



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**DOCTORADO EN CIENCIAS BIOMÉDICAS
INSTITUTO DE ECOLOGÍA**

**Genómica de la diversificación de las dietas y evolución de la
nectarivoría en murciélagos Filostómidos**

T E S I S

**QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS**

PRESENTA:

YOCELYN TERESA GUTIÉRREZ GUERRERO

DIRECTOR DE TESIS:

**DR. LUIS E. EGUIARTE FRUNS
INSTITUTO DE ECOLOGÍA, UNAM**

COMITÉ TUTOR:

**DR. ENRIQUE IBARRA LACLETTE
INSTITUTO NACIONAL DE ECOLOGÍA A. C. (INECOL)
DR. RODRIGO A. MEDELLÍN LEGORRETA
INSTITUTO DE ECOLOGÍA, UNAM**

CIUDAD UNIVERSITARIA, CD. MX. DICIEMBRE 2020



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

1. Agradecimientos

2. Resumen / Abstract

3. Presentación

4. Introducción

4.1. Marcadores moleculares clásicos y genómicos

4.2 Evolución molecular en la genómica comparada

- Evolución de los genomas a partir de las familias génicas y pseudogenizaciones
- Paisaje adaptativo en genes funcionales a través de cambios en las tasas de sustitución (dN/dS)
- Cambios en las tasas de sustitución en el tiempo

4.3 Estudios genómicos en mamíferos

4.4 Avance en los estudios genómicos en Murciélagos

- Genomas completos
- Transcriptomas
- Epigenómica
- Viroma
- Estudios genómicos recientes en murciélagos

4.5 Macroevolución y diversificación de las dietas en animales

4.6 Historia evolutiva de la familia Phyllostomidae y especialización de la nectarivoría

4.7 Evolución convergente en vertebrados nectarívoros: Murciélagos y Colibríes

5. Objetivo General y Específicos

6. Capítulo I:

Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats

Yocelyn T. Gutiérrez-Guerrero¹, Enrique Ibarra-Laclette², Carlos Martínez del Río³, Josué Barrera-Redondo¹, Eria A. Rebollar⁴, Jorge Ortega⁵, Livia León-Paniagua⁶, Araxi Urrutia⁷, Erika Aguirre-Planter¹ and Luis E. Eguiarte*¹

7. Capítulo II:

How feeding habits shape the molecular evolution of sensory systems in bats

Yocelyn T. Gutiérrez-Guerrero¹, Josué Barrera-Redondo¹, Michael Nachman², Jorge Ortega³, Luis E. Eguiarte^{*1}

8. Capítulo III:

Atlas transcripcional y mecanismos de regulación de glucosa en dos especies de vertebrados nectarívoros: el murciélago tequilero (*Leptonycteris yerbabuena*) y el colibrí (*Hylocharis leucotis*)

9. Discusión general

10. Conclusiones generales

11. Perspectivas

12. Referencia citada

13. Leyendas Tablas y Figuras

14. Material Suplementario

Agradecimientos

Al Posgrado en Ciencias Biomédicas, Universidad Nacional Autónoma de México.

Al tutor principal Dr. Luis E. Eguiarte Fruns, y el comité tutorial Dr. Enrique Ibarra Laclette, Instituto de Ecología A. C., y Dr. Rodrigo Medellín, Instituto de Ecología, UNAM, por el seguimiento y apoyo durante el doctorado.

Al Consejo Nacional de Ciencia y Tecnología, CONACyT, por brindarme el apoyo económico durante mi doctorado y mis estancias. Que se abran oportunidades y se fortalezca la ciencia básica.

Beca CONACyT (CV. 580116, Folio Beca. 580116) y Beca Mixta (No. 291250) CONACyT y proyecto de Fronteras de la Ciencia: “Genómica de la Diversidad de Vertebrados Mexicanos 1: Leptonycteris y la evolución de la nectarivoría en Murciélagos y Aves”, No. 177, por el apoyo financiero, que me permitió hacer un extenso análisis en mi proyecto de doctorado y aprender nuevas herramientas teóricas y prácticas.

A los miembros del jurado, Dr. Daniel Piñero Dalmau (Presidente), Dr. Luis E. Eguiarte Fruns (Secretario), Dra. Ella Vázquez Domínguez (Vocal), Dr. Diego Cortez Quezada (Vocal) y Dr. Luis David Alcaráz (Vocal), por sus comentarios y retroalimentación de la presente tesis.

Agradecimientos a título personal

A mi madre, mi padre, mis hermanos y mi abuela, por su apoyo incondicional, empujándome a seguir avanzando y abriendo caminos. Por escucharme, cuestionarme y su interés de aprender sobre murciélagos y genómica. Por estar allí, siempre.

A mi tutor Dr. Luis E. Eguiarte Fruns, por todo el apoyo y dedicación que me ha brindado. Toda la guía para avanzar, concluir proyectos y abrir nuevos. Grandes aprendizajes me llevo de lo que significa ser un buen tutor.

Una vez más y por siempre a la Universidad Nacional Autónoma de México, UNAM, mi *alma mater*. Por darme la oportunidad de aprender, estudiar y enseñar entre sus pasillos. Y enseñarme sobre la colectividad, empatía y dedicación, para construir un país mejor.

Un especial agradecimiento al Dr. Carlos Martínez del Río, por su apoyo, enseñanzas y entusiasmo en este proyecto.

Al Dr. Enrique Ibarra Laclette, por sus apoyo y enseñanzas. Al Dr. Jorge Ortega, por todo el apoyo que me ha brindado.

Al Laboratorio de Evolución Molecular y Experimental del Instituto de Ecología, más que un espacio físico es un lugar donde he cimentado mi preparación profesional, compartiendo experiencias y aprendiendo de mis compañeros que hacen esa comunidad.

Agradezco especialmente a la Dra. Valeria Souza, la Dra. Erika Aguirre-Planter y la Dra. Laura Espinosa-Asuar. Por su apoyo técnico, teórico y ser fundamentales en avanzar y concluir mi proyecto de doctorado.

Al M. en C. Emmanuel Villafán, Mat. Ernesto Campos, Ing. Rodrigo García, por su apoyo técnico en recursos computacionales. Al cluster Huitzillin, del INECOL A. C, al servidor Keri de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), y al servidor Patung del Laboratorio Nacional de Ciencias de la Sustentabilidad (LANCIS), Instituto de Ecología, por los recursos computacionales.

Agradezco al equipo que apoyaron la salida a campo, a la cueva el El Salitre, al M. en C. Roberto Trejo-Salazar, Dra. Osiris Gaona, Biol. Ana Galicia.

Un agradecimiento especial para el Dr. Daniel I. Piñero Dalmau, por su apoyo en la elaboración de propuestas y comentarios que han enriquecido esta tesis. Agradezco a la Dra. Nadia Santini (Lab. Ecología Evolutiva de las Interacciones Bióticas) y al Dr. Daniel Piñero, por el apoyo financiero para la secuenciación y publicación del artículo (Capítulo I).

A mis compañeros y amigos del Laboratorio. Al Dr. Josue Barrera-Redondo, por ser como un hermano en la ciencia. A la Dra. Gabriela Castellanos, por su constancia e interés en mi formación profesional y personal. Al M. en C.. Guillermo Sánchez, Dra. Eria Rebollar, Dra. Valerie de Anda, Dr. Jaime Gasca, Dr. Enrique Sheinvar y Dr. Jonás Aguirre, por todas sus enseñanzas. A la Señora Silvia Barrientos, por preocuparse en mantenernos en un lugar

agradable y su ayuda consiguiendo reactivos. A mis compañeros y todos los miembros del lab e Instituto de Ecología.

Un especial agradecimiento a Rodrigo Callado, Lilí Mata, Noé Pacheco, Laura Figueroa, Felicitas Lagunes, Karina Mendoza, Manuel Otero, Alejandro Benítez, Alejandra Moreno, Manuel Guerrero y Mago Vizuet. Gracias por su apoyo, escucharme y los alegres momentos.

Resumen

La evolución de nuevos hábitos alimenticios es uno de los eventos más importantes en la historia natural de muchos grupos de animales, ya que implica nuevas oportunidades ecológicas, innovación en la adquisición de nutrientes, estrategias de forrajeo y eventos de diversificación y radiación de especies. Dentro de los murciélagos, la mayoría de las especies tienen una dieta insectívora. Sin embargo, dos familias presentan dietas diferentes, la familia Pteropodidae, murciélagos frugívoros del Viejo Mundo y la familia Phyllostomidae del continente Americano, que exhiben un diverso espectro de hábitos alimenticios los cuales influyen en la evolución de la familia. En particular, algunas especies de filostómidos han evolucionado una especialización alimenticia, acompañada de adaptaciones a nivel morfológico y fisiológico, especialmente en los vampiros y en las especies nectarívoras. En esta tesis se secuenció el genoma completo de cinco especies de murciélagos filostómidos: el insectívoro *Macrotus waterhousii*, el frugívoro *Artibeus jamaicensis* y tres nectarívoros *Leptonycteris yerbabuenae*, *L. nivalis* y *Musonycteris harrisoni*. Se utilizó un enfoque de genómica comparada para investigar los cambios genómicos asociados con las transiciones y especializaciones de la dieta. Adicionalmente, se incluyó la información genómica disponible en bases de datos de otros murciélagos, entre ellos el vampiro común (*Desmodus rotundus*) y otros mamíferos. Se analizaron los cambios genómicos asociados con las transiciones en la dieta, describiendo la evolución de familias génicas y genes bajo selección positiva que son importantes en rutas metabólicas. Se encontraron duplicaciones en familias génicas y señales de adaptación en genes relevantes en el consumo de alimento. En los murciélagos especialistas, el vampiro presentó señales de selección positiva en genes que participan en mecanismos de coagulación y los murciélagos nectarívoros mostraron fuertes señales de selección positiva en genes involucrados en el metabolismo de azúcares, lípidos y almacenamiento de hierro. Se detectaron señales de selección y sustituciones radicales de amino ácidos paralelas en tres genes (*AACS*, *ALKBH7*, *UNC-45 B*), cambios exclusivos para los filostómidos nectarívoros y los murciélagos pteropodidos. El cambio de dieta está acompañado de nuevos comportamientos y mecanismos de forrajeo, por lo que analizamos la evolución de receptores sensoriales de olfato y sabor, y su relación con el hábito alimenticio. La mayoría de los receptores sensoriales responden a diferentes fuerzas evolutivas, y no siempre están correlacionados con la dieta. Los receptores olfativos, han presentado eventos de pseudogenización, sin embargo, el receptor *OR51E2*, muestra señales de selección positiva para los murciélagos frugívoros. Los murciélagos omnívoros, frugívoros y nectarívoros pueden detectar sabores dulces, mientras que los insectívoros tienen presiones de selección para identificar sabores amargos. Se secuenció el RNA de diferentes tejidos de *L. yerbabuenae*, y del colibrí *Hylocharis leucotis*. Con estos transcriptomas e incluyendo la información transcripcional de hígado de especies de aves y murciélagos insectívoros, se analizaron patrones génicos de

expresión diferencial y su posible relación en una dieta dominante en azúcares. Observamos niveles de expresión similares y altos en genes que participan en el metabolismo del azúcar y grasas en hígado para ambos nectarívoros, en comparación con otras especies de murciélagos y aves insectívoras. También, se identificaron diferencias en la regulación de glucosa dependiente de mecanismos de señalización de insulina, entre aves y murciélagos.

Abstract

The shift into new food habits is considered one of the most important events in the natural history evolution for several vertebrate species because it involves new ecological challenges, innovations in the nutrient acquisition strategies, and it has been accompanied by diversification events and radiation of species. In the Chiroptera order, most of the living species are insectivorous. However, two families exhibit dietary changes, the Old World fruit-bats from the Pteropodidae family, and the Phyllostomidae family (Leaf-nosed bats), which display a diverse spectrum of feeding behaviors, which underpin the evolution of the family. Interestingly, some Phyllostomid species have evolved extreme dietary specialization, accompanied by morphological and physiological adaptations, especially in vampire and the nectar-pollen feeder species. We sequenced the whole genome of five Phyllostomid bat species: the insect-feeder *Macrotus waterhousii*, the fruit-feeder *Artibeus jamaicensis*, and three nectar-pollen feeders *Leptonycteris yerbabuena*, *L. nivalis* and *Musonycteris harrisoni*. We performed genomic comparative approaches to investigate the genomic features associated with dietary transitions and feeding specialization. We included in the analyses the genomic public information from other bat species, such as the vampire *Desmodus rotundus*, and other mammalian species. We then traced the genomic changes associated with dietary shifts, analyzing the evolution of gene families, and identifying genes under positive selection. We found duplication in gene families and adaptation signatures in relevant genes involved in food intake. Among the specialist bats, the vampire exhibited positive selection in genes linked to coagulation mechanisms, whereas the nectar-feeder bats showed a strong positive selection in genes involved in sugar, lipid, and iron metabolic pathways. Additionally, we detected three genes (*AACS*, *ALKBH7*, *UNC-45 B*) showing positive selection and parallel amino-acid substitutions in conserved positions in the nectar-pollen feeders Leaf-nosed bats, and Pteropodid bats. Dietary shifts represent new challenges in feeding behaviour and foraging strategies. We analyzed the evolution of the olfactory and taste sensory gene receptors, and their association with feeding behavior. We observed that most of the chemosensory receptors are shaped by different evolutionary forces. We identified a high number of pseudogenes, but most of the olfactory receptors were functional. We observed that the *OR51E2* showed strong positive selection in frugivore bats, and an acceleration in the substitution rates in the vampire lineage, which could be due to relax purifying selection. Taste receptors of sugar perception were intact in frugivore and nectarivore bat species, but non-functional in the vampire and in some insect-feeder lineages. Finally, we sequenced the RNAseq of different tissues from two specialist nectar-pollen feeder vertebrates: the bat *L. yerbabuena* and the hummingbird *H. leucotis*. We used a transcriptomic approach to analyze patterns of gene expression involved in a diet dominated by high sugar intake. We found that both species exhibited high expression levels of genes involved in sugar and fat metabolic pathways in the

liver, relative to other insectivorous bat and bird species. Moreover, we observed changes in the transcriptomic expression involved in the signaling insulin-dependent pathway between avian and bats.

Presentación

Esta tesis se centra en el estudio de la evolución y especialización de las dietas en murciélagos, utilizando información de genómica comparada. El objetivo general fue secuenciar y ensamblar *de novo* los genomas de cinco murciélagos de la familia Phyllostomidae importantes para la biodiversidad mexicana, incluyendo especies endémicas o clave para los ecosistemas: *Macrotus waterhousii*, *Artibeus jamaicensis*, *Leptonycteris yerbabuena*, *Leptonycteris nivalis* y *Musonycteris harrisoni*. La elección de estas especies de filostómidos es relevante para el estudio de evolución de las dietas, ya que nos permite tener un espectro de los hábitos alimenticios dentro de la familia. *M. waterhousii* tiene una dieta insectívora y es un murciélagos que pertenece al grupo basal de los filostómidos, dentro de la subfamilia Macrotinae. *A. jamaicensis*, es un murciélagos frugívoro, que también puede consumir semillas e insectos. Los murciélagos *L. yerbabuena*, *L. nivalis* y *M. harrisoni*, tienen una dieta nectarívora que incluye también el consumo de polen, estas especies pertenecen a la subfamilia Glossophaginae. A partir de la información genómica y utilizando herramientas bioinformáticas, estadísticas y evolutivas, se estudió la evolución de las dietas y especialización de la nectarivoría.

La tesis inicia con una Introducción General, seguida por tres capítulos que abordan los objetivos principales del proyecto de doctorado, los cuales se presentan en formato de artículo científico, una sección de Discusión, las Conclusiones Generales y finalmente las Perspectivas que derivan del estudio.

En la Introducción General se proporciona información para justificar los objetivos del proyecto de doctorado. En esta sección se expone un panorama general de los estudios genómicos que se han realizado en mamíferos y cómo han ayudado las herramientas de secuenciación masiva (o de nueva generación) y los alcances que se han tenido. Posteriormente, se presenta una revisión más detallada de los avances en los estudios genómicos en murciélagos, que van desde la genómica y la transcriptómica, hasta la metagenómica para el análisis de virus y microbiomas. Finalmente, se abordan estudios evolutivos de la familia Phyllostomidae y la especialización de las dietas, haciendo particular énfasis en los murciélagos nectarívoros.

El Primer Capítulo (*Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats*) incluye los resultados principales sobre los genomas completos secuenciados, ensamblados y anotados de cinco murciélagos filostómidos. También incluye análisis detallados de genómica comparada, haciendo hincapié en los cambios genéticos involucrados en diferentes rutas metabólicas relevantes para sus diferentes hábitos alimenticios. El objetivo de este Primer Capítulo fue analizar los cambios genómicos relacionados con las transiciones y especializaciones del hábito alimenticio, haciendo especial énfasis en la evolución de la nectarivoría. Se incluyó adicionalmente información genómica de

otros murciélagos y mamíferos disponible en bases de datos. Se realizaron análisis de expansión y contracción de familias génicas, pruebas de selección, reconstrucción ancestral de secuencias, identificación de sustituciones paralelas en secuencias de amino ácidos y el modelado de proteínas.

En el Segundo Capítulo (*Feeding behaviour or less is more: Adaptive signatures and molecular substitutions rates in olfactory and taste sensory gene receptors across Chiroptera order*) el objetivo fue analizar la evolución de los sistemas sensoriales y su relación con el hábito alimenticio en 17 especies de murciélagos (representando cinco familias). Sabemos que las transiciones en el tipo de dieta involucran cambios a nivel metabólico, pero también adaptaciones y nuevas estrategias de forrajeo. Este capítulo se enfocó en estudiar la dinámica evolutiva de los receptores sensoriales olfativos y de sabor. Se realizó un “minado” de datos genómicos para diferenciar genes funcionales y no funcionales. Se hizo una identificación de ortólogos de copia única de receptores sensoriales. A partir de estos genes ortólogos, se hicieron pruebas de selección y se estimó las tasas de sustituciones moleculares y los tiempos de divergencia entre los diferentes grupos de murciélagos analizados.

El Tercer Capítulo (Paralelismos en patrones de expresión transcripcional en dos especies de vertebrados nectarívoros: el murciélago magueyero menor *Leptonycteris yerbabuenae* y el colibrí de oreja blanca *Hylocharis leucotis*) tuvo como objetivo principal identificar genes relevantes para una dieta dominada por azúcares, como es la nectarivoría. Estudiamos a dos vertebrados nectarívoros especialistas: el murciélago *L. yerbabuenae* y el colibrí *H. leucotis*. La especialización del hábito nectarívoro ha estado acompañada de adaptaciones morfológicas, fisiológicas y ecológicas que han sido similares en estas dos especies de vertebrados nectarívoros, a pesar de que son especies filogenéticamente distantes. Este capítulo empleó un enfoque transcriptómico donde, se secuenció, ensambló *de novo* y anotó el transcriptoma de diferentes tejidos del murciélago magueyero menor (cerebro, riñón, páncreas, hígado y pulmón), y del colibrí de oreja blanca (cerebro, lengua, corazón, páncreas, hígado y pulmón). Se realizaron análisis de expresión diferencial, enriquecimiento de genes por tejido y se identificaron los genes que se expresan paralelamente en ambas especies. Adicionalmente, se incorporó información de RNAseq de otras especies insectívoras de murciélagos (*Mormoops megalophylla*) y del pájaro mosquero (*Ficedula albicollis*), que presentan una dieta principalmente insectívora completada en menor proporción por frutas.

En la sección de Discusión y Conclusiones Generales se evalúan las ideas derivadas de los resultados obtenidos en los tres capítulos de acuerdo con el objetivo general de este proyecto. Finalmente, se incluye una sección de Perspectivas donde se proponen futuros objetivos e hipótesis que serán útiles para continuar con el estudio de la evolución de las dietas y entender otros mecanismos involucrados en la especialización de una dieta nectarívora en diferentes especies de vertebrados.

Introducción

Marcadores moleculares clásicos y genómicos

A partir de los años 70's inició un gran avance en estudios entorno a la genética y evolución (Templeton, 1992), gracias al desarrollo y creciente uso de los marcadores moleculares (Al-Samarai y Al-Kazaz, 2015) y las técnicas de amplificación del DNA, como la reacción en cadena de la polimerasa (PCR, por sus siglas en inglés).

La información genética derivada de marcadores moleculares obtenidos por la fragmentación del DNA por enzimas de restricción o por la amplificación selectiva de regiones del DNA, combinada con la teoría de evolución molecular, de la genética de poblaciones y métodos estadísticos han permitido explorar diferentes preguntas relacionadas a la evolución y ecología de las poblaciones. Dentro de estos estudios podemos mencionar en particular los análisis de la variación a partir de polimorfismos y de la biodiversidad generada por eventos de especiación, que involucran estudios de filogeografía, de análisis de evolución molecular y de reconstrucción de filogenias moleculares. Y especialmente, el estudio de los procesos evolutivos dentro de las poblaciones mediados por las fuerzas evolutivas y mecanismos genéticos, como la recombinación y rearrreglos cromosómicos (Schlötterer, 2004; Arif et al., 2011).

Si bien las investigaciones con marcadores moleculares tradicionales han permitido avanzar en estudios tanto evolutivos, ecológicos, de conservación, biomédicos y de mejoramiento genético, una gran cantidad de preguntas siguen abiertas, principalmente debido a la limitante que representa en el reducido número de fragmentos o loci que se analizan y al tipo de marcador molecular (neutral o codificante) (Ouborg et al., 2010; Sorkheh et al., 2017).

En los últimos 20 años, gracias al desarrollo de las nuevas tecnologías de secuenciación paralela masiva, también llamadas de “nueva generación” (NGS, por sus siglas en inglés) (Slatko et al., 2018; Kumar et al., 2019) y al avance en el desarrollo de herramientas bioinformáticas (Magi et al., 2010) se ha generado una gran cantidad de datos provenientes de genomas completos o transcriptomas, incluyendo miles de fragmentos genómicos de individuos de diferentes poblaciones (Ouborg et al., 2010; Sorkheh et al., 2017). También se ha obtenido información genómica sobre las comunidades virales y bacterianas que co-habitan en un ecosistema, como puede ser tapetes microbianos o microbiomas intestinales o de la epidermis (Ouborg et al., 2010; Pereira et al., 2017; Cao et al., 2017).

La información obtenida usando las plataformas NGS ha permitido abordar preguntas relacionadas con los procesos evolutivos de la historia natural de las especies, poblaciones, comunidades y ecosistemas. También han sido útiles para el estudio de las dinámicas moleculares en los patrones de expresión y regulación, bajo diferentes condiciones ambientales o entre diferentes tipos celulares (Anamika et al., 2016) para analizar patrones de plasticidad que son resultado de las interacciones genómicas y ambientales (Meaburn y Schulz, 2012) y

recientemente para analizar modificaciones generadas a partir de la edición genética con la técnica CRISPR (Yau y Rana, 2017).

Uno de los resultados más importantes de esta transición de la genética a la genómica, es un mayor conocimiento de los mecanismos genómicos y dinámicas evolutivas en especies modelo, como por ejemplo el humano, el ratón, el gusano *C. elegans*, *Arabidopsis thaliana* o *E. coli*. Pero también han sido útiles en especies no-modelo, donde los estudios genéticos estaban limitados a la información disponible de marcadores moleculares previamente desarrollados para la especie o especies filogenéticamente cercanas (Ekblom y Galindo, 2010) o por la posibilidad de tener especies cultivables, en caso de microorganismos (Dheilly et al., 2014). En particular, los estudios genómicos derivados del uso de las herramientas NGS en mamíferos no-modelo, han tenido un importante avance (Fig. 1).

Evolución molecular en la genómica comparada

Evolución de los genomas a partir de las familias génicas y pseudogenizaciones

Uno de los principales enfoques para estudiar la evolución de los genomas entre diferentes linajes son los análisis de la evolución de las familias génicas. Las familias génicas son grupos de genes que tienen un ancestro en común, por lo que comparten un alto porcentaje de identidad a nivel de secuencia de nucleótidos y/o amino ácidos (Liberles & Dittmar, 2008; Demuth & Hahn, 2009) lo que se traduce en una estructura similar que confiere en cierto grado una función parecida. La evolución de las familias esta sujeta principalmente a eventos de duplicación de genes o regiones genómicas completas, pero también hay otros mecanismos que tienen un impacto en la ganancia o pérdida de genes, como son los procesos de especiación y de transferencia horizontal (Demuth & Hahn, 2009; Ohta 2010). La ganancia (expansiones) o pérdidas (contracciones) de genes cambia a lo largo del tiempo evolutivo y es diferente para cada linaje, generalmente los cambios en el número de genes de una familia génica es resultado del azar y de la selección natural (Wang et al., 2018). Sin embargo, se han identificado tasas similares de ganancias y pérdidas de genes entre levaduras, mamíferos y moscas (0.0020, 0.0016 y 0.0012 ganancias y pérdidas de genes por millón de años, respectivamente) (Hahn et al., 2005; Demuth et al., 2006).

Para los genomas eucariontes, Ohta (2010) propone que la ganancia y pérdida de genes puede analizarse con un modelo de genética de poblaciones, asumiendo que cualquier gen miembro de una familia génica puede cambiar aleatoriamente la tasa de ganancia y/o pérdida de otro gen y estará sujeta a diferentes fuerzas evolutivas. Los eventos de duplicación son aleatorios y pueden ocurrir en cualquier región del genoma, donde la fracción de genoma duplicado varía de tamaño. Pueden ser pequeñas regiones de unos cuantos nucleótidos o grandes regiones que contienen genes codificantes (Gu et al., 2002; Demuth & Hahn, 2009). De esta manera se generan nuevas copias funcionales, las cuales pueden acumular mutaciones y estarán bajo presiones de selección y deriva (Ohta 2010, 2013). Ohta (2010) propone que hay varios tipos de duplicaciones de familias génicas: aquellas que están integradas por un número uniforme de genes o las que varían en el número de miembros, donde la homología de las secuencias es mantenida por procesos de un desigual *cross-over* (cruzamiento) durante la meiosis, por conversión génica o por la amplificación de un segmento del cromosoma que contiene genes codificantes (llamado en inglés *salutory replication*). En principio la conversión génica entre los miembros de genes de una familia puede verse como un tipo de mutación que puede ser deletérea o ventajosa (Ohta 2010). En eucariontes se propone que la rápida expansión de familias génicas está sujeta a presiones de selección adaptativa en las nuevas copias generadas (Demuth et al., 2006). De acuerdo con Demuth y Hahn (2009) los modelos de duplicación de genes involucran cambios adaptativos en el número de copias que se expanden,

donde la selección actúa en el número de copias que se expandieron y para una fracción de miembros de genes de familia actúa a nivel de secuencia de nucleótidos.

Por otra parte, las presiones evolutivas responsables de la contracción de familias y pérdida de genes son el resultado de mutaciones sin sentido (*nonsense*, por sus siglas en inglés) durante la fijación ya que estas mutaciones no son deletéreas cuando las presiones de selección están relajadas (Demuth & Hahn 2009; Wang et al., 2018). Dentro de un contexto adaptativo, la pérdida de genes puede ser explicada por cambios ambientales que promueven que un gen a la larga deje de ser esencial para la adecuación (*fitness*) de un organismo (Liberles & Dittmar 2008). Una estrategia que se ha utilizado para estudiar la pérdida de genes es cuantificar el número de pseudogenes, ya que puede ser una medida directa del impacto en la pérdida de genes en el tamaño de una familia génica (Hahn et al., 2007).

De acuerdo a lo anterior, la evolución de las familias génicas está estrechamente ligada a eventos de pseudogenización (Demuth & Hahn 2009). Los pseudogenes están presentes en un amplio número de especies, incluidas procariontes y eucariontes. Su abundancia en los genomas está generalmente sujeta a las tasas de duplicación y pérdida de genes (Gilbert et al., 1997). Dentro de los vertebrados, los mamíferos son el grupo que tienen un mayor número de pseudogenes con más de 8,000 detectados (Zhang et al., 2014), mientras que en *Drosophila* se han identificado alrededor de 20 pseudogenes producto de retrotransposones (Harrison et al., 2003).

La pseudogenización es resultado de eventos de duplicación y pérdida de genes, donde un gen adquiere y acumula mutaciones espontáneas, ya sea en el marco de lectura o introduciendo codones de paro prematuros afectando la función original (Mahmudi et al., 2015). La mayoría de los pseudogenes sin ninguna función evolucionan bajo neutralidad. Sin embargo, la información de genomas completos ha permitido identificar que algunos pseudogenes están bajo presiones evolutivas (Lynch & Force 2000; Petrov & Harti 2000) con una mayor frecuencia de mutaciones sinónimas que las mutaciones no sinónimas, por ejemplo, en los dedos de zinc que forman parte del complejo mayor de histocompatibilidad II (Sudbrack et al., 2003). También se han encontrado pseudogenes que están conservados a nivel de secuencia entre diferentes linajes, por ejemplo, pseudogenes ancestrales que tienen en común el humano y el ratón y que se originaron antes del evento de especiación (Mahmudi et al., 2015).

En vertebrados, los receptores sensoriales comprenden un amplio número de familias génicas y son un buen ejemplo para entender la evolución de las familias génicas en distintos linajes (Niimura & Nei 2007). En particular, los receptores olfativos (OR) son la familia génica más grande en cuanto a número de familias y número de genes identificadas en los genomas de mamíferos (Dong et al., 2009). Los OR son de gran interés, ya que han estado sujetos a eventos de expansiones, contracciones y pseudogenizaciones. Niimura y Nei (2007) encontraron una relación entre el número de OR funcionales, OR pseudogenes y la alta o poca sensibilidad del

olfato. Detectaron que los humanos tienen un pequeño número de OR funcionales y una alta abundancia de OR pseudogenizados en comparación con los roedores y los perros. Por otra parte, en un contexto adaptativo, en primates se ha identificado una pérdida de decenas de familias de OR, que coincide con la expansión de genes de opsinas que les confieren la visión tricromática (Gilad et al., 2007). Los OR son un buen modelo para estudiar las bases genéticas en la percepción del olor a lo largo del tiempo evolutivo en diferentes linajes. Sin embargo, los procesos de cómo y cuándo pseudogenes específicos se han originado y cómo han evolucionado en las familias génicas de OR requieren un análisis de mayor profundidad.

El análisis de expansión y contracción de familias génicas utiliza métodos que estudian la evolución de las familias génicas (Hahn et al., 2005), utilizando herramientas estadísticas y probabilísticas con la finalidad de cuantificar los cambios observados en el tamaño y número de familias de genes (Hahn et al., 2005; De Bie et al., 2006). Estos análisis requieren de información filogenética y de generar clusters de familias de genes ortólogos y parálogos (De Bie et al., 2006). Estos análisis integran dos elementos importantes para entender la evolución de los clusters de familias de genes: el nacimiento (*birth*) y muerte (*death*) de las familias génicas (Demuth & Hahn 2009). El nacimiento de familias de genes, se da a partir de los genes huérfanos (*orphan genes*), que tienen un origen a partir de duplicaciones de genes pre-existentes que tienen una neo-funcionalización, o a partir de transferencia horizontal (en bacterias y virus), y/o de genes que se originan *de novo* de secuencias previas no codificantes (Lynch & Force 2000; Zhang et al., 2003). En contraste, la muerte de una familia génica es resultado de la pérdida de genes que termina en la extinción de la familia génica completa (Hahn et al., 2016; Demuth & Hahn 2009). Algunas rutas que explican cómo una familia génica puede extinguirse, por delección o pseudogenización de todos los miembros de la familia, o puede retribuirse a un aceleramiento en las tasas de sustitución de los miembros de la familia (Lynch & Force 2000; Ohta et al., 2013).

La calidad de la información genómica es crucial en el estudio de expansiones, contracciones, así como en el *birth* y *death* de familias génicas, ya que errores en el ensamblaje, predicción de genes y errores en la anotación, frecuentemente generan falsos positivos y falsos negativos en los análisis de evolución de las familias génicas.

Paisaje adaptativo en genes funcionales a través de cambios en las tasas de sustitución (dN/dS)

Como se mencionó anteriormente, las presiones de selección son cruciales para entender la evolución de familias génicas o específicos genes entre los distintos linajes.

La aparición de una nueva mutación en el genoma de un individuo puede seguir dos caminos evolutivos, uno de ellos es distribuirse en todos los individuos de una población, aumentando su frecuencia y fijándose en el tiempo (Gilbert et al., 1997; Kosiol et al., 2008). El

segundo camino, implica la pérdida de la nueva mutación que va a estar sujeta a presiones de selección y deriva génica (Ohta 2010). Una de las formas de analizar las presiones de selección en el genoma entre distintos linajes es a partir de los genes que codifican para proteínas, estimando el ratio de las sustituciones no sinónimas (dN) entre las sustituciones sinónimas (dS) (Kosiol & Anisimova 2019). Las mutaciones sinónimas en su mayoría son neutrales y tienen tasa de sustitución neutral $\mu_S = \mu$, ya que no cambia el amino ácido en la secuencia de proteína y no se ve afectada la función (Sharp et al., 2005; Yang et al., 2008; Jeffares et al., 2015). Mientras que las mutaciones no sinónimas pueden estar bajo presiones de selección, y tienen una tasa de sustitución diferente a la neutral $\mu_N \neq \mu$ (Yang et al., 2008; Jeffares et al., 2015). La mayoría de los cambios no sinónimos afectan negativamente las regiones codificantes, alterando la estructura y función de las proteínas (Jeffares et al., 2015; Kosiol & Anisimova 2019). Por lo tanto, el cociente $\omega = dN / dS$ permite cuantificar la magnitud de la fuerza de la selección natural que está actuando a nivel funcional en el genoma, donde $\omega > 1$ indica selección positiva o adaptativa, $\omega = 1$ selección neutral y $\omega < 1$ indica selección purificadora (Yang et al., 2008; dos Reis & Wernish 2009; Jeffares et al., 2015; Kosiol & Anisimova 2019). El cambio adaptativo a nivel de secuencia de amino ácidos no necesariamente se traduce en un cambio en la función de la proteína, sino que puede reflejar un pico en el paisaje adaptativo con la optimización de la función de la proteína para ciertos linajes o como respuesta a cambios ambientales (dos Reis & Wernish 2009).

Para detectar selección y estimar el cociente ω se requiere información de secuencias de genes ortólogos de copia única, la filogenia correspondiente del gen, métodos probabilísticos de Máximo Likelihood (ML) o Bayesianos (Jeffares et al., 2015). También se utilizan modelos probabilísticos de sustitución, que explican la evolución de la secuencia. Los principales supuestos de estos modelos son: i) futuros cambios en la secuencia dependen del estado actual; ii) cada sitio en una secuencia evoluciona independientemente de otros sitios; y iii) un estado puede cambiar a cualquier otro (Bofkin & Goldman 2007; Kosiol & Anisimova 2019).

Una de las ventajas de los modelos basados en amino ácidos (*codon models*), es que permiten analizar a mayor profundidad la divergencia de los genes que están evolucionando rápidamente y no están influenciados por el efecto del sesgo en el uso de codones (*codon bias*) (Jeffares et al., 2015; Kosiol & Anisimova 2019). Los *codon models* permiten distinguir los cambios no sinónimos de los sinónimos, ya que durante episodios de evolución adaptativa, solo una pequeña fracción de sitios en una proteína tienen la capacidad de incrementar su fitness a través de cambios no sinónimos (Kosiol et al., 2017). Por lo anterior, estos modelos incorporan diferentes escenarios de variación en las presiones de selección, donde se ponen a prueba dos modelos: un modelo restrictivo que no permite la selección positiva ($\omega \leq 1$, hipótesis nula) y otro modelo que permite la selección positiva ($\omega \geq 1$, hipótesis alternativa). La selección positiva es detectada si el modelo $\omega > 1$, es significativamente mejor comparado con el modelo restrictivo ω

≤ 1 (Yang & Nielsen 2000; Koisol et al., 2017; Kosiol & Anisimova 2019). Para ello, es necesaria la construcción del *Likelihood Ratio Test* (LTR) que permite comparar la probabilidad entre las dos hipótesis y detectar aceleraciones en las tasas de sustituciones en el linaje de interés (Anisimova et al., 2001; Jeffares et al., 2015; Kosiol & Anisimova 2019).

La detección de selección adaptativa puede analizarse bajo tres escenarios, uno son los *branch-models*, se basan en identificar presiones de selección positiva en las ramas de una filogenia, donde ω puede variar entre los diferentes linajes de una filogenia; los *site-models*, donde los valores de ω los pueden variar a lo largo de las posiciones del gen; y los *branch-site models*, los valores ω varían en sitios específicos del gen, solo para ciertas ramas (Kosakovsky et al., 2005; Yang et al., 2007; Jeffares et al., 2015; Kosiol & Anisimova 2019).

Con la información de los genes anotados a partir de los genomas secuenciados, se han tenido avances importantes en el estudio e identificación de genes bajo presiones de selección y entender cómo evoluciona la parte funcional de los genomas en diferentes especies. Asimismo, existen diferentes programas para detectar presiones de selección y estimar ω , como son CODEML de la paquetería PAML, HYPHY, entre otros (Kosakovsky et al., 2005; Yang et al., 2007). Sin embargo, la calidad de las anotaciones de los genes y proteínas puede ser un factor limitante en la identificación de ortólogos de copia única y en la estimación de las tasas de sustitución. Es por ello que se recomienda revisar las secuencias ortólogas, descartando aquellas secuencias incompletas o con errores en su ensamble, para evitar obtener falsos positivos de selección positiva en sitios o ramas (Jeffares et al., 2015; Gutiérrez-Guerrero YT et al., 2020).

Cambios en las tasas de sustitución en el tiempo

El reloj molecular es una herramienta que permite estimar el tiempo de eventos filogenéticos, las tasas de especiación y extinción, el tiempo de eventos de dispersión y datar el origen de familias génicas (Kumar 2005; Ho 2008).

La hipótesis del reloj molecular supone que las secuencias de DNA y proteínas evolucionan a una tasa que es relativamente constante a lo largo del tiempo entre los linajes (Kumar 2005; Ho et al., 2016). Esta hipótesis, fue propuesta hace más de 50 años por Zuckerkandl y Pauling (1962). A partir de observaciones empíricas estimaron la tasa de evolución de la hemoglobina, contando las diferencias en las secuencias de amino ácidos entre el humano y caballo y suponiendo, que la tasa es constante entre los linajes, infirieron los tiempos de divergencia de la hemoglobina (~100 - 160 millones de años) (Gillespie 1991). Posteriormente, la teoría neutral de evolución molecular propuesta por Kimura demostró que la tasa de mutaciones neutrales que llegan a fijarse en una población (tasa de sustitución) es equivalente a la tasa en la que aparecen las nuevas mutaciones en cada miembro de una población (i.e. la tasa de mutación) (Kimura 1965, 1968, 1983). El reloj molecular supone una

tasa constante a lo largo del tiempo, una tasa homogénea a través de los linajes, y que se tiene buena información tanto taxonómica como del registro fósil (Kimura 1983; Huttley et al., 2007; Ho et al., 2016) Sin embargo, las tasas de evolución molecular están influenciadas por diferentes factores, como por ejemplo el tiempo generacional y la eficiencia en los mecanismos de reparación del DNA (Huttley et al., 2007; Ho et al., 2016), esto hace que las tasas sean significativamente diferentes entre las especies, por ello han propuesto dos modelos de reloj molecular. El primer modelo, llamado reloj estricto considera que aunque las tasas varían en el tiempo y entre especies hay un promedio de esta variación. El segundo modelo, llamado reloj relajado, supone que las tasas cambian, es decir evolucionan en el tiempo y considera que hay factores biológicos (como la tasa metabólica), que influyen en la evolución (Martin & Palumbi 1993; Ho et al., 2016).

Una de las consideraciones más importantes es la calibración del reloj molecular, ya sea utilizando un reloj estricto o relajado. Para calibrar el reloj molecular se debe de conocer el tiempo absoluto de algún evento de divergencia, por ejemplo el tiempo en el que se separan los linajes de primates y humano (Sarich & Wilson 1969). La estimación de los tiempos de divergencia se hace a partir de la información del registro fósil o utiliza los tiempos de eventos geológicos, como la formación de montañas (Kumar et al., 2005).

Con los avances en los métodos de secuenciación y el crecimiento en el número de genomas y genes de especies modelo y no-modelo, la hipótesis del reloj molecular ha resultado ser una robusta herramienta evolutiva que ha permitido estudiar a mayor profundidad la relación de los cambios genéticos entre linajes en una escala de tiempo (Huttley et al., 2007).

Estudios genómicos en mamíferos

Los mamíferos comprenden aproximadamente el 9% del total de especies de vertebrados (Koepfli et al., 2015) y se estima que actualmente existen más de 5,416 especies de mamíferos conocidos en la Tierra (Jones y Safi, 2011). En este momento, se tiene estudios de índole genómico y provenientes de *NGS* para mamíferos en aproximadamente 300 diferentes especies que van del ornitorrinco hasta el humano, las cuales incluye tanto especies modelo como no-modelo (Fig. 1) (Koepfli et al., 2015). A la fecha, en bases de datos como *National Center for Biotechnology Information* (NCBI, por sus siglas en inglés), se encuentran depositados alrededor de 3,000 genomas de animales, de los cuales más de 1,000 son de mamíferos (Fig. 2a). Por otra parte, en la base de datos de *Sequence Read Achievement*, del NCBI (SRA, por sus siglas en inglés), se encuentran albergados un total de 3,173,334 archivos proveniente de diferentes plataformas de secuenciación *NGS*, principalmente de la plataforma Illumina (Fig. 2b), donde los estudios de tipo genómico son los más abundantes (Fig. 2b, c, d).

Existe una diferencia importante en la profundidad y cantidad de estudios genómicos en mamíferos, ya que hay numerosas investigaciones para especies modelo y un menor número para las especies no-modelo. De acuerdo con la información en bases de datos, hay datos genómicos disponibles para dos especies monotremas (*Ornithorhynchus anatinus* y *Tachyglossus aculeatus*). Para la subclase Theria se tiene información de 15 órdenes, que comprenden 39 familias, para más de 200 especies de mamíferos (Fig. 1). Los ordenes para los que hay mayor información genómica son Quiróptera, Primates, Carnívora, Cetacea, y Artiodactyla. Sin embargo, el número de genomas secuenciados, representa un porcentaje muy bajo en relación a la riqueza de especies de cada grupo.

En particular para especies de mamíferos modelo, el genoma humano, fue el primer genoma de mamífero secuenciado (International Human Genome Sequencing Consortium, 2001). El proyecto inicial para la secuenciación del genoma humano surge en 1990 a partir del proyecto de investigación: “*Understanding Our Genetic Inheritance: The Human Genome Project, The First Five Years, FY 1991-1995*”, promovido por el *National Center for Human Genome Research* (NCHGR, por sus siglas en inglés). Este primer genoma de tamaño relativamente grande revolucionó las técnicas de secuenciación y bioinformáticas, con un avance en la construcción y armado de ensamblajes de genomas de tamaño grande a partir de la teoría de Euler y Bruijn (Compeau et al., 2011).

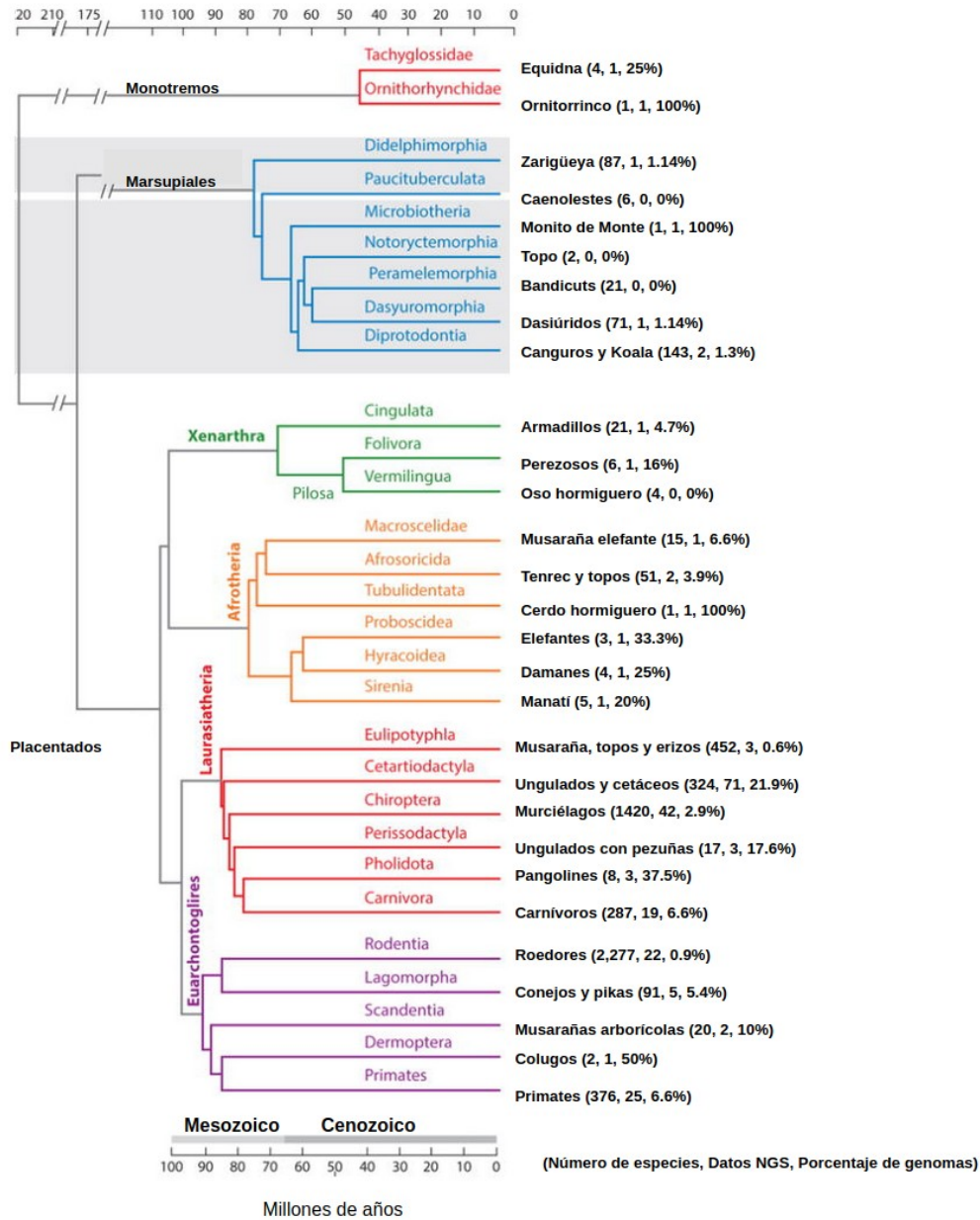


Figura 1.

Filogenia que comprende los órdenes de la clase Mamalia, donde se muestra por orden, el número de especies registradas, el número de datos genómicos con NGS disponibles en bases de datos y el porcentaje de datos genómicos en relación al número de especies descritas. Figura tomada y modificada de Koepfli et al., 2015.

Se tiene conocimiento de la arquitectura genómica de las 3,000,000,000 pares de bases (3Gbp) que constituyen el genoma humano y se estima que el genoma está compuesto por ~100,000 genes, de ellos entre 25 - 28,000 son genes que codifican para proteínas (International Human Genome Sequencing Consortium, 2014). También existe desde hace tiempo información

a detalle de sus mapas cromosómicos (Cohen et al., 1993) de sus rearrreglos genómicos (Tsai y Lieber, 2010) y han descrito sus rutas y redes de regulación génicas y transcripcionales que participan en diferentes procesos fisiológicos. Por ejemplo, se ha analizado la aparición de canas, donde por medio de estudios de RNA se identificaron 112 genes con expresión diferencial asociados con rutas de pigmentación (Bian et al., 2019). Y en particular existe amplia información sobre los genes y transcritos involucrados en el desarrollo del cáncer donde a partir de la secuenciación de 2,158 diferentes tumores cancerígenos, se identificaron mutaciones que pueden acelerar la evolución del tumor (*The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium*, 2020). Otro tipo de estudios son las modificaciones en la composición del microbioma, inducida por cambios fisiológicos como por ejemplo el estrés o embarazo (Edwards et al., 2017), o sobre la composición del microbioma relacionado a vivir en zonas urbanas o rurales (Tyakht et al., 2013).

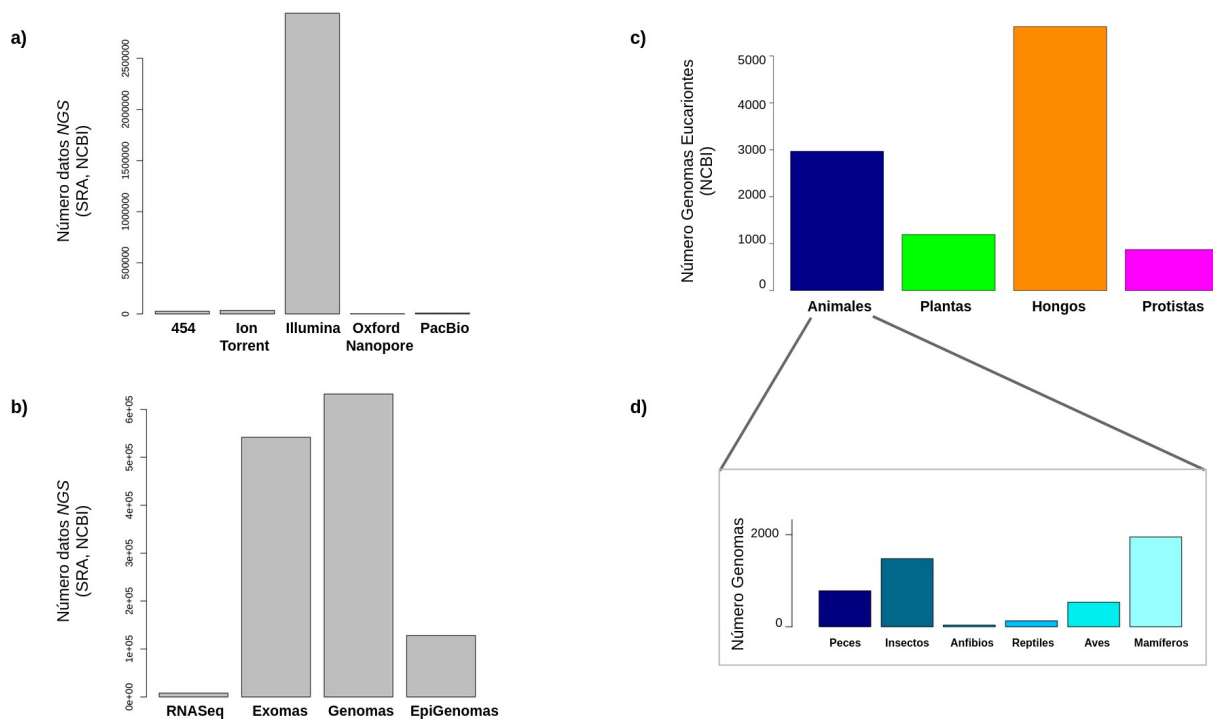


Figura 2. Datos genómicos de mamíferos obtenidos con herramientas NGS, disponibles en la plataforma de *National Center for Biotechnology Information* (NCBI). a) Número de datos para mamíferos depositados en la base *Sequence Read Achievement* (SRA), con diferentes plataformas de secuenciación. b) Número de datos para mamíferos depositados en SRA para diferentes herramientas genómicas. c) y d) Número de genomas a nivel contig, scaffold y cromosoma para animales depositados en la base de NCBI. Información consultada 02-2020.

A la fecha se tienen más de 100,000 genomas de diferentes individuos de distintas razas de poblaciones humanas bajo proyectos como 1000 Genomes Project (Devuyst, 2015), 100, 000 Genomes Project (Turnbull et al., 2018), PopHuman (Casillas et al., 2017), entre otros.

A partir de la generación de estos datos se han obtenido importantes resultados. Por ejemplo, se ha avanzado en entender los mecanismos genéticos de la esquizofrenia con estudios de asociación (GWAS, por sus siglas en inglés), donde se identificaron 13 locus relacionados con el trastorno bipolar y señales de inducción de calcio (Ripke et al., 2013). Un estudio clásico es el de las adaptaciones que permiten a las poblaciones humanas vivir a grandes elevaciones, como las poblaciones del Tibet, que presentan mutaciones únicas en 7 loci fuertemente asociadas con fenotipos de hemoglobinas y homocisteínas (Yang et al., 2017).

Para la mayoría de los mamíferos, que son especies no-modelo y no pueden manipularse, ni mantenerse en el laboratorio, hay un reducido número de datos provenientes de NGS. Sin embargo, el estudio de sus genomas nos ha permitido expandir el entendimiento de las bases genéticas de la adaptación, de los mecanismos evolutivos involucrados en la radiación de especies y sobre su organización genómica (Allendorf et al., 2010; McMahon et al., 2014).

Entre los estudios más recientes para estos mamíferos no-modelo podemos señalar al de Chein et al. (2019), donde secuenciaron los genomas de 44 diferentes especies del orden Artiodactyla, incorporando también la información transcriptómica de 516 muestras provenientes de 68 tejidos. Realizaron una reconstrucción filogenómica con un total de 51 especies, obteniendo una mejor resolución para entender las radiaciones para orden. Asimismo, detectaron una reducción en los tamaños efectivos poblacionales de muchas especies entre ~100,000 a 50,000 años, que coincide con un incremento en la población humana y sus migraciones. Por otro lado, elementos regulatorios importantes en el desarrollo del cráneo (formación de cuernos o astas), locomoción y sistema digestivo (Wang et al., 2019).

También existen importantes avances en estudios genómicos para mamíferos marinos. Árnason et al. (2018) encontraron una rápida radiación en ballenas que coincide con un periodo de cambios en la circulación oceánica (10.5 a 7.5 millones de años) y con procesos de especiación que ocurrieron bajo flujo génico, identificando señales de *incomplete lineage sorting*. Por otra parte, estudios de evolución de familias génicas encuentran importantes cambios en el número de genes en familias involucradas en procesos de respiración y almacenamiento de lípidos (Foote et al., 2015).

Por otra parte, se ha obtenido evidencia genómica de la domesticación del perro a partir de su ancestro, el lobo. Se han identificado dos procesos de domesticación independientes entre los perros del Este de Asia y el Oeste Euroasiático a partir de poblaciones de lobo genéticamente diferenciadas utilizando tanto información genómica poblacional como datos arqueológicos (Botigué et al., 2017; Frantz et al., 2016). Recientemente, Plassais y colaboradores (2019), con información de 722 genomas de perros y herramientas de asociación

GWAS, identificaron genes y regiones genómicas bajo selección, que están fuertemente relacionadas con 16 fenotipos, que explican el 90% de la variación en el tamaño del cuerpo.

Otras investigaciones están involucradas en la evolución del sistema inmune, como el estudio genómico en dos especies de pangolín. A partir de esta información genómica, se identificó una pseudogenización en el Interferon Epsilon (*IFNE*). En ratones con deficiencia de *IFNE* en células epiteliales, se ha observado un incremento en la susceptibilidad a infecciones. Sin embargo, en los genomas de los pangolines se identificó una posible compensación, gracias a la expansión de otros genes como la familia septin (relacionada en rutas de inmunidad) y adaptaciones morfológicas de la piel que caracteriza a esta especie, la cual puede ser un mecanismo de defensa (Choo et al., 2016). Estos estudios genómicos son relevantes para entender los procesos, adaptaciones y mecanismos del sistema inmune en mamíferos y pueden permitirnos elaborar herramientas para el manejo de epidemias, como el caso del virus SARS-Cov2, causante de la enfermedad COVID19.

Con la información genómica disponible en bases de datos se han realizado estudios de genómica comparada, por ejemplo, explorar a lo largo de los genomas las tasas de sustitución para 5,669 genes en 112 especies (Kumar y Subramanian, 2002). Otro tipo de análisis de genómica comparada, con información de RNASeq han permitido avanzar en el estudio de convergencias moleculares en ciertos genes que codifican a proteínas y que presentan cambios radicales en sustituciones de amino-ácidos para ballenas y murciélagos en genes involucrados en la locomoción (Lee et al., 2018).

Tener acceso a información de miles o millones de fragmentos del genoma (codificante y no codificante) ha proporcionado suficiente resolución para estudiar con mayor detalle fuerzas evolutivas como la endogamia, la deriva génica, el flujo génico y adaptaciones para especies no-modelo (Ekblom y Wolf, 2014). Estos resultados han permitido elaborar estrategias para programas de conservación, sobre todo para aquellas especies que se encuentran en alguna categoría de riesgo (Ouborg et al., 2010).

Los principales retos para los estudios genómicos en mamíferos son el desarrollo de las herramientas bioinformáticas eficientes y prácticas que pueden manejar la cantidad masiva de información que representan sus grandes genomas (que van de 2 a 4.5 Gbp) y la secuenciación de muestras que permitan tener una buena calidad, cantidad y cobertura del genoma; afortunadamente el desarrollo de estas herramientas bioinformáticas y genómicas, continúan avanzando. Estos datos nos van a proporcionar información para entender finamente las dinámicas evolutivas y moleculares de los genomas y nos ayudaran al estudio de la biodiversidad, permitiendo tener un detallado registro del archivo evolutivo que alberga toda la información genómica. También serán críticos para el desarrollo de programas de conservación de genes, poblaciones y especies.

Avance en los estudios genómicos en murciélagos

Los murciélagos (orden Chiroptera), son el segundo grupo de mamíferos más diversos (después de los roedores) con 18 familias y más de 1,400 especies descritas (Jones et al., 2005; Wu et al., 2018) distribuidas en todo el mundo (excepto en la Antártida) (Simmons, 2005). La capacidad de volar es una característica que los hace particularmente únicos dentro de los mamíferos. El vuelo controlado y activo en vertebrados ha evolucionado tres veces en la historia de la vida: en el linaje de los *Archaeopteryx* (hoy extinto), en las aves y en los murciélagos, lo que les ha permitido una mayor dispersión y colonización de diferentes nichos (Rayner, 1988). Por otra parte, la ecolocalización (o biosonar) es una característica de los murciélagos, aunque no es exclusiva del grupo (Jones y Telling, 2006) ha sido fundamental para entender la evolución y el éxito del orden.

El papel ecológico de muchas especies de murciélagos es fundamental en diversos ecosistemas donde son considerados controladores de plagas (Kunz et al., 2011; Kolkert et al., 2020). Los murciélagos con hábitos frugívoros o nectarívoros son fundamentales en la ecología de muchas comunidades vegetales al promover la polinización cruzada (Muscarella y Fleming, 2007) y también ayudar en actividades económicas como la producción de frutas incluyendo el durián en Asia y la fecundación de las plantas usadas para la producción de los diferentes mezcales en México (Havens y Vitt, 2016). Por otra parte, los murciélagos son considerados reservorios virales, siendo hospederos de un amplio número de virus, como pueden ser los ebolavirus ó el Sars-Cov2 (Calisher et al., 2006).

Los murciélagos son los mamíferos con tamaño de genoma más pequeños y con el menor número de cromosomas, con un rango de tamaño de genoma que va de 1.59 Gbp a 2.1 Gbp y $2n=16$ a 62 (en promedio para mamíferos 3.4 Gbp, $2n=60$) (Gregory, 2005; Smith y Gregory, 2009). Cantrell et al. (2008), proponen que la reducción del tamaño genómico está asociado a la pérdida de la familia de elementos móviles *LINE 1*, el elemento nuclear más largo y el más común en mamíferos (Cantrell et al., 2008). Gregory (2005) y Organ y Shedlock (2009) sugieren que la reducción del tamaño del genoma en murciélagos se debe a los daños al material genético que genera las especies de oxígeno reactivo (ROS), producto a su vez de las altas demandas metabólicas que requiere el vuelo (Organ y Shedlock, 2009). También se ha identificado una asociación entre tamaños genómicos pequeños y tamaño de las células. Un menor tamaño de las células representa una mayor área de superficie relativa al volumen de las células, permitiendo un eficiente intercambio de gases para satisfacer las demandas metabólicas (Szarski, 1983).

A la fecha, para el orden Quiróptera se tienen diferentes avances importantes en la obtención de información genómica con herramientas NGS, como se detallará a continuación.

Genomas completos de murciélagos

Se tiene información genómica de 43 especies de murciélagos, representando nueve familias (Fig. 3) que han sido secuenciadas utilizando plataformas de Illumina, PacBio y BioNano. Los tamaños de genomas reportados van desde los 1.73 Gbp a 2.79 Gbp. De estos 43 genomas depositados en NCBI, pero sólo 18 de ellos tienen disponible la anotación de los genomas, donde se reportan más de 17,000 genes que codifican a proteínas (Seim et al., 2013; Zhang et al., 2012; Parker et al., 2013; Pavlovich et al., 2018; Eckalbar et al., 2016; Ledesma-Mendoza et al., 2018; Jebb et al., 2019; Broad Institute, 2019; Gutiérrez-Guerrero et al., 2020).

Están disponibles los genomas de ocho especies de murciélagos frugívoros del Viejo Mundo pertenecientes a la familia Pteropodidae (8 genomas): *Pteropus alecto* (110 x), *Pteropus vampyrus* (2.65 x), *Pteropus giganteus* (63 x), *Rousettus aegyptiacus* (169.2 x), *Eonycteris spelea* (80 x), *Macroglossus sobrinus* (52.7 x), *Eidolon helvum* (18 x), y *Cynopterus brachyotis* (108 x). Dentro de las especies insectívoras, se encuentran 13 genomas de representantes de la familia Vespertilionidae: *Myotis lucifugus* (7 x), *Myotis davidii* (110 x), *Myotis brandtii* (120x), *Myotis myotis* (47.8 x), *Miniopterus natalensis* (77 x), *Eptesicus fuscus* (84 x), *Aeorestes cinereus* (42 x), *Miniopterus schreibersii* (24.8 x), *Lasiurus borealis* (42.1 x), *Pipistrellus pipistrellus* (52.8 x), *Murina aurata* (33 x), *Nycticeius humeralis* (41.7 x) y *Antrozous pallidus* (35.7 x). En la familia Craseonycteridae, está el genoma de la especie: *Craseonycteris thonglongyai* (52.2 x). Un genoma representante de la familia Rhinolophidae, *Rhinolopus ferrumequinum* (52.8 x). Para la familia Hipposideridae (2 genomas): *Hipposideros armiger* (218.6 x), *Hipposideros galeritus* (46 x). Para la familia Mormoopidae (3 genomas): *Pteronotus parnellii* (18x), *Mormoops blainvillei* (32 x) y *Taradira brasiliensis* (30 x). Dentro de la familia Molossidae: *Molossus molossus*. En la familia Noctonilidae: *Noctilio leporinus* (42.9 x). Para las especies de murciélagos carnívoros-insectívoros, se cuenta con un genoma representante de la familia Megadermatidae: *Megaderma lyra* (18 x). Para la familia Phyllostomidae, se encuentran disponibles 13 genomas de las especies: *Macrotus waterhousii* (49 x), *Macrotus californicus* (47.7 x), *Phyllostomus discolor* (65.8 x), *Desmodus rotundus* (72 x), *Anoura caudifer* (52.7 x), *Tonatia saurophila* (46.1 x), *Micronycteris hirsuta* (39.4 x), *Artibeus jamaicensis* (25 x), *Carollia perspicillata* (27 x), *Musonycteris harrisoni* (30 x), *Leptonycteris nivalis* (54 x), *Leptonycteris curasoae* (43 x) y *Leptonycteris yerbabuenae* (103 x).

A partir de esta información genómica, se han reconstruido filogenias con información de cientos de genes se han hecho análisis de la evolución de familias génicas y análisis de selección, e identificado genes relacionados con la ecolocación, respuesta del sistema inmune, hibernación, reparación del DNA y metabolismo de lípidos y azúcares (Parker et al., 2013; Seim et al., 2013; Zhang et al., 2013; Eckalbar et al., 2016; Jebb et al., 2019; Gutiérrez-Guerrero et al., 2020).

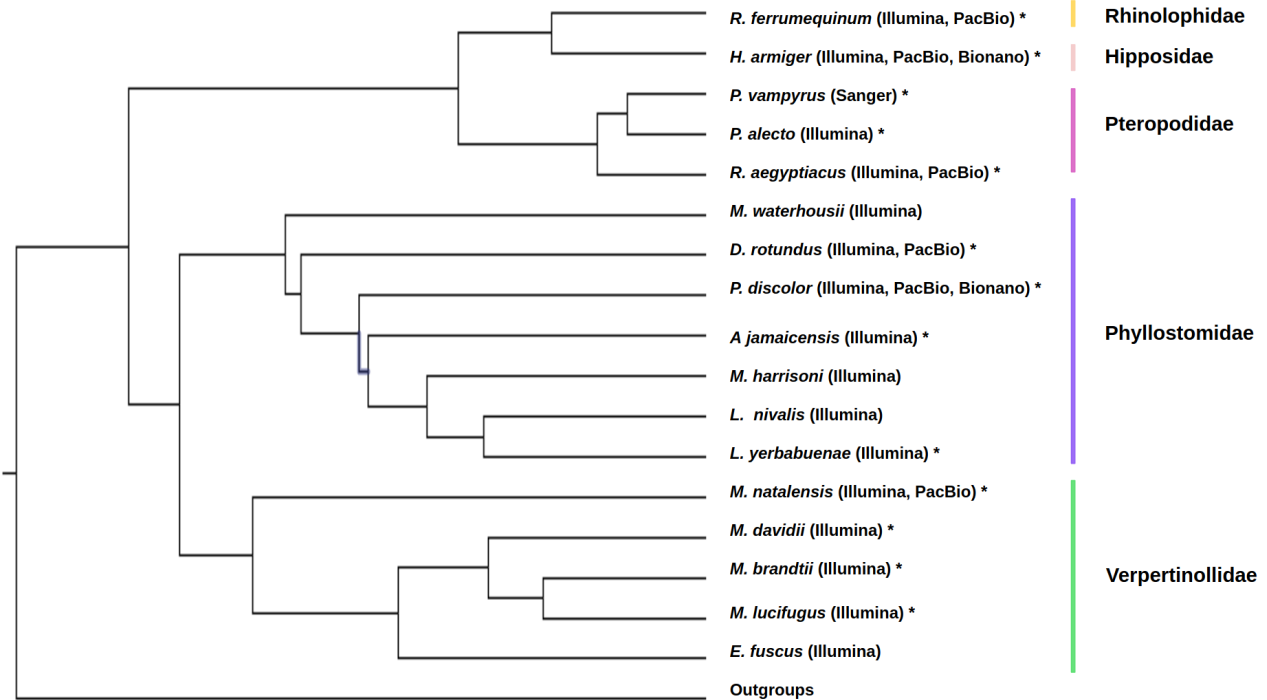


Figura 3. Genomas ensamblados y con anotación disponible en bases de datos de NCBI, ENSEMBL y GigaDatabase. Filogénia realizada en la presente tesis, utilizando genes de copia única. *Especies para las que hay información genómica y transcripcional.

A partir de datos genómicos, Zhang et al. (2013) propusieron una serie de genes que participan en mecanismos de reparación del DNA y pueden ser importantes en la aparición del vuelo. Por otra parte, diferentes estudios genómicos con varias especies de murciélagos, identifican genes con señales de selección y duplicación relacionados con ecolocalización, por ejemplo los genes *FOXP2*, *SERPINB6* y *TJP2* (Parker et al., 2013; Seim et al., 2013; Zhang et al., 2013; Jebb et al., 2019; Gutiérrez-Guerrero et al., 2020).

Otro de los aspectos más estudiados con la información genómica de estos murciélagos, es la evolución del sistema inmune, donde la mayoría de los estudios revela genes bajo selección y con expansiones en familias génicas involucradas en rutas de defensa (*SEPTIN*, *Natural killers*), restricción de infecciones virales (*APOBEC*), señalización e inflamación (*INAVA*, *NF-Kb*). Estos resultados son cruciales para entender como estos mamíferos pueden ser hospederos de una gran cantidad de virus, que pueden ser mortales para otros mamíferos (Zhang et al., 2013; Eckalbar et al., 2016; Jebb et al., 2019; Gutiérrez-Guerrero et al., 2020).

Estudios recientes han prestado atención a la evolución y especialización de las dietas en estos mamíferos. Zepeda-Ledesma et al. (2018) identificaron genes específicos relacionados con la especialización de la sanguivoría, por ejemplo, el gen *REG4*, involucrado en epitelio gastrointestinal y metabolismo anticoagulante y el gen *FFRA1*, que participa en el metabolismo de carbohidratos. Asimismo, identificaron la importancia del microbioma en la síntesis de

vitaminas, lo cual compensa los bajos valores nutricionales de una dieta exclusiva en consumo de sangre. Nuestro estudio (Gutiérrez-Guerrero et al., 2020) exploró la evolución de las dietas en la familia Phyllostomidae en un contexto de genómica comparada, analizando un espectro de hábitos alimenticios que van del insectívoro, a sanguívoro, frugívoro o nectarívoro, identificando una serie de genes bajo presiones de selección positiva, que participan en importantes rutas metabólicas (glucólisis y ketogénesis) para murciélagos nectarívoros, o por ejemplo genes bajo selección involucrados en evitar la coagulación exclusivos para el vampiro.

Transcriptomas

Para la mayoría de las especies de murciélagos con genoma secuenciado, también existe información transcriptómica, principalmente generada para tener evidencia transcripcional en la anotación de los genomas. A la fecha, hay datos disponibles de RNA-Seq para 28 especies de murciélagos (Fig. 3). Para los Pteropodidos: *P. alecto*, *P. vampyrus*, *R. aegyptiacus*, *C. sphinx* y *H. monstrosus*. En la familia Vespertilionidae: *M. lucifugus*, *M. davidii*, *M. brandtii*, *M. keaysi*, *M. leucogaster*, *M. ricketti*, *M. natalensis*, *M. schreibersii*, *E. fuscus*, *M. leucogaster*. Para la familia Rhinolophidae, *Rhinolopus ferrumequinum*. Para la familia Hipposideridae, *H. armiger*. En la familia Mormoopidae *T. brasiliensis* y *M. megalophylla*. Dentro de la familia Molossidae: *M. molossus* y *N. laticaudatus*. En la familia Emballonuridae, *P. macrotis*. Finalmente, en la familia Phyllostomidae: *M. californicus*, *P. discolor*, *D. rotundus*, *A. jamaicensis*, *C. brevicauda* y *L. yerbabuena*.

A partir de estudios transcriptómicos para el murciélago *M. brandtii*, se identificaron cambios en la expresión génica en condiciones de hibernación, por ejemplo, genes con baja-regulación que participan en la síntesis de proteínas y glucólisis, mientras que otros exhiben una sobre-expresión involucrados en el estado hipometabólico de la hibernación (*FGF21* y *ATG101*) (Seim et al., 2013).

Otros estudios se han enfocado en análisis de transcriptómica comparada. Moreno-Santillán et al., (2019), secuenciaron el transcriptoma de hígado de 5 especies de murciélagos tropicales (insectívoros y un frugívoro) identificando una alta expresión de transcritos tejido-específico (hepáticos), que participan en el metabolismo de los lípidos, como por ejemplo el gen *Apolipoprotein E*. También reportan una expresión diferencial de en el murciélago frugívoro, en genes como la glucosa-6-fosfato deshidrogenasa (Moreno-Santillán et al., 2019).

Epigenómica

Uno de los estudios más innovadores empleando diferentes estrategias y técnicas NGS, ha sido el análisis de la evolución del ala en la especie *M. natalensis*. Eckalbar et al., (2016), utilizaron datos de RNASeq y ChipSeq para analizar en diferentes estadios embrionarios la expresión génica diferencial y acetilación de histonas en genes que regulan la formación y desarrollo del

ala. A partir del CHIP-seq identificaron elementos génicos regulatorios que son importantes promotores de adaptaciones morfológicas, como el ala (*Tbx5* y *Pitx1*). Encontraron un total de 2,475 regiones diferenciadas y expresadas en las extremidades y con la información transcriptómica identificaron una expresión diferencial (DE, por sus iniciales en inglés) en tres periodos de desarrollo y formación de ala: en el primer periodo se observó un incremento en la expresión de genes catabólicos y de división celular; en el segundo periodo se organiza la cromatina; y en el tercer periodo se detectó una alta expresión diferencial en el metabolismo de colágeno, en el desarrollo del sistema esquelético y en el desarrollo de epidermis. En este trabajo también se reporta la participación de RNA largos no codificantes (lnc-RNA, por sus siglas en inglés), con una DE en los tres periodos de desarrollo, y su posible participación como reguladores (Eckalbar et al., 2016).

Viroma

Dada la importancia de los murciélagos como reservorios virales, Wu et al., (2016) analizaron con técnicas NGS en metagenomas, la composición viral de 40 diferentes especies de murciélagos (4,440 individuos) distribuidos en 29 provincias de China, tomando una muestra biológica de laringe y ano. En este trabajo reportan un total de 79 familias de virus, de las cuales los géneros Coronaviridae, Parvoviridae y Circoviridae son de los más abundantes (Wu et al., 2015). Por otra parte, Hu y colaboradores (2017) analizaron 235 murciélagos distribuidos en dos provincias de China, tomando muestras biológicas de intestino y pulmón. En este trabajo encontraron un total de 25 familias de virus los géneros con mayor abundancia fueron Herpesviridae, Coronaviridae, Adenoviridae y Parvoviridae, dando soporte a los resultados previamente mencionados (Hu et al., 2017).

Estudios genómicos recientes en murciélagos

Existen proyectos e iniciativas recientes que buscan obtener un mayor número de información genómica de murciélagos empleando herramientas NGS y bioinformáticas. A nivel nacional, se ha secuenciado el genoma de 6 especies de filostómidos, a través del proyecto “Genómica de la Diversidad de Vertebrados Mexicanos 1: *Leptonycteris* y la evolución de la nectarivoría en aves y murciélagos”, Fronteras de la Ciencia, CONACyT (Figura S1), a cargo del Dr. Luis E. Eguiarte, del Instituto de Ecología, UNAM, donde surge esta tesis y una propuesta recientemente aprobada por el CONACyT para estudiar los genomas de murciélagos endémicos de México, liderada por el Dr. Jorge Ortega de la Escuela Nacional de Ciencias Biológicas (ENCB) del Instituto Politécnico Nacional (IPN).

A nivel internacional, el consorcio Bat 1K, es una iniciativa que busca secuenciar los genomas a nivel cromosoma de ~ 1,200 especies de murciélagos, con un progreso de 15 genomas secuenciados a diferente resolución (Teeling et al., 2018). Por otra parte, el Broad

Institute encabeza un proyecto para la secuenciación del genoma de especies de mamíferos, incluidas más de 10 genomas de murciélagos, que están disponibles en bases de datos de NCBI.

Macroevolución y diversificación de las dietas en animales

La evolución de los diferentes hábitos alimenticios es un evento fundamental en la biología de las especies, ya que involucra una serie de procesos de diversificación de linajes (Román-Palacios et al. 2019), adaptaciones a nuevos nichos ecológicos como los marinos o terrestres (Gómez et al. 2010), cambios a nivel morfológico como la dentición (Price et al. 2012) y a nivel fisiológico están las innovaciones en la adquisición de nutrientes (Shondube et al. 2000). Asimismo, las estrategias dietéticas han sido cruciales para diversas interacciones bióticas que van desde el establecimiento de jerarquías tipo presa-cazador, la competencia por el alimento, interacciones mutualistas o de cooperación y también procesos de coevolución (Ehrlck & Rave 1964; Gómez et al. 2010; Wiens et al. 2015). Muchos de los procesos evolutivos mencionados, son una respuesta o consecuencia directa del cambio de dieta y que se ve reflejado en la riqueza de especies y biodiversidad (Burin et al. 2016).

De acuerdo con Román-Palacios et al. (2019), analizando más de 1,000 taxones en 21 Phyla del reino animal, identificaron que el 63% de los taxones tienen una dieta carnívora, el 32% son herbívoros y apenas el 3% tiene una dieta omnívora, y se sugiere que la carnivoría es probablemente la dieta ancestral del reino animal (Román-Palacios et al. 2019).

Uno de los grupos donde el papel de la dieta ha sido extensamente estudiado son los insectos, donde se ha observado que la herbivoría es el hábito alimenticio predominante (Strong et al. 1984). Sin embargo, en los hexápodos existe una diversidad de nichos ecológicos y la dieta, como son el parasitismo, la detritivoría, la fungivoría y la fitogágia, siendo este último el nicho dietético más abundante (Wiens et al. 2015). La diversidad de estos nichos ecológicos y dietéticos, tienen una fuerte relación con la radiación de especies de hexápodos, que comienza a finales del Paleozoico (durante la era Permiana) y durante el Mesozoico hay un incremento en la radiación de especies con hábitos fitofágios, predatorios y detritívoros (Wiens et al. 2015).

Dentro de los crustáceos, los Copepodos, Amphipodos, Isopoda, Caridea, Astacidea, Brachyura, entre otros (quienes conforman el grupo más grande de los artrópodos y que habitan ecosistemas terrestres, marinos y de agua dulce), se han identificado 31 eventos independientes de transición de una dieta carnívora y detritívora a consumidores de plantas y algas (Poore et al. 2017).

Otro ejemplo de como el cambio de dieta ha promovido dinámicas de especiación y extinción lo encontramos en las aves (O' Donnell et al. 2012). En el clado de las aves hay una enorme diversidad taxonómica ~10, 300 especies, diversidad ecológica y un amplio espectro de adaptaciones dietéticas (Kissling et al. 2012), como los insectívoros (por ejemplo, los vencejos y cisnes), los frugívoros (como los guacharos), los carroñeros (por ejemplo, los buitres) o los nectarívoros (como los colibríes) (Bruin et al. 2016). Bruin et al. (2016), identificaron a lo largo de la filogenia de las aves que la tasa de extinción es más frecuente en la transición y adaptación a

una nueva dieta, con la excepción de las especies con un hábito alimenticio omnívoro e insectívoro, donde se observa altas tasas de especiación. Para las aves, se sugiere que la transición a la omnivoría ha sido un evento macroevolutivo importante en la diversificación de especies (Bruin et al. 2016).

En los mamíferos, se observa lo contrario a las aves, una transición a la omnivoría involucra bajas tasas de especiación y altas tasas de extinción (Fritz et al. 2009). Price et al. (2011) identificaron que la transición a una dieta herbívora o carnívora son uno de los principales promotores en la diversificación de los linajes. Asimismo, el hábito alimenticio en mamíferos ha promovido la evolución de caracteres morfológicos como la dentición, mandíbula y tamaño del cuerpo, por ejemplo el tipo de dientes ha facilitado la diversificación del grupo Theria (Luo et al. 2007). En los mamíferos, el cambio y las transiciones de la dieta ha promovido también un gran número de eventos de duplicación y pérdida de genes relacionados con los sistemas sensoriales (Hughes et al. 2018), que tienen un papel importante en la obtención y detección del alimento.

Historia evolutiva de la familia Phyllostomidae y especialización de la nectarivoría

Dentro de los quirópteros, la familia Phyllostomidae es un grupo monofilético y el segundo grupo más grande después de los Vespertilionidos (Rojas et al., 2011; Dávalos et al., 2012; Dumont et al., 2012). La familia Phyllostomidae agrupa siete subfamilias, con más de 160 especies distribuidas en todo el continente Americano (Baker et al., 2010; Fleming y Geisemal 2008; Datzmann et al., 2010; Trejo-Salazar 2012). De acuerdo con reconstrucciones filogenéticas, el ancestro en común de los filostómidos fue datado para el Eoceno medio [42 millones de años (Ma), intervalo de confianza entre 49 y 37 Ma] (Datzmann et al., 2010; Rojas et al., 2011). Los cambios geológicos y climáticos del Oligoceno hasta el Mioceno temprano (27 – 12 Ma), promovieron que la mayoría de sus subfamilias se diversificaron rápidamente (Datzmann et al., 2010).

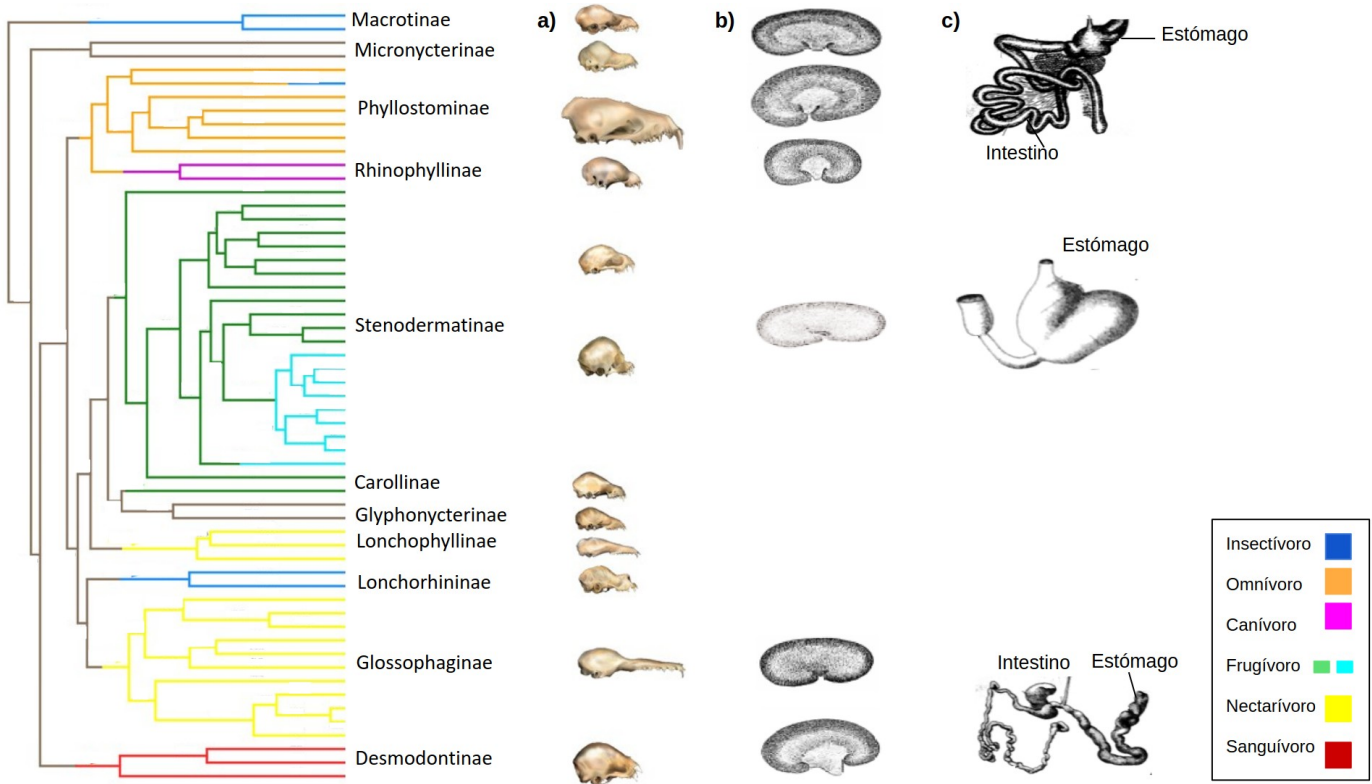


Figura 4. Filogenia de la familia Phyllostomidae. a) Morfología craneal para cada subfamilia. b) Morfología del sistema renal para las subfamilias Macrotoninae, Phyllostominae, Rhinophyllinae, Stenodermatinae, Glossophaginae y Desmodontinae. c) Morfología del sistema digestivo para insectívoros, omnívoros, frugívoros y sanguívoros. Figura tomada y modificada de Rossoni et al. (2017).

La familia Phyllostomidae ha sido ampliamente estudiada, principalmente porque dentro de los vertebrados es la única familia cuyos representantes exhiben una gran diversidad de hábitos alimenticios, ya que se tienen registradas especies insectívoras, carnívoras,

hematófagas, picívoras, frugívoras y nectarívoras (Schondube et al., 2001; Monteiro y Nogueira, 2011). Reconstrucciones de caracteres ancestrales sugieren que la insectivoría es la dieta ancestral (Yohe et al., 2015) y los cambios en las diferentes dietas han sido promovidos por factores climáticos y ecológicos, acompañados de adaptaciones morfológicas y fisiológicas en las funciones de diferentes tejidos y órganos (Fig. 4) (Schondube et al., 2001; Rossoni et al., 2017). La insectivoría es la fuente de alimentación para la mayoría de los murciélagos (~88% de las especies), excepto para la familia Pteropodidae del Viejo Mundo que tienen una dieta frugívora y para la familia Phyllostomidae (aproximadamente 80% son insectívoros – omnívoros) (Arrizabalaga-Escudero et al., 2019). De las especies insectívoras de filostómidos, el género *Macrotus* son los representantes vivos de la familia que divergieron más tempranamente (Jones et al., 2000; Trejo-Salazar 2012). Por otra parte, se tiene reportado dos extraordinarias especializaciones en el cambio de dieta, una es la sanguivoría en la subfamilia Desmodontinae integrada por tres especies que incluye al vampiro *Desmodus rotundus*, capaz de mantener una dieta exclusiva en el consumo de sangre (Carter et al., 2020).

Otra de las más importantes especializaciones es la nectarivoría. Reconstrucciones filogenéticas sugieren cambios en el tipo de dieta en dos ocasiones distintas, como es el caso de la nectarivoría–frugivoría–nectarivoría (Rojas et al., 2011; Trejo-Salazar, 2012; Dávalos et al., 2012; Dumont et al., 2012). Se tiene registro que este hábito alimenticio ha evolucionado en dos linajes independientemente, para las subfamilias Glossophaginae y Lonchophyllinae (Dávalos et al., 2012). Rojas et al. (2011) sugieren que el género *Phyllostomus* es el último ancestro común de la tribu Glossophagini (20.1 Ma) y Lonchophyllini (12.9 Ma). Jones et al., (2002) y Rojas et al. (2011) proponen una relación entre los eventos de diversificación en murciélagos y eventos de diversificación en las plantas con flor (25 - 40 Ma), durante el último calentamiento del Oligoceno, promoviendo la evolución mutualista en las tasas de diversificación.

En la subfamilia Glossophaginae los tres géneros de nectarívoros más especializados son *Leptonycteris*, *Glossophaga* y *Monophylus* (Datzmman et al., 2010; Rojas et al., 2011; Flores-Abreu et al., 2019). Por otra parte, la subfamilia Lonchophyllinae alberga dos géneros de nectarívoros *Lonchophylla* y *Lionycteris*, aunque la posición de estos géneros es controversial, ya que de acuerdo a algunos análisis están más relacionados con otros grupos dentro de la familia (omnívoros, insectívoros y frugívoros) que con la subfamilia Glossophaginae (Datzmman et al., 2010). Las diferencias de estas dos subfamilias que podría apoyar una evolución independiente son principalmente morfológicas. Por ejemplo, los nectarívoros de la subfamilia Lonchophyllinae tienen un profundo surco en su lengua y pelos papilares dorsales y ventrales; este surco no se encuentra en la tribu Glossophagini, donde los pelos papilares están distribuidos en forma de cepillo en la parte ante-dorsal de la lengua (Datzmman et al., 2010).

En la subfamilia Glossophaginae se tienen registradas alrededor de 40 especies y el cambio a una dieta frugívora y nectarívora inició hace 23 – 19 Ma (Trejo-Salazar 2012; Flores-

Abreu et al., 2019). Varias especies de esta subfamilia son consideradas generalistas, al consumir néctar y otras fuentes de alimento como frutas y/o insectos (Schondube et al., 2001; Monteiro y Nogueira, 2011), mientras que un pequeño grupo son nectarívoros estrictos, las cuales pueden pasar toda su vida con una dieta dominada por azúcares, consumiendo néctar y como segunda fuente de alimento el polen o algunos insectos para satisfacer sus requerimientos de nitrógeno y proteínas (Fleming et al., 1993; Braker et al., 2010; Flores-Abreu et al., 2019; Göttliger et al., 2019).

Estos nectarívoros estrictos presentan adaptaciones únicas que van desde la morfología, a la fisiología y diferencias genéticas. En cuanto a las adaptaciones morfológicas más notables, se observan principalmente en los miembros del género *Leptonycteris*, y en las especies *Musonycteris harrisonii* y *Choeronycteris mexicana* (Ayala-Berdon y Schondube, 2011; Gonzalez-Terrazas et al., 2012). Una característica es el alargamiento del hocico y cambio en la estructura de la lengua, que permite recoger el néctar que se encuentra a poca profundidad de las corolas de las flores (Gonzalez-Terrazas et al., 2012). La lengua es posible alargarla gracias a los músculos que controlan la extensión y retracción, donde el alargamiento de la mandíbula da soporte al peso y actividad de la lengua durante la alimentación (Freeman, 1995).

Las adaptaciones fisiológicas relacionadas con la nectarivoría han sido ampliamente exploradas. Ayala-Berdon et al. (2011), analizaron como afecta la variación en las concentraciones de azúcares en el néctar (sacarosa y hexosas) y la tasa de ingesta de comida en la fisiología de los murciélagos nectarívoros. Dado que la producción de néctar de cada planta varía mucho en su concentración de azúcar (entre el 3% al 67%), se identificó que los murciélagos tienden a incrementar su consumo de alimento cuando la concentración de sacarosa decrece. Hernández y Martínez del Río (1992) encontraron que los murciélagos nectarívoros y frugívoros exhiben una mayor actividad de disacaridasas en el área intestinal, en comparación con murciélagos insectívoros. Se estima que los murciélagos ingieren alrededor de 15.5 ml de néctar por noche y se calcula que un individuo puede visitar hasta 1000 flores en una noche (V-Helversen y Reyer, 1984; Tschapka y Dressler, 2002). Para la especie *L. yerbabuena*, se encontró una relación positiva entre el azúcar consumida y la concentración de azúcar en el néctar, lo cual indica que hay un límite energético cuando se alimentan de concentraciones bajas en azúcares (Ayala-Berdon et al., 2011).

Existen pocos estudios que traten aspectos genéticos de la evolución de la nectarivoría en murciélagos. Jiao et al. (2019) encontraron que el gen de la *trehalasa*, importante para obtener carbohidratos a partir de síntesis de trehalosa proveniente del consumo de insectos, se ha pseudogenizado o perdido en varias especies de Filostómidos, incluidas especies nectarívoras. Gutiérrez-Guerrero et al. (2020) analizando genomas de tres especies de filostómidos nectarívoros, encontraron genes con señales de selección que participan en rutas metabólicas de glucólisis, glucogénesis y el metabolismo ketogénico. También reportan

paralelismos moleculares entre murciélagos Glossophaginos y frugívoros del viejo mundo para dos genes: *ALKBH7* y *ACCS*, involucrados en la obtención de ATP a partir de lípidos. En murciélagos frugívoros Pteropodidos y filostómidos, Shen et al., (2012), estudiaron el gen *Slc2a4* que codifica para la proteína transmembranal GLUT4, involucrada en mantener la homeostasis de glucosa en sangre e identificaron señales de selección positiva y cambios estructurales en la proteína GLUT4 en los murciélagos del Viejo Mundo y del filostómido *A. jamaicensis*. Zhao et al. (2010), analizaron la evolución del receptor de sabores dulces, *Tas1r2s* (que se expresa como un heterodímero conformado por los genes: *Tas1r2* – *Tas1r3*), en 42 especies de quirópteros e identificaron que el gen *Tas1r2*, está bajo una fuerte selección purificadora tanto en frugívoros como insectívoros y se observa una pseudogenización de este receptor en las especies de murciélagos hematófagas (Jiao et al., 2018; Hong & Zhao, 2014). Aunque hay pocos estudios, los avances en esta área son relevantes, ya que algunos de estos genes que muestran señales de selección positiva en los murciélagos nectarívoros, en humanos, las mutaciones en estos genes están relacionados con trastornos metabólicos como hiperglucemia o diabetes (Gutiérrez-Guerrero et al., 2020).

Un grupo representativo y particularmente interesante de filostómidos nectarívoros estrictos, son las especies del género *Leptonycteris*, integrado por tres especies: el murciélago magueyero mayor *L. nivalis*, el murciélago magueyero menor *L. yerbabuena* y el murciélago magueyero u hocicudo de Curazao *L. curasoae* (Gonzalez-Terrazas et al., 2016; Rodriguez-Peña et al., 2016; Frick et al., 2018; Gómez-Ruiz & Lacher, 2019; Flores-Abreu et al., 2019). Se caracterizan por presentar un rostro relativamente alargado, que le proporciona un mayor acceso al tubo floral (Arita y Martínez del Río, 1990; Slauson 2000; Gonzalez-Terrazas et al., 2016). El tamaño de su lengua es aproximadamente igual al cuerpo (alrededor de 8 cm), con alta densidad de papilas que permite la absorción de néctar por capilaridad (Greenbaum y Phillips, 1974; Nassar et al., 2008; Gonzalez-Terrazas et al., 2012). Presenta un tracto digestivo corto comparado con otros murciélagos con dietas frugívoras e insectívoras (Cruz-Neto et al., 2001; Casotti et al., 2006). También muestra una reducción en el tamaño de sus riñones comparado con el sistema renal de otros murciélagos con otro tipo de dieta, esto se correlaciona con evitar la pérdida de electrolitos presentando una orina diluida (Carpenter 1969). Los cambios en su sistema de vuelo les confieren la capacidad de mantener un vuelo estacionario y/o de revoloteo (*hovering* en inglés), mientras visitan más de 1,000 flores por noche (Martínez-Río et al. 1997; 1999; Valiente-Banuet et al. 1996; Von et al., 2012; Gonzalez-Terrazas et al., 2016).

Estas especies son clave para sus ecosistemas, ya que su hábito nectarívoro es fundamental en la polinización de especies centrales para el paisaje y el funcionamiento de diversos ecosistemas de México, como son las especies de plantas del género *Agave* y muchas cactáceas columnares; promoviendo la fecundación cruzada y el flujo génico entre individuos (Arizaga et al., 2000; Molina-Freaner y Eguiarte, 2003; Valiente-Banuet et al., 2004; Rodríguez-

Oseguera et al., 2013; Flores-Abreu et al., 2019). Diferentes estudios filogenéticos sugieren que esta interacción mutualista ha promovido procesos de coevolución difusa, ya que por análisis de reloj molecular se observa una coincidencia en el tiempo de origen y diversificación entre *Leptonycteris* - agaves y cactus columnares (entre 8 - 6 millones de años) (Gentry 1982; Molina-Freaner et al., 2003; Good-Ávila et al. 2006; Fleming 2009; Flores-Abreu 2019). Además de ser especie clave para el ecosistema, *L. yerbabuena* y *L. nivalis* son los principales polinizadores de diferentes especies de agave mezcalero o pulqueros, como el *Agave angustifolia*, *A. inaequidens*, *A. potatorum*, *A. macroacantha*, entre otros (Rocha et al., 2006; Trejo-Salazar et al., 2016).

La industria del tequila y mezcal, constituyen la segunda actividad económica más importante dentro de las actividades de producción de bebidas alcohólicas en México (INEGI, 2018). Además, el mezcal es una bebida representativa de la historia y cultura mexicana, que cuenta con denominación de origen (NOM-079-SCFI-1994). De acuerdo con reportes del Consejo Regulador del Mezcal (2020), la producción de mezcal en 2011 fue de 980,000 litros, registrando un crecimiento de producción para el 2019 de 7,145,039 litros, incrementando un 37% la exportación, según su informe en el 2019 la industria mezcalera registró una derrama de 6.942 mdp, generando 23,000 empleos directos y más de 105,000 empleos indirectos.

Evolución convergente en vertebrados nectarívoros: murciélagos y colibríes

Las convergencias y paralelismos son dos formas de evolución independiente de rasgos idénticos o similares en diferentes linajes y en su mayoría son resultado de presiones de evolución adaptativa (Stewart 2006; Lossos 2011; Pickersgill 2018).

La evolución paralela y convergente originalmente fue estudiada utilizando caracteres fenotípicos; su estudio con los métodos de secuenciación y amplificación del DNA ha permitido extender a estos análisis a genes, secuencias de nucleótidos, secuencias de amino ácidos, estructuras de proteínas y regiones reguladoras (Sackton & Clark 2019). A nivel fenotipo, la evolución paralela se refiere a fenotipos similares en especies cercanamente relacionadas, mientras que la convergencia involucra similares fenotipos pero en especies filogenéticamente lejanas (Orr 2005; Lossos 2011). En genes los paralelismos ocurren cuando fenotipos similares se producen por genes ortólogos, es decir aquellos genes homólogos que han divergido a partir de un gen ancestral en común, mientras que en la convergencia los fenotipos similares son producto de genes no homólogos (Stewart 2006; Pickersgill 2018; Sackton & Clark 2019). A nivel de secuencia de nucleótidos los fenotipos similares son producto de cambios idénticos en el mismo gen, en contraste en la convergencia los fenotipos similares son producidos por diferentes cambios en el mismo gen (Pickersgill 2018; Sackton & Clark 2019).

Los eventos de evolución paralela se han descrito en virus, por ejemplo, el gen *hemagglutinin* del virus de la influenza aviar, donde la glucosilación de algunos sitios ha evolucionado independientemente y en diferentes ocasiones, volviendo a las especies de virus altamente patógenas (Gutiérrez et al., 2019). Otro ejemplo, son rasgos morfológicos en plantas, que han evolucionado en paralelo en especies de *Rheum* (Sun et al., 2012). Uno de los ejemplos más importantes es la evolución fenotípica de la placa de escudo del pez espinoso (*Gasterosteus aculeatus*) (Colosimo, 2005).

En general los eventos de evolución paralela y convergente son resultado de la selección adaptativa, lo que refleja que los taxones han estado bajo presiones ambientales o fisiológicas similares que promueven la especialización a un nicho (Arendt & Reznick 2008). Una de estas especializaciones que ha evolucionado de manera paralela y convergente, es la especialización de una dieta nectarívora.

En vertebrados, la evolución a una dieta nectarívora ha aparecido independientemente en dos linajes filogenéticamente lejanos, dentro de las aves la familia Trochilidae que alberga a los colibríes (Flemming et al., 2008; 2009). En los mamíferos, el orden de los murciélagos alberga a la subfamilia Glossophaginae, donde algunos de sus representantes se caracterizan por una dieta especializada en el consumo de néctar (Flores-Abreu et al., 2019).

Tanto los colibríes como los murciélagos nectarívoros han evolucionado de un ancestro con dieta insectívora (Bleiweiss, 1998). La diversificación de estos grupos, coincide con la radiación de grupos de angiospermas en la mitad del Mioceno (Bleiweiss, 1998; Dávalos et al.,

2012; Rojas et al., 2011). McGuire y colaboradores (2014), estimaron que el ancestro en común de los colibríes vivió hace unos 22.4 millones de años (con un intervalo de 20.3 - 24.7 Ma). En los murciélagos nectarívoros de la subfamilia Glossophaginae, diversos autores han estimado un ancestro en común hace 21 Ma (con un intervalo de 17 a 25 Ma) (Flores-Abreu et al., 2019). Se estima que actualmente hay ~338 especies de colibríes, con una distribución restringida al continente Americano (Licona-Vera & Ornelas, 2017). Mientras que para los murciélagos nectarívoros de la subfamilia Glossophaginae, existen alrededor de 45 especies, que solo se distribuyen en América (Fleming et al., 2009; Flores-Abreu et al., 2019).

Para un gran número de especies de plantas, los colibríes y los murciélagos nectarívoros fungen como polinizadores primarios (Muchhala et al., 2009), siendo especies críticas para mantener la salud genética de las plantas por medio de la polinización y fecundación cruzada (Rocha et al., 2006; Trejo-Salazar et al., 2018). Hernández y Martínez del Río (1992) identificaron que los murciélagos nectarívoros y colibríes difieren en los componentes y abundancias de los azúcares del néctar que prefieren. En promedio los néctares que prefieren los murciélagos nectarívoros consisten en 20% sacarosa, con una preferencia a visitar flores con altas concentraciones de glucosa + fructosa (Martínez del Río, 1990; Coleman & Downs, 2012; Rodríguez-Peña et al., 2016). En los colibríes el forrajeo ocurre preferencialmente en la mañana, cuando las temperaturas ambientales son más bajas, por lo que sus mecanismos termogénicos son activados, resultando una elevación en las tasas de energía gastada (Lotz et al., 2003; Powers et al., 2017). Los colibríes prefieren visitar flores que producen néctares ricos en azúcares, hasta 65% de azúcares (Martínez del Río et al., 1992; Roberts, 1996) y muestran una preferencia en soluciones de sacarosa + hexosa (Martínez del Río, 1990).

A pesar de ser especies filogenéticamente lejanas, colibríes y murciélagos nectarívoros han evolucionado una serie de caracteres convergentemente (Ingersoll et al., 2018). Una de las características extraordinarias de ambos grupos es la capacidad de mantener un vuelo estacionario mientras forrajean, el cual implica un costoso gasto energético (Suarez et al., 2010; Ingersoll et al., 2018). Para colibríes y murciélagos nectarívoros, la mayor cantidad de energía obtenida a partir del néctar será gastada durante el vuelo estacionario, ya que implica altas frecuencias en el aleteo y el incremento en las tasas metabólicas (Welch et al., 2011; Welch & Chen, 2014). Welch et al., (2011), encontraron que la azúcar ingerida es la primera fuente de carbohidratos que se oxida durante el vuelo estacionario, estimando para murciélagos nectarívoros una tasa de oxidación de 3.80 mg/min (para *G. soricina*, con masa corporal 10.2 ±0.1 g) y en colibríes de 3.49 mg/min (para *C. anna*, masa corporal 4.98 ±0.48 g). En contraste, Jentsens et al. (2004) reportan en humanos tasas de oxidación máxima de azúcares exógenas de ~1,220 mg/min durante el ejercicio (en un periodo de 120-150 min), de manera que la tasa de oxidación de masa específica de azúcares exógenas en *G. soriana*, es aproximadamente 18

veces más grande que la estimada en humanos y la más alta reportada en mamíferos (Jentjens et al., 2005; Jentjens et al., 2004; Jentjens et al., 2004; Pirnay et al., 1995).

Estos estudios apoyan la hipótesis de una evolución convergente en características fisiológicas y bioquímicas entre animales nectarívoros con vuelo estacionario, ya que la alta capacidad de oxidación exógena de azúcar es muy similar entre murciélagos y colibríes (Suárez et al., 2014). En vertebrados nectarívoros con vuelo estacionario, la variación en ingesta rápida de azúcar activa el metabolismo, la azúcar ingerida en los órganos de almacenamiento y el intestino corto, donde ocurre la absorción (Chen & Welch, 2013). La incorporación de azúcar exógena activa los substratos metabólicos en murciélagos y colibríes, los cuales exceden las tasas observadas en humanos, ya que un factor que limita el uso de azúcares exógenas durante el ejercicio en humanos es la capacidad de absorber carbohidratos por el intestino (Chen & Welch, 2013). Las tasas de absorción reportadas para colibríes y *G. sorciana* son excepcionalmente altas y una gran parte de absorción de azúcar ocurre en rutas paracelulares, mientras esta ruta de absorción representa una mínima contribución en otros mamíferos pequeños, como en ratones (Welch & Chen, 2014).

Recientemente, con los métodos de secuenciación masiva, para estos vertebrados nectarívoros se han identificado eventos de evolución paralela utilizando herramientas genómicas y transcriptómicas. Lim et al. (2019) secuenciaron los transcriptomas de hígado de 12 especies de colibríes y detectaron señales de evolución paralela en especies de colibríes que se distribuyen en altas y bajas elevaciones. Se identificaron eventos de evolución paralela a nivel molecular en genes involucrados en rutas de respiración, metabolismo y reparación de DNA en las especies que habitan en altas elevaciones.

También se han estudiado eventos de evolución paralela en murciélagos nectarívoros, que involucran presiones de selección en genes involucrados en la dieta. Gutiérrez-Guerrero et al. (2020), secuenciaron los genomas de los murciélagos glossophaginos *L. yerbabuena*, *L. nivalis* y *M. harrisoni*. A partir de la información genómica se identificaron señales de selección en genes que participan en rutas metabólicas de glucólisis, glucogénesis y ketogénesis, así como expansiones en familias génicas involucradas en producción y almacenamiento de hierro, lo cual es importante en especies con una dieta baja en nutrientes y puede estar involucrado en evitar anemia. Asimismo, se detectaron paralelismos moleculares entre murciélagos Glossophaginos y los Pteropodidos, frugívoros del viejo mundo para dos genes: *ALKBH7* y *ACCS*, involucrados en la obtención de ATP a partir de lípidos. En estos genes, se observaron cambios paralelos en la identidad del amino ácido en sitios específicos en las secuencias.

Interesantemente, Workman et al. (2018) a partir de la secuenciación del transcriptoma de hígado del colibrí *Archilochus colubris*, observaron una alta expresión en las enzimas *ACSM4* y *ACC*, involucradas en el proceso de lipogénesis. Este resultado es interesante, ya que estas enzimas pertenecen a la misma familia de las Aceto-acetil coenzima A, que evolucionaron con

cambios paralelos en los murciélagos nectarívoros (Gutiérrez-Guerrero YT et al., 2020). Estudios a nivel genómico, permitirán identificar eventos de evolución molecular convergente en colibríes y murciélagos nectarívoros, si bien son especies lejanas filogenéticamente, han estado sujetos a presiones adaptativas en respuesta a la especialización de la dieta nectarívora.

Objetivo General

Secuenciar, ensamblar y anotar los genomas de especies de murciélagos de la familia Phyllostomidae, y utilizar herramientas bioinformáticas, análisis estadísticos, de evolución molecular y genómica comparada para estudiar la evolución y especialización de las dietas.

Objetivos Específicos

- Analizar los cambios genómicos relacionados con las transiciones y especializaciones del hábito alimenticio, haciendo especial énfasis en la evolución de la nectarivoría en murciélagos Filostómidos.
- Estudiar la evolución de los receptores sensoriales (olfativo y de sabor), y su relación con el hábito alimenticio en especies de murciélagos, representando cinco familias del orden Quiróptera.
- Identificar genes relevantes para una dieta dominada por azúcares como es la nectarivoría, analizando la información transcripcional de dos vertebrados nectarívoros especialistas: el murciélago magueyero (*Leptonycteris yerbabuena*) y el colibrí de oreja blanca (*Hylocharis leucotis*).

Capítulo I

Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats

Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats

Yocelyn T. Gutiérrez-Guerrero¹, Enrique Ibarra-Laclette², Carlos Martínez del Río³, Josué Barrera-Redondo¹, Eria A. Rebollar⁴, Jorge Ortega⁵, Livia León-Paniagua⁶, Araxi Urrutia⁷, Erika Aguirre-Planter¹ and Luis E. Eguiarte^{*1}

GigaScience, 9(6):giaa059. Doi: 10.1093/gigascience/giaa059

Publicado: Junio 6, 2020

Abstract

Background

The New World Leaf-Nosed bats (Phyllostomids) exhibit a diverse spectrum of feeding habits and innovations in their nutrient acquisition and foraging mechanisms. However, the genomic signatures associated with their distinct diets are unknown.

Results

We conducted a genomic comparative analysis to study the evolutionary dynamics related to dietary diversification and specialization. We sequenced, assembled and annotated the genomes of five Phyllostomid species: one insect-feeder (*Macrotus waterhousii*), one fruit-feeder (*Artibeus jamaicensis*), and three nectar-feeders from the Glossophaginae subfamily (*Leptonycteris yerbabuenae*, *Leptonycteris nivalis* and *Musonycteris harrisoni*), also including the previously sequenced vampire *Desmodus rotundus*. Our phylogenomic analysis based on 22,388 gene families displayed differences in expansion and contraction events across the Phyllostomid lineages. Independently of diet, genes relevant for feeding strategies and food intake experienced multiple expansions and signatures of positive selection. We also found adaptation signatures associated with specialized diets: the vampire exhibited traits associated with a blood diet (i.e., coagulation mechanisms), whereas the nectarivore clade shares a group of positively selected genes involved in sugar, lipid, and iron metabolism. Interestingly, in fruit-nectar feeding Phyllostomid and Pteropodids bats, we detected positive selection in two genes: *AACS* and *ALKBH7*, which are crucial in sugar and fat metabolism. Moreover, in these two proteins we found parallel amino-acid substitutions in conserved positions exclusive to the tribe Glossophagini and to Pteropodids.

Conclusions

Our findings illuminate the genomic and molecular shifts associated with the evolution of nectarivory and shed light on how nectar-feeding bats can avoid the adverse effects of diets with high glucose content.

Keywords

Adaptation, Comparative genomics, Diet, Parallel evolution, Phyllostomid, Specialization

Background

Evolutionary shifts related to changes in feeding habits are considered one of the most important events in animal evolution (1). Diet changes open new ecological and physiological opportunities (1, 2). These shifts often involve changes in feeding behavior, dramatic innovations in the mechanism by which nutrients are assimilated and metabolized, and sometimes drastic morphological modifications (3). Evolutionary diet shifts are sometimes accompanied by species diversification and adaptive functional trait radiation (4).

The New World leaf-nosed bats (family Phyllostomidae) are one of the most species-rich mammalian taxa, with 216 species in 60 genera (5, 6). Leaf-nosed bats evolved from an insect-feeding common ancestor and now display a large and diverse spectrum of feeding habits that include insectivory, carnivory, frugivory, blood-feeding, nectar-pollen feeding and omnivory (5-7). Moreover, dietary specializations and species diversification seem to be correlated in these bats (7,6).

Although most extant Phyllostomids are insectivorous or omnivorous (6, 7), two lineages have extreme dietary specialization: blood-feeding within the subfamily Desmodotinae (including *Desmodus*, *Diphylla* and *Diademus*) and the nectar-pollen feeding species within the subfamily Glossophaginae (including *Leptonycteris*, *Glossophaga*, *Choeronycteris* and *Musonycteris*), that feed primarily on nectar and pollen (5, 6). Among these nectar-pollen feeding species, *Leptonycteris yerbabuena* (Lesser long-nosed bat) is notable due to its tight co-evolutionary interactions with plants and seeming specialization to nectarivory/pollinivory (8-10). The blood-feeder *Desmodus rotundus*, and the two other species in the Desmodotinae subfamily have a feeding mode unique among mammals (11). Data on the genome and microbiome of *D. rotundus* have revealed remarkable adaptive changes genes associated with blood diet (12).

Many studies have demonstrated evidence of evolutionary novelties associated with feeding diversification in leaf-nosed bats (6, 13, 14, 15). These include morphological traits involved in nectar extraction (10), and physiological characteristics related to the processing of a diet high in sugars (13, 14). However, the genomic signatures associated with dietary diversification and specialization during the evolution of Phyllostomid bats from an insect-feeder common ancestor remains largely unknown.

We investigated the genomic and evolutionary dynamics associated with the dietary diversification and nectar-pollen feeding specialization of Phyllostomid bats. We sequenced and assembled the whole genomes of five Phyllostomid bat species, including ecologically and economically important species. We sequenced the genomes of three nectar-pollen feeders *Leptonycteris yerbabuena*, *Leptonycteris nivalis* and *Musonycteris harrisonii*; the fruit-feeder *Artibeus jamaicensis*; and the insect-feeder *Macrotus waterhousii* (Fig. 1). For comparative purposes our analyses incorporated genomic data of the vampire *Desmodus rotundus* (12) and

other mammals.

Our research was guided by three sets of predictions. First, we predicted that the dietary diversification from insectivory, which is the ancestral condition in the group, to derived diets would be accompanied by evolutionary changes in relevant genes involved in food uptake and the metabolic pathways associated with the processing of assimilated nutrients. Second, we predicted that the dietary specializations observed in the subfamilies Desmodotidae and Glossophaginae would be correlated with evidence of selection in genes that facilitate the assimilation and metabolism of components of blood and nectar, respectively. More specifically, we expected the nectar-pollen feeder lineage to show adaptive signals in genes involved in sugar assimilation and metabolism. Our third prediction was that we should detect convergent evolution between the New World fruit and nectar feeders and the Old World fruit-bats in genes important for carbohydrate metabolism.

We adopted a hierarchical approach: we examined our predictions in deep nodes of the phylogeny, then we identified the nodes that represent dietary transitions and investigated the changes that accompanied these transitions. We conducted a genomic comparative approach and performed a phylogenomic reconstruction to identify expansions/contractions of gene families across the Phyllostomids lineage. We also evaluated orthologous protein sequences that have been targets of selection and their relation to dietary diversification and specialization. Finally, in order to identify convergent evolutionary signals associated with the diet, we carried out a comparison between the genomes of the nectar-pollen feeders Phyllostomid bats and the Old World fruit-feeder bats (family Pteropodidae), analyzing radical amino-acid substitutions in conserved positions.

Data Description

We sequenced the genome of one adult male Lesser long-nosed bat (*Leptonycteris yerbabuena*) by a whole high-throughput shotgun strategy and obtained a high quality *de novo* assembly (104 x) (Table 1; see Additional file 1, Table S1, S11). Additionally, we sequenced with medium coverage (~24 – 56 x) the genomes of four Phyllostomid bats: *M. waterhousii* (insect-feeder), *A. jamaicensis* (fruit-feeder), and the nectar-pollen feeders: *M. harrisoni* and *L. nivalis*. (Fig. 1; see Additional file 1, Table S1).

The genomic landscape of New World Leaf-nosed bats

The size of *L. yerbabuena*'s genome was similar to those reported for other bats (2.05 Gb), with an N50 scaffold length of 14, 735,151 bp, and L50 of 38 scaffolds (Table 1). Evaluation of the genome assembly for completeness based on BUSCO identified 94% of complete and 2.5% of fragmented genes from the mammalian database (Mammalia odb9). The genome contained 24,074 inferred coding sequences from an *ab initio* prediction, the transcript evidence and

homology evidence obtained from a set of proteins of several mammalian species (Table 1; see Additional file 1, Table S2). Approximately 26% of the genome assembly was composed of repetitive elements (547 Mbp length) (see Additional file 1, Table S3).

We constructed a reference guide genome assembly based on *L. yerbabuena* for the other four Phyllostomid bats, where we annotated from 18,000 to 24,471 coding sequences and proteins for each Phyllostomid (Table 2; see Additional file 1, Table S4-S5; Additional file 2, Fig. S1).

Analyses

Gene family evolution reflects distinct dietary needs

To understand genomic evolution and to trace changes associated with dietary diversification and specialization, we reconstructed a phylogenomic tree using 132 single-copy orthologous genes (61,331 amino acids sites), which was calibrated using two fossil dates (16-18). Based on the phylogenomic tree, we analyzed the dynamics (expansion and contractions) for 22,388 gene families across the Phyllostomid bat genomes (Fig. 2 and Table 3).

For all the Phyllostomid bats, the significant gene family enrichment functions were related to the cellular repair process and genetic make-up for protein synthesis. Furthermore, across the Phyllostomid bats many gene families exhibited changes with feeding habits, for example, the Phyllostomid node had a contraction related to the lipid metabolism. The blood-feeder lineage had a significant gain on gene families involved in the regulation of appetite and process for nitrogen acquisition, but also this lineage showed many contraction events involved in calcium metabolism (Table 3). The fruit and nectar feeding bats exhibited many expansion events in iron metabolism regulation pathways (Table 3).

Rapidly evolving genes across the whole genome

For all Phyllostomid bats, we identified 42 genes with robust signals of positive selection (FDR $p < 0.05$). According with the enrichment analysis, most of the adaptive genes are related to immune response, DNA repair, inflammatory response, RNA catalytic process and genes that mediate muscle function (such as *Myoblast* and *PAMR1*) (Fig. 3; see Additional file 1, Table S6-TableS8) (19).

Ecological and feeding behavior adaptations across the Phyllostomid lineages

To understand shifts associated with dietary diversification, we analyzed genes under positive selection involved in the mechanisms of carbohydrate digestion and lipid metabolism in each Phyllostomid species (Fig. 3).

In *M. waterhousii*, an insect-feeder, we found evidence of positive selection in *Chitinase*, which codes for proteins in the degradation of insect exoskeleton (20) (Fig. 3) Interestingly, for *M. waterhousii*, the *Trehalase* is a partial gene that exhibited signals of positive selection, but for the rest of the Phyllostomid species, *Trehalase* is a pseudogene. This finding is relevant, because trehalose is the principal sugar in insect's blood.

The vampire's genome revealed a complex set of genes crucial for maintaining a blood feeding diet under positive selection, including *THBD* (hematopoietic cell pathway) and *A2M* (complement and coagulation cascade pathway) (21). Only in the vampire we found under positive selection genes involved in feeding and lipid – cholesterol metabolism such as *MGAT2*, *PLAS2G16*, and *GFOD1* (22, 23) (Fig. 3). Interestingly, the vampire was the only genome where the *Trehalase* gene was completely missing.

In the fruit bat *A. jamaicensis* most of the enzymes analyzed involved in lipid and carbohydrates metabolic pathways showed positive selection pressures (Fig. 3; see Additional file 1, Table S8). The importance of these enzymes for this fruit bat might reflect the diversity of its diet, as *A. jamaicensis* has been documented to eat, besides insects and fruits, seeds and leaves (24).

Every nectar-pollen feeder bat species (*M. harrisoni*, *L. nivalis* and *L. yerbabuena*) showed positive selection signatures for genes involved in insulin secretion (*UCN 3*) (25), calcium and iron storage (*CALP2*, *CD248* and *FTL*) (26-27), bone morphogenetic regulation (*FsIt1*) (28), and in *IAP*, the gene coding for the mucosa defense factor involved in proper gut homeostasis (Fig. 3) (29). Interestingly, we found adaptive signatures for genes crucial for carbohydrates and lipid metabolic pathways, such as pancreatic secretion, glycolysis / gluconeogenesis, glycogen, glycerophospholipid, citrate acid metabolism and ketone metabolism (Fig. 4; see Additional file 1, Table S8) (30-32).

Finally, for the nectar-pollen and fruit bat species, we detected strong selective pressures in four enzymes, *AACS* (which appears to participate in the regulation of lipid metabolism) (32), *ALKBH7* (which codes for a protein that appears to be involved in the regulation of body mass and fat content) (33), *FABP1* (regulates fatty acid trafficking and prevents lipotoxicity) (34) and *AMPK* (major regulator of cellular energy homeostasis) (Fig. 3 and 4) (35).

Adaptation and convergent signatures in fruit and nectar-pollen feeding bats

The sugar and lipid metabolism genes detected under positive selection for the fruit and nectar-pollen feeding Phyllostomid bats, were also analyzed in the genomes of three Old World bat species available in databases, including *Pteropus alecto*, *Pteropus vampyrus* and *Rousettus aegyptiacus* (Family Pteropodidae). We found evidence of positive selection in *AACS* in *P. alecto*, *P. vampyrus* and *R. aegyptiacus*. In *R. aegyptiacus*, we also found evidence of positive selection in *ALKBH7* (see Additional file 1: Table S8).

We analyzed 1,918 orthologous sequences and reconstructed the ancestral sequences states, in order to identify some genes with unique and exclusively parallel substitutions for the Phyllostomid fruit bat, the Glossophagini and Pteropodid lineages, in a conserved position for the rest of the bats and mammal species (Fig. 5b). We found three genes with parallel signatures in a specific amino-acid position. Most of these parallel changes presumably led to changes in the physicochemical properties of the expressed protein (Fig. 5a). In *ACCS*, we identified six radical amino-acid substitutions along the sequence, from an ancestral glycine (non-polar) to a derived arginine (+ charged), alanine (non-polar) to threonine (polar), glycine to serine (polar), alanine to proline (non-polar), serine to proline and serine to leucine (non-polar) (Fig. 5a). In *ALKBH7*, we found a parallel amino-acid substitution from glutamic acid (- charged) to lysine (+ charged) and arginine to glutamine (polar). The latter change was found in the Glossophaginae and *R. aegyptiacus* (Fig. 5b). Finally, for the gene *UNC-45 B*, which codes for a protein involved in muscle cell development, we identified a substitution from leucine to arginine (Fig. 5a) (36). Moreover, we calculated the posterior probabilities for each amino acid reconstruction state for each node along the three. For all nodes in the Glossophagini and Pteropodids clades, the probabilities of each amino acid derived state were > 80% (Fig. 5c; see Additional file 1, Table S9).

Finally, in order to evaluate if the radical amino acid substitutions had affected the protein structure of *ACCS*, we modeled its 3D- protein structure using Hidden Markov Models, for the nectar-feeders: *M. harrisoni*, *L. nivalis* and *L. yerbabuena*; the fruit bat *P. alecto*; and *D. rotundus* and *M. waterhousii* (Fig. 6a; see Additional file 1, Table S10). *ACCS* protein structure is composed by 662 amino acids, two domains and 96 atoms of beta strand, 181 atoms of alpha helix and 4,391 loop atoms (Fig. 6a). Moreover, we performed a multi-comparison of the 3D structure for all the species mentioned above (see Additional file 1: Table S11). We found high similarity in the 3D protein structure for all bats (RSMD values from 0 to 0.003). However, we identified three residues of alpha-helix shared only for the Glossophaginae clade, and a beta strand shared only between *M. harrisoni* and *P. alecto* (Fig. 6b; see Additional file 1, Table S11).

Discussion

Our study provides unprecedented knowledge on the genomic signatures behind the dietary diversification and specialization in Phyllostomid bats. Surprisingly, and contrary to our first prediction, we found that many of the genomic characteristics of the ancestral Phyllostomid diet, remained functional in all lineages of the family. For example, the *Chitinase* gene was functionally conserved in most genomes (with the exception of the vampire), highlighting the relevance of digestion and nutrient uptake from insects in all lineages, including those that mainly feed on fruit, nectar-pollen and even in some cases, those that feed only on blood (Fig. 3 and 4) (9, 10, 20).

However, as previously studies suggest, we found that *Trehalase* (involved in the digestion of trehalose from insect blood, comprising approximately to 7% of their dry mass) is missing or a pseudogene for all the blood, fruit, and nectar-feeding bats, which may be a result of dietary diversification in the family (37).

Our second prediction was supported: we found unique genomic specializations in bats with obligate and restrictive diets, such as the vampire and the nectar-pollen feeders (Fig. 3 and 4, Table 3). The vampire's genome has unique characteristics associated with the ability to consume blood, including genes that play a crucial role in the down-regulation of fibrinolysis and those that control the production of blood cells (Fig. 4) (12, 21, 38). On the other hand, we also detected strong positive selection in genes crucial for carbohydrate oxidation, ATP production, and in genes involved in ketone metabolism in the three Glossophagini bats. These are associated with the extreme energetic feat of feeding on the wing (14, 30, 39) (Fig. 4). Our analyses also highlight the importance of genes involved in iron storage for animals that feed on iron deficient sources. These results might be related to the avoidance of metabolic disorders such as anemia (Fig. 4) (28). The results also help to explain how bats that feed on nectar-pollen can avoid the potentially adverse effects of their peculiar diet. In humans, loss-of-function due to mutations in some of these genes are associated with nutrient malabsorption and metabolism disorders including diabetes, hyperglycemia and obesity (30, 39).

In support of our third prediction, we identified signatures of molecular parallel evolution shared by fruit-feeding Pteropodids and nectar-pollen feeding Glossophagini bats (Fig. 5ab). The ancestral sequence reconstruction provided us with insights into the mechanisms of molecular adaptation and functional divergence. Signals of parallel evolution and adaptative selection for the proteins *ACCS* and *ALKBH7*, shed light on the importance of the storage of fatty fuels necessary to meet the energy demands of an expensive mode of foraging and pollinator ecology of these specialist bats. Protein function is more likely to be affected if genes show many radical substitutions in conserved positions, and signal of positive selection. In spite of the evolutionary changes detected for *ACCS* in the Glossophagini species, their tertiary protein structure exhibited high similarity when we compared it with other bats (see Additional file 1, Table S11). However, we identified exclusive differences in alpha-helix and beta-strand regions (Fig. 6b), that may be important in the protein function for the nectar-pollen feeders and the Old World fruit bat (Fig. 6). We consider that it is very probable that *ACCS* is up or down regulated. Future studies must evaluate the expression levels for this gene, and its regulation, including replicating and analyzing more tissues, such as the gut and liver (40, 41).

Surprisingly, the protein *UNC-45 B*, exhibited the same amino acid substitution between Glossophagini, Pteropodids and the dolphin. Lee *et al.*, (2018) have identified genes involved in muscle skeletal function and movement, with parallel substitutions shared between bats and

marine mammals (such as dolphin, whale and Baiji). We suggest that the protein *UNC-45 B* may be implicated in an efficient mobility and superfast muscle physiology for this species (42).

On the other hand, the molecular traits that we infer as the result of parallel evolution, were not found in the fruit-eating bat *A. jamaicensis*. We hypothesize that this species should be considered more omnivorous than strictly frugivorous (24, 41). Omnivory-frugivory might have been an important step in the transition to a more restricted fruit diet and to a nectar-pollen diet (Fig. 4 and 5) (5-8). To explore this hypothesis further, it will be necessary to expand our sample of genomes to include more Phyllostomid species that have more exclusively frugivorous fruits than *A. jamaicensis*.

Our findings suggest that parallel evolution due to nectar-feeding dietary specialization is likely a consequence of high metabolic demands required for foraging on flowers and fruits. These results are notable, given that the Pteropodidae and Glossophaginae lineages are separated by over 60 Mya (8) (see Additional file 2, Fig. S2) (17). Moreover, our results shed light on the evolutionary mechanisms and genomic shifts that take place in the transition to novel feeding habits. They also illuminate on the genomic changes that take place when animals adopt a diet dominated by sugar consumption, and with low levels of lipids and proteins. The generality of inferences can be tested in other nectar specialized vertebrate taxa, such as hummingbirds (40).

Finally, we found differences in the evolution of gene families and genes that are not necessarily or only related to diet among Phyllostomid lineages (Fig. 2 and Table 3). These differences are likely associated to other lineage-specific aspects in physiology, ecology (niche resources, interactions, immune system, etc.) and microbiomes (5, 6, 12, 43). As an example, positive selection in *IAP* enzyme was only detected in the nectar feeding bats, which supports a strong relationship between dietary specialization and the bacterial communities, that are involved in providing vitamins and aiding digestive processes (Fig. 4) (12). Future analyses should address the relationship between host diet-intestinal bacterial community, and the evolution of microbiomes across dietary diversification and specialization.

Methods

Animal sampling and genome sequencing

An adult male *Leptonycteris yerbabuena* (NCBI txid: 700936) was collected and processed on site at the cave “El Salitre” in Morelos state, Mexico (18°44'28" N, 99°10'46" W). All procedures were carried out in accordance with Federal Mexican Procedures (Guidelines of Secretaría de Medio Ambiente y Recursos Naturales, SEMARNAT), permit SGPA/DGVS/07161/15. The Zoology Museum “Alfonso L. Herrera” (Facultad de Ciencias, UNAM), donated the tissue samples from four leaf-nosed bats: *Macrotus waterhousii* (NCBI txid: 124750), *Artibeus*

jamaicensis (NCBI txid: 9417), *Leptonycteris nivalis* (NCBI txid: 59456) and *Musonycteris harrisoni* (NCBI txid: 148053) (Fig. 1; see Additional file 1, Table S1).

For all leaf-nosed bats species, we isolated their DNA using Phenol-Chloroform protocol and DNA Blood and Tissue Kit (Qiagen). We used the Illumina HiSeq 4000 150 PE platform to sequence the genomes (see Additional file 1, Table S1). We paid special attention to *L. yerbabuenae*, in order to use it as a reference to help in the assembly construction of the other genomes. In this species, we performed high whole genome sequencing (the DNA sample was sequenced on two lanes). Additionally, we used the fresh samples collected from *L. yerbabuenae* to obtain transcriptional evidence for the genome annotation, we extracted the RNA-Seq from five tissues: brain, pancreas, kidney, lung, and liver (reserved in a buffer storage of RNA stabilization) using the RNeasy Mini Kit (Qiagen). All tissues with RIN values ≥ 8 were sequenced on Illumina HiSeq 4000 150 PE platform.

***Leptonycteris yerbabuenae* genome assembly**

De novo genome assembly

The genome assembly was constructed *de novo* with Platanus v. 2.4.3 (44), using a heterozygous value = 0.04 (-u 0.04) and an initial kmer=32. To optimize and extend the genome assembly, we performed a scaffolding with MeDuSa software (45, 46). Finally, we used Pilon for correcting bases and polish the genome assembly (47).

We evaluated the genome assembly metrics (total length, number of scaffolds, number of contigs, L50, N50, and others). Moreover, with BUSCO v3 and the Mammalia odb9 database (48) we evaluated the measure for quantitative assessment of the genes content into the genome assembly.

Gene prediction

We performed TEdenovo from the REPET package (49) to predict, identify and annotate the transposable elements (TE) using the repetitive elements database Replibase (50). We masked the TEs across the genome using RepeatMasker v4.0.7 (see Additional file 1, Table S3) (51).

We also cleaned, filtered and assembled the RNA-Seq data of five tissues (brain, pancreas, kidney, liver, and lung) for the same individual with Trinity v4.4.7 (52). Based on the transcriptome assembly, we identified open-reading frames, coding sequences and their corresponding proteins (53).

We generated an *ab initio* gene prediction using Augustus v2.5.5 (54). To train Augustus, we used the gene structures of *E. fuscus* bat, and the RNA-Seq evidence (transcripts annotated from *L. yerbabuenae*). We performed a functional annotation by blastp using the UniProtKb SwissProt database and InterProScan (55, 56).

NW Leaf-Nosed bats reference genome construction (assembly and annotation)

We used the *L. yerbabuenae* assembly as a reference genome to construct the assembly of *M. waterhousii*, *A. jamaicensis*, *M. harrisoni* and *L. nivalis*.

All the raw data were filtered and cleaned (using a PHRED score ≥ 30) (see Additional file 1, Table S1). We followed the GATK v2.07 pipeline to identify Single Nucleotide Variants (SNVs) (57). First, to find all the SNVs along the genomic information from each Phyllostomid, based on the *L. yerbabuenae* genome assembly, all the high-quality genomic reads of each Phyllostomid were mapped to *L. yerbabuenae* genome assembly with BWA mem (58). Second, we used the GATK and Picard tools to recalibrate the genome mapping and identified the SNVs for each Phyllostomid (SortSam, MarkDuplicates, AddOrReplaceReadGroups, BuildBamIndex and CreateSequenceDictionary, RealignerTargetCreator, IndelRealigner, HaplotypeCaller, VariantFiltration, SelectVariants, BaseRecalibrator, AnalyzeCovariates, PrintReads, and VariantFiltration) (see Additional file 2, Fig. S3; Additional file 3, Methods) (56, 59). We constructed the consensus sequence based on the SNVs identified (with SAMtools, BCFtools, vcfutils.pl and SeqTK) (60-61). We evaluated for each consensus genome constructed, the assembly metrics and the integrity and gene content with BUSCO.

We predicted, identified and masked the TEs for each consensus genome using RepeatMasker v4.0.7 tool (50, 51). We performed the gene prediction with Augustus. The genomes were annotated with blast and InterProScan (see Additional file 3, Methods).

Phylogenomic and gene family analysis

A total of 132 single-copy orthologous genes (61,331 amino acids sites), across 18 mammals were concatenated to reconstruct a phylogenomic tree (best-fit model distribution JTT, +G +I+G and 80% consensus threshold) using PhyML3 (62) (see Additional file 3, Methods). We estimated molecular substitution rates with CODEML from the Phylogenetic Analysis by Maximum Likelihood, PAML package (63). Based on a Bayesian phylogenetic method, with MCMCtree tool, we estimated the species divergence times using fossil records from *Icaronycteris* (~50 Mya) and *Tachypteran* taxon with molecular ages of 64 Mya (16-18).

We used CAFE (64) to analyse the statistical changes in the gene family sizes using a birth and death estimator (λ and μ). Based on the distribution of observed family sizes, we calculated the *p-values* for gene family expansions and contractions. We carried out the gene family annotation with shell and PERL scripts (see Additional file 3, Methods).

dN/dS analysis using a branch-site model

Orthologous single copy genes and filtering

We used the proteins annotated of *M. waterhousii*, *A. jamaicensis*, *M. harrisoni*, *L. nivalis* and *L. yerbabuena* to create a database, incorporating the complete set of protein's of all Laurasiatheria species available in the ENSEMBL database (35 species), and the protein information of all bat species available in NCBI database (7 species) (see Additional file 1, Table S12). Based on a multi-species genome comparison with this database, we inferred the orthologous single copy genes using DIAMOND and Proteinortho programs (65, 66). We extracted all the single copy genes (scg) shared for each Phyllostomid bat, obtaining more than 9,637 scg clusters. We checked and removed all potential paralogous sequences, and ambiguous amino acids (letter X). Each single copy gene cluster was composed from 8 to maximum 20 sequences.

Each cluster was aligned with MAFFT aligner tool (67), we retained alignment sequences where the length is within 80 to 120% relative to the human and mouse sequences, and poorly aligned regions were removed by a visual inspection. We used the alignments and their corresponding coding sequences to perform a robust conversion of protein multi alignment into their corresponding codon alignments with PAL2NAL (see Additional file 3, Methods) (68). We reconstructed the phylogenetic tree for each cluster with RAXML (parameters -m GTRGAMMA -p 12345) (69).

dN/dS Test

We used the codon multi-alignment files and their corresponding phylogenetic tree to calculate synonymous sites and nonsynonymous sites (dN/dS) rates, using two bioinformatic tools: CODEML and HYPHY (63, 70).

With CODEML, we used a branch-site model, specifying the *foreground* branch (the species of our interest) and incorporating a null model [that assumes that background and the foreground branches share the same ratio (ω)]. We designated each Phyllostomid species as the foreground branch of our interest, and we performed independently the CODEML analysis. For assigning significance, we constructed the Likelihood Ratio Test (LRT) for each Phyllostomid result, using the likelihood values from the null and test model, and calculated the *p-value* ≤ 0.05 under a chi-square distribution. We also performed a *p-value* adjust, using the False Discovery Rate (FDR) correction, based on the likelihood ratio. Additionally, we considered under positive selection all those sequence sites with a posterior probability $> 95\%$ (by Bayes Empirical Bayes method).

With Hyphy, we used aBSREL (adaptive Branch-Site Random Effects Likelihood) (71). aBSREL infers the optimal number of ω to test if positive selection has occurred on a proportion of branches. The LRT is performed at each branch and compares the test model and null model. We inferred the optimal ω for all the branches for each single copy gene cluster (including bats and non-bats species).

We retained and classified those adaptive genes that were identified in both programs (CODEML and HYPHY), with a *p-value* ≤ 0.05 .

GO enrichment

We performed an enrichment analysis, using the weight01 algorithm with topGO v2.26 package from Bioconductor project in R (72-73). We obtained the statistical significance for the GO enrichment terms by performing the Fisher's exact test (*p-value* ≤ 0.01).

Radical amino-acid substitution in conserved positions

Ancestral sequence reconstruction

Based on the previously inferred orthologous genes, we extracted all single copy genes shared by the Glossophagini (Glsp): *M. harrisoni*, *L. nivalis*, *L. yerbabuena*; and the Pteropodids (Ptrp): *P. alecto*, *P. vampyrus* and *R. aegyptiacus*. Each single copy gene cluster was composed of between 12 and a maximum 30 sequences.

We obtained 1,918 clusters of orthologous sequences (including at least one Glsp and one Ptrp). Each cluster was aligned using PRANK (74) and we constructed their corresponding phylogenetic tree with RAxML (parameter -m PROTCATLG) (69). We checked all alignments for gaps and premature stop codons. We performed an ancestral sequence reconstruction using the protein alignments and phylogenetic trees. To dismiss incorrectly inferred residues and only retain the accurate ones for the reconstructed ancestral sequence, we used two different programs CODEML, that assumes a Markov process model and calculates a Bayesian empirical likelihood for each character at each sequence position (63), and FastML, that assumes a continuous time Markov process model and provides the posterior probabilities for each character at each sequence position (75). Both programs provide the ancestral sequence and the posterior probabilities distribution. For CODEML, we fixed the parameters: model =2, fix_alpha = 0, alpha = 0.5 and RateAncestor = 1.

Based on the ancestral sequence information, we identified pairs of branches for Glsp and Ptrp species that exhibited a parallel amino acid substitution. We also checked the ancestral state at each node for these substitutions. We classified each parallel substitution as a radical amino acid substitution in a conserved position, assuming two criteria: 1) parallel substitution is exclusive in the branches of Glsp, Ptrp and their corresponding nodes; and 2) different physicochemical properties between the most frequent amino acid state and the derived (parallel substitution).

We checked the Bayesian empirical likelihood at each ancestral state for all parallel substitution position. We retained only those parallel substitution with a posterior probability > 85%.

Finally, we constructed a phylogeny using the information of those genes with parallel evolution. We concatenated the amino acid sequence for these three genes. We aligned the sequences with MAFFT tool. We used ProtTest3 to select the best-fit model of protein evolution (76). The phylogenetic tree was constructed using a Maximum Likelihood method with RAxML (-p 12345 -m PROTCATLG).

Drivers of parallel evolution

In independent branches, mutation and selection can have equal impacts on patterns of parallel substitutions. For those proteins that exhibited parallel evolution, we also evaluated four variables at DNA and protein level: length, GC percent, rates at synonymous sites and nonsynonymous sites (dN/dS) and isoelectric point (77-78).

Protein modeling

To explore the effects of positive selection and the radical amino acid substitutions, we modeled the second and tertiary structure of the protein *Acetoacetyl CoA Synthetase* (ACCS) for *M. waterhousii*, *D. rotundus*, *M. harrisoni*, *L. nivalis*, *L. yerbabuena* and *P. alecto*. We used Phyre2 software (79), which compares the profile of the protein of our interest, with a protein database, using Hidden Markov Models and predicting secondary structure for each residue. To identify differences in the protein structure, we compared the secondary and tertiary structure between the nectar- fruit bats and *M. waterhousii* and *D. rotundus*, using the software PyMOL (80-81). To calculate the RMSD score, we aligned the PDB protein model between pairs.

Availability of data and materials

Supporting data including genome assemblies, genome and protein annotation, TE prediction and annotation, multifasta gene families, phylogenomic alignment, ancestral sequence reconstruction, and tertiary protein modelling (PDB) files, are available via the GigaScience database GigaDB [Gutiérrez-Guerrero YT, Ibarra-Laclette E, Martínez del Río C, Barrera-Redondo J, Rebollar EA; Ortega J, León-Paniagua L, Urrutia A, Aguirre-Planter E, Eguiarte LE. Supporting data for “Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats”. GigaScience Database. 2020. <http://dx.doi.org/10.5524/100746>] (82).

The whole genome and RNA-Seq sequence information analysed during the current study are available in the National Center for Biotechnology Information (NCBI): whole genome assembly for *L. yerbabuena* within BioProject: PRJNA542899 and SRA: SRR9076597. The RNA-Seq data is available within Bioproject: PRJNA543325. Raw genome data of *M. waterhousii*, *A. jamaicensis*, *M. harrisonii* and *L. nivalis* are available in the SRA: SRR908760, SRR9087866, SRR9089318 and SRR9089325.

Declarations

Ethics approval and consent to participate

The use of animals in this study was performed in accordance with the Federal Mexican Procedures: Guidelines of Secretaría de Medio Ambiente y Recursos Naturales, SEMARNAT, with permit SGPA/DGVS/07161/15.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Y.T.GG is supported by a doctoral scholarship from the Comisión Nacional de Ciencia y Tecnología (CONACyT, Beca Mixta grant no. 291250). The study was supported by a grant from Fronteras de la Ciencia, "Genómica de la Diversidad de Vertebrados 1: *Leptonycteris* y la evolución de la nectarivoría en murciélagos y aves" (CONACyT, project no. 177) to L.E.E.

Author's Contributions

YT.G.G and L.E.E designed and performed research. YT.G.G and L.E.E wrote the paper with contributions from C.M.R, E.A.R, A.U, J.O, E.A.P and E.I.L helped with computational resources and bioinformatic analyses. J.B.R helped with bioinformatic analyses. L.L.P donated the samples. All authors revised and edited the manuscript.

Acknowledgments

This manuscript constitutes part of the doctoral project of the first author, who thanks the Posgrado en Ciencias Biomédicas (Universidad Nacional Autónoma de México, UNAM) and acknowledges the doctoral scholarship supported by Comisión Nacional de Ciencia y Tecnología (CONACyT, grant no. 580116; Beca Mixta grant no. 291250), and the grant Fronteras de la Ciencia (CONACyT, project no. 177). The authors wish to acknowledge the Instituto de Ecología (UNAM), Comisión Nacional para el Conocimiento de la Biodiversidad (CONABIO) and Instituto de Ecología A.C. (INECOL, A.C) for computing resources. Special thanks to MSc. R. Trejo-Salazar, MSc. O. Gaona, Biol. A. Galicia for their assistance during fieldwork in the El Salitre cave. We thank Dra. L. Espinosa Asuar, Dra. V. Souza, S. Barrientos and all the Lab. Evolución Molecular y Experimental (UNAM) for their help in lab work. To MSc. E. Villafán, Ing. Rodrigo

García and Mat. E. Campos for their computational assistance. We gratefully acknowledge Dr. S. Ramírez-Barahona for their constructive comments on the manuscript. We acknowledge MSc. D. Zamora-Mejías, R. Medellín-Legorreta and Biol. Melissa E. Rodríguez for providing us pictures of Phyllostomid bat species. We gratefully acknowledge BioCurator Chris Hunter for his technician support to prepare our GigaDB dataset.

References

1. Hunter JP. Key innovation and ecology of macroevolution. *Trends Ecol Evol.* 1998;3:31–36.
2. Yoder JB, Clancey E, Des Roches S, Eastman JM, Gentry L, et al. Ecological opportunity and the origin of adaptive radiations. *J Evol Biol.* 2010;23:1581–1596.
3. Palm W, Thompson CB. Nutrient acquisition strategies of mammalian cells. *Nature.* 2017;546:234–242.
4. Borstein SR, Fordyce JA, O'Meara BC, Wainwright PC, McGee MD. Reef fish functional traits evolve fastest at trophic extremes. *Nat Ecol Evol.* 2018;3:191-199.
5. Rojas D, Vale A, Ferrero V, Navarro L. When did plants become important to leaf-nosed bats? Diversification of feeding habits in the family Phyllostomidae. *Mol Ecol.* 2011;20:2217-28.
6. Rossoni DM, Assis APA, Giannini NP, Marroig G. Intense natural selection preceded the invasion of new adaptive zones during the radiation of New World leaf-nosed bats. *Sci Rep.* 2017;7:1–11.
7. Yohe LR, Velazco PM, Rojas D, Gerstner BE, Simmons NB, Dávalos LM. Bayesian hierarchical models suggest oldest known plant-visiting bat was omnivorous. *Biol Lett.* 2015;11:20150501.
8. Fleming TH, Geiselman C, Kress WJ. The evolution of bat pollination: a phylogenetic perspective. *Ann Bot.* 2009;104:1017–1043.
9. Cole FR, Wilson DE. *Leptonycteris yerbabuena*. *Mamm Species.* 2006;797:1–7.
10. Gonzalez-Terrazas TP, Medellín RA, Knörnschild M, Tschapka M. Morphological specialization influences nectar extraction efficiency of sympatric nectar-feeding bats. *J Exp Biol.* 2012;215:3989–3996.
11. Hong W, Zhao H. Vampire bats exhibit evolutionary reduction of bitter taste receptor genes common to other bats. *Proc R Soc B Biol Sci.* 2014;281:20141079.
12. Zepeda-Mendoza ML, Xiong Z, Escalera-Zamudio M, Runge AK, Thézé J, et al. Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat Ecol Evol.* 2018;2:659–668.
13. Ayala-Berdon AJ, Schondube JE. A physiological perspective on nectar-feeding adaptation in Phyllostomid bats. *Physiol Biochem Zool.* 2015;84:661541.

14. Schondube JE, Herrera-M, LG, Martínez del Rio C. Diet and the evolution of digestion and renal function in phyllostomid bats. *Zoology (Jena)*. 2001;104:59–73.
15. Zhao H, Zhou Y, Pinto CM, Charles-Dominique P, Galindo-González J, et al. Evolution of the Sweet Taste Receptor Gene *Tas1r2* in Bats Research article. *Mol Biol Evol*. 2010;27:2642–2650.
16. Simmons NB, Geisler JH. Phylogenetic relationships of *Icaronycteris*, *Archaeonycteris*, *Hassianycteris*, and *Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. In: Bull Am Museum Nat Hist. 1998. <http://hdl.handle.net/2246/1629>. Accessed 01 Sep 2019.
17. Teeling EC, Springer MS, Madsen O, Bates P, O'brien SJ, et al. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science*. 2005;307:580-4.
18. Simmons NB, Seymour KL, Habersetzer J, Gunnell GF. Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature*. 2008;451:818-821.
19. Chen EH. Invasive Podosomes and Myoblast Fusion. *Curr Top Membr*. 2011;68:235-258.
20. Nardelli A, Vecchi M, Mandrioli M, Manicardi GC. The evolutionary history and functional divergence of trehalase (*treh*) genes in insects. *Front Physiol*. 2019:00062.
21. Okamoto T, Tanigami H, Suzuki K, Shimaoka M. Thrombomodulin: A Bifunctional Modulator of Inflammation and Coagulation in Sepsis. *Crit Care Res Pract*. 2012(Dic):1–10.
22. Tsuchida T, Fukuda S, Aoyama H, Taniuchi N, Ishihara T, et al. MGAT2 deficiency ameliorates high-fat diet-induced obesity and insulin resistance by inhibiting intestinal fat absorption in mice. *Lipids Health Dis*. 2012;11:1–10.
23. Xiong S, Tu H, Kollareddy M, Pant V, Li Q, et al. Pla2g16 phospholipase mediates gain-of-function activities of mutant p53. *Proc Natl Acad Sci USA*. 2014;30:11145-11150.
24. Kunz TH, Díaz CA. Folivory in Fruit-eating Bats, with New Evidence from *Artibeus jamaicensis* (Chiroptera: Phyllostomidae). *Biotropica*. 1995;27:106–120.
25. Li C, Chen P, Vaughan J, Lee K-F, Vale W. Urocortin 3 regulates glucose-stimulated insulin secretion and energy homeostasis. *Proc Natl Acad Sci USA*. 2007;104:4206–4211.
26. Teicher B. CD248: A therapeutic target in cancer and fibrotic diseases. *Oncotarget*. 2019;10:993-1009.
27. Drysdale J, Arosio P, Invernizzi R, Cazzola M, Volz A, et al. Mitochondrial ferritin: A new player in iron metabolism. *Blood Cells Mol Dis*. 2002;29:376–383.
28. Geng Y, Dong Y, Yu M, Zhang L, Yan X, et al. Follistatin-like 1 (*Fstl1*) is a bone morphogenetic protein (BMP) signaling antagonist in controlling mouse lung development. *Prot Natl Acad Sci USA*. 2011;17:7058-7063.
29. Lallès JP. Intestinal alkaline phosphatase: Multiple biological roles in maintenance of intestinal homeostasis and modulation by diet. *Nutr Rev*. 2010;68:323–332.

30. Sylow L, Kleinert M, Richter EA, Jensen TE. Exercise-stimulated glucose uptake - regulation and implications for glycaemic control. *Nat Rev Endocrinol*. 2016;13:133-148.
31. Griffin BA. Lipid metabolism. *Surgery (Oxford)*. 2013;31:267-272.
32. Shi L, Tu PB. Acetyl-CoA and the regulation of metabolism: mechanisms and consequences. *Curr Opin Cell Biol*. 2015;33:125-131.
33. Solberg A, Robertson AB, Aronsen JM, Rognmo O, Sjaastad I, et al. Deletion of mouse Alkbh7 lead to obesity. *J Mol Cell Biol*. 2013;5:194-203.
34. Guzmán C, Benet M, Pisonero-Vaquero S, Moya M, García-Mediavilla M, et al. The human liver fatty acid binding protein (FABP1) gene is activated by FOXA1 and PPAR α ; and repressed by C/EBP α : Implications in FABP1 down-regulation in nonalcoholic fatty liver disease. *BBA Mol Cell Biol L*. 2013;4:803-818.
35. Srivastava R, Pinkosky SL, Filippov S, Hanselman JC, Cramer CT, et al. AMP-activated protein kinase: an emerging drug target to regulate imbalances in lipid and carbohydrate metabolism to treat cardio-metabolic diseases. *J Lipids Res*. 2012;53:2490-2514.
36. Kachur TM, Pilgrim DB. Myosin assembly, maintenance and degradation in muscle: Role of the chaperone UNC-45 in myosin thick filament dynamics. *Int J Mol Sci*. 2008;9:1863-1875.
37. Jiao H, Zhang L, Xie HW, Simmons NB, Lui H, et al. Trehalase gene as a molecular signature of dietary diversification in mammals. *Mol Biol Evol*. 2019;10:2171-2183.
38. Qu J, Ko CW, Tso P, Bhargava A. Apolipoprotein A-IV: A multifunctional protein involved in protection against atherosclerosis and diabetes. *Cells*. 2019;4:319.
39. Jensen J, Rustad PI, Kolnes AJ, Lai YC. The role of skeletal muscle glycogen breakdown for regulation of insulin sensitivity by exercise. *Front Physiol*. 2011;2:1–11.
40. Workman RE, Myrka M, Wong GM, Tseng E, Welch KC Jr. et al. Single-molecule, full length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird *Alchilochus colubris*. *GigasScience*. 2018;3:1-12.
41. Moreno-Santillán DD, Machain-Williams C, Hernández-Montes G, Ortega J. De Novo Transcriptome Assembly and Functional Annotation in Five Species of Bats. *Sci Rep*. 2019;9:6222.
42. Lee JH, Lewis KM, Mural TW, Kirilenko B, Borgonovo B, et al. Molecular parallelism in fast-twitch muscle proteins in echolocating mammals. *Science*. 2018;4:eaat9660.
43. Song SJ, Sanders JG, Delsuc F, Metcalf J, Amato K, et al. Comparative analyses of vertebrates gut microbiomes reveal convergence between birds and bats. *MBio*. 2020;1:e02901-19.
44. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res*. 2014;24:1384–1395.

45. Lowe T. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 1997;25:955–964.
46. Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, et al. MeDuSa: A multi-draft based scaffold. *Bioinformatics.* 2015;31:2443–2451.
47. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 2014;0112963.
48. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31:3210–3212.
49. Flutre T, Duprat E, Feuillet C, Quesneville H. Considering transposable element diversification in de novo annotation approaches. *PLoS One.* 2011;6:0016526.
50. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA.* 2015;6:4–9.
51. Tarailo-Graovac, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics.* 2009;4:bi0410s25.
52. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 2013;8:1494–1512.
53. Evans T, Loose M. AlignWise: a tool for identifying protein-coding sequence and correcting frame-shifts. *BMC Bioinformatics.* 2015;1:376.
54. Stanke M, Morgenstern B. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* 2005;33:465–467.
55. The UniProt Consortium. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* 2018;46:2699.
56. Jones P, Binns D, Chang HY, Fraser M, Li W, et al., InterProScan 5: Genome-scale protein function classification. *Bioinformatics.* 2014;30:1236–1240.
57. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, et al. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics.* 2013;43:11.10.1-33.
58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;14:1754-1760.
59. Broad Institute. Picard Tools. Broad Institute, Github repository. <https://github.com/broadinstitute/picard/issues/808>.
60. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;16:2078–2079.

61. Danecek P, Auton A, Abecasis G, Alberts CA, Banks E, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;15:2156–2158.
62. Guindon S, Gascuel O. PhyML: “A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood”. *Syst Biol*. 2003;52:696-704.
63. Yang Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol*. 2007;24:1586–1591.
64. De Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: A computational tool for the study of gene family evolution. *Bioinformatics*. 2006;22:1269–1271.
65. Lechner M, Findeib SS, Steiner L, Marz M, Stadler PF. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics*. 2011;12:124.
66. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;1:59-60.
67. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;14:3059-3066.
68. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 2006;34:W609-W612.
69. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30:1312-1313.
70. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005;5:676-9.
71. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K. et al. Less is More: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol*. 2015;5:132-1353.
72. Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. 2006;22:1600-7.
73. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2008. <http://www.R-project.org>.
74. Löytynoja A. Phylogeny-aware alignment with PRANK. *Methods Mol Biol*. 2014;1079:155-70.
75. Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, et al. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res*. 2012;40:W580-4.
76. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 2011;8:1164-1165.
77. Kozłowski LP. IPC – Isoelectric Point Calculator. *Biol Direct*. 2016;11:55.

78. Bailey SF, Guo Q, Bataillon T. Identifying drivers of parallel evolution: A regression model approach. *Genome Biol Evol.* 2018; 10:2801-2812.
79. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10:845-858.
80. DeLano WL. PyMOL: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography.* 2002;40:82-92.
81. The PyMOL Molecular Graphics System, Version 2, Schrödinger, LLC. <https://pymol.org/2/>
82. Gutiérrez-Guerrero YT, Ibarra-Laclette E, Martínez del Río C, Barrera-Redondo J, Rebolgar EA; Ortega J, León-Paniagua L, Urrutia A, Aguirre-Planter E, Eguiarte LE. Supporting data for “Genomic consequences of dietary diversification and parallel evolution due to nectarivory in Leaf-nosed bats”. *GigaScience Database.* 2020. <http://dx.doi.org/10.5524/100746>.

Figures

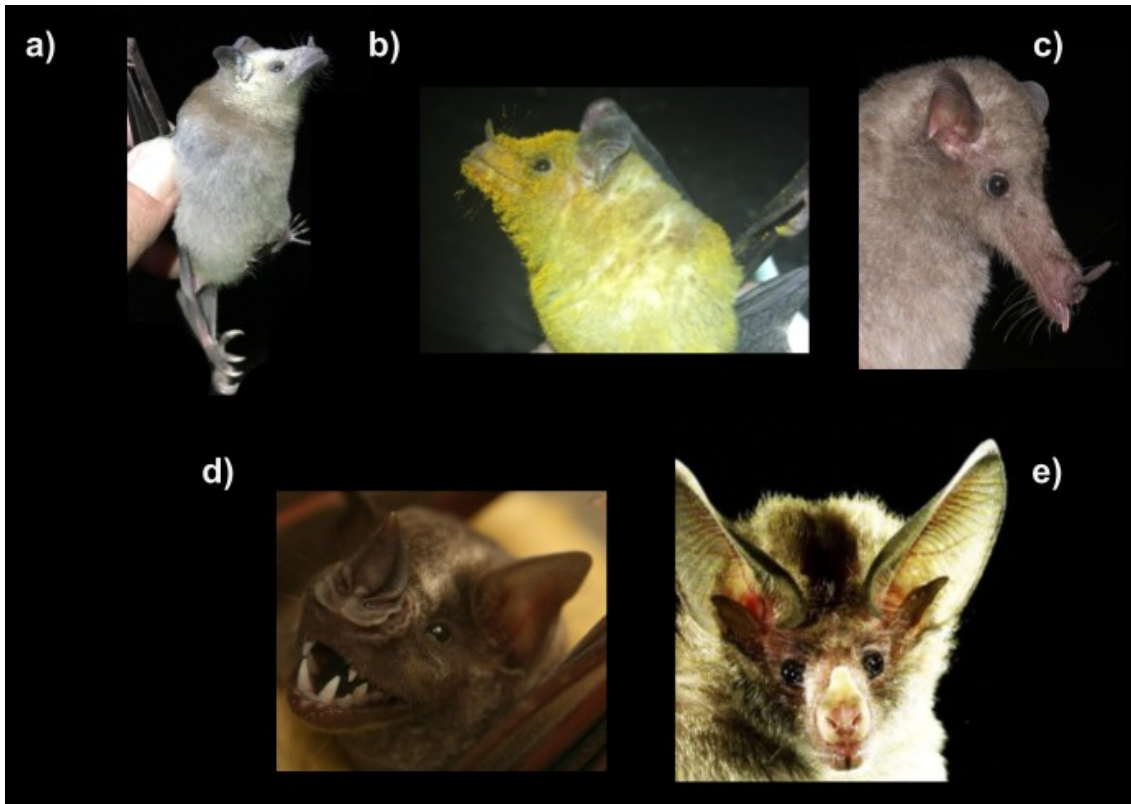


Figure 1. Phyllostomid species. a) *Leptonycteris yerbabuena* , b) *Leptonycteris nivalis*, c) *Musonycteris harrisoni*, d) *Macrotus waterhousii*, and e) *Artibeus jamaicensis*. Photo credits: a - b) Daniel Zamora-Mejías, c) Rodrigo Medellín-Legorreta, d) Melissa E. Rodríguez, and e) Wikimedia, public domain.

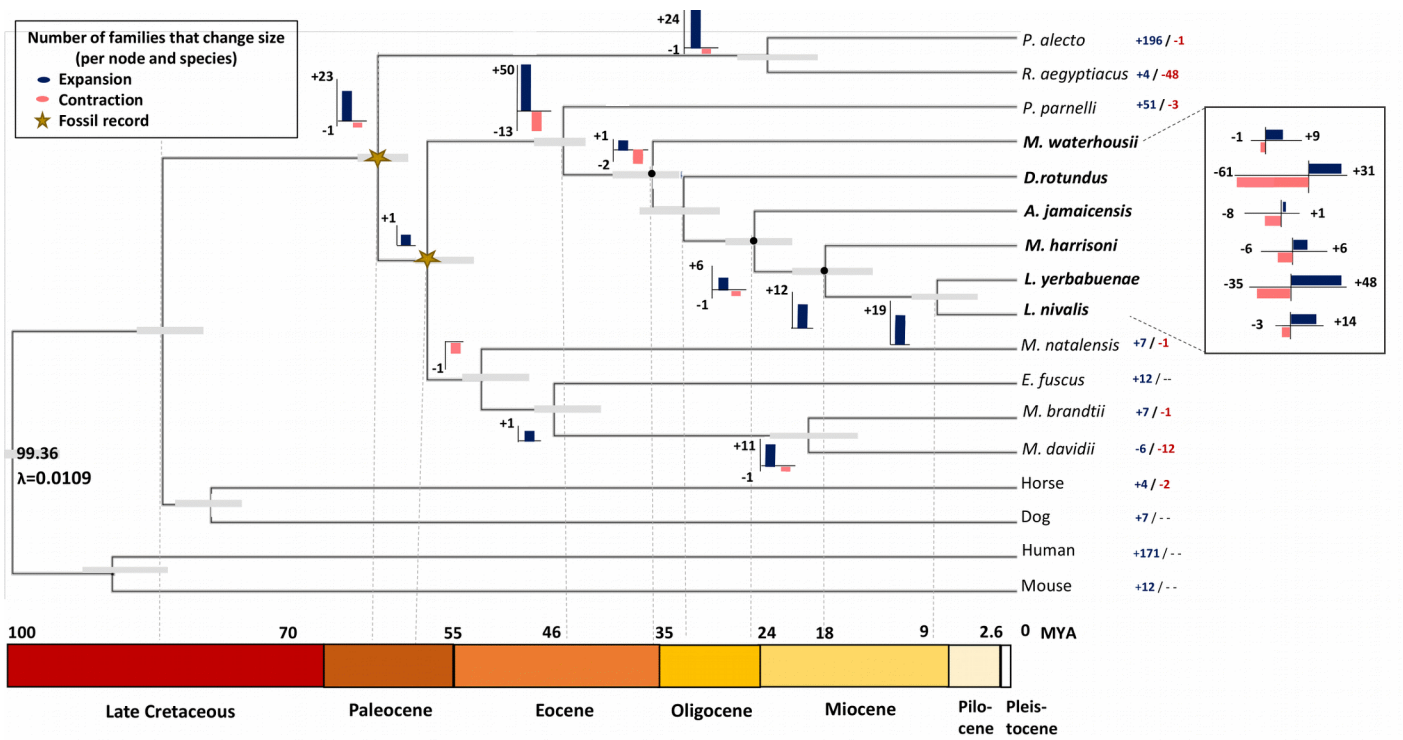


Figure 2. Phylogenetic tree constructed with 132 single copy genes and estimates of divergence times based on two fossil records (yellow stars) (see Methods). Based on 22,388 gene families we analyzed the number of orthologous families expanded (+ blue) and contracted (- red) across the phylogeny: per node (bars) and per species branch (right), with a p -value ≤ 0.01 . Gray bars reflect the divergence time interval based on 95% HPD.

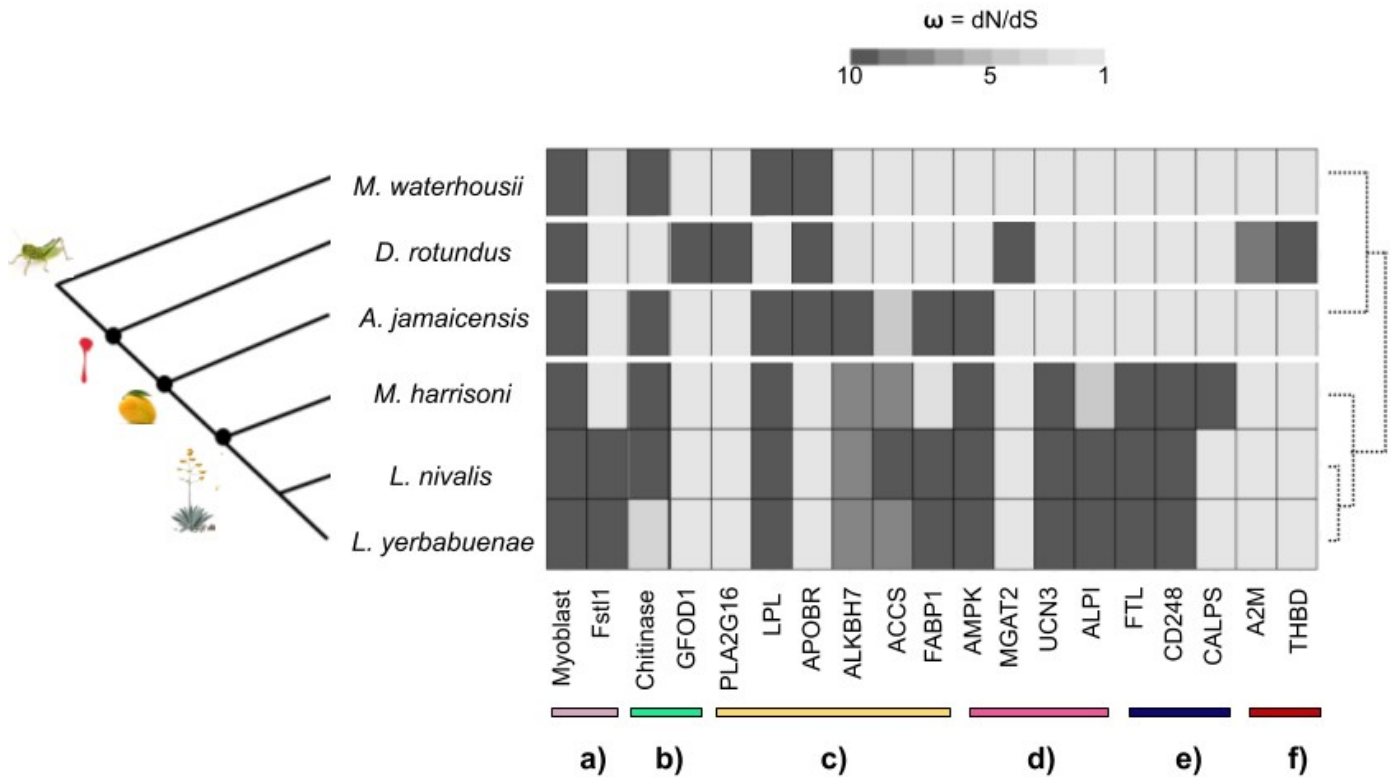


Figure 3. Positive selection in genes and proteins across the phylogeny of Phyllostomid bats, in comparison to the insect-feeder bat *Macrotus waterhousii*. Most of positive selected genes likely contribute to the regulation and processing of: a) muscle and bone development; b) carbohydrates; c) lipids; d) nutrients and food uptake; e) iron storage and calcium sources; and f) blood regulation (see genes and proteins abbreviations in Additional file 1, Table S14).

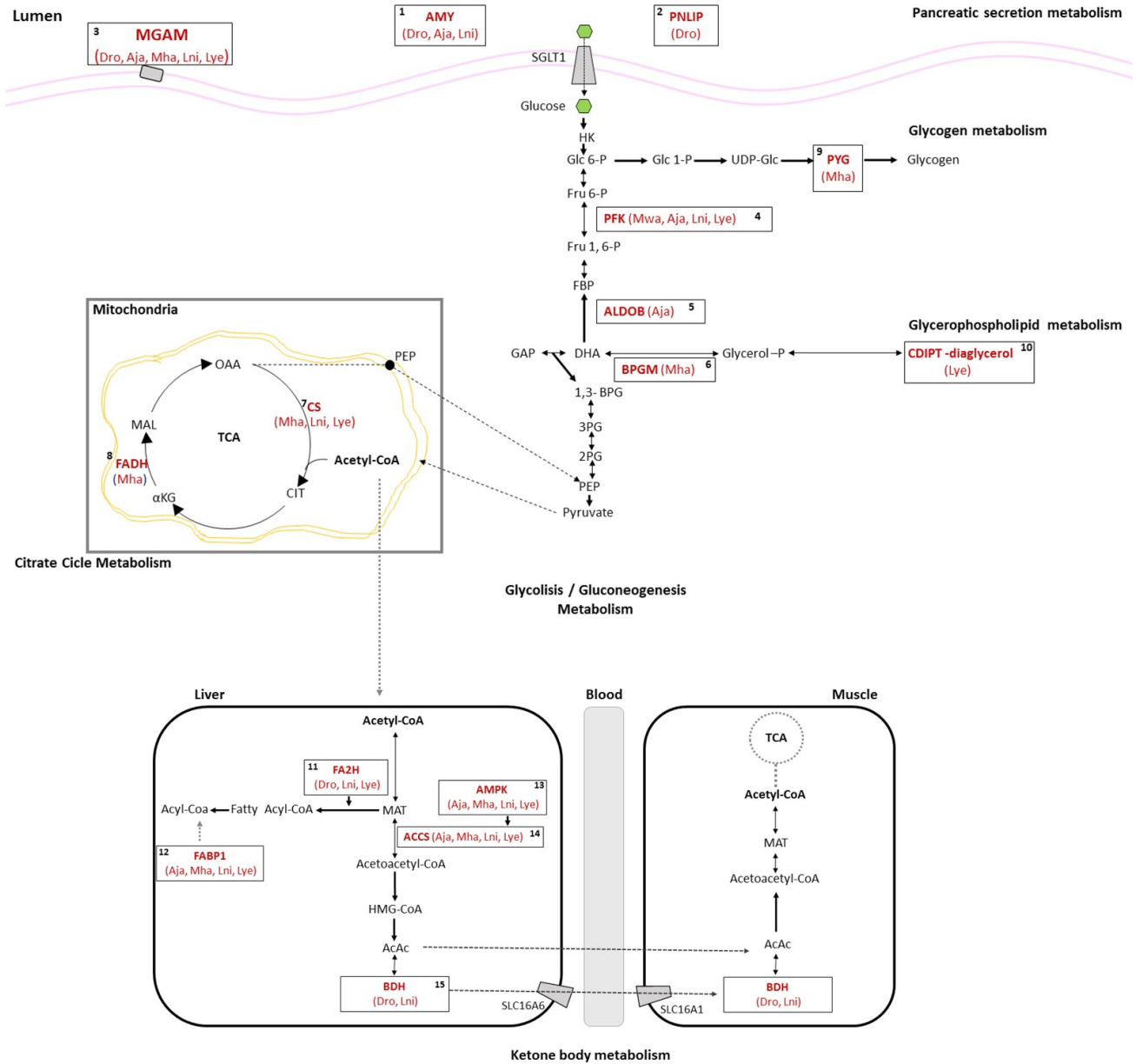


Figure 4. A subset of genes under positive selection (in red bold) that are involved in glucose and ketogen in the frugivorous (Aja: *A. jamaicensis*) and nectar-pollen bats (Mha: *M. harrisoni*, Lni: *L. nivalis*, and Lye: *L. yerbabuena*). The diagram also identifies adaptative signals for some genes in the vampire *D. rotundus* (Dro) and the insectivore *M. waterhousii* (Mwa). The diagram is based on the KEGG metabolic pathways database and a review of the literature (see gene and proteins abbreviations in Additional file 1, Table S8).

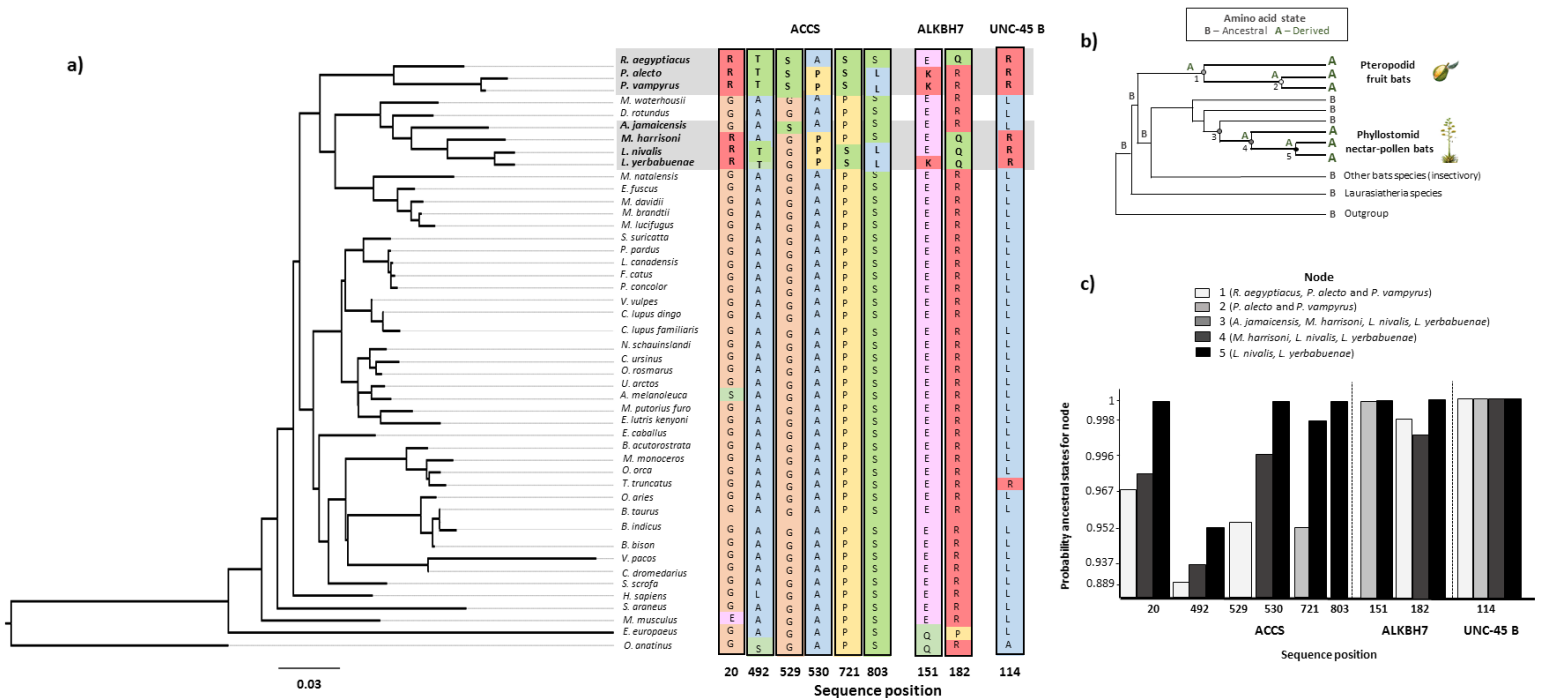


Figure 5. Parallel molecular evolution between Pteropodids (Old World) and Glossophagini (New World) bats, in three genes: Acetate-CoA Synthetase (AACs); Alpha-Ketoglutarate-Dependent Dioxygenase Homolog 7, mitochondrial (*ALKBH7*); and UNC-45 homolog B, mitochondrial (*UNC-45 B*). a) Phylogeny reconstruction for these three genes by Maximum likelihood (using 1,827 amino acids), for 47 mammal species. b) Ancestral sequence reconstruction (for branches and nodes) to infer parallel substitutions in conserved positions for the three genes. c) Probability of replacement at each ancestral state node for each sequence position. Amino acid abbreviations: A – Alanine (non-polar); T – Threonine (polar); Q – Glutamine (polar); R – Arginine (basic-charged); K – Lysine (basic-charged); E – Glutamic acid (acidic + charged); S – Serine (polar) and L – Leucine (non-polar).

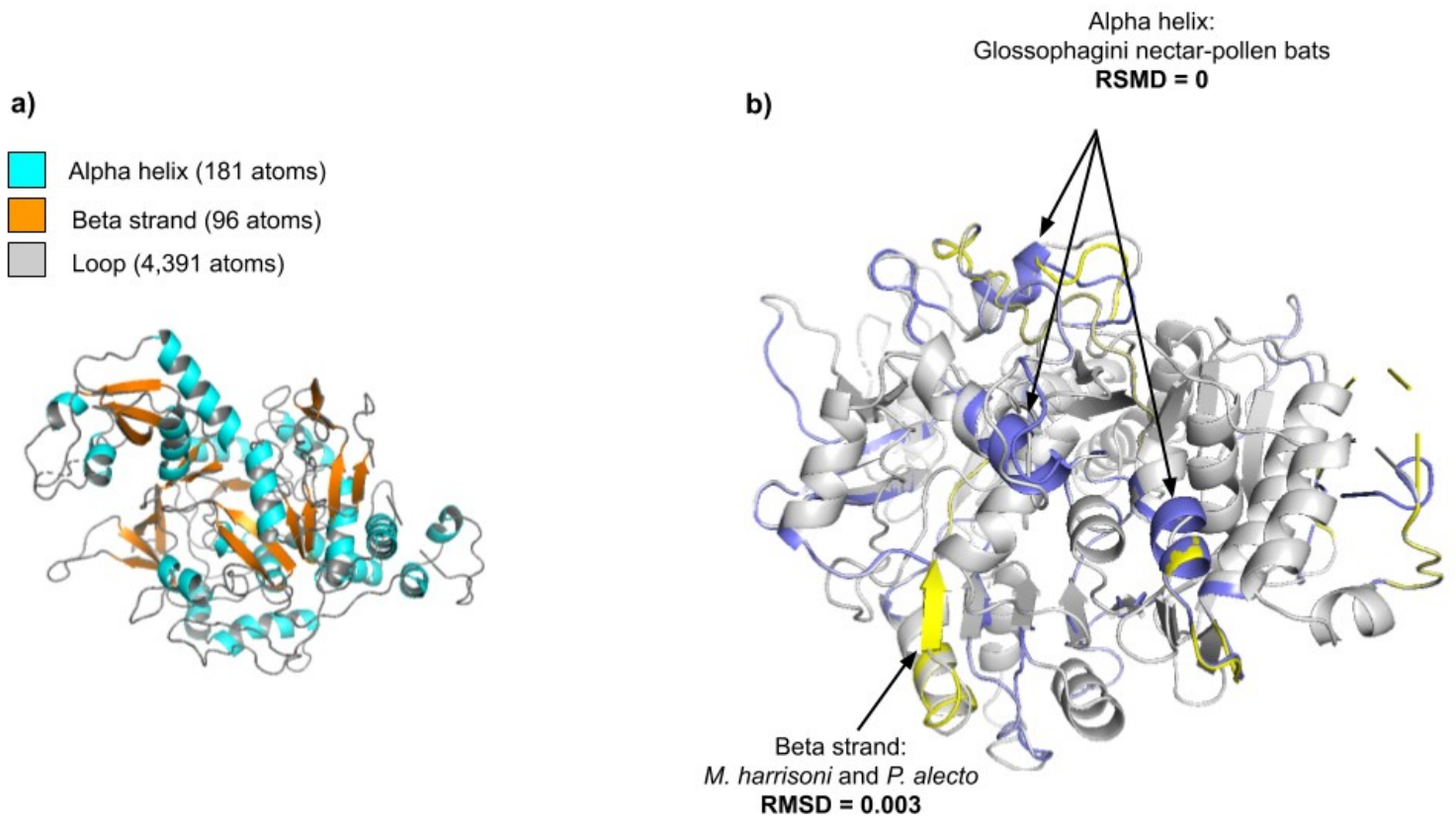


Figure 6. ACCS protein structure. a) 3D- structure of ACCS protein for *L. yerbabuena*. b) In gray: ACCS 3D- structure consensus (*M. waterhousii* and *D. rotundus*). In blue: alpha-helix structures shared only for the three Glossophagini nectar-pollen feeders (*M. harrisoni*, *L. nivalis* and *L. yerbabuena*). In yellow: beta-strand shared only between *M. harrisoni* and *P. alecto* (Pteropodid bat). RMSD score (protein 3D superposition and alignment) between pairs of species (see Additional file 1, Table S10).

Tables

Table 1. Global statistics for the nectar-pollen feeding bat *L. yerbabuena* genome assembly.

i) Sequencing	Total raw data (Gb)	Number Reads > PHRED 30	Coverage (x)
	254.4	690,759,531	103.6
ii) Assembly	N50 (Kb) - L50	Number – Longest (Mb)	Total Size (Gb)
Contig	69.49 - 8,805	78,626 – 0.55	2.05
Scaffold	14,735.1 – 38	34,419 – 70.81	2.05
BUSCO	Completed	Fragmented	Missing
	3,864 (94.1%)	103 (2.5%)	141 (3.4%)
iii) Annotation	Number	Length (Mb)	Percent
Exons	119,036	--	--
CDS / Proteins	24,074	--	--
Repeats	3,010,348	547.05	26.64%

Table 2. Mapping statistics and SNP identification in Phyllostomid bat genomes (based on *L. yerbabuenae* genome assembly).

	Phyllostomidae			
Species	<i>Leptonycteris nivalis</i>	<i>Musonycteris harrisoni</i>	<i>Artibeus jamaicensis</i>	<i>Macrotus waterhousii</i>
Diet	Nectar-pollen	Nectar-pollen	Fruits	Insects
Total data (Gb)	131.4	69.2	56.6	128.4
Coverage (x)	54.8	30.45	25	56.34
BUSCO				
Complete	93.7%	94.0%	93.0%	93.3%
Fragmented	2.8%	2.4%	3.5%	3.4%
Missing	3.5%	3.6%	3.5%	3.3%
CDS/Proteins	24,471	20,135	18,756	19,171

Table 3. GO enrichment for significant gene families per node and habit food. Gene Ontologies (GO) annotations involved in metabolism and diet are in bold (+ Gene families expansions; - contractions).

Specie and Nodes	Function and Metabolic Pathway	GO	p-value < 0.01
Phyllostomid node (expansions)	+ Structural constituent of ribosome + Translation	GO:0003735 GO:0006412	<1e-30 <1e-30
Phyllostomid node (contractions)	- Hydrolase activity - Lipid metabolic process - Aspartic-type endopeptidase activity	GO:0016788 GO:0006629 GO:0004190	0.00099 0.001 1.8e-06
<i>D. rotundus</i> (expansions)	+ Response to biotic stimulus + Defense response + Signal transduction + Nitrate assimilation + Regulation of appetite + Protein glycosylation in Golgi + GTPase activity + Molybdenum ion binding	GO:0009607 GO:0006952 GO:0007165 GO:0042128 GO:0032098 GO:0033578 GO:0003924 GO:0030151	5.7e-12 7.8e-12 1.3e-08 8.8e-08 0.00034 0.00313 < 1e-30 4.5e-05
(contractions)	- Translation - Calcium ion transmembrane - Cellular calcium ion homeostasis - Neuron development - Microtubule-based process - Homophilic cell adhesion via plasma - Peptidyl-prolyl cis-trans isomerase - Ephrin receptor activity - Ryanodine-sensitive calcium channel - Inorganic anion exchanger activity - Ionotropic glutamate receptor activity - Voltage-gated calcium channel activity	GO:0006412 GO:0070588 GO:0006874 GO:0048666 GO:0007017 GO:0007156 GO:0003755 GO:0005003 GO:0005219 GO:0005452 GO:0004970 GO:0005245	1e-30 1.3e-16 3.5e-09 4.2e-09 5.5e-08 1.7e-07 < 1e-30 2.7e-21 6.2e-21 2.1e-20 6.3e-20 5.6e-16
<i>A. jamaicensi</i> , <i>M. harrisoni</i> , <i>yerbabuena</i> and <i>L. nivalis</i> (expansions)	+ Translation + Integral component of membrane + Immune response 458 + Iron ion import membrane + HFE-transferrin receptor complex + Transferrin receptor binding	GO:0006412 GO:0016021 GO:0006955 GO:0098711 GO:1990712 GO:1990459	<1e-30 6.7e-06 1.1e-05 0.00018 5.7e-05 7.3e-05
(contractions)	- Protein peptidyl-prolyl isomerization	GO:0000413	< 1e-30
<i>M. harrisoni</i> , <i>L.</i>	+ Protein deubiquitination	GO:0016579	2.8e-09

<i>yerbabuenae</i> and <i>L. nivalis</i> (expansions)	+ Virion assembly 24 + Structural constituent of ribosome + Thiol-dependent ubiquitinyl hydrolase + Transferrin receptor binding + HFE-transferrin receptor + Iron ion import membrane	GO:0019068 GO:0003735 GO:0036459 GO:1990459 GO:1990712 GO:0098711	0.00027 < 1e-30 3.6e-11 0.0099 0.0099 0.0056
---	--	---	--

Additional file

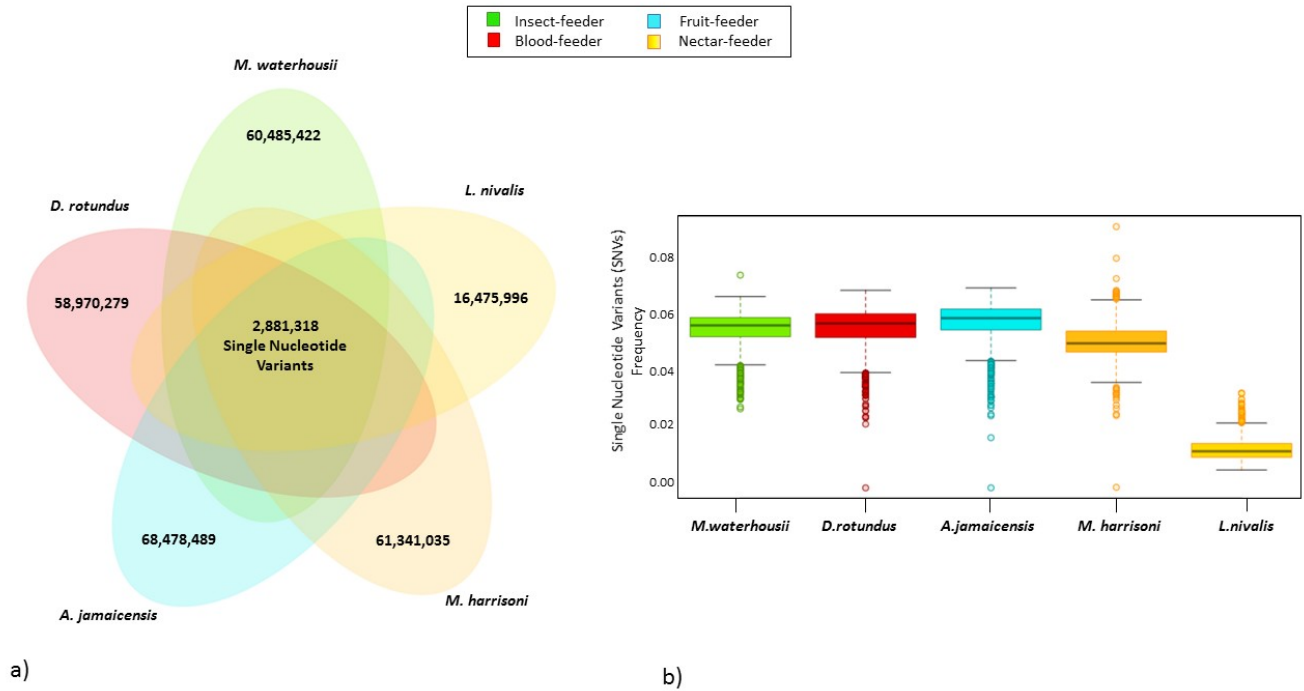


Figure S1. a) Venn diagram showing the number of shared and unique *Single Nucleotide Variants* for each Phyllostomid bat based on *L. yerbabuenae* genome. b) Nucleotide diversity estimated across sliding windows of 1 Mbp for each Phyllostomid, based on *L. yerbabuenae* genome assembly.

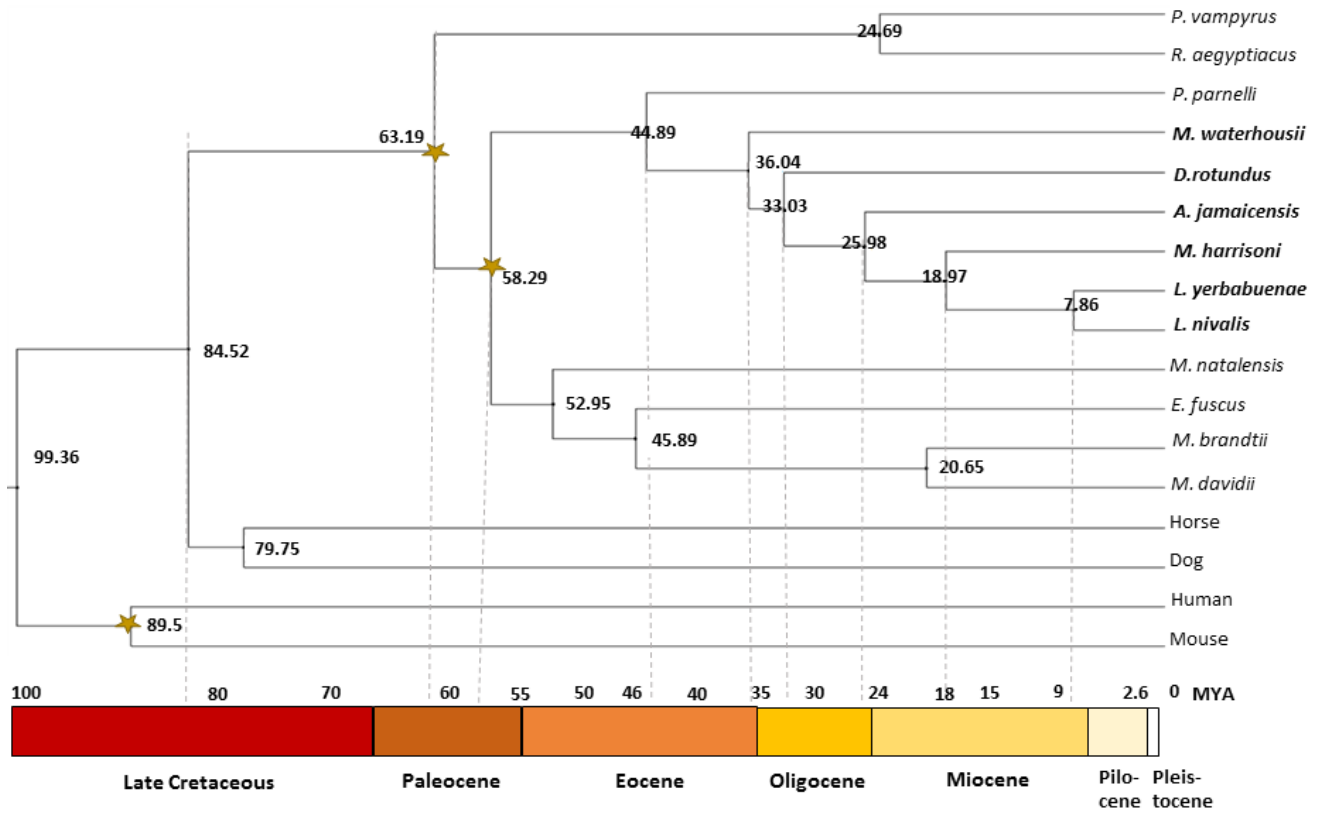


Figure S2. Phylogeny constructed based on 132 single copy genes and calibrated using three fossil records.

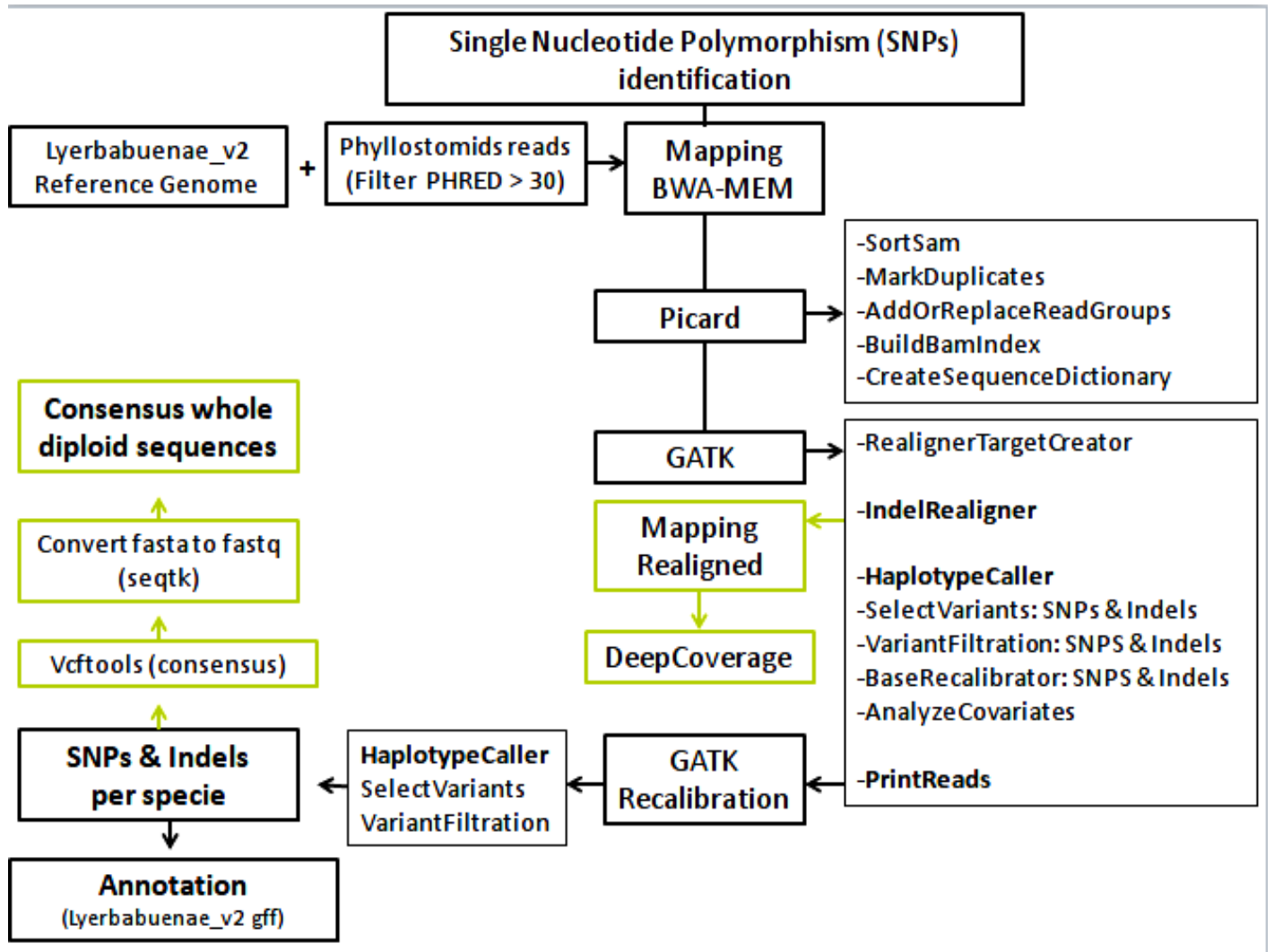


Figure S3. Pipeline for the reference-genome assembly construction.

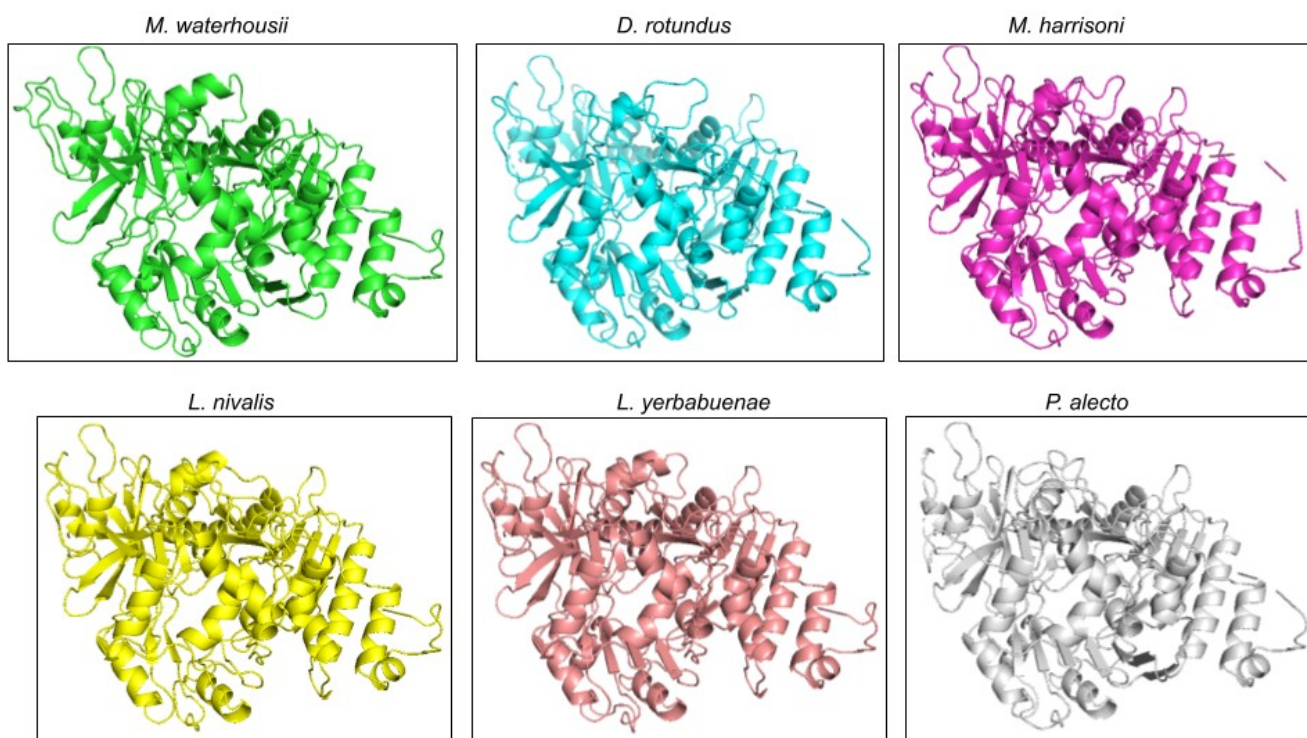


Figure S4. ACCS 3D protein structure for each bat species.

Capítulo II

How feeding habits shape the molecular evolution of sensory systems in bats

How feeding habits shape the molecular evolution of sensory systems in bats

Yocelyn T. Gutiérrez-Guerrero¹, Josué Barrera-Redondo¹, Michael Nachman², Jorge Ortega³,
Luis E. Eguiarte*¹

¹Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico.

²Museum of Vertebrate Zoology, Department of Integrative Biology, University of California Berkeley, CA 94720

³Departamento de Zoología, Laboratorio de Bioconservación y Manejo, Posgrado en Ciencias Químico-biológicas, Instituto Politécnico Nacional-ENCB, 11340, Mexico City, Mexico.

email: yoss_279@comunidad.unam.mx, josue_barrera@comunidad.unam.mx,
mnachman@berkeley.edu, jortegare@ipn.mx, fruns@unam.mx

Abstract

Bats have adapted to a wide range of niches, a process that involves the evolution of chemosensory gene repertoires that greatly varies among lineages and might reflect the species life history. We studied the evolution of olfactory and taste gene receptors in 17 bat species, using whole genome information available in databases. We evaluated the number of functional genes, signals of positive selection, estimated molecular substitution rates and their association with feeding behavior. We observed that most olfactory receptors were functional. Meanwhile *Tas1r2* and *Tas1r3* that are involved in sugar perception were functional in frugivore and nectarivore bat species, but non-functional in the vampire and in some insect-feeder lineages. Based on 45 highly conserved gene receptors, we identified that the olfactory *OR51E2* showed strong positive selection in frugivore bats. *Myotis* insect feeder species exhibited positive selection for the bitter receptor *Tas2r38*. Interestingly, some receptors exhibited an acceleration in the substitution rates in specific lineages, which might be the result of relaxed purifying selection. Our findings shed light on the importance of some olfactory and taste receptors with respect to the dietary niche, although most of the chemosensory receptors are shaped by kinds of selection and were not always correlated with diet.

Keywords

Bats, Diet, Evolution, Olfactory, Selection, Taste

Introduction

One of the most interesting adaptations in mammals is the evolution of sensory systems: vision, olfaction, taste, hearing, and balance (Hodos and Butler 1997). These systems are the result of different selection pressures, including strong natural selection, gained or lost gene receptors, pseudogenization and neofunctionalization events (Šestak et al. 2013; Hughes et al. 2018). In particular, olfactory and taste gene receptors have been extensively studied in some mammals (i. e. human, dog and mice) (Niimura et al. 2014; Antinucci and Risso 2017). Several studies show evidence of the evolution of hundreds of genes associated with olfactory and taste receptors (Smith and Bhatnagar 2019), which involve changes in the foraging, feeding behavior of the species, mating systems, physiology, and interactions (Gittleman 1991; Eisthen 1997; Nilsson 2009; Yu et al. 2010; Jiang et al. 2012).

Olfactory receptors (ORs) are the largest group of gene families in mammals, that encompasses from hundreds to ca. 4,230 OR functional and non-functional gene repertoires (Hadley et al. 2004; Šestak et al. 2013; Hughes et al. 2018; Tsuji et al. 2019). They are classified in chemosensory Class I and Class II (Zhang and Firestein 2002). This classification is based on structural features that may be related to the ligand specificity (Freitag et al., 1998) and the phylogenetic relationship of taxa, since Class I receptors are more similar among frogs and fishes, while Class II receptors are more similar among frogs and mammals (Zhang and Firestein 2002). A single ORs is recognized by multiple olfactory neurons and a single odorant can be detected by a combination of receptors (Nef 1993; Hadley et al. 2004).

Meanwhile, taste receptors (TAR) can recognize specific taste sensitivities (Adler et al. 2000; Hadley 2004). In mammals, TARs are classified into two families: *Tas1r* family plays a crucial role in distinguishing sweet taste (*Tas1r2* + *Tas1r3* heterodimer) and umami taste (*Tas1r1* + *Tas1r3* heterodimer) (Chéron et al. 2017; Kinnamon and Finger 2019); while *Tas2r* family contains several gene members important in detecting bitter taste (Go 2006; Lee et al. 2017).

The Chiroptera order is considered a successful group as it has exploited different ecological opportunities, colonized, and adapted to a wide range of niches (Speakman 2001; Yunpeng and Li 2015). The evolution of the chemosensory traits in bats has been shaped by a dynamic process of gene gains and losses within gene families (Jones et al. 2013; Hughes et al. 2018). Previous studies in ORs provided evidence of multiples changes in the number of functional, non-functional and neofunctionalized ORs genes, and their correlation with adaptations to an ecological niche and feeding behaviors, such as frugivory or herbivory (Hayden et al. 2014; Tsagkogeorga et al. 2017; Yohe et al. 2018). The *Tas1r1* gene is important in detecting umami flavors but was considered a pseudogene in different bat species (Zhao et al. 2012). Furthermore, the vampire lineage (*Desmodus rotundus*) showed pseudogenization and reduction of sweet taste receptor, in comparison with other bat species (Zhao et al. 2010; Hong and Zhao 2014).

Also, several studies have provided evidence of an association between gene sensory families and animal ecology (Liu et al. 2016; Tsagkogeorga et al. 2017; Hughes et al. 2018) but the relation between sensory receptors and feeding behavior in mammals is not completely clear. The order Chiroptera represents a good model to study the potential impact of diet on the evolution of sensory receptors, because of two main reasons: 1) All bat species evolved from an insect-feeder common ancestor (Dumont 2007); and 2) Bats of the Phyllostomidae family include highly diversified dietary niches and feeding specializations, such as blood-feeding, nectar-pollen, fruit feeding, and insect feeding bats, specializations that have implied changes in their foraging strategies (Schondube et al. 2001; Davies et al. 2020; Gutiérrez-Guerrero et al. 2020).

In this study, we investigated the evolutionary changes in the ORs and TARs receptors using genomic datasets available for 17 bat species from five families (Pteropodidae, Rhinolophidae, Hipposideridae, Phyllostomidae and Vespertilionidae). These species represent the different diets and feeding specializations in bats. We carried out exhaustive analyses to discriminate between functional and non-functional ORs and TARs. Moreover, we analyzed adaptive signatures, estimated substitution rates and divergence times for a group of orthologous single copy sensory genes to find evidences of common responses among lineages undergoing similar dietary changes.

Results

Genome data mining and gain/loss receptors

We constructed and filtered a database of olfactory and taste receptors based on the genomic information available in the database of 17 bat species and five additional mammal lineages (Fig. S1, See “Methods”). We found 361 different OR members across bats, the majority of the ORs belong to Class II and exhibited a diverse repertoire of members (See “Methods”). Additionally, we classified the functional and pseudogene olfactory receptors and assigned each bat species to a different dietary guild based on its main diet: insectivores (*Hipposideros armiger*, *Rhinolophus ferrumequinum*, *Macrotus waterhousii*, *Miniopterus natalensis*, *Eptesicus fuscus*, *Myotis davidii*, *Myotis brandtii* and *Myotis lucifugus*); omnivores (*Phyllostomus discolor*); frugivores (*Rousettus aegyptiacus*, *Pteropus alecto*, *Pteropus vampyrus* and *Artibeus jamaicensis*); sanguivores (*Desmodus rotundus*); and nectarivores (*Musonycteris harrisoni*, *Leptonycteris nivalis* and *Leptonycteris yerbabuena*) (See “Methods”). We found that specialized bats from nectar and sanguivory guilds showed fewer functional OR members (Fig. 1). Moreover, we identified a range of 15 to 35 putative OR pseudogenes, where the insectivore guild contained the highest number (Fig. 1). We also identified a high sequence similarity between some putative OR pseudogenes sequences and the non-coding genomic region (i. e. OR 52B4 and 3A1 in *M. davidii*; 8K3 in *R. aegyptiacus*; 4E1, 52N4 and 1P1 in *P. discolor*; and 52B4 and 52H1 in the vampire; Table S2).

Most of the OR genes were present in all the bat lineages, with few exceptions. For example, *OR14* gene was only present in three bat families (Rhinolophidae, Hipposideridae and Phyllostomidae). Furthermore, we detected that some OR genes were intact within specific lineages but were pseudogenized in the rest of the bat species (Fig. S2). Also, OR families 51 and 52 (Class I) contained less gene receptors, but more functional copies than Class II. Additionally, we identified that some OR families contained multiple gene copies due to duplication, but most of them were disrupted (i. e. *OR51H1* and *OR52P1*).

On the other hand, we found a range from 13 to 18 different taste receptor members, most of these TARs were shared among all bat species. Also, all receptors of the *Tas1r* family (*Tas1r1*, *Tas1r2* and *Tas1r3*) were pseudogenes in the vampire genome (Supplementary Table S2). In contrast, *Tas1r1* was only present as a pseudogene in *H. armiger*, *R. aegyptiacus*, *P. alecto* and *P. discolor* and was missing in the rest of the bat species. Interestingly, *M. brandtii*, *M. lucifugus* and *R. ferrimequinum* bat species lack the *Tas1r3* in their genomes, and was only present as a pseudogene in *M. davidii*, *D. rotundus* and *M. waterhousii* (Fig. 3). The *Tas2r* family contained a diverse repertory of gene receptors, but some members were missing for specific lineages. For example, *Tas2r9* was missing in all the Vespertilionid bats, while it was identified as a pseudogene in all Phyllostomid bats.

Our results show that gene families *OR51-A*, *OR51-I*, *OR51-G* and *OR6N* exhibited significant contractions in the number of copies for most of the bat species ($p < 0.05$). However, the Pteropodid species: *R. aegyptiacus* and *P. alecto* exhibited expansions in the *OR51L* and *OR6N* gene families. For TAR, we found changes in the gene family size of *Tas2r60* for *P. discolor* and *H. armiger*. The differences in the number of olfactory and taste receptors observed in each bat species may be related to evolutionary events of duplication and loss, but an error model correction suggests that the changes in the gene family sizes could be an artifact of the genome assembly and annotation.

Adaptive selection

After performing an exhaustive search, we found 44 OR (Class I and Class II) and 12 TAR orthologous single-copy genes shared in at least five bat species. We carried out a careful inspection of these OR and TAR genes to retain high-quality coding sequence alignments corresponding to one-to-one orthologous genes. We conducted a branch-site test for positive selection by estimating the ratio of non-synonymous to synonymous mutations ($w = dN / dS$) using Hyphy (Smith et al. 2015; Kosakovsky et al. 2020). We conducted the *Likelihood Ratio Tests* (LRT) and estimated a *False Discovery Rate* test (FDR) to find the OR and TAR genes under positive selection ($FDR < 0.01$ and $w > 1$).

We detected positive selection in 15 OR genes for specific bat lineages (Fig. 2). Within Class I, we found that *OR51* is the gene family containing more bat species presenting signals of

positive selection (representing the subfamilies: I1 and E2). Regarding Class II, we detected under positive selection the following gene families: *OR1L -3 / 8*; *OR2-A5 / Z1*; *OR4-C3 / D6*; *OR5D14*, *OR13C3 / OR1030* (Fig. 2). The positive selection signals in the ORs were different for each bat family and species. For instance *R. ferrumequinum* did not show evidence of positive selection in any of these receptors, while in *H. armiger* we found signals of positive selection only in *OR51I1*. The three Pteropodid species showed positive selection in members of the *OR52* and *OR51* receptors, but in none for the Class II receptors (Fig. 2; Table 1). Phyllostomid bats showed at least one OR gene under positive selection per species, where the *OR4D6* was exclusively under positive selection for the insect-feeder *M. waterhousii* and the vampire (Table 1).

The *OR52I1* and *OR2Z1* (from Class I and II, respectively) were under positive selection in most of the bat species, independently of their phylogenetic relationships. Interestingly, we observed parallelisms in positive selection for the *OR51E2* receptor (from Class I) in the bat species within the frugivore guild (Fig. 2).

We observed signatures of positive selection in the taste receptors: *Tas1r2*, *Tas2r3*, *Tas2r9* and *Tas2r38* within the Pteropodidae and Vespertilionidae families. The Pteropodid bat *P. vampyrus*, showed evidence of positive selection for the receptors *Tas1r2* and *Tas2r9*. The taste receptors *Tas2r* showed adaptive signals in the *Myotis* clade and *E. fuscus* (Fig. 2); whereas the *Tas2r38* was under positive selection in all *Myotis* species. For the rest of the analyzed bat species, we did not detect statistical significance at the 0.05 level.

Molecular substitution rates and divergence times

We estimated the divergence time and absolute rate of molecular substitution (substitutions per site per million years) of seven olfactory receptors and nine taste receptors using the Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in BEAST v2.6.1 (Bouckaert et al. 2019). We applied an uncorrelated relaxed log-normal clock model, a birth-death tree model, and a random starting tree, incorporating the fossil information from six fossil records (Table S1).

For the 10 olfactory gene receptors analyzed in the present study (Class I and Class II), we estimated a median rate from 0.0006 to 0.005 substitutions/site/million years (95% HPD 0.0002 – 0.006 substitutions/site/million years). The divergence time for Chiroptera was close to previously reported at ~60 Mya (Telling et al., 2005), with a 95% credibility interval at around 55 - 64 Mya. In accordance with our estimates, the New World bats split from the Old World bats around 50 Mya (95% HPD, 41.48 – 53 Mya). Despite the lower substitution rates in most of the receptors, we found rate accelerations in two OR genes for two bat species: the *OR51E2* (Class I) in the vampire branch, exhibiting a higher substitution rate of 0.0011 substitutions/site/million years (95% HPD 0.001 – 0.002 substitutions/site/million years) with a divergence-time at 13 Mya (*D. rotundus* – *M. waterhousii*; 95% credibility interval at 7 – 19.5 Mya) (Table 1; Fig. 3). Also, the *OR4D6* (Class II) for *P. vampyrus* showed a substitution rate of 0.005 substitutions/site/million

years (95% HPD 0.003 – 0.006 substitutions/site/million years), higher than in other bats (Table 1; Fig. 4a). Interestingly, the *OR52N2* receptor is represented by an apparently functional copy in *P. vampyrus* and *M. davidii*, but we observed that the copy from the *M. davidii* branch had a higher substitution rate of 0.0033 substitutions/site/million years (95% HPD 0.001 – 0.004 substitutions/site/million years) in comparison with the other copies and evolutionary branches of bats (Table 1; Fig. 3b).

In general, we observed higher molecular substitution rates in all the analyzed taste receptors than in the olfactory receptors, with an estimated range of 0.001 – 0.004 substitutions/site/million years (95% HPD 0.001 – 0.006 substitutions/site/million years), with a Bayesian posterior probability ≥ 0.7 (Table 1).

We observed that the *Tas1r3* receptors showed the highest substitution rates in the *M. harrisoni* branch, with a rate of 0.006 substitutions/site/million years (95 % HPD 0.003 – 0.009 Mya; posterior probability = 1) (Fig. 3c). Furthermore, we found for *Tas2r40* higher substitution rates (0.0045 substitutions/site/million years; 95% HPD 0.003-0.0055 Mya), in the insectivorous branches of *M. natalensis*, *M. davidii* and *P. discolor* (Fig. 3d).

Discussion

Our results provide an opportunity to understand how olfactory and taste gene receptors underwent recurrently gene loss and pseudogenization events, based on the genomic information from 18 bat genomes. Furthermore, we observed that chemosensory receptors are shaped in each lineage by positive selection, where certain olfactory and taste genes showed a common response in bat lineages with similar diet, independently of their phylogenetic relationships. Unexpectedly, we found that most gene receptors were not necessarily correlated with dietary preferences.

The remarkable diversity of olfactory receptors in mammals reflects a birth-death evolutionary model (Hughes et al. 2018), through functional duplication or pseudogenization across bat lineages (Niimura and Nei 2007; Liu et al. 2019). We identified more than 300 OR members in bats, from which a high number of copies were partial genes (Fig. 1, 2 and Fig. S1, S2). Our results support the previous findings of Tsagkogeorga *et al.*, (2017), who identified several contraction events in OR gene families across the Chiroptera order, but also an expansion in Old World Pteropodid bats (*R. aegyptiacus* and *P. alecto*).

Despite the diversity and large number of olfactory receptors for these nocturnal and flying mammals, we only identified 15 functional OR genes conserved in all species, with differences in selective signatures among lineages (Table 1, Fig. 2). Most of these conserved ORs exhibited low molecular substitution rates (a range from 0.0004 – 0.006 substitutions/site/million years), and reflecting the effect of the purifying selection. We hypothesize that these genes might participate in an enormous array of combinatorial receptors,

as heterodimers, which are important to detect specific odors, and are crucial for several biological and physiological mechanisms as for example in the cell-cell recognition, proliferation, and also apoptotic cycle (MaßBerg and Hatt 2018)

Previous studies in olfactory receptors suggested a link between sensory gene repertoires and organism diet (Niimura and Nei 2006; Hayden et al. 2014). Hayden *et al.*, (2014) and Hughes *et al.*, (2018) observed an association between the 1/3/7 and 2/13 OR gene families, in both, New World and Old World fruit bats (Phyllostomidae and Pteropodidae families). We detected novel signatures of positive selection in the *OR51E2* gene exclusively for New-World fruit bat *A. jamaicensis* (Phyllostomid) and for the Old-World fruit bats *P. alecto* and *R. aegyptiacus* (Pteropodid). We hypothesize that the *OR51E2* receptor might be relevant in the identification of compounds involved with the finding of fruits in their ecosystems. In contrast, we found a remarkable increase in the substitution rate of the vampire bat for *OR51E2*. In accordance with the *dN/dS* test, this increase is not a result of positive selection acting at the protein level, so we suggest that the *OR51E2* gene is acquiring and accumulating spontaneous mutations because of relaxed selective pressures (Fig. 3).

The evolution of the taste receptors family 1 is particularly important to understand how bat species can detect specific flavors (Yarmolinsky et al. 2009; Beauchamp and Jiang 2015). The observed pseudogenization of *Tas1r1* for all studied bat species supports previous findings regarding the inability of bats to detect the umami taste (Zhao et al. 2010; Zhao et al. 2011). Additionally, we provide evidence that some insectivorous bat lineages and the vampire lineage are not able to distinguish sweet flavors, because they lack a functional *Tas1r3* gene (Fig. 4), which makes part of the sweet taste receptor by the heterodimer *Tas1r2* + *Tas1r3* (Feng and Zhao 2013). Meanwhile, the rest of bat species with an omnivore, frugivore and nectarivore diet have a structurally conserved *Tas1r3* gene receptor. Murovets *et al.*, (2015) identified that the *Tas1R3* receptor is involved in the regulation of glucose homeostasis in pancreas and brain (Murovets et al. 2015). Our result suggests the importance of the *Tas1R3* receptor to detect not only the sweet taste of fruits and nectar, but also reflects the adaptation of some bat species to a diet with high sugar consumption.

The evolution of bitter taste receptors reflects the ecological adaptation and dietary shapes of each bat lineage (Zhao and Zhang, 2012; Liu et al., 2016). The vampire lineage exhibited multiple pseudogenization events in their bitter taste receptors (Fig. 2), as previous studies showed (Hong and Zhao 2014). The signals of positive selection identified in the *Tas2r38* gene for members of the Vespertilionid family (Fig. 2) were relevant to address the evolutionary role of this taste receptor in the insectivore diet (Li and Zhang 2013), as it is crucial to detect toxic components in food sources, such as leaves and insect exoskeletons (Antinucci and Risso 2017; Malik et al. 2019).

Additionally, we observed accelerations in the substitution rates of the *Tas2r40* and *Tas2r60* genes for Vespertilionid species. These results can be explained by a process of gene duplication and neofunctionalization in the *Tas2r* family, as reported previously for *Myotis* species (Jiao et al. 2018). In accordance with our results, we suggest that other insectivorous bat species should be analyzed to explore an adaptive response involved with this feeding behavior, and the evolutionary mechanism that led to the avoidance of ingesting potentially toxic compounds.

Finally, we suggest that during the Chiroptera evolution, the function of a small group of olfactory and taste receptors function became obsolete or were lost due to relaxed selection that has been influenced by new ecological challenges as well as by their dietary diversification.

Conclusions

This study has revealed the importance of certain sensory systems for specific diet preferences. Nonetheless, the hypothesis that diet shapes the evolution of sensory receptors is complex, because a great number of these genes are involved in other physiological pathways, and the evolution of many olfactory and taste receptors seems to be independent for each bat lineage. It is important to continue studying the evolutionary processes underlying these changes and perform physiological experiments to understand how dietary habits shape the evolution of sensory systems. Future studies should analyze the complete chemosensory landscape, and their responses to different sensorial stimulus, to better understand their link with ecological niches.

Methods

Genomic data

We used the protein and nucleotide information from 17 bat species genomes belonging to the families: Pteropodidae, Rhinolophidae, Hipposideridae, Vespertilionidae and Phyllostomidae. The genome assemblies and annotations were retrieved from the NCBI database: *Rousettus aegyptiacus* (Raegyp2.0), *Pteropus alecto* (ASM32557v1), *Hipposiderus armiger* (ASM189008v1), *Eptesicus fuscus* (EptFus1.0), *Minioterus natalensis* (Mnat.v1), *Myotis davidii* (ASM32734v1), *Myotis brandtii* (ASM41265v1), *Phyllostomus discolor* (mPhyDis1) and *Desmodus rotundus* (ASM294091v1) (O'Leary et al. 2016); GigaDataBase: *Macrotus waterhousii* (Mwaterhousii), *Artibeus jamaicensis* (Ajamaicensis), *Musonycteris harrisoni* (Mharrisoni), *Leptonycteris nivalis* (Lnivalis) and *Leptonycteris yerbabuenae* (Lyerbabuenae) (Gutiérrez-Guerrero et al. 2020); ENSEMBLE: *Pteropus vampyrus* (ENSPVA), *Rhinolophus ferrumequinum* (ENSRFE) and *Myotis lucifugus* (ENSMLU) (Yates et al. 2020). Moreover, we incorporated the genomic data of other mammals as outgroup for the analyses: horse (EquCab3), dog (CanFam3.1), human (GRCh38.p12), mouse (GRCm38.p6) and cow (Btau) and rat (Rnor_6.0).

Functional and putative pseudogene filtering

We built a nucleotide database with the olfactory receptors (OR) and taste receptors (TAR) annotated in the ENSEMBL database, as well pseudogenes for the bats: *P. vampyrus*, *R. ferrumequinum*, *M. lucifugus*, and the human. The first step to detect a potential pseudogene was to retrieve all the proteins with a function associated with olfactory and taste receptors for the 18 bat species. We used tblastn to identify for each bat species all the regions that share sequence similarity with the known OR and TAR proteins and the pseudogene database, using a cut off of $E\text{-value} \leq 1 \times 10^{-8}$, percent identity (*pid*) > 75% and percent coverage (*qcov*) > 90%. Those proteins that showed sequence homology based on our criteria were extracted, and we used tfasty from the FASTA program (Pearson, 1997) to obtain an accurate sequence alignment and calculate sequence similarities with frame shifts using an $E\text{-value} \leq 1 \times 10^{-5}$ and *pid* > 80%. Exclusively, for those bat species with available information of non-annotated genomic regions, we performed a tblastn using the putative pseudogene OR and TAR sequences previously detected. Finally, we evaluated the sequence completeness, length, open-reading frames and indels of all the putative ORs and TARs to differentiate between functional genes and pseudogenes identified with respect to the homology human sequence using Gblocks tool (Castresana 2000), bash and perl scripts and a visual inspection.

Positive selection test (*dN/dS*)

Highly conserved olfactory and taste orthologous single-copy genes

We incorporated the complete set of proteins from 35 Laurasiatheria species available in the ENSEMBL database (35 species) (Yates et al. 2020) to our dataset of functional bat ORs and TARs. We performed a multi-species genome comparison to obtain orthologous clusters genes with DIAMOND and Proteinortho (Buchfink et al. 2015; Lechner et al. 2011). We extracted the single copy gene clusters that were composed from 8 to maximum 22 sequences (including at least five sequences of bat species). We kept all the orthologous gene clusters that were annotated as taste or olfactory receptors.

Each cluster was aligned with MUSCLE (Edgar 2004; Katoh and Standley 2013), retaining only those sequences alignment where the total length is within 80 to 120% relative to the length of the human and mouse sequences, and all the partial sequences were removed. We used PAL2NAL to obtain the corresponding codon alignment for each orthologous cluster (Suyama et al. 2006). We reconstructed the phylogenetic tree for each cluster with RAxML (parameters -m GTRGAMMA -p 12345) (Stamatakis 2014).

We implemented a branch-site model to find evidence of positive selection in 3 or more branches on the phylogenetic tree, using the program aBSREL from Hyphy tool (Smith et al. 2015). We performed this analysis on the high quality gene clusters with conserved gene sequences found in more than five bat species and other mammals. We estimated the *Likelihood*

Ratio Test (LTR) and calculated the Chi-square to test for the assigned significance of positive selection. Additionally, we performed a *False Discovery Rate* test (FDR) (R Core Team) to correct for multiple testing. All branches with a *FDR* of 5% and a *dN/dS* (*w*) ≥ 5 , were considered to be under positive selection.

Molecular clock analyses

We estimated the divergence time and the molecular substitution absolute rate (substitutions per site per million years) of each olfactory and taste receptor using the Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in BEAST v2.6.1 (Bouckaert et al. 2019). We incorporated an uncorrelated relaxed log-normal clock model, a birth-death tree model, and a random starting tree for each cluster.

We used the substitution models and gamma shapes estimated with the Akaike Information Criterion implemented in jModeltest v2.1.10 (Darriba et al. 2012), as priors for this analysis (Table S1). Additionally, we included six calibration points based on fossils to set minimum dates of divergence for its corresponding node in each phylogeny, using the absolute age of the fossil on its uppermost boundary within the stratigraphic interval (Gradstein et al. 2012):

- 1) *Icaronycteris* is an extinct bat genus and the oldest known fossil, its records date at 52 Mya in the Eocene period (55.8 Mya). *Icaronycteris* exhibits clear features of microchiropteran bats, such as basicranial skeleton structures, suggesting that *Icaronycteris* was already capable of echolocate (Gunnell and Simmons 2005). Thus, we used this fossil as an upper limit to the divergence time for the Chiroptera Order.
- 2) *Hipposideros africanum* (41.3 Mya), is the oldest fossil record of the Hipposideridae family (Teeling et al. 2005; Ravel et al. 2016).
- 3) Vespertilionidae indet (41.3 Mya) (Eiting and Gunnell 2009; Gunnell et al. 2017)
- 4) Phyllostomidae indet (30 Mya) (Ubilla et al. 2019).
- 5) *Myotis* (20 Mya), is the oldest fossil record reported for the genus (Gunnell et al. 2017).
- 6) *Palynephyllum* (11.8 Mya), this fossil is the crown of the Lonchophylline subfamily (Phyllostomidae). Its original description is hypothesized to be a nectar-polen lineage (Santana et al. 2011; Dávalos et al. 2014).

We only used a subset of the fossil calibrations in each olfactory and taste receptor based on the availability of its assigned node within each gene tree, because there were some underlying conflicts between fossil calibrations and tree topologies obtained for each receptor (Table S1). The prior ages of each fossil calibration were set as a log-normal distribution with a mean of 1, a standard deviation of 1, and an offset equivalent to the age of the fossil in units of million years. For each receptor, we ran two independent MCMC analyses for ten million generations, sampling every 1,000 generations and specifying a burn-in of 10%. We assessed

the convergence between chains for all the estimated parameters using Tracer v1.6 (Rambaut et al. 2018). We merged the sampled trees from the two independent MCMC analyses of each gene receptor using LogCombiner (Suchard et al. 2018) to obtain maximum credibility trees, with mean node heights equivalent to divergence times and substitution rates, using TreeAnnotator (Bouckaert et al. 2019). All the trees were visualized in FigTree (Rambaut 2014).

Competing interests

The authors declare that they have no competing interests.

Funding

The study was supported by a grant from Fronteras de la Ciencia, “Genómica de la Diversidad de Vertebrados 1: *Leptonycteris* y la evolución de la nectarivoría en murciélagos y aves” (CONACyT, project no. 177) to L.E.E.

Acknowledgments

This manuscript is part of the doctoral project of the first author, who thanks the Posgrado en Ciencias Biomédicas (Universidad Nacional Autónoma de México, UNAM) and acknowledges the scholarship supported by Comisión Nacional de Ciencia y Tecnología (CONACyT, grant no. 580116; Beca Mixta grant no. 291250), and the grant Fronteras de la Ciencia (CONACyT, project no. 177). We gratefully acknowledge Dr. E. Aguirre-Planter and Dr. C. Martínez del Río for their constructive comments on the manuscript and help in lab work. We thank Drs. L. Espinosa Asuar and V. Souza. Also, S. Barrientos and all the team members of the Lab. Evolución Molecular y Experimental (UNAM) for their help in lab work. The authors acknowledge the Instituto de Ecología (UNAM), Comisión Nacional para el Conocimiento de la Biodiversidad (CONABIO) and Instituto de Ecología A.C. (INECOL, A.C) for computing resources, and MSc. E. Villafán, Ing. R. García and Mat. E. Campos for their computational assistance.

References

- Adler, E., M. A. Hoon, K. L. Mueller, J. Chandrashekar, N. J. Ryba, and C. S. Zuker. 2000. A novel family of mammalian taste receptors. *Cell* 100:693–702.
- Antinucci, M., and D. Risso. 2017. A Matter of Taste: Lineage-Specific Loss of Function of Taste Receptor Genes in Vertebrates. *Front Mol Biosci* 4:81.
- Beauchamp, G. K., and P. Jiang. 2015. Comparative biology of taste: Insights into mechanism and function. *Flavour* 4.

- Bouckaert, R., T. G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N. D. Maio, M. Matschiner, F. K. Mendes, N. F. Müller, H. A. Ogilvie, L. du Plessis, A. Poppinga, A. Rambaut, D. Rasmussen, I. Siveroni, M. A. Suchard, C.-H. Wu, D. Xie, C. Zhang, T. Stadler, and A. J. Drummond. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 15:e1006650.
- Buchfink, B., C. Xie, and D. H. Huson. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenomics analysis. *Mol Biol Evol* 17:540-552.
- Chéron, J. B., J. Golebiowski, S. Antonczak, and S. Fiorucci. 2017. The anatomy of mammalian sweet taste receptors. *Proteins* 85:332–341.
- Consortium, T. U. P. 2018. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models new heuristics and parallel computing. *Nat Methods* 9:772–772.
- Dávalos, L. M., Velazco, P., Warsi, O. M., Smiths, P. D., Simmons, N. B. 2014. Integrating incomplete fossils by isolating conflicting signal in saturated and non-independent morphological characters. *63:582-600*.
- Davies, K. T. J., L. R. Yohe, J. Almonte, M. K. R. Sánchez, E. M. Rengifo, E. R. Dumont, K. E. Sears, L. M. Dávalos, and S. J. Rossiter. 2020. Foraging shifts and visual preadaptation in ecologically diverse bats. *Mol Ecol* 29:1839–1859.
- Dumont, E. R. 2007. Feeding mechanisms in bats: variation within the constraints of flight. *Integr Comp Biol* 47:137–146.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–7.
- Eisthen, H. L. 1997. Evolution of Vertebrate Olfactory Systems. *Brain Behav and Evolut* 50:222–233.
- Eiting, T. P., and G. F. Gunnell. 2009. Global Completeness of the Bat Fossil Record. *J Mamm Evol* 16:151–173.

- Feng, P., and H. B. Zhao. 2013. Complex evolutionary history of the vertebrate sweet/umami taste receptor genes. *Chinese Science Bulletin* 58:2198–2204.
- Gittleman, J. L. 1991. Carnivore olfactory bulb size: allometry phylogeny and ecology. *J Zool* 225:253–272.
- Go, Y. 2006. Lineage-Specific Expansions and Contractions of the Bitter Taste Receptor Gene Repertoire in Vertebrates. *Mol Biol Evol* 23:964–972.
- Gradstein, F. M., J. G. Ogg, and F. J. Hilgen. 2012. On The Geologic Time Scale. *Newsl Stratigr* 45:171–188.
- Gunnell, G. F., R. Smith, and T. Smith. 2017. 33 million years old *Myotis* (Chiroptera Vespertilionidae) and the rapid global radiation of modern bats. *PLoS ONE* 12:e0172621.
- Gunnell, G. F., and N. B. Simmons. 2005. Fossil Evidence and the Origin of Bats. *J Mamm Evol* 12:209–246.
- Gutiérrez-Guerrero, Y. T., E. Ibarra-Laclette, C. Martínez del Río, J. Barrera-Redondo, E. A. Rebollar, J. Ortega, L. León-Paniagua, A. Urrutia, E. Aguirre-Planter, and L. E. Eguiarte. 2020. Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. *GigaScience* 9.
- Hadley, K., R. R. Orlandi, and K. J. Fong. 2004. Basic anatomy and physiology of olfaction and taste. *Otolaryn Clin N AM* 37:1115–1126.
- Hayden, S., M. Bekaert, A. Goodbla, W. J. Murphy, L. M. Dávalos, and E. C. Teeling. 2014. A Cluster of Olfactory Receptor Genes Linked to Frugivory in Bats. *Mol Biol Evol* 31:917–927.
- Hodos, W., and A. B. Butler. 1997. Evolution of Sensory Pathways in Vertebrates. *Brain Behav Evolut* 50:189–197.
- Hong, W., and H. Zhao. 2014. Vampire bats exhibit evolutionary reduction of bitter taste receptor genes common to other bats. *Proc Biol Sci* 281:20141079.
- Hughes, G. M., E. S. M. Boston, J. A. Finarelli, W. J. Murphy, D. G. Higgins, and E. C. Teeling. 2018. The Birth and Death of Olfactory Receptor Gene Families in Mammalian Niche Adaptation. *Mol Biol Evol* 35:1390–1406.
- Jiang, P., J. Josue, X. Li, D. Glaser, W. Li, J. G. Brand, R. F. Margolskee, D. R. Reed, and G. K. Beauchamp. 2012. Major taste loss in carnivorous mammals. *PNAS* 109:4956–4961.
- Jiao, H., Y. Wang, L. Zhang, P. Jiang, and H. Zhao. 2018. Lineage-specific duplication and adaptive evolution of bitter taste receptor genes in bats. *Mol Ecol* 27:4475–4488.

- Jones, G., E. C. Teeling, and S. J. Rossiter. 2013. From the ultrasonic to the infrared: molecular evolution and the sensory biology of bats. *Front Phys* 4.
- Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–80.
- Kinnamon, S. C., and T. E. Finger. 2019. Recent advances in taste transduction and signaling. *F1000Res* 8.
- Kosakovsky, P. S. L., A. F. Y. Poon, R. Velazquez, S. Weaver, N. L. Hepler, B. Murrell, S. D. Shank, B. R. Magalis, D. Bouvier, A. Nekrutenko, S. Wisotsky, S. J. Spielman, S. D. W. Frost, and S. V. Muse. 2020. HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol Biol Evol* 37:295–299.
- Lechner, M., S. Findeiss, L. Steiner, M. Marz, P. F. Stadler, and S. J. Prohaska. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124.
- Lee, H., L. J. Macpherson, C. A. Parada, C. S. Zuker, and N. J. P. Ryba. 2017. Rewiring the taste system. *Nature* 548:330–333.
- Li, D., and J. Zhang. 2013. Diet Shapes the Evolution of the Vertebrate Bitter Taste Receptor Gene Repertoire. *Mol Biol Evol* 31:303–309.
- Liu, A., F. He, L. Shen, R. Liu, Z. Wang, and J. Zhou. 2019. Convergent degeneration of olfactory receptor gene repertoires in marine mammals. *BMC Genomics*.
- Liu, Z., G. Liu, F. Hailer, P. Orozco-terWengel, X. Tan, J. Tian, Z. Yan, B. Zhang, and M. Li. 2016. Dietary specialization drives multiple independent losses and gains in the bitter taste gene repertoire of Laurasiatherian Mammals. *Front Zool*.
- Malik, B., N. Elkaddi, J. Turkistani, A. I. Spielman, and M. H. Ozdener. 2019. Mammalian Taste Cells Express Functional Olfactory Receptors. *Chemical Senses* 44:289–301.
- Maßberg, D., Hatt, H. 2018. Human olfactory receptors: novel cellular functions outside of the nose. *Physiol Rev*. 98:1739-1763.
- Murovets, V. O., A. A. Bachmanov, and V. A. Zolotarev. 2015. Impaired Glucose Metabolism in Mice Lacking the *Tas1r3* Taste Receptor Gene. *PLoS ONE* 10:e0130997.
- Nef, P. 1993. Early events in olfaction: diversity and spatial patterns of odorant receptors. *Receptors Channels* 1:259–66.

- Niimura, Y., A. Matsui, and K. Touhara. 2014. Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome Res* 24:1485–1496.
- Niimura, Y., and M. Nei. 2007. Extensive Gains and Losses of Olfactory Receptor Genes in Mammalian Evolution. *PLoS ONE* 2:e708.
- Niimura, Y., and M. Nei. 2006. Evolutionary dynamics of olfactory and other chemosensory receptor genes in vertebrates. *J Hum Genet* 51:505–517.
- Nilsson, D.-E. 2009. The evolution of eyes and visually guided behavior. *Philos T R Soc B* 364:2833–2847.
- O’Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O’Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, and K. D. Pruitt. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44:D733–45.
- R Development Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rambaut, A. 2014. FigTree v1.4.2, A Graphical Viewer of Phylogenetic Trees. Available from <http://tree.bio.ed.ac.uk/software/figtree/>
- Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* 67:901–904.
- Ravel, A., M. Adaci, M. Bensalah, A.-L. Charruault, E. M. Essid, H. K. AMmar, W. Marzougui, M. Mahboubi, F. Mebrouk, G. Merzeraud, M. Vianey-Liaud, R. Tabuce, and L. Marivaux. 2016. Origine et radiation initiale des chauves-souris modernes: nouvelles découvertes dans l’Éocène d’Afrique du Nord. *Geodiversitas* 38:355–434. Museum National d’Histoire Naturelle Paris, France.
- Santana, S. E., S. Strait, and E. R. Dumont. 2011. The better to eat you with: functional correlates of tooth structure in bats. *Funct Ecol* 25:839–847.

- Schondube, J. E., L. G. Herrera-M, and C. M. del Rio. 2001. Diet and the evolution of digestion and renal function in phyllostomid bats. *Zoology* 104:59–73.
- Šestak, M. S., V. Božičević, R. Bakarić, V. Dunjko, and T. Domazet-Lošo. 2013. Phylostratigraphic profiles reveal a deep evolutionary history of the vertebrate head sensory systems. *Front Zool* 10:18.
- Smith, M. D., J. O. Wertheim, S. Weaver, B. Murrell, K. Scheffler, and P. S. L. Kosakovsky. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* 32:1342–53.
- Smith, T. D., and K. P. Bhatnagar. 2019. Anatomy of the olfactory system. Pp. 17–28 *in* Handbook of Clinical Neurology.
- Speakman, J. R. 2001. The evolution of flight and echolocation in bats: another leap in the dark. *Mammal Rev* 31:111–130.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–3.
- Suchard, M. A., P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, and A. Rambaut. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4:vey016.
- Suyama, M., D. Torrents, and P. Bork. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 34:W609–12.
- Teeling, E. C., M. S. Springer, O. Madsen, P. Bates, S. J. O'Brien, and W. J. Murphy. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* 307:580–4.
- Tsagkogeorga, G., S. Müller, C. Dessimoz, and S. J. Rossiter. 2017. Comparative genomics reveals contraction in olfactory receptor genes in bats. *Sci Rep* 7:259.
- Tsuji, T., C. Tsuji, M. Lozic, M. Ludwig, and G. Leng. 2019. Coding of odors in the anterior olfactory nucleus. *Physiol Rep* 7:e14284.
- Ubilla, M., P. Gaudioso, and D. Perea. 2019. First fossil record of a bat (Chiroptera Phyllostomidae) from Uruguay (Plio-Pleistocene, South America): a giant desmodontine. *Historical Biology* 1-9.
- Yarmolinsky, D. A., C. S. Zuker, and N. J. P. Ryba. 2009. Common Sense about Taste: From Mammals to Insects. *Cell* 139:234–244.

- Yates, A. D., P. Achuthan, W. Akanni, J. Allen, J. Allen, J. Alvarez-Jarreta, M. R. Amode, I. M. Armean, A. G. Azov, R. Bennett, J. Bhai, K. Billis, S. Boddu, J. C. Marugán, C. Cummins, C. Davidson, K. Dodiya, R. Fatima, A. Gall, C. G. Giron, L. Gil, T. Grego, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, M. Kay, I. Lavidas, T. Le, D. Lemos, J. G. Martinez, T. Maurel, M. McDowall, A. McMahon, S. Mohanan, B. Moore, M. Nuhn, D. N. Oheh, A. Parker, A. Parton, M. Patricio, M. P. Sakthivel, S. A. I. Abdul, B. M. Schmitt, H. Schuilenburg, D. Sheppard, M. Sycheva, M. Szuba, K. Taylor, A. Thormann, G. Threadgold, A. Vullo, B. Walts, A. Winterbottom, A. Zadissa, M. Chakiachvili, B. Flint, A. Frankish, S. E. Hunt, G. Ilesley, M. Kostadima, N. Langridge, J. E. Loveland, F. J. Martin, J. Morales, J. M. Mudge, M. Muffato, E. Perry, M. Ruffier, S. J. Trevanion, F. Cunningham, K. L. Howe, D. R. Zerbino, and P. Flicek. 2020. Ensembl 2020. *Nucleic Acids Res* 48:D682–D688.
- Yohe, L. R., S. Hoffmann, and A. Curtis. 2018. Vomeronasal and Olfactory Structures in Bats Revealed by DiceCT Clarify Genetic Evidence of Function. *Front Neuroanat* 12.
- Yu, L., W. Jin, J.-xin Wang, X. Zhang, M.-meng Chen, Z.-hai Zhu, H. Lee, M. Lee, and Y.-ping Zhang. 2010. Characterization of TRPC2 an Essential Genetic Component of VNS Chemoreception, Provides Insights into the Evolution of Pheromonal Olfaction in Secondary-Adapted Marine Mammals. *Mol Biol Evol* 27:1467–1477.
- Yunpeng, L., and Y. Li. 2015. Advances on molecular mechanism of the adaptive evolution of Chiroptera (bats). *Yi Chuan* 37:25–33.
- Zhang, X., and S. Firestein. 2002. The olfactory receptor gene superfamily of the mouse. *Na Neurosci* 5:124–133.
- Zhao, H., D. Xu, S. Zhang, and J. Zhang. 2012. Genomic and genetic evidence for the loss of umami taste in bats. *Genome Biol Evol* 4:73–9.
- Zhao, H., D. Xu, S. Zhang, and J. Zhang. 2011. Genomic and Genetic Evidence for the Loss of Umami Taste in Bats. *Genome Biol Evol* 4:73–79.
- Zhao, H., Y. Zhou, C. M. Pinto, P. Charles-Dominique, J. Galindo-Gonzalez, S. Zhang, and J. Zhang. 2010. Evolution of the Sweet Taste Receptor Gene *Tas1r2* in Bats. *Mol Biol Evol* 27:2642–2650.
- Zhao, H., Y. Zhou, C. M. Pinto, P. Charles-Dominique, J. Galindo-González, S. Zhang, and J. Zhang. 2010. Evolution of the sweet taste receptor gene *Tas1r2* in bats. *Mol Biol Evol* 27:2642–50.

Table and Figure legends

Table 1. Ratio dN/dS and rate of molecular substitution for olfactory and taste gene receptors.

Figure 1. Functional genes and putative pseudogenes of olfactory receptor (OR) members identified in five dietary guilds.

Figure 2. Highly conserved olfactory and taste receptors under positive selection for each bat species.

Figure 3. Single-copy gene receptors: olfactory (a-b) and taste (c-d) that exhibited accelerations in their molecular substitution rates for specific bat lineages (red branch). a) *OR4D6*; b) *OR52N2*; c) *Tas1r3*; and d) *Tas2r40*.

Figure 4. Left. Phylogenetic tree of taste receptor 1 member 3 (*Tas1r3*) gene in mammals. Grey circles highlight the pseudogenization events. **Right.** *Tas1r3* sequence alignment for the Chiroptera species.

Supplementary Material

Supplementary Table 1. Priors used to calculate divergence times and absolute substitution rates for each taste and olfactory receptor.

Supplementary Table 2. Olfactory receptors located in non-coding regions.

Figure S1. Number of Olfactory receptors (Class I and Class II) and Taste receptors (Type I and Type II) identified in the genomic dataset available for all bat species.

Figure S2. Functional gene copies of 16 olfactory receptor families for each bat genome.

Tabla 1. Ratio dN/dS and rate of molecular substitution for olfactory and taste gene receptors.

		dN / dS and substitution rate (substitution/site/million years)													
		Olfactory Receptors					Taste Receptors								
		Class I			Class II		Type I			Type 2					
		<i>51E2</i>	<i>51I1</i>	<i>52N2</i>	<i>1L3</i>	<i>4D6</i>	<i>1R2</i>	<i>1R3</i>	<i>2R3</i>	<i>2R4</i>	<i>2R5</i>	<i>2R38</i>	<i>2R39</i>	<i>2R40</i>	<i>2R60</i>
Vespertilionidae	<i>M. lucifugus</i>	-- 0.0005	-- 0.0005	-- --	-- 0.0007	-- 0.001	-- --	-- --	-- 0.002	-- 0.003 --	-- 0.001	2.7 0.002	-- 0.003	-- 0.001	-- --
	<i>M. brandtii</i>	1 0.0005	-- 0.0004	10 0.0008	-- 0.0007	-- 0.001	-- --	-- --	-- --	-- --	-- --	10 0.002	-- 0.002	-- 0.001	-- 0.002
	<i>M. davidii</i>	-- 0.0005	10 0.0004	-- 0.0009	-- 0.001 --	-- 0.001	-- 0.001	-- --	5.9 0.002	-- 0.003	-- 0.002	29 0.004	-- 0.004	-- 0.004	-- 0.003
	<i>E. fuscus</i>	1 0.0005	-- 0.0005	-- --	-- 0.0006	-- --	-- 0.001	-- --	-- 0.002	-- --	-- 0.001	10 0.002	-- 0.001	-- 0.001	-- 0.001
	<i>M. natalensis</i>	-- 0.0004	-- 0.0004	1.1 0.0009	-- 0.0007	-- 0.002	-- 0.001	-- 0.001	-- 0.003	-- 0.004	-- 0.003	-- 0.004	-- 0.002	-- 0.004	-- 0.004
	<i>L. yerbabuena</i>	1 0.0006	-- --	-- 0.001 --	-- --	-- 0.001	-- 0.001	-- 0.002	-- --	-- 0.003 --	-- 0.003	-- 0.004	-- 0.006	-- 0.002	-- 0.002

Phyllostomidae	<i>L. nivalis</i>	--	--	--	10	--	--	--	--	--	--	--	--	--					
		0.0004	0.0005	--	0.0008	0.001	0.001	0.002	--	--	0.002	0.002	0.001	0.001	0.002				
	<i>M. harrisoni</i>	--	--	--	--	0.001	--	--	--	--	--	--	--	--					
		0.0006	0.0004	0.0008	--	0.002	0.002	0.006	--	--	0.004	0.005	0.003	0.002	0.002				
	<i>A. jamaicensis</i>	10	2.5	--	--	--	--	--	--	--	--	--	--	--	--				
	0.0005	0.0005	0.0005	0.0007	0.002	0.002	0.003	--	--	0.002	0.003	0.004	0.002	0.003					
	<i>D. rotundus</i>	--	0.001	--	10	--	0.001	10	--	--	--	--	--	--					
			0.0004	0.001	--	0.001	--	0.001	--	0.002	--	0.005	--0.003	0.003	0.003	0.004			
	<i>P. discolor</i>	--	10	--	0.001	--	--	--	--	0.003	--	--	--	--	--				
		0.0006	0.0005	--	0.0006	--	--	--	--	0.001	0.002	--	--0.003	0.004	0.002	0.02	0.004	0.004	--
Hipposideridae	<i>M. waterhousii</i>	--0.0006	--	--	9.8	10	--	--	--	--	--	--	--	--	--				
			0.0004	0.001	0.0006	0.001	0.001	--	--	--	--	0.002	0.003	0.001	0.002	--			
		--	10	2.1	--	--	--	--	--	--	--	--	--	--	--	--			
	<i>H. armiger</i>	0.0004	0.0004	0.0009	--	0.001	0.001	--	0.002	0.002	0.002	0.002	--	--	0.001				
	<i>P. vampyrus</i>	--0.0004	--	--	--	--	--	--	--	--	--	--	--	--	--				
			0.0004	0.0006	0.0009	0.005	0.001	0.002	--	0.002	0.003	0.003	--	0.001	--				

Pteropodidae	<i>P. alecto</i>	10	10	--	--	--	2.2	--	--	--	--	--	--	--
		0.0005	0.0005	0.0009	0.0006	--	0.001	0.002	0.001	0.003	0.002	0.002	0.002	0.002
	<i>R. aegyptiacus</i>	10	--	--	--	0.001	--	--	--	--	--	--	--	--
		0.0006	0.0004	0.0003	--	--	0.002	0.003	0.003	0.003	0.003	0.003	0.004	0.001

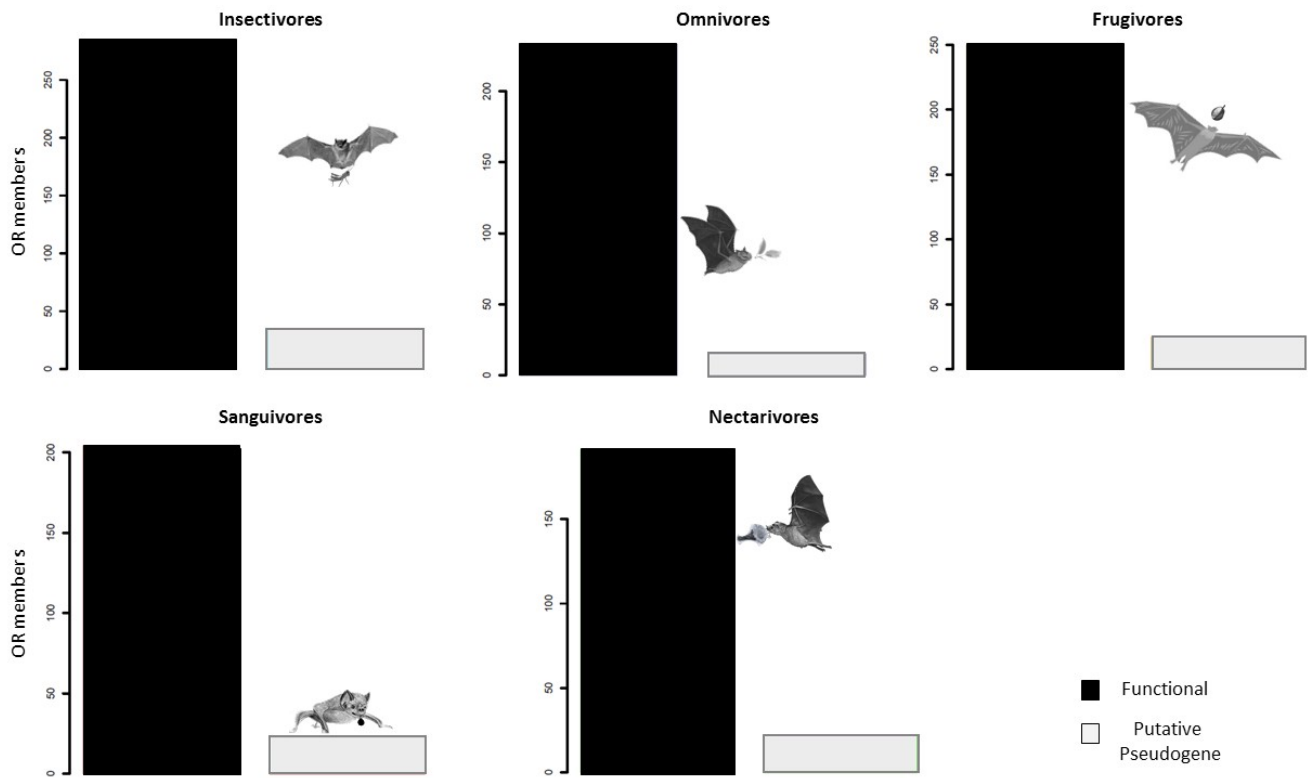


Figure 1. Functional genes and putative pseudogenes of olfactory receptor (OR) members identified in five dietary guilds.

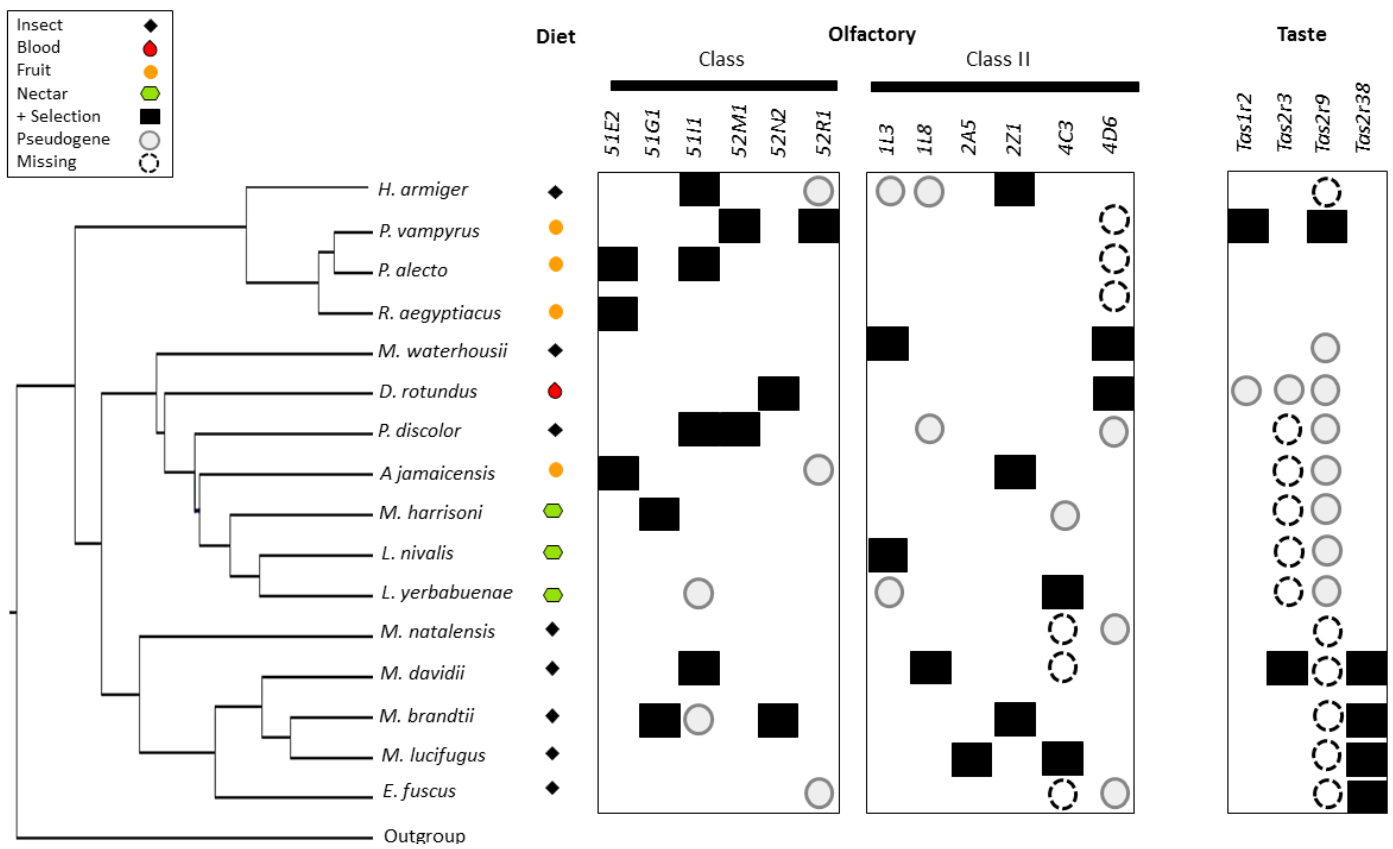


Figure 2. Highly conserved olfactory and taste receptors under positive selection for each bat species.

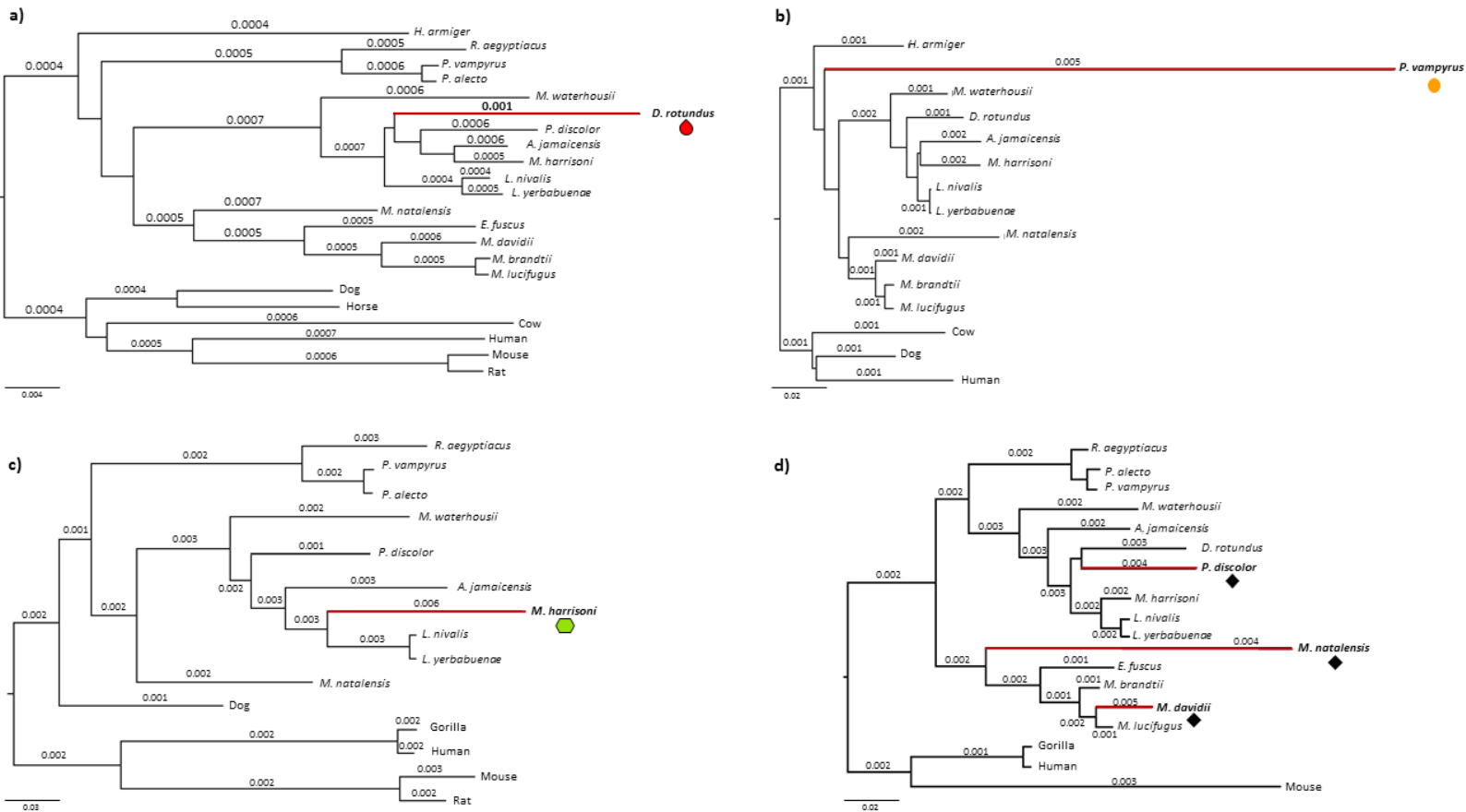


Figure 3. Single-copy gene receptors: olfactory (a-b) and taste (c-d) that exhibited accelerations in their molecular substitution rates for specific bat lineages (red branch). a) *OR4D6*; b) *OR52N2*; c) *Tas1r3*; and d) *Tas2r40*.

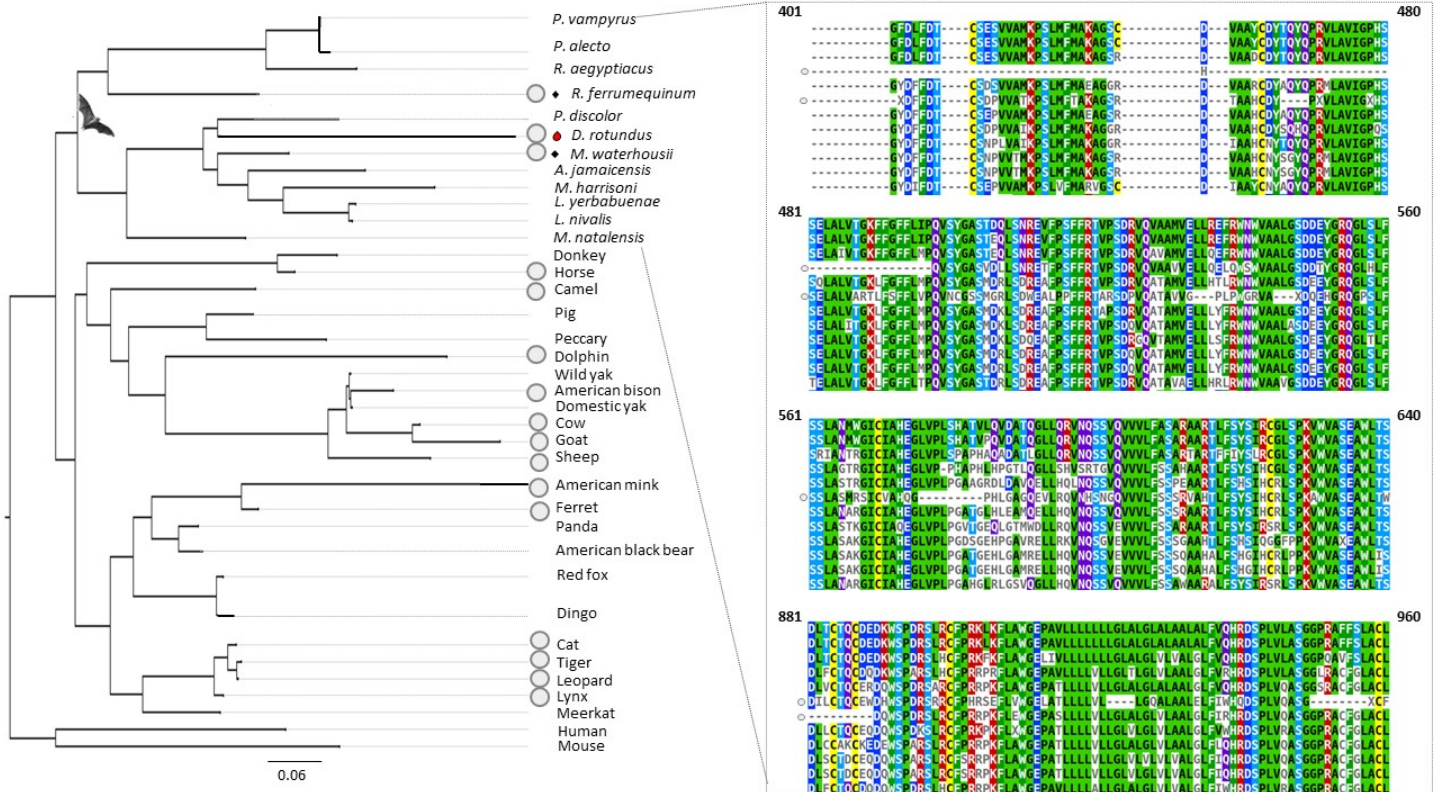


Figure 4. Left. Phylogenetic tree of taste receptor 1 member 3 (*Tas1r3*) gene in mammals. Grey circles highlight the pseudogenization events. **Right.** *Tas1r3* sequence alignment for the Chiroptera species.

Additional material

Table Supplementary 1. Priors used to calculate divergence times and absolute substitution rates for each taste and olfactory receptor.

Sensory receptor	Substitution model	Gamma shape	Invariant sites	Fossil calibrations (Mya)
<i>Tas1r2</i>	GTR+I+G	0.614	0.187	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
<i>Tas1r3</i>	GTR+G	0.535	NA	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30) <i>Palynephyllum</i> (11.8)
<i>Tas2r3</i>	GTR+G	0.805	NA	<i>Icaronycteris</i> (55.8) <i>Myotis</i> (20)
<i>Tas2r4</i>	GTR+I+G	3.193	0.267	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
<i>Tas2r5</i>	GTR+G	0.619	NA	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
<i>Tas2r38</i>	GTR+G	0.936	NA	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30) <i>Palynephyllum</i> (11.8)
<i>Tas2r39</i>	GTR+G	1.172	NA	<i>Icaronycteris</i> (55.8) <i>Phyllostomidae</i> indet. (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
<i>Tas2r40</i>	GTR+G	0.557	NA	<i>Icaronycteris</i> (55.8) <i>Vespertilionidae</i> indet. (41.3) <i>Phyllostomidae</i> indet. (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
<i>Tas2r60</i>	GTR+G	2.446	NA	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae</i> indet. (30)
<i>OR1L3</i>	GTR+G	0.471	NA	<i>Icaronycteris</i> (55.8) <i>Vespertilionidae</i> indet. (41.3) <i>Pterop</i> (3.6) <i>Myotis</i> (20)
<i>OR4D6</i>	GTR+G	0.740	NA	<i>Icaronycteris</i> (55.8)

				<i>Phyllostomidae indet.</i> (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
OR51E2	GTR+I+G	0.868	0.605	<i>Icaronycteris</i> (55.8) <i>Phyllostomidae indet.</i> (30) <i>Myotis</i> (20) <i>Palynephyllum</i> (11.8)
OR51I1	GTR+I+G	0.824	0.521	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae indet.</i> (30)
OR52M1	GTR+G	0.481	NA	<i>Icaronycteris</i> (55.8) <i>Hipposideros africanum</i> (41.3) <i>Phyllostomidae indet.</i> (30) <i>Myotis</i> (20)
OR52N2	GTR+I+G	0.897	0.347	<i>Icaronycteris</i> (55.8) <i>Phyllostomidae indet.</i> (30) <i>Myotis</i> (20)

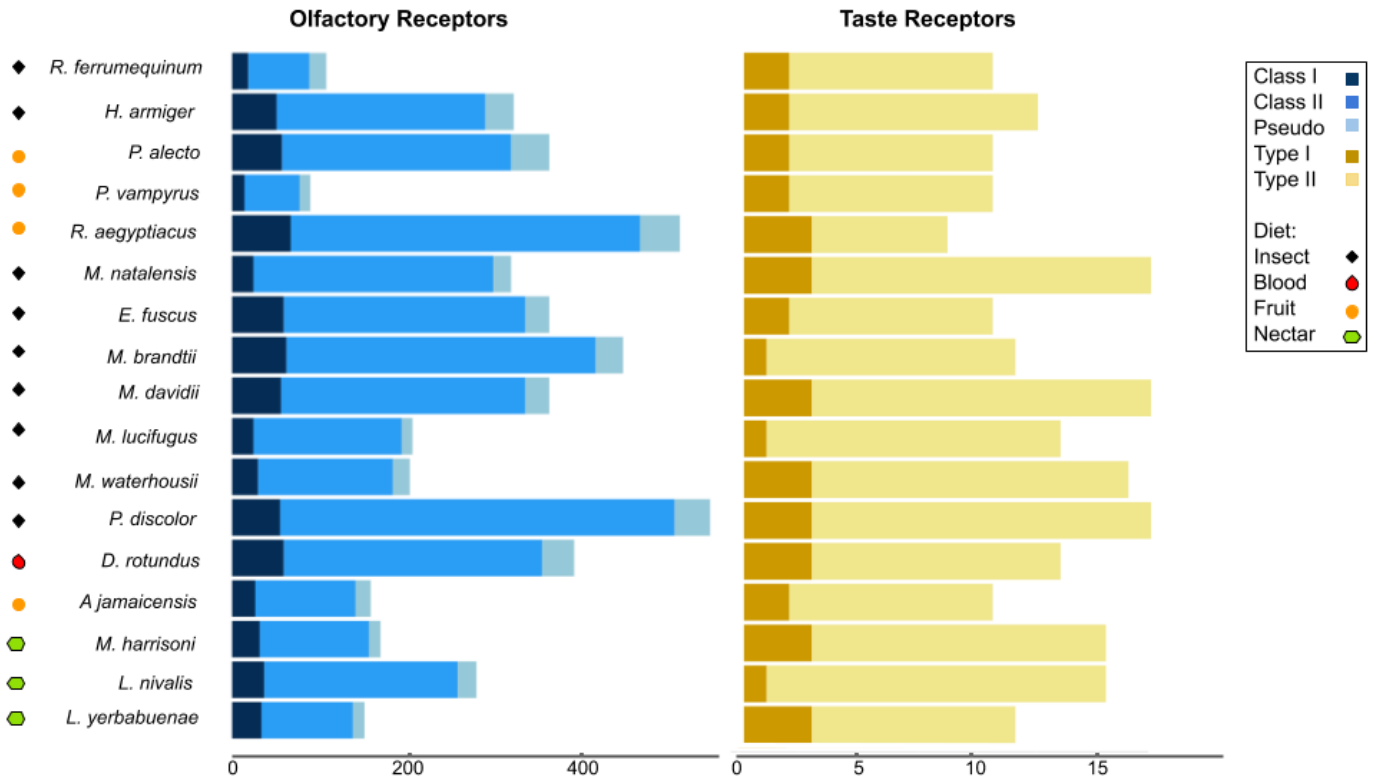
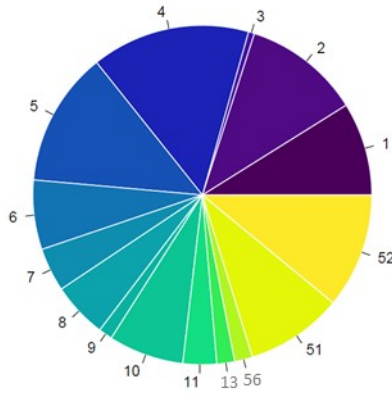


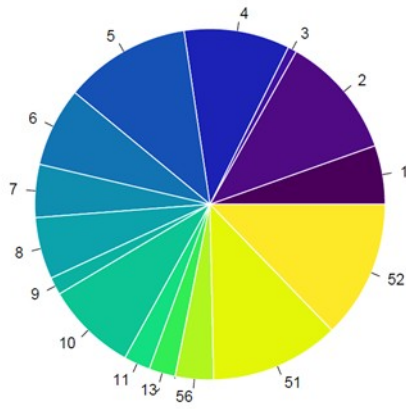
Figure S1. Number of Olfactory receptors (Class I and Class II) and Taste receptors (Type I and Type II) identified in the genomic dataset available for all bat species.

Pteropodidae

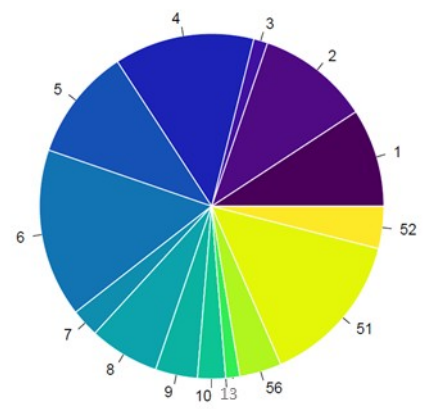
a) *R. aegyptiacus*



b) *P. alecto*

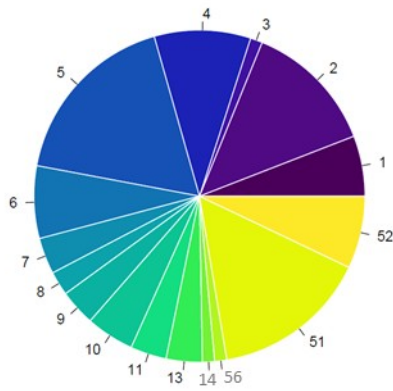


c) *P. vampyrus*



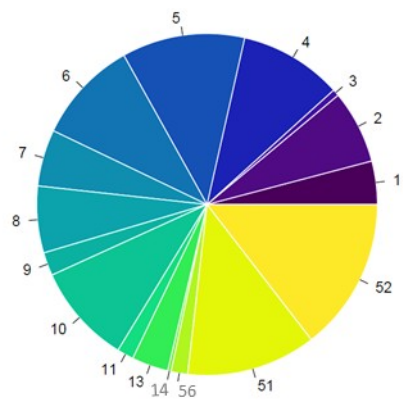
Rhinolophidae

d) *R. ferrumequinum*

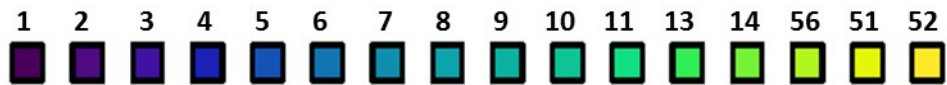


Hipposideridae

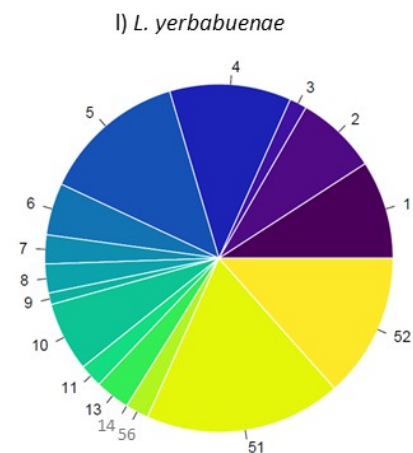
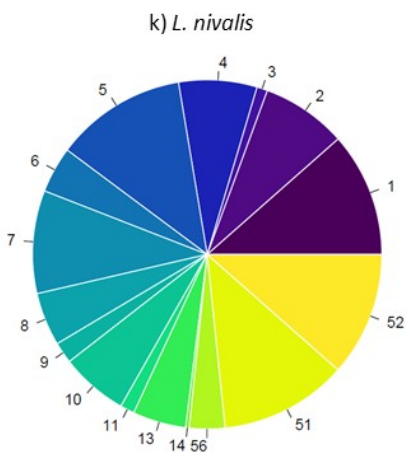
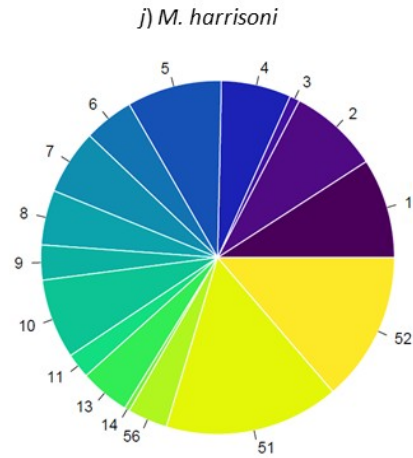
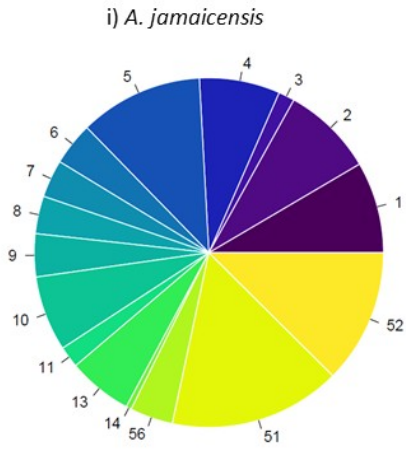
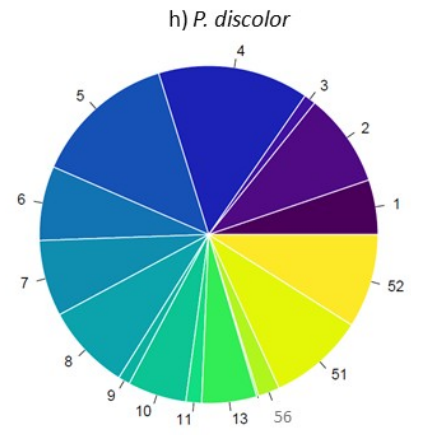
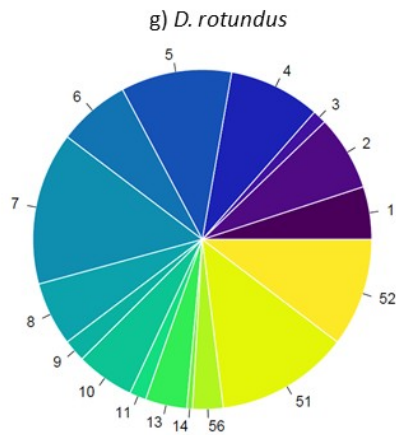
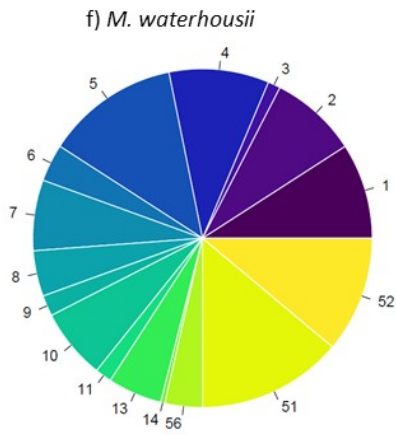
e) *H. armiger*



Olfactory Receptor



Phyllostomidae



Vespertilionidae

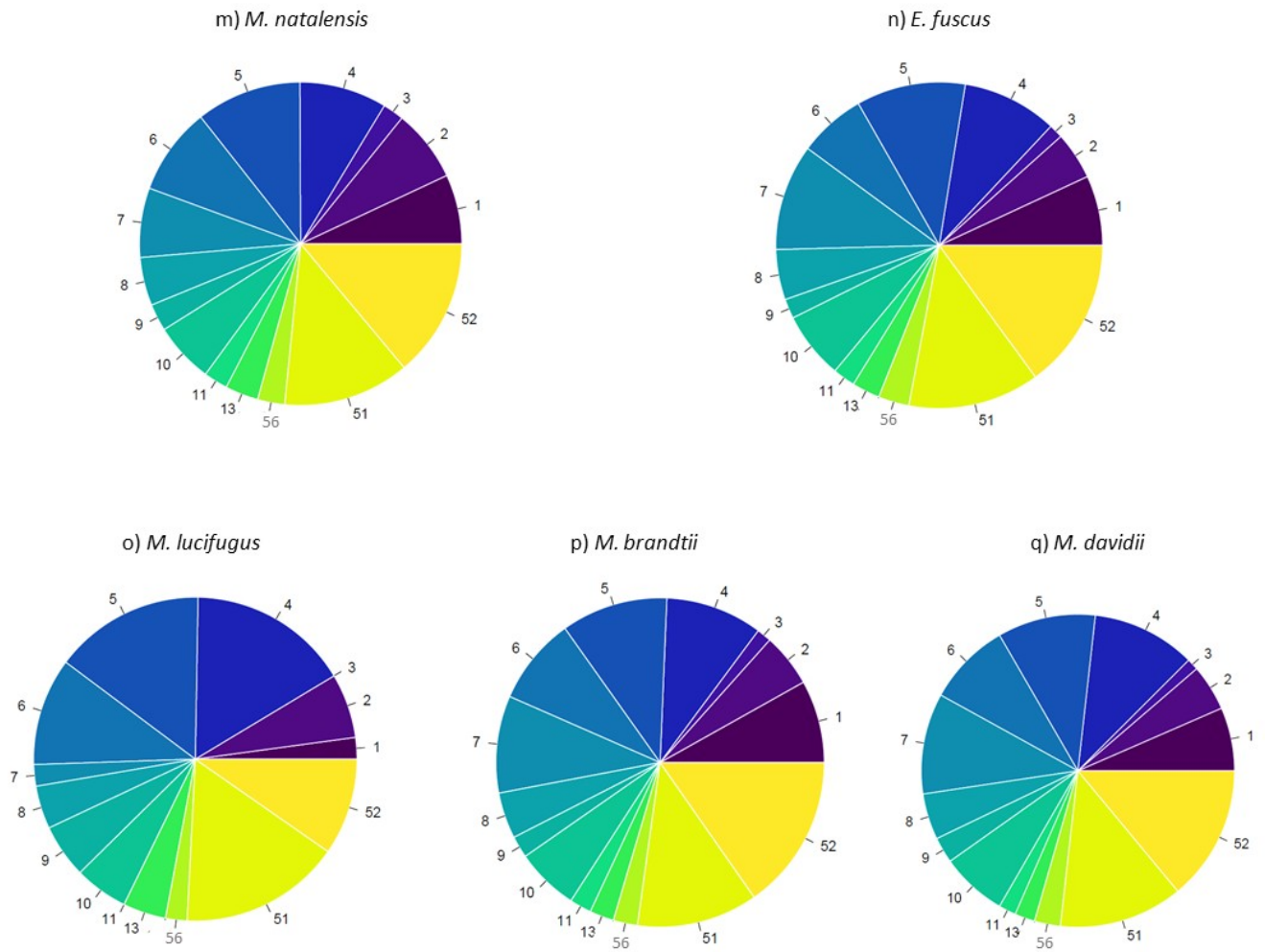


Figure S2. Functional gene copies of 16 olfactory receptor families for each bat genome.

Capítulo III

Atlas transcripcional y mecanismos de regulación de glucosa en dos especies de vertebrados nectarívoros: el murciélago magueyero menor (*Leptonycteris yerbabuenae*) y el colibrí de oreja blanca (*Hylocharis leucotis*)

Introducción

Los niveles de azúcar en sangre (BGL: *Blood Glucose Levels*, por sus siglas en inglés) que se definen como la diferencia entre la tasa de azúcar que entra y la que se remueve por circulación (Song & Schmidt 2012), son mantenidas por mecanismos de homeóstasis y otros sistemas que permiten sentir y controlar las variaciones de BGLs, evitando los diferentes daños patológicos asociados a una mala regulación, como por ejemplo la *diabetes mellitus* (Beuchat & Chong, 1998; Martínez del Río & Gutiérrez-Guerrero, 2020). En mamíferos, los mecanismos de regulación de glucosa, están principalmente controlados por mecanismos dependientes de insulina, en el que intervienen un gran número de reacciones, enzimas y genes (Nagarajan et al. 2017). En particular, el hígado es considerado como la predominante fuente de producción de glucosa (ya que la glucosa se mueve libremente a través de la membrana en cualquier dirección), con una pequeña contribución de otros órganos como el riñón y el intestino, así como otros tejidos (por ejemplo, el músculo esquelético) (Polakof et al. 2011).

El control de la glicemia (concentraciones de azúcar en sangre) en animales depende de diferentes factores que a lo largo de la historia evolutiva de las especies han experimentado cambios en los sistemas de control, sentir y regular las concentraciones de glucosa en sangre (Polakof et al. 2011). Uno de estos cambios, involucra las transiciones y especializaciones del hábito alimenticio, como es el caso de la nectarivoría (Martínez del Río & Gutiérrez-Guerrero, 2020).

Desde una perspectiva macroevolutiva, el surgimiento de la nectarivoría ha sido uno de los eventos más importantes en la evolución (Fleming et al., 2009), promoviendo la radiación e incremento en las tasas de diversificación de especies de plantas y animales, junto con una serie de adaptaciones en las plantas y en los animales que interactúan (Kay & Sargent, 2009; Serrano-Serrano et al., 2017) y también ha promovido adaptaciones únicas a una dieta con alto contenido de azúcares (Tremlett et al., 2019; Suarez et al., 2017). En particular, dos grandes grupos de especies de vertebrados nectarívoros, los colibríes (en las aves) y los murciélagos nectarívoros (en los mamíferos), se han adaptado a un hábito alimenticio dominado por el consumo de néctar.

La convergencia a una dieta dominante en azúcares en ambos grupos de vertebrados voladores (colibríes y murciélagos nectarívoros) ha estado acompañada de conductas y adaptaciones convergentes (Ingersoll et al., 2018). Por ejemplo, los rasgos morfológicos como el largo del pico o el alargamiento del hocico de las especies les permiten tener un mejor acceso al néctar de las flores (Martínez del Río et al., 1992; Muchala et al., 2009). Dentro sus rasgos fisiológicos se ha documentado una eficiente oxidación del azúcar, conocida como "cascada de oxidación del azúcar" (análoga a cascada de fosforilación oxidativa), promoviendo un rápido flujo de azúcar en los músculos activos (Suárez et al., 2011, Welch et al., 2014). Durante el ejercicio en humanos, alrededor del 30% de la tasa máxima de O^2 , proviene de la ingesta oral de glucosa-

fructosa y contribuyen directamente al 15 - 12 % de la producción energética (Jetsen et al., 2014). Este mecanismo es relevante, ya que en murciélagos nectarívoros y colibríes el consumo de carbohidratos es el combustible principal y la oxidación del azúcar provee del ~95% al ~80% de la energía requerida durante el vuelo (Chen et al., 2013).

La tolerancia a la ingesta de alimentos con altas concentraciones de azúcares que han evolucionado colibríes y murciélagos nectarívoros, sin presentar un efecto patológico, como sucede con otras especies, es en parte resultado de una serie de adaptaciones que surgen desde la divergencia de ambos grupos de vertebrados voladores (Martínez del Río & Gutiérrez-Guerrero, 2020). Las aves son el grupo que presentan los niveles más altos de azúcar en sangre reportados en vertebrados (Fig. 1) (Polakof et al., 2011). El estado normal y saludable de las aves, es similar al estado de pacientes diabéticos, pero sin presentar los efectos y daños patológicos que se observa en otros mamíferos con esta comorbilidad (Beuchat & Chong 1998). En general, las aves presentan valores de glicemia entre ~15.6mM a 28mM (estimadas para 511 especies; Polakof et al., 2011), con BGLs que pueden llegar hasta los ~450 mmg/ml, siendo valores que son de 2 a 5 veces mayores que en mamíferos (Fig. 1) (Martínez del Río & Gutiérrez-Guerrero, 2020).

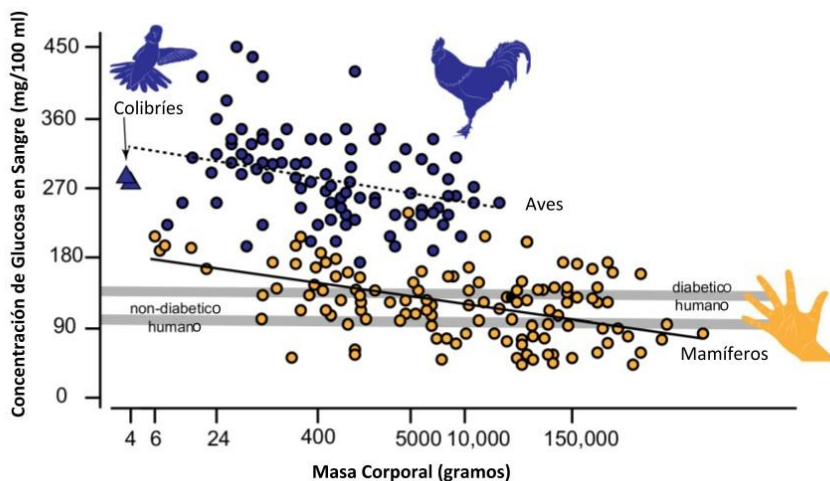


Figura 1. Relación entre las concentraciones de glucosa en sangre reportada en aves y mamíferos en relación a su masa corporal (Figura tomada y modificada de Martínez del Río & Gutiérrez-Guerrero, 2020).

Dada esta condición, tanto colibríes como aves con diferentes tipos de dieta son consideradas hiperglucémicas (Sitbon et al., 1980; Beuchat & Chong 1998), con mecanismos de regulación de insulina diferente a los mamíferos (Dupont et al., 2008) y con mecanismos para evitar el daño de los productos de la glicación (AGEs, por sus siglas en inglés *Advanced Glycation End Products*), los cuales se ha demostrado causan daño a diferentes órganos y pacientes con diabetes presentan un incremento en la acumulación de los AGEs (Rondeae & Bourdon 2010). Diferentes estudios han documentado que las aves han evolucionado mecanismos de resistencia a la glicación, que involucra desde pérdida de genes (como el

receptor de AGEs: RAGE) y bajas concentraciones de diferentes moléculas (por ejemplo, taurina, lisinas o metilglioxal), que previenen la formación de productos AGEs (Swergold & Miller 2014). Aunado a lo anterior, las aves han perdido algunos genes cruciales para el metabolismo y homeostasis de glucosa en sangre, como es el caso de la proteína transportadora de azúcar en sangre *GLUT4* (*SLC2A4*) (Braun & Sweazea 2008). Sin embargo, Coudert et al. (2015) identificaron en aves que la proteína transportadora *GLUT12* presenta actividad similar a *GLUT4* y que actúa en respuesta a la insulina. La regulación de la glucosa por medio de insulina-glucagon en aves parece ser diferente a los mamíferos y otros vertebrados, donde los niveles de glucagon exceden los niveles de insulina (~2 veces), siendo el glucagon el principal regulador de glucosa en sangre en aves (Braun & Sweazea 2008; Polakof et al., 2011), mientras que en mamíferos la insulina es el principal sistema de regulación de glucosa (Polakof et al., 2011).

Los murciélagos, a pesar de su reducida masa corporal, mantienen BGLs relativamente normales en comparación con otros mamíferos (de 7 a 8.33 mM) (Polakof et al., 2011; Peng et al., 2017; Martínez del Río & Gutiérrez-Guerrero, 2020). Sin embargo, algunos autores han observado un incremento significativo en BGLs en murciélagos nectarívoros después de alimentarse de hasta ~25mM, siendo niveles considerados patológicos para otras especies de mamíferos (Tabla1) (Kelm et al., 2011; Peng et al., 2017). Por otra parte, Peng et al. (2017), observaron en murciélagos nectarívoros y frugívoros del Viejo Mundo (*Eonycteris spelea* y *Cynopterus sphinx*) que el vuelo es un mecanismo regulador para mantener la homeóstasis de azúcar durante el forrajeo. Mientras que Kelm et al. (2011) sugieren que los murciélagos nectarívoros del Nuevo Mundo (*Glossophaga soricina*) después de alimentarse, usan el vuelo como una estrategia para reducir sus niveles de glucosa en sangre y así prevenir un estado de hiperglucemia y procesos de glicación de proteínas.

Tabla 1. Rangos y tasas estimados de glicemia y relacionadas con los niveles de glucosa en aves y mamíferos. Colibríes*; Murciélagos nectarívoros**

	Aves	Mamíferos
Glicemia (mg/100ml)	280 - 504	126 81 - 126**
Rango concentración glucosa en sangre (mg/100ml)	180 - 450	40 - 230
BGLs posterior alimentación (mg/100ml)	> 243 > 306*	180 450**
Tasa oxidación de glucosa durante vuelo ($\mu\text{mol g}^{-1} \text{min}^{-1}$)	14.8*	9.1**
Principal sistema regulación glucosa	Glucagon	Insulina
Sensor glucosa en sangre	GLUT2 / GLUT8	GLUT2
Homeostasis glucosa	GLUT12	GLUT4

Se han identificado mecanismos evolutivos con respecto a ciertas características fisiológicas, morfológicas y conductuales, pero aún no se conocen a profundidad los detalles de estas convergencias a nivel genético entre los murciélagos nectarívoros y los colibríes, y se conoce muy poco de los mecanismos de regulación de la glucosa entre estos vertebrados voladores.

En este capítulo analizamos el atlas transcripcional del colibrí de oreja blanca (*Hylocharis leucotis*) y el murciélago magueyero menor (*Leptonycteris yerbabuena*). A partir de la secuenciación del RNA de diferentes tejidos, evaluamos la expresión diferencial del RNA mensajero. Asimismo, incorporamos información de RNAseq de hígado para especies insectívoras (el ave *Ficedula albicollis* y el murciélago *Mormoops megalophylla*), con tres principales objetivos: 1) Identificar genes sobre-expresados involucrados en mecanismos de regulación de glucosa en los vertebrados nectarívoros; 2) Estimar la expresión de genes que son producto de AGEs, dentro y entre murciélagos y aves; y 3) Analizar los mecanismos de regulación y señalización de glucosa, dependiente de insulina entre murciélagos y aves.

Metodología

Obtuvimos datos del transcriptoma de un individuo macho adulto del murciélago filostómido *Leptonycteris yerbabuena*, conocido como murciélago magueyero menor y del colibrí *Hylocharis leucotis*, comúnmente llamado colibrí de oreja blanca. Los transcriptomas se obtuvieron inicialmente para tener evidencia transcripcional *ab inito* para la anotación del genoma.

Obtención de muestras y secuenciación de RNA

Se realizó el muestreo de campo. El sitio de captura del individuo del murciélago magueyero menor fue en la cueva “El Salitre” (Ejido Emiliano Zapata, Morelos, México), y del colibrí de oreja blanca se realizó en Mineral del Monte, Pachuca, ambos con permisos de colecta SGPA/DGVS/07161/15 por la Secretaría de Medioambiente y Recursos Naturales (SEMARNAT).

Para la obtención de la muestra se utilizaron redes de niebla, y una vez capturado los individuos fue mantenido en un costal de tela para disminuir el estrés. Siguiendo todas las recomendaciones y protocolos de ética y técnicos nacionales por la SEMARNAT, los individuos fueron procesados humanitariamente. Se realizaron las disecciones utilizando material esterilizado para obtener muestras de los tejidos: riñón, páncreas, cerebro, pulmón, hígado, corazón, lengua y músculo. Todas las muestras fueron depositadas en buffer *RNAlater Stabilization Reagent*, para mantener la estabilización del RNA y evitar la degradación, cambios en los patrones de expresión e inducción transcripcional.

Las extracciones del RNA se realizaron siguiendo el protocolo del kit RNeasy Tissue (Qiagen), con modificaciones en la cantidad de tejido. La disrupción y homogenización de tejidos

se realizó con un *TissueLyser* (ThermoScientific). Se verificó la extracción del RNA de cada uno de los tejidos por medio de electroforesis. Asimismo, para todas las extracciones se calculó la calidad y cantidad de RNA por nanodrop y Qubit (ThermoScientific).

Las muestras se enviaron al laboratorio y servicios de secuenciación de “Vincent Coates Genomics, University of Berkeley”. Las extracciones de RNA mensajero para cada tejido fueron evaluadas por Bioanalyzer, así como la integridad del RNA. Aquellos tejidos que tuvieron un valor de RIN (RNA Integrity) ≥ 8 , fueron secuenciadas, con la construcción de librerías ribo-depleted de 150 pb y la secuenciación de las muestras en una línea de Illumina HiSeq 4000 PE 150, para cada especie.

Para el murciélago magueyero menor, se secuenciaron cinco tejidos: cerebro, páncreas, hígado, riñón y pulmón, mientras que para el colibrí ojo blanco se secuenciaron seis tejidos: cerebro, páncreas, hígado, pulmón, corazón y lengua (Tabla S1).

Ensamble y anotación del transcriptoma

Inicialmente, se verificó la calidad de las lecturas de cada uno de los tejidos con el programa Fastqc (Patel & Jain, 2012), y se removieron aquellas de baja calidad, reteniendo aquellas lecturas con un valor PHRED ≥ 30 con un *script* de python (*qualityControl.py*). Utilizando el programa SeqPrep (Zhang et al., 2016), se unieron las lecturas *paired-end* (R1 y R2) que empalman 20 nucleótidos (generando lecturas más largas, llamadas en inglés *single-longer reads*).

Para obtener el ensamble del transcriptoma se utilizó el programa Trinity v 2.1.1 (Grabherr et al., 2011). Para el ensamble del transcriptoma (construido con Trinity), se realizó un filtrado y limpieza de calidad utilizando el programa SeqClean (*SeqClean*.<http://seqclean.sourceforge.net>), el cual nos permite remover contigs de baja calidad y residuos de cadenas polyA. Posteriormente, se utilizó el programa DeconSeq (Schmieder & Edwards, 2011) para eliminar posible contaminación dentro del transcriptoma.

Con el transcriptoma ensamblado (filtrado por SeqClean y DeconSeq), se identificaron los marcos de lectura más largos y completos (codón de inicio y término), utilizando el *script* de python “get-orfs-or -cds” y el programa Alignwise (Evans & Loose, 2015). La anotación se realizó por *blastp*, utilizando un e-value de 1×10^{-5} y un hit como mínimo (HSP). Para la anotación del transcriptoma de *L. yerbabuena*, se utilizaron los proteomas de cinco especies de murciélagos: *Myotis brandtii*, *Myotis lucifugus*, *Miniopterus natalensis* (Familia *Vespertilionidae*); *Roussseatus aegyptiacus* y *Pteropus vampyrus* (Familia *Pteropodidae*), y las cuatro especies de mamíferos mejor anotados (*Homo sapiens*, *Equus caballus*, *Canis lupus familiaris* y *Mus musculus*). Para la anotación de *H. leucotis*, se utilizaron los proteomas de cuatro especies de aves: el colibrí (*Calypte anna*), el vencejo (*Chaetura pelagica*), el pinzon diamante cebra (*Tanupia guttata*), y el pollo (*Gallus gallus*).

Además, se realizó la anotación de los transcriptomas utilizando el programa InterProScan, con el cuál se anotaron también los términos *Gene Ontology* (GO) (Ashburner et al., 2000; Jones et al., 2014).

Análisis Expresión Diferencial

Con el programa RSEM (RNA-Seq by Expectation Maximization) (Li & Dewey, 2011) se realizó el análisis de expresión diferencial para cada tejido utilizando el transcriptoma ensamblado y las lecturas filtradas por calidad de RNAseq, siguiendo los parámetros recomendados para muestras que no cuentan con réplicas biológicas.

Como primer paso se indexó el transcriptoma ensamblado con Bowtie 2 (*script rsem-prepare-reference*; Langmead & Salzberg, 2012). Posteriormente se realizó un mapeo de lecturas, utilizando como referencia el transcriptoma ensamblado e indexado y las lecturas filtradas por calidad de cada uno de los tejidos (*script rsem-calculate-expression*). A partir del número de fragmentos alineados, se generó una matriz de abundancia relativa y la normalización de las lecturas (FPKM: *fragments per kilobase transcript length per million fragments mapped* y TPM: *transcripts per million*; *script abundance_estimates_to_matrix.p*). Posteriormente, se realizó el análisis de expresión diferencial (DE), empleando el método edgeR v2 (Robinson et al., 2009).

En el análisis de expresión diferencial con edgeR se calculó la dispersión de las muestras con un análisis que describe la varianza de los datos y liga tanto a la varianza como a la media a través de una distribución binomial negativa (Anders & Huber, 2010; McCarthy et al., 2012). Algunos autores proponen que para calcular la dispersión cuando no se tienen réplicas biológicas se puede utilizar cada uno de los tejidos como muestras independientes (Gentleman et al., 2004). La dispersión se calculó utilizando el paquete de R: Bioconductor y las librerías de Bioconductor: *edgeR* y las funciones *DGEList*, y *estimateGLMCommonDisp*, utilizando el método “*deviance*” (Gentleman et al., 2004).

Se realizó un análisis de agrupamiento, para conocer la agrupación de los tejidos y la expresión diferencial entre ellos. Con el objetivo de identificar los genes con expresión diferencial, se realizaron pruebas de tipo *quasi-likelihood F-test*, que permiten tener un control sobre los falsos positivos cuando el número de réplicas es pequeño, ajustando los datos a modelos lineales. Posteriormente con la función *makecontrast* de la librería edgeR, se encontraron los genes específicos de cada tejido y con las funciones *gmlQLFTest.default* y *decideTestDGE* de la misma librería se identificaron los genes que tienen los mayores niveles de expresión evaluando aquellos con valores de *log fold changes* (logFC) mayor a 2, y un *p-value* ajustado por *False Discovery Rate* (FDR) del 5%.

Enriquecimiento de funciones Gene Ontology (GO)

Para aquellos transcritos que mostraron una expresión diferencial se hizo un análisis de enriquecimiento de funciones, evaluando las categorías de *Gene Ontology* (Biological Process, Molecular Functions y Cellular Components). El análisis de enriquecimiento se realizó utilizando la paquetería de R “topGO” y la anotación obtenida por InterProScan (Alexa et al., 2020). Este análisis se hizo independiente para los datos del murciélago y colibrí.

Análisis transcripcional del hígado entre vertebrados nectarívoros e insectívoros

A partir de una revisión de la literatura, se obtuvieron las secuencias de RNA de hígado del murciélago insectívoro *Mormoops megalophylla* (Tabla S1). Se descargaron los datos de RNA-seq de *M. megalophylla* (tres individuos), que se encuentran en la base *Sequence Read Achivement* (SRA) del NCBI con códigos de acceso: SRR7910149, SRR7910150 y SRR7910152.

Se utilizó la información de RNA del hígado del ave insectívora *Ficedula albicollis*, conocida como el pájaro mosquero. La secuenciación se realizó por la plataforma Illumina HiSeq 2500, 150PE, con librerías de enriquecimiento de colas de poliA (Uebbing et al., 2013). Los datos crudos de RNA-Seq de hígado de 3 individuos se descargaron de la base SRA con códigos de accesión: SRR9617529, SRR9617541 y SRR9617548 (Tabla S1).

Los datos crudos de RNA-Seq de *M. megalophylla* y *F. albicollis*, fueron filtrados por análisis de calidad y se removieron adaptadores (utilizando los programas que se describen en la sección de *Ensamble y anotación del transcriptoma*). Se realizaron de manera independiente, dos ensambles de transcriptoma. Por un lado se ensamblaron los datos de RNAseq de hígado del murciélago insectívoro y nectarívoro (*M. megalophylla* y *L. yerbabuena*). En segundo término, se ensambló el transcriptoma de los datos de RNAseq de hígado, de los tres individuos de *F. albicollis* y la muestra de hígado de *H. leucotis*. Ambos ensambles se generaron con el programa Trinity v2, fueron filtrados, se obtuvieron las secuencias codificantes y proteínas con los programas y parámetros que se describen en la sección

Con estos ensambles, se hicieron análisis de expresión diferencial con los programas RSEM y edgeR, generando los archivos de contraste y de réplicas. Para los datos de los hígados de murciélagos se utilizó como contraste los transcritos expresados en hígado entre el murciélago insectívoro y nectarívoro (para el murciélago insectívoro, se utilizando tres réplicas). Para el análisis de expresión diferencial en aves el contraste se hizo utilizando la categoría nectarívora (*H. leucotis*) e insectívora (*F. albicollis*, se utilizaron las tres réplicas de hígado).

Los transcritos que presentaron una sobre-expresión se obtuvieron utilizando un criterio de $\log_{2}FC \geq 5$ y un valor de p ajustado por medio del test FDR para que sea ≤ 0.05 . Estos transcritos que mostraron expresión diferencial, fueron anotados por medio de *blastp* utilizando la base de datos de UniProt y por medio de InterProScan.

Finalmente, a partir de una revisión de la literatura (García-Vallejo et al., 2006; Gkogkolou & Böhm 2012; Swergold & Miller 2014; Adams et al., 2016), se hizo una lista de los genes y enzimas que participan en las reacciones de glicosilación y los productos AGEs (Tabla S2). Estos genes se buscaron en las anotaciones de los transcritos sobre-expresados.

Resultados

Ensamble transcriptomas

El ensamble *de novo* del murciélago magueyero menor esta representado por un total de 3,555,094 contigs con porcentaje de GC de 42.65%. Se identificaron 2,496,926 genes, de los cuales 642,636 son secuencias codificantes con marco de lectura completo. Con InterProScan, se anotó un total de 68,549 de transcritos, y con *blastp* 38,972 secuencias de proteínas, de las cuales el 95.6% de las secuencias están compartidas entre murciélagos y el 4.3% restante comparte el proteoma de *L. yerbabuena* con los otros mamíferos (humano, caballo, perro y ratón).

El ensamble *de novo* del colibrí de oreja blanca contiene 1,100,580 contigs, con un porcentaje de GC de 42.8%. Se identificaron un total de 605,204 genes, de los cuales 242,094 son secuencias codificantes, con marco de lectura completo. Por medio de InterProScan, se anotaron 76,882 secuencias de proteínas y con *blastp* un total de 68,997 proteínas.

Atlas transcriptómico en vertebrados nectarívoros

Se analizó la expresión diferencial de muestras de diferentes tejidos de *L. yerbabuena* y *H. leucotis*, utilizando un valor de $\log_{2}FC > 2$, y el valor de p ajustado a $FDR \leq 0.05$. Bajo estos criterios, el hígado y el cerebro son los tejidos que mostraron una mayor sobre-expresión de transcritos en ambas especies (Fig. 2; Fig. S1). De acuerdo con la anotación realizada por *blastp*, KO (base de datos KEGG) y GO se observa que hay una especificidad en los RNA mensajeros que están sobre-expresados en cada tejido (Tabla S3-S4).

Murciélago Magueyero Menor

Para *L. yerbabuena* en la mayoría de los tejidos se identificó una alta expresión de genes involucrados en defensa y sistema inmune, como la familia de genes Serpin o genes del complejo mayor de histocompatibilidad MHC (Fig. 2).

También se observó la sobre-expresión de ciertos genes relacionados a funciones propias de cada órgano (Fig. 2; TablaS3). Por ejemplo, en el cerebro hay una alta proporción de transcritos sobre-expresados asociados a funciones neuroactivas de ligandos-receptores y receptores de glutamato. En el páncreas, los transcritos DE participan en rutas de señalización de receptores de células B y otros en el sistema inmune (*MHC clase II antigen*) y receptores de ácidos grasos. El pulmón es el tejido con los menores niveles de expresión de transcritos (Fig.

S1), con unos cuantos activos como pulmón-surfactantes. En el riñón del murciélago identificamos transcritos con expresión tejido-específico, como el gen *SLC2A5*, que codifica para la proteína transportadora de hexosas GLUT5. También identificamos genes que participan en el metabolismo de carbohidratos, colesterol y ácidos grasos (*FASN* - ácidos grasos en adipocitos), y otros grupos de genes con funciones relacionadas a la síntesis de hierro y familias de esteroides.

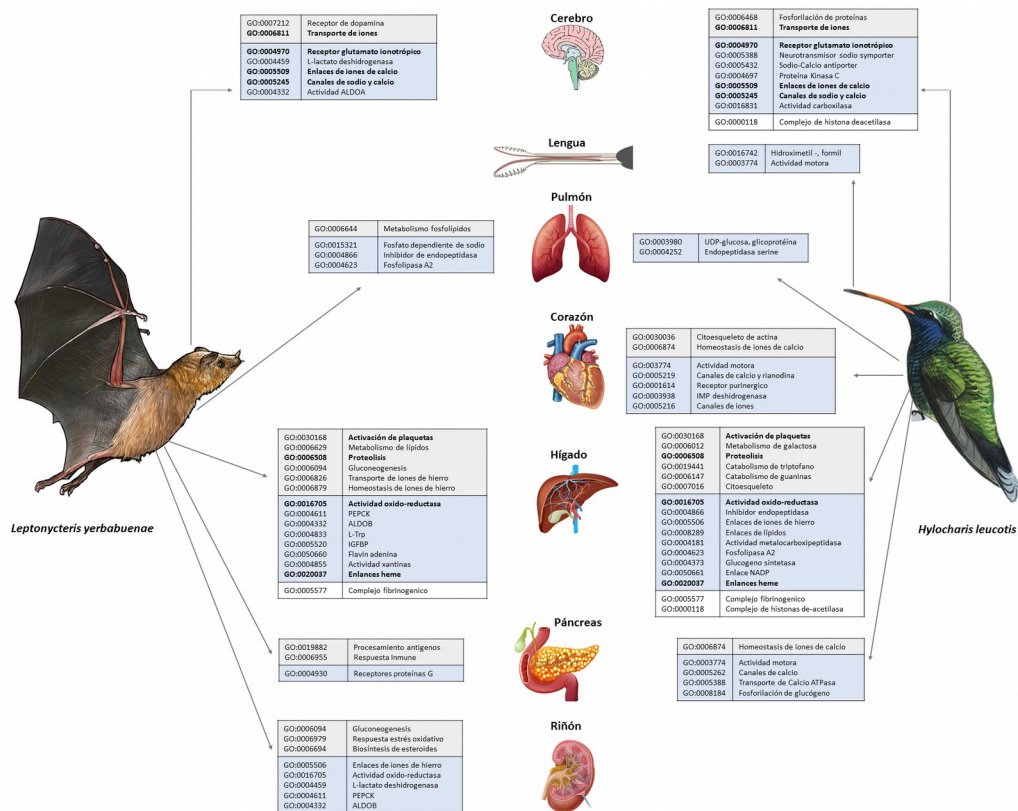


Figura 2. Enriquecimiento de funciones para aquellos transcritos con expresión diferencial en cada tejido. En gris, se presentan los GO de la categoría Biological Process; en azul, los GO de la categoría Molecular Function; y en blanco, los GO de la categoría Cellular Components. En negritas, son los transcritos que se expresan en ambas especies para cada tejido.

En el hígado se identificaron transcritos sobre-expresados en al menos 65 rutas metabólicas. Para el metabolismo de la digestión y absorción de carbohidratos se identificaron transcritos con expresión tejido-específico como el transportador de glucosa, *solute carrier family 2 (SLC2A2)*, que codifica para la proteína *GLUT2*. También se observó una alta expresión de la enzima *alpha-amilasa (AMY)*, quitinasas (*CHIA*), *glucogeno sintetasa (GYS)*, la enzima *glucosa-6-fosfato (GGPC)*, aldosa, *fructose-bisphosphate B (ALDOB)*, *fructose-1,6-bisphosphatase 1 (FBP)*, *carboxypeptidasa*. En el hígado también se identificó un alto número de transcritos que participan en la digestión y absorción de grasas, como son las enzimas de la familia *acyl-coA oxidasa* y *sintetasa*, *glucogeno sintetasa*, *lipopolisacaridos (LBP)*, *fosfolipasas A2*,

*keto*hexoquinasas y ácidos grasos (*FASN*). Otro grupo de transcritos DE, están relacionados en el almacenamiento del hierro, como es el gen *ferritin*, *light polypeptide* (*FTL*).

Colibrí de Oreja Blanca

De acuerdo con los transcritos sobre-expresados en el cerebro, identificamos una alta actividad en funciones neuroactivas, y genes que participan en vías metabólicas de sinápsis y regulación circadiana (Fig. 2). En la lengua encontramos altos niveles de transcritos DE involucrados en el movimiento muscular-esquelético. En el corazón identificamos un gran número de genes regulados negativamente. Sin embargo, los genes activos están relacionados con proteínas heat shock, miosinas, y alta expresión del gen *Titin*, que es importante para la contracción de músculos cardíacos. En el páncreas se observó una alta expresión de transcritos que participan en señales de calcio y adrenérgicos. Se identificó transcritos activos en el metabolismo de la sacarasa [*glucógeno fosforilasa* (*PYG*) y *UTP-glucosa-1-fosfato uridililtransferasa*].

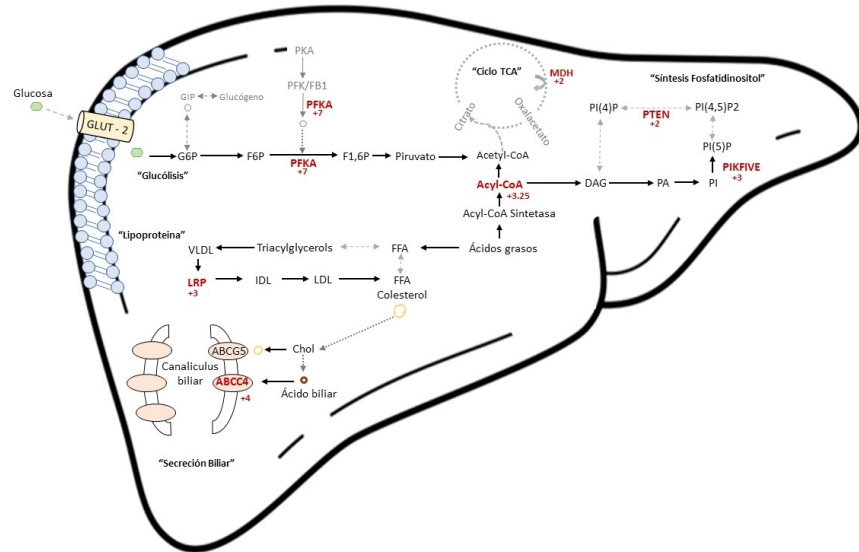
En el hígado del colibrí, se identificaron transcritos sobre-expresados con actividad en 72 rutas metabólicas. En el metabolismo de digestión y absorción de carbohidratos, se identificó al transportador de glucosa, (*SLC2A2*), *acetyl-CoA sintetasa* (*ACSS*), *hexoquinasas*, 6-fosfofructoquinasa y *ALDOB*. En el metabolismo del colesterol, se identificaron transcritos activos como la *colesterol 7alfa-mono oxigenasa* (*CYP7A1*) y *apolipoproteínas* (*APOA*). En el metabolismo de digestión y absorción de grasas, se observaron a las enzimas *keto*hexoquinasas y ácidos grasos sintetasa (*FASN*), enzimas involucradas en la síntesis de ácidos grasos de cadena larga. Asimismo, hubo alta DE de transcritos involucrados en exportar sales biliares.

Genes sobre-expresados paralelamente en hígado entre vertebrados insectívoros y nectarívoros

A partir del transcriptoma ensamblado con los datos de secuenciación de hígado (Tabla S1) realizamos los análisis de expresión diferencial entre murciélagos [*M. megalophylla* (insectívoro) y *L. yerbabuena* (nectarívoro)]. Independientemente, evaluamos la expresión diferencial en aves [*F. albicollis* (insectívoro) y el colibrí *H. leucotis* (nectarívoro)]. Las muestras de RNAseq de hígado para cada especie presentó consistencia en su agrupamiento (Fig. S2).

Para todas las especies observamos una sobre-expresión de genes involucrados en el metabolismo de lípidos, carbohidratos, sistema inmune, reparación de DNA y termoregulación, funciones que son propias del hígado (Tabla 2; Fig. S3). Ambas especies insectívoras presentaron un mayor número de transcritos sobre expresados ($\log_{2}FC \geq 5$; $FDR \leq 0.05$), con funciones relacionadas principalmente a sistema inmune, defensa y un gran número de genes que participan en el metabolismo de lípidos y grasas (Fig. 3).

Figura 3. Rutas metabólicas del azúcar y grasas en hígado. En rojo, genes con expresión diferencial paralela en hígado ambos vertebrados nectarívoros.



En las especies nectarívoras identificamos un total de 8 transcritos activamente expresados ($\log_{FC} \geq 5$; $FDR \leq 0.05$) en rutas metabólicas de azúcares, como en la ruta de la glucólisis - glucogénesis, fructosa-manosa y glucógeno, como son la *glucogeno sintetasa*, la *fructokinasa* (*PFKA*). Otros genes y enzimas que participan en el metabolismo ketogénico y de lipogénesis, como es el gen *PTEN*, genes de la familia *aldo-keto reductasas*, las enzimas *acyl-coenzyme A synthetase* (*ACSM4*) y *acyl-CoA synthetase cadena-larga* (involucradas en la conversión de lípidos a ATP). También en este órgano se encontraron transcritos sobre-expresados que participan en el metabolismo del colesterol y regulación de ubiquitinas (Fig. 3). También en alta expresión se detectaron genes involucrados en la exportación de sales biliares.

Tabla 2. Expresión diferencial y enriquecimiento de funciones de hígado (en gris: Molecular function; azul: Biological Process; blanco: Cellular Components).

Especies	GO	Anotación por GO	Prueba Fisher
<i>F. albicollis</i>	GO:0004523	RNA-DNA hybrid ribonuclease activity	8.1e-06
	GO:0004866	endopeptidase inhibitor activity	0.00015
	GO:0016853	isomerase activity	0.00213
	GO:0003964	RNA-directed DNA polymerase activity	0.00234
	GO:0008378	galactosyltransferase activity	0.01206
	GO:0004623	phospholipase A2 activity	0.01265
	GO:0006298	mismatch repair	0.0023
	GO:0006278	RNA-dependent DNA biosynthetic process	0.0040
	GO:0019318	hexose metabolic process	0.0041
	GO:0006644	phospholipid metabolic process	0.0111
GO:0009086	methionine biosynthetic process	0.0115	
<i>H. leucotis</i>	GO:0003777	microtubule motor activity	0.00011
	GO:0034597	phosphatidylinositol-4,5-bisphosphate	0.00125
	GO:0004402	histone acetyltransferase activity	0.00253
	GO:0003872	6-phosphofructokinase activity	0.00365
	GO:0003677	DNA binding	0.00385
	GO:0004842	ubiquitin-protein transferase activity	0.00640
	GO:0005524	ATP binding	0.00861

	GO:0003918 GO:0005085 GO:0003700 GO:0004525 GO:0004016 GO:0005086 GO:0004672	DNA topoisomerase type II (ATP-hydrolyze) guanyl-nucleotide exchange factor activity) DNA-binding transcription factor activity) ribonuclease III activity adenylate cyclase activity ARF guanyl-nucleotide exchange factor protein kinase activity	0.01162 0.01480 0.01889 0.02327 0.03538 0.04651 0.04695
	GO:0007018 GO:0007219 GO:0006355 GO:0006621 GO:0007275 GO:0006171 GO:0006891 GO:0006511 GO:0006888	microtubule-based movement Notch signaling pathway regulation of transcription, DNA-template protein retention in ER lumen multicellular organism development cAMP biosynthetic process intra-Golgi vesicle-mediated transport ubiquitin-dependent protein catabolic ER to Golgi vesicle-mediated transport	7.8e-05 0.0012 0.0077 0.0104 0.0121 0.0334 0.0342 0.0352 0.0352
	GO:0005634 GO:0000786 GO:0030117	nucleus nucleosome membrane coat	0.0086 0.0203 0.0296
<i>M. megaphylla</i>	GO:0008289 GO:0003854 GO:0017048 GO:0003700 GO:0003779 GO:0046873 GO:0020037	lipid binding 3-beta-hydroxy-delta5-steroid dehydrogenase Rho GTPase binding DNA-binding transcription factor activity actin binding metal ion transmembrane transporter activity heme binding	5.4e-06 0.00050 0.00061 0.00278 0.00358 0.00985 0.01322
	GO:0016043 GO:0006869 GO:0006694 GO:0030001 GO:0050852 GO:0019882	cellular component organization lipid transport steroid biosynthetic process metal ion transport T cell receptor signaling pathway antigen processing and presentation	0.00030 0.00052 0.00072 0.00172 0.01090 0.01338
	GO:0005576	extracellular region	0.0049
<i>L. yerbabuena</i>	GO:0005388 GO:0004370 GO:0042800 GO:0019781 GO:0004838 GO:0003729 GO:0004348 GO:0004014 GO:0004420 GO:0005096 GO:0005525 GO:0003924	calcium-transporting ATPase activity glycerol kinase activity histone methyltransferase activity (H3-K) NEDD8 activating enzyme activity L-tyrosine:2-oxoglutarate aminotransferase mRNA binding glucosylceramidase activity adenosylmethionine decarboxylase activity hydroxymethylglutaryl-CoA reductase (NADH) GTPase activator activity GTP binding GTPase activity	0.00093 0.00418 0.00418 0.01180 0.01200 0.02296 0.02298 0.02298 0.02298 0.02641 0.02744 0.03886
	GO:0006457 GO:0006886 GO:0016192 GO:0006597 GO:0015936 GO:0006680 GO:0006348 GO:0016579 GO:0006606 GO:0006525 GO:0006511	protein folding intracellular protein transport vesicle-mediated transport spermine biosynthetic process coenzyme A metabolic process glucosylceramide catabolic process chromatin silencing at telomere protein deubiquitination protein import into nucleus arginine metabolic process ubiquitin-dependent protein catabolic	0.00089 0.00199 0.00222 0.02079 0.02079 0.02079 0.03326 0.03364 0.04145 0.04789 0.04933
	GO:0035097	histone methyltransferase complex	0.0035

	GO:0008290	F-actin capping protein complex	0.0035
	GO:0005685	U1 snRNP	0.0059
	GO:0005694	chromosome	0.0419
	GO:0030117	membrane coat	0.0427
	GO:0000785	chromatin	0.0438

Mecanismos de regulación y control de glucosa dependiente de insulina en hígado

En los hígados de las especies de murciélagos y aves identificamos cambios en la expresión de genes involucrados en la ruta de señalización de la insulina entre aves y mamíferos, donde los genes expresándose en mamíferos se encuentran sin expresarse en aves (Fig. 4). Identificamos 15 genes/enzimas, de los cuales 9 genes están activos en los murciélagos y sin expresarse en las aves, y los 6 genes restantes están activos en las aves y apagados en murciélagos (Tabla S5). En particular los dos murciélagos presentaron una sobre-expresión de los genes *SOS*, *PPARGC1A* y *PKLR*.

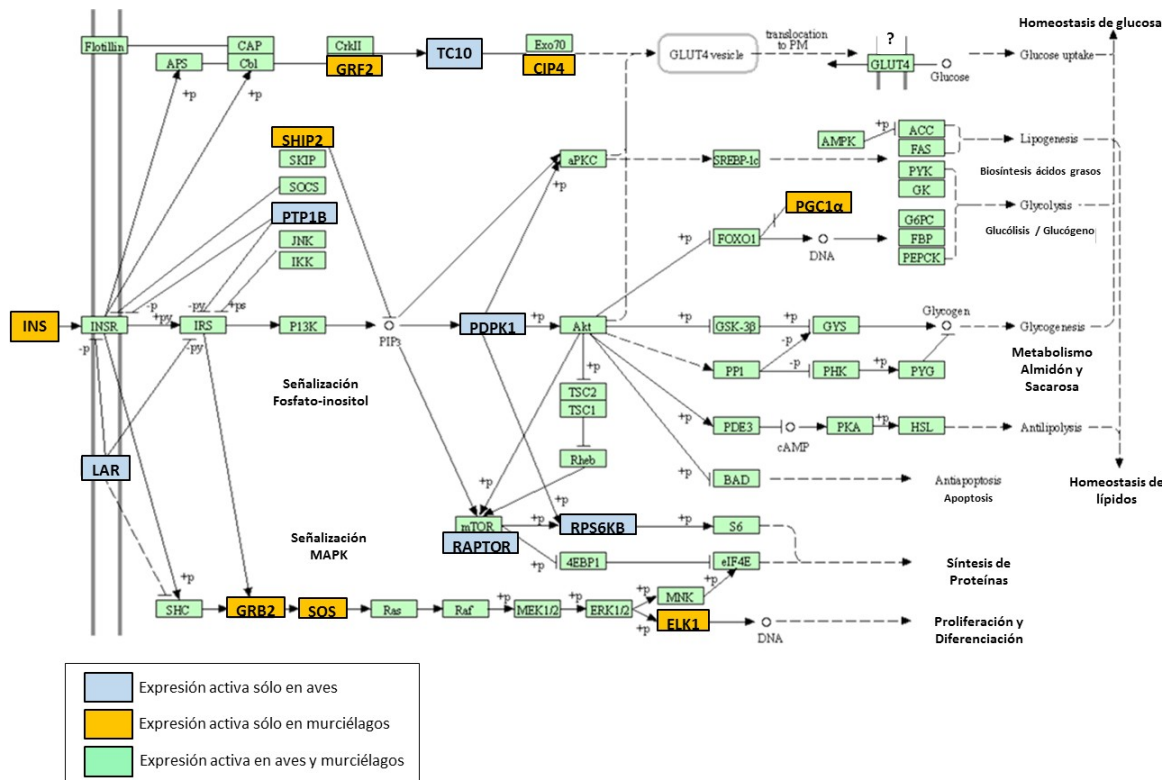


Figura 4. Genes expresados activamente en la ruta de señalización de la Insulina en aves y murciélagos. Mapa de señalización de insulina para humanos (KEGG).

Sobre-expresión de transcritos involucrados en glicación

A partir de la lista de algunos genes involucrados en el proceso de glicación, observamos que sólo algunos genes estaban activamente expresándose en los transcriptomas de hígado para los murciélagos y aves. En los murciélagos identificamos activamente expresándose al gen *DDOST*, el resto de los genes no se encontró en los transcritos anotados (Fig. S4). En el caso de las

aves, *F. albicollis* presenta una expresión activa de la mayoría de los genes identificados en el proceso de glicación. Interesantemente, los genes sobre-expresados en *F. albicollis* ($\log_{2}FC \geq 2$), se encuentran reprimidos en el colibrí (Fig. S4).

Discusión

La adaptación a un nuevo nicho ha promovido la evolución convergente en diferentes características morfológicas, fisiológicas, bioquímicas y genómicas en diferentes especies. En este capítulo describimos patrones de expresión transcripcional paralela en una especie de murciélago nectarívoro *L. yerbabuena* y el colibrí *H. leucotis*, que se caracterizan por una dieta dominante en azúcares.

En la descripción del atlas transcripcional de los dos vertebrados nectarívoros, observamos que un gran número de transcritos tienen una función tejido-específico (Fig. 2; Tabla S3-S4), principalmente en el cerebro, hígado y riñón (este último órgano, sólo en el murciélago maguero menor). En general, encontramos una alta expresión en transcritos involucrados en el transporte, enlaces de iones de hierro y calcio, y procesos de óxido-reducción.

Para *L. yerbabuena* identificamos una alta actividad de funciones relacionadas con metabolismo de lípidos, hierro, conversión de glucosa a ATP y funciones *heat shock*. Estos resultados se compararon con los reportados por Moreno-Santillán et al., (2019) donde analizan los patrones de expresión en transcriptomas de hígados de cinco especies de murciélagos insectívoros [*Mormoops megalophylla* (Mormoopidae), *Myotis keaysi* (Vespertilionidae), *Nyctinomops laticaudatus* (Molossidae) y *Peropteryx macrotis* (Emballonuridae)]; y un frugívoro-omnívoro [*A. jamaicensis* (Phyllostomidae)]. Los autores reportaron para los murciélagos insectívoros una sobre-expresión predominante de genes reguladores y otros involucrados en sistema inmune. Sin embargo, entre el murciélago nectarívoro (*L. yerbabuena*) y el murciélago frugívoro-omnívoro (*A. jamaicensis*), se expresa activamente el gen ALDOB, importante en metabolismo de carbohidratos (Leong, 2018).

En *H. leucotis* un gran número de transcritos sobre-expresados están involucrados en funciones motoras (con alta expresión en lengua), citoesqueleto, procesos de catabolismo, *heat shock*, metabolismo de glucógeno, azúcares y lípidos. En el estudio de Lim et al. (2019) utilizando la información transcriptómica de hígado de 12 especies de colibríes, detectaron que algunas especies presentan señales de selección positiva en genes *heat shock* y otros involucrados en actividad motora y metabolismo de ácidos grasos. También en el estudio del transcriptoma de hígado del colibrí *Archilochus colubris* (Workman et al., 2018), encontraron la expresión de transcritos involucrados en la ruta de triglicéridos y lipogénesis, siendo algunos de estos genes, los mismo que encontramos en *H. leucotis*, como es la enzima *FASN*.

¿Cómo las especies estudiadas pueden mantener una dieta dominante en azúcares? Nuestro análisis comparativo de la expresión de hígado entre especies insectívoras (el

murciélago *M. megalophylla* y el pajarito mosquero *F. albicollis*) y los dos vertebrados nectarívoros, nos permitió avanzar en el estudio del eficiente metabolismo de azúcares y grasas para la obtención de energía que tienen las especies nectarívoras (Fig. 3). En este estudio observamos que el hígado tiene una alta actividad de enzimas que participan en rutas de fosforilación del azúcar expresadas en el murciélago magueyero menor, el colibrí de oreja blanca, y que ha sido reportado para otras especies de colibríes como *Archilochus colubris* (Myrka & Welch, 2018). Este resultado apoya los trabajos de Peng et al. (2017) sobre el eficiente mecanismo que tienen los murciélagos nectarívoros para regular la azúcar ingerida y poder obtener energía durante el vuelo, reduciendo las concentraciones de azúcar en sangre. También Suarez y Welch (2017) reportan en tejido muscular una rápida oxidación del azúcar para obtener ATP en colibríes y murciélagos nectarívoros, ya que el tipo de vuelo flotante es energéticamente muy costoso.

A diferencia de los mamíferos, diferentes autores han identificado que las aves regulan los niveles de glucosa a partir del metabolismo del glucagón y con una menor contribución por medio de insulina (Braun & Sweazea 2008; Polakof et al., 2011). En el presente análisis transcripcional, identificamos que las aves (tanto nectarívora e insectívora) están expresando genes involucrados en la ruta de señalización de insulina (Fig. 4). Sin embargo, un gran número de genes importantes para esta ruta de insulina, no están expresándose (Tabla S5). Este resultado apoya los trabajos de Sitbon et al., (1980), Braun & Sweazea (2008) y Dupont et al (2010), que sugieren que las aves han cambiado sus mecanismos de regulación de glucosa, dependientes de insulina y que posiblemente presentan otros genes que intervienen en esta ruta, con una función similar a los que se han reportado en mamíferos, un ejemplo es la proteína transportadora de glucosa *GLUT12*. Para mamíferos, la principal proteína transportadora de glucosa dependiente de insulina es el *GLUT4*, principalmente expresada en músculo (Suárez et al., 2010; 2011); esta proteína durante la divergencia de las aves se perdió y no se ha reportado en ninguna especie de ave (Braun & Sweazea 2008; Martínez del Río & Gutiérrez-Guerrero, 2020). Sin embargo, Coudert et al. (2015) documentan en pollos que la proteína *GLUT12* ha evolucionado funciones similares a las de *GLUT4*. En este estudio identificamos sólo en el pájaro mosquero una alta expresión de la proteína *GLUT12*.

¿Cómo las especies nectarívoras pueden mantener altos BGLs, sin presentar daños patológicos? (Fig. 1). A pesar de las altas tasas metabólicas que implica el vuelo y los niveles de azúcar en aves y murciélagos (Umminger 1975; Beuchat & Chong 1998; Daan et al., 1990; Peng et al., 2017; Martínez del Río & Gutiérrez-Guerrero, 2020), no observamos transcritos expresándose asociados al proceso de glicación y acumulación de AGEs (productos finales de la glicación) en las especies nectarívoras. Diferentes autores sugieren que las aves tienen mecanismos para evitar o reducir la acumulación de AGE en la hemoglobina. Recientemente, Anthony-Regnitz et al. (2020) observaron que la albumina, es resistente a la glicación en

comparación con mamíferos (vacas). Esta resistencia a la glicación en especies hiperglucémicas está relacionado a diferentes factores. Por ejemplo, las aves perdieron el receptor de AGEs (RAGE o AGER) importante para la producción de AGEs y la concentración de grupos carbonilos para la formación de AGEs es menor en aves, así como el porcentaje de residuos de lisina (la cual tiene alta afinidad a el proceso de glicación). Kelm et al (2011), reportó que murciélagos nectarívoros (como *G. soricina*) tienen una alta tolerancia a los altos niveles de azúcar en sangre, ya que a pesar de los altos BGLs que pueden alcanzar durante la ingesta de néctar, la producción de AGEs es similar a la reportada para otros mamíferos con diferentes dietas, a pesar de que las concentraciones de azúcar que pueden mantener estos murciélagos nectarívoros son similares a los de pacientes diabéticos.

Conclusión

La evolución a un nicho dietético dominante en azúcares, está acompañado de mecanismos con un combustible energético en común.

Los resultados obtenidos en el Capítulo I de esta tesis muestran señales de selección positiva y paralelismos moleculares en sitios específicos de amino ácidos para el gen *Acetyl-CoA sintetasa (ACCS)*, exclusivo de murciélagos nectarívoros de la subfamilia Glossophaginae y los Pteropodidos (Gutiérrez-Guerrero et al., 2020). En el estudio de Workman et al., (2018), identificaron para *A. colubris*, señales de selección positiva y altos niveles de expresión para *Acetyl-CoA carboxilasa (ACC)*. En este análisis transcriptómico, observamos una sobre-expresión de enzimas de las familias *acyl-CoA* y *Acetyl-CoA*, lo que nos sugieren que el metabolismo de ácidos grasos y procesos de lipogénesis, en los dos vertebrados nectarívoros voladores (*H. leucotis* y *L. yerbabuena*) son importantes para mantener una dieta dominante en azúcares y eficientizar el metabolismo de lípidos, de manera que pueden obtener altas cantidades de ATP y cubrir los requerimientos energéticos que implica el forrajeo y mantener el vuelo estacionario. Es importante mencionar que mutaciones en los genes mencionados están relacionados con desordenes metabólicos en humanos, como obesidad, hiperglucemia y *diabetes mellitus II*.

La evolución convergente al hábito de la nectarivoría, que involucra restricciones nutricionales en lípidos y proteínas en colibríes y murciélagos nectarívoros, ha estado acompañada de una serie de adaptaciones paralelas fisiológicas y bioquímicas. Por lo anterior, sugerimos que futuros estudios incorporen la información transcriptómica de otros tejidos como el músculo, intestino y riñón, de manera que podamos identificar otros transcritos involucrados en la fosforilación y oxidación del azúcar y metabolismo de lípidos. También será importante incluir información del microbioma de estas especies especialistas a la nectarivoría y analizar como las comunidades bacterianas pueden contribuir en obtener los nutrientes necesarios para satisfacer las demandas nutricionales.

Material Adicional

Tabla S1. Información de los datos de RNAseq de murciélagos y aves insectívoras utilizados de hígado.

ID	Especie	Tejido	SRA accesión	Secuenciación	Reads >= PHRED 30	Referencia
1	<i>Mormoops megalophylla</i>	Hígado	SRR7910149	Illumina HiSeq 4000 PE Librería ribodepleted	29,129,191	Moreno-Santillán et al., 2019
2	<i>Mormoops megalophylla</i>	Hígado	SRR7910150	Illumina HiSeq 4000 PE Librería ribodepleted	15,300,629	Moreno-Santillán et al., 2019
3	<i>Mormoops megalophylla</i>	Hígado	SRR7910152	Illumina HiSeq 4000 PE Librería ribodepleted	22,793,577	Moreno-Santillán et al., 2019
1	<i>Artibeus jamaicensis</i>	Hígado	SRR7910151	Illumina HiSeq 4000 PE Librería ribodepleted	32,393,565	Moreno-Santillán et al., 2019
2	<i>Artibeus jamaicensis</i>	Hígado	SRR7910153	Illumina HiSeq 4000 PE Librería ribodepleted	30,849,601	Moreno-Santillán et al., 2019
3	<i>Artibeus jamaicensis</i>	Hígado	SRR7910154	Illumina HiSeq 4000 PE Librería ribodepleted	19,886,556	Moreno-Santillán et al., 2019
1	<i>Ficedula albicollis</i>	Hígado	SRR9617529	Illumina HiSeq 2500 PE Librería poliA	47,440,544	Uebbing et al., 2013
1	<i>Ficedula albicollis</i>	Corazón	SRR9617531	Illumina HiSeq 2500 PE Librería poliA	50,758,299	Uebbing et al., 2013
1	<i>Ficedula albicollis</i>	Riñón	SRR9617528	Illumina HiSeq 2500 PE Librería poliA	52,205,491	Uebbing et al., 2013
2	<i>Ficedula albicollis</i>	Hígado	SRR9617548	Illumina HiSeq 2500 PE Librería poliA	46,489,126	Uebbing et al., 2013
2	<i>Ficedula albicollis</i>	Corazón	SRR9617550	Illumina HiSeq 2500 PE Librería poliA	50,740,714	Uebbing et al., 2013
2	<i>Ficedula albicollis</i>	Riñón	SRR9617547	Illumina HiSeq 2500 PE Librería poliA	50,982,077	Uebbing et al., 2013
3	<i>Ficedula albicollis</i>	Hígado	SRR9617541	Illumina HiSeq 2500 PE Librería poliA	52,554,069	Uebbing et al., 2013
3	<i>Ficedula albicollis</i>	Corazón	SRR9617543	Illumina HiSeq 2500 PE Librería poliA	54,865,423	Uebbing et al., 2013
3	<i>Ficedula albicollis</i>	Riñón	SRR9617544	Illumina HiSeq 2500 PE Librería poliA	47,652,854	Uebbing et al., 2013

Tabla S2. Lista de genes involucrados en procesos de glicación, reportados en literatura.

Gen (abreviación)	Nombre completo del gen
RAGE / AGER	Receptor for Advanced Glycation End Products
MT1A	Metallothionein 1A
AGE-R1	Oligosaccharyl Transferase-48
AGE-R3 / LGALS3	Galectin-3
SR-A	macrophage scavenger receptors type I and type II
DDOST	Dolichyl-Diphosphooligosaccharide-Protein Glycosyltransferase Subunit
AGE-R2/ PRKCSH	Protein Kinase C substrate 80K-H
MSR1	Macrophage Scavenger Receptor 1
CHIA	Acidic Chitinase
DDX19A	Box polypeptide 19A
MYOM3	Myomesin 3
ISCA2/NCP2	NPC intracellular cholesterol transporter 2
FBXO33	<i>F-box Protein 33</i>
MACROD2	MACRO domain containing 2
ZBTB38	<i>BTB domain containing 38</i>
PALLD	palladin, cytoskeletal associated protein
B3GALT β4GALT	Galactosyltransferases
MGAT IGNT RFNG C2GNT	N-Acetylglucosaminyltransferases
FUT[1-10]	Fucosyltransferase
ST3GAL ST6GAL ST8SIA	Sialyltransferases

Tabla S3. Transcritos expresados diferencialmente en los cinco tejidos para el Murciélago Maguerero Menor.

Tejido	GO	Función	p-value
Cerebro	GO:0006811	<i>Ion transport</i>	2.6e-06
	GO:0007212	<i>Dopamine receptor signaling pathway</i>	0.00013
	GO:0005230	<i>Extracellular ligand-gated ion channel</i>	6.5e-09
	GO:0004970	<i>Ionotropic glutamate receptor activity</i>	5.7e-07
	GO:0004888	<i>Transmembrane signaling receptor</i>	7.9e-06
	GO:0032051	<i>Clathrin light chain binding</i>	0.00018
	GO:0005509	<i>Calcium ion binding</i>	0.00067
	GO:0004459	<i>L-lactate dehydrogenase activity</i>	0.00559
	GO:0004949	<i>Cannabinoid receptor activity</i>	0.00559
GO:0016021	<i>Integral component of membrane</i>	7e-05	
Pulmón	GO:0015321	<i>Sodium-dependent phosphate transmembrane</i>	0.0036
	GO:0004623	<i>Phospholipase A2 activity</i>	0.0060
Hígado	GO:0030168	<i>Platelet activation</i>	1e-05
	GO:0006559	<i>L-phenylalanine catabolic process</i>	1e-05
	GO:0006508	<i>Proteolysis</i>	0.00015
	GO:0009074	<i>Aromatic amino acid family catabolic</i>	0.00041
	GO:0006570	<i>Tyrosine metabolic process</i>	0.00048
	GO:0006869	<i>Lipid transport</i>	0.00076
	GO:0006629	<i>Lipid metabolic process</i>	0.00358
	GO:0009086	<i>Methionine biosynthetic process</i>	0.00456
	GO:0006094	<i>Gluconeogenesis</i>	0.00930
	GO:0006879	<i>Cellular iron ion homeostasis</i>	0.00930
	GO:0004252	<i>Serine-type endopeptidase activity</i>	5.0e-12
	GO:0004866	<i>Endopeptidase inhibitor activity</i>	2.1e-10
	GO:0005506	<i>Iron ion binding</i>	4.4e-09
	GO:0020037	<i>Heme binding</i>	2.1e-07
	GO:0016705	<i>Oxidoreductase activity, acting</i>	2.2e-06
	GO:0004867	<i>Serine-type endopeptidase inhibitor</i>	0.00014
	GO:0004611	<i>Phosphoenolpyruvate carboxykinase</i>	0.00020
	GO:0004332	<i>Fructose-bisphosphate aldolase activity</i>	0.00035
	GO:0016491	<i>Oxidoreductase activity</i>	0.00035
	GO:0004411	<i>Homogentisate 1,2-dioxygenase activity</i>	0.00048
	GO:0004833	<i>Tryptophan 2,3-dioxygenase activity</i>	0.00048
	GO:0004838	<i>L-tyrosine:2-oxoglutarate aminotransferase</i>	0.00048
	GO:0005520	<i>Insulin-like growth factor binding</i>	0.00320
	GO:0050660	<i>Flavin adenine dinucleotide binding</i>	0.00507
	GO:0004499	<i>N,N-dimethylaniline monooxygenase</i>	0.00684
	GO:0005577	<i>Fibrinogen complex</i>	1.1e-05
Páncreas	GO:0019882	<i>Antigen processing and presentation</i>	0.0043
Riñón	GO:0005506	<i>Iron ion binding</i>	5.2e-11
	GO:0020037	<i>Heme binding</i>	1.2e-10
	GO:0016705	<i>Oxidoreductase activity, acting on paire...</i>	1.6e-10
	GO:0004459	<i>L-lactate dehydrogenase activity</i>	0.0033
	GO:0004185	<i>Serine-type carboxypeptidase activity</i>	0.0099
	GO:0004602	<i>Glutathione peroxidase activity</i>	0.0099

Tabla S4. Transcritos expresados diferencialmente en los seis tejidos del colibrí de oreja blanca.

Tejido	GO	Función	p-value
Cerebro	GO:0006811	Ion transport	3.4e-12
	GO:0007156	Homophilic cell adhesion via plasma memb...	0.00052
	GO:0005230	Extracellular ligand-gated ion channel a...	3.9e-12
	GO:0004888	Transmembrane signaling receptor activit...	2.2e-11
	GO:0004970	Ionotropic glutamate receptor activity	2.1e-09
	GO:0005216	Ion channel activity	0.00026
	GO:0015293	Symporter activity	0.00065
	GO:0004930	G protein-coupled receptor activity	0.00076
	GO:0005248	Voltage-gated sodium channel activity	0.00104
	GO:0005509	Calcium ion binding	0.00105
	GO:0005388	Calcium-transporting ATPase activity	0.00113
	GO:0016020	Membrane	0.00073
	GO:0001518	Voltage-gated sodium channel complex	0.00152
Lengua	GO:0016742	Hydroxymethyl-, formyl- and related	0.0029
Corazón	GO:0006874	Cellular calcium ion homeostasis	7.2e-05
	GO:0003774	Motor activity	2.3e-10
	GO:0005262	Calcium channel activity	7.7e-05
	GO:0005219	Ryanodine-sensitive calcium-release	8.4e-05
Pulmón	GO:0003980	UDP-glucose:glycoprotein glucosyltransferase	0.00051
Hígado	GO:0006012	Galactose metabolic process	0.00070
	GO:0006559	L-phenylalanine catabolic process	0.00092
	GO:0030168	Platelet activation	0.00315
	GO:0006570	Tyrosine metabolic process	0.00520
	GO:0004252	Serine-type endopeptidase activity	1.1e-08
	GO:0004866	Endopeptidase inhibitor activity	2.9e-06
	GO:0020037	Heme binding	0.00011
	GO:0004867	Serine-type endopeptidase inhibitor	0.00023
	GO:0004499	N,N-dimethylaniline monooxygenase	0.00032
	GO:0005506	Iron ion binding	0.00038
	GO:0016705	Oxidoreductase activity, acting	0.00088
	GO:0004181	Metalloprotease activity	0.00431
	GO:0008289	Lipid binding	0.00453
	GO:0022857	Transmembrane transporter activity	0.00574
	GO:0004411	Homogentisate 1,2-dioxygenase activity	0.00608
	GO:0005509	Calcium ion binding	0.00688
	GO:0005507	Copper ion binding	0.00711
	GO:0004335	Galactokinase activity	0.00915
	GO:0004047	Aminomethyltransferase activity	0.00915
	GO:0005576	Extracellular region	2.9e-12
	GO:0005577	Fibrinogen complex	0.0076
	GO:0016021	Integral component of membrane	0.0092
	Páncreas	GO:0006874	Cellular calcium ion homeostasis
GO:0003774		Motor activity	7.0e-07
GO:0005219		Ryanodine-sensitive calcium-release	3.7e-06
GO:0005262		Calcium channel activity	5.3e-05
GO:0005388		Calcium-transporting ATPase activity	0.0083

Tabla S5. Transcritos activamente expresados (on) y sin expresarse (off) en los murciélagos (*L. yerbabuena* y *M. megalophylla*) y aves (*H. leucotis* y *F. albicollis*), involucrados en la ruta de señalización de insulina y resistencia a insulina.

Ruta de Resistencia de Insulina	Gen (abreviación)	Enzima [EC]	On/Off
NF-kappa-B inhibitor alpha	NFKBIA		On
Nitric-oxide synthase, endothelial	NOS3	1.14.13.39	Off
Insulin receptor substrate 1	IRS1		Off
Tribbles homolog 3	TRIB3		Off
Ribosomal protein S6 kinase beta	RPS6KB	2.7.11.1	On
Acetyl-CoA carboxylase / biotin carboxylase 2	ACACB	6.4.1.2 6.3.4.14 2.1.3.15	Off
Carnitine O-palmitoyltransferase 1, muscle isoform	CPT1B	2.3.1.21	On
Insulin	INS		Off
Alpha peroxisome proliferator-activated receptor gamma coactivator 1-alpha	PGC1A		Off
Forkhead box protein O1	FOXO1		Off
Solute carrier family 27 (fatty acid transporter), member 5	SLC27A5	6.2.1.7	Off
Novel protein kinase C epsilon type	PRKCE	2.7.11.13	On
3-phosphoinositide dependent protein kinase-1	PDPK1	2.7.11.1	On
Protein O-GlcNAcase / histone acetyltransferase	NCOAT/MGEA5	3.2.1.169 2.3.1.48	Off
Protein O-GlcNAc transferase	OGT	2.4.1.255	On
CREB-regulated transcription coactivator 2	CRTC2 TORC2		Off
Ruta de Señalización de Insulina			
Receptor-type tyrosine-protein phosphatase F	LAR	3.1.3.48	On
Rap guanine nucleotide exchange factor 1	RAPGEF1 GRF2		Off
Ras homolog gene family, member Q	RHOQ TC10		On
Thyroid hormone receptor interactor 10	TRIP10 CIP4		Off
Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 2	INPPL1 SHIP2	3.1.3.86	Off
Tyrosine-protein phosphatase non-receptor type 1	PTP1B PTPN1	3.1.3.48	On

Alpha peroxisome proliferator-activated receptor gamma coactivator 1-alpha	PGC1A; PPARGC1A		Off
Pyruvate kinase isozymes R/L	PKLR	2.7.1.40	Off
Insulin	INS		Off
3-phosphoinositide dependent protein kinase-1	PDPK1	2.7.11.1	On
Regulatory associated protein of mTOR	RAPTOR		On
Ribosomal protein S6 kinase beta	RPS6KB	2.7.11.1	On
Growth factor receptor-binding protein 2	GRB2		Off
Son of sevenless	SOS		Off
ETS domain-containing protein Elk-1	ELK1		Off

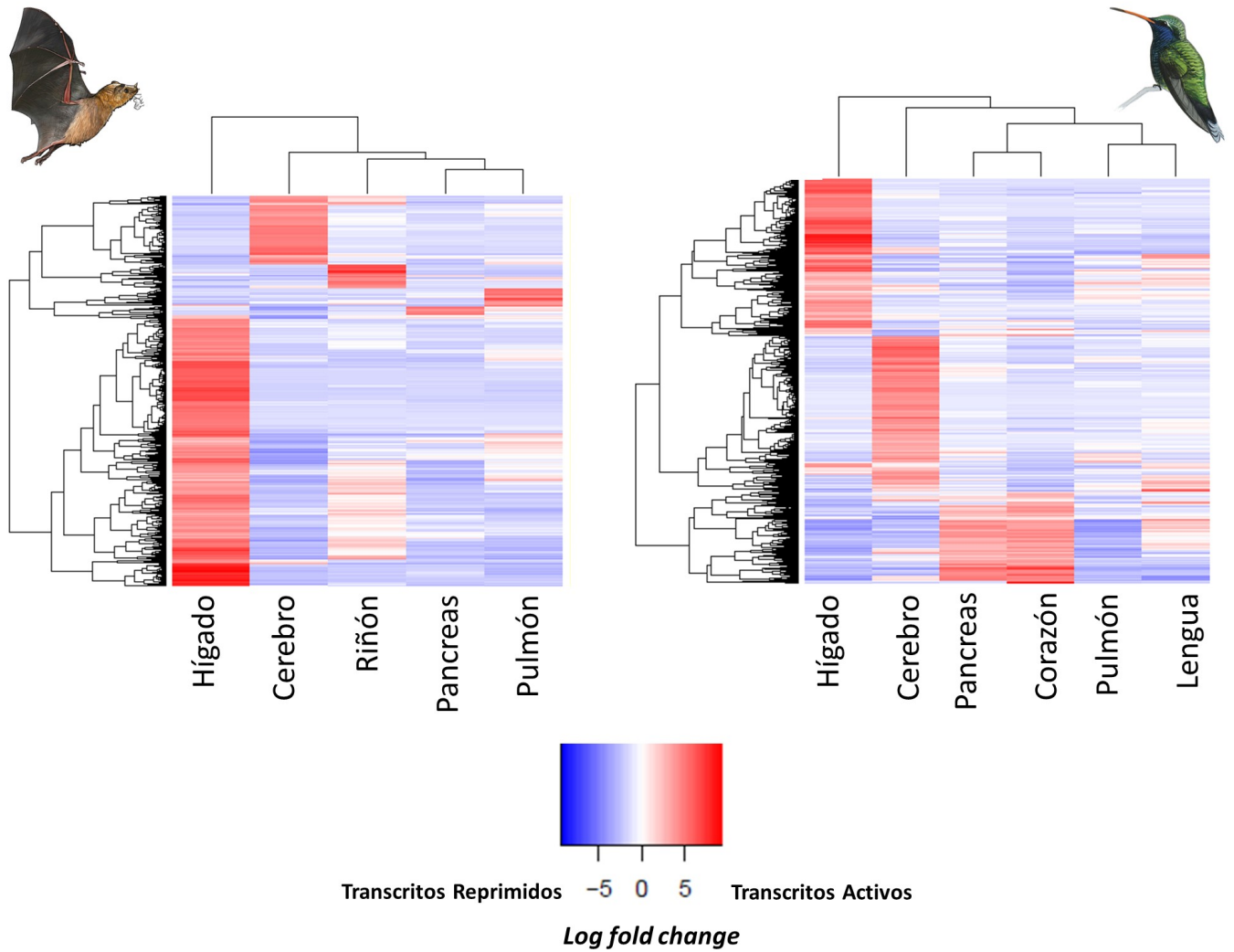


Figura S1. *Heatmap* y agrupamiento de distancias del análisis de expresión diferencial de los transcritos sobre-expresados y regulados negativamente de cada tejido. A la derecha, se presentan los resultados cinco tejidos de *L. yerbabuena*. A la izquierda, se muestran los resultados para los 6 tejidos de *H. leucotis*.

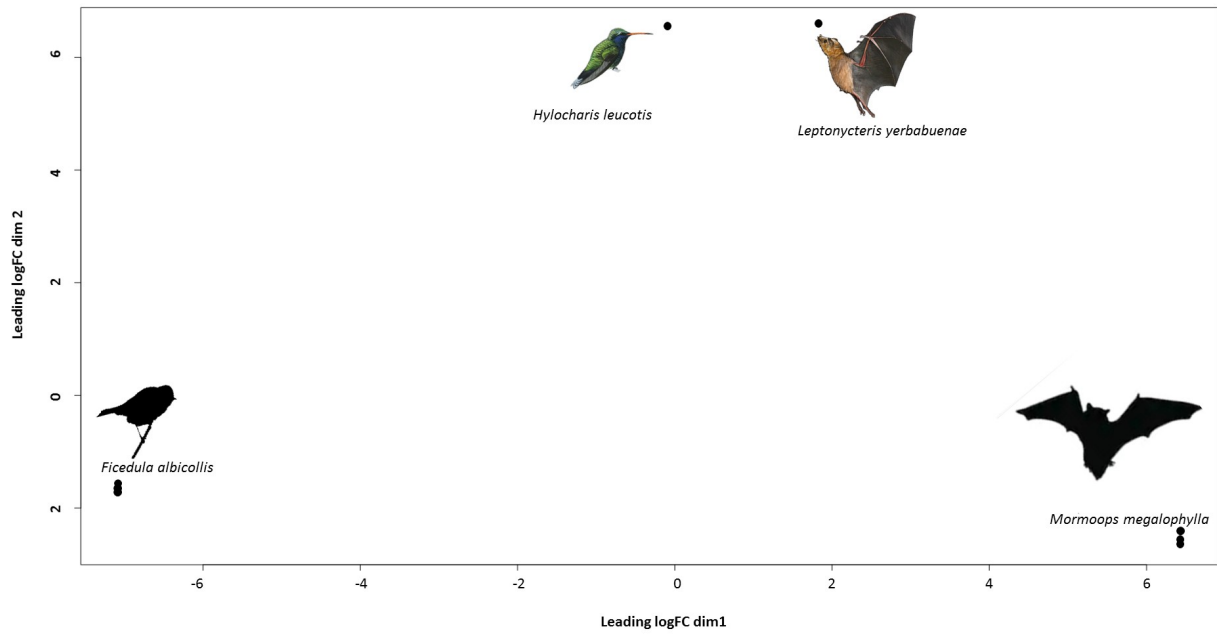


Figura S2. Gráfica *Multi-Dimensional Scaling*, que agrupa consistentemente las muestras de transcritos para cada especie y sus correspondientes réplicas.

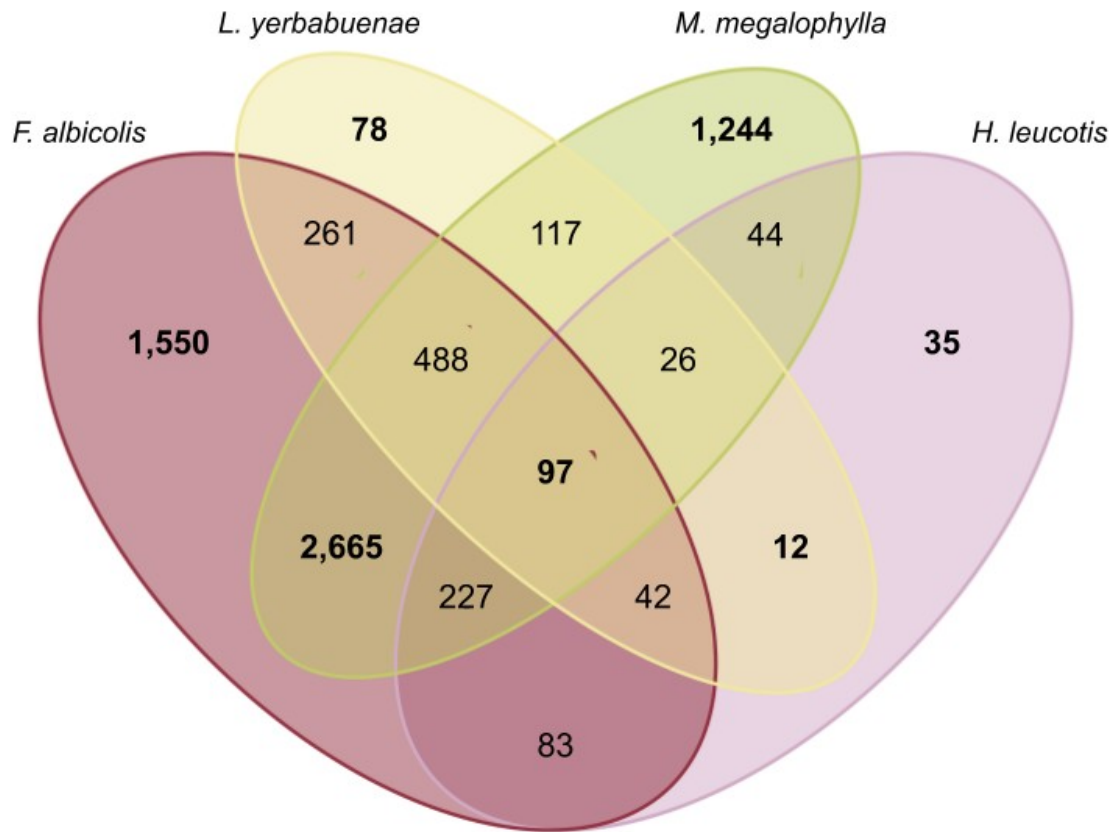


Figura S3. Anotación funcional utilizando la base de datos KEGG, de los genes con una sobre-expresión en el hígado.

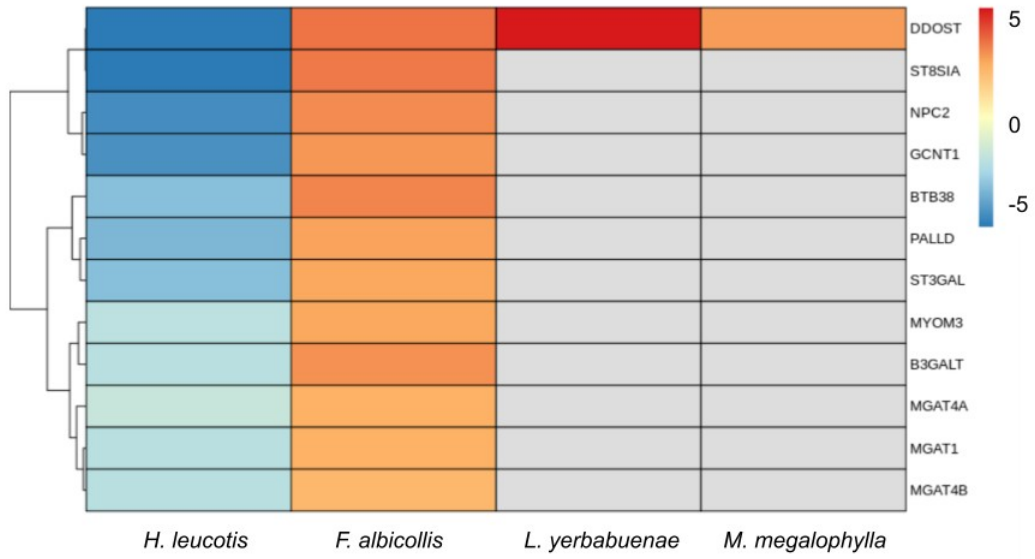


Figura S4. Heatmap de los genes involucrados en procesos de glicación y acumulación de AGEs identificados en los transcriptomas de hígados de los murciélagos y aves.

Discusión general

La evolución y la plasticidad de las diferentes dietas son eventos fundamentales en la biología de los animales. Los cambios y transiciones de hábitos alimenticios han permitido identificar su relación con eventos de diversificación de linajes y riqueza de especies a lo largo de la historia evolutiva de los animales, (Cattin et al. 2014; Olla-Tárraga et al. 2017). Por ejemplo, la diversidad de nichos dietéticos como la fitofagia en los insectos, promovió la radiación de especies de hexápodos a finales del Paleozoico (Wiens et al. 2010; 2015), pero también ha involucrado una serie de interacciones bióticas con las plantas, que pueden ser recíprocas o no. Un ejemplo clásico es el estudio de Ehrlich y Raven (1964) donde describieron como la presión de insectos herbívoros sobre algunas especies de plantas promovió la evolución de nuevos fitoquímicos (por ejemplo, los alcaloides), producto de metabolitos secundarios como respuesta de defensa a la herbivoría, lo que a su vez promovió la radiación de linajes de mariposas (por ejemplo, los grupos Papilioninae-Troidini y Nymphalinae-Heliconiini), quienes evolucionaron mecanismos fisiológicos y bioquímicos para descomponer esos fitoquímicos, desintoxicar a estas especies de plantas y alimentarse de ellas.

En el cambio de dieta también ha estado acompañado de radiaciones adaptativas a nuevos nichos ecológicos, fisiológicos, involucrando cambios genómicos (Gómez et al., 2010; Price et al., 2012; Román-Palacios et al. 2019) como son los procesos de optimización de genes, duplicaciones, pérdida de genes o pseudogenizaciones (Ledezma et al., 2018; Gutiérrez-Guerrero et al. 2020). Recientemente, Román-Palacios et al. (2019) han sugerido que la carnivoría es probablemente la dieta ancestral en el reino animal, e identificaron que el 63% de los taxones tienen una dieta carnívora, el 32% son herbívoros y sólo un 3% de los linajes son omnívoros. Las especies que se han adaptado a una dieta especialista son menos abundantes, pero representan importantes eventos de especiación y radiación, ya que la transición a una dieta especialista, ha estado sujeta a diferentes fuerzas evolutivas, dejando huellas genómicas con señales de adaptación en específicas rutas metabólicas (Wäckers et al. 2007; Flemming et al. 2009; Boulton et al. 2018).

Una de las especializaciones de la dieta más extremas es la hematofagia (Boulton et al. 2018), donde la alimentación está dominada o restringida al consumo de sangre. Esta dieta ha evolucionado independientemente en diferentes linajes del reino animal y se remonta a hace unos 220 millones de años, con un fósil de protomosquito (Schutt et al., 2008; Boulton et al. 2018). Dentro de los insectos y arácnidos, existen aproximadamente 15,000 especies que se alimentan de sangre, como los mosquitos (familia Culicidae), las chinches (familia Cimicidae) y los piojos (familia Pulicidae) (Mans et al., 2002). En los agnatos se encuentran las lampreas (género *Petromyzon*), en los anélidos están las sanguijuelas (la subclase Hirudinae). En aves la hematofagia evolucionó recientemente en el pinzón vampiro (*Geospiza septentrionalis*), quién

también consume semillas e insectos (Grant et al. 2000). En los mamíferos, los hematófagos estrictos son un grupo de murciélagos (Schutt et al., 2008), conocidos como vampiros verdaderos, que se encuentran dentro de la familia Desmodontinae conformado por tres especies, de las cuales sólo *Desmodus rotundus* es un hematófago obligado (Schutt et al., 2008; Boulton et al. 2018). En esta tesis, detectamos que el vampiro *D. rotundus*, exhibe fuertes señales de selección positiva ($dN/dS > 1$; $pval = 0.001$), en genes involucrados en rutas que regulan la coagulación de la sangre. La mayoría de las especies hematófagas, tienen en común la presencia de componentes anticoagulantes en la saliva (Boulton et al. 2018; Ledezma et al. 2018; Gutiérrez-Guerrero et al. 2020) y también una serie de pseudogenizaciones en sus sistemas sensoriales de olfato y sabor y un incremento en su percepción térmica, como es el caso del vampiro. Si bien, la hematofagia ha evolucionado independientemente en diferentes linajes, la evolución de estos mecanismos fisiológicos y sensoriales claramente sugieren eventos de evolución convergente (Boulton et al. 2018).

Otra de las especializaciones en el hábito alimenticio es la nectarivoría, que ha promovido una fuerte radiación tanto de especies de animales como de angiospermas (Feinsinger et al. 1987). El consumo de néctar implica una ingesta elevada de azúcares, ya que cerca del 100% de su peso seco son azúcares, como la sacarosa, la fructosa y la glucosa (Koutsots et al. 2016). Si bien las especies nectarívoras obtienen la mayor cantidad de energía a partir del néctar, esta dieta por sí sola contiene muy bajos niveles de amino ácidos, vitaminas y minerales (Martínez del Río et al. 1982; Koutsots et al. 2016; Gutiérrez-Guerrero et al. 2020). Es por ello, que la forma en que las especies compensan estos requerimientos nutricionales ausentes en el néctar, es con el consumo directo e indirecto del polen (Koutsots et al. 2016). El polen concentra un alto contenido de amino ácidos esenciales (que van de 9 hasta 16 diferentes tipos de amino ácidos), lípidos, un rico contenido de nitrógeno y minerales, por lo que el consumo combinado de néctar y polen, permite a las especies nectarívoras cubrir los requerimientos nutricionales esenciales (Braker et al. 1962).

Desde el Cretácico la dieta nectarívora y polínica ha promovido interacciones mutualistas, de cooperación y también procesos de evolución recíproca (Sun et al. 1998). Por ejemplo, Schoen y Clegg (1985) observaron que el efecto del color de la flor *Ipomoea purpurea*, tiene un efecto importante en la tasa del éxito reproductivo de sus polinizadores, los abejorros, quienes incrementan la frecuencia de visita a las flores púrpuras. Otro de los procesos macroevolutivos que ha promovido la nectarivoría es la coevolución entre ciertos grupos de plantas y animales polinizadores, quienes simultáneamente han estado bajo presiones de selección recíprocas (Feinsinger et al. 1987; Flemming et al. 2008: 2009). Un ejemplo, es la coevolución difusa entre murciélagos nectarívoros y los agaves y cactus columnares (Braker et al. 1982; Feinsinger et al. 1987; Flores-Abreu et al. 2019), estas plantas han experimentado cambios en los rasgos florales como respuesta al síndrome de polinización de la quiropterofilia y

algunos murciélagos glossofaginos presentan adaptaciones morfológicas, como el alargamiento del hocico, y una especialización de sistemas sensoriales para localizar a las flores, entre otros (Braker et al. 1982; Feinsinger et al. 1987). Además, Flores-Abreu (2019) reporta similares tiempos de radiación entre ambos grupos (aproximadamente 8 a 12 millones de años).

Los insectos son el grupo que alberga un mayor número de especies nectarívoras y polínicas, principalmente de las familias Lepidóptera, Díptera, Coleóptera Heteróptera y Thysanóptera (Wäckers et al. 2007). Recientemente, Sendi et al. (2020) se encontraron fósiles en ámbar (*Lepidopterix vegrandis*, de la familia Lepidóptera) que datan del Cretácico y donde también quedaron preservadas muestras de polen, indicando que los insectos son las especies nectarívoras y polínicas más antiguas en la historia de la vida. En los invertebrados, el cambio de una dieta ancestral carnívora a nectarívora/polínica ha tenido varias transiciones: de carnívoro – herbívoro – néctar/polén, de carnívoro a néctar/polén; o de carnívoro – omnívoro – néctar/polén (Wäckers et al. 2007). En vertebrados, las especies con una dieta nectarívora/polínica son menos abundantes, pero esta dieta ha promovido una importante radiación de especies en dos grupos: los colibríes en las aves y en dos grupos de murciélagos en los mamíferos (Fleming et al. 2008; 2009).

La evolución a la nectarivoría, en especies filogenéticamente distantes como son los murciélagos nectarívoros y colibríes (Fleming et al. 2008; Muschala et al. 2009) involucra adaptaciones convergentes como, por ejemplo, el vuelo estacionario y las estrategias de forrajeo implican una alta demanda energética, para las cuales, estos dos nectarívoros han evolucionado mecanismos involucrados en una eficiente conversión del azúcar en energía (Suárez et al., 2009; 2012). Pero también, involucra adaptaciones en genes relevantes en el metabolismo de azúcares y principalmente de lípidos. Ambos grupos animales han evolucionado mecanismos, que les permite compensar los requerimientos en el metabolismo de las grasas, dado una dieta que obtiene principalmente las fuentes de proteínas y de lípidos solo del polen.

En particular la evolución de la nectarivoría en murciélagos de la familia Phyllostomidae, nos ha permitido estudiar a mayor profundidad como esta dieta ha promovido la radiación de grupos como los Glossofaginos (Martínez del Río et al. 1982; Schondube et al. 2000; Gutiérrez-Guerrero et al., 2020), donde se encuentran especies ecológicamente importantes para el ecosistema, como son los murciélagos nectarívoros que hemos estudiado en esta tesis, el murciélago magueyero menor (*L. yerbabuena*), especie clave para el ecosistema y económicamente importante para la industria del mezcal, el murciélago magueyero mayor (*L. nivalis*) y el murciélago platanero (*M. harrisoni*), estos últimos son polinizadores, especies endémicas e importantes para la biodiversidad.

El estudio de la evolución de la nectarivoría utilizando herramientas de secuenciación de nueva generación, bioinformáticas y de evolución molecular en estos murciélagos con una dieta

dominante en azúcares, nos ha permitido detectar señales de selección positiva ($dN/dS > 1$; $pval = 0.001$), que participan en rutas metabólicas de carbohidratos y lípidos. Esto nos sugiere que dichos murciélagos metabolizan de una manera eficiente los azúcares que obtienen del néctar y los escasos lípidos que obtienen del polen, para librar los requerimientos energéticos necesarios que implica el vuelo estacionario y forrajeo (Suárez et al., 2010). También identificamos que duplicaciones en familias génicas involucradas en el almacenamiento de hierro pueden estar relacionadas en prevenir desordenes metabólicos, como son la anemia, dado que la nectarivoría es una dieta con escasos nutrientes, como es el hierro.

La información de los genes codificantes nos permitió identificar como la especialización de la dieta dominante en el consumo de azúcares, ya sea por néctar o frutos maduros (que concentran altos contenidos de azúcares), ha promovido la evolución paralela en linajes filogenéticamente lejanos entre los murciélagos nectarívoros del Nuevo Mundo (*M. harrisoni*, *L. nivalis* y *L. yerbabuena*) y los murciélagos frugívoros del Viejo Mundo (*R. aegyptiacus*, *P. alecto* y *P. vampyrus*). Estas señales de evolución paralela exclusiva en estos murciélagos pueden ser una respuesta a las altas demandas energéticas que implica el forrajeo y que permiten eficientizar la conversión de grasas a energía. Consideramos relevante que futuros análisis incluyan la información genómica de otras especies de murciélagos de la familia Pteropodidae, con una dieta nectarívora como *Cynoterus sphinx* (Elangovan 2000), y otros murciélagos de la familia Phyllostomidae con una dieta frugívora estricta, como especies de los géneros *Sturnira* y *Anoura* (Willing et al. 1993). El tener esta información también nos permitirá estudiar la evolución sensorial para esta dieta en la percepción de las diferentes formas, olores y texturas de las frutas, identificando la relevancia de ciertos receptores sensoriales como el *receptor olfativo 51E2 Clase I*.

La adaptación a un nicho dietético dominado por azúcares que han evolucionado colibríes y murciélagos nectarívoros, está acompañado de mecanismos con un combustible energético en común. En el análisis transcriptómico de dos vertebrados nectarívoros voladores (*H. leucotis* y *L. yerbabuena*) (Capítulo III) observamos una sobre-expresión de genes cruciales para el metabolismo de ácidos grasos y procesos de lipogénesis. Estos resultados apoyan estudios fisiológicos y bioquímicos previos sobre el eficiente mecanismo de metabolizar azúcares y lípidos de manera que se pueda obtener altas cantidades de ATP y cubrir los requerimientos energéticos que implica el forrajeo y mantener el vuelo estacionario. Asimismo, a pesar de la condición hiperglucémica de las aves y los altos niveles de azúcar en sangre que pueden llegar a alcanzar los murciélagos nectarívoros, ambas especies voladoras tienen mecanismos para evitar y reducir los procesos de glicación y la acumulación de AGEs.

La evolución de las dietas, principalmente la nectarivoría, ha sido ampliamente estudiada desde la perspectiva de la ecología, taxonomía, filogenética, filogeografía y la genética de

poblaciones. Incorporar a estas investigaciones, herramientas genómicas, bioinformáticas y la teoría de la evolución molecular, nos ha permitido avanzar en esta tesis en el estudio de los genes y las presiones adaptativas relacionadas con la dieta. En específico, identificamos en el clado de los murciélagos nectarívoros genes con una función optimizada en el metabolismo del azúcar. En humanos se ha documentado que mutaciones en algunos de estos genes están asociadas con desórdenes metabólicos, como es la *diabetes mellitus*, obesidad, anemia e hiperglucemia.

Finalmente, entender la evolución de las dietas en la historia biológica de los diferentes linajes de animales requiere del estudio paralelo de sus comunidades microbianas que las que co-habitan e interactúan. Futuros estudios para conocer las comunidades bacterianas intestinales nos permitirá estudiar cómo los procesos de pseudogenización en los genomas de los animales son compensados (o no) con el papel que tiene su microbiota intestinal. Por ejemplo, las aves y algunos murciélagos no tienen la capacidad de digerir la trehalosa (principal azúcar de los exoesqueletos de los insectos), ya que han perdido o se ha pseudogenizado el gen de la *Trehalasa* (Brun et al. En preparación). Sin embargo, las comunidades bacterianas son las encargadas en descomponer la *Trehalosa* en azúcares más simples. El estudio del microbioma intestinal nos permitirá tener mayor información de como las especies adaptadas a una dieta especialista, obtienen a partir de diversos grupos de bacterias los requerimientos nutricionales esenciales.

Conclusiones generales

1. Se secuenció a mediana y alta cobertura los genomas de cinco especies de murciélagos filostómidos incluyendo al murciélago magueyero menor (*L. yerbabuena*), especie clave para el ecosistema y económicamente importante para la industria del mezcal; al murciélago magueyero mayor, y al murciélago platanero (*L. nivalis* y *M. harrisoni*, respectivamente), ambos polinizadores y especies endémicas; al murciélago frugívoro y dispersor de semillas (*A. jamaicensis*); y al murciélago orejudo (*M. waterhousii*), especie insectívora y tal vez importante como controlador de plagas para la agricultura.
2. Las transiciones y especializaciones del hábito alimenticio han estado acompañadas de cambios genómicos asociados al metabolismo y a la percepción sensorial, que sugieren la acción de diferentes fuerzas evolutivas y la presencia de diferentes adaptaciones.
3. El consumo de insectos (estado ancestral para el orden Quiróptera), es actualmente relevante para los filostómidos independientemente su hábito alimenticio. Los genes involucrados en el metabolismo y obtención de nutrientes a partir del consumo de insectos, muestran claras señales de selección.
4. La transición a una dieta especialista, ha estado sujeta a diferentes fuerzas evolutivas, dejando huellas genómicas, con señales de adaptación en específicas rutas metabólicas. Por ejemplo, debido a su dieta especialista en el consumo de sangre, el vampiro (*D. rotundus*) exhibe fuertes señales de selección positiva ($dN/dS > 1$; $pval = 0.001$), en genes involucrados en rutas que regulan la coagulación de la sangre.
5. En el clado de murciélagos nectarívoros de la subfamilia Glossophaginae los genes y enzimas bajo señales de selección positiva ($dN/dS > 1$; $pval = 0.001$) que participan en rutas metabólicas de carbohidratos y lípidos, sugiere que estos murciélagos metabolizan de una manera eficiente los azúcares que obtienen del néctar, y los escasos lípidos que obtienen del polen, para así cubrir los requerimientos energéticos necesarios que implica el vuelo estacionario y forrajeo, así como evitar altas concentraciones de azúcar en sangre. Asimismo, para este clado las duplicaciones en familias génicas involucradas en el almacenamiento de hierro, pueden estar relacionadas en prevenir desordenes metabólicos como la anemia, dado que la nectarivoría es una dieta con bajo consumo de hierro.
6. Una dieta omnívora-frugívora, fue un paso importante en la transición a la especialización de la nectarivoría en murciélagos filostómidos.
7. Los murciélagos con una dieta dominante en azúcares, provenientes del néctar o de frutas, han evolucionado mecanismos para eficientizar la conversión de grasas a

- energía, a partir de los escasos lípidos que obtienen del alimento. Las señales de evolución paralela exclusiva en los murciélagos nectarívoros del Nuevo Mundo y los murciélagos frugívoros del Viejo Mundo, es una consecuencia de la evolución de la dieta y las altas demandas energéticas que implica el forrajeo.
8. Los genes con señales de selección positiva identificados en el clado de los murciélagos nectarívoros son cruciales en el metabolismo del azúcar. En humanos se ha documentado que mutaciones en algunos de estos genes están asociados con desórdenes metabólicos, como es la diabetes, obesidad, anemia e hiperglucemia.
 9. La evolución de receptores olfativos y de sabor ha estado sujeta a recurrentes eventos de pseudogenización, principalmente en especies de murciélagos con un nicho alimenticio especialista, como es el caso del vampiro.
 10. La mayoría de los murciélagos insectívoros representantes de diferentes familias en el orden Quiróptera presentan la pérdida total o pseudogenización de los receptores de sabor de la familia *Tas1r*, lo cual implica que no tienen la capacidad de identificar sabores dulces, ni umami. En cambio, para estas especies insectívoras la percepción de los sabores amargos está sujeta a presiones de selección positiva y purificadora.
 11. La evolución a la nectarivoría en especies filogenéticamente distantes como son los murciélagos nectarívoros y colibríes involucra genes relevantes en el metabolismo de lípidos con una sobre-expresión transcripcional. Ambos grupos han evolucionado mecanismos que les permite compensar los requerimientos en el metabolismo de las grasas, dado una dieta baja en nutrientes protéicos y de lípidos.
 12. El vuelo estacionario y las estrategias de forrajeo implican una alta demanda energética, para la cual los colibríes y los murciélagos nectarívoros han evolucionado mecanismos involucrados en una eficiente conversión del azúcar en energía.

Perspectivas

Con los resultados obtenidos en esta tesis y las recientes investigaciones publicadas sobre la evolución de los genomas en murciélagos y otros mamíferos (Koepli et al., 2015), considero que en particular va a ser importante realizar futuros estudios que permitan entender la reducción de los genomas en los murciélagos, ya que dentro de los mamíferos los murciélagos son las especies con tamaño de genoma más pequeños. Esta reducción del genoma está relacionado con la disminución en el contenido de su DNA no repetitivo y con pérdidas de elementos repetitivos largos (Gregory et al., 2005). En diferentes especies de vertebrados, se ha sugerido que el tamaño de genoma es proporcional al contenido de elementos transponibles (Hanh et al., 2007), por lo que sugerimos que a partir de los 42 genomas de murciélagos disponibles en este momento en las bases de datos, e incluyendo los de otros mamíferos, se debe cuantificar la proporción de elementos transponibles, estimar el total que comprenden estos elementos en los genomas y hacer la clasificación de los elementos repetitivos. Tener la secuencia y clasificación de estos elementos permitirá diseñar sondas o primers para la amplificación de estos elementos en otras especies de murciélagos y con esta información realizar análisis de evolución molecular para detectar eventos de nacimiento y muerte (*birth/death*) de estas regiones del genoma, así como estimar finalmente las fechas de cuando ocurrieron estos eventos.

Para estudiar a mayor profundidad las transiciones y especializaciones de la dieta desde una perspectiva genómica y evolutiva, considero que futuros estudios deben incluir un mayor número de especies de murciélagos frugívoros de la familia Phyllostomidae, como pueden ser los géneros *Sturnira* y *Anoura*. Va a ser muy importante incorporar otras especies nectarívoras de la subfamilia Lonchophyllinae, como *Lonchophylla thomasi* y *Lyonycteris spurrelli*, pero también especies de murciélagos nectarívoros del Viejo Mundo de la familia Pteropodidae, como *Syconycteris australis* y *Eonycteris spelaea*. La información genómica de las especies mencionadas, nos permitirá analizar la evolución de la nectarivoría en dos subfamilias especialistas dentro de los Filostómidos y eventos de convergencia entre los murciélagos nectarívoros del Viejo y Nuevo Mundo. Esto nos permitirá estudiar con detalle la evolución a una dieta a base de frutas y semillas entre los frugívoros especialistas, y tener mayor información de las transiciones genómicas de una dieta frugívora a la nectarívora.

Para entender a mayor profundidad los diferentes mecanismos involucrados en la diversificación y especialización de las dietas recomendamos realizar estudios del microbioma intestinal en especies de murciélagos insectívoros (como *Macrotus waterhousii* o especies del género *Myotis*), omnívoros (como *Phyllostomus discolor*) y nectarívoros (como *L. yerbabuena* y *E. spelaea*). Proponemos utilizar herramientas de secuenciación con la plataforma Illumina, para obtener los metagenomas de diferentes secciones del intestino: duodeno, yeyuno e íleon. Previos estudios de microbioma en el vampiro han identificado comunidades bacterianas,

importantes en el metabolismo y obtención de nutrientes (Zepeda-Mendoza et al., 2018). Analizar el microbioma de murciélagos nectarívoros, permitirá identificar comunidades de bacterias que están participando en la obtención de nutrientes y que pueden ser importantes en romper azúcares como la trehalosa, la cual se encuentra en los exoesqueletos de insectos.

Para el caso del estudio de la nectarivoría en murciélagos y colibríes, proponemos realizar estudios de genómica funcional utilizando herramientas transcriptómicas, de *exon-capture* y de edición génica optimizada (Anamika et al., 2016). Esta información permitirá estudiar la regulación genes y otros mecanismos involucrados con esta dieta. Proponemos que este estudio genómico se realice en al menos tres especies de murciélagos nectarívoros, como *L. yerbabuenae*, *L. thomasi* y *E. spelaea*. En el caso de los colibríes las especies a estudiar pueden ser *Calypte anna*, *Archilochus colibrus*, y *Hylocharis leucottis*. Para la parte transcriptómica, sugerimos explorar las diferencias en la expresión de genes en condiciones de ingesta de néctar y sin proporcionar alimento. Para este experimento recomiendo tener al menos tres réplicas por cada muestra de tejido. Los tejidos importantes para analizar serán el páncreas, el riñón y el hígado (los cuales tienen un papel fundamental en rutas de carbohidratos y lípidos), el músculo (crucial en la ruta de la oxidación del azúcar) y diferentes secciones del intestino (duodeno, yeyuno e íleo). Esta información permitirá explorar la expresión de ciertos genes como la enzima Sac, que es importante en la transformación de disacáridos en azúcares simples. La secuenciación de los exomas de las especies de murciélagos y colibríes, permitirá identificar aquellas regiones que están bajo presiones de selección.

Con la información de los transcriptomas y los exomas, se podrán realizar análisis para detectar regiones *cis*-QTL y la expresión de alelos-específicos (llamados en inglés *allele-specific expression*). Estos análisis proporcionarán información puntual de los genes y las regiones reguladoras que están controlando los procesos del metabolismo del azúcar, entre otros. Conociendo los genes y regiones reguladoras (i.e., su secuencia y posición en los genomas), se podrían utilizar técnicas de edición genética como la introducción de genes utilizando las herramientas optimizadas (*Prime Editing Gene*, su nombre en inglés), como son: CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeat*) y las endonucleasas cas. El conocer la secuencia de genes y regiones reguladoras permitirá diseñar guías más específicas para la edición génica (CRISPR-cas9 o CRISPR-cas12). Estos experimentos se podrían realizar en ratones sometidos a una dieta con alta ingesta de azúcares (similando las concentraciones naturales que tiene el néctar) y se evaluaría su respuesta metabólica.

Estudios de genómica funcional como los enunciados en el párrafo anterior no permitirán validar los mecanismos de regulación del metabolismo del azúcar y proporcionarán información genética para posibles tratamientos de enfermedades en humanos, como hiperglucemia, diabetes u otros trastornos metabólicos.

Referencia citada

1. Adams, J. N. Raffield, L. M. Martelle, S. E. Genetic analysis of advanced glycation end products in the DHS MIND study. *Gene* 15, 173-179 (2016)
2. Alcorn, S. M. McGregor, S. E. Olin, G. Pollination of Saguaro Cactus by Doves, Nectar-Feeding Bats, and Honey Bees. *Science* 133, 1594–5 (1961).
3. Allendorf, F. W. Hohenlohe, P. A. Luikart, G. Genomics and the future of conservation genetics. *Nat Rev Gen* 11, 697–709 (2010).
4. Al-Samarai, F. R. Al-Kazaz, A. A. Molecular Markers: an Introduction and Applications. *J Mol Biotech* 9, 118–130 (2015).
5. Anamika, K. Verma, S. Jere, A. Desai, A. Transcriptomic Profiling Using Next Generation Sequencing - Advances Advantages, and Challenges. En: *Next Generation Sequencing - Advances, Applications and Challenges*, InTech US (2016). doi:10.5772/61789.
6. Anders, S. Huber, W. Differential expression analysis for sequence count data. *Genome Biol* 11, (2010).
7. Anisimova, M. Bielawski, JP. Yang, Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol.* 18:1585–1592 (2001).
8. Anthony-Regnitz, C. M. Wilson, A. E. Sweazea, K. L. Braun, E. J. Fewer exposed lysine residues may explain relative resistance of chicken serum albumin to in vitro protein glycation in comparison to bovine serum albumin. *J Mol Evol* 88, 653-661 (2020).
9. Arif, I. A. et al. DNA marker technology for wildlife conservation. *J Biol Sci* 18, 219–225 (2011).
10. Arizaga, S. et al. Pollination ecology of *Agave macroacantha* (Agavaceae) in a Mexican tropical desert. II. The role of pollinators. *Am J Bot* 87, 1011–7 (2000).
11. Arendt, J. Reznick, D. Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol* 23, 26–32 (2008).
12. Árnason, Ú. Lammers, F. Kumar, V. Nilsson, M. A. Janke, A. Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. *Sci Adv* 4, eaap9873 (2018).
13. Arrizabalaga-Escudero, A. et al. Trait-based functional dietary analysis provides a better insight into the foraging ecology of bats. *J Anim Ecol* 88, 1587–1600 (2019).

14. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. *Nat Gen* 25, 25–29 (2000).
15. Ayala-Berdon, J. et al. Foraging behavior adjustments related to changes in nectar sugar concentration in phyllostomid bats. *Comp Biochem Physiol A Mol Integr Physiol* 160, 143–8 (2011).
16. Ayala-Berdon, J. Schondube, J. E. A physiological perspective on nectar-feeding adaptation in phyllostomid bats. *Physiol Biochem Zool* 84, 458–66 (2011).
17. Baldwin, M. W. et al. Sensory biology. Evolution of sweet taste perception in hummingbirds by transformation of the ancestral umami receptor. *Science* 345, 929–33 (2014).
18. Berger, C. Zdzieblo, D. Glucose transporters in pancreatic islets. *Pflugers Arch* (2020).
19. Beuchat, C. A. Chong, C. R. Hyperglycemia in hummingbirds and its consequences for hemoglobin glycation. *Comp Biochem Physiol A Mol Integr Physiol* 120, 409–16 (1998).
20. Bian, Y. et al. Global downregulation of pigmentation associated genes in human premature hair graying. *Exp Ther Med* (2019) doi:10.3892/etm.2019.7663.
21. Bleiweiss, R. Tempo and mode of hummingbird evolution. *Biol J Linn Soc* 65, 63–76 (1998).
22. Bofkin, L. & Goldman, N. Variation in evolutionary processes at different codon positions. *Mol Biol Evol.* 24:513–521 (2007).
23. Botigué, L. R. et al. Ancient European dog genomes reveal continuity since the Early Neolithic. *Nat Commun* 8, (2017).
24. Boulton, R.A. Hematophagy. En: Vonk, J. Shackelford, T. *Encyclopedia of Animal Cognition and Behavior*. Springer, US (2018).
25. Braun, E. J. Sweazea, K. L. Glucose regulation in birds. *Comp Biochem Physiol B* 159, 1-9 (2008).
26. Brun, A. et al. Morphological bases for intestinal paracellular absorption in bats and rodents. *J Morphol* 280, 1359–1369 (2019).
27. Burin, G. et al. Omnivory in birds is a macroevolutionary sink. *Nat. Commun.* 7, 11250 (2016).

28. Calisher, C. et al. Bats: Important Reservoir Hosts of Emerging Viruses. *Cli Mic Rev* 19, 531–545 (2006).
29. Cantrell, M. et al. A. Loss of LINE-1 Activity in the Megabats. *Genetics* 178, 393–404 (2008).
30. Cao, Y. et al. A Review on the Applications of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies. *Front Microbiol* 8, (2017).
31. Carter, G. G. et al. Development of New Food-Sharing Relationships in Vampire Bats. *Curr Biol* 30, 1275–1279.e3 (2020).
32. Casillas, S. et al. PopHuman: the human population genomics browser. *Nucleic Acids Research* 46, D1003–D1010 (2017).
33. Cattin, M. F. et al. Phylogenetic constraints and adaptation explain food structure. *Nature* 427, 835–839 (2014).
34. Chen, C. W. Welch, K. C. Hummingbirds can fuel expensive hovering flight completely with either exogenous glucose or fructose. *Func Ecol* 28, 589–600 (2013).
35. Choo, S. W. et al. Pangolin genomes and the evolution of mammalian scales and immunity. *Genome Res* 26, 1312–1322 (2016).
36. Cohen, D. Chumakov, I. Weissenbach, J. A first-generation physical map of the human genome. *Nature* 366, 698–701 (1993).
37. Coleman, J. C. Downs, C. T. The sweet side of life: Nectar sugar type and concentration preference in Wahlbergs epauletted fruit bat. *Comp Bioch Physiol Part A: Molecular & Integrative Physiology* 162, 431–436 (2012).
38. Colosimo, P. F. Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science* 307, 1928–1933 (2005).
39. Compeau, P. E. C. Pevzner, P. A. Tesler, G. How to apply de Bruijn graphs to genome assembly. *Nat Bioch* 29, 987–991 (2011).
40. Daan, S. Masman, D. Groenewold, A. Avian basal metabolic rates: their association with body composition and energy expenditure in nature. *Am J Physiol Regulatory Integrative Comp Physiol* 259, 333-340 (1990).
41. Datzmann, T. Helversen, O. von. Mayer, F. Evolution of nectarivory in phyllostomid bats (Phyllostomidae Gray 1825, Chiroptera: Mammalia). *BMC Evolutionary Biology* 10, 165 (2010).

42. Dávalos, L. M. Cirranello, A. L. Geisler, J. H. Simmons, N. B. Understanding phylogenetic incongruence: lessons from phyllostomid bats. *Biol Rev Camb Philos Soc* 87, 991–1024 (2012).
43. De Bie, T. et al. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22, (2006).
44. Demuth, J. P. et al. The evolution of mammalian gene families. *PLoS ONE* 1:e85 (2006).
45. Demuth, P. Hahn, M. The life and death of gene families. *Bio Essays*. 31, 29-39 (2009).
46. Devuyst, O. The 1000 Genomes Project: Welcome to a New World. *Peritoneal Dialysis International: Journal of the International Society for Peritoneal Dialysis* 35, 676–677 (2015).
47. Dheilly, N. M. et al. No more non-model species: The promise of next generation sequencing for comparative immunology. *Developmental & Comparative Immunology* 45, 56–66 (2014).
48. Dong, D. et al. Evolution of Olfactory Receptor Genes in Primates Dominated by Birth-and-Death Process. *Genom Biol Evol.* 1:258-264 (2009).
49. dos Reis, M. & Wernisch, L. Estimating translational selection in eukaryotic genomes. *Mol Biol Evol.* 26:451–461 (2009).
50. Dumont, E. R. et al. Morphological innovation, diversification and invasion of a new adaptive zone. *Proc Biol Sci* 279, 1797–805 (2012).
51. Dupont, J. Tesseraud, S. Simon, J. Insulin signaling in chicken liver and muscle. *Gen Comp Endocrinol* 163, 52-57 (2008).
52. Eckalbar, W. L. et al. Transcriptomic and epigenomic characterization of the developing bat wing. *Nature Genetics* 48, 528–536 (2016).
53. Edwards, S. et al. The Maternal Gut Microbiome During Pregnancy. *MCN The American Journal of Maternal/Child Nursing* 1 (2017) doi:10.1097/nmc.0000000000000372.
54. Ehrlich, P. R. Raven, P. H. Butterflies and plants: a study in coevolution. *Evolution* 18, 586–608 (1964).
55. Ekblom, R. Galindo, J. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107, 1–15 (2010).
56. Ekblom, R. Wolf, J. B. A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl* 7, 1026–42 (2014).

57. Evans, T. Loose, M. AlignWise: a tool for identifying protein-coding sequence and correcting frame-shifts. *BMC Bioinformatics* 16, (2015).
58. Feinsinger, P. Approaches to nectatvore plant interactions in the New Wolrd. *Rev Chilena Hist Nat* 60, 285-319 (1987).
59. Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945 (2004).
60. Fleming, T. H. Geiselman, C. Kress, W. J. The evolution of bat pollination: a phylogenetic perspective. *Annals of Botany* 104, 1017–1043 (2009).
61. Fleming, T. H. Nuñez, R. A. Sternberg, L. D. Seasonal changes in the diets of migrant and non-migrant nectarivorous bats as revealed by carbon stable isotope analysis. *Oecologia* 94, 72–75 (1993).
62. Flores-Abreu, I. N. et al. Tempo and mode in coevolution of *Agave sensu lato* (Agavoideae Asparagaceae) and its bat pollinators, Glossophaginae (Phyllostomidae). *Mol Phyl Evol* 133, 176–188 (2019).
63. Foote, A. D. et al. Convergent evolution of the genomes of marine mammals. *Nature Genetics* 47, 272–275 (2015).
64. Frantz, L. A. et al. Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* 352, 1228–31 (2016).
65. Freeman, P. W. Nectarivorous feeding mechanisms in bats. *Biol J Linn Soc* 56, 439–463 (1995).
66. Frick, W. F. et al. Seasonal ecology of a migratory nectar-feeding bat at the edge of its range. *J Mammal* 99, 1072–1081 (2018).
67. Fritz, S.A. Bininda-Emonds, O. R. Purvis, A. Geographical variation in predictors of mammalian extinction risk: Big is bad, but only in the tropics. *Ecol Lett* 12, 538–549 (2009).
68. García-Vallejo, J. J. Gringhuis, S. I. van Dijk, W. van Die, Irma. Gene expression analysis of glycosylation-related by Real-Time Polymerase Chain Reaction. En: Brockhausen, I. *Methods in Molecular Biology: Glycobiology Protocols*. Human Press Inc, Totowa, NJ (2006).
69. Gentleman, R. C. et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5, R80 (2004).

70. Gilad, Y. et al. Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biol* 5:e148 (2007).
71. Gilbert, W. et al. Origin of genes. *Proc Natl Acad Sci* 94, 7698–7703 (1997).
72. Gillespie, J. H. *The Causes of Molecular Evolution*. Oxford University Press, UK (1991).
73. Gkogkolou, P. Böhm, M. Advanced glycation end products Key players in skin aging? *Dermato Endocrinol* 4, 259-270 (2012).
74. Gómez, J. M. Verdu, M. Perfectti, F. Ecological interactions are evolutionarily conserved across the entire tree of life. *Nature* 465, 918–921 (2010).
75. Gómez-Ruiz, E. P. Lacher, T. E. J. Climate change, range shifts, and the disruption of a pollinator-plant complex. *Sci Rep* 9, 14048 (2019).
76. Gonzalez-Terrazas, T. et al. Morphological specialization influences nectar extraction efficiency of sympatric nectar-feeding bats. *J Exp Biol* 215, 3989–96 (2012).
77. Gonzalez-Terrazas, T. P. et al. Finding flowers in the dark: nectar-feeding bats integrate olfaction and echolocation while foraging for nectar. *R Soc Open Sci* 3, 160199 (2016).
78. Gonzalez-Terrazas, T. P. et al. How Nectar-Feeding Bats Localize their Food: Echolocation Behavior of *Leptonycteris yerbabuenae* Approaching Cactus Flowers. *PLoS One* 11, e0163492 (2016).
79. Göttinger, T. et al. What Do Nectarivorous Bats Like? Nectar Composition in Bromeliaceae With Special Emphasis on Bat-Pollinated Species. *Frontiers in Plant Science* 10, (2019).
80. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech* 29, 644–652 (2011).
81. Grant, P. R. Grant, B. R. Petren, K. The allopatric phase of speciation: The sharp-beaked ground finch (*Geospiza difficilis*) on the Galápagos islands. *Biol J Linnean Soc* 69, 287–317 (2010).
82. Greenbaum, I. F. Phillips, C. J. Comparative anatomy and general histology of tongues of long-nosed bats (*Leptonycteris sanborni* and *L. nivalis*) with reference to infestation of oral mites. *J Mammal* 55, 489–504 (1974).
83. Gregory T. R. *Genome Size Evolution in Animals*. En: *The Evolution of the Genome* 3–87 Elsevier, (2005). doi:10.1016/b978-012301463-4/50003-6.

84. Gu, Z. et al. Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. *Mol Biol Evol* 19, 256–262 (2002).
85. Gutierrez, B. Escalera-Zamudio, M. Pybus, O. G. Parallel molecular evolution and adaptation in viruses. *Curr Opin Virol* 34, 90–96 (2019).
86. Gutiérrez-Guerrero, Y. T. et al. Genomic consequences of dietary diversification and parallel evolution due to nectarivory in leaf-nosed bats. *Gigascience* 9, (2020).
87. Hahn, M. et al. Estimating the tempo and mode of gene family from comparative genomic data. *Genomic Res* 15, 8 (2005).
88. Hahn, M. et al. Accelerated rate of gene gain and loss in primates. *Genetics*. 177, 3:1941-1949 (2007).
89. Havens, K. Vitt, P. The Importance of Phenological Diversity in Seed Mixes for Pollinator Restoration. *Nat Areas J* 36, 531 (2016).
90. Helversen, O. v. Reyer, H.-U. Nectar intake and energy expenditure in a flower visiting bat. *Oecologia* 63, 178–184 (1984).
91. Hernandez, A. Martínez del Río, C. Intestinal disaccharides in five species of phyllostomid bats. *Comp Bioch Physiol Part B: Comparative Biochemistry* 103, 105–111 (1992).
92. Ho, S. The molecular clock and estimating species divergence. *Nat Edu.* 1, 1:168 (2008).
93. Hong, W. Zhao, H. Vampire bats exhibit evolutionary reduction of bitter taste receptor genes common to other bats. *Proc Biol Sci* 281, 20141079 (2014).
94. Howell, D. J. Bats and pollen: Physiological aspects of the syndrome of chiropterophily. *Comp Bioch Physiol Part A: Physiology* 48, 263–276 (1974).
95. Hu, D. et al. Virome analysis for identification of novel mammalian viruses in bats from Southeast China. *Sci Rep* 7, (2017).
96. Huttley, G. et al. Rates of genome evolution and branching order from whole genome analysis. *Mol Biol Evol.* 24, 1722-1730 (2007).
97. Ingersoll, R. Haizmann, L. Lentink, D. Biomechanics of hover performance in Neotropical hummingbirds versus bats. *Sci Adv* 4, eaat2980 (2018).
98. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921 (2001).

99. Jebb, D. et al. Six new reference-quality bat genomes illuminate the molecular basis and evolution of bat adaptations. *Nature* 583, 578-584 (2020).
100. Jeffares, D. et al. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. *Methods Mol Biol.* 1201:65-90 (2015).
101. Jentjens, R. L. et al. Oxidation of combined ingestion of glucose and fructose during exercise. *J Appl Physiol* (1985) 96, 1277–84 (2004).
102. Jentjens, R. L. Venables, M. C. Jeukendrup, A. E. Oxidation of exogenous glucose, sucrose, and maltose during prolonged cycling exercise. *J Appl Physiol* (1985) 96, 1285–91 (2004).
103. Jiang, Y. K. et al. Upregulation Of Renal GLUT2 And SGLT2 Is Involved In High-Fat Diet-Induced Gestational Diabetes In Mice. *Diabetes Metab Syndr Obes* 12, 2095–2105 (2019).
104. Jiao, H. et al. Lineage-specific duplication and adaptive evolution of bitter taste receptor genes in bats. *Mol Ecol* 27, 4475–4488 (2018).
105. Jiao, H. et al. Trehalase Gene as a Molecular Signature of Dietary Diversification in Mammals. *Mol Biol Evol* 36, 2171–2183 (2019).
106. Jjones, G. Telling, E. The evolution of echolocation in bats. *Trends Ecol Evol* 21, 149–156 (2006).
107. Jones, K. E. Bininda-Emonds, O. R. Gittleman, J. L. Bats, clocks, and rocks: diversification patterns in Chiroptera. *Evolution* 59, 2243–55 (2005).
108. Jones, K. E. Safi, K. Ecology and evolution of mammalian biodiversity. *Philos T R Soc A B: Biological Sciences* 366, 2451–2461 (2011).
109. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240 (2014).
110. Karasov, W. H. Integrative physiology of transcellular and paracellular intestinal absorption. *J Exp Biol* 220, 2495–2501 (2017).
111. Kay, K. Sargent, R. D. The Role of Animal Pollination in Plant Speciation: Integrating Ecology Geography, and Genetics. *Annual Review of Ecology, Evolution, and Systematics* 40, 637–656 (2009).

112. Kelm, D. H. Simon, R. Kuhlow, D. et al. High activity enables life on a high-sugar diet: blood sugar regulation in nectar-feeding bats. *Proc Biol Sci* 7, 3490-6 (2011).
113. Kimura, M. The rate of molecular evolution considered from the standpoint of population genetics. *Proc Nat Acad Sci*. 63, 1181-1188 (1969).
114. Kimura, M. Evolutionary rate at the molecular level. *Nature*. 217, 624-626 (1968).
115. Kimura, M. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, UK (1983).
116. Kissling, W. D. Sekercioglu, C. H. Jetz, W. Bird dietary guild richness across latitudes, environments and biogeographic regions. *Global Ecol Biogeogr* 21, 328–340 (2012).
117. Koepfli, K. P. Paten, B. The Genome 10K Project: A Way Forward. *Annual Review of Animal Biosciences* 3, 57–111 (2015).
118. Kolkert, H. et al. Insectivorous bats selectively source moths and eat mostly pest insects on dryland and irrigated cotton farms. *Ecol Evol* 10, 371–388 (2020).
119. Kosiol, C. Holmes, I. Goldman, N. An empirical codon model for protein sequence evolution. *Mol Biol Evol*. 24:1464–1479 (2007).
120. Kosiol, C. et al. Patterns of positive selection in six Mammalian genomes. *PLoS Genet*. 4:e1000144 (2008).
121. Kosiol, C. & Anisimova, M. Selection acting on genomes. *Evol Genom*. 373-397 (2019).
122. Kosakovsky, P. & Muse, S. Site-to-site variation of synonymous substitution rates. *Mol Biol Evol*. 22:2375–2385 (2005).
123. Kumar, S. Subramanian, S. Mutation rates in mammalian genomes. *Proc Nat Acad Sci* 99, 803–808 (2002).
124. Kumar, S. Molecular clocks: Four decades of evolution. *Nat Rev Gen*. 6, 654-662 (2005).
125. Kumar, K. R. Cowley, M. J. Davis, R. L. Next-Generation Sequencing and Emerging Technologies. *Seminars in Thrombosis and Hemostasis* 45, 661–673 (2019).

126. Kunz, T. et al. Ecosystem services provided by bats. *Ann N Y Acad Sci* 1223, 1–38 (2011).
127. Landau, W. M. Liu, P. Dispersion Estimation and Its Effect on Test Performance in RNA-seq Data Analysis: A Simulation-Based Comparison of Methods. *PLoS ONE* 8, e81415 (2013).
128. Langmead, B. Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Met* 9, 357–359 (2012).
129. Laurindo, R. S. et al. Drivers of bat roles in Neotropical seed dispersal networks: abundance is more important than functional traits. *Oecologia* 193, 189–198 (2020).
130. Lee, J. H. et al. Molecular parallelism in fast-twitch muscle proteins in echolocating mammals. *Sci Adv* 4, eaat9660 (2018).
131. Leong, I. ALDOB promotes liver metastases development. *Nat Rev Endoc* 14, 380–380 (2018).
132. Leturque, A. et al. The role of GLUT2 in dietary sugar handling. *J Physiol Biochem* 61, 529–37 (2005).
133. Li, B. Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, (2011).
134. Liberles, D. Dittmar K. Characterizing gene family evolution. *Biol Proced.* 10, 1:66.73 (2008).
135. Licona-Vera, Y. Ornelas, J. F. The conquering of North America: dated phylogenetic and biogeographic inference of migratory behavior in bee hummingbirds. *BMC Evol Biol* 17, 126 (2017).
136. Lim, M. et al. Parallel Molecular Evolution in Pathways, Genes, and Sites in High-Elevation Hummingbirds Revealed by Comparative Transcriptomics. *Genome Biol Evol* 11, 1552–1572 (2019).
137. Losos, J. B. Review Convergence, adaptation, and constraint. *Evolution.* 65, 7 1827-40 (2011).
138. Lotz, Martínez del Río, C. Nicolson, S. W. Hummingbirds pay a high cost for a warm drink. *J Comp Physiol B: Biochemical Systemic, and Environmental Physiology* 173, 455–462 (2003).

139. Luo, Z. X. Transformation and diversification in early mammal evolution. *Nature* 450, 1011–1019 (2017).
140. Lynch, M. Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459–473 (2000).
141. Magi, A. et al. Bioinformatics for Next Generation Sequencing Data. *Genes* 1, 294–307 (2010).
142. Mahmudi, O. et al. Gene-pseudogene evolution: a probabilistic approach. *BMC Genomics*. 16, S12 (2015).
143. Mans, B. J. Louw, A. I. Neitz, A. W. Evolution of hematophagy in ticks: Common origins for blood coagulation and platelet aggregation inhibitors from soft ticks of the genus *Ornithodoros*. *Mol Biol Evol* 19, 1695–1705 (2002).
144. Martin, A. & Palumbi, S. Body size, metabolic rate, generation time, and the molecular clock. *Proc Nat Acad Sci*. 90, 4087-4091 (1993).
145. Martínez del Río, C. Sugar Preferences in Hummingbirds: The Influence of Subtle Chemical Differences on Food Choice. *The Condor* 92, 1022 (1990).
146. Martínez del Río, C, Baker, H. G. Baker, I. Ecological and evolutionary implications of digestive processes: Bird preferences and the sugar constituents of floral nectar and fruit pulp. *Experientia* 48, 544–551 (1992).
147. Martínez del Río, C. Gutiérrez-Guerrero, Y. T. An evolutionary remedy for an abominable physiological mystery: Benign hyperglycemia in birds. *J Mol Evol*, (2020).
148. McCarthy, D. J. Chen, Y. Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40, 4288–97 (2012).
149. McGuire, J. A. et al. Molecular Phylogenetics and the Diversification of Hummingbirds. *Curr Biol* 24, 910–916 (2014).
150. McMahon, B. J. Teeling, E. C. Höglund, J. How and why should we implement genomics into conservation? *Evol Appl* 7, 999–1007 (2014).
151. Meaburn, E. Schulz, R. Next generation sequencing in epigenetics: Insights and challenges. *Seminars in Cell & Developmental Biology* 23, 192–199 (2012).
152. Mendoza, M. et al. Hologenomic adaptations underlying the evolution of sanguivory in the common vampire bat. *Nat Ecol Evol* 2, 659–668 (2018).

153. Molina-Freaner, F. Eguiarte, L. E. The pollination biology of two paniculate agaves (Agavaceae) from northwestern Mexico: contrasting roles of bats as pollinators. *Am J Bot* 90, 1016–24 (2003).
154. Monteiro, L. R. Nogueira, M. R. Evolutionary patterns and processes in the radiation of phyllostomid bats. *BMC Evol Biol* 11, (2011).
155. Moreno-Santillán, D. et al. De Novo Transcriptome Assembly and Functional Annotation in Five Species of Bats. *Sci Rep* 9, (2019).
156. Muchhala, N. Adaptive trade-off in floral morphology mediates specialization for flowers pollinated by bats and hummingbirds. *Am Nat* 169, 494–504 (2007).
157. Muchhala, N. et al. A generalized pollination system in the tropics: bats, birds and *Aphelandra acanthus*. *Ann Bot* 103, 1481–7 (2009).
158. Muchhala, N. Thomson, J. D. Fur versus feathers: pollen delivery by bats and hummingbirds and consequences for pollen production. *Am Nat* 175, 717–26 (2010).
159. Muscarella, R. Fleming, T. H. The Role of Frugivorous Bats in Tropical Forest Succession. *Biol Rev* 82, 573–590 (2007).
160. Myrka, A. M. Welch, K. C. Evidence of high transport and phosphorylation capacity for both glucose and fructose in the ruby-throated hummingbird (*Archilochus colubris*). *Comp Bioch Physiol Part B: Biochemistry and Molecular Biology* 224, 253–261 (2018).
161. Nagarajan, S. R. Brandon, A. E. McKenna, J. A. et al. Insulin and diet-induced changes in the ubiquitin modified proteome of rat liver. *PLoS ONE* 12, e017431 (2011).
162. Nassar, J. M. et al. Seasonal sebaceous patch in the nectar-feeding bats *Leptonycteris curasoae* and *L. yerbabuena* (Phyllostomidae: Glossophaginae): phenological, histological, and preliminary chemical characterization. *Zoology (Jena)* 111, 363–76 (2008).
163. Nicolson, S. W. Fleming, P. A. Drinking problems on a simple diet: physiological convergence in nectar-feeding birds. *J Exp Biol* 217, 1015–1023 (2014).
164. O'Donnell, S. Logan, C. J. Clayton, N. S. Specializations of birds that attend army ant raids: an ecological approach to cognitive and behavioral studies. *Behav Processes* 91, 267–274 (2012).

165. Olalla-Tárraga, M. A. et al. Contrasting evidence of phylogenetic trophic niche conservatism in mammals worldwide. *J Biogeogr* 44, 99–110 (2010).
166. Ohta, T. Gene conversion and evolution of gene families: An overview. *Genes*. 10.3390/genes1030349 (2010).
167. Ohta, T. Evolution of gene families. *Gene* 2, 666-669 (2013).
168. Organ, C. L. Shedlock, A. M. Palaeogenomics of pterosaurs and the evolution of small genome size in flying vertebrates. *Biol Lett* 5, 47–50 (2009).
169. Orr, H. A. The probability of parallel evolution. *Evolution*. 59,1 216-20 (2005).
170. Ouborg, N. J. et al. Conservation genetics in transition to conservation genomics. *Trends Gen* 26, 177–187 (2010).
171. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* 578, 82–93 (2020).
172. Parker, J. et al. Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502, 228–231 (2013).
173. Patel, R. K. Jain, M. NGS-QC Toolkit: A Toolkit for Quality Control of Next Generation Sequencing Data. *PLoS ONE* 7, e30619 (2012).
174. Pavlovich, S. S. et al. The Egyptian Rousette Genome Reveals Unexpected Features of Bat Antiviral Immunity. *Cell* 173, 1098–1110.e18 (2018).
175. Peng, X . Xiangyang, H. Liu, Q. et al. Flight is the key to postprandial blood glucose balance in the fruit bats *Eonycteris spelaea* and *Cynopterus sphinx*. *Ecol Evol* 7, 8804-8811 (2017).
176. Pereira, M. A. et al. Application of Next-Generation Sequencing in the Era of Precision Medicine. En: *Applications of RNA-Seq and Omics Strategies - From Microorganisms to Human Health*, US (2017). doi:10.5772/intechopen.69337.
177. Petrov, D. & Hartl, D. Pseudogene evolution and natural selection for a compact genome. *J Heredity* 91, 221–227 (2000).
178. Pickersgill, B. Parallel vs Convergent evolution in domestication and diversification of crops in the Americas. *Front Ecol Evol* 6, 56 (2018).
179. Pirnay, F. et al. Exogenous glucose oxidation during exercise in relation to the power output. *Int J Sports Med* 16, 456–60 (1995).

180. Plassais, J. et al. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat Comm* 10, (2019).
181. Polakof, S. Mommsen, T. P. Soengas, J. L. Glucosensing and glucose homeostasis: From fish to mammals. *Comp Biochem Physiol B* 160, 123-149 (2011).
182. Poore, A. G. Ah Yong, S. T. Lowry, J. Sotka, E. Plant feeding promotes diversification in the Crustacea. *Proc Natl Acad Sci USA* 114, 8829–8834 (2017).
183. Powers, D. R. et al. Hovering in the heat: effects of environmental temperature on heat regulation in foraging hummingbirds. *R Soc Open Sci* 4, 171056 (2017).
184. Price, S. A. et al. Tempo of trophic evolution and its impact on mammalian diversification. *Proc Natl Acad Sci USA* 109, 7008–7012 (2012).
185. Rainford, J. L. Mayhew, P. J. Diet evolution and clade richness in hexapoda: A phylogenetic study of higher taxa. *Am Nat* 186, 777–791 (2015).
186. Rayner, J. The evolution of vertebrate flight. *Biol J Linn Soc* 34, 269–287 (1988).
187. Ripke, S. et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics* 45, 1150–1159 (2013).
188. Roberts, MW. Hummingbirds nectar concentration preferences at low volume: the importance of time scale. *Animal Behaviour* 52, 361–370 (1996).
189. Robinson, M. D. McCarthy, D. J. Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2009).
190. Rocca, M. A. Sazima, M. Beyog hummingbird-flowers: the other side of onitophyly in the neotropics. *Oecologia Australis* 14, 67–99 (2010).
191. Rocha, M. et al. Pollination Biology and Adaptive Radiation of Agavaceae with Special Emphasis on the Genus *Agave*. *Aliso* 22, 329–344 (2006).
192. Rodríguez-Oseguera, A. et al. Effect of habitat disturbance on pollination biology of the columnar cactus *Stenocereus quevedonis* at landscape-level in central Mexico. *Plant Biol* 15, 573–82 (2013).
193. Rodríguez-Peña, N. et al. Factors affecting nectar sugar composition in chiropterophilic plants. *Revista Mexicana de Biodiversidad* 87, 465–473 (2016).

194. Rodríguez-Peña, N. et al. Intestinal paracellular absorption is necessary to support the sugar oxidation cascade in nectarivorous bats. *J Exp Biol* 219, 779–82 (2016).
195. Rojas, D. Vale, A. Ferrero, V. Navarro, L. When did plants become important to leaf-nosed bats? Diversification of feeding habits in the family Phyllostomidae. *Mol Ecol* 20, 2217–28 (2011).
196. Román-Palacios, C. Scholl, J. P. Wiens, J. J. Evolution of diet across the animal tree of life. *Evol Lett* 3, 339–347 (2019).
197. Rondeau, P. Bourdon, E. The glycation of albumin: Structural and functional impacts. *Biochimie* 93, 645e658 (2010).
198. Rossoni, D. et al. A multiple peak adaptive landscape based on feeding strategies and roosting ecology shaped the evolution of cranial covariance structure and morphological differentiation in phyllostomid bats. *Evolution* 73, 961–981 (2019).
199. Sackton, T. B. Clark, N. Convergent evolution in the genomics era: new insights and directions. *Philos Trans R Soc Lond B Biol Sci.* 22, 374 (2019).
200. Sarich, V. Wilson, A. Generation time and genomic evolution in primates. *Science*. 179, 1144- 1147 (1969).
201. Schlötterer, C. The evolution of molecular markers just a matter of fashion?. *Nat Rev Gen* 5, 63–69 (2004).
202. Schmieder, R. Edwards, R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One* 6, e17288 (2011).
203. Schondube, J. E. Herrera-M, L. Martínez del Río, C. Diet and the evolution of digestion and renal function in phyllostomid bats. *Zoology* 104, 59–73 (2001).
204. Schutt, W. A. et al. *Dark Banquet: Blood and the curious lives of blood-feeding creatures.* Harmony Books, NY (2008).
205. Seim, I. et al. Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. *Nat Commun* 4, (2013).
206. Sendi, H. et al. Roach nectarivory, gymnosperm and earliest flower pollination evidence from Cretaceous ambers. *Biology* (2020).
207. Serrano-Serrano, M. et al. Hummingbird pollination and the diversification of angiosperms: an old and successful association in Gesneriaceae. *Proc R Soc B: Biological Sciences* 284, 20162816 (2017).

208. Sharp, P. et al. Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res.* 33:1141–1153 (2005).
209. Shen, B. et al. Adaptive evolution in the glucose transporter 4 gene *Slc2a4* in Old World fruit bats (family: Pteropodidae). *PLoS One* 7, e33197 (2012).
210. Simmons, N. B. Evolution. An Eocene big bang for bats. *Science* 307, 527–8 (2005).
211. Sitbon, G. Laurent, F. Mialhe, A. et al. Diabetes in birds. *Horm Metab Res* 12, 1-9 (1980).
212. Slatko, B. E. Gardner, A. F. Ausubel, F. M. Overview of Next-Generation Sequencing Technologies. *Curr P Mol Biol* 122, e59 (2018).
213. Smith, J. Gregory, T. R. The genome sizes of megabats (Chiroptera: Pteropodidae) are remarkably constrained. *Biol Lett* 5, 347–351 (2009).
214. Song, F. Schmidt, A. M. Glycation & Insulin Resistance: Novel mechanism unique and targets? *Arterioscler Thromb Vasc Biol* 32, 1760-1765 (2013).
215. Sorkheh, K. et al. Comparison of traditional and new generation DNA markers declares high genetic diversity and differentiated population structure of wild almond species. *Sci Rep* 7, (2017).
216. Stewart, CB. Evolution: Convergent and parallel evolution. *Encyclopedia of Life Sciences.* John Wile & Sons. (2006).
217. Strong, D. R. Lawton, J. H. Southwood, T. R. E. *Insects on Plants: Community Patterns and Mechanisms.* Harvard University (1984).
218. Suarez, R. Welch, K. C. The sugar oxidation cascade: aerial refueling in hummingbirds and nectar bats. *J Exp Biol* 214, 172–178 (2010).
219. Suarez, R. Welch, K. Sugar Metabolism in Hummingbirds and Nectar Bats. *Nutrients* 9, 743 (2017).
220. Sudbrak, R. et al. Comparative and evolutionary analysis of the Rhesus macaque extended MHC class II region. *Immunogenetics* 54(10), 699-704 (2003).
221. Sun, G. Dilcher, D. L. Zheng, S. Zhou, Z. In search of the first flower: a Jurassic angiosperm, *Archaeofructus*, from northeast China. *Science* 282, 1692–1695 (1988).

222. Szwegold, B. S. Miller, C. B. Potential of birds to serve as pathology-free models of type 2 diabetes, part 2: Do high levels of Carbonyl-Scavenging amino acids (e.g. Taurine) and low concentrations of Methylglyoxal limit the production of advanced glycation end products?. *Rejuvenation Res* 17, 347-358 (2014).
223. Szarski, H. Cell size and the concept of wasteful and frugal evolutionary strategies. *J Theor Biol* 105, 201–9 (1983).
224. Teeling, E. C. et al. Bat Biology Genomes, and the Bat1K Project: To Generate Chromosome-Level Genomes for All Living Bat Species. *Annual Review of Animal Biosciences* 6, 23–46 (2018).
225. Templeton, N. S. The Polymerase Chain Reaction History Methods and Applications. *J Diag Mol Path* 1, 58–72 (1992).
226. Tremlett, C. et al. Pollination by bats enhances both quality and yield of a major cash crop in Mexico. *J App Ecol* 57, 450–459 (2019).
227. Tsai, A. G. Lieber, M. R. Mechanisms of chromosomal rearrangement in the human genome. *BMC Genomics* 11, S1 (2010).
228. Tschapka, M. Dressler, S. Chiropterophily: On bat-flowers and flower-bats. *Botanical Magazine* 19, 114–125 (2002).
229. Turnbull, C. et al. The 100000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ* k1687 (2018) doi:10.1136/bmj.k1687.
230. Tyakht, A. V. et al. Human gut microbiota community structures in urban and rural populations in Russia. *Nat Commun* 4, (2013).
231. Umminger, B. L. Body size and whole blood sugar concentrations in mammals. *Comp Biochem Physiol* 52A, 455-458 (1975).
232. Valiente-Banuet, A. et al. Geographic differentiation in the pollination system of the columnar cactus *Pachycereus pecten -aboriginum*. *Am J Bot* 91, 850–5 (2004).
233. Von, B. R. Hedenström, A. Winter, Y. Johansson, L. C. Kinematics and wing shape across flight speed in the bat, *Leptonycteris yerbabuena*. *Biol Open* 1, 1226–38 (2012).
234. Wäckers, F. L. Romeis, J. Van Rijn, P. Nectar and pollen feeding by insect herbivores and implications for multitrophic interactions. *Annu Rev Ento* 52, 301-323 (2007).

235. Wang, B. Chen, L. Wang, W. Genomic insights into ruminant evolution: from past to future prospects. *Zool. Res.* 40, 476–487 (2019).
236. Warren, W. et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453, 175–183 (2008).
237. Welch, K. C. et al. Glucose Transporter Expression in an Avian Nectarivore: The Ruby-Throated Hummingbird (*Archilochus colubris*). *PLoS ONE* 8, e77003 (2013).
238. Welch, K. J. Chen, C. C. Sugar flux through the flight muscles of hovering vertebrate nectarivores: a review. *J Comp Physiol B* 184, 945–59 (2014).
239. Welch, K. Suarez, R. Oxidation of dietary sugar during hovering flight in small vertebrate nectarivores. *J Exp Biol* 214, 3324–3324 (2011).
240. Wiens, J. et al. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* 13, 1310–1324 (2010).
241. Wiens, J. Lapoint, R. Whiteman, N. Herbivory increases diversification across insect clades. *Nat. Commun.* 6, 8370 (2015).
242. Workman, R. E. et al. Single-molecule, full-length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird *Archilochus colubris*. *Gigascience* 7, 1–12 (2018).
243. Wu, J. et al. Testing the sensory trade-off hypothesis in New World bats. *Proc Biol Sci* 285, (2018).
244. Wu, Z. et al. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *ISME* 10, 609–620 (2015).
245. Yang, Z. Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol.* 17:32–43 (2000).
246. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591 (2007).
247. Yang, Z. Nielsen, R. Mutation-selection models of codon substitution and their use to estimate selective strengths on codon usage. *Mol Biol Evol.* 25:568–579 (2008).
248. Yang, J. et al. Genetic signatures of high-altitude adaptation in Tibetans. *Proc Natl Acad Sci U S A* 114, 4189–4194 (2017).

249. Yau, E. H. Rana, T. M. Next-Generation Sequencing of Genome-Wide CRISPR Screens. En: *Methods in Molecular Biology* 203–217, Springer New York, (2017). doi:10.1007/978-1-4939-7514-3_13.
250. Yohe, L. R. et al. Bayesian hierarchical models suggest oldest known plant-visiting bat was omnivorous. *Biol Lett* 11, 20150501 (2015).
251. Zhang, Z. et al. Millions of years of evolution preserved: A comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res* 13, 2541–2558 (2003).
252. Zhang, G. et al. Comparative Analysis of Bat Genomes Provides Insight into the Evolution of Flight and Immunity. *Science* 339, 456–460 (2012).
253. Zuckerkandl, E. & Pauling, L. Evolutionary divergence and convergence in proteins. En *Evolving Genes and Proteins*. Eds. Bryson, V. & Vogel, H. Academic Press. 97-166 New York (1965).

Tablas y Figuras

Figura 1. Filogenia que comprende los órdenes de la clase Mamalia, donde se muestra por orden, el número de especies registradas y el número de datos genómicos con NGS disponibles en bases de datos. Figura tomada y modificada de Koepfli et al., 2015.

Figura 2. Datos genómicos de mamíferos obtenidos con herramientas NGS, disponibles en la plataforma de *National Center for Biotechnology Information* (NCBI). a) Número de datos para mamíferos depositados en la base *Sequence Read Achievement* (SRA), con diferentes plataformas de secuenciación. b) Número de datos para mamíferos depositados en SRA para diferentes herramientas genómicas. c) y d) Número de genomas a nivel contig, scaffold y cromosoma para animales depositados en la base de NCBI. Información consultada 02-2020.

Figura 3. Genomas ensamblados y con anotación disponible en bases de datos de NCBI, ENSEMBL y GigaDatabase. * especies para las que hay información genómica y transcripcional.

Figura 4. Filogenia de la familia Phyllostomidae. a) Morfología craneal para cada subfamilia. b) Morfología del sistema renal para las subfamilias Macrochinae, Phyllostominae, Rhinophyllinae, Stenodermatinae, Glossophaginae y Desmodontinae. c) Morfología del sistema digestivo para insectívoros - omnívoros, frugívoros y sanguívoros. Figura tomada y modificada de Rossoni et al., 2017.

Capítulo I

Table 1. Global statistics for the nectar-pollen feeding bat *L. yerbabuena* genome assembly.

Table 2. Mapping statistics and SNP identification in Phyllostomid bat genomes (based on *L. yerbabuena* genome assembly).

Table 3. GO enrichment for significant gene families per node and habit food. Gene Ontologies (GO) annotations involved in metabolism and diet are in bold (+ Gene families expansions; - contractions).

Figure 1. Phyllostomid species. a) *Leptonycteris yerbabuena* b) *Leptonycteris nivalis*, c) *Musonycteris harrisoni*, d) *Macrotus waterhousii*, and e) *Artibeus jamaicensis*. Photo credits: a - b) Daniel Zamora-Mejías, c) Rodrigo Medellín-Legorreta, d) Melissa E. Rodríguez, and e) Wikimedia, public domain.

Figure 2. Phylogenetic tree constructed with 132 single copy genes and estimates of divergence times based on two fossil records (yellow stars) (see Methods). Based on 22,388 gene families we analyzed the number of orthologous families expanded (+ blue) and contracted (- red) across the phylogeny: per node (bars) and per species branch (right), with a *p-value* ≤ 0.01 . Gray bars reflect the divergence time interval based on 95% HPD.

Figure 3. Positive selection in genes and proteins across the phylogeny of Phyllostomid bats, in comparison to the insect-feeder bat *Macrotus waterhousii*. Most of positive selected genes likely

contribute to the regulation and processing of: a) muscle and bone development; b) carbohydrates; c) lipids; d) nutrients and food uptake; e) iron storage and calcium sources; and f) blood regulation (see genes and proteins abbreviations in Additional file 1, Table S14).

Figure 4. A subset of genes under positive selection (in red bold) that are involved in glucose and ketogen in the frugivorous (Aja: *A. jamaicensis*) and nectar-pollen bats (Mha: *M. harrisoni*, Lni: *L. nivalis*, and Lye: *L. yerbabuena*). The diagram also identifies adaptative signals for some genes in the vampire *D. rotundus* (Dro) and the insectivore *M. waterhousii* (Mwa). The diagram is based on the KEGG metabolic pathways database and a review of the literature (see gene and proteins abbreviations in Additional file 1, Table S8).

Figure 5. Parallel molecular evolution between Pteropodids (Old World) and Glossophagini (New World) bats, in three genes: Acetacetyl CoA Synthetase (AACCS); Alpha-Ketoglutarate-Dependent Dioxygenase Homolog 7, mitochondrial (*ALBKH7*); and UNC-45 homolog B, mitochondrial (*UNC-45 B*). a) Phylogeny reconstruction for these three genes by Maximum likelihood (using 1,827 amino acids), for 47 mammal species. b) Ancestral sequence reconstruction (for branches and nodes) to infer parallel substitutions in conserved positions for the three genes. c) Probability of replacement at each ancestral state node for each sequence position. Amino acid abbreviations: A – Alanine (non-polar); T – Threonine (polar); Q – Glutamine (polar); R – Arginine (basic-charged); K – Lysine (basic-charged); E – Glutamic acid (acidic + charged); S – Serine (polar) and L – Leucine (non-polar).

Figure 6. ACCS protein structure. a) 3D- structure of ACCS protein for *L. yerbabuena*. b) In gray: ACCS 3D- structure consensus (*M. waterhousii* and *D. rotundus*). In blue: alpha-helix structures shared only for the three Glossophagini nectar-pollen feeders (*M. harrisoni*, *L. nivalis* and *L. yerbabuena*). In yellow: beta-strand shared only between *M. harrisoni* and *P. alecto* (Pteropodid bat). RSMD score (protein 3D superposition and alignment) between pairs of species (see Additional file 1, Table S10).

Figure S1. a) Venn diagram showing the number of shared and unique *Single Nucleotide Variants* for each Phyllostomid bat based on *L. yerbabuena* genome. b) Nucleotide diversity estimated across sliding windows of 1 Mbp for each Phyllostomid, based on *L. yerbabuena* genome assembly.

Figure S2. Phylogeny constructed based on 132 single copy genes and calibrated using three fossil records.

Figure S3. Pipeline for the reference-genome assembly construction.

Figure S4. ACCS 3D protein structure for each bat species.

Capítulo II

Table 1. Ratio dN/dS and rate of molecular substitution for olfactory and taste gene receptors.

Figure 1. Functional genes and putative pseudogenes of olfactory receptor (OR) members identified in five dietary guilds.

Figure 2. Highly conserved olfactory and taste receptors under positive selection for each bat species.

Figure 3. Single-copy gene receptors: olfactory (a-b) and taste (c-d) that exhibited accelerations in their molecular substitution rates for specific bat lineages (red branch). a) *OR4D6*; b) *OR52N2*; c) *Tas1r3*; and d) *Tas2r40*.

Figure 4. Left. Phylogenetic tree of taste receptor 1 member 3 (*Tas1r3*) gene in mammals. Grey circles highlight the pseudogenization events. **Right.** *Tas1r3* sequence alignment for the Chiroptera species.

Supplementary Material

Supplementary Table 1. Priors used to calculate divergence times and absolute substitution rates for each taste and olfactory receptor.

Supplementary Table 2. Olfactory receptors located in non-coding regions.

Figure S1. Number of Olfactory receptors (Class I and Class II) and Taste receptors (Type I and Type II) identified in the genomic dataset available for all bat species.

Figure S2. Functional gene copies of 16 olfactory receptor families for each bat genome.

Capítulo III

Tabla 1. Rangos y tasas estimados de glicemia y relacionadas con los niveles de glucosa en aves y mamíferos.

Colibríes*; Murciélagos nectarívoros**

Tabla 2. Expresión diferencial y enriquecimiento de funciones de hígado (en gris: Molecular function; azul: Biological Process; blanco: Cellular Components).

Figura 1. Relación entre las concentraciones de glucosa en sangre reportada en aves y mamíferos en relación a su masa corporal (Figura tomada y modificada de Martínez del Río & Gutiérrez-Guerrero, 2020).

Figura 2. Enriquecimiento de funciones para aquellos transcritos con expresión diferencial en cada tejido. En gris, se presentan los GO de la categoría Biological Process; en azul, los GO de la categoría Molecular Function; y en blanco, los GO de la categoría Cellular Components. En negritas, son los transcritos que se expresan en ambas especies para cada tejido.

Figura 3. Rutas metabólicas del azúcar y grasas en hígado. En rojo, genes con expresión diferencial paralela en hígado entre ambos vertebrados nectarívoros.

Figura 4. Genes expresados activamente en la ruta de señalización de la Insulina en aves y murciélagos. Mapa de señalización de insulina para humanos (KEGG).

Tabla S1. Información de los datos de RNAseq de murciélagos y aves insectívoras utilizados de hígado.

Tabla S2. Lista de genes involucrados en procesos de glicación, reportados en literatura.

Tabla S3. Transcritos expresados diferencialmente en los cinco tejidos para el Murciélago Maguerero Menor.

Tabla S4. Transcritos expresados diferencialmente en los seis tejidos del colibrí de oreja blanca.

Tabla S5. Transcritos activamente expresados (on) y sin expresarse (off) en los murciélagos (*L. yerbabuenae* y *M. megalophylla*) y aves (*H. leucotis* y *F. albicollis*), involucrados en la ruta de señalización de insulina y resistencia a insulina.

Figura S1. *Heatmap* y agrupamiento de distancias del análisis de expresión diferencial de los transcritos sobre-expresados y regulados negativamente de cada tejido. A la derecha, se presentan los resultados cinco tejidos de *L. yerbabuenae*. A la izquierda, se muestran los resultados para los 6 tejidos de *H. leucotis*.

Figura S2. Gráfica *Multi-Dimensional Scaling*, que agrupa consistentemente las muestras de transcritos para cada especie y sus correspondientes réplicas.

Figura S3. Anotación funcional utilizando la base de datos KEGG, de los genes con una sobre-expresión en el hígado.

Figura S4. *Heatmap* de los genes involucrados en procesos de glicación y acumulación de AGEs identificados en los transcriptomas de hígados de los murciélagos y aves.

Material Suplementario

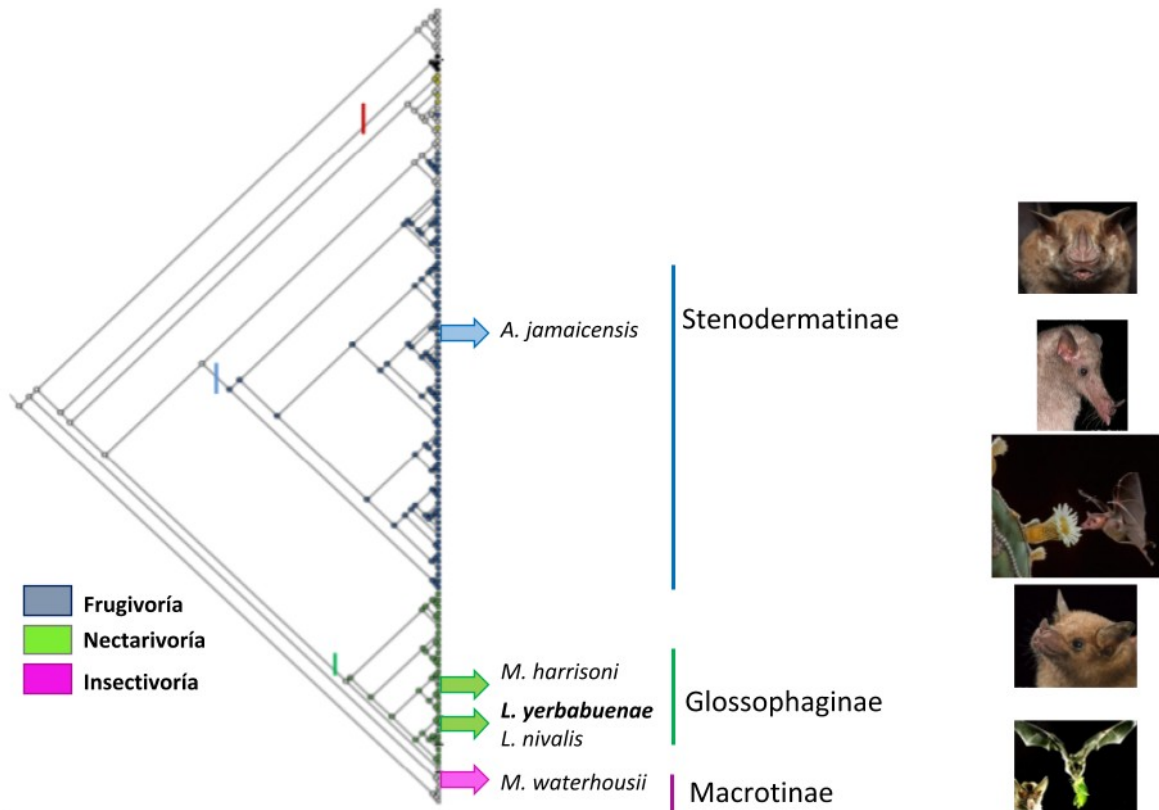


Figura S1. Especies de Filistómidos secuenciados en el proyecto de Fronteras de la Ciencia. Filogenia tomada y modificada de Trejo-Salazar, 2013. Fotos: Merlin Tuttle, Rodrigo Medellín, Fong G, Tschapka M.