en Pinnípedos por las diferencias que existen en si este sueño es obligado o facultativo.
- Existirá mayor presión de selección en las ramas clave de la evolución de los Cetáceos y en los Cetáceos actuales, en comparación con las ramas clave de los Pinnípedos y en los Pinnípedos actuales debido a que los Cetáceos presentan el sueño unihemisférico obligado.

1.4 Objetivos

1.4.1 Objetivo General

Identificar cuáles son los posibles genes relacionados al sueño unihemisférico en Cetáceos y Pinnípedos mediante el uso de la genómica comparativa.

1.4.2 Objetivos Específicos

- Identificar los genes relacionados con el sueño unihemisférico en Pinnípedos y Cetáceos, llevando a cabo un análisis genómico comparativo basado en inferencia por grafos.
- Analizar la historia evolutiva de los genes identificados llevando a cabo una reconciliación de los árboles de genes relacionados con el sueño y el árbol de especies de Cetáceos y Pinnípedos.
- Evaluar la tasa de selección de los genes relacionados con el sueño en ramas clave del árbol obtenido de todas las secuencias de Cetáceos y Pinnípedos.
- Inferir el estado ancestral de los genes en distintos puntos de la historia evolutiva de Cetáceos y Pinnípedos, con base en el registro fósil.

Capítulo 2. Metodología

2.1 Obtención de datos genómicos

Para obtener los datos de los genes relacionados al sueño primero se descargaron las secuencias de nucleótidos codificantes (CDS) de los genomas de 18 especies de mamíferos del repositorio de genomas del NCBI (del inglés, National Center for Biotechnology Information) (<https://www.ncbi.nlm.nih.gov>), donde se consideraron 10 especies de Cetáceos, 5 especies de Pinnípedos, el humano, el ratón y el perro (Tabla I). El humano y el ratón fueron seleccionados con el fin de utilizarlos como grupos externos. Al tener genomas de organismos que no están emparentados con los grupos principales del estudio (Cetáceos y Pinnípedos) es posible etiquetar a estas tres especies que no presentan sueño unihemisférico como organismos testigo, debido a que en teoría no compartirán los genes relacionados al sueño unihemisférico con Cetáceos y Pinnípedos.

Tabla 1. Lista de especies con los respectivos números de acceso a sus genomas en el NCBI.

	Familia	Especie	Ensamblaje GenBank	Ensamblaje RefSeq
Cetacea	Balaenopteridae	<i>Balaenoptera acutorostrata</i>	GCA_000493695.1	GCF_000493695.1
	Phocoenidae	<i>Neophocaena asiaorientalis</i>	GCA_003031525.1	GCF_003031525.1
	Physeteridae	<i>Physeter catodon</i>	GCA_002837175.2	GCF_002837175.2
	Iniidae	<i>Lipotes vexillifer</i>	GCA_000442215.1	GCF_000442215.1
	Monodontidae	<i>Monodon monoceros</i>	GCA_005190385.2	GCF_005190385.2
	Monodontidae	<i>Delphinapterus leucas</i>	GCA_002288925.3	GCF_002288925.3
	Delphinidae	<i>Globicephala melas</i>	GCA_006547405.1	GCF_006547405.1
	Delphinidae	<i>Orcinus orca</i>	GCA_000331955.2	GCF_000331955.2
	Delphinidae	<i>Tursiops truncatus</i>	GCA_001922835.1	GCF_001922835.1
	Delphinidae	<i>Lagenorhynchus obliquidens</i>	GCA_003676395.1	GCF_003676395.1
Pinnipedia	Phocidae	<i>Leptonychotes weddellii</i>	GCA_000349705.1	GCF_000349705.1
	Phocidae	<i>Neomonachus schauinslandi</i>	GCA_002201575.1	GCF_002201575.1
	Phocidae	<i>Mirounga leonina</i>	GCA_011800145.1	GCF_011800145.1
	Odobenidae	<i>Odobenus rosmarus divergens</i>	GCA_000321225.1	GCF_000321225.1
	Otariidae	<i>Zalophus californianus</i>	GCA_900631625.1	GCF_900631625.1
Carnivora	Canidae	<i>Canis lupus dingo</i>	GCA_012295265.1	GCF_012295265.1
Primates	Hominidae	<i>Homo sapiens</i>	GCA_000001405.28	GCF_000001405.39
Rodentia	Muridae	<i>Mus musculus</i>	GCA_000001635.8	GCF_000001635.26

La selección de estos genomas se llevó a cabo bajo el criterio de que tuvieran ensamble tanto de GenBank como de RefSeq, con el fin de analizar los genomas con mejor anotación disponibles y evitar que existieran errores en el análisis consecuencia de errores previos de anotación de los genomas. Para el ensamble RefSeq se llevan a cabo distintos procesos de curación, por lo que presentan menos errores de anotación en comparación con los ensamblajes GenBank. En los ensamblajes RefSeq existe un conjunto no redundante de estándares de referencia derivados de las bases de datos INSDC (del inglés, International Nucleotide Sequence Database Collaboration) (Pruitt *et al.* 2020).

2.2 Identificación de homólogos basada en inferencia por grafos

Con el objetivo de identificar los genes con relaciones de homología entre las 18 especies del estudio (Tabla I), se comparó de manera pareada a las especies de Cetáceos y Pinnípedos, el humano, el ratón y el perro y se identificaron todos los genes homólogos que comparten las 18 especies. Llevar a cabo este procedimiento permite analizar las diferencias en número de genes compartidos y determinar si existen genes ortólogos compartidos únicamente entre los grupos de interés, los Cetáceos y Pinnípedos.

Para este análisis se utilizó el software Proteinortho (Lechner *et al.* 2011). Este software fue diseñado para el análisis de bases de datos de gran escala (genomas) pudiendo aplicarlo a estudios con múltiples especies (Hellmuth y Wieseke 2016). Este análisis consiste en dos etapas: La etapa de agrupamiento (BLAST) y la etapa de construcción del grafo. La etapa de agrupamiento consiste en un apareamiento por similitudes en las secuencias, basado en el algoritmo BLAST (Altschul *et al.* 1990). Este apareamiento se calcula entre todas las secuencias utilizando su similitud (Lechner *et al.* 2011). Se obtiene un grafo donde los vértices representan los genes y las aristas representan las relaciones de ortología. Por último, se lleva a cabo un agrupamiento espectral de ortólogos utilizando un algoritmo de agrupamiento basado en Cadenas de Markov, el cual clasifica en clusters a cada posible familia de genes. Los grafos finales son construidos a partir de los resultados obtenidos de la etapa de construcción del grafo, sobre el cual se lleva a cabo la etapa de agrupamiento para determinar los grupos de ortólogos (Lechner *et al.* 2011; Hellmuth y Wieseke 2016). Este análisis se realizó utilizando las secuencias de nucleótidos y estimando el valor E de Blast para la similitud entre las secuencias.

2.3 Filtrado e identificación de genes relacionados al sueño

Para realizar el filtrado de los genes que están relacionados al sueño, del repositorio AmiGO 2 se extrajo la lista de todos los genes relacionados al sueño en dos distintas categorías de ontología de genes (Carbon *et al.* 2009). La categoría de ontología de genes GO:0030431 donde se categorizan todos los genes relacionados con el sueño, es decir, genes relacionados con cualquier proceso en el que un organismo ingresa y mantiene un estado periódico, fácilmente reversible, de conciencia y actividad metabólica reducida. También, la categoría GO:0050802, con todos los genes relacionados con ciclos de sueño circadianos. De la lista obtenida, se llevó a cabo un filtrado eliminando duplicados (etiqueta del gen) en la lista, es decir, el mismo gen, pero reportado en diferentes especies, con el fin de tener únicamente un registro y etiqueta por gen.

Una vez obtenida la lista de todos los genes relacionados al sueño, para obtener únicamente los ortólogos relacionados al sueño, se realizó a cabo un filtrado utilizando la lista obtenida de AmiGO y la lista de todos los genes ortólogos entre las especies que se obtuvo con el software Proteinortho en el paso anterior. El software ProteinOrtho realiza el agrupamiento y nombramiento de ortólogos con base en la anotación de los genes de la secuencia codificante (CDS) obtenido del NCBI para cada especie. La nomenclatura utilizada para etiquetar esos genes, es la misma que se utiliza en el repositorio AmiGO, por lo tanto, esa nomenclatura fue la utilizada para filtrar por nombre únicamente los ortólogos relacionados con el sueño.

La función de cada uno de los genes se obtuvo del repositorio AmiGO. En la base de datos de este repositorio se encuentran las descripciones y las funciones de cada uno de los genes obtenidos, dichas descripciones y funciones han sido obtenidas de distintos estudios que llevan a cabo la anotación de los genes en distintas especies.

2.4 Reconciliación del árbol de genes con el árbol de especies

2.4.1 Árbol filogenético

Para la obtención del árbol de especies con el cual se llevó a cabo la reconciliación del paso siguiente, se realizó un análisis filogenético de las 18 especies. Para la construcción de este árbol, se concatenaron un total de 13 genes, 4 mitocondriales (CYTB, COX1, ND1 y ND2) y 9 nucleares (ADORA, ADRB2, APOB, APP,

BCHE, BDNF, BMI1, CNR1 y PNO). La selección de estos genes estuvo basada en la filogenia de mamíferos en la que se ha utilizado mayor número de genes tanto nucleares como mitocondriales realizada hasta la fecha, por lo tanto, se seleccionaron los mismos genes que estos autores utilizaron para tener la filogenia más acertada posible (Upham *et al.* 2019). El concatenado de los 13 genes de las 18 especies se ejecutó en el software Biopython y se alineó utilizando el algoritmo MAFFT v7.273 (Katoh y Standley 2013). Una vez alineadas las secuencias, se realizó análisis filogenético de secuencias concatenadas como lo describe Gadagkar y Rosenberg (2005), donde se estimó el modelo de sustitución nucleotídica que mejor las explicara, siendo este el modelo general de tiempo reversible con variación gamma y sitios invariantes. Posteriormente se llevó a cabo el análisis filogenético por máxima verosimilitud utilizando el software PAML 4, con un bootstrap de 10000 repeticiones y calibrando el árbol con el registro fósil (Pakicetus hace 55 ma, Puijila hace 24 ma) para obtener con precisión la divergencia de las ramas en millones de años (ma).

2.4.2 Reconciliación

Para inferir la historia evolutiva de cada uno de los genes analizados, se llevó a cabo una reconciliación de los árboles de genes relacionados al sueño con el árbol de especies obtenido en el paso anterior. El objetivo fue determinar las ramas del árbol de especies donde se dieron los eventos de pérdida, modificaciones y duplicaciones de los genes. Esto se realizó utilizando la herramienta REvolutionH-tl (<https://gitlab.com/jarr.tecn/revolutionh-tl>), la cual infiere la historia evolutiva de familias de genes entre distintas especies, utilizando el grafo de ortólogos obtenido de ProteinOrtho. El grafo de ortólogos es transformado en un árbol de genes utilizando un algoritmo de descomposición modular para el grafo y un algoritmo “min-cut” que asegura que se añadan relaciones de ortología de interés y eliminen relaciones de ortología sin buen soporte. Esta herramienta nos permite obtener un árbol de reconciliación global y árboles individuales de genes.

El paso final del análisis es la reconciliación global entre los árboles de genes obtenidos a partir del grafo de ortólogos, con el árbol de especies. En esta reconciliación, se mapea cada uno de los árboles de genes individuales al árbol de especies (Lechner *et al.* 2011; Hernandez-Rosales *et al.* 2012). Este método proporciona un árbol de reconciliación global donde se representa el número de genes totales que existen en un nodo o rama, el número de genes que se ganan y los genes que se pierden. Además, en este mismo paso, se obtienen los árboles de genes individuales donde las hojas representan a las secuencias (genes) y

los nodos internos representan a los eventos evolutivos (divergencia de genes (especiación), duplicación o pérdida).

2.4.3 Árboles de genes individuales y selección de genes candidatos

Para obtener los genes candidatos a ser los relacionados con el sueño unihemisférico en Cetáceos y Pinnípedos, se realizó un análisis de los árboles de genes individuales obtenidos. Estos árboles representan la evolución de la familia de genes, que tienen alguna relación de homología con un gen particular dentro de los grupos de especies e indican cómo un gen o grupo de genes relacionados se comparten entre distintas especies y cómo han presentado cambios en la historia evolutiva de éstas.

El criterio de selección de los genes candidatos a ser reguladores del sueño unihemisférico se basó en los patrones de homología de los distintos genes en las especies del estudio que presentan el rasgo de interés (sueño unihemisférico). Para esta búsqueda y selección de genes candidatos, se analizaron todos los árboles de genes individuales (N=39, ver resultados), explorando cuáles presentan relaciones de homología entre las especies de Pinnípedos y Cetáceos, cuáles solo entre los Cetáceos y cuáles solo entre los Pinnípedos, descartando aquellos árboles donde un gen es compartido entre todas las especies del estudio. El descarte de estos árboles de genes individuales compartidos entre todas las especies se realizó debido a que el objetivo del trabajo es encontrar los genes que tienen patrones de homología únicamente en Cetáceos y Pinnípedos.

Además, se analizaron los patrones de homología de los distintos genes en especies que no presentan sueño unihemisférico, con el fin de evaluar si en los árboles de genes obtenidos se encuentran genes relacionados con el sueño bihemisférico.

2.5 Análisis filogenéticos y selección dN/dS

Para realizar el análisis de selección en cada una de las ramas de cada árbol filogenético para los distintos genes candidatos (CASP1 y PER3), se extrajeron las secuencias codificantes de esos genes de los genomas de las 18 especies. Además, se obtuvieron secuencias codificantes de otras especies de mamíferos,

relacionadas filogenéticamente con las 18 especies iniciales, del repositorio del NCBI para cada uno de estos genes. Esta descarga de secuencias se realizó con el fin de aumentar la cantidad de especies y aumentar la cantidad de información de cada gen para los análisis subsecuentes. Para el gen CASP1, se obtuvieron secuencias de especies relacionadas cercanamente con las especies los distintos grupos de su árbol, teniendo secuencias de especies de Primates (*Pan paniscus* y *Pongo abelii*), de Roedores (*Castor canadensis* y *Rattus rattus*), de Perisodáctilos (*Equus caballus*) y de Cetartiodáctilos (*Sus crofa*, *Bos taurus*, *Balaenoptera musculus* y *Phocoena sinus*). Por último, para el gen PER3, también se obtuvieron secuencias de especies cercanas a las de los grupos de su árbol, teniendo secuencias de Carnívoros (*Panthera pardus* y *Panthera tigris*), de Pinnípedos (*Callorhinus ursinus*, *Eumetopias jubatus*, *Halichoeres grypus* y *Phoca vitulina*), de Perisodáctilos (*Equus caballus*) y de Cetartiodáctilos (*Bos taurus*, *Balaenoptera musculus* y *Phocoena sinus*).

Las secuencias codificantes de las 18 especies del estudio y de otras especies de mamíferos de cada uno de los genes CASP1 y PER3, se alinearon por separado utilizando el algoritmo MAFFT v7.273 (Katoh y Standley 2013) y se tradujeron a aminoácidos para el posterior análisis de la estructura de aminoácidos de cada gen en las reconstrucciones de las proteínas de los ancestros y en las especies actuales. Se determinó el modelo de sustitución nucleotídica (modelo evolutivo) y de sustitución de aminoácidos para cada uno de los genes utilizando el software JModelTest 2.1.4 (Darriba *et al.* 2012). El modelo evolutivo se utilizó para llevar a cabo el análisis filogenético basado en nucleótidos de cada uno de los genes por Máxima Verosimilitud utilizando el software PAML 4.8 (Yang 2007) con un bootstrap de 10000 repeticiones. El método para calcular la tasa de selección en distintas ramas de un árbol filogenético está basado en máxima verosimilitud, por lo tanto, es necesario llevar a cabo los árboles filogenéticos para este análisis con este método, con el fin de que concuerden los cálculos de la distancia entre las ramas y del bootstrap. El análisis por nucleótidos se realizó con el fin de analizar toda la variación posible en las secuencias y el análisis por aminoácidos se realizó para comparar la composición de aminoácidos entre cada una de las secuencias.

Utilizando los alineamientos y árboles filogenéticos de cada uno de los genes obtenidos en el paso anterior, se evaluó si existe evidencia de selección positiva ($dN/dS > 1$) sobre estos dos genes, en ramas clave del árbol obtenido de todas las secuencias de Cetáceos, Pinnípedos y otros mamíferos, con el fin de comparar el efecto de la selección natural sobre distintas ramas y grupos del árbol filogenético. Para esto, se implementó un modelo de sitios-ramas utilizando la paquetería *codeml* de PAML donde se analizó el nivel de selección en distintas ramas interiores del árbol definidas por cada grupo de especies dentro del análisis (Cetacea, Pinnipeda, Primates, Rodentia). El análisis de presión de selección sobre estos nodos

determinará si han existido cambios potenciales en genes asociados al sueño durante la historia evolutiva de cada grupo taxonómico analizando la presión que ha existido en las secuencias de especies actuales a partir del último ancestro común de cada rama (UAC). Las ramas evaluadas variaron según el gen analizado según las especies de cada árbol filogenético y según los distintos eventos de divergencia clave en dichos árboles. Para el gen CASP1, se evaluó la selección en la rama de los Primates a partir del UAC-Hominidae, Perisodactyla-Cetartiodactyla y en la rama de los Cetáceos a partir del UAC-Cetacea. Por último, para el gen PER3, se analizó la selección sobre los Pinnípedos, a partir del UAC-Pinnipedia, y en la rama de los Cetáceos, a partir del UAC-Cetacea.

2.6 Inferencia del estado ancestral

La inferencia del estado ancestral de los genes nos permite analizar cómo eran estos genes en los ancestros y cómo han evolucionado en las especies actuales. Para esta reconstrucción del estado ancestral de las secuencias de cada uno de los genes (CASP1 y PER3) se utilizaron como base los análisis filogenéticos por máxima verosimilitud obtenidos en el paso anterior y los alineamientos de secuencias de aminoácidos para analizar la composición de aminoácidos de cada gen. Este análisis también está basado en inferencia por máxima verosimilitud, por lo tanto, también es requerido que los árboles filogenéticos hayan sido obtenidos mediante este método.

La reconstrucción del estado ancestral se llevó a cabo utilizando el software PAML 4. El análisis calcula la correlación entre la edad de los últimos ancestros comunes de cada rama, las distancias evolutivas entre las especies y las probabilidades de las distribuciones posteriores de los nucleótidos y aminoácidos variantes (Yang 2007). En este análisis, se determina la precisión, definida como la probabilidad promedio de la ocurrencia de cada base y aminoácido. También se determina el Valor P estadístico de la probabilidad de que la secuencia completa obtenida sea correcta. Estos dos estadísticos permiten evaluar la probabilidad estadística de cada nucleótido y aminoácido dentro de la secuencia y también la probabilidad estadística de la secuencia completa.

La reconstrucción del estado ancestral fue diferente para cada uno de los dos genes, dependiendo de las especies incluidas en los árboles filogenéticos y de los ancestros comunes a los cuales había interés de hacerles la reconstrucción. Para el gen CASP1, se reconstruyó la secuencia para el UAC-Hominidae-Perisodactyla-Cetartiodactyla, UAC-Perisodactyla-Cetartiodactyla, UAC-Cetartiodactyla, UAC-Cetacea y

UAC-Odontoceti. Por último, para el gen PER3, se reconstruyeron las secuencias de los ancestros UAC-Carnivora-Perisodactyla-Cetartiodactyla, UAC-Carnivora-Cetartiodactyla, UAC-Carnivora, UAC-Pinnipedia-Canidae, UAC-Pinnipedia, UAC-Phocidae, UAC-Cetartiodactyla, UAC-Cetacea y UAC-Odontoceti. La selección de estos ancestros, se realizó con base en los eventos de divergencia de especies, en donde se haya llevado a cabo la divergencia entre los distintos grupos de especies obtenidos para cada gen.

Una vez obtenidas las secuencias de los genes para los últimos ancestros comunes en las diferentes ramas, se llevaron a cabo los alineamientos de las secuencias de cada gen ancestral con las secuencias de las especies actuales y se tradujeron a aminoácidos utilizando nuevamente el algoritmo MAFFT v7.273. Posteriormente, una vez traducidas y alineadas las secuencias, se visualizaron en el software MEGA X (Kumar *et al.* 2018), con el fin de comparar la composición de aminoácidos de los genes de las especies actuales y los genes de los ancestros, analizando las pérdidas, ganancias y sustituciones de aminoácidos dadas en los ancestros comunes y en las especies actuales.

Capítulo 3. Resultados

3.1 Obtención de datos genómicos

Para las especies consideradas, se obtuvieron los genomas codificantes descargando los archivos fasta del repositorio del GenBank. Posteriormente, se obtuvo la base de datos de las secuencias codificantes a analizar en busca de homólogos, que consistió en 912,045 secuencias totales (Tabla 2) las cuales son el total de secuencias no repetidas entre las especies.

Tabla 2. Número de secuencias utilizadas por archivo fasta para el análisis de Proteinortho.

Archivo de especie	Secuencias
Homo-sapiens.fasta	120689
Mus-musculus.fasta	88610
Canis-lupus-dingo.fasta	71098
Zalophus-californianus.fasta	61096
Lagenorhynchus-obliquidens.fasta	54651
Globicephala-melas.fasta	54607
Physeter-catodon.fasta	50636
Delphinapterus-leucas.fasta	50540
Mirounga-leonina.fasta	45230
Monodon-monoceros.fasta	45097
Leptonychotes-weddellii.fasta	40334
Tursiops-truncatus.fasta	38872
Neophocaena-asiaeorientalis.fasta	38173
Balaenoptera-acutorostrata.fasta	37684
Odobenus-rosmarus-divergens.fasta	31443
Neomonachus-schauinslandi.fasta	28431
Orcinus-orca.fasta	27924
Lipotes-vexilifer.fasta	26930
Total	912045

3.2 Identificación de genes ortólogos

En el análisis de identificación de ortólogos en el software Proteinortho, se obtuvo un total de 8,998,942 relaciones de ortología comparando en pares todas las especies. Una vez filtrados los genes, se obtuvieron un total de 14,090 relaciones de ortología, las cuales corresponden únicamente a relaciones de ortología de genes relacionados al sueño.

3.3 Filtrado e identificación de genes relacionados al sueño

La lista de genes relacionados con el sueño obtenida del repositorio AmiGO fue de 385 genes. Una vez eliminados las etiquetas de genes duplicadas, se obtuvieron 198 genes únicos los cuales se usaron para llevar a cabo el análisis de ortólogos. Al ser analizadas estas relaciones de ortología, se infirieron 57 ortogrupos, cada ortogrupo corresponde al set de genes que descienden de un solo gen que se encontraba en un ancestro común. De esos ortogrupos inferidos, los que presentaron consistencia con el árbol de especies fueron los seleccionados para analizarse, siendo un total de 39 (Tabla 4). La consistencia con el árbol de especies se da cuando el set de genes del ortogrupo inferido sigue con exactitud la topología del árbol de especies. Existen genes en los cuales la evolución no necesariamente es consistente con el árbol de especies. Debido a esto, se analizaron los árboles de genes inconsistentes con el árbol de especies con el fin de determinar si existe alguna relación de ortología o paralogía entre algún gen relacionado al sueño en las distintas especies, sin importar si esta relación es consistente con el árbol de especies.

Tabla 3. Estadísticas del análisis de genes ortólogos utilizando la herramienta revolutionhl.

Ortogrupos inferidos	57
Ortogrupos totales	59
Árboles de genes consistentes	39

Una vez filtrados los genes relacionados al sueño (de los resultados de Proteinortho) y analizados por la herramienta de reconciliación de árbol de genes con árbol de especies, se pudo determinar que existen 39 genes relacionados al sueño que tienen relaciones de homología entre las 18 especies y que, además, son consistentes con las relaciones filogenéticas inferidas (Tabla 5). De estos 39 genes consistentes en

especies con sueño unihemisférico, por las relaciones homología observadas en Cetáceos y Pinnípedos, se seleccionó a CASP1 y PER3 como posibles reguladores del sueño unihemisférico. Según la anotación de genes en el repositorio AmiGO, el gen CASP1 sintetiza la caspasa 1 y también es regulador positivo del ciclo NREM. La regulación positiva se refiere a cualquier proceso que activa o incrementa la calidad y/o duración del sueño REM o NREM. Por su parte, el gen PER3 sintetiza una proteína reguladora del ciclo circadiano y es, en general, un regulador del sueño y de los ciclos circadianos.

En cuanto al sueño bihemisférico, por las relaciones de homología observadas entre las especies con esta característica, se seleccionaron como candidatos a los genes CST3 y PTGER4 como reguladores del este tipo de sueño. El gen CST3 sintetiza la proteína cistatina C que actúa como reguladora positiva del sueño bihemisférico, particularmente en la fase REM. PTGER4 es un receptor de Prostaglandina reportado como regulador del sueño. De los demás genes, existen varias funciones, siendo su mayoría, genes reguladores del ciclo REM o reguladores negativos de NREM, sin embargo, en estos genes no se observaron patrones de homología en las especies de Cetáceos y Pinnípedos. La regulación negativa se refiere a cualquier proceso que inhibe, previene o reduce la duración y/o calidad del sueño (Carbon *et al.* 2009) (Tabla 5).

Tabla 4. Genes relacionados al sueño con relaciones de ortología entre las 18 especies y su función. Se indican los genes seleccionados como candidatos a regular el sueño unihemisférico en Cetáceos y Pinnípedos (en negrita).

Gen	Producto	Función relacionada al sueño
GHRH	Neuropéptido familia glucagón	Regulación + del ciclo circadiano, sueño REM
STAR	Proteína reguladora aguda esteroideogénica	Regulación del ciclo circadiano, sueño REM
PER3	Proteína del período circadiano	Regulación del ciclo circadiano
KCNA2	Canal de potasio por voltaje	Regulación del ciclo circadiano, sueño NREM
ADORA2A	Receptor de adenosina	Regulación del ciclo circadiano
DRD2	Receptor adrenérgico de dopamina	Regulación del ciclo circadiano
DRD3	Receptor adrenérgico de dopamina	Regulación del ciclo circadiano
CHRN2	Receptor ionotrópico	Regulación del ciclo circadiano, sueño REM
BTBD9	Dominio BTB	Regulación del ciclo circadiano NREM
SLC29A1	Transportador de nucleósidos	Sueño

FOS	Proto-oncogeno c-Fos	Sueño
CSF2	Factor de estimulación de macrófagos	Regulación del ciclo circadiano
IL18	Interleucina 18	Regulación del ciclo circadiano
HTR2A	Receptor de serotonina	Sueño
CASP1	Caspasa 1	Regulación + del ciclo circadiano, sueño NREM
ADORA1	Receptor de adenosina	Regulación - del ciclo circadiano, sueño NREM
GHRL	Precursor de ghrelina	Regulación - del ciclo circadiano, sueño REM
MRGPRX1	Proteína G relacionada a mas	Ruta de señalización
NMU	Neuromedina U	Regulación del ciclo circadiano
UTS2R	Receptor de urotensina	Regulación + del ciclo circadiano, sueño REM
ALB	Albumina	Regulación + del ciclo circadiano, sueño NREM
PTGER4	Receptor 4 de Prostaglandina E	Regulación del sueño bihemisférico
CST3	Cistatina C	Regulación del ciclo REM

3.4 Reconciliación del árbol de genes con el árbol de especies

3.4.1 Árbol filogenético base

El árbol filogenético de máxima verosimilitud obtenido a partir de los 13 genes concatenados de las 18 especies presentó buen soporte para las ramas (Figura 1). Cabe mencionar, que existen dos ramas del árbol filogenético que no tuvieron ese soporte, sin embargo, esto no afectará los análisis subsecuentes debido a que estas incongruencias se encuentran dentro de la rama de los Cetáceos y los análisis del trabajo se centran en todo el grupo de los Cetáceos y no en ramas internas de géneros o especies. La distribución de las ramas y los estimados de tiempo de divergencia coinciden con lo reportado en la literatura, indicando que los Cetáceos surgen hace aproximadamente 55 ma y se van diversificando en las

distintas familias. Por su parte, el grupo de los Carnívoros analizados, surge hace aproximadamente 40 ma, donde el grupo de los Pinnípedos diverge de la familia Canidae hace aproximadamente 25 ma (Fig 1).

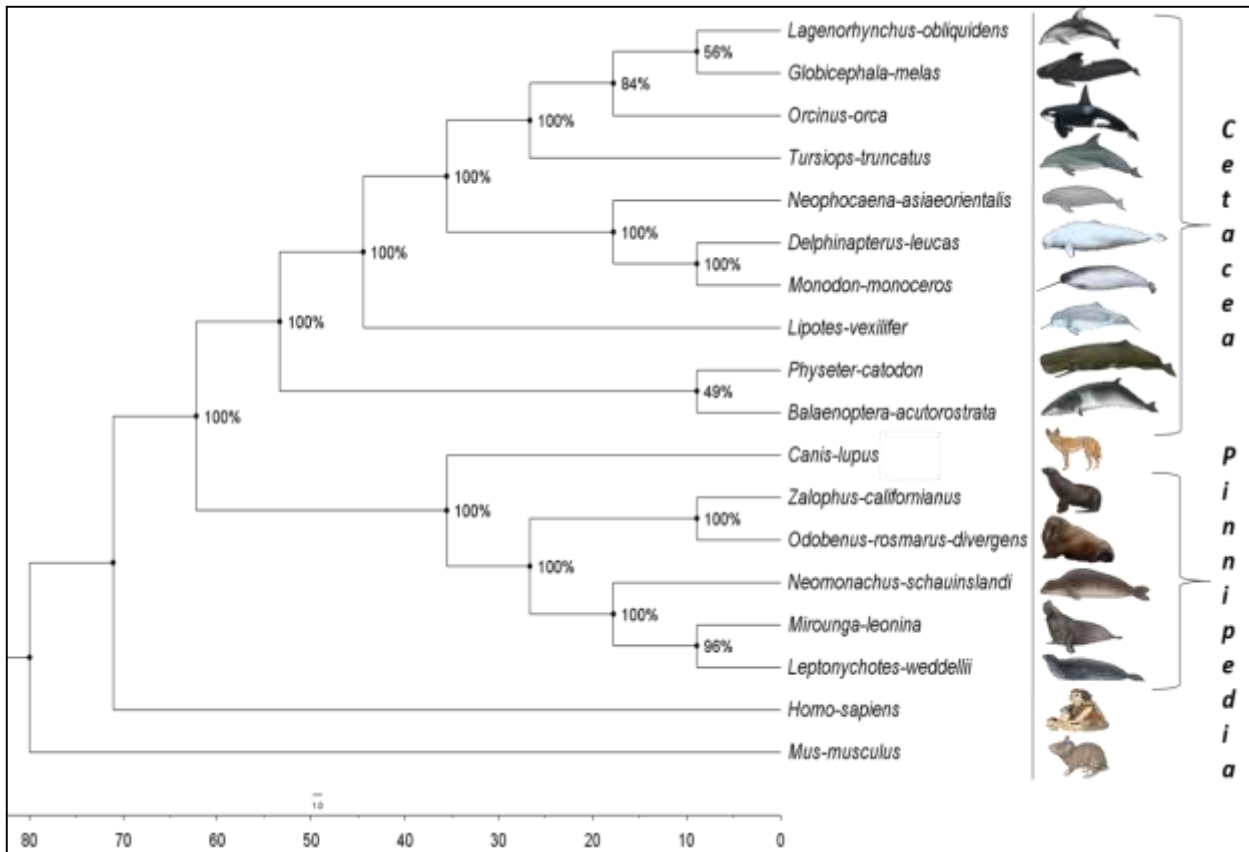


Figura 1. Árbol filogenético por máxima verosimilitud de la filogenia de mamíferos utilizando 13 genes (9 nucleares y 5 mitocondriales) y calibrado con el registro fósil de Cetáceos y Pinnípedos. Los números de la escala representan los millones de años. En cada rama se muestra el valor de bootstrap.

En el árbol de reconciliación global (Fig. 2) se observan los 198 genes evaluados con relaciones de ortología entre las 18 especies, originados en el último ancestro común de estas especies. En el clado correspondiente a los Cetáceos, hubo una ganancia de 9 genes relacionados al sueño. En el clado de los Carnívoros se observa una pérdida de 157 genes. Comparando el número de genes que hay en los clados de Cetáceos y Pinnípedos, se observa que en el último ancestro de los Cetáceos hubo 180 genes de los 189 analizados, de los cuales algunos se fueron perdiendo conforme diversificaron las especies, hasta tener en promedio 30 genes (24-34) en todas las especies de Cetáceos. Por su parte, en el clado de los Pinnípedos hubo 110 genes que se fueron perdiendo con el paso del tiempo, hasta tener en promedio 30 genes (28-31) en las especies. Estas observaciones en el árbol de reconciliación, permiten inferir que existe mayor

número de genes relacionados al sueño en los Cetáceos y en Pinnípedos, en comparación con el humano y el ratón, habiendo mayor número de genes en los Cetáceos.

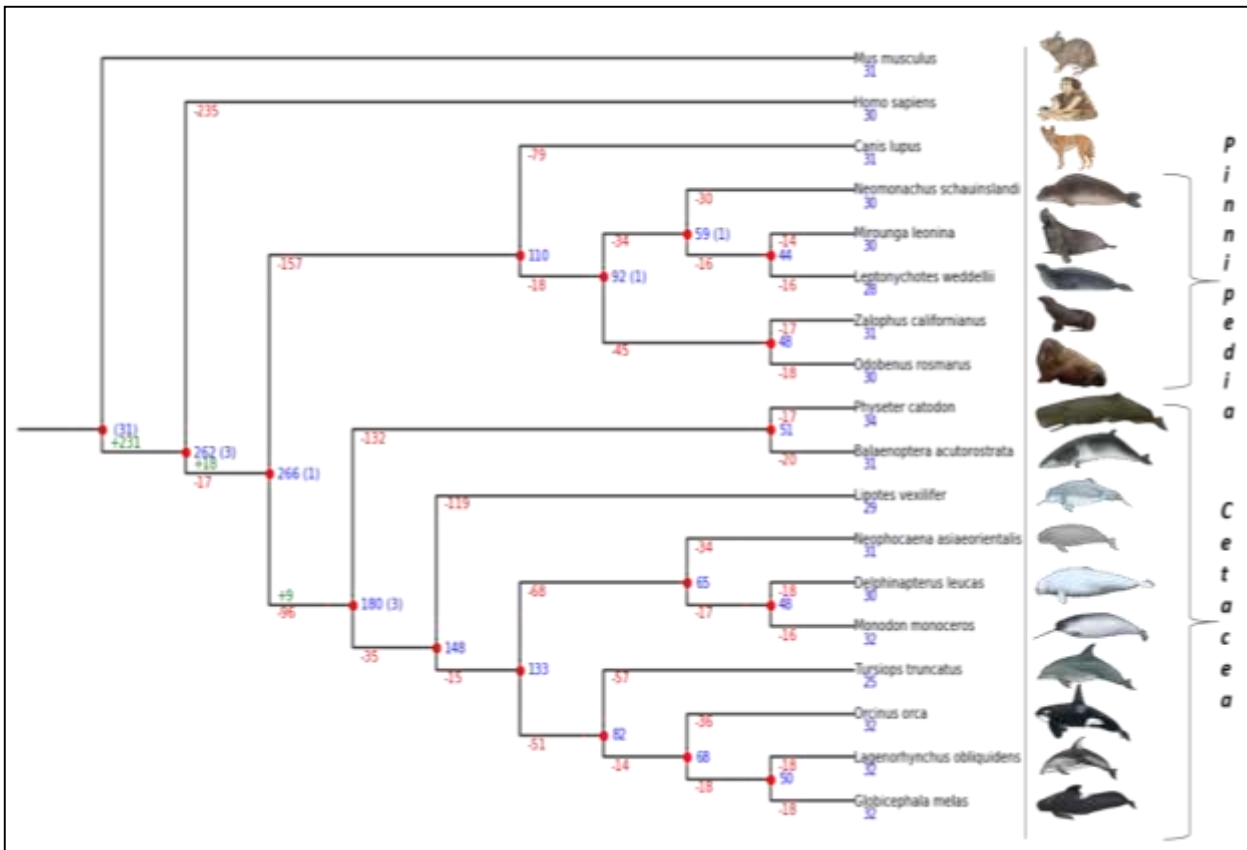


Figura 2. Reconciliación del árbol de genes relacionados al sueño con el árbol de especies. Se indica el número de genes en el nodo (azul), número de genes perdidos (rojo), número de genes ganados (verde).

En el agrupamiento de los 39 ortogrupos entre las distintas especies, existen genes que son homólogos entre las 18 especies del análisis. Sin embargo, existen agrupaciones y relaciones de homología en grupos en particular, que son las agrupaciones analizadas para elegir a los genes candidatos. Las agrupaciones más importantes que se observan son la agrupación entre los miembros de la familia Phocidae y la agrupación entre los Cetáceos y el ratón (Fig. 3).

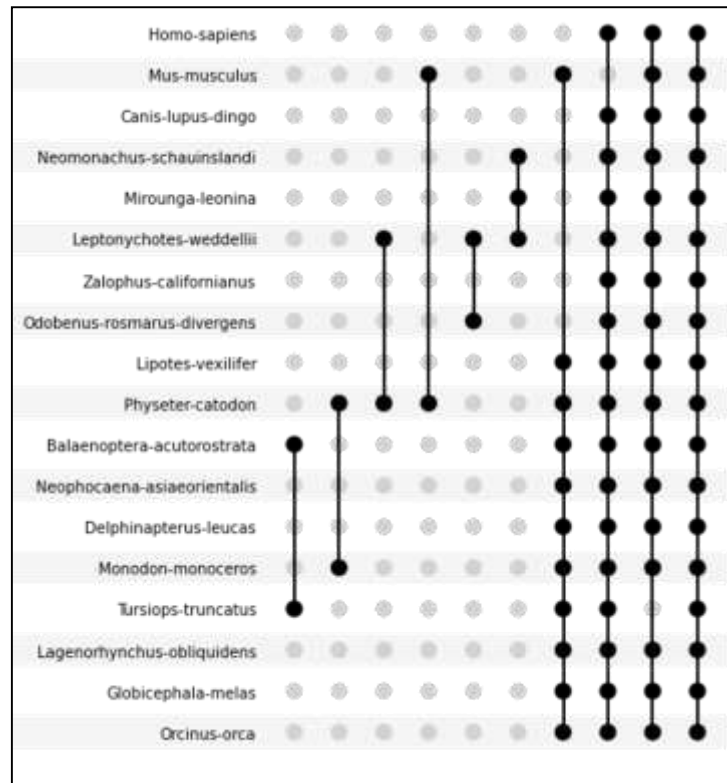


Figura 3. Agrupamiento de genes entre las distintas especies, donde cada columna representa un subconjunto de genes en las especies y cada círculo negro representa en qué especies se encuentra ese gen. En esta figura solo se observan los genes que no se comparten en todas las especies.

3.4.2 Árboles de genes individuales

Utilizando la herramienta REvolutionH-tl (<https://gitlab.com/jarr.tecn/revolutionh-tl>) se obtuvieron 39 árboles de genes individuales. De los 39 árboles de genes, en 35 tienen relación de homología entre todas las especies o no presentan relaciones en las especies de Cetáceos y Pinnípedos. Por ejemplo, el gen ADORA2A es ortólogo en todas las especies, presenta una alta similitud entre ellas y es originado a partir de un evento de especiación en el ancestro común de todas las especies consideradas (Fig 4a). En el árbol de reconciliación de la familia del gen ADORA2A no se presenta ningún evento de duplicación o pérdida de genes (Fig 4b), únicamente eventos de especiación en cada una de las ramas, y, por lo tanto, es un gen compartido y ortólogo entre todas las especies. El resultado de este árbol indica que este gen es ortólogo entre todas las especies del análisis, sin ninguna distinción en si estas tienen sueño unihemisférico o bihemisférico.

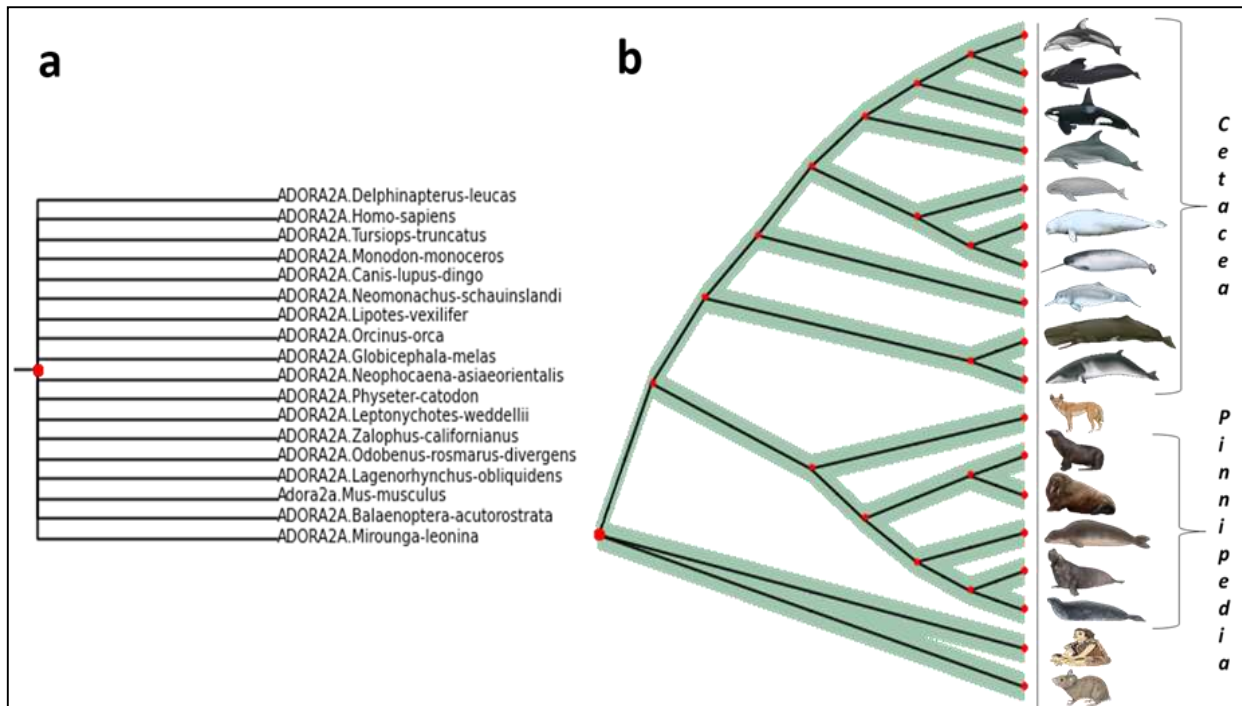


Figura 4. a) Árbol del gen ADORA2A. b) Reconciliación del árbol de la familia del gen ADORA2A con el árbol de especies. Los puntos rojos representan eventos de especiación.

Se observaron patrones de ortología entre especies con sueño unihemisférico o sueño bihemisférico en 4 de los 39 genes. De estos 4 genes seleccionados, 2 presentan patrones de homología entre especies con sueño bihemisférico y están ausentes en especies con sueño unihemisférico, lo que los hace candidatos a ser reguladores del sueño bihemisférico o del sueño tipo REM. Estos genes son PTGER4 y CST3, los cuales son muy similares en cuanto a las relaciones de ortología, paralogía, eliminaciones y duplicaciones en el árbol de reconciliación (Fig 5). Para ambos genes existe un evento de duplicación en el último ancestro entre Carnívoros, Cetáceos y Primates, sin embargo, existe una pérdida completa en el grupo de los Cetáceos (Fig. 5). Estos resultados son evidencia de que estos genes no son reguladores del sueño unihemisférico por estar ausentes en los Cetáceos que solo presentan este tipo de sueño, indicando que regulan alguna característica del sueño bihemisférico, ya que se presentan en las especies que presentan este comportamiento.

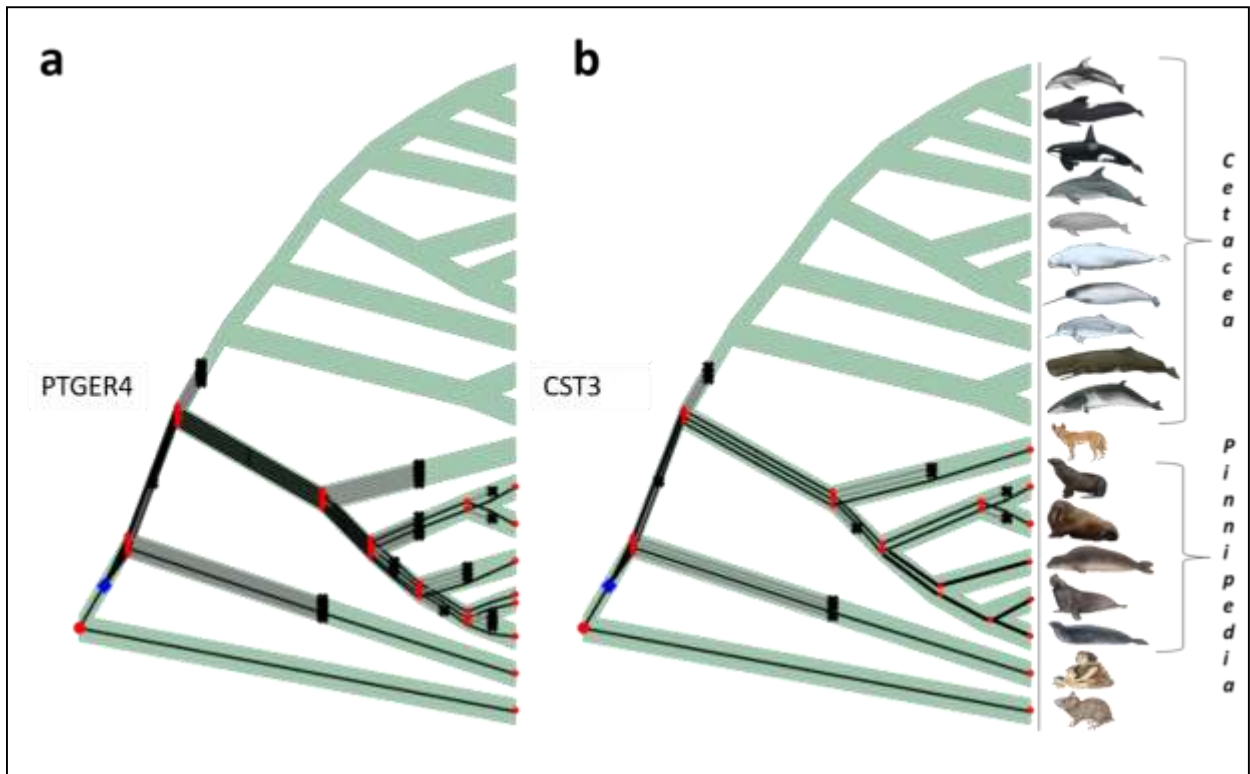


Figura 5. a) Reconciliación del árbol de genes de la familia PTGER4 con el árbol de especies. b) Reconciliación del árbol de genes de la familia CST3 con el árbol de especies. Los puntos rojos representan eventos de especiación, los puntos azules representan eventos de duplicación de genes y los símbolos "X" representan pérdida de genes.

Uno de los genes relacionados con el sueño unihemisférico es el gen CASP1, ya que presenta relaciones de homología entre todas las especies de Cetáceos consideradas en el estudio y que también tiene similitud con el gen CASP1 del ratón. Según el árbol de reconciliación, se observa que hubo un evento de especiación en el último ancestro común de Cetáceos y el ratón, y posteriormente en el clado de los Cetáceos se presentó un evento de duplicación de genes (Fig 6a). En la reconciliación con el árbol de especies se observa que el gen CASP1 se pierde en el humano y en el clado de los Carnívoros. Además, se observa un evento de duplicación de estos genes ocurrido en el clado de los Cetáceos. Sin embargo, todos los Cetáceos comparten un gen de CASP1, siendo este un gen out-parálogo producto de un evento de duplicación que se dio antes del evento de especiación en el ancestro de estas especies (Fig 6b). Los resultados del árbol de este gen indican que este gen sufrió eventos evolutivos particulares en el clado de los Cetáceos, donde a partir de un evento de duplicación se heredó el gen a las especies de Cetáceos actuales. Este patrón de homología particular dentro del grupo de los Cetáceos podría indicar que este

gen es el posible regulador del sueño unihemisférico en los Cetáceos, y el posible causante de que los Cetáceos presenten esta característica de manera obligada.

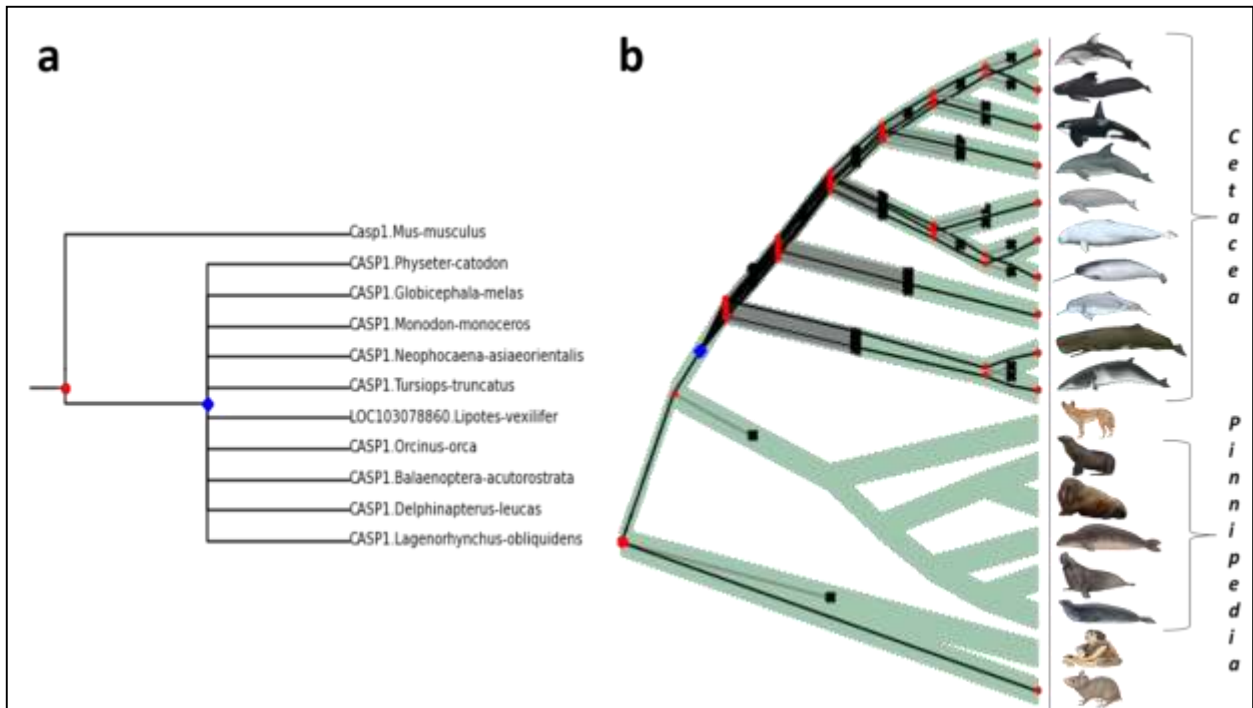


Figura 6. a) Árbol del gen CASP1. b) Reconciliación del árbol de genes de la familia CASP1 con el árbol de especies. Los puntos rojos representan eventos especiación, los puntos azules representan eventos de duplicación de genes y los símbolos "X" representan pérdida de genes.

En cuanto a PER3, se observó que este gen es conservado únicamente en la familia Phocidae (Fig 7a). En el árbol de reconciliación de la familia de genes de PER3 con el árbol de especies se observa que solo existe relación de ortología de estos genes dentro de las especies de la familia Phocidae, originándose dicha relación de ortología en el último ancestro común de estas especies. Cabe destacar que no se observa relación, es decir, que no es compartido con otros miembros de los Pinnípedos, con otros Cetáceos, ni con otras especies de mamíferos y que no existen eventos de pérdida de estos genes en otros clados (Fig 7b). Es importante recalcar que los resultados observados no indican que ese gen no está presente en las demás especies, sino que, en cuestión de similitud del gen y de relaciones de ortología, ese gen presente en la familia Phocidae es muy similar entre esas tres especies y a su vez, muy diferente del resto de las especies del estudio.

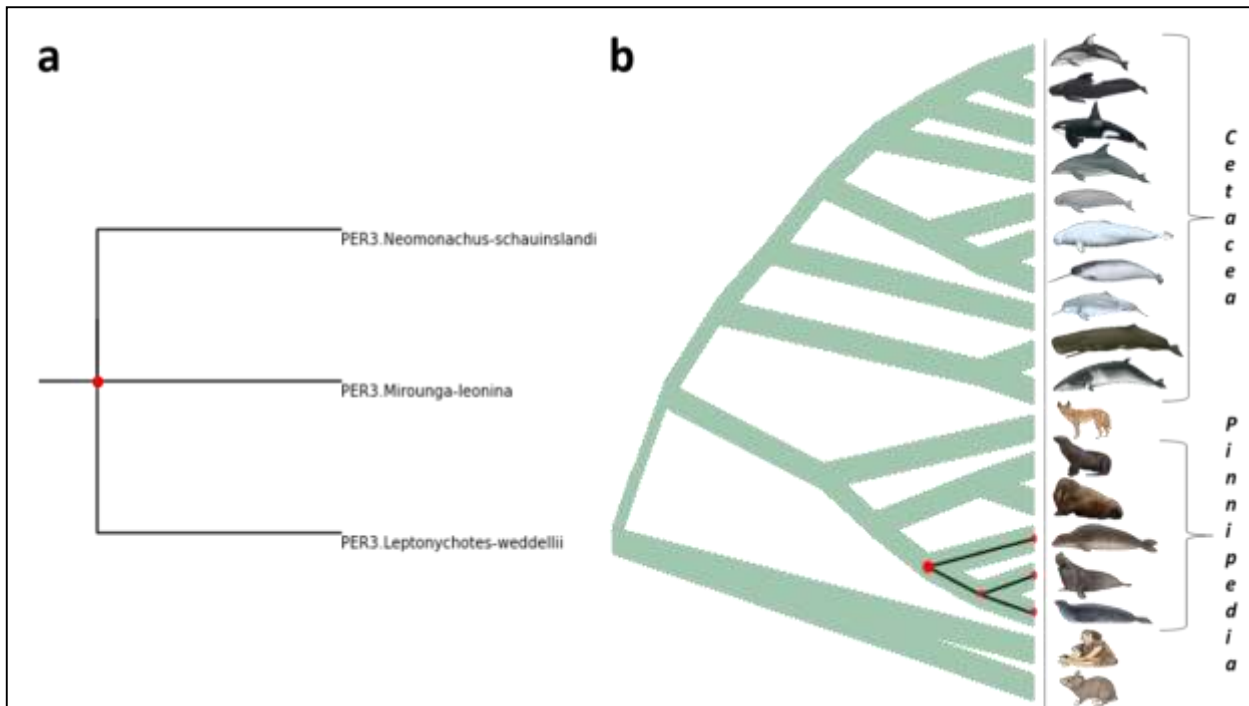


Figura 7. a) Árbol del gen PER3 de Phócidos. b) Reconciliación del árbol de genes de la familia PER3 con el árbol de especies. Los puntos rojos representan eventos de especiación, los puntos azules representan eventos de duplicación de genes y los símbolos “X” representan pérdida de genes.

Al analizar los resultados de árboles de genes inconsistentes con el árbol de especies se logró identificar al resto de genes PER3 que no se observaron en el árbol anterior. En el árbol del gen PER3 se observa que, a partir de un evento de duplicación en el último ancestro de Cetáceos y Carnívoros, se origina el gen PER3 de *Zalophus californianus* (Otariidae) y el gen PER3 de Cetáceos haciéndolos genes parálogos, y que el gen PER3 es ortólogo entre las especies de Cetáceos (Fig 8a). Además, en la red de relaciones de homología entre los genes PER3, se observa que los genes de los miembros de la familia Phocidae, están relacionados con el gen PER3 de la familia Otariidae de Pinnípedos, pero no con el gen PER3 de Cetáceos. A su vez, el gen PER3 de Cetáceos está relacionado con PER3 de la familia Otariidae pero no tiene relación con el gen de la familia Phocidae (Fig 8b). Estos resultados indican que PER3 tiene una relación de homología cercana entre la familia Otariidae y los Cetáceos y que la relación no es cercana entre los Cetáceos y los miembros de la familia Phocidae.

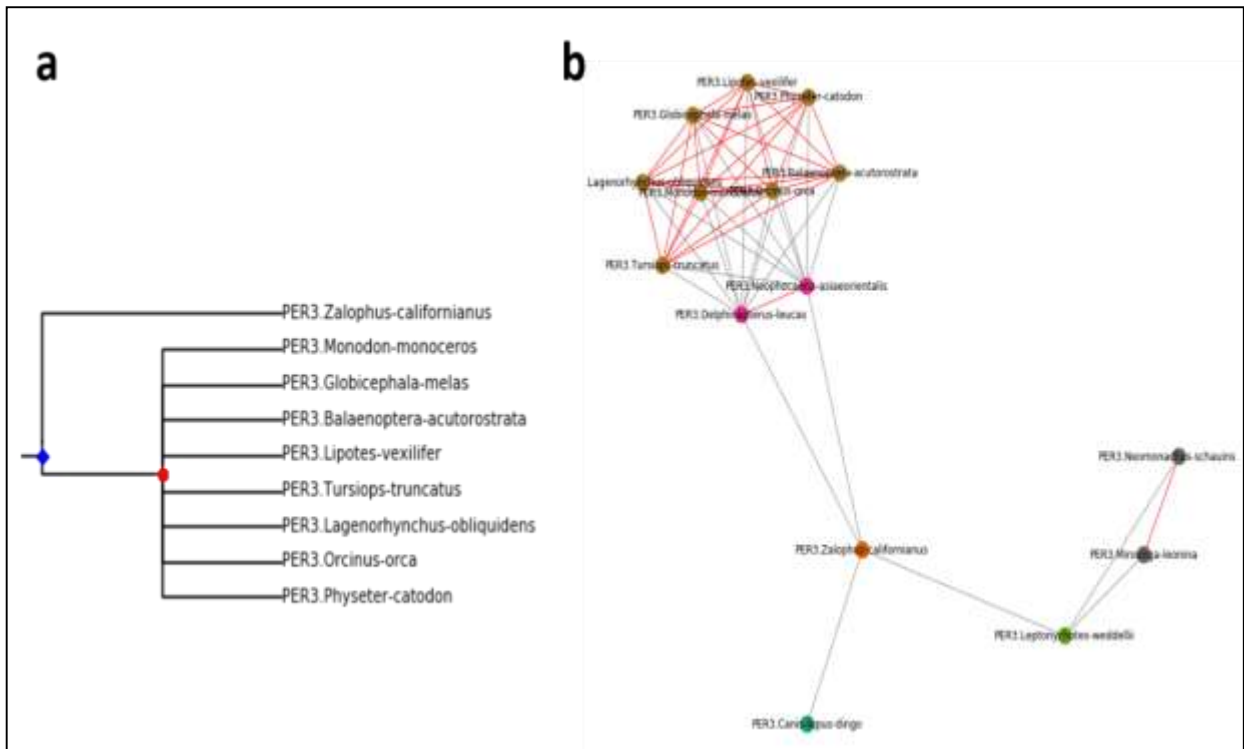


Figura 8. a) Árbol del gen PER3. b) Red de ortólogos relacionados a PER3. Para la Fig 8a, el punto rojo representa un evento de divergencia de genes (especiación), el punto azul representa un evento de duplicación de genes.

3.5 Análisis de selección dN/dS para CASP1 y PER3

El análisis de selección demostró que para el gen CASP1, existe selección negativa o purificadora en todos los grupos analizados (Cetáceos, Primates, Roedores), siendo los Cetáceos donde hay un menor valor de dN/dS, seguido de los Roedores y por último los Primates. Para el gen PER3, se presentó selección purificadora en los Cetáceos y los Carnívoros, grupos que obtuvieron los valores más bajos de dN/dS. Por su parte, los Pinnípedos presentaron valores superiores a 1, por lo que este gen en esas especies está sujeto a selección positiva. Estos resultados demuestran que los genes están sujetos a distintos tipos de selección variando entre cada grupo, teniendo un tipo de selección para el gen CASP1 en cada uno de los grupos y un tipo de selección diferente para el gen PER3.

Tabla 5. Análisis de selección por grupo para los genes CASP1 y PER3.

CASP1	
Grupo	ω dN/dS
Cetáceos	0.29683
Primates	0.5354
Roedores	0.36869
PER3	
Grupo	ω dN/dS
Pinnípedos	1.2456
Cetáceos	0.7485
Carnívoros	0.6843
Phócidos	1.1524

3.6 Inferencia del estado ancestral para los genes candidatos CASP1 y PER3.

El análisis del modelo evolutivo de sustituciones de aminoácidos determinó que el mejor modelo para el gen CASP1 fue el Jones-Taylor-Thornton con sitios invariantes y para el gen PER3 fue el modelo Jones-Taylor-Thornton con distribución gamma y frecuencia de aminoácidos (Tabla 6).

Tabla 6. Modelos de sustitución nucleotídica y de aminoácidos utilizados en el análisis de máxima verosimilitud para la reconstrucción ancestral de las secuencias de nucleótidos de los genes CASP1 y PER3.

Nucleótidos				
Gen	Modelo	Parámetros	AICc	<i>lnL</i>
PER3	GTR+G+I	63	46559.645	-23216.779
CASP1	T92+G	42	13560.805	-6738.331
Aminoácidos				
PER3	JTT+G+F	73	31287.947	-15570.794
CASP1	JTT+I	40	8290.072	-4104.841

La reconstrucción de las secuencias de los ancestros y el árbol filogenético se llevó a cabo considerando 406 aminoácidos correspondientes a 1218 pb (Fig. 9). Las reconstrucciones del gen CASP1 en los ancestros, tuvieron alta precisión y todas, excepto la del UAC-Odontoceti, presentaron un valor P significativo, demostrando que estas reconstrucciones son correctas y estadísticamente probables (Tabla 7). Esto significa que la probabilidad de la ocurrencia de cada base y aminoácido dentro de las secuencias reconstruidas para cada ancestro es muy alta y también que hay alta probabilidad de que las secuencias completas sean correctas (Tabla 9).

Tabla 7. Estadísticas de la reconstrucción del estado ancestral del gen CASP1. Precisión es la probabilidad promedio de la ocurrencia de cada base y aminoácido. El Valor P significativo indica que la secuencia completa del ancestro es correcta dado el modelo de sustitución. Los números en color rojo indican que el Valor P no fue significativo, por lo tanto, hay mayor incertidumbre en que la secuencia completa esté correcta.

Ancestro	Precisión	Valor P
UAC-Hominidae-Perisodactyla-Cetartiodactyla	0.97042	1.95E-19
UAC-Perisodactyla-Cetartiodactyla	0.98393	1.74E-10
UAC-Cetartiodactyla	0.99181	1.08E-05
UAC-Cetacea	0.9994	0.045009
UAC-Odontoceti	0.9997	0.68363

En el árbol filogenético correspondiente a CASP1 diverge el grupo de los Primates, seguido del grupo de los Perisodáctilos. Por último, el árbol indica que el grupo de los Cetáceos es el más reciente, y divergió de los Cetartiodáctilos, donde se separen claramente las distintas familias de Cetáceos (Fig. 9a).

Para el grupo de los Cetartiodáctilos se reconstruyó la secuencia de aminoácidos de 3 ancestros (UAC-Cetartiodactyla, UAC-Cetacea y UAC-Odontoceti). También, se reconstruyó la secuencia para el último ancestro común entre Cetartiodáctilos y Perisodáctilos (UAC-Perisodactyla-Cetartiodactyla) y para el último ancestro común entre Primates, Cetartiodáctilos y Perisodáctilos (UAC-Hominidae-Perisodactyla-Cetartiodactyla). En cuanto a las sustituciones y eliminaciones de aminoácidos en las secuencias de CASP1, se observa que, a partir del último ancestro común de los Cetáceos (UAC-Cetacea), ocurren diversas sustituciones de aminoácidos en distintas posiciones de la proteína (Fig. 9b). En la posición 81 se observa que hay una sustitución en el grupo de Perisodáctilos y Cetartiodáctilos, donde se sustituye una Tirosina (Y) por una Histidina (H). En esta misma posición ocurre una nueva sustitución dentro del grupo de los Cetáceos, donde se sustituye esa H por una Leucina (L). En la posición número 241, ocurre una sustitución de un Ácido Glutámico (E) por una Alanina (A) en el grupo de los Cetartiodáctilos, donde posteriormente, en los Cetáceos se sustituye esa A por una Treonina (T). En la posición 300 se observa una sustitución originada en el grupo de los Cetáceos donde hay un cambio de una Glicina (G) por un Ácido Glutámico. Por último, en la posición 363 también se observa una sustitución exclusiva de Cetáceos donde se sustituye una Serina (S) por una Fenilalanina (F). Estas distintas sustituciones de aminoácidos observadas en el clado de los Cetáceos, demuestran que la proteína CASP1 de los Cetáceos tiene una composición de aminoácidos y que dicha composición se heredó a las especies actuales a partir de su último ancestro común.

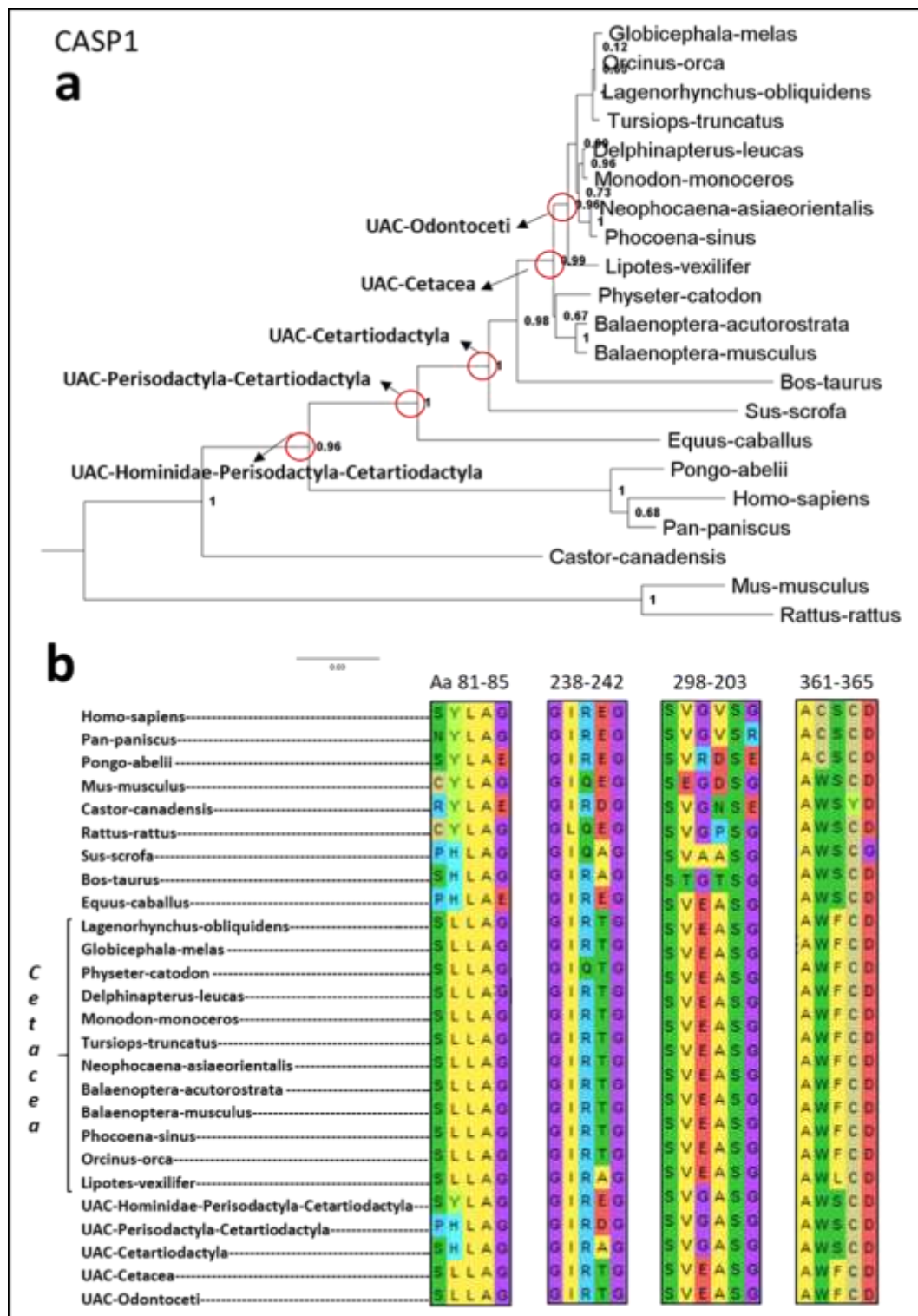


Figura 9. a) Árbol filogenético por máxima verosimilitud de las secuencias de aminoácidos del gen CASP1. b) Posición de las sustituciones y eliminaciones de aminoácidos dentro de la secuencia de la proteína CASP1 de las especies actuales y de los ancestros comunes. Las flechas indican los ancestros de los cuáles se llevó a cabo la reconstrucción del gen.

La reconstrucción de las secuencias de ADN de PER3 de los ancestros se llevó a cabo con un árbol filogenético considerando 1160 aminoácidos correspondientes a 3480 pb (Fig. 10). Las reconstrucciones

del gen PER3 para los distintos ancestros, también tuvieron alta precisión y todas, excepto la del UAC-Odontoceti, presentaron un valor P significativo (Tabla 8). Esto significa que, al igual que el gen CASP1, la probabilidad de la ocurrencia de cada base y aminoácido dentro de las secuencias del gen PER3 es muy alta y también, que hay alta probabilidad de que las secuencias completas sean correctas.

Tabla 8. Estadísticas de la reconstrucción del estado ancestral del gen PER3. Precisión es la probabilidad promedio de la ocurrencia de cada base y aminoácido. El Valor P significativo indica que la secuencia completa del ancestro es correcta dado el modelo de sustitución. Los números en color rojo indican que el Valor P no fue significativo, por lo tanto, hay mayor incertidumbre en que la secuencia completa esté correcta.

Ancestro	Precisión	Valor P
UAC-Phocidae	0.99816	0.000407
UAC-Pinnipedia	0.99511	3.14E-09
UAC-Pinnipedia-Canidae	0.98245	6.07E-32
UAC-Carnivora	0.97809	1.71E-39
UAC-Carnivora-Perisodactyla-Cetartiodactyla	0.98016	1.53E-35
UAC-Carnivora-Cetartiodactyla	0.98032	1.98E-35
UAC-Cetartiodactyla	0.98107	7.80E-34
UAC-Cetacea	0.99907	0.02693
UAC-Odontoceti	0.99963	0.237014

En el árbol filogenético de PER3 todas las ramas presentaron un buen soporte. Se observa la divergencia de los distintos clados de especies, siendo la primera rama en surgir la de los Perisodactylos. Posteriormente, diverge el grupo de los Carnívoros donde los primeros en separarse son los felinos del género Panthera, seguidos de la familia Canidae y por último todas las especies de Pinnípedos pertenecientes a la familia Phocidae, Odobenidae y Otariidae. El grupo de los Cetáceos diverge de los cetartiodactylos y se observa claramente la separación entre las distintas familias dentro de este grupo (Fig. 10a).

Para el grupo de los Cetartiodáctilos se reconstruyó la secuencia de aminoácidos de 3 ancestros (UAC-Cetartiodactyla, UAC-Cetacea y UAC-Odontoceti). También, se reconstruyó la secuencia para el último ancestro común entre Cetartiodáctilos y Carnívoros (UAC-Carnivora-Cetartiodactyla) y para el último ancestro común entre Carnívoros, Cetartiodáctilos y Perisodáctilos (UAC-Carnivora-Perisodactyla-Cetartiodactyla). Dentro del grupo de los Carnívoros se reconstruyó la secuencia para el último ancestro común de todos los Carnívoros (UAC-Carnivora), para el último ancestro común de los Pinnípedos y los

canes (UAC-Pinnipedia-Canidae), para el último ancestro común de los Pinnípedos (UAC-Pinnipedia) y para el último ancestro común de la familia Phocidae (UAC-Phocidae). En los miembros de la familia Phocidae se observan 7 eliminaciones de aminoácidos en las posiciones 45 a 51 donde se eliminan los aminoácidos Serina (S), Arginina (R), Lisina (K), Serina (S), Histidina (H), Leucina (L) y Glutamina (Q). Cabe destacar que estas eliminaciones solo se dan en especies de la familia Phocidae y que se dan a partir del último ancestro común de esta familia, ya que el último ancestro común de los Pinnípedos si presentaba los aminoácidos. En cuanto a las sustituciones de aminoácidos en la proteína de los Pinnípedos, en la posición 112, se sustituye un Ácido Glutámico (E) por una Glicina (G), en la posición 116 una Isoleucina (I) por una Treonina (T), en la posición 682 una Leucina (L) por una Prolina (P), en la posición 683 una Serina (S) por una Glicina (G) que solo presentan las especies actuales de miembros de la familia Phocidae, en la posición 755 una Treonina (T) por una Alanina (A), en la posición 761 una Lisina (K) por una Treonina (T) y luego esa Treonina (T) por una Metionina (M) en *Halichoeres grypus* y *Phoca vitulina*, y por último, en la posición 763 una Glicina (G) por una Lisina (K). En la posición número 753 se observa una sustitución exclusiva de la familia Phocidae donde se sustituye una Arginina (R) por una Glutamina(Q) (Fig. 10b). La mayoría de estas sustituciones de aminoácidos se dan dentro del grupo de los Pinnípedos. Sin embargo, es importante destacar las 7 eliminaciones de aminoácidos dentro de la familia Phocidae, así como las distintas sustituciones exclusivas presentes dentro de esta familia. Estos resultados indican que la composición de aminoácidos de la proteína PER3 de los Pinnípedos es muy particular, debido a las diferentes sustituciones de aminoácidos que han tenido a partir de su último ancestro común. Dentro de los Pinnípedos, la proteína PER3 de los miembros de la familia Phocidae sufre una eliminación de varios aminoácidos, siendo un posible indicador de que PER3 proteína en estas especies es diferente en su composición de aminoácidos y por lo tanto, podría ser diferente en su función.

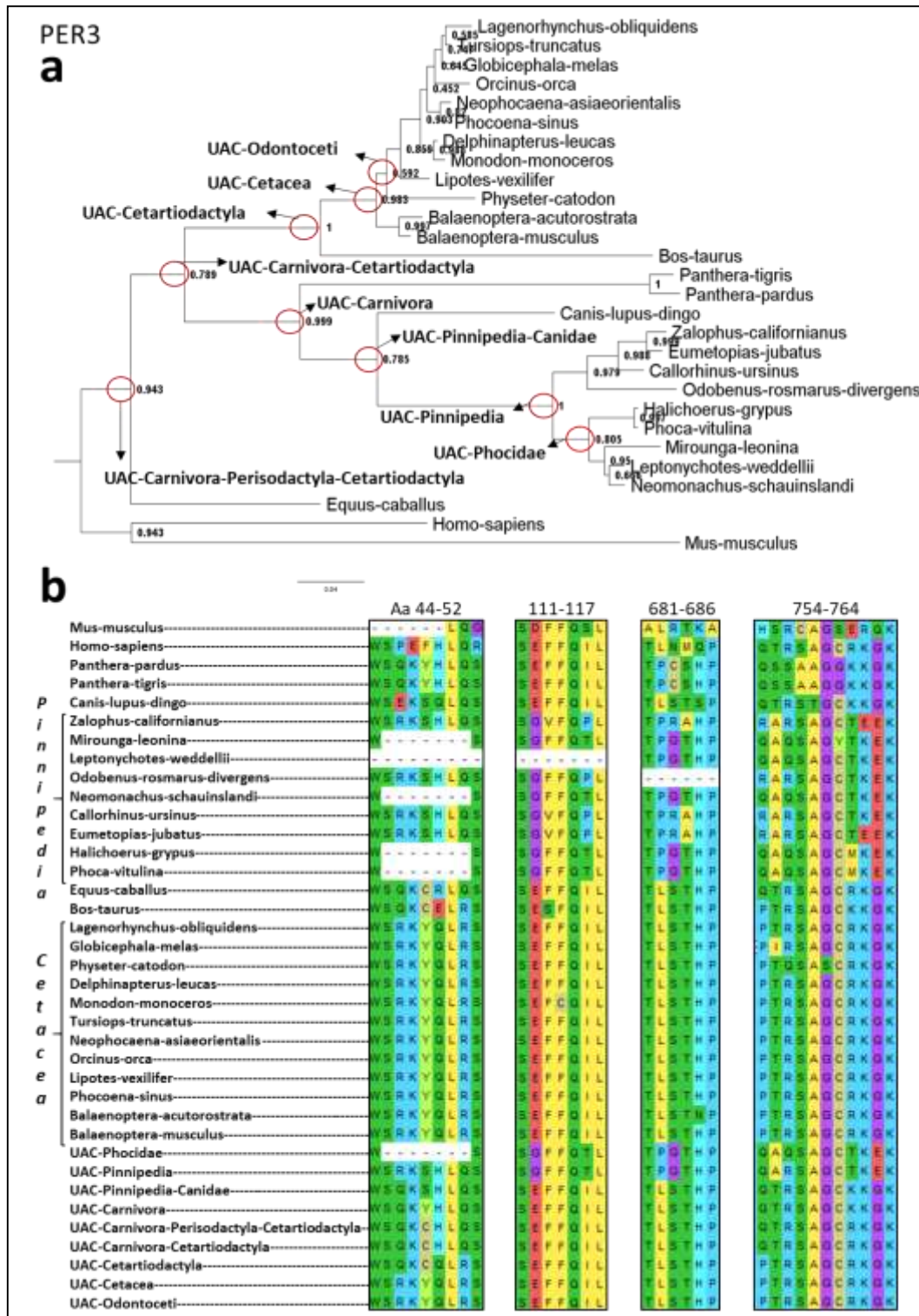


Figura 10. a) Árbol filogenético por máxima verosimilitud de las secuencias de aminoácidos del gen PER3. b) Posición de las sustituciones y eliminaciones de aminoácidos dentro de la secuencia de la proteína PER3 de las especies actuales y de los ancestros comunes. Las flechas indican los ancestros de los cuáles se llevó a cabo la reconstrucción del gen.

Capítulo 4. Discusión

4.1 Diferencias en número de genes relacionados al sueño entre Cetáceos y Pinnípedos

El objetivo general del presente estudio fue identificar los genes relacionados al sueño unihemisférico en Cetáceos y Pinnípedos. Al llevar a cabo esta identificación y filtrado de genes, se encontró que el número de genes relacionados con el sueño en general, varía considerablemente entre estos dos grupos de especies.

Al evaluar el número de genes relacionados al sueño en Cetáceos y Pinnípedos, se encontró que los Cetáceos fueron el clado donde se observa mayor cantidad de genes relacionados a esta característica, teniendo 180 genes comparado con los 92 del grupo de los Pinnípedos. Una de las maneras más comunes de incorporación de nuevos genes es la duplicación, ya sea de pequeñas regiones del DNA o de genomas completos (Otto y Yong 2002). Estos eventos de duplicación son responsables de darle una alta variabilidad a los genomas de las especies, ya que estas copias de genes pueden tener papeles importantes en la adaptación de los organismos (Otto y Yong 2002). Se ha evidenciado que estos eventos de duplicación están relacionados con el tiempo, es decir, mientras más tiempo tenga un grupo de especies en la tierra, habrá más posibilidades de que se hayan llevado a cabo eventos de duplicación en su genoma, y por lo tanto, de que se hayan producido nuevos genes (Otto y Yong 2002; Demuth *et al.* 2006; Katju y Bergthorsson 2013). Por lo tanto, es posible que las diferencias en la cantidad de genes observadas entre Cetáceos y Pinnípedos sean producto del tiempo de aparición de estos grupos. Los Cetáceos se originaron hace aproximadamente 55 millones de años y los Pinnípedos hace aproximadamente 23 ma (Nyakatura y Bininda-Emonds 2012; Zurano et al. 2019).

Los genes duplicados están sujetos a las fuerzas evolutivas y la selección natural tiene un papel importante para determinar si estos genes se fijarán o no en las poblaciones, eliminando duplicaciones deletéreas o pseudogenes y manteniendo las dúplicas favorables (Otto y Yong 2002; Kondrashov y Kondrashov 2006). También, diferentes alelos de estos genes duplicados se fijan en la población debido a que pueden presentar una ventaja selectiva, causada por el incremento en la dosis protéica que estos diferentes alelos duplicados pueden dar (Kondrashov y Kondrashov 2006). Existe evidencia de la ocurrencia de estos fenómenos a nivel genómico en distintos organismos, donde la duplicación de genes tiene un papel muy importante en el desarrollo de rasgos nuevos que tengan influencia en la adaptación. En *Drosophila* se ha

evidenciado que cuando las duplicaciones se encuentran bajo selección positiva, es probable que tengan un efecto positivo en el organismo, debido al incremento en la dosis protéica, concluyendo que, en gran medida, hay alta probabilidad de que las duplicaciones puedan conferirles un beneficio a los organismos, y que estas duplicaciones se fijen en la población por selección (Cardoso-Moreira *et al.* 2016). En cuanto a mamíferos, se ha evidenciado que existe una alta contribución de los genes duplicados en la transcripción y expresión de genes en tejidos específicos, siendo de importancia en la evolución fenotípica. De estos genes de tejidos específicos, existe una alta cantidad de duplicaciones de genes expresados en el tejido cerebral, particularmente con funciones relacionadas a la transmisión sináptica, cognición, aprendizaje y memoria (Guschanski *et al.* 2017)

Es posible que la presión selectiva que presentan los Cetáceos en los genes relacionados al desarrollo cerebral y neuronal tenga una influencia en las diferencias en la cantidad de genes observadas entre Cetáceos y Pinnípedos. Esta selección pudo haber causado que se fijaran más duplicaciones que les hayan conferido una ventaja adaptativa, y por eso, se observa mayor número de genes relacionados al sueño que en Pinnípedos.

La tasa de evolución molecular varía considerablemente entre las especies, habiendo diferencias en la tasa de mutación entre grupos taxonómicos (Martin y Palumbi 1993). Esta variación depende de diversos factores fisiológicos y demográficos, tales como el tiempo entre cada generación, tasa reproductiva, esperanza de vida y tasa metabólica (Nikolaev *et al.* 2007; Bromham 2009). Los Cetáceos, por sus características fisiológicas y demográficas, son un grupo que presenta una tasa de evolución molecular muy lenta, debido a que son organismos con una esperanza de vida muy alta, tiempos de generación largos, complejidad social y rangos de movimiento muy largos, lo que influencia en que sus patrones de diversidad genética y evolución molecular sean menores que en otros grupos de mamíferos con características contrastantes (Vachon *et al.* 2018).

Como se mencionó anteriormente, la tasa de evolución molecular en el genoma de Cetáceos es significativamente menor que en otros mamíferos y comparable con las de organismos con características similares, tales como Primates y elefantes (Martin y Palumbi 1993; Bininda-emonds 2007). Sin embargo, se ha demostrado que los Cetáceos, por la presión de selección a la que se vieron sujetos en su transición al mar, tienen partes de su genoma donde hay alta selección positiva, mayor cantidad de genes y donde se incorporaron más mutaciones de las esperadas (McGowen *et al.* 2012). Estos genes donde se ha observado alta tasa de selección son genes relacionados a procesos neurológicos y del sistema nervioso, que están involucrados en procesos de sinapsis y metabolismo relacionados con las especializaciones que

adquirieron estos organismos, tales como el buceo profundo, el metabolismo y almacenamiento de grasa (McGowen *et al.* 2012; Montgomery *et al.* 2013). Es posible que, debido a esta presión de selección en estos genes particulares, se haya encontrado una mayor cantidad de genes relacionados al sueño en este grupo que en los otros grupos de mamíferos.

El sueño es un proceso neurológico y regulado por distintas neuronas y en distintas partes del cerebro, tales como las neuronas subcorticales neuromodulatorias presentes en el tronco encefálico, cerebro medio, hipotálamo, cerebro anterior, tálamo y corteza cerebral (Aulsebrook *et al.* 2016; Eban-Rothschild *et al.* 2017). El cerebro de los Cetáceos es de los más desarrollados y se ha demostrado que en comparación con otros mamíferos terrestres, han evolucionado hasta ser uno de los grupos con mayor coeficiente de encefalización, lo que está relacionado a muchos aspectos etológicos y fisiológicos, siendo uno de estos el sueño (Hickie 1986; McGowen *et al.* 2012). Los resultados obtenidos en este estudio demuestran que hay mayor cantidad de genes relacionados al sueño en Cetáceos de lo que se esperaría por sus características de historia de vida. Como se mencionó anteriormente, el sueño unihemisférico es un proceso neurológico y en los Cetáceos existe evidencia de que existe mayor presión de selección en genes relacionados con procesos neurológicos y del desarrollo del cerebro, por lo tanto, es posible que, debido a esta presión de selección en estos genes particulares, se haya encontrado una mayor cantidad de genes relacionados al sueño en este grupo que en los otros grupos de mamíferos.

En cuanto al grupo de los Pinnípedos, el desarrollo y coeficiente de encefalización es muy similar al de mamíferos terrestres, presentando diferencias considerables con el cerebro tan desarrollado de los Cetáceos (Hickie 1986). A diferencia de éstos, los Pinnípedos son semi-acuáticos y muy diferentes a los Cetáceos en algunas características metabólicas, fenotípicas y etológicas (Park *et al.* 2018). Se ha demostrado que en Pinnípedos, al igual que en Cetáceos, existen regiones del genoma donde existe mayor presión de selección sobre ciertos genes (Park *et al.* 2018). Estos genes bajo selección positiva están más relacionados a procesos auditivos porque, a diferencia de los Cetáceos y Sirenios, los Pinnípedos tuvieron que adaptarse a la comunicación y percepción de sonidos tanto en el agua, como en la tierra (Park *et al.* 2018). Estas diferencias entre las regiones del genoma de Cetáceos (procesos neurológicos) y Pinnípedos (procesos auditivos) que están bajo selección son evidencia de las diferencias en las adaptaciones moleculares de cada grupo frente a su ambiente y estilo de vida, siendo una de las posibles razones de que en el grupo de los Pinnípedos se haya observado menor número de genes relacionados al sueño.

4.2 Genes relacionados al sueño unihemisférico

Entre los genes analizados en este estudio, el gen CASP1 y el gen PER3 fueron elegidos por ser los candidatos a posibles reguladores del sueño unihemisférico en Cetáceos y Pinnípedos. El gen CASP1 solo se observó en los Cetáceos y dentro de este grupo este gen es out-parólogo entre las especies, ya que fue producto de un evento de duplicación antes del evento de especiación en el último ancestro común. Los Cetáceos solo presentan sueño NREM y el gen CASP1 tiene un papel importante en este tipo de sueño (Lyamin *et al.* 2008; Zielinski *et al.* 2018). La proteína sintetizada por este gen actúa en la ruta metabólica del inflammasoma NLRP3 y de la interleucina 1-Beta (IL-1B). El inflammasoma NLRP3 es un complejo proteico que se forma en respuesta a cambios en el ambiente local del cerebro y al ser formado, activa a Caspasa 1 para convertir a IL-1B en su forma activa y así inducir el sueño NREM (Zielinski *et al.* 2018). Se ha demostrado que, al inhibir la expresión de alguno de los componentes de esta ruta metabólica, se reduce la calidad y duración del sueño NREM y que, al inhibir en específico la expresión de CASP1, se reduce considerablemente el tiempo en que los organismos se encuentran en el ciclo NREM. (Zielinski *et al.* 2018). La actividad de estas proteínas es mayor durante el final del ciclo NREM, en el período en el que los organismos están despertándose y también se ha evidenciado que al incrementar los periodos en los que los organismos están despiertos, se aumenta la expresión de estas proteínas (Zielinski y Krueger 2013; Zielinski *et al.* 2018). Por lo tanto, esta ruta metabólica es fundamental en la regulación del sueño NREM, que es el único que presentan los Cetáceos (Imeri *et al.* 2006). Estos resultados demuestran la importancia de CASP1 en el ciclo de sueño NREM de los Cetáceos, ya que al solo presentar este tipo de sueño requieren de una mayor influencia de esta ruta metabólica para regular sus ciclos de sueño unihemisférico.

Como se mencionó anteriormente, el análisis de reconciliación demostró que el gen CASP1 en las especies de Cetáceos es producto de un evento de duplicación previo al evento de especiación en su ancestro común. Al llevarse a cabo este evento de duplicación y que todos los Cetáceos hayan heredado este gen, se puede especular que existió una evolución paralela de CASP1 en este grupo. Se ha demostrado que la evolución paralela de genes a partir de eventos de duplicación tiene un rol fundamental en la adaptación de distintos organismos al mismo ambiente (Hughes y Friedman 2003). Estos eventos de evolución paralela, provocan que dichos genes evolucionen y se fijen en las poblaciones por selección natural, teniendo una alta similitud en la secuencia de aminoácidos, a partir de su último ancestro común (Hughes y Friedman 2003; Pickersgill 2018). Con base en este resultado se podría apoyar la hipótesis de que el sueño unihemisférico fue producto de una evolución paralela en los distintos linajes de Cetáceos (Lyamin *et al.* 2008).

Los resultados de este estudio indican la ausencia del gen CASP1 en el grupo de los Carnívoros. Se ha demostrado que en el grupo de los Carnívoros ha habido pérdida evolutiva de diferentes inflamasomas. La ruta metabólica del inflamasoma NLRP3 es muy diferente a la explicada anteriormente, ya que en los Carnívoros hay una pérdida de la Caspasa 1 y solo se encuentra una proteína que es producto de la fusión de CASP1 y CASP4 (Digby *et al.* 2020). Sin embargo, esta proteína no tiene la misma función en el reclutamiento y maduración de IL-1B para inducir al sueño NREM, además, también actúa un inflamasoma mediado por Caspasa 8 que procesa ineficientemente la IL-1B (Digby *et al.* 2020). Por lo tanto, debido a la pérdida de inflamasomas y de CASP1 en Carnívoros y a la ineficiencia de las Caspasas que poseen para inducir el sueño NREM, la ruta metabólica que podría regular el sueño unihemisférico en Otáridos probablemente no es la misma que en Cetáceos.

La reconstrucción del estado ancestral de CASP1, evidenció que los cambios en la secuencia de aminoácidos se dan a partir del último ancestro común de Cetáceos, indicando que son exclusivos de este grupo y que se dieron en el proceso de especiación de estas especies, otorgándole a estas especies una proteína con una composición de aminoácidos distintas a la de las especies de mamíferos terrestres.

El ciclo de sueño circadiano y la homeostasis del sueño están regulados en mamíferos por los genes PER, y entre estos genes, el PER3 no es requerido para mantener el ritmo circadiano como sí lo son PER1 y PER2 (Dijk y Archer 2010). Sin embargo, se ha asociado el gen PER3 con fenotipos variantes (organismos con mutaciones y deficiencia parcial del sueño) en los que los ciclos de sueño se alteran, teniendo efecto en la duración y calidad de las fases de sueño REM y NREM (Hasan *et al.* 2011; Hida *et al.* 2014). Se ha evidenciado que en individuos homocigotos de PER3, es mayor el efecto de este gen en la fase NREM que en la REM, ya que permanecen menor tiempo despiertos y mayor tiempo en la fase NREM donde tienen mayor cantidad de ondas de sueño lentas (Viola *et al.* 2007, 2008; Hasan *et al.* 2011). Los resultados del análisis de selección sobre los distintos clados demostraron que los únicos grupos del estudio donde se observa que existe selección que beneficia a los homocigotos sobre este gen PER3 son los Pinnípedos y la familia Phocidae. Se observó que las características en la secuencia y en la composición de aminoácidos de PER3 de los Pinnípedos y en específico de la familia Phocidae son muy particulares, y es posible que la selección diversificadora que actúa sobre este gen les haya conferido la adaptación a los primeros Pinnípedos y a los Otáridos para poder mantenerse en sueño NREM y de onda corta, que son las características del sueño unihemisférico.

Los Fócidos no presentan sueño unihemisférico y son la familia de Pinnípedos más reciente, originándose hace aproximadamente 11 ma. Considerando las edades de las familias y la presencia del sueño

unihemisférico solo en algunas, es posible que esta característica haya estado en el ancestro de Pinnípedos, que se haya mantenido en Otáridos, los cuales se originaron hace aproximadamente 20 ma y que luego se haya perdido en el linaje de Phócidos (Nyakatura y Bininda-Emonds 2012). En ese caso, la selección diversificadora que favorece homocigotos debió haber favorecido el desarrollo del sueño unihemisférico en los primeros Pinnípedos existentes, y posteriormente, cambiado en el grupo de los Phócidos para perder esta característica y adquirir nuevamente el sueño bihemisférico. Los resultados de la reconstrucción ancestral de este gen demostraron que la mayoría de las sustituciones de aminoácidos para este gen se dieron dentro del grupo de los Pinnípedos. Teniendo dentro de este grupo el caso particular de los Phócidos donde es importante destacar las 7 eliminaciones de aminoácidos en su secuencia, así como las distintas sustituciones exclusivas. Esto podría ser resultado de esa pérdida del sueño unihemisférico en este grupo, siendo esas las sustituciones importantes para que estas especies adquirieran de nuevo el sueño bihemisférico.

En cuanto a los Cetáceos, los resultados obtenidos demuestran que existe una relación cercana entre su PER3 y el PER3 de Otáridos, sugiriendo que este gen tiene influencia en el sueño unihemisférico en ambos grupos de especies, lo que podría indicar una posible convergencia evolutiva de este gen en estos dos grupos de especies. El PER3 de los Phócidos aún tiene similitudes con el de Otáridos, sin embargo, no tiene similitudes ni relación de ortología cercana con el de Cetáceos. Estos resultados sugieren que PER3 tiene influencia en el sueño unihemisférico de Cetáceos y Pinnípedos, y que esta proteína sufrió cambios (pérdidas de aminoácidos) en Phócidos que pudieron estar relacionados con la pérdida del sueño unihemisférico en este grupo, obligándolos a mantener únicamente el bihemisférico.

Para la hipótesis de este trabajo referente a la presión de selección en las distintas ramas clave de Cetáceos y Pinnípedos, no se pudieron llevar a cabo las comparaciones entre los distintos clados debido a que los genes encontrados no fueron los mismos para ambos grupos. Los Cetáceos además de presentar esa relación de ortología cercana de PER3 con Otáridos, presentan las características antes mencionadas en el gen CASP1. Dado que los Cetáceos son los únicos organismos con sueño unihemisférico que presentan esas características en el gen CASP, es posible que CASP1 sea el gen con mayor influencia en el desarrollo del sueño unihemisférico obligado en Cetáceos, teniendo influencia de PER3 y las rutas metabólicas relacionadas con este gen. Por su parte, en Otáridos, PER3 es un posible responsable de conferirles el sueño unihemisférico facultativo. Estos resultados sugieren que existe una posible convergencia evolutiva molecular de PER3 en Cetáceos y Otáridos probablemente con una función relevante en la característica del sueño unihemisférico. Sin embargo, para poder demostrar esta convergencia, es necesario que se

realicen estudios posteriores con base en las características funcionales de las proteínas de las especies de cada grupo.

4.3 Genes relacionados al sueño bihemisférico

El análisis realizado sobre todos los genes reportados en la literatura como relacionados al sueño permitió encontrar resultados acerca de genes relacionados al sueño bihemisférico que solo se encontraron en las especies del estudio con este tipo de sueño. Se encontró que existen dos genes, el PTGER4 y el CST3, que son compartidos en organismos que presentan sueño bihemisférico, estando ausentes en Cetáceos, único grupo del estudio que no presenta sueño bihemisférico de ningún tipo. El receptor PTGER4 es un promotor del sueño bihemisférico que actúa en la ruta metabólica de los distintos receptores de prostaglandina (Yoshida *et al.* 2000). En esta ruta metabólica el receptor E2 es el principal responsable de regular los ciclos de sueño, principalmente de mantener despierto al organismo y esto se logra mediante la activación del sistema histaminérgico por el promotor PTGER4 que activa la transcripción y activación de los receptores E4 (Yoshida *et al.* 2000; Huang *et al.* 2003, 2006, 2011). Con base en esta información y en los resultados obtenidos, es evidente que el papel del receptor PTGER4 es muy importante en los ciclos de sueño bihemisférico, activando el sistema histaminérgico y por lo tanto activando a E2, promoviendo que los organismos puedan estar alerta y mantenerse despiertos. Debido a que este promotor no se encontró en Cetáceos, se sugiere que este receptor no actúa en el sueño unihemisférico.

Por su parte, el gen CST3 codifica a la proteína Cistatina C, la cual es un inhibidor de proteasas endógenas, que tiene un papel importante en el sueño bihemisférico, disminuyendo el estado de alerta e incrementando el tiempo y frecuencia de la fase de ciclo REM (González-Rivera *et al.* 2006). Además, también se ha encontrado que esta proteína puede tener alta influencia en la temperatura corporal durante los incrementos de ciclos REM. La temperatura corporal es muy importante en la regulación de los ciclos de sueño y el efecto de la Cistatina C es disminuir la temperatura durante estos ciclos REM prolongados (González-Rivera *et al.* 2006, 2012). Se ha especulado si los Cetáceos carecen de ciclos REM o si estos tienen ciclos REM diferentes que no han sido detectados (Mascetti 2016). Estos resultados pueden ser indicio de que los Cetáceos no presentan ningún tipo de ciclo REM, debido a que, como esta proteína tiene tanta relación con la temperatura corporal y con la inducción de estos ciclos de sueño, su ausencia en Cetáceos, puede ser evidencia de que no tienen esta característica o de que, en el caso que la presenten, esté regulada por otros genes. Estos resultados respaldan la hipótesis de que el sueño

unihemisférico se desarrolló en gran parte por la necesidad de termorregular en el ambiente acuático, ya que, además de la necesidad de tener un hemisferio despierto manteniendo las funciones normales del cuerpo y permitiendo a los organismos mantener la termogénesis muscular constante, también es importante evitar ciclos de sueño REM, en donde la temperatura corporal pueda disminuir (Lyamin *et al.* 2008; Aulsebrook *et al.* 2016; Mascetti 2016).

Capítulo 5. Conclusiones y perspectivas

- Los resultados del presente trabajo demuestran que existen diferencias en el número de genes relacionados al sueño entre Cetáceos y Pinnípedos, las cuales pueden estar relacionadas con la diferente presión de selección que existe en distintas regiones del genoma de cada grupo y con el tiempo de origen. Estas regiones del genoma con alta presión de selección son para los Cetáceos, regiones del genoma relacionadas con procesos neurológicos y para Pinnípedos, regiones del genoma relacionadas a procesos auditivos.
- La reconciliación del árbol de genes con el árbol de especies permitió identificar los posibles genes reguladores del sueño unihemisférico en Cetáceos y Pinnípedos, siendo estos genes CASP1 y PER3.
- CASP1 y PER3 presentan características diferentes en los patrones de ortología y paralogía en los distintos grupos. Por lo tanto, se sugiere que PER3 es un gen que influencia el sueño unihemisférico en Cetáceos y en Otáridos y que CASP1 es el posible responsable de que los Cetáceos presenten el sueño unihemisférico de manera obligada.
- En cuanto al sueño bihemisférico, se demostró que los Cetáceos no presentan genes que tienen una gran importancia en este tipo de sueño. Por lo tanto, hay una mayor evidencia para poder inferir que estos organismos carecen completamente de este tipo de sueño.
- Para poder llegar a conclusiones más certeras acerca de la evolución de estos genes, es importante que se realicen estudios acerca de las características de estas proteínas en Cetáceos, Otáridos y Phócidos, así como en los ancestros comunes, con el fin de evidenciar las características estructurales de la proteína que les confieren estas similitudes y diferencias en el sueño unihemisférico, así como las diferentes ventajas adaptativas que estas proteínas les pudieran dar.
- Con el fin de demostrar si existe una convergencia evolutiva molecular en el gen PER3 de Cetáceos y Pinnípedos, es importante realizar comparaciones y análisis posteriores con base en las características funcionales de las proteínas.

Literatura citada

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, y D. J. Lipman. 1990. Basic local alignment search tool, *Journal of Molecular Biology*, 215(3):, pp. 403–410.
- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, y G. Sherlock. 2000. Gene Ontology: tool for the unification of biology, *Nature Genetics*, 25(1):, pp. 25–29.
- Aulsebrook, A. E., T. M. Jones, N. C. Rattenborg, T. C. Roth, y J. A. Lesku. 2016. Sleep Ecophysiology: Integrating Neuroscience and Ecology, *Trends in Ecology and Evolution*. Elsevier Ltd, 31(8):, pp. 590–599.
- Bininda-emonds, O. R. P. 2007. Fast Genes and Slow Clades : Comparative Rates of Molecular Evolution in Mammals, *Evolutionary Applications*, 3, pp. 59–85.
- Bromham, L. 2009. Why do species vary in their rate of molecular evolution ?, *Biology Letters*, (April):, pp. 401–404.
- Carbon, S., A. Ireland, C. J. Mungall, S. Shu, B. Marshall, S. Lewis, J. Lomax, C. Mungall, B. Hitz, R. Balakrishnan, M. Dolan, V. Wood, E. Hong, y P. Gaudet. 2009. AmiGO: Online access to ontology and annotation data, *Bioinformatics*, 25(2):, pp. 288–289.
- Cardoso-Moreira, M., J. R. Arguello, S. Gottipati, L. G. Harshman, J. K. Grenier, y A. G. Clark. 2016. Evidence for the fixation of gene duplications by positive selection in *Drosophila*, *Genome Research*, 26(6):, pp. 787–798.
- Casewell, N. R. ... S. T. Turvey. 2019. Solenodon genome reveals convergent evolution of venom in eulipotyphlan mammals, *Proceedings of the National Academy of Sciences of the United States of America*, 116(51):, pp. 25745–25755.
- Darriba, D., G. L. Taboada, R. Doallo, y D. Posada. 2012. JModelTest 2: More models, new heuristics and parallel computing, *Nature Methods*. Nature Publishing Group, 9(8):, p. 772.
- Demuth, J. P., T. De Bie, J. E. Stajich, N. Cristianini, y M. W. Hahn. 2006. The Evolution of Mammalian Gene Families, *PLoS ONE*, (1):, p. e85.
- Digby, Z., J. Rooney, J. P. Boyle, B. Bibo-Verdugo, R. Pickering, S. J. Webster, T. P. Monie, L. Hopkins, N. Kayagaki, G. Salvesen, S. Warming, L. Weinert, y C. E. Bryant. 2020. Evolutionary loss of inflammasomes in carnivores to facilitate carriage of zoonotic infections, *bioRxiv*.
- Dijk, D. y S. N. Archer. 2010. PERIOD3, circadian phenotypes, and sleep homeostasis, *Sleep Medicine Reviews*. Elsevier Ltd, 14(3):, pp. 151–160.
- Doolittle, R. F. 1994. Convergent evolution : the need to be explicit, *Trends Biochem. Sci*, 19(January):, pp. 15–18.

- Eban-Rothschild, A., W. J. Giardino, y L. de Lecea. 2017. To sleep or not to sleep: neuronal and ecological insights, *Current Opinion in Neurobiology*. Elsevier Ltd, 44, pp. 132–138.
- Futuyma, D. J. 2005. *Evolution*. Sinauer Associates, Inc.
- Gabaldón, T. y E. Koonin. 2013. Functional and evolutionary implications of gene orthology, *Nat Rev Genet*, 14(5);, pp. 360–366.
- Gadagkar, S. R. y M. S. Rosenberg. 2005. Inferring Species Phylogenies From Multiple Genes: Concatenated Sequence Tree Versus Consensus Gene Tree, *Journal of Experimental Zoology*, 304(B);, pp. 64–74.
- González-Rivera, R., M. Martínez-Vargas, E. Tabla-Ramon, F. Estrada-Rojo, B. Solis-Luna, A. Perez-Arredondo, y L. Navarro. 2012. Cystatin C Decreases the Body Temperature and Pain Perception in Rats, *WebMedCentral*, pp. 1–9.
- González-Rivera, R., L. Navarro, M. Martínez-Vargas, K. Guzman-Vasquez, P. León-Rosario, A. Landa, y O. Prospero-Garcia. 2006. Potential participation of cystatin C in rapid eye movement sleep (REMS) modulation, *Neuroscience Letters*, 408, pp. 178–182.
- Guschanski, K., M. Warnefors, y H. Kaessmann. 2017. The evolution of duplicate gene expression in mammalian organs, *Genome Research*, 27(9);, pp. 1461–1474.
- Hardison, R. 2003. Comparative Genomics, *PLoS Biology*, 1(2);, pp. 156–160.
- Hasan, S., D. R. Van Der Veen, R. Winsky-sommerer, D. Dijk, y S. N. Archer. 2011. Altered sleep and behavioral activity phenotypes in PER3-deficient mice, *Am J Physiol Regul Integr Comp Physiol* 301; 301, pp. 1821–1830.
- Hernandez-Rosales, M., M. Hellmuth, N. Wieseke, K. Huber, V. Moulton, y P. F. Stadler. 2012. From event-labeled gene trees to species trees, *BMC Bioinformatics*, 13(19);, p. S6.
- Hickie, J. P. 1986. Relative Brain Size in Marine Mammals, *The American Naturalist*, 128(4);, pp. 445–459.
- Hida, A., S. Kitamura, Y. Katayose, M. Kato, H. Ono, H. Kadotani, M. Uchiyama, T. Ebisawa, Y. Inoue, Y. Kamei, M. Okawa, K. Takahashi, y K. Mishima. 2014. Screening of Clock Gene Polymorphisms Demonstrates Association of a PER3 Polymorphism with Morningness –, *Scientific Reports*, 4, pp. 1–6.
- Huang, Z., Y. Sato, T. Mochizuki, T. Okada, W. Qu, A. Yamatodani, Y. Urade, y O. Hayaishi. 2003. Prostaglandin E 2 Activates the Histaminergic System via the EP 4 Receptor to Induce Wakefulness in Rats, *The Journal of Neuroscience*, 23(14);, pp. 5975–5983.
- Huang, Z., Y. Urade, y O. Hayaishi. 2006. Prostaglandins and adenosine in the regulation of sleep and wakefulness, *Current opinion in pharmacology*, 7, pp. 33–38.

- Huang, Z., Y. Urade, y O. Hayaishi. 2011. Key roles of the histaminergic system in sleep – wake regulation, *Sleep and Biological Rhythms*, 9, pp. 34–37.
- Hughes, A. L. y R. Friedman. 2003. Parallel evolution by gene duplication in the genomes of two unicellular fungi, *Genome Research*, 13(5):, pp. 794–799.
- Imeri, L., S. Bianchi, y M. R. Opp. 2006. Inhibition of caspase-1 in rat brain reduces spontaneous nonrapid eye movement sleep and nonrapid eye movement sleep enhancement induced by lipopolysaccharide, *Am J Physiol Regul Integr Comp Physiol*, 291, pp. 197–204.
- Katju, V. y U. Bergthorsson. 2013. Copy-number changes in evolution : rates , fitness effects and adaptive significance, *Frontiers in Genetics*, 4, pp. 1–12.
- Katoh, K. y D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7 : Improvements in Performance and Usability Article Fast Track, *Mol. Biol. Evol.*, 30(4):, pp. 772–780.
- Kawashima, T. 2019. Comparative and Evolutionary Genomics, *Encyclopedia of Bioinformatics and Computational Biology*, pp. 257–267.
- Kedziora, D. J., R. G. Abeysuriya, A. J. K. Phillips, y P. A. Robinson. 2012. Physiologically based quantitative modeling of unihemispheric sleep, *Journal of Theoretical Biology*. Elsevier, 314, pp. 109–119.
- Kendall-Bar, J. M., A. L. Vyssotski, L. M. Mukhametov, J. M. Siegel, y O. I. Lyamin. 2019. Eye state asymmetry during aquatic unihemispheric slow wave sleep in northern fur seals (*Callorhinus ursinus*), *PLoS ONE*, 14(5):, pp. 1–13.
- Kondrashov, F. A. y A. S. Kondrashov. 2006. Role of selection in fixation of gene duplications, *Journal of Theoretical Biology*, 239, pp. 141–151.
- Koonin, E. V. 2005. Orthologs, paralogs, and evolutionary genomics., *Annual review of genetics*, 39, pp. 309–38.
- Kumar, S., G. Stecher, M. Li, C. Knyaz, y K. Tamura. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms, *Molecular Biology and Evolution*, 35(6):, pp. 1547–1549.
- Lechner, M., S. Findeiß, L. Steiner, M. Marz, P. F. Stadler, y S. J. Prohaska. 2011. Proteinortho : Detection of (Co-) orthologs in large-scale analysis, *Bioinformatics*, 12(124):, pp. 1–9.
- Liu, Z., F. Qi, D. Xu, X. Zhou, y P. Shi. 2018. Genomic and functional evidence reveals molecular insights into the origin of echolocation in whales, *Science advances*, 4, pp. 1–9.
- Lyamin, O. I., P. O. Kosenko, S. M. Korneva, A. L. Vyssotski, L. M. Mukhametov, y J. M. Siegel. 2018. Fur Seals Suppress REM Sleep for Very Long Periods without Subsequent Rebound, *Current Biology*. Elsevier Ltd., 28(12):, pp. 2000–2005.

- Lyamin, O. I., P. R. Manger, S. H. Ridgway, L. M. Mukhametov, y J. M. Siegel. 2008. Cetacean sleep: An unusual form of mammalian sleep, *Neuroscience and Biobehavioral Reviews*, 32, pp. 1451–1484.
- Lyamin, O. I., L. M. Mukhametov, J. M. Siegel, y E. A. Nazarenko. 2002. Unihemispheric slow wave sleep and the state of the eyes in a white whale, *Behavioural Brain Research*, 129, pp. 125–129.
- Lynch, M. y J. S. Conery. 2003. The Origins of Genome Complexity, *Science*, 302(5649):, pp. 1401–1404.
- Martin, A. P. y S. R. Palumbi. 1993. Body size , metabolic rate , generation time , and the molecular clock, *Proceedings of the National Academy of Sciences*, 90(May):, pp. 4087–4091.
- Mascetti, G. G. 2016. Unihemispheric sleep and asymmetrical sleep: behavioral, neurophysiological, and functional perspectives, *Nature and Science of Sleep*, 8, pp. 221–238.
- McGowen, M. R., L. I. Grossman, y D. E. Wildman. 2012. Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown, *Proceedings of the Royal Society B: Biological Sciences*, 279(1743):, pp. 3643–3651.
- McGowen, M. R., G. Tsagkogeorga, J. Williamson, P. A. Morin, y S. J. Rossiter. 2020. Positive selection and inactivation in the vision and hearing genes of cetaceans, *Molecular Biology and Evolution*.
- Merkl, R. y R. Sterner. 2016. Ancestral protein reconstruction: Techniques and applications, *Biological Chemistry*, 397(1):, pp. 1–21.
- Montgomery, S. H., J. H. Geisler, M. R. MCGowen, C. Fox, L. Marino, y J. Gatesy. 2013. The evolutionary history of cetacean brain and body size, *Evolution*, 67(11):, pp. 3339–3353.
- Mukhametov, L. M., O. I. Lyamin, I. S. Chetyrbok, A. A. Vassilyev, y R. P. Diaz. 1992. Sleep in an Amazonian manatee, *Trichechus inunguis*, *Experientia*, 48, pp. 417–419.
- Nikolaev, S. I., J. I. Montoya-burgos, K. Popadin, L. Parand, y E. H. Margulies. 2007. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements, *Proceedings of the National Academy of Sciences*, 104(51):, pp. 20443–20448.
- Nyakatura, K. y O. R. Bininda-Emonds. 2012. Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates, *BMC Biology*, 10(1):, p. 12.
- Otto, S. P. y P. Yong. 2002. The Evolution of Gene Duplicates, *Advances in Genetics*, 46(604):, pp. 451–483.
- Park, J. Y., K. Kim, H. Sohn, H. W. Kim, Y. An, J. Kang, E. Kim, W. Kwak, C. Lee, D. Yoo, J. Jung, S. Sung, J. Yoon, y H. Kim. 2018. Deciphering the evolutionary signatures of pinnipeds using novel genome sequences : The first genomes of *Phoca largha*, *Callorhinus ursinus*, and *Eumetopias jubatus*, *Scientific Reports*, (February):, pp. 1–11.

- Pickersgill, B. 2018. Parallel vs. Convergent evolution in domestication and diversification of crops in the Americas, *Frontiers in Ecology and Evolution*, 6, pp. 1–15.
- Rattenborg, N. C., B. Voirin, S. M. Cruz, R. Tisdale, G. D. Omo, H. Lipp, M. Wikelski, y A. L. Vyssotski. 2016. Evidence that birds sleep in mid-flight, *Nature Communications*, pp. 1–9.
- Rybczynski, N., M. R. Dawson, y R. H. Tedford. 2009. A semi-aquatic Arctic mammalian carnivore from the Miocene epoch and origin of Pinnipedia, *Nature*. Nature Publishing Group, 458(7241):, pp. 1021–1024.
- Schmidt, M. H. 2014. The energy allocation function of sleep: A unifying theory of sleep, torpor, and continuous wakefulness, *Neuroscience and Biobehavioral Reviews*, 47, pp. 122–153.
- Upham, N. S., J. A. Esselstyn, y W. Jetz. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation, *PLoS Biology*.
- Vachon, F., H. Whitehead, y T. R. Frasier. 2018. What factors shape genetic diversity in cetaceans ?, *Ecology and Evolution*, (November 2017):, pp. 1–19.
- Vallender, E. J. y B. T. Lahn. 2004. Positive selection on the human genome, 13(2):, pp. 245–254.
- Viola, A. U., S. N. Archer, L. M. James, J. A. Groeger, J. C. Y. Lo, D. J. Skene, M. Von Schantz, y D. Dijk. 2007. PER3 Polymorphism Predicts Sleep Structure and Waking Performance, *Current Biology*, 17, pp. 613–618.
- Viola, A. U., L. M. James, S. N. Archer, D. Dijk, V. Au, J. Lm, A. Sn, y D. Dj. 2008. PER3 polymorphism and cardiac autonomic control : effects of sleep debt and circadian phase, *Am J Physiol Heart Circ Physiol*, 295, pp. 2156–2163.
- Vyazovskiy, V. V y A. Delogu. 2014. NREM and REM Sleep : Complementary Roles in Recovery after Wakefulness, *The Neuroscientist*, 20, pp. 203–219.
- Wei, L., Y. Liu, I. Dubchak, J. Shon, y J. Park. 2002. Comparative genomics approaches to study organism similarities and differences, *Journal of Biomedical Informatics*, 35(2):, pp. 142–150.
- Yang, Z. 2007. PAML 4 : Phylogenetic Analysis by Maximum Likelihood, *Mol. Biol. Evol.*, 24(MI):, pp. 1586–1591.
- Yoshida, Y., H. Matsumura, C. A. T. Nakajima, M. Mandai, y T. Urakami. 2000. Prostaglandin E (EP) receptor subtypes and sleep: promotion by EP4 and inhibition by EP1/EP2, *Neurophysiology*, 11(10):, pp. 2127–2131.
- Zakon, H. H. 2002. Convergent Evolution on the Molecular Level, *Brain, Behavior and Evolution*, 59, pp. 250–261.

- Zielinski, M. R., D. Gerashchenko, S. A. Karpova, V. Konanki, W. Mccarley, F. S. Sutterwala, R. E. Strecker, y R. Basheer. 2018. The NLRP3 inflammasome modulates sleep and NREM sleep delta power induced by spontaneous wakefulness, sleep deprivation and lipopolysaccharide, *Brain Behav Immun*, 62, pp. 137–150.
- Zielinski, M. R. y J. M. Krueger. 2013. Sleep and innate immunity, *Front Biosci*, (3):, pp. 632–642.
- Zurano, J. P., F. M. Magalhães, A. E. Asato, G. Silva, C. J. Bidau, D. O. Mesquita, y G. C. Costa. 2019. Cetartiodactyla: Updating a time-calibrated molecular phylogeny, *Molecular Phylogenetics and Evolution*. Elsevier, 133(May 2018):, pp. 256–262.