

1

Evolutionary sparse learning reveals the *shared genetic basis of convergent traits.*

Lucas Casanova Retamal

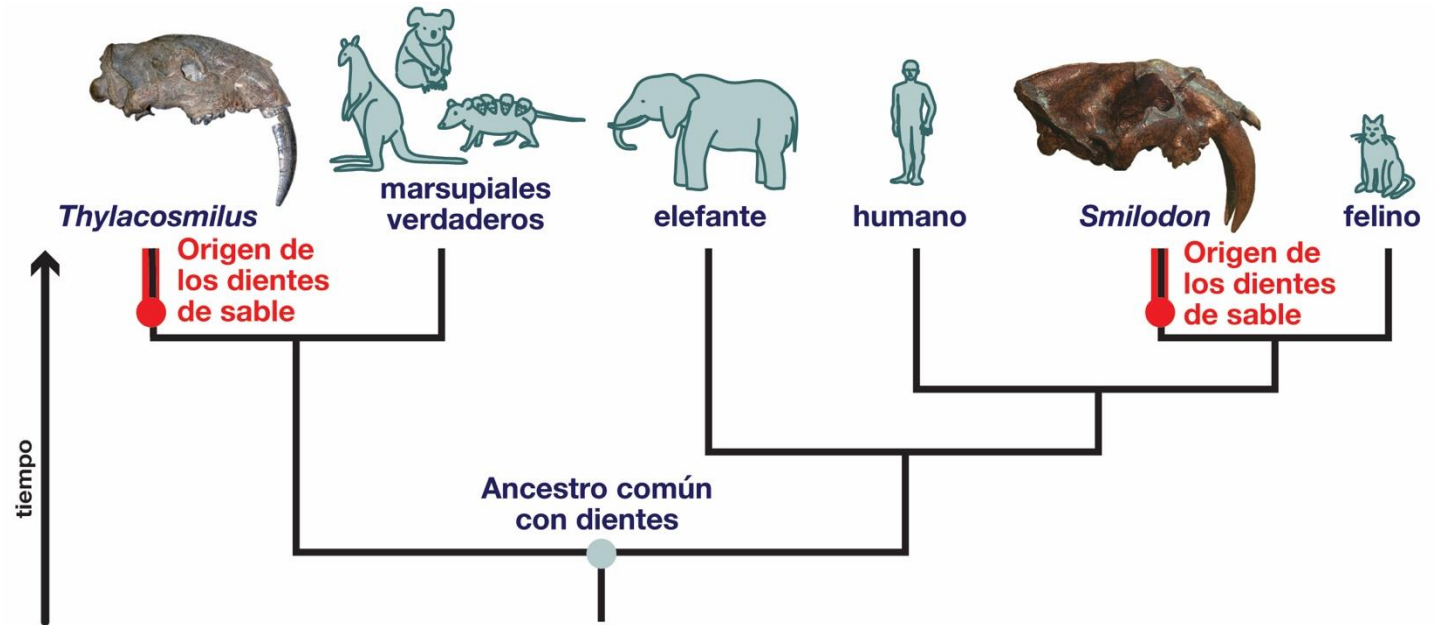
Talca, 31 de julio de 2026

1

Introducción

- Evolución convergente.
- Hipótesis.
- Base genética.

La evolución convergente ocurre cuando especies distintas desarrollan rasgos similares de forma independiente.



Hipótesis: La evolución convergente de 2 características, fotosíntesis C4 de las gramíneas y la ecolocalización en mamíferos, comparten cambios en los mismos genes e incluso en los mismos sitios de esos genes.

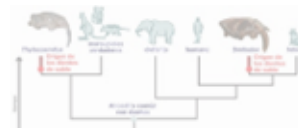
¿La naturaleza reutiliza las mismas soluciones genéticas cuando enfrenta problemas similares?

1

Introducción

- Alguien lo pensó antes.
- Falta de resultados de peso.

La evolución convergente ocurre cuando especies distintas desarrollan rasgos similares de forma independiente.



Hipótesis: La evolución convergente de características, fotosíntesis C4 de las gramíneas y la ecolocalización en mamíferos, comparten cambios en los mismos genes e incluso en los mismos sitios de esos genes.

¿La naturaleza reutiliza las mismas soluciones genéticas cuando enfrenta problemas similares?

A functional enrichment test for molecular convergent evolution finds a clear protein-coding signal in echolocating bats and whales

Amir Marcovitz, Yatish Turakhia, Heidi I. Chen ,  +3, and Gill Bejerano  [Authors Info & Affiliations](#)

Genomic and functional evidence reveals molecular insights into the origin of echolocation in whales

ZHEN LIU , FEI-YAN QI , DONG-MING XU , XIN ZHOU , AND PENG SHI  [Authors Info & Affiliations](#)

P=0.0486

SCIENCE ADVANCES • 3 Oct 2018 • Vol 4, Issue 10 • DOI: 10.1126/sciadv.aat8821

Molecular parallelism in fast-twitch muscle proteins in echolocating mammals

JUN-HOE LEE , KEVIN M. LEWIS, TIMOTHY W. MOURAL, BOGDAN KIRILENKO, BARBARA BORGONOVO, GISA PRANGE, MANFRED KOESSL, STEFAN HUGGENBERGER,

CHULHEE KANG , AND MICHAEL HILLER  [Authors Info & Affiliations](#)

1

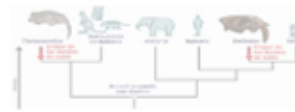
Introducción

- Explicación a la falta de resultados.

La evolución convergente ocurre cuando especies distintas desarrollan rasgos similares de forma independiente.

A functional enrichment test for molecular convergent evolution finds a clear protein-coding signal in echolocating bats and whales

[bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint \(which was not certified by peer review\) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.](#)



Genomic and functional evidence reveals molecular insights into the origin of echolocation in whales

[bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint \(which was not certified by peer review\) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.](#)

Molecular parallelism in fast-twitch muscle proteins in echolocating mammals

[bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint \(which was not certified by peer review\) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.](#)

Hipótesis: La evolución convergente de 2 características, fotosíntesis C4 de las gramíneas y la ecolocalización en mamíferos, comparten cambios en los mismos genes e incluso en los mismos sitios de esos genes.

¿La naturaleza reutiliza las mismas soluciones genéticas cuando enfrenta problemas similares?

¿Por qué ocurre esta falta de resultados

- Bases genéticas no compartidas.
- Insuficiente poder estadístico en los métodos utilizados.
- La incapacidad para excluir la convergencia no adaptativa.
- Los métodos clásicos son principalmente retrospectivos.

2

Modelo

- ESL.
- PSC.

Solución que proponen a estos problemas

Modelos genéticos predictivos.

- Evolutionary Sparse Learning (ESL):
 - Clasifica especies en función de si tienen o no una cierta característica convergente evolutiva.
- No calcula parámetros evolutivos.
- Busca concordancia entre los cambios en las secuencias con la presencia de una característica.
- No todos los genes son útiles para esta clasificación.
- Le da un peso a cada aminoácido, la mayoría es 0.

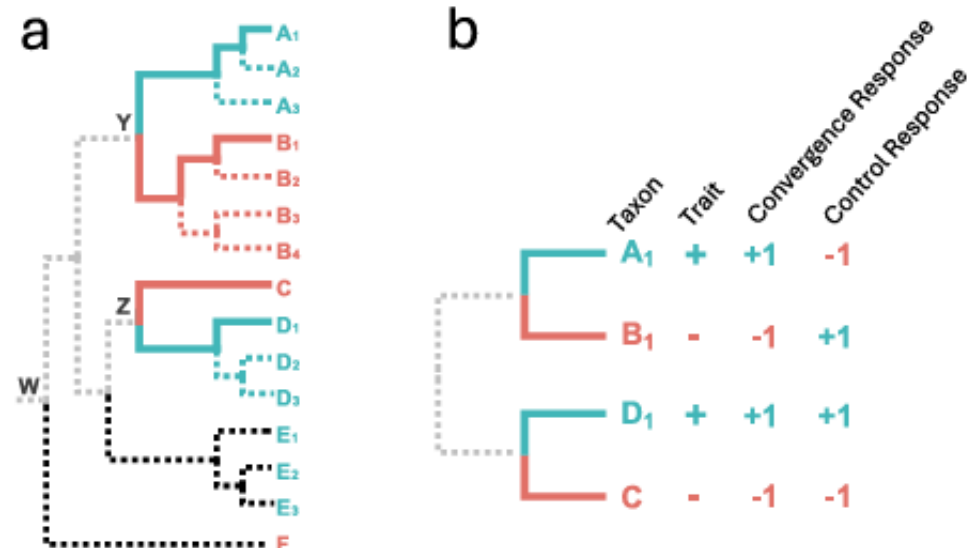
Modelo

- ESL.
- PSC.

- Paired Species Contrast (PSC):
Selecciona pares de especies, una que cuenta con la característica y la otra no.

Existen reglas de selección:

- Los pares deben ser de clados independientes evolutivamente.
- El ancestro común más reciente de cada especie seleccionada no puede ser ancestro de ninguna otra especie en el análisis.



3

Ecolocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.

La adquisición independiente de la ecolocalización en murciélagos y ballenas es uno de los casos de convergencia más estudiados.



Murciélago marrón
pequeño

Lil' Brown

Ecolocalización



Delfín nariz de
botella

3

Ecolocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.

La adquisición independiente de la ecolocalización en murciélagos y ballenas es uno de los casos de convergencia más estudiados.



Ecolocaliza



Murciélago marrón
pequeño
Lil' Brown

Delfin nariz de
botella



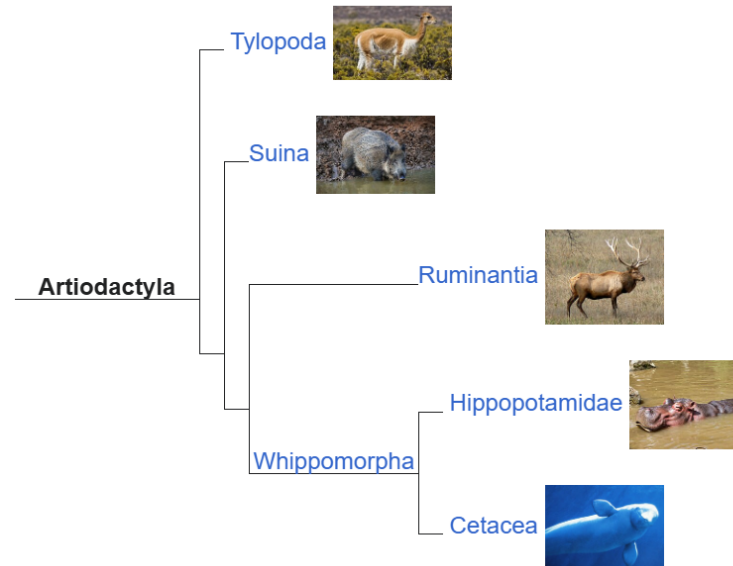
turns out to be a completely frugivorous bat a
feeds on flower nectar and pollen. The HGN de

Gran zorro volador

Sin ecolocalización

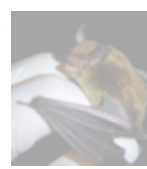


Oveja



Ecocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.



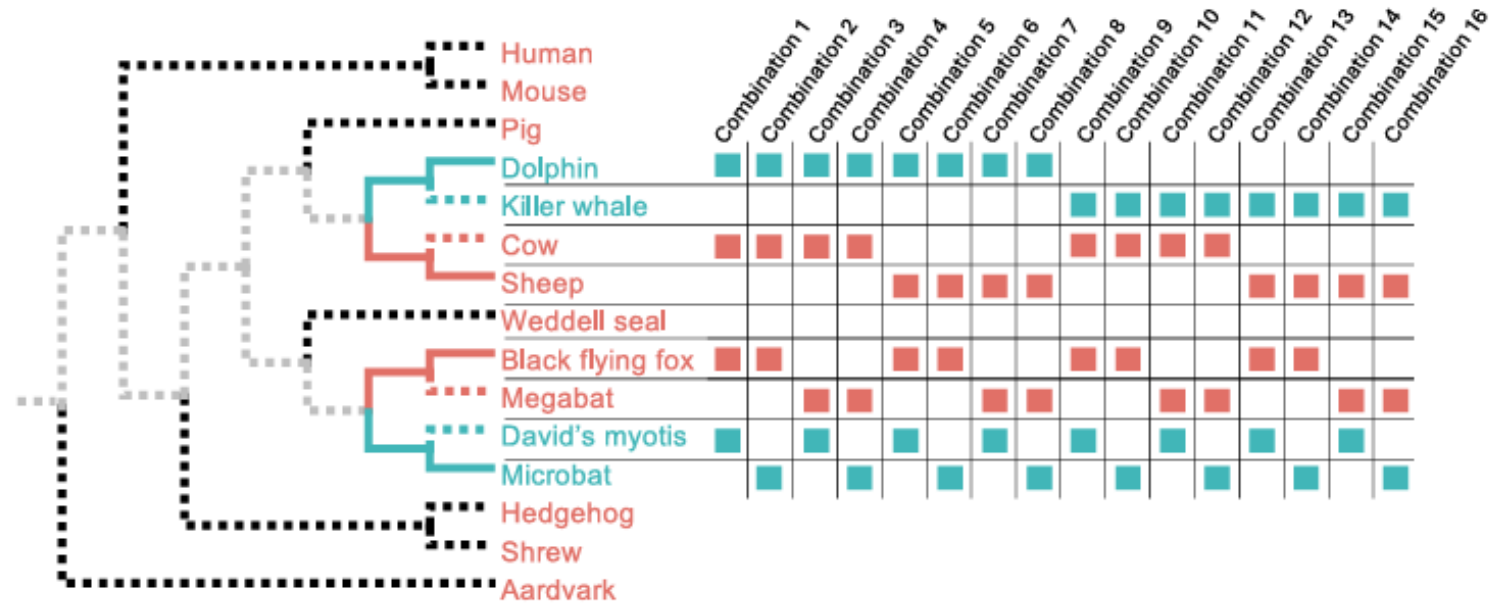
Ecocaliza



Murciélago marino
pequeño
Lil' Brown



- Analizaron 14.509 proteínas ortólogas, cada una estaba alineada entre muchas especies de mamíferos.
- Se genera un ranking de proteínas relacionadas con la característica.



3

Ecolocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.

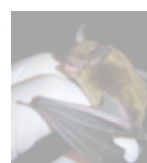


- **Prestin (SLC26A5)** → participa en la amplificación mecánica de las células ciliadas del oído interno.
- **TMC1** → canal mecanosensible esencial para la audición.
- **PJVK (DFNB59)** → asociado con sordera hereditaria.
- **CDH23** → importante para los enlaces entre estereocilios del oído.
- **CASQ1** → relacionado con el funcionamiento muscular especializado.
- **CABP2** → proteína que regula canales de calcio en células auditivas.

3

Ecolocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.



Ecolocaliza



Murciélago marrón pequeño
Lil' Brown

Delfín nariz de botella



Sin ecolocalización



Oveja

Gran zorro volador

Análisis de enriquecimiento funcional.

- Top 100 de ESL-PSC fuertemente enriquecidas en proteínas relacionadas con la percepción sensorial del sonido ($P < 10^{-4}$).
- Top 50, 150, 200.
- Comparación con otros modelos.



3

Ecolocalización

- Especies.
- Proteínas.
- Análisis.
- Predicción.



Plantearon la siguiente idea:

“Si una categoría de proteínas enriquecida ha experimentado realmente una evolución convergente en aminoácidos funcionalmente relevantes para la adquisición de un rasgo, sus secuencias deberían contener suficiente información genética para predecir con alta precisión la presencia o ausencia de dicho rasgo en otras especies.”

- 136 proteínas que entran en la categoría funcional de Cillium Movement (CCS) y realizaron predicciones, obteniendo un 75% de exactitud balanceada.
- 144 proteínas de la categoría funcional Sensory Perception of Sound (ESL-PSC), y realizaron predicciones, obteniendo un 96% de exactitud balanceada.

4

Conclusiones

¿Cómo separa este método la convergencia molecular genuina de la similitud heredada, y qué riesgos de falsos positivos persisten al buscar firmas de convergencia en todo el proteoma?

Con respecto a la separación:

- Con el diseño PSC, comparando pares de especies estrechamente emparentados, donde una posee el característica convergente y la otra no.
- Al compartir gran parte de su historia evolutiva, las mutaciones heredadas del ancestro común aparecen en ambas especies, no contribuyendo a diferencias la característica de interés.
- Se excluyen las ramas del árbol filogenético que no están relacionadas con la evolución del rasgo.

4

Conclusiones

¿Cómo separa este método la convergencia molecular genuina de la similitud heredada, y qué riesgos de falsos positivos persisten al buscar firmas de convergencia en todo el proteoma?

Con respecto a los riesgos:

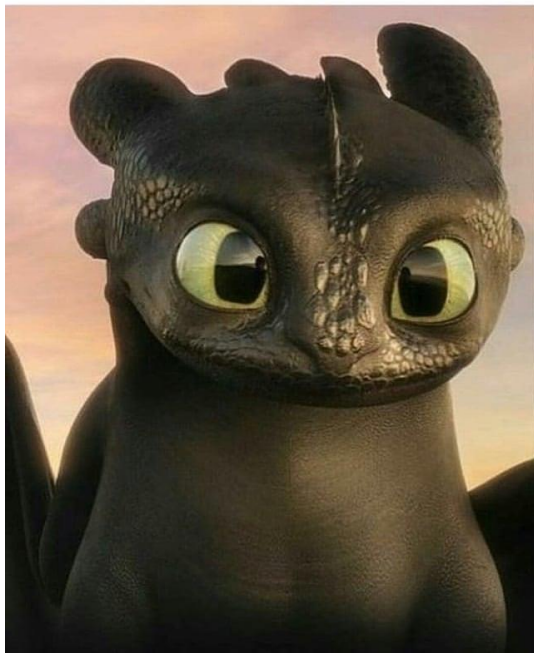
- Se mencionan casos donde proteínas de una misma “familia funcional”, pueden poseer tasas de mutaciones parecidas, haciendo que este sea un criterio de selección.
- Errores en los datos, contaminación de secuencias, errores de alineamiento, de anotación.

LOONEY TUNES



4

Personajes famosos con ecolocalización



Polémico