Aus der Abteilung für Wildtierkrankheiten
im Leibniz-Institut für Zoo- und Wildtierforschung
im Forschungsverbund Berlin e.V.

eingereicht über den Fachbereich Veterinärmedizin der Freien Universität Berlin

The Vampire Bat Virome: Evolutionary Implications in an Immunological Context

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie (PhD)
in den biomedizinischen Wissenschaften
im Dahlem Research School (DRS) Programms
an der Freien Universität Berlin

vorgelegt von
M.Sc. Marina Escalera Zamudio
aus Mexiko Stadt, Mexiko

Berlin 2016
Journal-Nr.: 3880

Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Jürgen Zentek
Erster Gutachter:	Prof. Dr. Alex D. Greenwood
Zweiter Gutachter:	Prof. Dr. Ralph Tiedemann
Dritter Gutachter:	Univ.-Prof. Dr. Klaus Osterrieder

Deskriptoren (nach CAB-Thesaurus):
Chiroptera, Betaretrovirus, haematophagy, disease transmission, infections, Toll like-receptors (MeSH), mutations, evolution

Tag der Promotion: 07.09.2016

Coverbild © ArtThailand - Fotolia.com

"Nothing in Biology Makes Sense Except in the Light of Evolution"
(Theodosius Dobzhansky, 1973)

Dedicated to all my colleagues, friends and family.
Thank you for your support.

„Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes“

Supervisors:

Univ.-Prof. Dr. Univ.-Prof. Dr. Alex D. Greenwood
Department of Wildlife Diseases,
Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

Supervisor Committee:

Univ.-Prof. Dr. Klaus Osterrieder
Department of Veterinary Medicine, Institut für Virologie,
Freie Universität Berlin, Berlin, Germany

Univ.-Prof. Dr. Marion East
Department of Evolutionary Ecology,
Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

List of Content

Preface

This thesis is divided in two chapters, each one of them based on two published manuscripts. In the first chapter, the molecular characterization of a novel endogenous retrovirus ubiquitously present in the common vampire bat population and its evolutionary dynamics among different species is described. In the second chapter the molecular characterization of the nucleic acid sensing Toll-like receptors of the common vampire bat and the comparative evolutionary analysis of these receptors within eight different bat species and among other mammals is presented.

List of Abbreviations

Abbreviation	Full term
mya	Million years ago
ERVs	Endogenous retrovirus
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
OTU	Operational taxonomic unit
PSS	Positive selected site
TLRs	Toll-like receptors
ECD	Ectodomain
LRRs	Leucine-rich tandem repeat motifs
LBS	Ligand binding sites

General Introduction

The order Chiroptera: diversity and classification

The order Chiroptera is one of the most diverse mammalian groups with an estimate of 1,200 species and new bat species still being described (Gorfol *et al.* 2014; Koubinova *et al.* 2013; Simmons 2005). Bats are world-wide distributed animals living in different habitats including temperate and tropical forests, deserts and urban and suburban areas (Burns 2014; Simmons 2005). The adaptation of bats to different environments has resulted in the evolution of unique phenotypic and genotypic traits such as flight and echolocation and other behavioral or physiological characteristics like torpor and hibernation (Patterson *et al.* 2003; Simmons 2005). The bat lineage (Chiroptera) is classified within the Laurasiatheria superorder (including shrews, hedgehogs, pangolins, whales, most hoofed mammals and carnivores) and is considered a sister lineage to Cetartiodactylans (whales, dolphins and even-toed ungulates like pigs and cows) and more distantly related to Carnivores (canids, felids and seals) (Murphy *et al.* 2001; Nery *et al.* 2012; Zhou *et al.* 2012). Bats are divided into two major groups: megabats (Megachiroptera) and microbats (Microchiroptera). While most microbats use echolocation, megabats have well-developed vision and thus have in general lost the ability to use echolocation (Holland *et al.* 2004; Teeling *et al.* 2000). Moreover, megabats are mainly frugivorous whereas microbats display a wide range of dietary adaptations including insectivory, nectarivory and hematophagy (Datzmann *et al.* 2010; Simmons 2005). The Megachiroptera (also called Yinpterochiroptera) are considered to be a monophyletic group with just one family (*Pteropodidae*), while the Microchiroptera are a polyphyletic group divided in two major sub-clusters: the Yangochiroptera (*Emballonuridae, Furipteridae, Miniopteridae, Molossidae, Mormoopidae, Mystacinidae, Myzopodidae, Natalidae Noctilionidae, Phyllostomidae, Thyropteridae,* and *Vespertilionidae* families) and the Rhinolophoidea (*Craseonycteridae, Hipposideridae, Megadermatidae, Rhinolophidae,* and *Rhinopomatidae*) (Agnarsson *et al.* 2011; Jones *et al.* 2005). The estimated divergence dates for the chiropteran sub-lineages have shown that microbats separated from pteropodids at least 60 million years ago (mya), whereas the diversification of most Yangochiropteran families

occurred approximately 53 mya (Agnarsson *et al.* 2011; Jones *et al.* 2005). Among the Yangochiroptera are the *Phyllostomidae* or New World leaf-nosed bats, representing one of the most morphologically diverse groups adapted to a wide range of environments and diets (Burns 2014). The diversification of the *Phyllostomidae* family took place at least 14 mya giving place to 49 genera with more than 150 extant species that are distributed from the southern United Sates to northern Argentina (Agnarsson *et al.* 2011; Burns 2014; Jones *et al.* 2005; Simmons 2005).

The Desmodontinae subfamily

Among the *Phyllostomidae* are the vampire bats (*Desmodontinae* subfamily) with three species (*Desmodus rotundus, Diaemus youngi,* and *Diphylla ecaudata*) known to be the only mammals to feed exclusively on the blood of other animals. It is estimated that vampire bats diverged from an insectivorous ancestor evolutionarily recently around 5 mya from which the transition from insectivory to sanguivory required specific morphological, behavioral and physiological adaptation occurring in the *Desmodontinae* subfamily (Phillips & Baker 2015). Thus, vampire bat species exhibit a unique set of behavioral, physiological, and morphological characteristics distinct among all other bats (Davis *et al.* 2010; Greenhall *et al.* 1983; Simmons 2005). Some of the morphological characteristics of vampire bats include a clawed thumb on each wing used to climb onto their prey, well developed incisors and canines to draw blood during a bite and an extraordinary sense of smell due to unique nasal structures as well as good sight and hearing abilities for prey detection. Vampire bats also rely to some extent on echolocation despite enhanced hearing ability (Greenhall *et al.* 1983). In terms of behavior, vampire bats are social animals that live in big colonies consisting of a of couple hundred to thousands of individuals in which meal sharing and social grooming occurs frequently (Greenhall *et al.* 1983). Vampire bats are unique among other chiropterans as they are the only bat species that can still use terrestrial locomotion (Forment *et al.* 1971). Regarding the physiological traits unique to vampire bats, transcriptomic analysis of the *D. rotundus* salivary glands has revealed than they possess a unique array of anticoagulants, vasodilators, anti-inflammatory proteins, neural-disruptors and antimicrobial agents involved in hematophagy. However, more than 17% of the described proteins with specific functions related to sanguinivory are classified as unknown (Francischetti *et al.* 2013; Phillips & Baker 2015). It has been shown by sequencing analysis that *D. rotundus* has a diet preference for mammalian blood (particularly from domestic swine and bovids) whereas poultry is employed as a secondary food source, indicating that vampire bats may selectively feed on the blood of domesticated animals as they represent an easily accessible feeding source (Bobrowiec *et al.* 2015).

Bats as natural reservoirs for viruses

The rich diversity of bats is thought to provide valuable ecosystem services such as pest suppression, seed dispersal, pollination and nutrient distribution (Kunz *et al.* 2011). However, bats have also been increasingly recognized as reservoirs for different viruses, some of which are able to cross species barriers (Brook & Dobson 2015; Calisher *et al.* 2006; Hayman *et al.* 2013). Molecular and phylogenetic analyses suggest that some of these viruses have co-evolved with their bat hosts for a long time and that cross-species transmission events have been frequent throughout evolutionary history (Cui *et al.* 2012a; Drexler *et al.* 2012a; Drexler *et al.* 2012b; Zhuo *et al.* 2013). However, viral research studies have focused mainly on species from Asia, Australia, Europe and Africa and to a lesser extent on bats from the Americas, resulting in a biased representation of the overall viral diversity in bats. Moreover, because hemathophagy involves the exchange of body fluids like blood and saliva, it may represent a direct transmission route for different viruses thus making vampire bat species likely to carry viruses present in other mammals. The common vampire bat (*D. rotundus)* is known to be a reservoir for rabies-causing lyssaviruses and it is considered a major constraint on the cattle industry in Latin America since bat-transmitted rabies is a critical problem for livestock (Belotto *et al.* 2005). However, the presence of other viruses other than *Lyssavirus* has hardly been explored.

Retroviruses

Retroviruses are a diverse and widely distributed group of RNA viruses that are primarily transmitted via mucosal contact and through the exchange of body fluids like blood (Kurth & Bannert 2010). These viruses are able to reverse transcribe or copy their RNA genetic material into DNA and subsequently integrate the copies into the genome of their host. Retroviruses exist in both exogenous and endogenous forms and while exogenous retroviruses remain infectious and are transmitted horizontally, endogenous retroviruses (ERVs) may become fixed in the population and can be inherited as Mendelian traits (Kurth & Bannert 2010). Many exogenous retroviruses have been implicated in the development of

diseases whereas most endogenous retroviruses (ERVs) are considered inactive ‘genetic fossils’ (Kurth & Bannert 2010). However, some may still retain the ability to transcribe active viral elements or can become reactivated with potential health implications for the host (Engel & Hiebert 2010). The Retroviridae viral family can be divided into those with simple (alpha, beta, gamma and epsilon retroviruses) or complex genomes (lenti, delta and spumaviruses). Simple retroviral genomes contain by genes encoding the structural and functional polyproteins (GAG, PRO, POL and ENV) while complex retroviral genomes encode additional proteins and RNA’s with diverse virulence-enhancing functions (Kurth & Bannert 2010). ERVs are present in the genomes of all vertebrate species studied so far and thus are considered remnants of ancestral infections (Dewannieux & Heidmann 2013; Hayward *et al.* 2015a). Since different species can share endogenous retroviral sequences, it is assumed that an exogenous viral form infected the last common ancestors and became fixed in the genome before species divergence, or that they were later dispersed by spill-over events (Hayward *et al.* 2013b). It is widely accepted that retroviral cross-species transmissions that led to host-switching occurred frequently throughout evolutionary history, especially among the gamma- and betaretroviruses (being the most common retroviral groups among higher vertebrates) (Hayward *et al.* 2015b). Retroviruses in chiropterans are highly diverse and display a complex evolutionary history, potentially representing some of the oldest retroviral lineages among mammalian taxa (Cui *et al.* 2012a; Cui *et al.* 2012e; Hayward *et al.* 2013c; Zhuo *et al.* 2013). However, retroviral characterization in bats has been restricted to species from Eurasia, Africa and Australia and little is known about retroviruses in bats from the Neotropics.

The bat Toll-like receptors

Although bats can be persistently infected with many viruses, they rarely show clinical symptoms. It has been suggested that bats display a 'viral tolerance phenotype' as they might have evolved specific immune strategies to control viral replication (Baker *et al.* 2013). Studies on bat immunology have shown that the basic processes and signaling cascades appear to be overall the same as in other mammals, but bats may exhibit an increased reliance on adaptive immunity in order to compensate for a diminished antibody-mediated immunity during certain life stages (Baker *et al.* 2013; Brook & Dobson 2015).Toll like-receptors (TLRs) are a class of innate immune receptors that recognize a wide variety of pathogen-associated molecular patterns, and thus are considered to be the first-line defense against invading pathogens (Boehme & Compton 2004). The mammalian TLRs 3, 7, 8 and 9 play an important role in triggering acquired immunity as they are activated by nucleic acid ligands such as unmethylated DNA, double-stranded RNA and single-stranded RNA (Boehme & Compton 2004). TLRs are type-I integral membrane proteins that have an N-terminal ectodomain (ECD) with several leucine-rich tandem repeat motifs (LRR) where ligand biding sites (LBS) occur. The ECD is followed by a highly conserved transmembrane domain and a cytoplasmic Toll/interleukin-1 receptor signaling motif responsible for initiating the signaling cascade that leads to immune gene expression after their activation (Boehme & Compton 2004). TLRs receptors are of interest from an evolutionary point of view since there is evidence that the ligand recognition properties of these receptors may vary among different species, thereby having an impact on the evolutionary ecology of infectious diseases (Tschirren *et al.* 2012; Werling *et al.* 2009). Therefore, the analysis of the immune variation at a molecular level has reveal patterns of resistance or susceptibility to pathogens within different species and at different taxonomic levels (Alcaide & Edwards 2011; Schroder & Schumann 2005; Tschirren *et al.* 2012; Wlasiuk & Nachman 2010). Nonetheless, the study of the genetic variability of the immune system in bats has been restricted to a few species and to a few genes, in part due to the lack of sequences available for comparative analyses (Baker *et al.* 2013; Cowled *et al.* 2011; He *et al.* 2010; Iha *et al.* 2010; Kepler *et al.* 2010; Omatsu *et al.* 2008; Sarkar & Chakravarty 1991).

Molecular evolution analysis: phylodynamics

A *phylogenetic tree* is a branching diagram that represents a statistically supported hypothesis on the evolutionary history (or *phylogeny*) of a group of operational taxonomic units (OTUs). OTU's can be defined as a number of individuals from the same species, a groups of species, groups of taxa or of molecular data such as sequences (Penny *et al.* 1991). Phylogenetic trees are built based on the presence or absence of traits (when comparing taxonomic groups) or in the similarities and differences within coding DNA or protein sequences (when comparing molecular data) that are weighed under different methods with an underlying mathematical model that describes the evolution of the observed characters (Penny *et al.* 1991). Molecular evolution analysis is often based on the alignment of orthologous sequences followed by a site-by-site comparison in order to infer the underlying evolutionary pattern that is then represented by a phylogenetic tree. The effects of positive selection can be further drawn from both the alignments and trees under the assumption that nucleotide substitutions within coding sequences can be either synonymous (not resulting in an amino acid change) or non-synonymous (involving an amino acid change). Usually, most non-synonymous changes are thought to be eliminated by purifying selection over time, but can be fixed in a population if these are favored under Darwinian positive selection (Nei & Jin 1989). Thus, calculating the ratio of non-synonymous to synonymous substitutions in a group of sequences may be helpful in determining if selection has occurred. Sequences evolving neutrally or under purifying selection are expected to have a ratio of non-synonymous substitutions per non-synonymous sites (Cadieux *et al.*) to synonymous substitutions per synonymous sites (dS) below one ($dN/dS = \omega \leq 1$), while significant deviations towards a value above one may be interpreted as evidence for positive selection ($dN/dS = \omega > 1$) (Pond & Muse 2005). The detection of particular lineages evolving under positive selection and lineage-specific positive selected codons or sites (PSS) can be further done by using the site-specific, branch-specific and branch-site models that allow for different evolutionary scenarios (purifying selection, neutral evolution or positive selection) to act within specific sites of an alignment or branches of a phylogenetic tree. The 'robustness' of each model in terms of how well it fits the data analyzed and the comparison among models is statistical

weighed under a maximum likelihood framework (Kosakovsky Pond *et al.* 2011; Pond & Muse 2005; Yang & dos Reis 2011).

The shape of a phylogenetic tree can be influenced by changes in the population size and in the spatio-temporal dynamics of the OTUs analyzed. Changes in a phylogenetic pattern over time are easily observed for fast evolving biological entities such as viruses, which have a higher evolutionary rate (defined as the rate of divergence between taxonomic groups measurable as amino acid or nucleotide substitutions over a period of time such as per million years) and short generation times, thus affecting the branch lengths of the tree (Frost & Volz 2013; Grenfell *et al.* 2004). *Phylodynamics* is an area of molecular evolution analysis that aims to describe how epidemiological, immunological, and evolutionary processes potentially interact to shape a pathogen's phylogeny, and often use a combination of phylogenetics, epidemiology, population genetics and molecular immunology data to associate both host and pathogen evolutionary processes (Frost & Volz 2013; Volz *et al.* 2013). *Phylodynamic*s can also estimate the effects of immune selection, host specificity or changes in the host population structure on the evolutionary pattern of a pathogen. Such occurrences can be drawn from the relative branch length and clustering patterns observed in the tree (Frost & Volz 2013; Grenfell *et al.* 2004; Volz *et al.* 2013). Molecular evolutionary analyses are also useful for estimating pathogen emergence dates and for detecting recombination among different strains (Frost & Volz 2013; Magiorkinis *et al.* 2013). Moreover, the comparison of both host and pathogen phylogenies is useful for determining host-pathogen co-divergence patterns (*e.g.* for associating the genetic relatedness of a group of pathogens to specific hosts) (Sironi *et al.* 2015).

Thesis outline

Hypothesis and study aims

The common vampire bat (*D. rotundus*) likely carries viruses common to other mammals but not yet described for this species, as its adaptation to hematophagy could have resulted in viral spill-over events among taxa throughout evolutionary history. Given that retroviruses are primarily transmitted via body fluid exchange, we postulate that these viruses were particularly prone to jump between vampire bats and other taxa. Thus, we aimed to detect novel and known retroviruses in sample populations of common vampire bats collected in different localities in Mexico and by using a molecular evolutionary approach we searched for evidence of historic cross-species transmission events among retroviruses in vampire bats and other taxa.

Given the unique adaptations within the Chiroptera, we predict that bats as a taxonomic group have acquired distinctive mutations fixed within the nucleic-acid sensing TLRs with potential consequences on the functional properties of these proteins. We characterized the nucleic acid sensing TLRs (3, 7, 8 and 9) of the common vampire bat (*D. rotundus*) and compared the genetic variation of these receptors among eight different bats species belonging to three different families (*Pteropodidae, Vespertilionidae* and *Phyllostomidae*) and among other mammals. In order to detect the evolutionary forces acting upon the bat TLRs over both a short and long timescales, we further tested if ongoing and episodic diversifying positive selection had acted upon the bat TLRs.

Animal samples

The use of all animal samples was approved by the Internal Committee for Ethics and Animal Welfare (No. approval: 2012-09-05). For the project purposes, two vampire bat individuals (*D. rotundus)* were donated by Berlin Zoological Garden in Berlin Germany while thirty-two free-ranging bat individuals from the same species were captured using 'mist' nets from different localities in Mexico (Soledad Doblado, Veracruz; Cotaxtla, Veracruz; Tuxpan, Veracruz, Mexico Paso del Toro, Veracruz; San Pablo Tlaltizapan, Morelos; San Luis Potosí; San Luis Potosí and Estado de Mexico) (Table 1). All free-ranging bats were collected meeting Mexican regulations under a sample collection permit and export certificate (Num/SGPA/DGVS: 03173/14; SAGARPA: 241111524599811488A467371). Animals were sacrificed with chloroform and/or submitted to necropsy for organ extraction (lung, spleen, stomach and intestines) stored in RNAlater at -80°C for further use. Because sampling was dependent on the seasonality of bats, we were only able to obtain a limited number of individuals from the different localities. Given the timing of the project, the captive cat samples were processed and sequenced before the free-ranging bat samples.

Table 1

Species, location, number of individuals and sample types collected

Species/Origin	Sample Type	No. Individuals
DRZ1[1]	Lung, Spleen, Intestine, Salivary gland	2
DR MOR[2]	Oral and rectal swabs, stomach, Spleen, Lung	7
DR EDO[3]	Stomach, Spleen, Lung	3
DR SD[4]	Oral and rectal swabs, Spleen	22

[1] *Desmodus rotundus*; Berlin Zoologischer Garten, Berlin, Germany

[2] *Desmodus rotundus*; San Pablo Tlaltizapan, Morelos, México

[3] *Desmodus rotundus*; Estado de Mexico, México

[4] *Desmodus rotundus*; Soledad Doblado, Veracruz, México

Collaborators

Supervisor: Dr. Prof. Alex D. Greenwood. Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany.

PhD Student: M.Sc Marina Escalera Zamudio. Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany.

Sample collection:

- Elizabeth Loza-Rubio. Centro Nacional de Investigación Disciplinaria en Microbiología Animal-INIFAP, Mexico.
- Edith Rojas-Anaya. Centro Nacional de Investigación Disciplinaria en Microbiología Animal-INIFAP, Mexico.
- Maria L. Méndez-Ojeda. Facultad de Medicina Veterinaria y Zootecnia, Universidad Veracruzana, Veracruz, Mexico.

Sample processing and exportation:

- Carlos F. Arias. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca Morelos, Mexico.
- Gábor Á. Czirják. Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany.

Bioinformatic analysis:

- Marie-Lisandra Zepeda Mendoza. Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark.
- Blanca Taboada. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca Morelos, Mexico.
- Felix Heeger. Berlin Center for Genomics in Biodiversity Research. Berlin, Germany.

A.D.G and M.E.Z designed research; M.E.Z performed research, M.E.Z and M.L.Z.M analyzed data; E.L.R, E.R.A, M.L.M.O provided samples while G.A.C and C.F.A contributed to sample processing. M.E.Z and A.D.G wrote the papers, with comments from all authors.

Summary of Publications

Publication 1: Escalera-Zamudio M, Mendoza MLZ, Heeger F, Loza-Rubio E, Rojas-Anaya E, Méndez-Ojeda ML, Taboada B, Mazzoni CJ, Arias CF, Greenwood AD. A novel endogenous betaretrovirus in the common vampire bat *(Desmodus rotundus)* suggests multiple independent infection and cross-species transmission events. J Virol. 2015 May;89(9):5180-4. doi: 10.1128/JVI.03452-14. Epub 2015 Feb 25

The common vampire bat (*D. rotundus*) is a known reservoir for viruses of medical importance and a potential host to other viruses because of its exclusive adaptation to hematophagy. As retroviruses are frequently blood transmitted, vampire bats could have been particularly prone to acquiring and transmitting retroviruses through evolutionary history; however, little is known about retroviruses in Neotropical bats. A combined high-throughput sequencing and PCR approach was used to look for novel retroviruses in *D. rotundus* and to search for evidence of historical cross-species transmission events. A novel vampire bat endogenous betaretrovirus (DrERV) fixed in the *D. rotundus* population was discovered. However, DrERV was not found to be ubiquitously present in other closely related bat species suggesting independent introduction events within the *Phyllostomidae* bat family. Moreover, DrERV is not phylogenetically related to other bat betaretroviruses and forms part of a recombinant lineage related to retroviruses from rodents and primates suggesting that multiple and evolutionary recent cross-species transmission events have taken place. Thus, the distribution of this lineage is not constrained by species boundaries and the young integration time estimated for some of its members indicates that these viruses are likely to be present in a wide range of New World species, with the possibility of an exogenous counterpart recently or currently active circulating in Latin America.

http://dx.doi.org/10.1128/JVI.03452-14

Publication 2: Escalera-Zamudio M, Mendoza MLZ, Loza-Rubio E, Rojas-Anaya E, Méndez-Ojeda ML, Arias CF and Greenwood AD. The evolution of bat nucleic acid sensing Toll-like receptors. Mol Ecol. 2015 Oct 26. doi: 10.1111/mec.13431. [Epub ahead of print]

The characterization of the nucleic acid sensing Toll-like receptors (TLR) of a New World bat species the common vampire bat (*D. rotundus*) was done, and through a comparative molecular evolutionary approach, the general adaptation patterns among the nucleic acid sensing TLRs of eight different bats species belonging to three families (*Pteropodidae, Vespertilionidae* and *Phyllostomidae*) were investigated. In general, the bat TLRs were found to be evolving slowly and mostly under purifying selection as other mammalian nuclear genes, with the divergence pattern of such receptors being overall congruent with the species tree. However, the chiropteran TLRs exhibited unique mutations fixed in ligand binding sites, some of which involved non-conservative amino acid changes and/or were detected to be targets of positive selection. Such changes could potentially modify protein function and ligand biding properties, as they were predicted to alter nucleic acid binding motifs in TLR 9. Moreover, evidence for episodic diversifying selection acting specifically upon the bat lineage and sub lineages was detected. Thus, the long-term adaptation of chiropterans to a wide variety of environments and ecological niches with different pathogen profiles is likely to have shaped the evolution of the bat TLRs at different taxonomic levels. The observed evolutionary patterns provide evidence for potential functional differences between bat and other mammalian TLRs in terms of resistance to specific pathogens or recognition of nucleic acids in general.

http://dx.doi.org//10.1111/mec.13431

Chapter 1: A novel endogenous betaretrovirus in the common vampire bat *(Desmodus rotundus)* suggests multiple independent infection and cross-species transmission events

Marina Escalera-Zamudio[a]#*, M. Lisandra Zepeda Mendoza[b], Felix Heeger[c], Elizabeth Loza-Rubio[d], Edith Rojas-Anaya[d], Maria L. Méndez-Ojeda[e], Blanca Taboada[f], Camila J. Mazzoni[a,c], Carlos F. Arias [f] and Alex D. Greenwood [a]#

Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany[a]; Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark[b]; Berlin Center for Genomics in Biodiversity Research. Berlin, Germany[c]; Centro Nacional de Investigación Disciplinaria en Microbiología Animal-INIFAP, Mexico[d]; Facultad de Medicina Veterinaria y Zootecnia, Universidad Veracruzana, Veracruz, Mexico[e]; Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca Morelos, Mexico[f]

Running title: Novel betaretrovirus in *D. rotundus*

Address correspondence to Alex D. Greenwood, greenwood@izw-berlin.de and Marina Escalera-Zamudio, escalera@izw-berlin.de

* Present address: Marina Escalera-Zamudio, Leibniz-Institut für Zoo- und Wildtierforschung (IZW). Alfred-Kowalke-Straße 17, 10315 Berlin, Germany

ABSTRACT

The *Desmodus rotundus* endogenous betaretrovirus (DrERV) is fixed in the vampire bat *D. rotundus* population and in other phyllostomid bats, but is not present in all species from this family. DrERV is not phylogenetically related to Old World bat betaretroviruses, but to betaretroviruses from rodents and New World primates suggesting recent cross-species transmission. A recent integration age estimation of the provirus in some taxa indicates that an exogenous counterpart might have been in circulation.

The common vampire bat (*Desmodus rotundus*), is a phyllostomid bat species with a broad geographical distribution and lives in close proximity with humans and domestic animals (Lee *et al.* 2012). Recently described retroviruses in chiropterans are diverse, some potentially representing the oldest viral lineages in mammalian taxa (Cui *et al.* 2012a; Cui *et al.* 2012b; Hayward *et al.* 2013a; Zhuo *et al.* 2013). However, retroviral characterization has been restricted to bat species distributed in Eurasia, Africa and Australia (Cui *et al.* 2012a; Hayward *et al.* 2013a; Zhuo *et al.* 2013). Little is known about retroviruses in bats from the Neotropics and nothing about those in vampire bats. Endogenous retroviruses (ERVs) are present in the genomes of all vertebrates examined (Anai *et al.* 2012; Baillie *et al.* 2004; Baillie & Wilkins 2001; Chiu *et al.* 1983; Cui *et al.* 2012a; Cui *et al.* 2012b; Hayward *et al.* 2013b; Mang *et al.* 1999; Mayer *et al.* 2013; Roca *et al.* 2004; van der Kuyl *et al.* 1999; Zhuo *et al.* 2013). As different species may share ERV sequences, it is assumed that in many cases, an exogenous retrovirus infected the common ancestor of multiple species and became fixed in the genome prior to species divergence (Belshaw *et al.* 2004; Bonnaud *et al.* 2005; Hayward *et al.* 2013b). Most ERVs are inactive '*genetic fossils*', whereas some may still retain the ability to transcribe active elements or can become reactivated having potential health implications for the host (Anai *et al.* 2012; Engel & Hiebert 2010; Ruprecht *et al.* 2008; Wootton *et al.* 2005; Yi *et al.* 2004). As retroviruses are primarily transmitted via blood to blood contact, we postulate that vampire bat retroviruses are particularly prone to jumping from one species to another. In this study we characterized an endogenous betaretrovirus present in *D. rotundus* and searched for evidence indicating cross species transmission events within its evolutionary history.

DrERV is a Type D endogenous betaretrovirus

Genomic Illumina MiSeq shotgun sequencing from a population of free-ranging and captive *D. rotundus* bats from Mexico and the Berlin Zoological Garden (Dataset S1) revealed the presence of a novel retrovirus, designated here DrERV (*Desmodus rotundus* endogenous retrovirus). Read assignment analysis showed that DrERV is homologous to SMRV, a type D retrovirus found in the New World squirrel monkeys (*Saimiri* genus) (Colcher *et al.* 1977), sharing a global percentage nucleotide similarity of 72% (E-value: 0.0) as determined by blast. To retrieve the complete DrERV full genome and integration sites, reads were assembled against the SMRV genome through a combined mapping and *de novo* assembly approach using Bowtie version 2 .2.2, BWA version 0.7.9 and Velvet 1.2.10 against the squirrel monkey retrovirus genome to build a consensus (SMRV-H; accession number: M23385) (Langmead & Salzberg 2012; Li & Durbin 2010; Zerbino 2010) and sequence gaps were covered by PCR (Dataset S2). The DrERV full-genome was confirmed by PCR and Sanger sequencing, while integration site variability and

copy number was determined by mapping the DrERV consensus to *D. rotundus* genomic sequences available for this work.

DrERV is a low-copy provirus that has typical Type D betaretroviral structural characteristics (Chiu *et al* 1983; Elder *et al.* 1992). The pol and env retroviral core elements have several stop codons and are thus expected to yield truncated gene products. However, the gag and protease ORFs are intact, suggesting they might code for functional proteins. The gag ORF contains a predicted zinc knuckle (zf-CCHC; pfam00098) and a zinc-finger domain (zf-CCHC_5; pfam14787), while the protease region has a trimeric dUTP diphosphatase domain (dUTPase; cd07557) (Fig. 1A). The pol gene has three ORFs coding for a ribonuclease H (RNase H; cd09273), a transposase (Tra5; COG2801) and a Jaagsiekte sheep retrovirus-like ORF (Orf-X; cl04426). The env gene has two ORFs coding for a coat protein (pfam00429) and a transmembrane subunit (cd09851) (Fig. 1A). The core viral elements are flanked by LTRs approximately 360 bp-long, with the 5'LTR having an intact primer binding site (PBS) complementary to tRNALys[1,2]. We detected four DrERV copies (DrERV_824, DrERV_216, DrERV_479 and DrERV_C53), with variants DrERV_216 and DrERV_824 being complete proviruses but with substantially high overall genetic divergence in the env and LTR regions (0.039 in Tamura-Nei distance), suggesting that they might have arisen from independent endogenization events or underwent recombination. In contrast, DrERV_479 and DrERV_C53 are truncated but share >99% similarity to the consensus, indicating that they likely arose via recent gene duplication (Fig. A2).

DrERV-related retroviruses detected in other bat and non-bat species

We identified DrERV-related sequences in another phyllostomid bat genome (*Carollia perspicillata),* corresponding to the betaretrovirus CpERV-β5_AC138156 described by Baillie et al., 2004 (Baillie *et al.* 2004) by sequential blastn and tblastx mining of the DrERV consensus sequence against the mammalian nucleotide collection and reference genomic sequence NCBI databases (NCBI 2015). DrERV shares a global nucleotide similarity of 75% to CpERV (E-value: 0.0), and an amino acid similarity of 65% in Gag and 84% in Env. CpERV has a complete deletion of the pol and pro region and therefore could not be analyzed. DrERV homology was also detected within the common brown rat genome (*Rattus norvegicus*), matching the betaretrovirus RnERV_AC243170 described by Baillie *et al.* (Baillie *et al.* 2004). DrERV shares a global nucleotide similarity to RnERV of 67% (E-value: 6e-108) and up to 40% in Gag, 50% in Protease, 67% in Pol, and 43% in transmembrane subunit of Env amino acid sequences. However, the Env coat protein of DrERV and RnERV were not homologous, and further PSI-blastp analysis revealed that this region in DrERV is similar to the reticuloendotheliosis gammaretrovirus (REV)

Env suggesting that recombination has occurred.

We evaluated the presence of DrERV sequences by PCR in two other phyllostomid species including another vampire bat (*Diphylla ecaudata)* and a fruit-eating species (*Artibeus jamaicensis*) (Dataset S2). DrERV sequences could not be detected in either species (data not shown). The absence of DrERV in *D. ecaudata* but presence in *D. rotundus,* two species from the same subfamily *(Desmodontinae*) (Datzmann *et al.* 2010), suggests that the DrERV invasion of phyllostomid bats occurred via independent infections subsequent to species divergence and is supported by the same pattern observed in the *A. jamaicensis* and *C. perspicillata* species both belonging to the subfamily *Carolliinae* (Datzmann *et al.* 2010). However, the presence of DrERV-related sequences in other Phyllostomid bats cannot be ruled-out, as our sample size is small and overall represents less than 2% of the diversity found within this group.

DrERV is related to primate and rodent betaretroviruses

We inferred phylogenetic trees from the Gag, Pol and Env amino acid sequence alignments (Edgar 2004; Gouy *et al.* 2010) of several mammalian ERVs (Dataset S3) under the Maximum Likelihood criteria using PhyML (Guindon *et al.* 2010) under the LG+I+Γ model with both aLRT (approximate likelihood ratio tests SH-like) and bootstrap of 100 repetitions for branch support, while recombination events identified were further tested using GARD-SBP *(Kosakovsky Pond et al.* 2006). For the Gag and Pol trees, it was observed that DrERV does not group with other previously described bat betaretroviruses but clusters with SMRV (Figure 2A; Gag tree not shown). Because the pol region in CpERV is absent, we were not able to include it in the analysis. SMRV and DrERV form a sister clade to the β5-group rodent retroviruses, including RnERV_AC243170 (Figure 2A). A different phylogenetic pattern was observed for Env than in Gag and Pol. SMVR/DrERV/ CpERV form a discrete cluster with the Australian common brushtail possum retrovirus (TvERV; *Trichosurus vulpecula* endogenous betaretrovirus) at a basal position, all diverging from REV gammaretrovirus. In contrast, the rest of the β5 rodent and megabat ERVs diverge from the Gibbon ape leukemia gammaretrovirus (GALV), supporting the recombination observation made for the env gene (Figure 2B). Recombination tests confirmed a single breakpoint reflecting topological incongruence at position 258 in the DrERV Env protein (AIC value of 16025 and Δ AIC of 311.249; model averaged support 100% (data not shown). SMRV has a type C retroviral Env (Chiu *et al.* 1986), while other gammaretroviruses such as RD114 also have a Type D Env (van der Kuyl *et al.* 1999), suggesting that this region is not only prone to recombination but could potentially confer a selective advantage for the viruses in terms of their ability to bind to cells from different animals. Moreover, the long branch separating DrERV_216 from the other variants in the Env tree indicates that substantial time passed between the

introduction of DrERV_824 and DrERV_216. Therefore, an exogenous ancestor of DrERV might have circulated for a long time in the *D. rotundus* population.

Introduction of the SMRV/DrERV/CpERV lineage occurred via independent cross-species transmission events

At the time of integration retroviral LTR sequences are identical, but after integration they behave as paralog sequences evolving at the same rate as the host's genome. Therefore, it is possible to estimate the date of a provirus insertion based on the LTR's divergence using the following formula: T=(D/R)/2, were T is the invasion time (million years), D is the 5′ and 3′ LTR divergence given in number of differences, per nucleotide, per site (overall nucleotide divergence) (Tamura *et al.* 2007) and R is the genomic substitution rate per site per year (Kijima & Innan 2010), although there are some reservations about this dating method (Ishida *et al.* 2014; Mayer *et al.* 2013; Roca *et al.* 2004). For the DrERV_824 and DrERV_216 variants, an integration time of 17.9 and 16.0 Mya was estimated, while for CpERV a date of 7.7 Mya was obtained (Table 1). The insertion date for RnERV was estimated to have occurred 0.9 Mya, while for the SMRV, given that the LTRs are 100% identical, no insertion date estimate could be obtained and it is assumed that integration occurred recently (Table 1). The tMRCA date estimation for the SMRV/DrERV/CpERV cluster was inferred from the gag nucleotide alignment using the BEAST package under a GTR+I+G substitution model and a lognormal uncorrelated relaxed molecular clock (Drummond *et al.* 2012). The tMRCA estimation for the gag gene produced comparable results to the LTR dating, with a mean age for the SMRV/DrERV/CpERV cluster divergence of 17.8 Mya (95% HPD Interval: 10.5, 26.2) and a mean age for the tMRCA of the viral lineage of 24.4 Mya (95% HPD Interval: 15.85, 34.6). Finally, if the integration of an endogenous retrovirus occurred before species divergence, it can be assumed that the proviruses would have orthologous insertion sites. To test this, we compared the integration site sequences of DrERV and CpERV by sequence mapping to the *C. perspicillata* genome. For this analysis, we used the insertion sites of DrERV_824, as it is the oldest provirus and most likely ancestral. Our results reveal that the insertion sites between the CpERV and DrERV_824 are not homologous, as the corresponding insertion sites of DrERV could be mapped to other regions of the *C. perspicillata* genome (data not shown). The tMRCA date estimation for the SMRV/DrERV/CpERV lineage suggest that the exogenous ancestral virus must have circulated approximately 24.4 Mya, while the first invasion of *D. rotundus* occurred at least 17.9 Mya, pointing to DrERV as the oldest provirus within the SMRV/DrERV/CpERV lineage compared to the recent integration of the provirus in both rodents and squirrel monkeys. The integrity of SMRV and RnERV viral genomes strongly suggests that that these proviruses were 'recently' active and transmissible from an unknown host. Our results

suggests that exogenous member of this viral group was recently or still is in circulation in Latin America and has likely been transmitted to other species.

ACKNOWLEDGMENTS

Funding: Deutsche Forschungsgemeinschaft (DFG, GR 3924/9-1) and International Doctoral Scholarship Mexico and Germany (CONACyT-DAAD: 311664).

All free-ranging bats samples collection was approved by the Internal Committee for Ethics and Animal Welfare (No. approval: 2012-09-05) and meeting with Mexican regulations under the sample collection permit and export certificate (Num/SGPA/DGVS: 03173/14; SAGARPA: 241111524599811488A467371). Thanks to: Biol. Heiner Klös (Zoologischer Garten Berlin) and Biol. Carsten Kallasch (AG BUBO- Berlin) for providing samples from captive bats; all Mexican and German colleagues involved in the collection and exportation of samples, in particular to Gabor Czirjak, Luis Caballero, Omar Ríos and Alfredo Patraca; Tom Gilbert, Guojie Zhang and Mike Bunce (Centre for GeoGenetics, Copenhagen, Denmark) for granting access to their unpublished data; Joshua Hayward, Gilda Tachedjian (Burnet Institute, Melbourne, Australia) and Gregory Baillie (Wellcome Trust Sanger Institute, Cambridge, UK) for kindly providing the sequences from their work and Yasuko Ishida and Adam Brandt (University of Illinois, Urbana-Champaign USA) for their support in date estimation analysis. This is BeGenDiv publication number 14.

TABLE 1 Estimated insertion time of DrERV and related proviruses based on LTR divergence

Provirus	Divergence[a]	Average Length	Mutation Rate[b]	Insertion Date (Mya)[c]
DrERV_824	0,028	376	3,13E-09	17,9
DrERV_216	0,025	367	3,13E-09	16,0
CpERV	0,012	365	3,13E-09	7,7
RnERV	0,002	514	4,5E-09	0,9
SMRV	0	456	2,2E-09	ND[d]

[a]Divergence of LTRs for each provirus is defined as number of differences, per nucleotide, per site (overall mean distance).

[b]Estimated genomic mutation rates for host described in literature (23, 24, 2, 25)

[c]Calculated by T=(D/R)/2, where T is the invasion time (million years), D is the 5′ and 3′ LTR divergence and R is the genomic substitution rate per site per year.

[d]Given that the LTRs in this provirus show no differences, no accurate estimation can be obtained

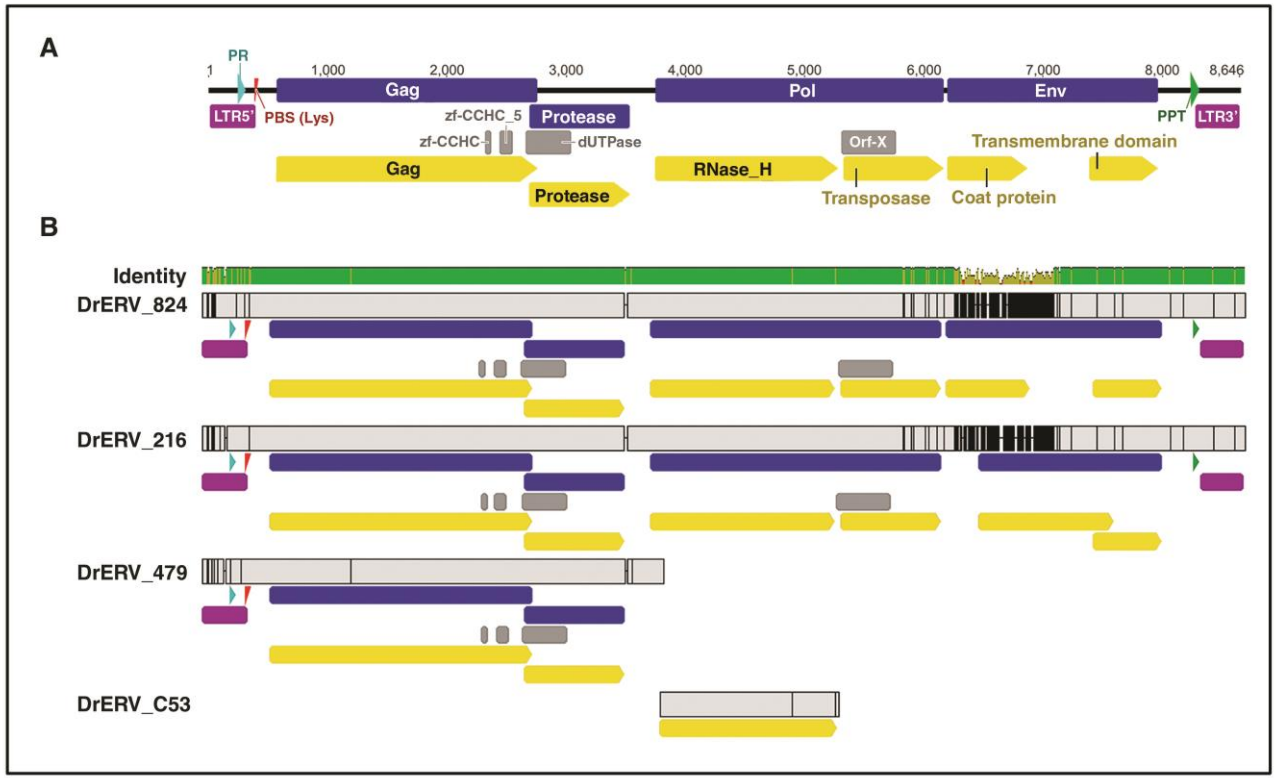

FIG 1
The genomic structure of DrERV. (A) Structural regions are shown as follows: 5' and 3' long terminal repeats (LTR; in purple), promoter region (PR; in turquoise), primer binding site (PBS; in red) and a polypurine tract upstream the env gene (PPT; in green). The core retroviral genes gag, protease, pol and env are depicted in blue. Coding regions are shown in yellow, while predicted conserved domains are shown in gray (zinc knuckle: zf-CCHC; zinc-finger domain: zf-CCHC_5; trimeric dUTP diphosphatase domain: dUTPase; Jaagsiekte sheep retrovirus-like ORF domain: Orf-X). (B) Alignment of the four DrERV variants (DrERV_216, DrERV_824, DrERV_479 and DrERV_C53) detected in the D. rotundus. Variants DrERV_216 and DrERV_824 are complete proviruses, while DrERV_479 and DrERV_C53 are truncated. Genome alignments are represented by the outlined bars in light gray with divergent sites highlighted in black. The green bar above indicates the percentage identity among the sequences (with green being the highest identity through red being the lowest). Structural elements are color coded as in panel A.

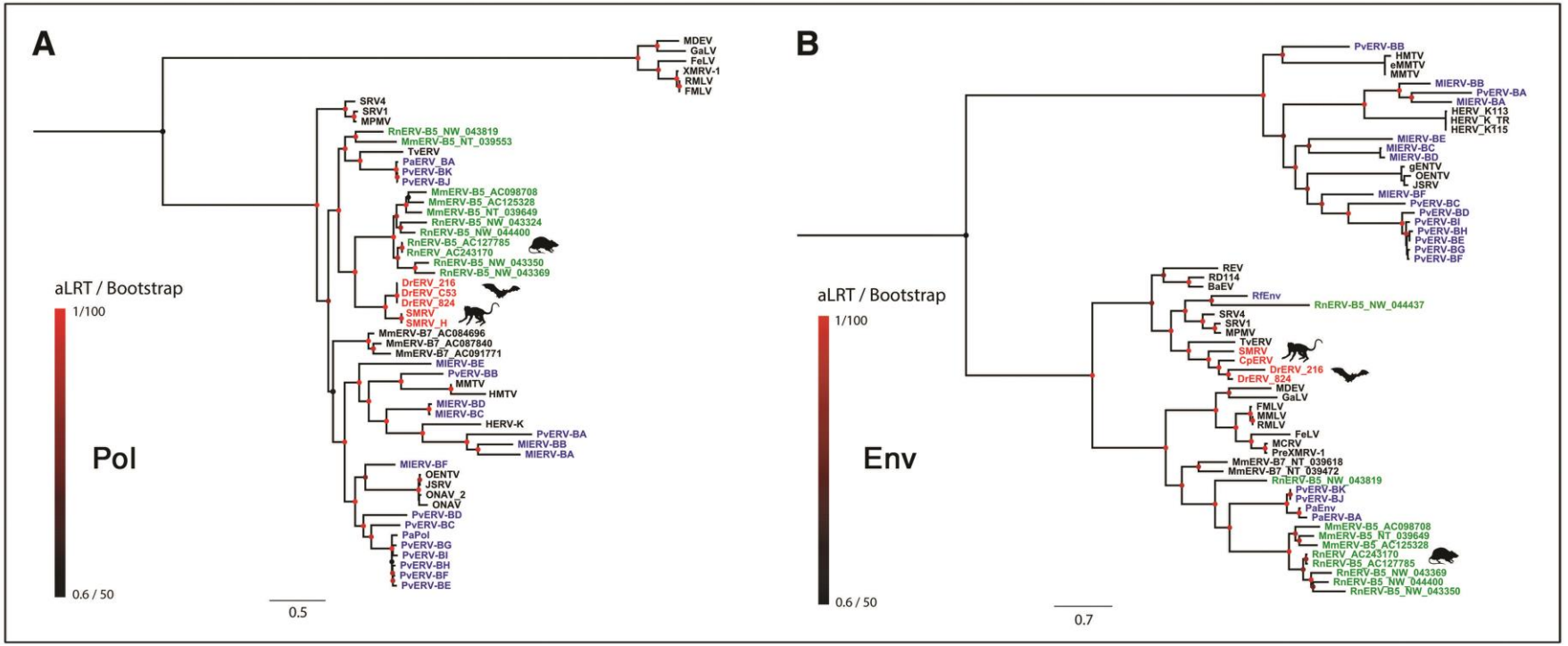

FIG 2
The Pol and Env DrERV phylogenetic trees. (A) The Pol tree (branch lengths scaled to amino acid substitutions per site) shows that DrERV (indicated by the bat pictogram) does not group with the previously described bat betaretrovirues (highlighted in blue), but forms a well-supported cluster (highlighted in red) with the SMRV (indicated by the primate pictogram) and the β5-group rodent ERVs (highlighted in green; indicated by the rodent pictogram). Support values are represented by circles in each node colored according to aLRT and bootstrap values in a gradient from black to red, with black being the lowest and red being the highest (1 in aLRT, 100 in bootstrap). (B) The Env tree (scaled as for the Pol tree) shows that the SMRV/DrERV/CpERV lineage (in red) is closely related to primate type D retroviruses diverging from REV gammaretroviruses, while the β5 rodent (in green) and bat ERVs (in blue) diverge from GALV gammaretroviruses, indicating recombination.

REFERENCES

1. **Lee DN, Papes M, Van den Bussche RA.** 2012. Present and potential future distribution of common vampire bats in the Americas and the associated risk to cattle. PloS one **7:**e42466.
2. **Cui J, Tachedjian G, Tachedjian M, Holmes EC, Zhang S, Wang LF.** 2012a. Identification of diverse groups of endogenous gammaretroviruses in mega- and microbats. The Journal of general virology **93:**2037-2045.
3. **Cui J, Tachedjian M, Wang L, Tachedjian G, Wang LF, Zhang S.** 2012b. Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. Journal of virology **86:**4288-4293.
4. **Hayward JA, Tachedjian M, Cui J, Field H, Holmes EC, Wang LF, Tachedjian G.** 2013a. Identification of diverse full-length endogenous betaretroviruses in megabats and microbats. Retrovirology **10:**35.
5. **Zhuo X, Rho M, Feschotte C.** 2013. Genome-wide characterization of endogenous retroviruses in the bat Myotis lucifugus reveals recent and diverse infections. Journal of virology **87:**8493-8501.
6. **Mayer J, Tsangaras K, Heeger F, Avila-Arcos M, Stenglein MD, Chen W, Sun W, Mazzoni CJ, Osterrieder N, Greenwood AD.** 2013. A novel endogenous betaretrovirus group characterized from polar bears (Ursus maritimus) and giant pandas (Ailuropoda melanoleuca). Virology **443:**1-10.
7. **Anai Y, Ochi H, Watanabe S, Nakagawa S, Kawamura M, Gojobori T, Nishigaki K.** 2012. Infectious endogenous retroviruses in cats and emergence of recombinant viruses. Journal of virology **86:**8634-8644.
8. **Baillie GJ, van de Lagemaat LN, Baust C, Mager DL.** 2004. Multiple groups of endogenous betaretroviruses in mice, rats, and other mammals. Journal of virology **78:**5784-5798.
9. **Baillie GJ, Wilkins RJ.** 2001. Endogenous type D retrovirus in a marsupial, the common brushtail possum (Trichosurus vulpecula). Journal of virology **75:**2499-2507.
10. **Chiu IM, Andersen PR, Aaronson SA, Tronick SR.** 1983. Molecular cloning of the unintegrated squirrel monkey retrovirus genome: organization and distribution of related sequences in primate DNAs. Journal of virology **47:**434-441.
11. **Mang R, Goudsmit J, van der Kuyl AC.** 1999. Novel endogenous type C retrovirus in baboons: complete sequence, providing evidence for baboon endogenous virus gag-pol ancestry. Journal of virology **73:**7021-7026.
12. **Roca AL, Pecon-Slattery J, O'Brien SJ.** 2004. Genomically intact endogenous feline leukemia viruses of recent origin. Journal of virology **78:**4370-4375.
13. **van der Kuyl AC, Dekker JT, Goudsmit J.** 1999. Discovery of a new endogenous type C retrovirus (FcEV) in cats: evidence for RD-114 being an FcEV(Gag-Pol)/baboon endogenous virus BaEV(Env) recombinant. Journal of virology **73:**7994-8002.
14. **Belshaw R, Pereira V, Katzourakis A, Talbot G, Paces J, Burt A, Tristem M.** 2004. Long-term reinfection of the human genome by endogenous retroviruses. Proceedings of the National Academy of Sciences of the United States of America **101:**4894-4899.
15. **Bonnaud B, Beliaeff J, Bouton O, Oriol G, Duret L, Mallet F.** 2005. Natural history of the ERVWE1 endogenous retroviral locus. Retrovirology **2:**57.

16. **Hayward A, Grabherr M, Jern P.** 2013b. Broad-scale phylogenomics provides insights into retrovirus-host evolution. Proceedings of the National Academy of Sciences of the United States of America **110:**20146-20151.
17. **Engel ME, Hiebert SW.** 2010. The enemy within: dormant retroviruses awaken. Nature medicine **16:**517-518.
18. **Ruprecht K, Mayer J, Sauter M, Roemer K, Mueller-Lantzsch N.** 2008. Endogenous retroviruses and cancer. Cellular and molecular life sciences : CMLS **65:**3366-3382.
19. **Wootton SK, Halbert CL, Miller AD.** 2005. Sheep retrovirus structural protein induces lung tumours. Nature **434:**904-907.
20. **Yi JM, Kim HM, Kim HS.** 2004. Expression of the human endogenous retrovirus HERV-W family in various human tissues and cancer cells. The Journal of general virology **85:**1203-1210.
21. **Colcher D, Heberling RL, Kalter SS, Schlom J.** 1977. Squirrel monkey retrovirus: an endogenous virus of a new world primate. Journal of virology **23:**294-301.
22. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nature methods **9:**357-359.
23. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26:**589-595.
24. **Zerbino DR.** 2010. Using the Velvet de novo assembler for short-read sequencing technologies. Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.] **Chapter 11:**Unit 11 15.
25. **Chiu IM, Skuntz SF.** 1986. Nucleotide sequence analysis of squirrel monkey retrovirus reveals a novel primer-binding site for tRNALys1,2. Journal of virology **58:**983-987.
26. **Elder JH, Lerner DL, Hasselkus-Light CS, Fontenot DJ, Hunter E, Luciw PA, Montelaro RC, Phillips TR.** 1992. Distinct subsets of retroviruses encode dUTPase. Journal of virology **66:**1791-1794.
27. **Coordinators NR.** 2014. Database resources of the National Center for Biotechnology Information. Nucleic acids research **42:**D7-17.
28. **Datzmann T, von Helversen O, Mayer F.** 2010. Evolution of nectarivory in phyllostomid bats (Phyllostomidae Gray, 1825, Chiroptera: Mammalia). BMC evolutionary biology **10:**165.
29. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics **5:**113.
30. **Gouy M, Guindon S, Gascuel O.** 2010. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution **27:**221-224.
31. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology **59:**307-321.
32. **Kosakovsky Pond SL PD, Gravenor MB, Woelk CH, Frost SD.** 2006. Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol. **23:**1891-1901.

33. **Chiu CN, Mitra R, Chiu IM.** 2003. Exchange of genetic sequences of long terminal repeat and the env gene by a promiscuous primate type D retrovirus. Virus research **96:**107-111.
34. **Tamura K, Dudley J, Nei M, Kumar S.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular biology and evolution **24:**1596-1599.
35. **Kijima TE, Innan H.** 2010. On the estimation of the insertion time of LTR retrotransposable elements. Molecular biology and evolution **27:**896-904.
36. **Ishida Y, Zhao K, Greenwood AD, Roca AL.** 2014. Proliferation of Endogenous Retroviruses in the Early Stages of a Host Germ Line Invasion. Molecular biology and evolution.
37. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution **29:**1969-1973.

DATASET S1 Samples used to build high-throughput sequencing libraries

Name	Sample Type	No. Individuals
DRBZ1[a]	Individual libraries: Lung, Spleen, Intestine, Salivary gland	1
DRMOR[b]	Pool: Stomach, Spleen, Lung	7
DREDO[c]	Pool: Stomach, Spleen, Lung	2
DRSD[d]	Pool: Spleen	22

[a] Desmodus rotundus; Berlin Zoologischer Garten, Berlin, Germany

[b] Desmodus rotundus; San Pablo, Tlaltizapan Morelos, Mexico

[c] Desmodus rotundus; Estado de Mexico, Mexico

[d] Desmodus rotundus; Soledad Doblado, Veracruz, Mexico

DATASET S2 Primers used in this study

Primer Sequence (5' to 3')	Name	Tm	Product length	Target
TTTGCAGCTTGATCAGAATCCT	F1_DrERV	58	1789	DrERV P[a]
TATTTTTGCAAGAGAAGTGGTTGC	R1_DrERV	58	1789	DrERV P
CCTATCGCCAAGGAGAGCT	F2_DrERV	58	1454	DrERV P
GGGTATCTAGAATAATTGGCAGGAA	R2_DrERV	58	1454	DrERV P
GATTCTTACCCGGCCAAGG	F3_DrERV	58	3500	DrERV P
GGCTGTCAAAATGTCTGTCGTT	R3_DrERV	59	3500	DrERV P
GTAATCTCTCATATGATCCACTGCT	F4_DrERV	58	1900	DrERV P
GACTACTTCTGCTAGGGAATCTAAC	R4_DrERV	58	1900	DrERV P
TTCATCCGTTACAGGTATATCG	F5_DrERV	55	563	DrERV P
TGTATGTCAGGGGAATTATTGT	R5_DrERV	55	563	DrERV P
ACAATAATTCCCCTGACATACA	F6_DrERV	55	930	DrERV P
GTGTGTTACAGCAGATTTGAC	R6_DrERV	55	930	DrERV P
ATCAGCCTGCATTTATGATCA	F7_DrERV	55	912	DrERV P
CAATCTTGGGCTAGTGTAGG	R7_DrERV	55	912	DrERV P
GGCAAACCTATCCAGATACAT	F8_DrERV	55	700	DrERV P
ATTCTGATCAAGCTGCAAATTT	R8_DrERV	55	700	DrERV P

AATTAAAAGAATCTCTCAAGGTGC	GagF	56	288	gag D[b]
GAGAATTTCTTGCCCTTCTTTG	GagR	56	288	gag D
ACAGACATCTATGGACAGCC	PolF	56	150	pol D[c]
ATTGATTACTGGGAGGGCA	PolR	56	150	pol D
CTGCAACACAAAAACAGCTTA	EnvF	56	339	env D[d]
CCGGCTTGTAAGTATTTGGT	EnvR	56	339	env D
AAAAGAGAGCTGCGGATACC	DrERV_WG_F	57	8768	DrERV WG[e]
CAGCAAGACAAAGGATTCTGA	DrERV_WG_R	56	8768	DrERV WG

[a]DrERV P: Primers for amplification of partial sequences of the DrERV genome

[b]GagD: Primers for amplification of a small fragment of the gag gene region

[c]PolD: Primers for amplification of a small fragment of the pol gene region

[d]EnvD: Primers for amplification of a small fragment of the env gene region

[e]DrERV WG: Primers for amplification of the complete DrERV genome

DATASET S3 Sequences and accession numbers used in this study

Abbreviation	Name	Accesion number
JSRV	Jaagsiekte sheep retrovirus	NC_001494.1
MPMV	Mason-Pfizer monkey virus	NC_001550.1
SMRV, SMRV-H	Squirrel monkey retrovirus	NC_001514.1 and M23385.1
MMTV	Mouse mammary tumor virus	NC_001503.1
eMMTV	Endogenous mouse mammary tumor virus	AF228552.1
SRV1	Simian retrovirus 1	M11841.1
SRV4	Simian retrovirus 4	NC_014474.1
oENTV	Ovine enzootic nasal tumor virus	NC_007015.1
gENTV	Enzootic nasal tumor virus of goats	NC_004994.2
HMTV	Human mammary tumor virus	AF248269.1 and AF346816.1
HERV-K113	Human endogenous retrovirus-K113	AY037928.1
HERV-K115	Human endogenous retrovirus-K115	AY037929.1
HERV-K-TR	Human endogenous retrovirus-K-TR	AF074086.2
TvERV	Trichosurus vulpecula endogenous betaretrovirus	AF224725.1
FMLV	Friend murine leukemia virus	NC_001362.1
RMLV	Rauscher murine leukemia virus	NC_001819.1
MMLV	Moloney murine leukemia virus	NC_001501.1
PreXMRV-1	Pre-Xenotropic murine leukemia virus-related virus-1	NC_007815.2
MCRV	Murine type C retrovirus	NC_001702.1
FeLV	Feline leukemia virus	NC_001940.1
GALV	Gibbon ape leukemia virus	NC_001885.2
MDEV	Mus dunni endogenous retrovirus	AF053745.1
BaEV	Baboon endogenous virus strain M7	D10032.1
RD114	RD114 feline endogenous retrovirus	NC_009889.1
REV	Reticuloendotheliosis virus	NC_006934.1
MlERV-βA	Myotis lucifugus endogenous retrovirus - betaretrovirus A	Scaffold GL429780; 11816573-11826438; provided by the author

MlERV-βB	Myotis lucifugus endogenous retrovirus - betaretrovirus B	Scaffold GL429905; 2902336-2910456; provided by the author
MlERV-βC	Myotis lucifugus endogenous retrovirus - betaretrovirus C	Scaffold AAPE02058399; 20007-28108; provided by the author
MlERV-βD	Myotis lucifugus endogenous retrovirus - betaretrovirus D	Scaffold AAPE02063220; 12691-3685; provided by the author
MlERV-βE	Myotis lucifugus endogenous retrovirus - betaretrovirus E	Scaffold GL429817; 2299709-2307595; provided by the author
MlERV-βF	Myotis lucifugus endogenous retrovirus - betaretrovirus F	Scaffold GL431456; 9711-1477; provided by the author
PvERV-βA	Pteropus vampyrus endogenous retrovirus - betaretrovirus A	Scaffold 22753; 8224-518; provided by the author
PvERV-βB	Pteropus vampyrus endogenous retrovirus - betaretrovirus B	Scaffold 15954; 22278-13022; provided by the author
PvERV-βC	Pteropus vampyrus endogenous retrovirus - betaretrovirus C	GeneScaffold 3132; 77401-84529; provided by the author
PvERV-βD	Pteropus vampyrus endogenous retrovirus - betaretrovirus D	GeneScaffold 2885; 129192-136214; provided by the author
PvERV-βE	Pteropus vampyrus endogenous retrovirus - betaretrovirus E	Scaffold 9648; 43764-35887; provided by the author
PvERV-βF	Pteropus vampyrus endogenous retrovirus - betaretrovirus F	Scaffold 17393; 20588-12812; provided by the author
PvERV-βG	Pteropus vampyrus endogenous retrovirus - betaretrovirus G	Scaffold 12793; 22017-29493; provided by the author
PvERV-βH	Pteropus vampyrus endogenous retrovirus - betaretrovirus H	GeneScaffold 1344; 357039-349197; provided by the author
PvERV-βI	Pteropus vampyrus endogenous retrovirus - betaretrovirus I	Scaffold 2273; 57766-7814; provided by the author
PvERV-βJ	Pteropus vampyrus endogenous retrovirus - betaretrovirus J	Scaffold 7237; 70384-61612; provided by the author
PvERV-βK	Pteropus vampyrus endogenous retrovirus - betaretrovirus K	Scaffold 10684; 4789-13495; provided by the author
PaERV-βA	Pteropus alecto endogenous retrovirus - betaretrovirus A	SRP008674; provided by the author
PaPol-01	Pteropus alecto Polymerase 01	SRP008674; provided by the author
PaEnv-01	Pteropus alecto Envelope 01	SRP008674; provided by the author
RfEnv-01	Rhinolophus ferrumequinum Envelope 01	Provided by the author
MmERV-β5_AC098708	Mus musculus β5 betaretrovirus	AC098708; provided by the author
MmERV-β5_AC125328	Mus musculus β5 betaretrovirus	AC125328; provided by the author
MmERV-β5_NT_039553	Mus musculus β5 betaretrovirus	NT_039553; provided by the author
MmERV-β5_NT_039649	Mus musculus β5 betaretrovirus	NT_039649; provided by the author
RnERV-β5_AC127785	Rattus norvegicus clone CH230-127O17	AC127785, provided by the author
RnERV-β5_NW_043324	Rattus norvegicus β5 betaretrovirus	NW_043324; provided by the author
RnERV-β5_NW_043350	Rattus norvegicus β5 betaretrovirus	NW_043350; provided by the author
RnERV-β5_NW_043369	Rattus norvegicus β5 betaretrovirus	NW_043369; provided by the author
RnERV-β5_NW_043819	Rattus norvegicus β5 betaretrovirus	NW_043819; provided by the author

RnERV-β5_NW_044400	Rattus norvegicus β5 betaretrovirus	NW_044400; provided by the author
RnERV-B5_NW_044437	Rattus norvegicus β5 betaretrovirus	NW_044437; provided by the author
CpERV-β5_AC138156	Carollia perspicillata clone 41M6, complete sequence	AC138156.3; from base 37196 to 43259.
RnERV_AC243170	Rattus norvegicus Y Chr BAC RNAEX-212I22 complete sequence	AC243170; from base 127494 to 137275.
DrERV_824[a]	Desmodus_rotundus_endogenous_retrovirus_DrERV_824	KP175580
DrERV_216[a]	Desmodus_rotundus_endogenous_retrovirus_DrERV_216	KP175581
DrERV_479[a]	Desmodus_rotundus_endogenous_retrovirus_DrERV_479	KP175582
DrERV_C53[a]	Desmodus_rotundus_endogenous_retrovirus_DrERV_C53	KP175583

[a]Sequences described in this work

Chapter 2: The evolution of bat nucleic acid sensing Toll-like receptors

Marina Escalera-Zamudio[*,1], M. Lisandra Zepeda-Mendoza[2], Elizabeth Loza-Rubio[3], Edith Rojas-Anaya[3], Maria L. Méndez-Ojeda[4], Carlos F. Arias[5] and Alex D. Greenwood[*,1]

[1]Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research (IZW), Alfred-Kowalke-Straße 17 10315 Berlin, Germany; [2]Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5-7, 1350 Copenhagen, Denmark; [3]Centro Nacional de Investigación Disciplinaria en Microbiología Animal-INIFAP, kM 15.5 carretera federal México-Toluca Col. Palo Alto 05110 México City, Mexico; [4]Facultad de Medicina Veterinaria y Zootecnia, Universidad Veracruzana, Miguel Ángel de Quevedo Col. Unidad Veracruzana 91710 Veracruz, Mexico; [5]Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Chamilpa, 62210 Cuernavaca Morelos, Mexico

Keywords: Molecular evolution, selection, nucleic acid-sensing Toll-like receptors, bats

*Corresponding authors: e-mail: greenwood@izw-berlin.de and escalera@izw-berlin.de Leibniz-Institut für Zoo- und Wildtierforschung (IZW). Alfred-Kowalke-Straße 17,10315 Berlin, Germany. Fax: 0049 (0) 30 5126 104
Running title: The bat TLR evolution

ABSTRACT

We characterized the nucleic acid sensing Toll-like receptors (TLR) of a New World bat species, the common vampire bat (*Desmodus rotundus*), and through a comparative molecular evolutionary approach searched for general adaptation patterns among the nucleic acid sensing TLRs of eight different bats species belonging to three families (*Pteropodidae, Vespertilionidae* and *Phyllostomidae*). We found that the bat TLRs are evolving slowly and mostly under purifying selection and that the divergence pattern of such receptors is overall congruent with the species tree, consistent with the evolution of many other mammalian nuclear genes. However, the chiropteran TLRs exhibited unique mutations fixed in ligand binding sites, some of which involved non-conservative amino acid changes and/or targets of positive selection. Such changes could potentially modify protein function and ligand biding properties, as some changes were predicted to alter nucleic acid binding motifs in TLR 9. Moreover, evidence for episodic diversifying

selection acting specifically upon the bat lineage and sub lineages was detected. Thus, the long-term adaptation of chiropterans to a wide variety of environments and ecological niches with different pathogen profiles is likely to have shaped the evolution of the bat TLRs in an order-specific manner. The observed evolutionary patterns provide evidence for potential functional differences between bat and other mammalian TLRs in terms of resistance to specific pathogens or recognition of nucleic acids in general.

INTRODUCTION

The order Chiroptera is one of the most diverse mammalian groups with an estimate of 1,200 species and new bat species still being described (Gorfol *et al.* 2014; Koubinova *et al.* 2013; Simmons 2005). Bats display unique traits among mammals such as flight, echolocation and exceptional dietary adaptations like hematophagy. Their worldwide distribution and adaptation to different environments has resulted in the evolution of diverse and unique phenotypic traits (Patterson *et al.* 2003; Simmons 2005). The study of the genetic changes associated with the development of bat-specific traits has been recently explored through a whole genome comparison approach of three bat species (*Pteropus alecto, Myotis davidii* and *Myotis brandtii*) (Seim *et al.* 2013; Zhang *et al.* 2013), revealing that positive selection is not only acting upon genes associated with physiological traits such as hibernation and vision but also on immunity-associated genes (Zhang *et al.* 2013). The analysis of the immune variation at a molecular level can reveal patterns of resistance or susceptibility to pathogens within different species and at different taxonomic levels (Alcaide & Edwards 2011; Schroder & Schumann 2005; Tschirren *et al.* 2012; Wlasiuk & Nachman 2010). However, the study of the genetic variability of the immune system in bats has been restricted to a few species (mainly belonging to the *Pteropus* and *Myotis* genus) and to a few genes, in part due to the lack of sequences available for comparative analyses (Baker *et al.* 2013; Cowled *et al.* 2011; He *et al.* 2010; Iha *et al.* 2010; Kepler *et al.* 2010; Omatsu *et al.* 2008; Sarkar & Chakravarty 1991).

Toll like-receptors (TLRs) are a class of innate immune receptors considered to be the first-line defense mechanism against invading pathogens by recognizing a wide variety of pathogen-associated molecular patterns (Boehme & Compton 2004). The TLRs 3, 7, 8 and 9 play an important role in triggering acquired immunity, as they are activated by nucleic acid ligands such as unmethylated DNA, dsRNA and ssRNA present in most pathogens including viruses (Boehme & Compton 2004). TLRs are of interest from an evolutionary point of view since there is evidence that the ligand recognition properties of these receptors may vary among species, thereby having an impact on the evolutionary ecology of infectious diseases (Tschirren *et al.* 2012; Werling *et al.* 2009). Given the unique adaptations present within the Chiroptera, we predict that

bats as a taxonomic group have acquired distinctive mutations fixed within the nucleic-acid sensing TLRs with potential consequences on their ligand recognition properties. In this study, we characterized the nucleic acid sensing TLRs (3, 7, 8 and 9) of the common vampire bat (*Desmodus rotundus*; Phyllostomidae family) and compared the genetic variation of these receptors within eight different bats species belonging to three different families (*Pteropodidae, Vespertilionidae* and *Phyllostomidae*). We further tested for ongoing and episodic diversifying selection acting upon the bat TLRs to describe their evolution over both a long and short timescales. Although the bat TLRs are generally evolving under similar functional constraints as in other mammals, we detected several lineage-specific mutations and found evidence for positive selection occurring at different taxonomic levels. These findings may reflect the broad spectrum of bat-specific adaptations to a wide variety of ecological niches and pathogens to which bats are exposed to.

MATERIALS AND METHODS

Sample collection and DNA extraction

Tissue samples from one *D. rotundus* individual captured in Mexico were used for nucleic acid extraction. The animal was sacrificed and submitted to necropsy in order to obtain the spleen, lung and intestine tissues. Approximately 25 mg of each tissue was homogenized using the Precellys 24 tissue homogenizer (Bertin technologies) and further prepared according to the protocol suggested by the DNeasy Blood & Tissue Kit (Qiagen) to extract genomic DNA. All samples were then kept at -20°C for further use.

TLR characterization

Primers derived from the orthologous mammalian TLR 3, 7, 8 and 9 sequence alignments were used to amplify partial sequences for each TLR from different tissues, as described in (Iha *et al.* 2010). All products were Sanger sequenced on both strands (StarSEQ; Mainz, Germany) and further analyzed by blastn to determine their homology and sequence identity to other bat TLRs (NCBI 2015). In order to obtain the complete consensus coding DNA sequences (CDS) for each TLR, the partial sequences obtained by PCR were mapped against *D. rotundus* genomic data using Bowtie version 2.2.2 with a stringency of 80% (Zepeda-Mendoza *et al.* unpublished data;(Langmead & Salzberg 2012). Complete coding regions were detected using Geneious (Geneious v7.5.5, Biomatters 2014) (Supporting Information: Table S1) and ORFs were translated into protein sequences and annotated by searching for conserved domains with LRRfinder (http://www.lrrfinder.com). To obtain the 3-D structures of each TLR ectodomain, the fully annotated protein sequences were 3 D modeled using the SWISS-MODEL automated

system (Biasini *et al.* 2014) using the crystal structures of the unliganded human TLR 3 and 8 and of the horse TLR 9 as templates (File S1) (Bell *et al.* 2005; Choe *et al.* 2005; Ohto *et al.* 2015; Tanji *et al.* 2013). All models yielded a good fit with the corresponding templates, as determined by QMEAN and GMQE (Supporting Information: Table S2) (Benkert *et al.* 2008). Visualization and image generation was performed with PyMOL Molecular Graphics System Version 1.5.0.4 (Schrödinger, LLC).

Phylogenetic analysis

The complete CDS for the TLR 3, 7, 8 and 9 of 35 representative Laurasiatherian and Euarchontoglires mammalian species were downloaded from the NCBI Database (Supporting Information: Table S3) (NCBI 2015). The available sequences retrieved included those from seven bat species: *Eptesicus fuscus, Myotis brandtii, Myotis davidii, Myotis lucifugus, Pteropus alecto, Pteropus vampyrus* and *Rousettus leschenaultii* (Supporting Information: Table S3) (Cowled *et al.* 2011; Iha *et al.* 2010). The TLR sequences for *D. rotundus* obtained in this study were further added to each dataset and sequences were aligned based on their protein translation using MUSCLE implemented in SeaView (File S2) (Edgar 2004; Gouy *et al.* 2010). Modeltest was applied to each dataset to determine the best-fit substitution model and phylogenetic analysis based on the nucleotide sequences was performed under maximum likelihood criteria using PhyML 3.0 (GTR+G; with approximate likelihood ratio tests SH-like for branch support) (File S2) (Guindon *et al.* 2010; Posada 2006). All trees were rooted with the marsupial *Monodelphis domestica* sequence determined to be the closest outgroup for eutherian mammals. The resulting topologies were compared to the previously reported mammalian and chiropteran species trees (Agnarsson *et al.* 2011; Cowled *et al.* 2011; Jones *et al.* 2005; Nery *et al.* 2012; Zhou *et al.* 2012).

Nucleotide diversity and rate estimations

The aligned bat and mammalian sequences were used to estimate the mean nucleotide diversity (π; defined as the average number of nucleotide substitutions per site for a group of DNA sequences) (Nei & Jin 1989). The π estimates between different taxonomic groups were obtained using the Nei method implemented in MEGA 6.0 under the maximum composite likelihood model with a gamma distribution for rate variation among sites (Nei & Jin 1989; Tamura *et al.* 2007). Units represent the percentage of nucleotide substitutions per site between sequences adjusted in Tamura-Nei distances, while the standard error was obtained with 500 bootstrap replicates. The evolutionary rates for the mammalian TLRs were inferred under a lognormal uncorrelated relaxed molecular clock using the BEAST package (Drummond & Rambaut 2007). The average

mutation rate for mammalian genomes (2.2E-9 substitutions per site per year) (Kumar & Subramanian 2002) and a yule tree assuming a constant speciation rate per lineage were used as priors (Drummond & Rambaut 2007). Analysis was run for 8E08 MCMC chains or until all relevant parameters converged (with an ESS ≥ 250 in all cases). Ten percent of the MCMC chains were discarded as burn-in.

Estimation of global dN/dS (ω) and detection of positive selected sites (PSS)

Under neutral evolution or purifying selection, coding sequences are expected to have a ratio of dN/dS close to 1 (ω ≤1), while significant rate deviations may be interpreted as evidence for positive selection (ω >1) (Pond & Muse 2005). The effect of positive selection acting upon the bat lineage compared to the rest of the mammalian tree was tested by estimating global ω values for the bat node and for the rest of the mammalian tree under the M1/M2 branch model with CODEML in PAML v4 (File S3) (Yang 2007). The M1 model allows for neutral evolution (ω =1) assuming identical ω ratios among all branches, while the M2 model constraints a specific node and allows for positive selection (ω >1) to act on the selected branches. Models were evaluated under a likelihood ratio test (LRT) using a χ^2 distribution with the number of *d.f.* obtained from the number of parameters used and testing under a p value <0.05 (File S4). For a conservative detection of sites evolving under ongoing positive selection (PSS) within the bat TLRs, datasets were tested under the M8a/M8 site model using CODEML in PAML v4 (File S3) (Yang 2007). The M8 model allows all sites to evolve under positive selection (ω > 1), while the M8a model uses the same parameters but with ω fixed to 1 (Yang 2007). Models were compared under a LRT as described above and PSS were scored under Naive Empirical Bayes (NEB) and/or Bayes empirical Bayes (BEB) with a P≥ 90% (File S4). Scored PSS were further confirmed under the Random Effect Likelihood method (REL) implemented in the datamonkey sever (http://www.datamonkey.org) using the GTR model and accepting sites with a posterior probability (PP) ≥80 (Delport *et al.* 2010; Pond & Muse 2005). The REL method fits a distribution of rates across sites and then infers the substitution rate for all individual sites. Thus, it has the advantage of improving the ω ratio estimation by incorporating variation in the synonymous substitution rate (Pond & Muse 2005). PSS detected were further mapped onto the previously obtained phylogenies using MEGA 6.0 by reconstruction of ancestral states under a maximum likelihood framework (Tamura *et al.* 2007).

Ligand binding site (LBS) mutations and detection of episodic positive selection

The positions of the previously described LBS in the human, mouse, horse and bovid nucleic acid-sensing TLRs were visually determined within each protein alignment and unique mutation

patterns in the bat sequences were detected by site-to-site comparison (Bell *et al.* 2006; Bell *et al.* 2005; Choe *et al.* 2005; Ohto *et al.* 2015; Pan *et al.* 2012; Tanji *et al.* 2013; Wei *et al.* 2009; Zhou *et al.* 2013). LBS mutations detected were mapped onto the *D. rotundus* TLR 3-D models using PyMOL Molecular Graphics System Version 1.5.0.4 (Schrödinger, LLC). To further test if the bat-specific LBS mutations had been a target of episodic positive selection, we evaluated each dataset by constraining the bat node from the rest of the mammalian tree under the branch-site model A (BSA) using CODEML in PAML v4 (File S3) (Yang 2007). The branch-site model A estimates ω values upon sites and specific branches, classifying sites into four different categories: Class 0 (purifying selection on all branches), Class 1 (neutral evolution on all branches), Class 2a (positive selection in selected branches and purifying selection for the rest of the tree) and Class 2b (purifying selection on the selected branch and neutral evolution for the rest of the tree) (Yang 2007). Thus, it is sensitive to lineage-specific PSS (e.g. a definite number of sites within defined lineages or branches) that evolved under positive selection at some point in evolutionary history (Murrell *et al.* 2012; Yang & dos Reis 2011). The BSA model was evaluated under a LRT against the null hypothesis (BSA with ω fixed to 1), while PSS were scored under BEB with a P≥ 95% (File S4). Finally, episodic diversifying selection acting within specific lineages was further tested under the branch site-random effects likelihood model (BS-REL) implemented in the Datamonkey server (http://www.datamonkey.org) using the GTR model with P values <0.1 (Delport *et al.* 2010; Kosakovsky Pond *et al.* 2011). The BS-REL model again considers three different evolutionary scenarios (purifying, neutral and diversifying selection), allowing evolutionary rate variation along both branches and sites simultaneously and without making any a priori assumptions about lineages (Kosakovsky Pond *et al.* 2011; Kosakovsky Pond SL 2006). The effect of selection is detected as deviations in the ω rate ratio across tree branches and tested under a LRT to identify branches with significant evidence of positive selection (Delport *et al.* 2010; Kosakovsky Pond *et al.* 2011; Yang & dos Reis 2011).

RESULTS

The structure of the *D. rotundus* nucleic acid sensing TLRs

The TLR 3, 7, 8 and 9 of *D. rotundus* exhibited the classic genetic characteristics of other mammalian TLRs including a signal peptide, an ectodomain with several Leucine-rich repeats (LRRs) where ligand binding sites (LBS) occur, and a highly conserved transmembrane and TIR domain towards the C-terminus of the protein. The 3-D structures of each ectodomain revealed the classic horseshoe-shaped solenoid with a concave and convex surface and a descending and ascending lateral surface (Figure 1, Panel A) (Bell *et al.* 2003; Botos *et al.* 2011; Uematsu & Akira 2007). Annotation of the protein sequences demonstrated the presence of 23-26 LRRs

within each TLR ectodomain exhibiting the typical consensus sequence LxxLxLxxN/CxL, followed by a number of hydrophobic residues spaced at distinct intervals and adopting a loop structure. Prediction of N-linked glycosylation sites in each TLR yielded 13 to 17 sites for C-mannose likely required for bioactivity as described for other mammalian TLRs (de Bouteiller *et al.* 2005; Sun *et al.* 2006).

Slow evolution and divergence of TLRs

The topologies obtained for each TLR were generally congruent with the chiropteran and mammalian species trees in which all bats form a well-supported clade with the microchiroptera (*D. rotundus, E. fuscus, M. brandtii, M. davidii* and *M. lucifugus*) diverging from the megachiroptera (*Pteropus species* and *R. leschenaultii*) (Agnarsson *et al.* 2011; Jones *et al.* 2005). For the TLR 3 tree, the bat lineage was closest to perissodactylans and more distantly related to carnivores, while for the TLR 7 and 8 trees bats formed a sister clade to both carnivores and euungulates. However, in the TLR 9 tree, bats formed a monophyletic clade positioned externally to all other eutherian mammals (Figure 2), as observed in previos studies for TLR 9 (Cowled *et al.* 2011). We further estimated the mean nucleotide diversisty (π) for the bat sequences and for the rest of the mammalian sequences and compared the mean overall diversity between groups. The π value estimated for all mammalian TLR sequences including bats ranged between 0.15-0.24% (SE 0.004-0.012), while the π estimate within mammalian sequences excluding bats was 0.14-0.21% (SE 0.004-0.02). The π estimate for only the bat sequences ranged between 0.10-0.19% (SE 0.004-0.012), while the π value between the mammalian and bat sequences was between 0.16 to 0.19% (SE 0.002-0.016). These results suggest that all TLR sequences exhibit roughly the same π value among and within groups, indicating no significant variation between estimates for the bat and the rest of the mammalian sequences. The π values were also consistent with those previously reported for other mammalian autosomal nuclear genes ranging between 0.01-0.5% (Leffler *et al.* 2012). Partial sequences for each TLR were determined for a total of 30 individual *D. rotundus* samples with an observed identity >99.9% among sequences. Thus, the consensus results based on a single sequence are representative of the *D. rotundus* species (data not shown). Finally, we estimated the relative speed at which such receptors are evolving. The mean substitution rate for the bat and the rest of the mammalian TLRs ranged between 1.3E-9 and 2.3E-9 (1.07E-9, 3.12E-9 95%HPD), which is comparable to the evolutionary rate for other nuclear genes in mammals (Kumar & Subramanian 2002). The values for the *coefficient of variation* and *ucld.stdev* parameters (standard deviation of the uncorrelated lognormal relaxed clock; σ) obtained ranged between 0.6 and 0.8 in all cases, indicating a substantial heterogeneity in the substitution rate

among branches, thus showing that the nucleic acid sensing TLRs of mammals are not evolving in a strict-clock like fashion (Drummond & Rambaut 2007).

Functional constraint but evidence for site-specific positive selection

The LRT for the branch-specific model (M1/M2) was significant for all TLRs yielding global estimates of ω =1 for the bat node and of 0.15≥ ω ≥0.2 for the rest of the branches (File S4), suggesting the bat TLRs are either evolving neutrally resulting from relaxation of functional constraint, or that both positive and purifying selection occur canceling each other out and yielding an ω value close to 1. We further sought to detect if there were PSS within the bat sequences under a site-specific model (M8a/M8). LRT were again significant for all TLRs, detecting a few sites evolving under positive selection (File S4) (Table 1). As expected, the proportion of PSS with an estimated ω >1 was very low in all cases (0.19 for TLR 3, 0.01 for TLR 7, 0.1 for TLR 8 and 0.007 for TLR 9), suggesting that most sites are evolving either under purifying selection or neutrally (File S4). A total of six PSS were detected for TLR 3, one for TLR 7, eight for TLR 8 and four for TLR 9, all of them located within the LRRs of the proteins (Table 1). The PSS detected were found to be moderately conserved among all mammalian taxa, but often showed high variability and represented non-conservative changes in the bat species studied. To determine the influence of the PSS on the divergence of the bat lineage, we mapped the detected sites onto the previously obtained phylogenies. The TLR 3 tree showed an accumulation of four PSS on the bat node with two sites mapping to the *Pteropus* genus branch, while for the TLR 7 and 8 trees most sites mapped to the *Pteropus* branch (the single site detected for TLR 7 and seven out of eight for the TLR 8, respectively). For TLR 9, three out of four sites mapped to the microbat lineage clustering mainly on the *D. rotundus* branch (Figure 2).

Fixed mutations in the LBS of the bat TLRs

We searched for mutations within the TLRs in sites interacting directly or indirectly with nucleic acid ligands as determined both *in vitro* and *in silico* for other mammalian species (Bell *et al.* 2006; Bell *et al.* 2005; Choe *et al.* 2005; Ohto *et al.* 2015; Pan *et al.* 2012; Tanji *et al.* 2013; Wei *et al.* 2009; Zhou *et al.* 2013). A total of four bat-specific LBS mutations compared to other mammalian sequences were found for TLR 3 (with the first letter being the most prevalent amino acid among other taxa and the second the mutation the one observed in bat species): K493Q and R331G present in the Pteropodidae family (*R. leschenaultii* and *Pteropus* spp), R331C in *E. fuscus* and N145D in the *Myotis* genus (Figure 1, Panel B; Table 2). No bat-specific LBS mutations were observed for TLR 7. For TLR 8 two different mutations in the same position were detected: N539H in the *Myotis* genus and N539S in *Pteropus* species (Figure 1, Panel B; Table 2). The

majority of mutations were detected within TLR 9, with a total of six changes: K51L/T, H76R and Q335K/T in microbats, K181Q/E in the Vespertilionidae family, K286S/N heterogeneously present within three different families (Vespertilionidae, Phyllostomidae and Pteropodidae) and L364Q found only in Pteropodids (Table 2). We further explored if the LBS mutations found within TLR 9 could potentially alter described binding motifs present on the secondary structure of the protein (Ohto *et al.* 2015). The binding motif S-N-R-I-H-H located within the LRR1 of the TLR 9 protein (positions 98 to 103 of the alignment; File S2) is highly conserved among different orders but is modified to C-N-R-I-R/H-H in bats (Figure 1, Panel B) (Ohto *et al.* 2015). Similarly, the conserved motif S-P-M-H-F-P located within the LRR2 of the protein (positions 130-135 of the alignment; File S2) is altered to S-I/R-M-H-W-A/D/N in all bat sequences, while the consensus H-T-L-L located within the LRR20 (positions 679-682 of the alignment; File S2) is H-A-V/I/L-L in bats (Figure 1, Panel B). Only the mutation R331C/G found in TLR 3 involving a non-conservative amino acid change was identified as a target of ongoing positive selection under the M1/M2 site model (Table 2).

Episodic positive selection targeting the bat lineage

As most of the LBS mutations were not detected to be targets of ongoing positive selection under the M1/M2 site models, we tested if they had been targets of positive selection at some point in their evolutionary history under the branch-site model (BSA). Our results confirmed our previous observations that most of the codon sites in the bat and mammalian nucleic acid sensing TLRs are evolving under purifying selection, with the highest proportion of sites evolving under $\omega < 1$ (0.74 for TLR 3, 0.70 for TLR 7, 0.84 for TLR 8 and of 0.65 for TLR 9). Only the LRT for TLR 9 was significant as approximately 0.1% of all sites were estimated to be evolving under a $\omega > 1$ when compared to the rest of the mammalian tree at $p<0.05$ (File S4). Two bat-specific LBS mutations (K181Q/E and K286S/N) were scored as PSS under BEB with a P>95% (Table 2). We then analyzed what lineages had been a target of episodic diversifying selection within each TLR tree under the Branch-site REL model. Analysis of TLR 3 failed to identify any lineages evolving with a significant proportion of sites under $\omega > 1$ tested at $p < 0.05$, while for TLR 7 only the bat node was detected to have evolved under episodic positive selection with a mean $\omega = 1.5$ ($p <0.0001$) (Figure 2). For TLR 8, the bat ($\omega = 1.76$; $p <0.03$) and *Pteropus* nodes ($\omega = 0.66$; $p =0.00$) as well as the *M. davidii* branch *(*$\omega = 0.63$; $p=0$) were scored, together with the *Panthera tigris* branch ($\omega = 0.30$; $p <0.0001$), the Primates ($\omega = 0.67$; $p <0.0001$) and Cetartiodactyla nodes ($\omega = 0.54$; $p <0.0001$) (Figure 2). Similarly, for the TLR 9 tree the *Myotis (*$\omega = 1.6$; $p<0.0001$) and *Vespertilionidae* nodes *(*$\omega = 0.26$; $p<0.0001$) as well as the *D. rotundus* branch *(*$\omega = 0.42$; $p=0.1$) were detected for episodic positive selection (Figure 2). Additionally, the Laurasiatheria *(*$\omega = 0.28$;

p=0.01) and Perissodactyla nodes *(*ω = 0.37; p=0.1) together with the *Tursiops truncatus* branch *(*ω = 0.23; p=0.0) were also detected. In all analyses only the bat node and sub lineage branches displayed a mean ω >1 in comparison to the rest of the mammalian taxa; nonetheless, a mean ω <1 does not necessarily indicate that episodic positive selection did not occur as the ω estimate for the subset of sites under selection and the proportion of sites is considered for the analysis.

DISCUSSION

As observed for other mammalian nuclear genes, all four nucleic acid sensing TLRs showed comparable diversity levels and evolutionary rates (Leffler *et al.* 2012), while the resulting phylogenies were generally congruent with the species tree in which chiropterans were positioned closely to ungulates and carnivores (Agnarsson *et al.* 2011; Jones *et al.* 2005). Only the TLR 9 tree displayed an incongruent phylogenetic pattern likely the result of an accumulation of unique sites within the bat lineage. These 'unique sites' might have accumulated as a consequence of differential selective pressures exerted by long term pathogen-host interactions. While purifying selection was determined to be the main evolutionary force acting upon the bat receptors, a few PSS were detected within important domains of the proteins as described for other vertebrate species (Areal *et al.* 2011; Barreiro *et al.* 2009; Fornuskova *et al.* 2013; Lewis & Obbard 2014; Roach *et al.* 2005; Wlasiuk *et al.* 2009; Wlasiuk & Nachman 2010). Specifically, our analysis revealed a significant accumulation of PSS within the *Pteropus* node potentially representing family-specific adaptions to unique pathogen profiles not shared among other bats or mammals. In contrast with our negative results for ongoing positive selection under the branch-specific models, we found evidence for episodic diversifying selection acting upon both specific sites and branches within the bat lineage. Such results are not surprising, as natural selection affecting only a subset of lineages is often difficult to identify by branch-specific models as it can be masked by a large proportion of neutrally or negatively selected sites versus a small proportion of PSS found on the selected branches (Kosakovsky Pond *et al.* 2011; Murrell *et al.* 2012; Yang & dos Reis 2011). Thus, the evolution of the bat nucleic acid sensing TLRs may be characterized by general long-term purifying selection with bursts of episodic evolution, consistent with the idea of both transient and long-term pathogen-host interactions.

A number of bat-specific mutations fixed directly on LBS were also detected for the bat nucleic acid sensing TLR, most of them representing conservative changes between charged to polar amino acids or *vice versa*. However, mutation R331G within the TLR 3 was detected as a target of ongoing positive selection and involved a radical amino acid change between a basic/charged to a non-polar residue therefore possibly altering the electrostatic environment required for ligand binding (Choe *et al.* 2005). The bat TLR 9 also displayed a large number of

fixed mutations within LBS as compared to the other bat nucleic acid sensing TLRs, with mutations K181Q/E and K286S/N identified to have evolved under transient positive selection. Some of the mutations in TLR 9 were found to alter conformational DNA binding motifs previously described for the protein (motifs C-N-R-I-R/H-H, S-I/R-M-H-W-A/D/N and H-A-V/I/L-L) (Ohto *et al.* 2015). Though speculative, such structural changes could influence the ligand binding specificity for this receptor and should be further explored through functional analyses. Nonetheless, most of the bat-specific LBS mutations are unlikely to alter protein function and might have been fixed long ago as a result of order and genus and species-specific constraints. Our observations on the bat nucleic acid sensing TLRs could reflect the reservoir status described for some bat species carrying specific pathogens, such as *D. rotundus* as a host for rabies lyssaviruses and *Pteropus* species for Hendra- and Nipahviruses (Calisher *et al.* 2006; Werling *et al.* 2009). In the case of *D. rotundus*, this pattern could also relate to its exclusive adaptation to heamatophagy, as feeding on blood may pose an increased exposure to blood-borne pathogens which in turn may have represented a source of selective pressure on such receptors. However, there are many other immune molecules and physiological traits beyond immunity (such as generational time, social behavior, flight and metabolic trade-offs) that could also explain to some extent such 'pathogen tolerance phenotype'. Such is the case of the *flight-as-fever* hypothesis that postulates that the high body temperatures reached by bats during flight may stimulate the immune system, and that an increased oxidative stress can induce autophagy and apoptosis mechanisms contributing to intracellular pathogen clearance (Brook & Dobson 2015). We conclude that bat TLR evolution is consistent the general expectation of slow evolution. However, unique fixed mutations in sites of functional importance together with evidence for episodic selection specifically acting upon the bat lineage suggests that bats have unique adaptations in their nucleic acid sensing TLRs.

ACKNOWLEDGMENTS

The authors thank all Mexican and German colleagues involved in the collection and exportation of samples (Gabor Czirjak, Luis Caballero, Omar Ríos and Alfredo Patraca). Sample collection was approved by the Liebniz Institute for Zoo and Wildlife Research Internal Committee for Ethics and Animal Welfare (No. approval: 2012-09-05) meeting with Mexican regulations under the following permit and export certificate (Num/SGPA/DGVS: 03173/14; SAGARPA: 241111524599811488A467371). The authors thank Lisandra Zepeda-Mendoza, Tom Gilbert, Guojie Zhang and Mike Bunce (Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen) for granting access to their unpublished data. We also thank Alexandre Courtiol (IZW-Berlin), Romain Studer (European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire), Nikolas Nikolaidis (Department of Biological

Science and Center for Applied Biotechnology Studies, College of Natural Sciences and Mathematics, California State University Fullerton) and Tiandi Wei (Cancer Research Center, School of Medicine, Shandong University) for their valuable comments. This project was supported by the Deutsche Forschungsgemeinschaft (DFG) (GR 3924/9-1 to ADG) and an international doctoral scholarship provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) and the German Academic Exchange Service (DAAD) (311664 to M.EZ).

REFERENCES

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ (2011) A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). *PLoS Curr* **3**, RRN1212.

Alcaide M, Edwards SV (2011) Molecular evolution of the toll-like receptor multigene family in birds. *Mol Biol Evol* **28**, 1703-1715.

Areal H, Abrantes J, Esteves PJ (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evol Biol* **11**, 368.

Baker ML, Schountz T, Wang LF (2013) Antiviral immune responses of bats: a review. *Zoonoses Public Health* **60**, 104-116.

Barreiro LB, Ben-Ali M, Quach H*, et al.* (2009) Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS Genet* **5**, e1000562.

Bell JK, Askins J, Hall PR, Davies DR, Segal DM (2006) The dsRNA binding site of human Toll-like receptor 3. *Proc Natl Acad Sci U S A* **103**, 8792-8797.

Bell JK, Botos I, Hall PR*, et al.* (2005) The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc Natl Acad Sci U S A* **102**, 10976-10980.

Bell JK, Mullen GE, Leifer CA*, et al.* (2003) Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol* **24**, 528-533.

Benkert P, Tosatto SC, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* **71**, 261-277.

Biasini M, Bienert S, Waterhouse A*, et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252-258.

Boehme KW, Compton T (2004) Innate sensing of viruses by toll-like receptors. *J Virol* **78**, 7867-7873.

Botos I, Segal DM, Davies DR (2011) The structural biology of Toll-like receptors. *Structure* **19**, 447-459.

Brook CE, Dobson AP (2015) Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends Microbiol* **23**, 172-180.

Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev* **19**, 531-545.

Choe J, Kelker MS, Wilson IA (2005) Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* **309**, 581-585.

Cowled C, Baker M, Tachedjian M*, et al.* (2011) Molecular characterisation of Toll-like receptors in the black flying fox Pteropus alecto. *Dev Comp Immunol* **35**, 7-18.

de Bouteiller O, Merck E, Hasan UA*, et al.* (2005) Recognition of double-stranded RNA by human toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J Biol Chem* **280**, 38133-38145.

Delport W, Poon AF, Frost SD, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455-2457.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113.

Fornuskova A, Vinkler M, Pages M*, et al.* (2013) Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE). *BMC Evol Biol* **13**, 194.

Gorfol T, Csorba G, Eger JL, Son NT, Francis CM (2014) Canines make the difference: a new species of Hypsugo (Chiroptera: Vespertilionidae) from Laos and Vietnam. *Zootaxa* **3887**, 239-250.

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224.

Guindon S, Dufayard JF, Lefort V*, et al.* (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321.

He G, He B, Racey PA, Cui J (2010) Positive selection of the bat interferon alpha gene family. *Biochem Genet* **48**, 840-846.

Iha K, Omatsu T, Watanabe S*, et al.* (2010) Molecular cloning and expression analysis of bat toll-like receptors 3, 7 and 9. *J Vet Med Sci* **72**, 217-220.

Jones KE, Bininda-Emonds OR, Gittleman JL (2005) Bats, clocks, and rocks: diversification patterns in Chiroptera. *Evolution* **59**, 2243-2255.

Kepler TB, Sample C, Hudak K*, et al.* (2010) Chiropteran types I and II interferon genes inferred from genome sequencing traces by a statistical gene-family assembler. *BMC Genomics* **11**, 444.

Kosakovsky Pond SL, Murrell B, Fourment M*, et al.* (2011) A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* **28**, 3033-3043.

Kosakovsky Pond SL PD, Gravenor MB, Woelk CH, Frost SD (2006) Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol.* **23**, 1891-1901.

Koubinova D, Irwin N, Hulva P, Koubek P, Zima J (2013) Hidden diversity in Senegalese bats and associated findings in the systematics of the family Vespertilionidae. *Front Zool* **10**, 48.

Kumar S, Subramanian S (2002) Mutation rates in mammalian genomes. *Proc Natl Acad Sci U S A* **99**, 803-808.

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359.

Leffler EM, Bullaughey K, Matute DR*, et al.* (2012) Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biol* **10**, e1001388.

Lewis SH, Obbard DJ (2014) Recent insights into the evolution of innate viral sensing in animals. *Curr Opin Microbiol* **20**, 170-175.

Murrell B, Wertheim JO, Moola S*, et al.* (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* **8**, e1002764.

NCBI (2015) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **42**, D7-17.

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. *Mol Biol Evol* **6**, 290-300.

Nery MF, Gonzalez DJ, Hoffmann FG, Opazo JC (2012) Resolution of the laurasiatherian phylogeny: evidence from genomic data. *Mol Phylogenet Evol* **64**, 685-689.

Ohto U, Shibata T, Tanji H*, et al.* (2015) Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature*.

Omatsu T, Bak EJ, Ishii Y*, et al.* (2008) Induction and sequencing of Rousette bat interferon alpha and beta genes. *Vet Immunol Immunopathol* **124**, 169-176.

Pan X, Yue J, Ding G*, et al.* (2012) Leucine-rich repeat 11 of Toll-like receptor 9 can tightly bind to CpG-containing oligodeoxynucleotides, and the positively charged residues are critical for the high affinity. *J Biol Chem* **287**, 30596-30609.

Patterson B, Willig M, Steevens R (2003) Bat Ecology. In: *Bat Ecology* (ed. MB KTaF), pp. 536–579. Univ. of Chicago Press, Chicago.
Pond SK, Muse SV (2005) Site-to-site variation of synonymous substitution rates. *Mol Biol Evol* **22**, 2375-2385.
Posada D (2006) ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. *Nucleic Acids Res* **34**, W700-703.
Roach JC, Glusman G, Rowen L*, et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proc Natl Acad Sci U S A* **102**, 9577-9582.
Sarkar SK, Chakravarty AK (1991) Analysis of immunocompetent cells in the bat, Pteropus giganteus: isolation and scanning electron microscopic characterization. *Dev Comp Immunol* **15**, 423-430.
Schroder NW, Schumann RR (2005) Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *Lancet Infect Dis* **5**, 156-164.
Seim I, Fang X, Xiong Z*, et al.* (2013) Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. *Nat Commun* **4**, 2212.
Simmons NB (2005) Order Chiroptera. In: *Mammal species of the world. A taxonomic and geographic reference* (ed. Wilson DE RDB), pp. 312-529. Johns Hopkins University Press.
Sun J, Duffy KE, Ranjith-Kumar CT*, et al.* (2006) Structural and functional analyses of the human Toll-like receptor 3. Role of glycosylation. *J Biol Chem* **281**, 11144-11151.
Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-1599.
Tanji H, Ohto U, Shibata T, Miyake K, Shimizu T (2013) Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science* **339**, 1426-1429.
Tschirren B, Andersson M, Scherman K, Westerdahl H, Raberg L (2012) Contrasting patterns of diversity and population differentiation at the innate immunity gene toll-like receptor 2 (TLR2) in two sympatric rodent species. *Evolution* **66**, 720-731.
Uematsu S, Akira S (2007) Toll-like receptors and Type I interferons. *J Biol Chem* **282**, 15319-15323.
Wei T, Gong J, Jamitzky F*, et al.* (2009) Homology modeling of human Toll-like receptors TLR7, 8, and 9 ligand-binding domains. *Protein Sci* **18**, 1684-1691.
Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ (2009) Variation matters: TLR structure and species-specific pathogen recognition. *Trends Immunol* **30**, 124-130.
Wlasiuk G, Khan S, Switzer WM, Nachman MW (2009) A history of recurrent positive selection at the toll-like receptor 5 in primates. *Mol Biol Evol* **26**, 937-949.
Wlasiuk G, Nachman MW (2010) Adaptation and constraint at Toll-like receptors in primates. *Mol Biol Evol* **27**, 2172-2186.
Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**, 1586-1591.
Yang Z, dos Reis M (2011) Statistical properties of the branch-site test of positive selection. *Mol Biol Evol* **28**, 1217-1228.
Zhang G, Cowled C, Shi Z*, et al.* (2013) Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* **339**, 456-460.
Zhou W, Li Y, Pan X*, et al.* (2013) Toll-like receptor 9 interaction with CpG ODN--an in silico analysis approach. *Theor Biol Med Model* **10**, 18.
Zhou X, Xu S, Xu J, et al. (2012) Phylogenomic analysis resolves the interordinal relationships and rapid diversification of the laurasiatherian mammals. Syst Biol 61, 150-164.

DATA ACCESSIBILITY

All sequences generated in this work were deposited in GenBank under the accession number listed in Supporting Information: Table S3, while sequences used for analysis are listed in Supporting Information: Table S1. All 3-D models are available in PBD format in File S1.

Alignments and trees in Fasta and Newick formats are available in File S2, while complete PAML results (mlc files) are available in File S3 (doi:10.5061/dryad.dq1j6). Finally, the LRT tables for the interpretation of the PAML results are available in File S4.

AUTHOR CONTRIBUTIONS

M.E.Z and A.D.G designed research; M.E.Z performed research, M.E.Z and M.L.Z.M analyzed data; E.L.R, E.R.A, M.L.M.O provided samples and C.F.A contributed to sample processing. M.E.Z and A.D.G wrote the paper, with comments from all authors.

TABLE 1. Positive selected sites detected in the bat TLR 3, 7, 8 and 9 sequences

TLR	Site	PAML/ Pr(w>1) [1]	REL/ PP [2]	Position Aln	Position Protein
TLR 3	177 R	0,990** (NEB)	0,987	188	LRR7
	256 M	0,988** (NEB)	0,983	267	LRR10
	321 G	0,990** (NEB)	0,965	332	LRR13/ LBS
	333 L	0,987** (NEB)	0,987	344	LRR13
	414 S	0,984** (NEB)	0,981	425	LRR16
	604 N	0,987** (NEB)	0,981	617	LRR24
TLR 7	391 Q	0,909* (NEB/BEB)	0,959	395	LRR10
TLR 8	104 S	0.961** (NEB/BEB)	--	107	LRR2
	146 K	0.956** (NEB/BEB)	--	149	LRR4
	363 R	0.971** (NEB/BEB)	0,969	368	LRR12
	410 K	0.973** (NEB/BEB)	0,966	415	LRR13
	496 G	0.951** (NEB/BEB)	0,959	506	LRR16
	666 C	0.995** (NEB/BEB)	0,994	676	LRR22
	678 L	0.960** (NEB/BEB)	--	688	LRR23
	701 S	0.980** (NEB/BEB)	0,978	711	LRR24
TLR 9	68 R	0.995** (NEB/BEB)	--	115	LRR2
	120 H	0.915* (BEB)	--	167	LRR3
	302 S	0.964** (NEB/BEB)	--	349	LRR10
	332 W	0.942* (NEB/BEB)	--	379	LRR11

[1] PSS detected by site-specific model M8 in PAML and scored by BEB and/or NEB (Pr(w>1): * P>90%, ** P>95%). For PSS scored by both NEB/BEB, only BEB values are shown.

[2] PSS detected by REL (GTR model) and scored with a PP >80.

TABLE 2. Mutations in the LBS of the bat nucleic acid sensing TLR

Site[1]	Position Aln	Mammals[2]	DR[3]	MD[3]	ML[3]	MB[3]	EF[3]	RL[3]	PA[3]	PV[3]	Protein Region	Pr(w>1) [4]
TLR3												
493	494	K/R/T	R	K	K	K	K	**Q**	**Q**	**Q**	LRR19	
331	332	R/H/W	R	H	R	R	**C**	**G**	**G**	**G**	LRR13/LBS	0.990** (NEB)
145	146	N/K/S	N	**D**	**D**	**D**	N	N	N	N	LRR5	
TLR8												
539	540	N/K/R	N	**H**	**H**	**H**	N	-	**S**	**S**	LRR17	
TLR9												
51	75	K	**T**	**L**	**T**	**T**	**T**	K	K	K	LRRNT	
76	102	H	**R**	**R**	**R**	**R**	**R**	H	H	H	LRR1	
181	207	K/M	K	**Q**	**Q**	**Q**	**E**	K	K	K	LRR5	0.979* (BEB)
286	318	K	**S**	K	**S**	**S**	**S**	**N**	K	K	LRR9	0.991** (BEB)
335	361	Q/R	**T**	**K**	**K**	**K**	R	Q	R	Q	LRR10	
364	390	L/V/T/R/K/A	T	R	R	R	R	**Q**	**Q**	**Q**	LRR11	

[1]As determined for human, mouse, horse and bovids (Choe *et al.* 2005; Bell *et al.* 2005; Bell *et al.* 2006; Wei *et al.* 2009; Zhou et al. 2013; Pan *et al.* 2012; Ohto *et al.* 2015; Tanji *et al.* 2013).

[2] Amino acids present in other mammalian taxa (Laurasiatheria and Euarchontoglires).

[3] Amino acids present in Chiropterans. DR= *D. rotundus,* MD= *M. davidii,* ML=*M. lucifugus,* MB=*M. branditii,* EF=*E. fuscus,* RL=*R. leschenaultii,* PA= *P. alecto,* PV= *P. vampyrus*. Unique mutations to bats are shown in bold letters.

[4] PSS detected by PAML under M8a/M8 model and scored by NEB for TLR 3. PSS detected by BSA and scored by BEB (Pr(w>1): * P>95%, ** P>99%) for TLR 9.

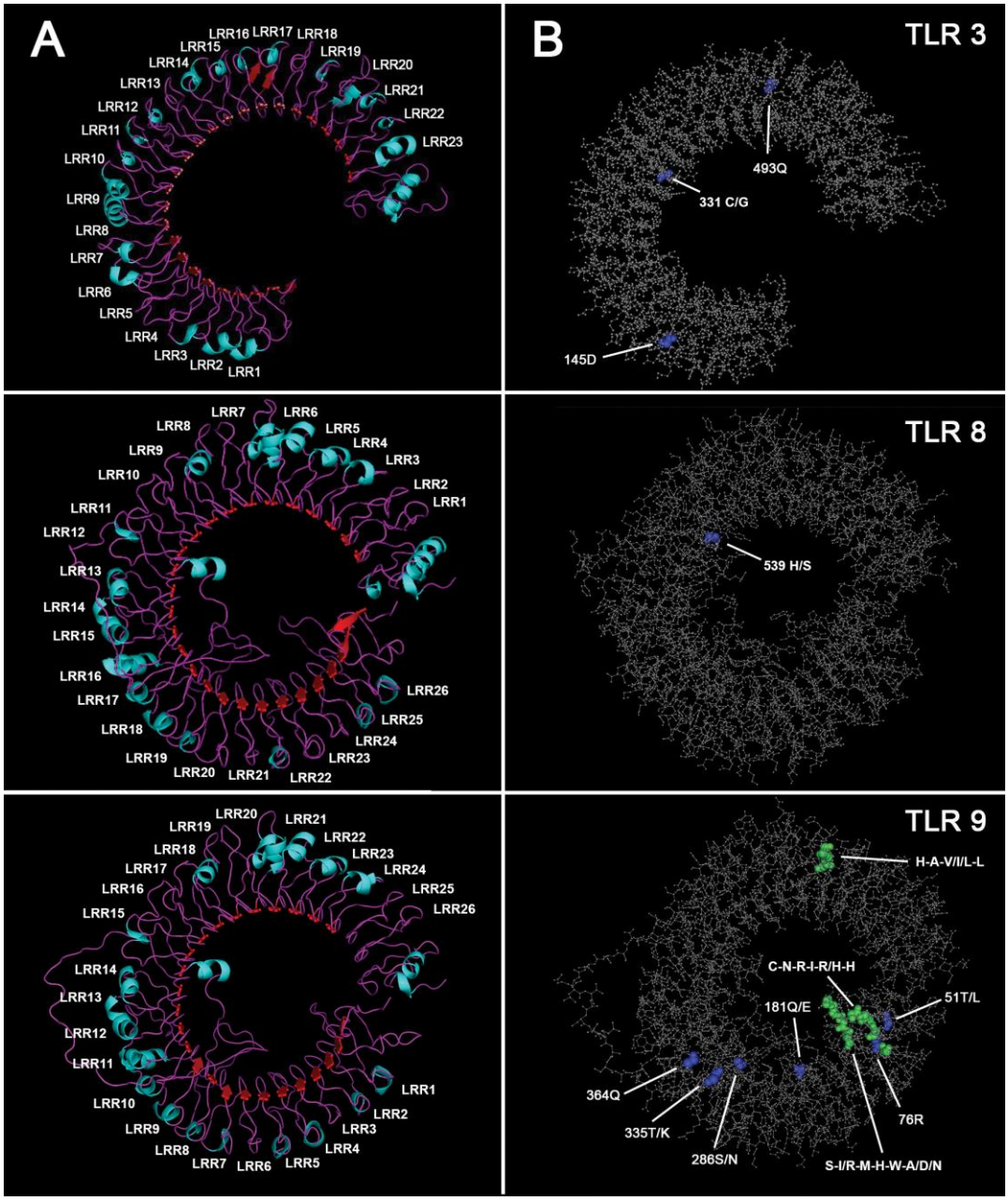

FIGURE 1

The structure of bat TLRs exhibits unique mutations fixed in ligand binding sites (LBS). (A) The 3-D protein models for the *D. rotundus* TLR 3, 8 and 9 ectodomains show the TLR typical horseshoe-shaped solenoid conformation with a concave and convex surface, and a descending and ascending lateral surface. Each LRR is characterized by the consensus sequence LxxLxLxxN/CxL followed by a number of hydrophobic residues forming a loop structure. Predicted LRRs are indicated with their corresponding position in the protein sequence. All models are shown as cartoon and colored by secondary structure, with helixes in blue, loops in purple and sheets in red. (B) The detected LBS mutations found in the bats TLR 3, 8 and 9. Models are shown as stick-and-ball with the protein back bone in gray. LBS mutations are highlighted under space fill in blue while altered binding motifs found on the secondary structure of TLR 9 are shown in green. All changes are indicated with the numbering corresponding to the original site in the protein sequences.

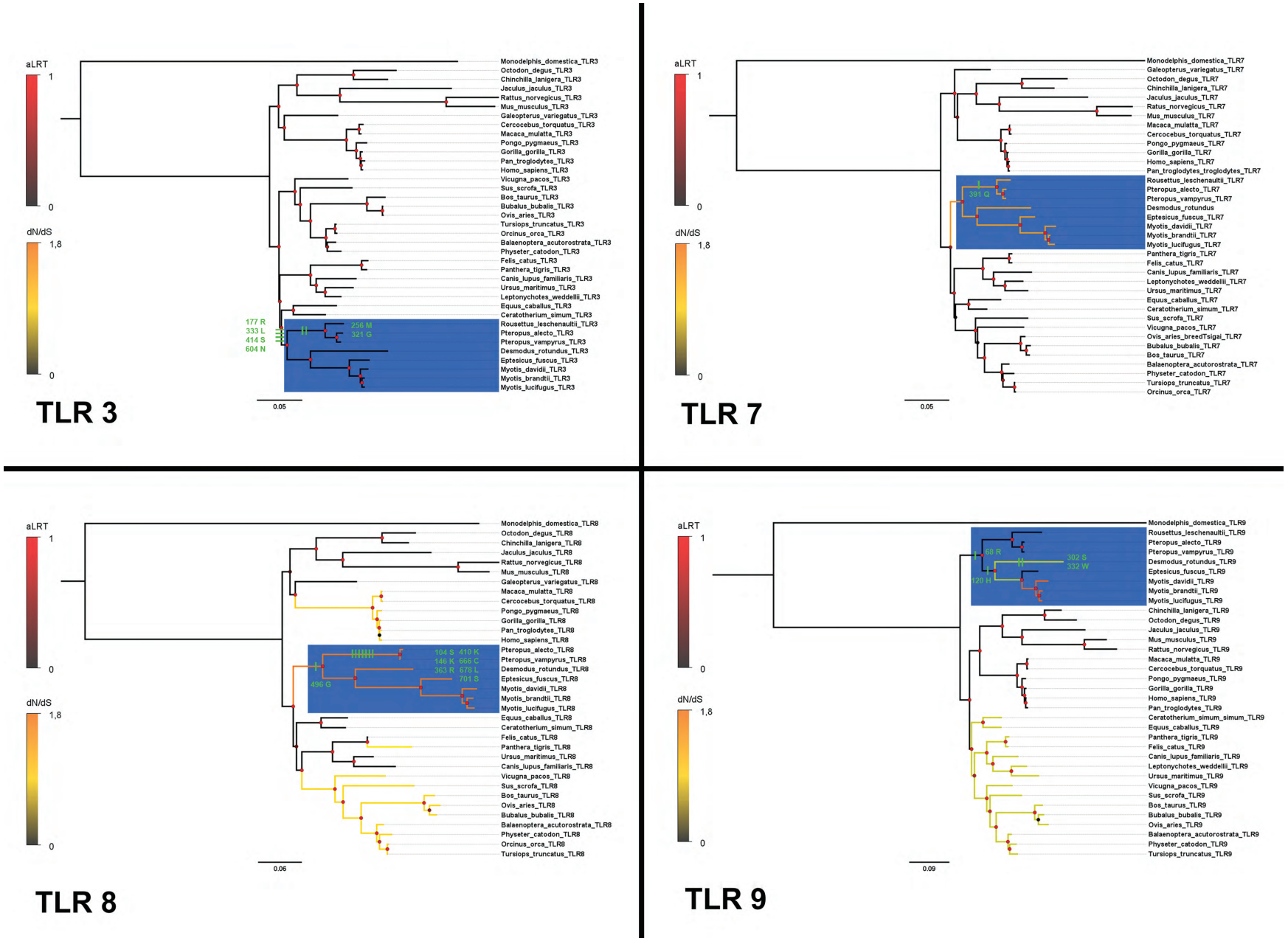
aLRT
1
0
dN/dS
1,8
0
Monodelphis_domestica_TLR3
Octodon_degus_TLR3
Chinchilla_lanigera_TLR3
Jaculus_jaculus_TLR3
Rattus_norvegicus_TLR3
Mus_musculus_TLR3
Galeopterus_variegatus_TLR3
Cercocebus_torquatus_TLR3
Macaca_mulatta_TLR3
Pongo_pygmaeus_TLR3
Gorilla_gorilla_TLR3
Pan_troglodytes_TLR3
Homo_sapiens_TLR3
Vicugna_pacos_TLR3
Sus_scrofa_TLR3
Bos_taurus_TLR3
Bubalus_bubalis_TLR3
Ovis_aries_TLR3
Tursiops_truncatus_TLR3
Orcinus_orca_TLR3
Balaenoptera_acutorostrata_TLR3
Physeter_catodon_TLR3
Felis_catus_TLR3
Panthera_tigris_TLR3
Canis_lupus_familiaris_TLR3
Ursus_maritimus_TLR3
Leptonychotes_weddellii_TLR3
Equus_caballus_TLR3
Ceratotherium_simum_TLR3
Rousettus_leschenaultii_TLR3
Pteropus_alecto_TLR3
Pteropus_vampyrus_TLR3
Desmodus_rotundus_TLR3
Eptesicus_fuscus_TLR3
Myotis_davidii_TLR3
Myotis_brandtii_TLR3
Myotis_lucifugus_TLR3
177 R
333 L
414 S
604 N
256 M
321 G
0.05
TLR 3
aLRT
1
0
dN/dS
1,8
0
Monodelphis_domestica_TLR7
Galeopterus_variegatus_TLR7
Octodon_degus_TLR7
Chinchilla_lanigera_TLR7
Jaculus_jaculus_TLR7
Ratus_norvegicus_TLR7
Mus_musculus_TLR7
Macaca_mulatta_TLR7
Cercocebus_torquatus_TLR7
Pongo_pygmaeus_TLR7
Gorilla_gorilla_TLR7
Homo_sapiens_TLR7
Pan_troglodytes_troglodytes_TLR7
Rousettus_leschenaultii_TLR7
Pteropus_alecto_TLR7
Pteropus_vampyrus_TLR7
Desmodus_rotundus
Eptesicus_fuscus_TLR7
Myotis_davidii_TLR7
Myotis_brandtii_TLR7
Myotis_lucifugus_TLR7
Panthera_tigris_TLR7
Felis_catus_TLR7
Canis_lupus_familiaris_TLR7
Leptonychotes_weddellii_TLR7
Ursus_maritimus_TLR7
Equus_caballus_TLR7
Ceratotherium_simum_TLR7
Sus_scrofa_TLR7
Vicugna_pacos_TLR7
Ovis_aries_breedTsigai_TLR7
Bubalus_bubalis_TLR7
Bos_taurus_TLR7
Balaenoptera_acutorostrata_TLR7
Physeter_catodon_TLR7
Tursiops_truncatus_TLR7
Orcinus_orca_TLR7
391 Q
0.05
TLR 7
aLRT
1
0
dN/dS
1,8
0
Monodelphis_domestica_TLR8
Octodon_degus_TLR8
Chinchilla_lanigera_TLR8
Jaculus_jaculus_TLR8
Rattus_norvegicus_TLR8
Mus_musculus_TLR8
Galeopterus_variegatus_TLR8
Macaca_mulatta_TLR8
Cercocebus_torquatus_TLR8
Pongo_pygmaeus_TLR8
Gorilla_gorilla_TLR8
Pan_troglodytes_TLR8
Homo_sapiens_TLR8
Pteropus_alecto_TLR8
Pteropus_vampyrus_TLR8
Desmodus_rotundus_TLR8
Eptesicus_fuscus_TLR8
Myotis_davidii_TLR8
Myotis_brandtii_TLR8
Myotis_lucifugus_TLR8
Equus_caballus_TLR8
Ceratotherium_simum_TLR8
Felis_catus_TLR8
Panthera_tigris_TLR8
Ursus_maritimus_TLR8
Canis_lupus_familiaris_TLR8
Vicugna_pacos_TLR8
Sus_scrofa_TLR8
Bos_taurus_TLR8
Ovis_aries_TLR8
Bubalus_bubalis_TLR8
Balaenoptera_acutorostrata_TLR8
Physeter_catodon_TLR8
Orcinus_orca_TLR8
Tursiops_truncatus_TLR8
104 S
410 K
146 K
666 C
363 R
678 L
701 S
496 G
0.06
TLR 8
aLRT
1
0
dN/dS
1,8
0
Monodelphis_domestica_TLR9
Rousettus_leschenaultii_TLR9
Pteropus_alecto_TLR9
Pteropus_vampyrus_TLR9
Desmodus_rotundus_TLR9
Eptesicus_fuscus_TLR9
Myotis_davidii_TLR9
Myotis_brandtii_TLR9
Myotis_lucifugus_TLR9
Chinchilla_lanigera_TLR9
Octodon_degus_TLR9
Jaculus_jaculus_TLR9
Mus_musculus_TLR9
Rattus_norvegicus_TLR9
Macaca_mulatta_TLR9
Cercocebus_torquatus_TLR9
Pongo_pygmaeus_TLR9
Gorilla_gorilla_TLR9
Homo_sapiens_TLR9
Pan_troglodytes_TLR9
Ceratotherium_simum_simum_TLR9
Equus_caballus_TLR9
Panthera_tigris_TLR9
Felis_catus_TLR9
Canis_lupus_familiaris_TLR9
Leptonychotes_weddellii_TLR9
Ursus_maritimus_TLR9
Vicugna_pacos_TLR9
Sus_scrofa_TLR9
Bos_taurus_TLR9
Bubalus_bubalis_TLR9
Ovis_aries_TLR9
Balaenoptera_acutorostrata_TLR9
Physeter_catodon_TLR9
Tursiops_truncatus_TLR9
68 R
128 H
302 S
332 W
0.09
TLR 9

FIGURE 2

Phylogenetic trees for the mammalian 3, 7, 8 and 9 TLRs. ML trees are shown with branches lengths scaled to amino acid substitutions per site, emphasizing the position of the bat lineage (in blue) with respect to other Larasiatherian (represented by Cetartiodactyla, Perissodactyla and Carnivora orders) and Euarchontoglires mammals (Primates and Rodentia orders). Support values are represented by circles within each node colored according to aLRT in a gradient from black to red, with black being the lowest and red being the highest. Positively selected sites (PSS) detected by the M8a/M8 site-specific model were mapped onto the topologies based on the reconstruction of ancestral states under a Maximum Likelihood frame. Each green bar represents one PSS with its corresponding amino acid state. The branches in each tree are colored according to the mean dN/dS (ω) estimates obtained by the Branch-site REL model to detect episodic diversifying selection under statistical significance of <0.05. Colored branches correspond to a gradient from black to orange, with black being the lowest (ω=0) and orange being the highest (ω>1). For TLR 7, 8 and 9, the bat lineage and/or sub-lineages were detected to have evolved under episodic diversifying selection.

Table S1. Complete coding sequences for the TLR 3, 7, 8 and 9 of *D. rotundus*

Species	Name	ORF lenght (bp)	Accession
D. rotundus	TLR 3	2712	KR349157
D. rotundus	TLR 7	3148	KR349160
D. rotundus	TLR 8	3123	KR349163
D. rotundus	TLR 9	3115	KR349164

Table S2. Templates used for the 3-D modeling of the *D. rotundus* TLR 3, 7, 8 and 9 [1]

TLR	Template	Identity	Oligostate	Resolution	Similarity	Range	Coverage	GMQE	QMEAN4	Description
TLR3	2a0z.1.A	81,6	monomer	2.40Å	0,5	27-695	0,8	0,7	-1.07	Human Toll-like Receptor 3
TLR7	3w3k.1.A	42,5	monomer	2.30Å	0,4	30-820	0,8	0,5	-6.67	Human Toll-like Receptor 8
TLR8	3w3g.1.A	69,0	monomer	2.30Å	0,5	30-818	0,8	0,7	-4.33	Human Toll-like Receptor 8
TLR9	3wpb.1.A	78,4	monomer	2.40Å	0,5	29 - 786	0,8	0,7	-4,61	Horse Toll-like receptor 9

[1] As determined by blastp. Models were built with ProMod Version 3.7

Table S3. Sequences and accession numbers used in this study

Species [1]	Accesion number [2]				Superorder	Order
	TLR3	TLR7	TLR8	TLR9		
Balaenoptera acutorostrata	XM_0071739781	XM_007180334	XM_007180333	XM_007174719	*Laurasiatheria*	*Artiodactyla*
Bos taurus	AY124007	NM_001033761.1	NM_001033937.1	NM_183081.1	*Laurasiatheria*	*Artiodactyla*
Bubalus bubalis	HQ3434171	NM_001290938.1	NM_001290928.1	HQ242779.1	*Laurasiatheria*	*Artiodactyla*
Canis lupus familiaris	XM_005629968	AB248956.1	XM_003435448	AB104899.3	*Laurasiatheria*	*Carnivora*
Ceratotherium simum	XM_004428766	XM_004435114	XM_004435115	XM_004419783	*Laurasiatheria*	*Perissodactyla*
Cercocebus torquatus	EU204935.1	EU204942.1	EU204945.1	EU204946.1	*Euarchontoglires*	*Primates*
Chinchilla lanigera	XM_005373438	XM_005411374	XM_005411373	XM_005410346	*Euarchontoglires*	*Rodentia*
Eptesicus fuscus	XM_008151907	XM_008156577	XM_008156576	XM_008155066	*Laurasiatheria*	*Chiroptera*
Equus caballus	NM_001081798	XM_005613953	NM_001111301	XM_005600476	*Laurasiatheria*	*Perissodactyla*
Felis catus	NM_001079829	NM_001080133	EF484949	NM_001009285	*Laurasiatheria*	*Carnivora*
Galeopterus variegatus	XM_008563573	XM_008571332	XM_00857134	NA	*Euarchontoglires*	*Dermoptera*
Gorilla gorilla	NM_001279752	KF321040	KF321277.1	KF321355	*Euarchontoglires*	*Primates*
Homo sapiens	AB4456311	AK313858	BC101075	NM_017442	*Euarchontoglires*	*Primates*
Jaculus jaculus	XM_00466637	XM_004663709.1	XM_004663710	XM_004664314	*Euarchontoglires*	*Rodentia*
Leptonychotes weddellii	XM_006741822	XM_006733443.1	XM_006752354	XM_006750986	*Laurasiatheria*	*Carnivora*
Macaca mulatta	NM_001036685	AB445665.1	AB445672.1	AB445679.1	*Euarchontoglires*	*Primates*
Mus musculus	XM_006509283	NM_001290755	XM_006528719	AF348140	*Euarchontoglires*	*Rodentia*
Myotis brandtii	XM_005863096	XM_005880946	XM_005880947	XM_005881923	*Laurasiatheria*	*Chiroptera*
Myotis davidii	XM_00677270	XM_006763796	XM_006763795	XM_006770629	*Laurasiatheria*	*Chiroptera*
Myotis lucifugus	XM_006092654.1	XM_006088607	XM_006088606	XM_006106454	*Laurasiatheria*	*Chiroptera*
Octodon degus	XM_004639797	XM_004635044	XM_004635045	XM_004625260.1	*Euarchontoglires*	*Rodentia*
Orcinus orca	XM_004277192	XM_004284914	XM_004284915	NA	*Laurasiatheria*	*Artiodactyla*
Ovis aries	NM_001135928	HQ529279	NM_001135929	AM981307	*Laurasiatheria*	*Artiodactyla*
Pan troglodytes	XM_009448543	KF321079	KF321298	KF321407	*Euarchontoglires*	*Primates*

Panthera tigris	XM_007098990	XM_007090078	XM_007099101.1	XM_007081936.1	*Laurasiatheria*	*Carnivora*
Physeter catodon	XM_007126456	XM_007121207	XM_007121208	XM_007105933	*Laurasiatheria*	*Artiodactyla*
Pongo pygmaeus	AB445635.1	AB445663	AB445670	AB44567	*Euarchontoglires*	*Primates*
Pteropus alecto	NM_001290169	NM_001290164	NM_001290163	GU045608	*Laurasiatheria*	*Chiroptera*
Pteropus vampyrus	XM_011363986.1	XM_011362910.1	XM_011362911.1	XM_011364642.1	*Laurasiatheria*	*Chiroptera*
Rattus norvegicus	NM_198791	XM_006256842	NM_001101009	NM_198131	*Euarchontoglires*	*Rodentia*
Rousettus leschenaultii	AB472355	AB472356	NA	AB472357	*Laurasiatheria*	*Chiroptera*
Sus scrofa	HQ4127961	NM_001097434	GU936184	XM_005669565	*Laurasiatheria*	*Artiodactyla*
Tursiops truncatus	XM_004321214	XM_004317713	XM_004317714.1	XM_004313247.1	*Laurasiatheria*	*Artiodactyla*
Ursus maritimus	XM_0087104651	XM_008696372.1	XM_008696373	XM_008704100.1	*Laurasiatheria*	*Carnivora*
Vicugna pacos	XM_006198002.1	XM_006212620	XM_006212621	XM_006220180.1	*Laurasiatheria*	*Artiodactyla*
Monodelphis domestica	XM_007496035.1	XM_007500926.1	XM_007500927.1	XM_007500513.1	*Marsupialia*	*Didelphimorphia*

[1] Names in bold letters indicate bats species

[2] NA= sequence not availllable in databases

File S4: LRT tables for the interpretation of the PAML results

M1/M2

Sequences/Test[a]	lnL	ΔLRT	2× (lnL1 - lnL0)	Parameters (np)	(np1 - np0)	ω (dN/dS)	Prob. density	Cum. distribution	1-P (<0.05)	Result
TLR3										
0 (M1)	-22958,9027			73		1,00				
1 (M2)	-22818,9139		279,978	74	1	0,25	0,00	1,000	0,000	Significant
TLR7										
0 (M1)	-26017,6175			73		1,00				
1 (M2)	-25883,1138		269,007	74	1	0,29	0,00	1,000	0,000	Significant
TLR8										
0 (M1)	-7215,7526			69		1,00				
1 (M2)	-7129,9506		171,604	70	1	0,16	0,00	1,000	0,000	Significant
TLR9										
0 (M1)	-22509,2476			69		1,00				
1 (M2)	-22392,4165		233,662	70	1	0,27	0,00	1,000	0,000	Significant

[a] Test applied to the complete TLR mammalian sequence alignment constraining the bat node

M8a/M8

Sequences/Test[a]	lnL	ΔLRT	2× (lnL1 - lnL0)	Parameters (np)	(np1 - np0)	p1[b]	ω (dN/dS)	Prob. density	Cum. distribution	1-P (<0.05)	Result
TLR3											
0 (M8a)	-6892,8323			17							
1 (M8)	-6755,3189		275,027	18	1	0,191	1,06	0,00	1,000	0,000	Significant
TLR7											
0 (M8a)	-7943,7432			17							
1 (M8)	-7905,0894		77,308	18	1	0,011	3,68	0,00	1,000	0,000	Significant
TLR8											
0 (M8a)	-9452,5549			15							
1 (M8)	-9304,1043		296,901	16	1	0,100	2,62	0,00	1,000	0,000	Significant
TLR9											
0 (M8a)	-6796,2737			17							
1 (M8)	-6787,7137		17,120	18	1	0,007	10,78	0,00	1,000	0,000	Significant

[a] Test applied only to the bat TLR sequence alignment

[b] p1 is the proportion of sites having ω > 1

BSAn/BSA

Sequences/ Test [a]	lnL	ΔLRT 2× (lnL1 - lnL0)	Parameters (np)	(np1 - np0)	p1[b]	Background ω (dN/dS)	Foreground ω (dN/dS)	Prob. density	Cum. distribution	1-P (<0.05)	Result
TLR3											
0 (w=1)	-22265,5191		75								
1 (BSA)	-22265,5191	0,000	76	1	0,140	0,08	1,00	∞	0,000	1,000	Not significant
TLR7											
0 (w=1)	-25081,4177		75								
1 (BSA)	-25081,4080	0,019	76	1	0,011	0,06	1,17	2,84	0,110	0,890	Not significant
TLR8											
0 (w=1)	-6971,4654		71								
1 (BSA)	-6971,4654	0,000	72	1	0,000	0,07	0,00	∞	0,000	1,000	Not significant
TLR9											
0 (w=1)	-21458,6784		71								
1 (BSA)	-21434,3713	48,614	72	1	0,090	0,05	1,30	0,00	1,000	0,000	Significant

[a] Test applied to the complete TLR mammalian sequence alignment constraining the bat node

[b] p1 is the proportion of sites under class 2a: positive selection in the foreground branches (bat node) and purifying selection for the rest of the tree (background branches).

Chapter 3: General Discussion

The retroviruses described in this work showed a complex evolution and evidence for cross-species transmission events among bats and different mammalian taxa. We determined that the *D. rotundus* endogenous retrovirus (DrERV) is an 'old retrovirus' that must have invaded the common vampire bat species at least 24 mya and was not found to be related to other bat retroviruses, but rather to retroviruses from rodents and New World primates. Although it is unlikely that *D. rotundus* is the original host for this retroviral group, the integrity of the genome of the related viruses present in rodents and monkeys suggests that this group of ERVs was 'recently' active and transmissible. Such evolutionary patterns have been observed for other retroviruses in which cross-species transmission has led to host-switching and adaptation (Hayward *et al.* 2013a). For example, some ERVs found in South East Asian species of mice are apparently ancestral to the GALV-related exogenous retrovirus (KoRV; Koala retrovirus) found currently circulating in captive gibbons and koalas (Callahan *et al.* 1977; Tarlinton *et al.* 2006). Thus, retroviral horizontal transmission may not only reflect a long history of spill-over events shaped by the host, virus and other ecological factors but may also have deeper implications in the evolution of the host's genomes as a significant portion of vertebrate genomic sequences have originated from retroviral-like elements (Jern & Coffin 2008).Further directions for this work are to determine the original reservoir for this retroviral lineage by sampling animals co-habiting with *D. rotundus,* such as rodent and other primate species, and to determine if there is an exogenous viral form still in circulation.

On the other hand, the bat TLR evolution was found to be consistent with the slow evolution generally observed in the receptors of other mammalian species, as all four TLR studied showed low diversity levels and relatively slow evolutionary rates comparable to other mammalian nuclear genes (Roach *et al.* 2005). However, the long-term adaptation of bats to specific environments and ecological niches with unique pathogen profiles is likely to have shaped the evolution of the bat nucleic acid sensing TLRs to some extent. We found evidence of diversifying selection acting specifically upon the bat lineage and detected unique fixed mutations in ligand binding sites of the bat TLRs. Some of these mutations were also determined to be targeted by ongoing and episodic diversifying positive selection and thus

may possibly modify protein function. The observed evolutionary pattern could be a reflection of the ‘viral tolerance phenotype’ observed in bats. However, such ‘tolerant phenotypes’ cannot be exclusively attributed to the variation within immune molecules, and it is likely a result of a combination of traits such as a long-life span, population size, behavior, flight and metabolic tradeoffs (Brook & Dobson 2015). Future perspectives for this work would be to do receptor expression assays to test if there are functional differences between the bat and other mammalian nucleic acid sensing TLRs and further evaluate the effects of site-directed mutagenesis of the previously detected PSS and LBS on protein function.

Concluding remarks

Bats have evolved over many millions of years and many of the viruses that they host are likely to have evolved with them. Despite the clear co-evolutionary pattern observed for some of the viral groups detected in bats, many other 'bat' viruses have been shown to have complex cross species transmission dynamics among other mammalian taxa in both a recent and long-term evolutionary timescales. Such is the case of the endogenous retrovirus (DrERV) described in this work, whose evolutionary pattern suggests historic multiple cross-species transmissions among vampire bats and other mammals such as rodents and primates. It has been suggested that bats may host more zoonotic viruses per species than other species considered to be natural reservoirs for viral zoonosis (Luis *et al.* 2013). Although such hypothesis is to be further tested, the reason behind why bats may carry so many viruses remains unsolved. In the case of the vampire bat species, it is possible that their unique adaptation to sanguinivory has played an important role in the transmission of viruses to other mammalian species. For other bat species, a direct association between specific bat adaptations and spill-over events is less obvious, but traits such as flight could play an important role in the dissemination of viruses. Although the answer to why bats are able to host and transmit many different viruses remains unanswered, it is believed that both physiological and immunological factors may contribute to the persistence of viruses within different bat species. However, despite being one of the most species-rich and abundant group of mammals, bats are one of the least studied animal groups in the area of immunology. The study of unique immunological features of bats in an evolutionary context may provide the basis for understanding why bats are able to carry and tolerate viruses often lethal to other mammalian species. Such knowledge can also provide the necessary basis for effective preventive strategies and contribute to the de-mystification of bats as disease-carrying pests.

References

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ (2011) A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). *PLoS Curr* **3**, RRN1212.

Alcaide M, Edwards SV (2011) Molecular evolution of the toll-like receptor multigene family in birds. *Mol Biol Evol* **28**, 1703-1715.

Baker ML, Schountz T, Wang LF (2013) Antiviral immune responses of bats: a review. *Zoonoses Public Health* **60**, 104-116.

Belotto A, Leanes LF, Schneider MC, Tamayo H, Correa E (2005) Overview of rabies in the Americas. *Virus Res* **111**, 5-12.

Bobrowiec PED, Maristerra R, Lemes RG (2015) Prey preference of the common vampire bat (Desmodus rotundus, Chiroptera) using molecular analysis. *Journal of Mammalogy* **96** 54-63.

Boehme KW, Compton T (2004) Innate sensing of viruses by toll-like receptors. *J Virol* **78**, 7867-7873.

Brook CE, Dobson AP (2015) Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends Microbiol* **23**, 172-180.

Burns LV, Hutzley; Z, Laubach (2014) *Phyllostomidae.* Animal Diversity Web. http://animaldiversity.ummz.umich.edu/accounts/Phyllostomidae/

Cadieux C, Kedinger V, Yao L*, et al.* (2009) Mouse mammary tumor virus p75 and p110 CUX1 transgenic mice develop mammary tumors of various histologic types. *Cancer Res* **69**, 7188-7197.

Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: important reservoir hosts of emerging viruses. *Clin Microbiol Rev* **19**, 531-545.

Callahan R, Sherr CJ, Todaro GJ (1977) A new class of murine retroviruses: immunological and biochemical comparison of novel isolates from Mus cervicolor and Mus caroli. *Virology* **80**, 401-406.

Cowled C, Baker M, Tachedjian M*, et al.* (2011) Molecular characterisation of Toll-like receptors in the black flying fox Pteropus alecto. *Dev Comp Immunol* **35**, 7-18.

Cui J, Tachedjian G, Tachedjian M*, et al.* (2012a) Identification of diverse groups of endogenous gammaretroviruses in mega- and microbats. *J Gen Virol* **93**, 2037-2045.

Cui J, Tachedjian M, Wang L*, et al.* (2012b) Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. *J Virol* **86**, 4288-4293.

Datzmann T, von Helversen O, Mayer F (2010) Evolution of nectarivory in phyllostomid bats (Phyllostomidae Gray, 1825, Chiroptera: Mammalia). *BMC Evol Biol* **10**, 165.

Davis JS, Nicolay CW, Williams SH (2010) A comparative study of incisor procumbency and mandibular morphology in vampire bats. *J Morphol* **271**, 853-862.

Dewannieux M, Heidmann T (2013) Endogenous retroviruses: acquisition, amplification and taming of genome invaders. *Curr Opin Virol* **3**, 646-656.

Drexler JF, Corman VM, Muller MA*, et al.* (2012a) Bats host major mammalian paramyxoviruses. *Nat Commun* **3**, 796.

Drexler JF, Seelen A, Corman VM*, et al.* (2012b) Bats worldwide carry hepatitis E virus-related viruses that form a putative novel genus within the family Hepeviridae. *J Virol* **86**, 9134-9147.

Engel ME, Hiebert SW (2010) The enemy within: dormant retroviruses awaken. *Nat Med* **16**, 517-518.

Forment WL, Schmidt U, Greenhall AM (1971) Movement and population studies of the vampire bat (Desmodus rotundus) in Mexico. *J Mammal* **52**, 227-228.

Francischetti IM, Assumpcao TC, Ma D*, et al.* (2013) The "Vampirome": Transcriptome and proteome analysis of the principal and accessory submaxillary glands of the vampire bat Desmodus rotundus, a vector of human rabies. *J Proteomics* **82**, 288-319.

Frost SD, Volz EM (2013) Modelling tree shape and structure in viral phylodynamics. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120208.

Gorfol T, Csorba G, Eger JL, Son NT, Francis CM (2014) Canines make the difference: a new species of Hypsugo (Chiroptera: Vespertilionidae) from Laos and Vietnam. *Zootaxa* **3887**, 239-250.

Greenhall AM, Joermann G, Schmidt U (1983) Desmodus rotundus. *Mammalian Species* **202**, 1–6.

Grenfell BT, Pybus OG, Gog JR*, et al.* (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327-332.

Hayman DT, Bowen RA, Cryan PM*, et al.* (2013) Ecology of zoonotic infectious diseases in bats: current knowledge and future directions. *Zoonoses Public Health* **60**, 2-21.

Hayward A, Cornwallis CK, Jern P (2015a) Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. *Proc Natl Acad Sci U S A* **112**, 464-469.

Hayward A, Grabherr M, Jern P (2013b) Broad-scale phylogenomics provides insights into retrovirus-host evolution. *Proc Natl Acad Sci U S A* **110**, 20146-20151.

Hayward A, Tachedjian M, Cui J*, et al.* (2013c) Identification of diverse full-length endogenous betaretroviruses in megabats and microbats. *Retrovirology* **10**, 35.

He G, He B, Racey PA, Cui J (2010) Positive selection of the bat interferon alpha gene family. *Biochem Genet* **48**, 840-846.

Holland RA, Waters DA, Rayner JM (2004) Echolocation signal structure in the Megachiropteran bat Rousettus aegyptiacus Geoffroy 1810. *J Exp Biol* **207**, 4361-4369.

Iha K, Omatsu T, Watanabe S*, et al.* (2010) Molecular cloning and expression analysis of bat toll-like receptors 3, 7 and 9. *J Vet Med Sci* **72**, 217-220.

Jern P, Coffin JM (2008) Effects of retroviruses on host genome function. *Annu Rev Genet* **42**, 709-732.

Jones KE, Bininda-Emonds OR, Gittleman JL (2005) Bats, clocks, and rocks: diversification patterns in Chiroptera. *Evolution* **59**, 2243-2255.

Kepler TB, Sample C, Hudak K*, et al.* (2010) Chiropteran types I and II interferon genes inferred from genome sequencing traces by a statistical gene-family assembler. *BMC Genomics* **11**, 444.

Kosakovsky Pond SL, Murrell B, Fourment M*, et al.* (2011) A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* **28**, 3033-3043.

Koubinova D, Irwin N, Hulva P, Koubek P, Zima J (2013) Hidden diversity in Senegalese bats and associated findings in the systematics of the family Vespertilionidae. *Front Zool* **10**, 48.

Kunz TH, Braun de Torrez E, Bauer D, Lobova T, Fleming TH (2011) Ecosystem services provided by bats. *Ann N Y Acad Sci* **1223**, 1-38.

Kurth R, Bannert N (2010) *Retroviruses: Molecular Biology, Genomics and Pathogenesis* Caister Academic Press.

Luis AD, Hayman DT, O'Shea TJ*, et al.* (2013) A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? *Proc Biol Sci* **280**, 20122753.

Magiorkinis G, Sypsa V, Magiorkinis E*, et al.* (2013) Integrating phylodynamics and epidemiology to estimate transmission diversity in viral epidemics. *PLoS Comput Biol* **9**, e1002876.

Murphy WJ, Eizirik E, Johnson WE*, et al.* (2001) Molecular phylogenetics and the origins of placental mammals. *Nature* **409**, 614-618.

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. *Mol Biol Evol* **6**, 290-300.

Nery MF, Gonzalez DJ, Hoffmann FG, Opazo JC (2012) Resolution of the laurasiatherian phylogeny: evidence from genomic data. *Mol Phylogenet Evol* **64**, 685-689.

Omatsu T, Bak EJ, Ishii Y*, et al.* (2008) Induction and sequencing of Rousette bat interferon alpha and beta genes. *Vet Immunol Immunopathol* **124**, 169-176.

Patterson B, Willig M, Steevens R (2003) Bat Ecology. In: *Bat Ecology* (ed. MB KTaF), pp. 536–579. Univ. of Chicago Press, Chicago.

Penny D, Hendy MD, Steel MA, Miyamoto MM, Cracraft J (1991) Phylogenetic Analysis of DNA Sequences, pp. 155–183. Oxford University Press, United Kingdom.

Phillips CD, Baker RJ (2015) Secretory gene recruitments in vampire bat salivary adaptation and potential convergences with sanguivorous leeches. . *Front. Ecol. Evol.* **3**.

Pond SK, Muse SV (2005) Site-to-site variation of synonymous substitution rates. *Mol Biol Evol* **22**, 2375-2385.

Roach JC, Glusman G, Rowen L*, et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proc Natl Acad Sci U S A* **102**, 9577-9582.

Sarkar SK, Chakravarty AK (1991) Analysis of immunocompetent cells in the bat, Pteropus giganteus: isolation and scanning electron microscopic characterization. *Dev Comp Immunol* **15**, 423-430.

Schroder NW, Schumann RR (2005) Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *Lancet Infect Dis* **5**, 156-164.

Simmons NB (2005) Order Chiroptera. In: *Mammal species of the world. A taxonomic and geographic reference* (ed. Wilson DE RDB), pp. 312-529. Johns Hopkins University Press.

Sironi M, Cagliani R, Forni D, Clerici M (2015) Evolutionary insights into host-pathogen interactions from mammalian sequence data. *Nat Rev Genet* **16**, 224-236.

Tarlinton RE, Meers J, Young PR (2006) Retroviral invasion of the koala genome. *Nature* **442**, 79-81.

Teeling EC, Scally M, Kao DJ*, et al.* (2000) Molecular evidence regarding the origin of echolocation and flight in bats. *Nature* **13**, 188-192.

Tschirren B, Andersson M, Scherman K, Westerdahl H, Raberg L (2012) Contrasting patterns of diversity and population differentiation at the innate immunity gene toll-like receptor 2 (TLR2) in two sympatric rodent species. *Evolution* **66**, 720-731.

Volz EM, Koelle K, Bedford T (2013) Viral phylodynamics. *PLoS Comput Biol* **9**, e1002947.

Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ (2009) Variation matters: TLR structure and species-specific pathogen recognition. *Trends Immunol* **30**, 124-130.

Wlasiuk G, Nachman MW (2010) Adaptation and constraint at Toll-like receptors in primates. *Mol Biol Evol* **27**, 2172-2186.

Yang Z, dos Reis M (2011) Statistical properties of the branch-site test of positive selection. *Mol Biol Evol* **28**, 1217-1228.

Zhou X, Xu S, Xu J*, et al.* (2012) Phylogenomic analysis resolves the interordinal relationships and rapid diversification of the laurasiatherian mammals. *Syst Biol* **61**, 150-164.

Zhuo X, Rho M, Feschotte C (2013) Genome-wide characterization of endogenous retroviruses in the bat Myotis lucifugus reveals recent and diverse infections. *J Virol* **87**, 8493-8501.

Summary

The adaptation of bats to different environments has resulted in the evolution of unique phenotypic and genotypic characteristics such as flight, echolocation and highly specialized diets. Bats have also been increasingly recognized as reservoirs for viruses which can cross species barriers. Among the Neotropical bats, vampire bats (*Desmodontinae* subfamily) are the only mammals that feed exclusively on the blood from other animals. Because of such an exceptional adaptation to hematophagy these species exhibit a unique set of behavioral, physiological, and morphological characteristics distinct among all other bats. The common vampire bat (*Desmodus rotundus*) is a known reservoir for rabies-causing lyssaviruses considered to be a major constraint on the cattle industry since bat-transmitted rabies is a primary problem in livestock from Latin America. Nonetheless, the presence of other viruses in this species has hardly been explored. Although bats can be persistently infected with many viruses, they rarely display clinical symptoms and it has been suggested that they might have evolved specific immune strategies to control viral replication. Toll like-receptors (TLRs) are a class of innate immune receptors considered to be the first-line defense mechanism against invading pathogens. The mammalian TLRs 3, 7, 8 and 9 play an important role in triggering acquired immunity as they are activated by nucleic acid ligands. TLRs are of interest from an evolutionary point of view since there is evidence that the ligand binding properties of these receptors may vary among different species thereby having an impact on the evolutionary ecology of infectious diseases. Thus, the analysis of the bat immune variation at a molecular level could reveal patterns of resistance or susceptibility to pathogens within different species and at different taxonomic levels. However, the study of the genetic variability of the immune system in bats has been restricted to a few species and to a few genes.

The hypothesis that vampire bats carry viruses common to other mammals was tested based on the premise that their exclusive adaptation to hematophagy could have resulted in viral spill-over events among taxa throughout evolutionary history. Particular focus was made on Retroviruses, given that this viral group is primarily transmitted via body fluid exchange, and thus might have been particularly prone to jump between vampire bats and other taxa. A novel endogenous betaretrovirus (DrERV) was described with an evolutionary pattern that

suggests multiple cross-species transmissions among different species throughout their evolutionary history. It was further hypothesized that given the unique adaptations within the Chiroptera, bats as a taxonomic group would have acquired distinctive mutations fixed within the nucleic-acid sensing TLRs with potential consequences on their ligand recognition properties. The nucleic acid sensing TLRs (3, 7, 8 and 9) of the common vampire bat were characterized and the genetic variation of these receptors within different bats species and among other mammals was compared by further testing for ongoing and episodic diversifying selection acting upon specific lineages. Our results provide evidence for potential functional differences between the bat and other mammalian TLRs in terms of recognition of foreign nucleic acids. This project was carried out in close collaboration with several European and Mexican institutions contributing to the development of research on emerging zoonotic diseases and wildlife surveillance

Zusammenfassung “Vampir Fledermaus Virom: Evolutions Auswirkungen in einem immunologischen Kontext„

Anpassungen von Fledermäusen an verschiedene Umgebungen haben zu der Entwicklung Die Anpassung von Fledermäusen an ihre Umwelt hat zu der Entwicklung von besonderen phänotypischen und genotypischen Eigenschaften wie Flugfähigkeit, Echoortung und spezifischen Ernährungsweisen geführt. Diese Anpassungen haben auch zur evolutionären Entwicklung potentieller Krankheitserreger beigetragen. So werden Fledermäuse zunehmend als Wirte für Viren erkannt, welche auch interspezifisch übertragen werden können. Unter den Fledermäusen der Neuen Welt sind die Vampirfledermäuse (Unterfamilie *Desmodontinae*) als die einzigen Säugetiere bekannt, die sich ausschließlich vom Blut anderer Tiere ernähren können. Wegen dieser ungewöhnlichen Anpassung an Hämatophagie, zeigen diese Arten besondere Eigenschaften in ihrem Verhalten, ihrer Physiologie und ihrer Morphologie, die sie von allen anderen Fledermausarten unterscheidet. Der Gemeine Vampir (*Desmodus rotundus*) ist ein bekannter Wirt für Tollwut-verursachende Lyssaviren und stellt eine große Einschränkung für die Rinderhaltung in Lateinamerika dar. Der Einfluss anderer Virentypen in dieser Art ist bisher kaum erforscht.

Obwohl Fledermäuse dauerhaft von vielen Viren infiziert sein können, zeigen sie selten klinische Symptome. Es wird deshalb angenommen dass sie spezifische Immunstrategien entwickelt haben um die Virusreplikation zu kontrollieren. Die *Toll like-receptors* (TLRs) gehören zu den Immunrezeptoren, die als Bestandteil der primäreren Abwehrantwort gegen eindringende Pathogene angesehen werden. Abhängig von der Tierart unterscheidet man bis zu 13 verschiedene TLRs. Insbesondere die Nukleinsäure-bindenden TLRs 3, 7, 8 und 9 spielen eine wichtige Rolle in der erworbenen Immunreaktion, weil sie diese durch Liganden in Form von Nukleinsäuren aktivieren. Die TLRs sind aus evolutionsbiologischer Sicht besonders interessant, da die Bindungseigenschaften dieser Rezeptoren unter verschiedenen Tierarten variieren können und deshalb einen Einfluss auf die Evolutionsökologie von Infektionskrankheiten haben. Daher könnte die Analyse der Immunvariation bei Fledermäusen auf molekularer Ebene Resistenz- oder Anfälligkeitsmuster gegenüber Pathogenen innerhalb verschiedener Arten und auf unterschiedlichen

taxonomischen Ebenen offenlegen. Bisher wurde die genetische Variabilität des Immunsystems von Fledermäusen nur bei wenigen Arten und Genen erforscht. Unserer Hypothese zur Folge könnten Vampirfledermäusen Viren in sich tragen, die auch in anderen Säugetieren vorkommen, da ihre Fähigkeit zur Hämatophagie zu einer Virusübertragung zwischen verschiedenen Taxa in der Evolutionsgeschichte geführt haben könnte. Beispielsweise werden Retroviren hauptsächlich über den Austausch von Körperflüssigkeiten übertragen, weswegen Viren dieser Gruppe leicht zwischen Vampirfledermäusen und Tieren andere Taxa übertragen werden können. Ziel unserer Arbeit war es, neue und bereits bekannte Retroviren in D. rotundus mit Hilfe molekularevolutionärer Methoden zu untersuchen, um interspezifische Übertragungen zu identifizieren. Wir konnten zeigen, dass ein neues endogenes Betaretrovirus (DrERV) in D. rotundus ein Evolutionsmuster aufweist, das auf mehrere interspezifische Übertragungen zwischen verschiedenen Taxa im Laufe der Stammesentwicklung hindeutet.

Außerdem sind wir davon ausgegangen dass sich in Fledermäuse, auf Grund ihrer Anpassungen, spezielle Mutationen in den Nukleinsäure-bindenden TLRs fixiert haben, welche wahrscheinlich Auswirkungen auf die Bindungseigenschaften von Liganden haben. Wir haben die Nukleinsäure-bindenden TLRs 3, 7, 8 und 9 des Gemeinen Vampirs grundlegend untersucht und die genetischen Variationen dieser Rezeptoren innerhalb verschiedener Fledermausarten und anderen Säugetieren verglichen. Wir untersuchten ob sich die diversifizierende Selektion fortlaufend oder punktuell auf die Fledermaus-TLRs ausgewirkt hat, um das evolutionäre Muster über lange und kurze Zeiträume beschreiben zu können. Anschließende Analysen der Evolution auf molekularer Ebene ließen spezielle Anpassungsmuster bei den TLRs verschiedener Fledermausarten erkennen. Unsere Ergebnisse konnten außerdem funktionelle Unterschiede in den Bindungseigenschaften von wirtsfremden Nucleinsäuren zwischen den TLRs von Fledermäusen und denen anderer Säugetiere zeigen und lieferten somit entscheidende Hinweise über die Anpassung von Fledermäuse an Viren als potentielle Krankheitserreger.

Dieses Projekt wurde in enger Zusammenarbeit mit einigen europäischen und mexikanischen Instituten durchgeführt, die zur Entwicklung von Forschungsbereichen wie Zoonosen und Wildtierüberwachung beitragen.

List of publications

IN PRESS

1. **Escalera-Zamudio M**, Mendoza MLZ, Loza-Rubio E, Rojas-Anaya E, Méndez-Ojeda ML, Arias CF and Greenwood AD. The evolution of bat nucleic acid sensing Toll-like receptors. Mol. Ecol. Manuscript MEC-15-0892 accepted 20.10.2015.
2. Espinosa-Cueto P, **Escalera-Zamudio M**, Magallanes-Puebla A, López-Marín LM, Segura-Salinas E, Mancilla R. Mycobacterial glycolipids di-O-acylated trehalose and tri-O-acylated trehalose downregulate inducible nitric oxide synthase and nitric oxide production in macrophages. BMC Immunol. 2015 Jun 23;16:38. doi: 10.1186/s12865-015-0102-3
3. **Escalera-Zamudio M**, Mendoza MLZ, Heeger F, Loza-Rubio E, Rojas-Anaya E, Méndez-Ojeda ML, Taboada B, Mazzoni CJ, Arias CF, Greenwood AD. A novel endogenous betaretrovirus in the common vampire bat (*Desmodus rotundus*) suggests multiple independent infection and cross-species transmission events. J Virol. 2015 May;89(9):5180-4. doi: 10.1128/JVI.03452-14. Epub 2015 Feb 25.
4. **Escalera-Zamudio M**, Nelson MI, Cobián Güemes AG, López-Martínez I, Cruz-Ortiz N, Iguala-Vidales M, García ER, Barrera-Badillo G, Díaz-Quiñonez JA, López S, Arias CF, Isa P; Members of Colegio de Pediatría del Estado de Veracruz. Molecular epidemiology of influenza A/H3N2 viruses circulating in Mexico from 2003 to 2012. PLoS One. 2014 Jul 30;9(7):e102453. doi: 10.1371/journal.pone.0102453.
5. Vazquez-Perez JA, Isa P, Kobasa D, Ormsby CE, Ramírez-Gonzalez JE, Romero-Rodríguez DP, Ranadheera C, Li Y, Bastien N, Embury-Hyatt C, González-Duran E, Barrera-Badillo G, Ablanedo-Terrazas Y, Sevilla-Reyes EE, **Escalera-Zamudio M**, Cobián-Güemes AG, Lopez I, Ortiz-Alcántara J, Alpuche-Aranda C, Perez-Padilla JR, Reyes-Terán G. A (H1N1) pdm09 HA D222 variants associated with severity and mortality in patients during a second wave in Mexico. Virol J. 2013 Jan 31;10:41. doi: 10.1186/1743-422X-10-41.
6. **Escalera-Zamudio M**, Cobián-Güemes G, de los Dolores Soto-del Río M, Isa P, Sánchez-Betancourt I, Parissi-Crivelli A, Martínez-Cázares MT, Romero P, Velázquez-Salinas L, Huerta-Lozano B, Nelson M, Montero H, Vinuesa P, López S, Arias CF. Characterization of an influenza A virus in Mexican swine that is related to the A/H1N1/2009 pandemic clade. Virology. 2012 Nov 10;433(1):176-82. doi:10.1016/j.virol.2012.08.003.
7. **Escalera-Zamudio M** and Greenwood AD. On the classification and evolution of ERVs: human endogenous retroviruses may not be 'human' after all. 2015. APMIS special issue: Human endogenous retroviruses and related sequences. APMIS. Manuscript APMIS-2015-45- (3623) accepted October 2015.
8. Arias CF, **Escalera-Zamudio M**, Soto-Del Río Mde L, Cobián-Güemes AG, Isa P, López S. Molecular anatomy of 2009 influenza virus A (H1N1). Arch Med Res. 2009 Nov;40(8):643-54. doi: 10.1016/j.arcmed.2009.10.007.

Acknowledgements

This work was carried out thanks to the following funding: Deutsche Forschungsgemeinschaft (DFG, GR 3924/9-1) and International Doctoral Scholarship Mexico and Germany (CONACyT-DAAD: 311664).

We thank: Biol. Heiner Klös (Zoologischer Garten Berlin) and Biol. Carsten Kallasch (AG BUBO- Berlin) for providing samples from captive bats and all the Mexican and German colleagues involved in the collection and exportation of samples (Gabor Czirjak, Luis Caballero, Omar Ríos and Alfredo Patraca). Thanks to Marie-Lisandra Zepeda Mendoza, Tom Gilbert, Guojie Zhang and Mike Bunce (Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen) for granting access to their unpublished *D. rotundus* genomic data.

In general, I thank all my collaborators in Mexico (Centro Nacional de Investigación Disciplinaria en Microbiología Animal-INIFAP, Facultad de Medicina Veterinaria y Zootecnia, Universidad Veracruzana, Mexico and Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca Morelos) and at the Leibniz Institute for Zoo and Wildlife Research (IZW) for their support. Special thanks to Susan Mbedi, Kirsten Richter and Camilla Mazzoni for their support on the high-throughput sequencing and to Karin Hoenig, Nicole Dinse and Katja Pohle for their technical support. Finally, we thank Dr. Jens Mayer (Department of Human Genetics, Center of Human and Molecular Biology, Medical Faculty, University of Saarland, Homburg, Germany), Yasuko Ishida and Adam Brandt (University of Illinois, Urbana-Champaign USA), Alexandre Courtiol (IZW-Berlin), Romain Studer (European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire), Nikolas Nikolaidis (Department of Biological Science and Center for Applied Biotechnology Studies, College of Natural Sciences and Mathematics, California State University Fullerton) and Tiandi Wei (Cancer Research Center, School of Medicine, Shandong University) for their valuable comments.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Dissertation selbständig, ohne unzulässige fremde Hilfe und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

I hereby confirm that I have made this work autonomously. I assure that I have read and used only the specified sources claimed in this work.

Marina Escalera Zamudio

7 September 2016

Berlin, Germany