Large Numbers of Novel miRNAs Originate from DNA Transposons and Are Coincident with a Large Species Radiation in Bats

Roy N. Platt II,[1,2] Michael W. Vandewege,[1] Colin Kern,[3] Carl J. Schmidt,[4] Federico G. Hoffmann,[1,2,*] and David A. Ray[*,‡,1,2]

[1]Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University

[2]Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University

[3]Department of Computer and Information Sciences, University of Delaware

[4]Department of Animal and Food Sciences, University of Delaware

[‡]Present address: Department of Biological Sciences, Texas Tech University, Lubbock, TX

***Corresponding author:** E-mail: david.4.ray@gmail.com, federico.g.hoffmann@gmail.com.

Associate editor: Emma Teeling

Abstract

Vesper bats (family Vespertilionidae) experienced a rapid adaptive radiation beginning around 36 Ma that resulted in the second most species-rich mammalian family (>400 species). Coincident with that radiation was an initial burst of DNA transposon activity that has continued into the present in some species. Such extensive and recent DNA transposon activity has not been seen in any other extant mammal. Indeed, retrotransposon activity is much more common in all other sequenced mammal genomes. Deep sequencing of the small RNA fraction from a vespertilionid bat, *Eptesicus fuscus*, as well as a dog and horse revealed large numbers of 17–24 bp putative miRNAs (p/miRNAs). Although the origination rate of p/miRNAs is similar in all three taxa, 61.1% of postdivergence p/miRNAs in *Eptesicus* are derived from transposable elements (TEs) compared with only 23.9% and 16.5% in the dog and horse, respectively. Not surprisingly, given the retrotransposon bias of dog and horse, the majority of TE-derived p/miRNAs are associated with retrotransposons. In *Eptesicus*, however, 58.7% of the TE-derived and 35.9% of the total p/miRNAs arose not from retrotransposons but from bat-specific DNA transposons. Notably, we observe that the timing of the DNA transposon expansion and the resulting introduction of novel p/miRNAs coincide with the rapid diversification of the family Vespertilionidae. Furthermore, potential targets of the DNA transposon-derived p/miRNAs are identifiable and enriched for genes that are important for regulation of transcription. We propose that lineage-specific DNA transposon activity lead to the rapid and repeated introduction of novel p/miRNAs. Some of these p/miRNAs are likely functional miRNAs and potentially influenced the diversification of Vespertilionidae. Our observations suggest a mechanism for introducing functional genomic variation rapidly through the expansion of DNA transposons that fits within the TE-thrust hypothesis.

Key words: **miRNA, transposon, bat, evolution.**

Article

Introduction

MicroRNAs (miRNAs) are short, ~22 nt long, noncoding RNAs that act as posttranscriptional regulators of gene expression (Bartel 2004; Chen and Rajewsky 2007; Carthew and Sontheimer 2009); thus, changes in miRNA repertoires have great potential to generate evolutionary novelties through expanding regulatory pathways. Expansions in miRNA repertoires have been linked to major innovations in the evolution of vertebrates (Heimberg et al. 2008). Many miRNAs arise from transposable elements (TEs), and in mammals, several TE-derived miRNAs have been identified (Piriyapongsa and Jordan 2007; Piriyapongsa et al. 2007; Devor et al. 2009; Ahn et al. 2013). miRNAs originate from transcribed hairpins (He and Hannon 2004; Chen and Rajewsky 2007) suggesting that DNA sequences tending to form such structures would be a valuable source of novel miRNA loci. Class II TEs, that is, DNA transposons, commonly harbor palindromic sequences that form such hairpins (Feschotte et al. 2002; Sijen and Plasterk 2003; Ahn et al. 2013). Thus, fortuitous transcription initiated by nearby promoters would be all that is needed to generate pri-miRNAs from DNA transposon insertions. Given that these elements are high in copy number, they represent a particularly robust potential source of novel miRNAs, a hypothesis supported by the observation that miRNAs derived from ancient DNA transpositions are overrepresented in the human genome (Piriyapongsa et al. 2007).

Mammals in general have not experienced DNA transposon activity in the recent past (Pace and Feschotte 2007; Hellen and Brookfield 2013), and instead, retrotransposons (Class I TEs) dominate. This has disallowed direct, detailed study of ongoing miRNA origination via DNA transposon deposition. Vesper bats are unique among mammals in that they have experienced a substantial increase in DNA transposon activity, led by the Helitron superfamily. This

shift was first observed in the genome of *Myotis lucifugus*, the little brown bat, where multiple DNA transposon families are present in high copy numbers, have been deposited within the last 40 My (Pritham and Feschotte 2007; Ray et al. 2007, 2008; Pace et al. 2008; Thomas et al. 2011; Pagan et al. 2012) and are likely still actively transposing (Mitra et al. 2013).

Why and how DNA transposon activity has increased so substantially in vesper bats is still unknown. Introduction of the DNA transposons was likely via horizontal transfer (Pace et al. 2008; Gilbert et al. 2010; Pagan et al. 2010; Thomas et al. 2011), but a direct source of the transfer has yet to be identified. DNA transposon families have been identified in other vespertilionids (Thomas et al. 2011; Pagan et al. 2012), indicating that the expansions of DNA transposons can be traced to a common ancestor. Our knowledge of TEs in nonvesper bats is limited but informative. In the Pteropodidae (a yinpterochiropteran family), nearly all published work has focused on the retrotransposon extinction seen in several genera (Cantrell et al. 2008; Gogolevsky et al. 2009). Our own studies of *Pteropus vampyrus* (Pteropodidae) and studies from others (Ray DA, Platt III RN, unpublished data; Feschotte C, personal communication) suggest that there has been no recent activity from any TE class. Genomic survey sequencing in *Artibeus lituratus* suggests a TE history in Phyllostomidae (a yangochiropteran family more closely related to Vespertilionidae) that is more typical of mammals, i.e., extensive retrotransposon activity and little to no DNA transposon activity (Oliveira et al. 2012; Pagan et al. 2012; Ray DA, Pagan HJT, Platt III RN, Schaack S, unpublished data). Although this observation is made based on data from a single taxon, efforts to survey and or fully sequence the genomes of other phyllostomid bats will likely support it. Although there is the hint of horizontal transfer of a single family of Mariner transposons in Phyllostomidae (Oliveira et al. 2012), reports by Thomas et al. (2011) and Ray et al. (2008) suggest that no significant DNA transposon activity had occurred in a close relative of the vesper bats, the family Miniopteridae, or any phyllostomid, further supporting the limited distribution of the activity to vesper bats.

Given that Class II TEs are disproportionately likely to harbor secondary structures characteristic of miRNA loci (Feschotte et al. 2002; Sijen and Plasterk 2003; Ahn et al. 2013) and that this single clade of bats has experienced a recent expansion of DNA-based TEs in their genome, the genomes of vesper bats provide a natural system to explore the potential of DNA transposons as sources of novel miRNAs. We hypothesized that the increase in DNA transposon activity could have resulted in an upsurge in miRNA origination. To test this hypothesis, we compared patterns of TE activity and miRNA birth among *Eptesicus fuscus*, a vespertilionid bat, horse, and dog. Dog and horse were chosen as our comparison taxa because they 1) exhibit the "typical" mammalian TE landscape (i.e., retrotransposon rich, with no evidence of recent DNA transposon activity), 2) are relatively closely related to *E. fuscus*, and 3) are taxa for which tissues are readily available. More specifically, we compared the TE landscapes and predicted miRNA repertoires in each genome, estimated miRNA origination rates in each lineage, discriminating between DNA transposon, retrotransposon, and non-TE-derived miRNAs, and compared the observed patterns among species.

Validating individual miRNAs can be accomplished in several ways (Kuhn et al. 2008; Vasudevan 2012). For example, miRNAs are processed via the Dicer–Drosha pathway. Thus, association of the candidate RNAs to either protein would serve to confirm their status as miRNAs. Furthermore, the silencing of the mRNA target should result in a phenotypic change, and detection of that change would indicate miRNA function. Finally, a miRNA requires an mRNA target and that mRNA target should be expressed simultaneously with the candidate miRNA. Unfortunately, the tools to manipulate *Eptesicus* at the genome level in a manner that would meet all of these requirements do not yet exist. In an effort to be conservative in our language and to distinguish between experimentally validated or known miRNAs and those identified here, based on our expression and computational predictions, we prefer to refer to the latter by a term other than miRNA. The term milRNA (miRNA-like) is already used to refer to a class of small RNAs in fungi that are generated by a distinct biological pathway (Kang et al. 2013; Lau et al. 2013). We will therefore refer to our proposed putative miRNAs as p/miRNAs.

Results and Discussion

Genus *Eptesicus* Experienced Bursts of DNA Transposon Activity in the Recent Past

Our first step was to compare patterns of TE activity among *Eptesicus*, horse, and dog. To this end, we performed a de novo characterization of the TE landscape using the current genome draft of *Eptesicus* to verify that its genome has experienced similar bursts of Class II transposon activity when compared with *Myotis*. Our results confirm that over the past 40 My, around half of the TE content of the two bat genomes is derived from Class II TEs compared with only minor fractions in dog and horse (fig. 1). As expected, given that DNA transposons invaded a common ancestor of *Eptesicus* and *Myotis*, many DNA transposon families identified in *Myotis* are also present in *Eptesicus*. Finally, we identified multiple TE families found in *Eptesicus* that were missing or at very low copy number in the genome of *Myotis* (and vice versa), confirming that both lineages have experienced their own unique set of Class II TE expansions after their divergence from a common ancestor (fig. 1).

A Large Number of Putative miRNAs Emerged after the Divergence of *Eptesicus*, Dog, and Horse

A previous study of human miRNAs uncovered a burst of miRNA origination shortly before the split between Laurasiatheria and Euarchontoglires (Iwama et al. 2013). Accordingly, the genomes of *Eptesicus*, dog, and horse, all members of Laurasiatheria, would be expected to share several miRNAs. Through comparison with 19 other mammal genomes (supplementary table S1, Supplementary Material online), we identified 217, 345, and 382 putative miRNA (p/miRNAs) in *Eptesicus*, dog, and horse, respectively, which

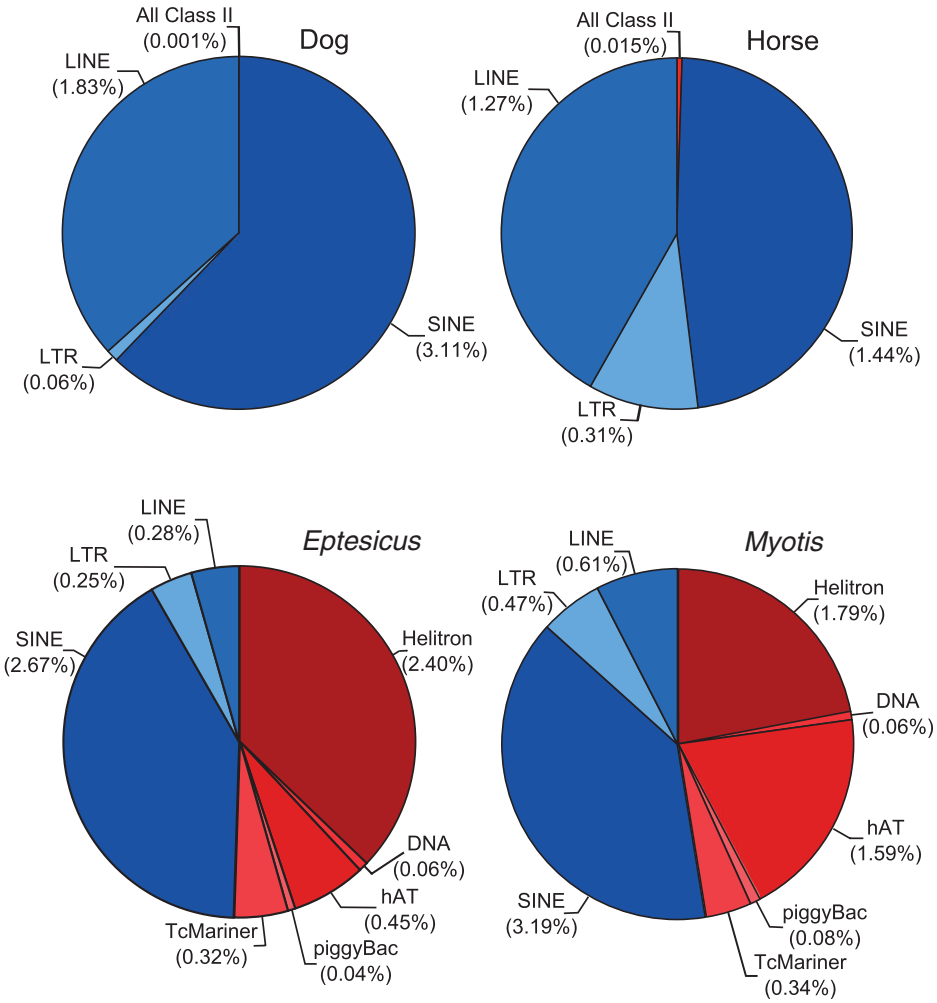

Fig. 1. Genome contributions (% genome) from TE classes over the past 40 My. Class I elements (retrotransposons) are shaded blue and Class II elements (DNA transposons) are shaded red. Genome coverage estimates were obtained by RepeatMasking the genome drafts with custom TE libraries for the two bats, and RepBase-derived libraries, dog and horse. Only hits spanning at least 50% of the consensus were considered. Divergences from the consensus were calculated as described in the Materials and Methods, and ages were calculated by applying a mutation rate of 2.366×10^{-9}.

originated prior to their most recent common ancestor (fig. 2, supplementary table S2 and fig. S1, Supplementary Material online). The lower values for the bat are to be expected as we are only including p/miRNAs identified during our analysis compared with all available sources for the dog and horse. To verify the initial p/miRNAs predicted in *Eptesicus*, additional small RNA reads from the testes of a separate individual were sequenced at a greater depth. These reads were then mapped back to the predicted pre-p/miRNAs hairpins from the original data set. An 11-fold increase in sequencing depth using the second individual confirmed 613 (93%) of the 661 p/miRNAs originally predicted in the smaller data set, suggesting that our depth of coverage was adequate.

Our analyses reveal that a large proportion of expressed p/miRNAs arose after the divergence of all three taxa from a common ancestor ~82 Ma (Murphy et al. 2007; Meredith et al. 2011). 58% (355 out of 613), 34% (184 out of 535), and 47% (339 out of 729) of the total p/miRNAs are unique to the lineages of *Eptesicus*, dog, and horse, respectively (table 1).

Most Recent *Eptesicus* Putative miRNAs Are Derived from Vespertilionid-Specific DNA Transposons

Even though TEs are an important source of novel p/miRNAs in all three species analyzed, *Eptesicus* stands out in two respects. Close to two-thirds (62%) of the 355 postdivergence p/miRNAs in *Eptesicus* are derived from TEs compared with only 23.9% in dog and 16.2% in horse. The sources of these TE-derived p/miRNAs in dog and horse are also distinct. The majority of TE-derived p/miRNAs, 90.9% and 65.4% are associated with retrotransposons, which are the dominant TEs in the genome of dog and horse. In *Eptesicus*, however, 60% of the TE-derived and 37.2% of all p/miRNAs arose from DNA transposons (table 1). When one considers that DNA transposons comprise only 10.3% of the *Eptesicus* genome compared with the 16.5% that is derived from retrotransposons, it is clear that DNA transposon-derived p/miRNAs are overrepresented. Yet, analysis at finer scales reveals even more interesting patterns.

Overall rates of p/miRNA origination are similar when comparing non-TE-derived p/miRNAs in *Eptesicus*, dog, and

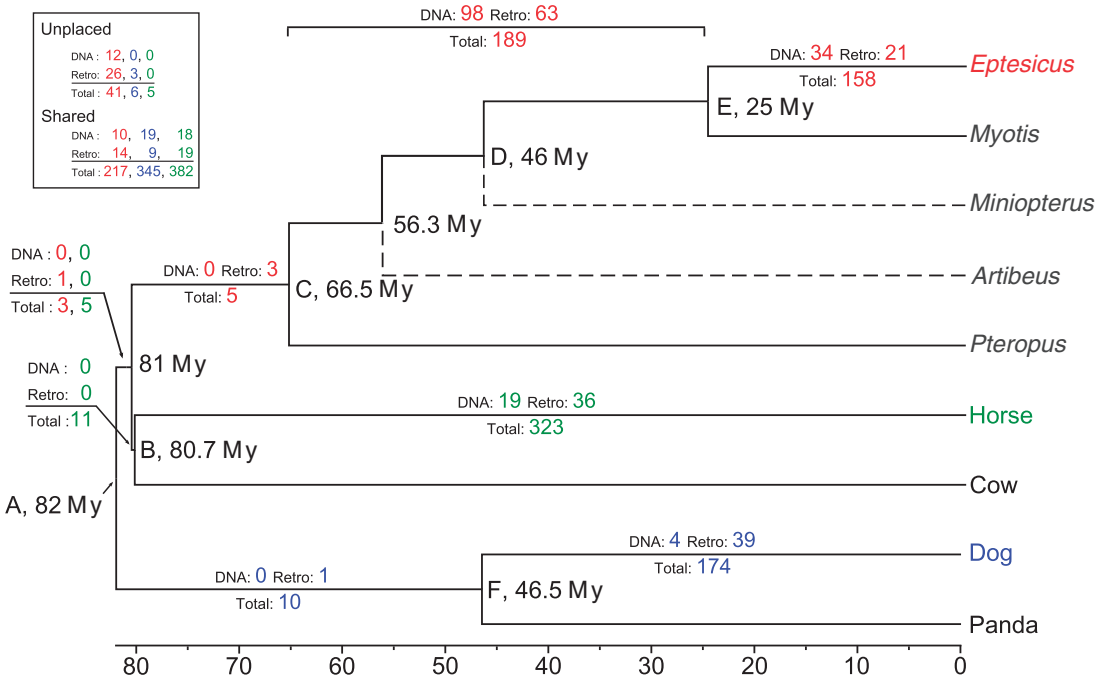

Fig. 2. Phenogram illustrating relationships and approximate divergence times for relevant taxa. Divergence times are as described in Meredith et al. (2011). Broken branches indicate taxa for which we do not have whole genome data, yielding holes in our knowledge with regard to TE complement, but which are relevant to the analyses (see the text). Total p/miRNA originations are identified below each branch for *Eptesicus* (red), dog (blue), and horse (green) (and on selected internal branches). Numbers of TE-derived p/miRNAs are found above each branch based on their origin. Class I (retrotransposon)- and Class II (DNA transposon)-derived p/miRNAs are shown above branches. p/miRNAs that originated prior to the common ancestor of bats, carnivores, and cetartiodactyls based on queries against 19 additional mammal genomes (Supplementary table S1, Supplementary Material online) are considered "Shared" and would have originated along the branch leading to node A (not depicted). p/miRNAs predicted by miRanalyzer but which cannot be assigned to a node are referred to as "Unplaced."

horse (table 1; A-*Eptesicus*, A-dog, and A-horse). However, the TE-derived p/miRNA origination rate in *Eptesicus* is around four times higher than either horse or dog, and the rate of DNA-transposon-derived p/miRNA birth is between 32 and 8 times higher in *Eptesicus* than dog or horse, respectively. Even more notable, high rates of p/miRNA origination map to branches with high DNA transposon activity. A conservative estimate based on data derived only from available genome drafts limits DNA transposon activity to after the divergence of *Pteropus* from the other bats (fig. 2, branch C-*Eptesicus*). The overall p/miRNA origination rate during this ~66 My period is ~5.3 p/miRNAs/My, with many being derived from DNA transposons (~2 p/miRNAs/My).

However, because previous data suggest that it is unlikely that DNA transposons were active in either *Artibeus* or *Miniopterus* (Thomas et al. 2011; Pagan et al. 2012), more closely related bats, a more realistic rate of DNA transposon-derived p/miRNA birth can be calculated. In particular, after the divergence from *Miniopterus* ~46 Ma (fig. 2, node D), we can surmise that DNA transposon-derived p/miRNAs arose in the lineage leading to *Eptesicus* (D-*Eptesicus*) at an increased rate of ~2.9 p/miRNAs/My. Notably, most p/miRNAs we identified were shared between *Myotis* and *Eptesicus*, indicating that they arose in a common ancestor and narrowing their origin to a window of ~21 My (fig. 2, branch D-E). We estimate that during this period, the p/miRNA origination rate reached a peak and that Class II-derived p/miRNAs were arising at 4.7 p/miRNAs/My.

A peak of Class II TE activity in *Eptesicus* and *Myotis* occurred during this same window, at ~26 Ma (fig. 3), just before the estimated vespertilionid divergence. After the *Eptesicus*/*Myotis* divergence, TE activity in general has been declining in *Eptesicus* and relatively few DNA transposon-derived p/miRNAs localize to that terminal branch (~1.4 p/miRNAs/My), leading us to conclude that the rise and fall of the p/miRNAs origination rate in *Eptesicus* is explained by the rise and fall of DNA transposon activity in this species.

DNA Transposon-Derived Putative miRNAs Target Genes Expressed in *Eptesicus* Testes

Mammalian miRNAs often regulate mRNA expression through targeted binding of the 3′-untranslated region (UTR) (Lai 2002), and recent data have shown that Helitrons in the 3′-UTR of *M. lucifugus* could serve as potential miRNA-binding sites (Thomas J, Philips C, Baker R, Pritham E, submitted for publication). To investigate the potential regulatory roles of testes p/miRNAs, we identified candidate targets by mapping the p/miRNAs to the annotated *Eptesicus* genome draft. Subsequently, we identified potential target genes that were expressed in the testes using gene predictions as described in the Materials and Methods. Forty-five p/miRNAs appear to target genes expressed in testes (supplementary table S3, Supplementary Material online). We then

Table 1. Summary of p/miRNA Data for *Eptesicus*, Dog, and Horse.

	Eptesicus		Dog	Horse
Previously identified miRNAs[a]	—		323	341
Predicted testis p/miRNAs	613		212	388
Previously identified miRNAs sequenced from testis	—		229	295
Total p/miRNAs analyzed from each taxon	613		535	729
Branch	A-*Eptesicus*		A-Dog	A-Horse
MRCA[b] to terminal taxa				
Total p/miRNAs	355		184	339
Non-TE-derived p/miRNAs	135		140	283
TE-derived p/miRNAs	220		44	56
Class II-derived p/miRNAs	132		4	20
% Class II-derived p/miRNAs	37.2%		2.2%	5.9%
p/miRNA origination rate (per My)	4.3		2.2	4.1
Non-TE p/miRNA origination rate (per My)	1.6		1.7	3.5
TE-derived p/miRNA origination rate (per My)	2.7		0.5	0.7
Class II-derived p/miRNA origination rate (per My)	1.6		0.05	0.2
Branch[c]	E-*Eptesicus*	C-E	F-Dog	B-Horse
Branches of interest				
Total p/miRNAs	158	189	174	323
Non-TE-derived p/miRNAs	103	28	131	268
TE-derived p/miRNAs	55	161	43	55
Class II-derived p/miRNAs	34	98	4	19
% Class II-derived p/miRNAs	21.5%	51.8%	2.3%	5.9%
p/miRNA origination rate (per My)	6.3	4.7	3.7	4
Non-TE p/miRNA origination rate (per My)	4.1	0.7	2.8	3.3
TE-derived p/miRNA origination rate (per My)	2.2	3.9	0.9	0.7
Class II-derived p/miRNA origination rate (per My)	1.4	2.4	0.1	0.2

[a]From mirBase.

[b]MRCA = most recent common ancestor, in this case Variamana. Data taken from global mean provided in Meredith et al. (2011).

[c]Time estimates for all nodes are global means from Meredith et al. (2011) except D-*Eptesicus* and C-D. D-*Eptesicus* was calculated using Time Tree (*Eptesicus* vs. *Myotis*). C-D is the difference between the dates on nodes C and D.

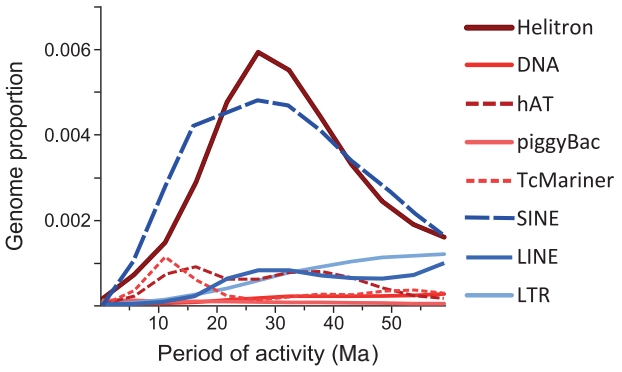

Fig. 3. Line graph illustrating contributions to the *Eptesicus* genome by TEs over the past ~60 My (just prior to the estimated divergence of *Eptesicus* from *Artibeus*). As with figure 1, shades of blue indicate retrotransposons and shades of red indicate DNA transposon superfamilies.

grouped the p/miRNAs into categories based on their age. Four expressed genes appear to be targeted by p/miRNAs that arose prior to the Boreotherian divergence. Three of these are involved in signal transduction (GPR135, GNG10, and TAC3) while the fourth (BCO2) mediates beta-carotene oxidation. Only four expressed p/miRNAs with identifiable targets arose after the boreotherian split, but prior to the divergence of Chiroptera. Two (PVRL3 and TNP1) play a direct role in spermatogenesis, while the others (TOP1 and RECQL4) function in telomere maintenance.

Using DAVID (Huang et al. 2009) to categorize the bat-specific p/miRNAs targets ($n = 38$) reveals enrichment for two ontological terms: transcription (22 genes, P-value = 6.8e−6, FDR = 0.01) and ubiquitin-dependent protein catabolic processes (seven genes, P-value = 1e−4, FDR = 0.1). Of the 22 genes affecting transcription, 11 are targeted by p/miRNAs derived from DNA transposons, while two of the seven genes involved in ubiquitin-dependent processes arise from DNA transposons. The remaining p/miRNAs are derived from retrotransposons or non-TE-derived sequences.

The Introduction of DNA Transposon-Derived Putative miRNAs Coincides with the Rapid Diversification of Vesper Bats

Vespertilionidae radiated rapidly into ~400 species (and ~48 genera) after splitting from their close relatives, Cistugidae, ~34 Ma (Miller-Butterworth et al. 2007; Lack and Van Den Bussche 2010; Meredith et al. 2011), making it the second most species-rich mammalian family (Simmons 2005).

Lineage-through-time plots indicate an increased rate of diversification (Lack and Van Den Bussche 2010) when comparing vesper bats to the phyllostomid bats, which accumulated lineages more gradually but did not experience a similar Class II expansion (Thomas et al. 2011; Pagan et al. 2012). In fact, the increased rate of cladogenesis is most apparent between 20 and 30 Ma, a period that coincides with the massive initial expansion and eventual peak of DNA transposon, primarily Helitron, activity (fig. 3). Of the top 20 TE families that have contributed testis p/miRNAs, 12 are DNA transposons and 10 of these belong to the Helitron superfamily (supplementary table S4, Supplementary Material online). Of the top 10, 5 p/miRNAs sources are Helitrons. These data strongly suggest that the unique DNA transposon activity observed in vesper bats has contributed uniquely and significantly to the repertoire of small regulatory RNAs.

How Might the Burst of Putative miRNAs Have Impacted Vespertilionid Bat Diversity?

Multiple authors have suggested that TEs may impact diversification (Furano et al. 1994; Jurka et al. 2011, 2012; Oliver and Greene 2011, 2012), and in fact, a connection between DNA transposon activity and early primate evolution has been previously noted (Pace and Feschotte 2007; Zeh et al. 2009; Oliver and Greene 2011). Additional authors have suggested a link between the unique TE activity in vesper bats and their rapid diversification (Pritham and Feschotte 2007; Ray et al. 2008; Zeh et al. 2009; Thomas et al. 2011; Oliver and Greene 2012; Mitra et al. 2013), but no mechanism has yet been supported by experimental data.

We note the following observations: 1) vesper bats experienced increased rates of diversification centering 20–30 Ma, 2) the same bats experienced a massive increase in lineage-specific DNA transposon activity during the same period, 3) a large proportion of p/miRNAs are derived from the lineage-specific DNA transposons, and 4) those p/miRNAs are enriched for terms related to transcriptional regulation. We also note one additional feature in the history of these bats and their evolution. Approximately 33–34 Ma, there was a rapid and significant shift in Earth's climate, the Eocene–Oligocene transition. It was during this transition that the climate changed from warm tropical conditions to a more temperate regime, with ice sheets covering the poles (Zanazzi et al. 2007; Liu et al. 2009). Selection would likely favor organisms with the ability to rapidly respond to a changing environment.

We show that the genome of the ancestral vespertilionid was successfully invaded by Class II TEs just prior to and during this period and that the invasion resulted in the introduction of a large number of p/miRNAs with the potential to influence regulatory pathways. Given the importance of transcriptional regulation to evolution and the number of transcription regulators that appear to be targeted by DNA transposon-derived p/miRNAs, it is reasonable to hypothesize that the recently accumulated p/miRNAs led to changes in gene regulation that were acquired in a lineage-specific manner and that could influence the adaptive radiation of vesper bats.

All this being said, we must point out that vesper bats are behaviorally, morphologically, and karyotypically highly conserved (Simpson 1944; Corbet and Hill 1991; Fitch and Ayala 1994; Koopman 1994; Nowak 1999; Simmons 2005), and it is this characteristic homogeneity that has made relationships within the family difficult to resolve (Lack and Van Den Bussche 2010). The same authors suggest that vesper bats diversified and dispersed rapidly, obviating the need for morphological or behavioral specialization. Such a scenario is supported by a recent study of South American *Myotis* from Larsen et al. (2012). Prior to that study, 15 species of South American *Myotis* were recognized. Using nuclear and mitochondrial markers, the authors suggest that the actual number of species may be as high as 34. One "species," *M. nigricans*, exhibited at least 12 species-level lineages (i.e., cytochrome b variation greater than 5%) and many of these were paraphyletic, suggesting that a single group of essentially identical specimens might actually represent multiple cryptic species.

By contrast, the phyllostomid bats are one of the most morphologically diverse mammalian families but contain comparable numbers of lineages (genera) to Vespertilionidae. There are many aspects to phenotypic diversity beyond morphology, however. In fact, gene expression is often considered a quantitative trait on par with morphology (Cheung and Spielman 2002). In applying this understanding, it is difficult to say that one family of bats is more phenotypically diverse than another. Due to the complex and rapid acquisition and expression of novel TE-derived p/miRNAs in vespertilionids, we hypothesize that the opportunity for genetic incompatibilities (isolating mechanisms) was increased. As the opportunity for such incompatibilities increased, species diversification would likely follow but might not have been accompanied by morphological or behavioral variation. Indeed, the lack of morphological diversity in spite of increased species diversity seems to implicate a significant role for genetic rather than morphological or behavioral isolating mechanisms in vespertilionids.

Coincidence of DNA Transposon with Other Species Diversifications

The expansion of novel p/miRNAs via DNA transposon activity proposed here provides several avenues for future research and a number of testable hypotheses. For example, there has been a recent increase in activity by hAT and piggyBac DNA transposons in *Myotis* (Ray et al. 2007; Pace et al. 2008; Ray et al. 2008; Mitra et al. 2013). The activity of these two families is independent of the peak of Helitron activity shared by all vesper bats and is not found in *Eptesicus*. Stadelmann et al. (2007) found that genus *Myotis* experienced a species radiation that includes more than 100 species during the last 10–15 My, and this coincides with the rapid rise in hAT activity (Ray et al. 2008). By contrast, *Eptesicus* consists of 23 species that emerged over the same period (Roehrs et al. 2010). If our hypotheses are correct, one

would expect to find an increase in hAT-derived p/miRNAs over the same time in *Myotis*. Furthermore, we would expect these p/miRNAs to be restricted to genus *Myotis* and not be shared by other vesper bats. Mitra et al. (2013) point out that a single New World *Myotis* clade has radiated very recently within the last 1–2 My. This radiation is coincident with the expansion of a piggyBac family (*piggyBat*) and evidence to date suggests that *piggyBat* is restricted to New World members of the genus (Ray et al. 2008; Mitra et al. 2013). Like most DNA elements, piggyBacs and hATs contain palindromes forming hairpins that could lead to novel miRNAs.

All of these scenarios fit well with the recently proposed TE-thrust hypothesis of Oliver and Greene (2011, 2012), which suggests that "[l]ineages with an abundant and suitable repertoire of TEs have enhanced evolutionary potential and, if all else is equal, tend to be fecund, resulting in species-rich adaptive radiations." Indeed, other eukaryotic clades that are more prone to DNA transposon activity (especially from Helitrons) may have experienced phenomena similar to what we observe in Vespertilionidae. For example, the green anole (*Anolis*), which exhibits an incredible array of mostly young TEs including both retrotransposons and DNA transposons (Carmell et al. 2002; Novick et al. 2009, 2010, 2011; Alfoldi et al. 2011), is particularly diverse with more than 400 recognized species having come about through rapid speciation. Future studies of *Anolis* might provide further insight into the relationship between DNA transposon activity and miRNA origination as well as TE-thrust in general.

Conclusions

The data presented here explain the acquisition of novel p/miRNAs via lineage-specific TE activity and indicate a mechanism that may have impacted the radiation of a highly diverse mammalian clade. The expansion of DNA transposons and the resulting introduction of p/miRNA loci in the ancestral vespertilionid genome correspond with the rapid diversification of this group. The p/miRNAs that resulted could have provided the raw material to generate novel regulatory pathways, as evidenced by the targeting of expressed testis genes. Introduction of novel gene expression phenotypes could have promoted the rapid diversification of this family.

Materials and Methods

TE Landscape Characterization in *Eptesicus* and *Myotis*

We characterized the TE landscape of *E. fuscus* (EptFus1.0, GenBank accession ALEH00000000) using RepeatModeler and a combination of existing tools and custom perl scripts (Smit and Hubley 2008–2010, and available upon request; Smith and Ray 2011). The methods used were similar to those described previously (Alfoldi et al. 2011; Consortium 2012). Briefly, RepeatModeler was used to identify potential TE family consensus sequences. We then used BLAST to query the entire WGS draft (Altschul et al. 1990). Hits of 100 bases or more were extracted along with up to 1,000 bases of flanking sequence using custom perl scripts. MUSCLE (Edgar 2004) was used to align the extracted hits with the consensus and 50% majority rules consensus sequences were constructed. We then examined the 5′- and 3′-ends of the alignments. Consensus sequences were considered "complete" when single copy sequence could be identified at both ends of the alignment in the component sequences. If single copy sequences were not identifiable, the process was repeated. The resulting library of elements was submitted to CENSOR (Kohany et al. 2006), BLASTN, and BLASTX to identify potential designations, which were confirmed through the identification of key sequence landmarks such as terminal inverted repeats, target site duplications, and coding sequences. All elements have been submitted to RepBase.

The TE landscape of *M. lucifugus* has been previously characterized (Pritham and Feschotte 2007; Ray et al. 2007, 2008). We combined our novel library of *Eptesicus* TEs with elements previously identified in *Myotis* and other bats (Pritham and Feschotte 2007; Ray et al. 2007, 2008; Pace et al. 2008; Thomas et al. 2011; Pagan et al. 2012) as well as elements characterized independently and available from RepBase. Using the combined library, the *Eptesicus* and *Myotis* genome drafts were analyzed using a locally implemented version of RepeatMasker to estimate the TE content of the genome and to generate genome-wide TE annotations.

The TE landscapes of dog and horse have been previously characterized as described at http://www.repeatmasker.org/species/canFam.html (last accessed April 2, 2014) and http://www.repeatmasker.org/species/equCab.html (last accessed April 2, 2014). RepeatMasker output files were obtained from these sites.

To calculate approximate periods of activity for all taxa, we restricted ourselves to RepeatMasker hits that spanned at least 50% of any given consensus sequence. To estimate divergences, we used a modified version of the calcDivergenceFromAlign.pl script that is included in the RepeatMasker package to calculate Kimura two-parameter distances between each insertion and its respective consensus (Pagan et al. 2010). The -noCpG option was invoked. We applied the mutation rate estimated by Ray et al. (2008), 2.366×10^{-9} substitutions per site/My.

Small RNA Isolation and Sequencing

A horizontal cross section from the testes of a wild caught *E. fuscus* was taken, and tissues were snap frozen in liquid nitrogen. Testis samples from dog and horse were collected as a byproduct of veterinarian preformed castrations. Total RNA was extracted in TRIzol following the manufacturer's recommended protocol. Small RNAs were isolated, indexed, pooled, and prepared for sequencing using the Illumina TruSeq Small RNA Sample Preparation kit. Small RNAs were sequenced on a single Illumina HiSeq lane (1×50 nt reads). Reads counts for bat, dog, and horse were 18,310,167, 65,893,540, and 69,420,477, respectively. 3′-Adapter sequences were removed, and reads <15 bp were discarded. Every base call within reads was required to have a Phred quality scores >25. Reads falling short of this threshold were removed. Each of the above steps

was accomplished via FASTX-toolkit (version 0.0.13; http://hannonlab.cshl.edu/fastx_toolkit/, last accessed April 9, 2014).

p/miRNAike Sequence Prediction and Evolution

To predict p/miRNAs, we used a local installation of the miRanalyzer package and completed all searches against EptFus1.0, canFam3.1, and the equCab2 assemblies for the *Eptesicus*, dog, and horse genomes, respectively. The miRanalyzer platform is a multistep approach to miRNA discovery. In the first step, reads were mapped to all known species-specific pre-miRNAs from the current release of miRBase (Release 19; Kozomara and Griffiths-Jones 2011). After this comparison, all reads mapped to known pre-miRNAs were removed, and the remaining reads were mapped to the ab initio transcriptome libraries from Ensembl (Release 70; Flicek et al. 2012) or to the Augustus-based gene predictions of the *E. fuscus* assembly. Any reads mapping to the transcriptome were considered to be degraded mRNA and removed. In the final step, miRanalyzer, via Bowtie, mapped all reads to the genome draft. The region surrounding these reads was analyzed for the potential to form stable pre-miRNA hairpins. In each step of our analyses, we chose conservative parameters. All mapping steps did not allow any mismatches, and p/miRNA predictions were required to meet a probability score of 0.95 and be positively predicted by four of five miRNA models in miRanalyzer. In some instances, miRanalyzer predicted multiple pre-p/miRNAs from a single sequence read. These multiple hairpins were usually offset by a single nucleotide. In these cases, one pre-p/miRNA was kept while other overlapping pre-p/miRNAs were excluded from additional analyses. Predicted pre-p/miRNA loci were intersected with RepeatMasker TE annotations using Bedtools (Quinlan and Hall 2010). p/miRNAs were considered TE-derived if the mature p/miRNAs overlapped a TE by 10 nt or more. When a region was annotated as multiple TEs or when a hairpin p/miRNA intersected two TEs, the TE overlapping the p/miRNA the most was retained.

Verification of Predicted Putative miRNAs in *Eptesicus*

Because of the relatively low read count in *Eptesicus* in our initial analysis, we chose to verify the initial p/miRNA predictions, with additional small RNA data from a second individual. The small RNA pool from a unique male *E. fuscus* from a separate population was isolated and prepared for sequencing using the methods described above. The small RNA library was sequenced on an Illumina HiSeq (1×50 nt reads), yielding 208,193,690 reads. After the adapter sequences were trimmed and reads falling outside of a 17–24 bp size range were excluded, the remaining reads ($n = 11{,}553{,}083$) were mapped to the original miRanalyzer predictions for *Eptesicus*. Mapping reads were considered confirmation of the p/miRNAs in *Eptesicus*.

Phylogenetic Inference of Putative miRNA Origins

To infer the likely origination period of each p/miRNA, we identified the presence of the mature p/miRNAs in 22 mammalian genomes (supplementary table S1, Supplementary Material online) using the Mapmi package (Guerra-Assuncao and Enright 2010). By default Mapmi uses the RepeatMasked version of Ensemble genomes. To circumvent this feature, we used a local installation of Mapmi and supplied non-RepeatMasked genomes (Ensemble release 70). After querying each genome with all predicted mature p/miRNAs, an estimated origination period was inferred for each. p/miRNA origination was mapped onto a reduced version (22 taxa) of the mammalian phylogeny provided by Meredith et al. (2011) using the Dollo parsimony function in Phylip (supplementary table S2, Supplementary Material online).

De Novo Gene Prediction in *Eptesicus*

Gene predictions on EptFus1.0 were accomplished using Augustus (Stanke and Waack 2003). In addition to the assembled genome scaffolds, two reference species, *P. vampyrus* and *M. lucifugus*, were incorporated into the Augustus analysis. Tophat version 2.0.6 and Bowtie version 2.0.5 were used to align the RNAseq reads to the assembled genome scaffolds. These alignments were then used by Augustus to improve gene predictions. Note that the supplementary scripts and programs required that are packaged with Augustus were used from Augustus version 2.7, when Augustus itself was run, version 2.5.5 was used. The reference genomes from the *P. vampyrus* and *M. lucifugus were* added in the second iteration of running Augustus. The sequences were downloaded from the Ensembl 70 database, and alignments were created using the genblastg version 1.38 with the "-gff" argument to generate alignments in General Feature Format (GFF). The results were concatenated to the hints file generated from the RNA-Seq data and the first iteration of Augustus predictions. For the final run of Augustus, the "-UTR = on" argument was added to generate UTR predictions.

Putative miRNAs Target Prediction

Mammalian miRNAs regulate mRNA expression through targeted binding within the 3′-UTR of testis transcripts as per our Augustus gene predictions. Perfect complementarity within the seed region of the miRNA (nt 2–8) is necessary for miRNA binding. To identify potential binding sites, p/miRNAs were mapped to the 3′-UTRs using Bowtie with the following parameters: the first base on the 5′-end was skipped, the seed was 7-bp long, no mismatches were allowed within the seed, and up to two mismatches were allowed in the remainder of the p/miRNAs.

Supplementary Material

Supplementary tables S1–S4 and figure S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank the Broad Institute Genomics Platform and Genome Sequencing and Analysis Program, Federica Di Palma, and Kerstin Lindblad-Toh for making the data for

E. fuscus and *M. lucifugus* available. The mammal phylogeny was kindly provided by R. Meredith. This work was supported by the National Science Foundation (EPS-0903787 to F.G.H., DBI-1147029 to C.J.S., MCB-0841821 to D.A.R., and DEB-1020865 to D.A.R.). Additional support was provided by the College of Agriculture and Life Sciences at Mississippi State University. Tissues were provided by the College of Veterinary Medicine at Mississippi State University, Jeremiah Dumas, and Dr. Robert Baker and Dr. Caleb Phillips of the Natural Science Research Laboratory at Texas Tech University. RNA-Seq and small RNA reads have been deposited in the Gene Expression Omnibus archive (GSE47593).

References

Ahn K, Gim JA, Ha HS, Han K, Kim HS. 2013. The novel MER transposon-derived miRNAs in human genome. *Gene* 512:422–428.

Alfoldi J, Di Palma F, Grabherr M, Williams C, Kong L, Mauceli E, Russell P, Lowe CB, Glor RE, Jaffe JD, et al. 2011. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* 477:587–591.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.

Cantrell MA, Scott L, Brown CJ, Martinez AR, Wichman HA. 2008. Loss of LINE-1 activity in the megabats. *Genetics* 178:393–404.

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* 16: 2733–2742.

Carthew RW, Sontheimer EJ. 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell* 136:642–655.

Chen K, Rajewsky N. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet.* 8:93–103.

Cheung VG, Spielman RS. 2002. The genetics of variation in gene expression. *Nat Genet.* 32(Suppl), 522–525.

Consortium HG. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98.

Corbet GB, Hill JE. 1991. A world list of mammalian species. New York: Oxford University Press.

Devor EJ, Peek AS, Lanier W, Samollow PB. 2009. Marsupial-specific microRNAs evolved from marsupial-specific transposable elements. *Gene* 448:187–191.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Feschotte C, Jiang N, Wessler SR. 2002. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet.* 3:329–341.

Fitch WM, Ayala FJ. 1994. Tempo and mode in evolution. *Proc Natl Acad Sci U S A.* 91:6717–6720.

Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, et al. 2012. Ensembl 2012. *Nucleic Acids Res.* 40:D84–D90.

Furano AV, Hayward BE, Chevret P, Catzeflis F, Usdin K. 1994. Amplification of the ancient murine Lx family of long interspersed repeated DNA occurred during the murine radiation. *J Mol Evol.* 38: 18–27.

Gilbert C, Schaack S, Pace JK 2nd, Brindley PJ, Feschotte C. 2010. A role for host–parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 464:1347–1350.

Gogolevsky KP, Vassetzky NS, Kramerov DA. 2009. 5S rRNA-derived and tRNA-derived SINEs in fruit bats. *Genomics* 93:494–500.

Guerra-Assuncao JA, Enright AJ. 2010. MapMi: automated mapping of microRNA loci. *BMC Bioinformatics* 11:133.

He L, Hannon GJ. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet.* 5:522–531.

Heimberg AM, Sempere LF, Moy VN, Donoghue PC, Peterson KJ. 2008. MicroRNAs and the advent of vertebrate morphological complexity. *Proc Natl Acad Sci U S A.* 105:2946–2950.

Hellen EH, Brookfield JF. 2013. The diversity of class II transposable elements in mammalian genomes has arisen from ancestral phylogenetic splits during ancient waves of proliferation through the genome. *Mol Biol Evol.* 30:100–108.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 4:44–57.

Iwama H, Kato K, Imachi H, Murao K, Masaki T. 2013. Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Mol Biol Evol.* 30:613–626.

Jurka J, Bao W, Kojima KK. 2011. Families of transposable elements, population structure and the origin of species. *Biol Direct.* 6:44.

Jurka J, Bao W, Kojima KK, Kohany O, Yurka MG. 2012. Distinct groups of repetitive families preserved in mammals correspond to different periods of regulatory innovations in vertebrates. *Biol Direct.* 7:36.

Kang K, Zhong J, Jiang L, Liu G, Gou CY, Wu Q, Wang Y, Luo J, Gou D. 2013. Identification of microRNA-like RNAs in the filamentous fungus *Trichoderma reesei* by solexa sequencing. *PLoS One* 8: e76288.

Kohany O, Gentles A, Hankus L, Jurka J. 2006. Annotation, submission and screening of repetitive elements in Repbase: RebaseSubmitter and Censor. *BMC Bioinformatics* 7:474.

Koopman KF. 1994. Chiroptera: Systematics. In: Niethammer J, Schliemann H, Starck D, editors. Handbuch der Zoologie. Berlin: de Gruyter. p. 100–109.

Kozomara A, Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39: D152–D157.

Kuhn DE, Martin MM, Feldman DS, Terry AV Jr, Nuovo GJ, Elton TS. 2008. Experimental validation of miRNA targets. *Methods* 44:47–54.

Lack JB, Van Den Bussche RA. 2010. Identifying the confounding factors in resolving phylogenetic relationships in Vespertilionidae. *J Mammal.* 91:1435–1448.

Lai EC. 2002. Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet.* 30:363–364.

Larsen RJ, Knapp MC, Genoways HH, Khan FA, Larsen PA, Wilson DE, Baker RJ. 2012. Genetic diversity of neotropical *Myotis* (chiroptera: vespertilionidae) with an emphasis on South American species. *PLoS One* 7:e46578.

Lau SK, Chow WN, Wong AY, Yeung JM, Bao J, Zhang N, Lok S, Woo PC, Yuen KY. 2013. Identification of microRNA-like RNAs in mycelial and yeast phases of the thermal dimorphic fungus *Penicillium marneffei*. *PLoS Negl Trop Dis.* 7:e2398.

Liu Z, Pagani M, Zinniker D, Deconto R, Huber M, Brinkhuis H, Shah SR, Leckie RM, Pearson A. 2009. Global cooling during the eocene–oligocene climate transition. *Science* 323:1187–1190.

Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TL, Stadler T, et al. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524.

Miller-Butterworth CM, Murphy WJ, O'Brien SJ, Jacobs DS, Springer MS, Teeling EC. 2007. A family matter: conclusive resolution of the taxonomic position of the long-fingered bats, *Miniopterus*. *Mol Biol Evol.* 24:1553–1561.

Mitra R, Li X, Kapusta A, Mayhew D, Mitra RD, Feschotte C, Craig NL. 2013. Functional characterization of piggyBat from the bat *Myotis lucifugus* unveils an active mammalian DNA transposon. *Proc Natl Acad Sci U S A.* 110:234–239.

Murphy WJ, Pringle TH, Crider TA, Springer MS, Miller W. 2007. Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Res.* 17:413–421.

Novick P, Smith J, Ray D, Boissinot S. 2010. Independent and parallel lateral transfer of DNA transposons in tetrapod genomes. *Gene* 449: 85–94.

Novick PA, Basta H, Floumanhaft M, McClure MA, Boissinot S. 2009. The evolutionary dynamics of autonomous non-LTR retrotransposons in the lizard *Anolis carolinensis* shows more similarity to fish than mammals. *Mol Biol Evol.* 26:1811–1822.

Novick PA, Smith JD, Floumanhaft M, Ray DA, Boissinot S. 2011. The evolution and diversity of DNA transposons in the genome of the lizard *Anolis carolinensis*. *Genome Biol Evol.* 3:1–14.

Nowak RM. 1999. Walker's mammals of the world. Baltimore (MD): The Johns Hopkins University Press.

Oliveira SG, Bao WD, Martins C, Jurka J. 2012. Horizontal transfers of Mariner transposons between mammals and insects. *Mob DNA.* 3:14.

Oliver KR, Greene WK. 2011. Mobile DNA and the TE-thrust hypothesis: supporting evidence from the primates. *Mob DNA.* 2:8.

Oliver KR, Greene WK. 2012. Transposable elements and viruses as factors in adaptation and evolution: an expansion and strengthening of the TE-Thrust hypothesis. *Ecol Evol.* 2:2912–2933.

Pace JK 2nd, Feschotte C. 2007. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res.* 17:422–432.

Pace JK 2nd, Gilbert C, Clark MS, Feschotte C. 2008. Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci U S A.* 105:17023–17028.

Pagan HJ, Macas J, Novak P, McCulloch ES, Stevens RD, Ray DA. 2012. Survey sequencing reveals elevated DNA transposon activity, novel elements, and variation in repetitive landscapes among vesper bats. *Genome Biol Evol.* 4:575–585.

Pagan HJ, Smith JD, Hubley RM, Ray DA. 2010. PiggyBac-ing on a primate genome: novel elements, recent activity and horizontal transfer. *Genome Biol Evol.* 2:293–303.

Piriyapongsa J, Jordan IK. 2007. A family of human microRNA genes from miniature inverted-repeat transposable elements. *PLoS One* 2: e203.

Piriyapongsa J, Marino-Ramirez L, Jordan IK. 2007. Origin and evolution of human microRNAs from transposable elements. *Genetics* 176: 1323–1337.

Pritham EJ, Feschotte C. 2007. Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc Natl Acad Sci U S A.* 17:422–432.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.

Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL. 2008. Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res.* 18: 717–728.

Ray DA, Pagan HJT, Thompson ML, Stevens RD. 2007. Bats with hATs: evidence for recent DNA transposon activity in genus *Myotis*. *Mol Biol Evol.* 24:632–639.

Roehrs ZP, Lack JB, Van Den Bussche RA. 2010. Tribal phylogenetic relationships within Vespertilioninae (Chiroptera:Vespertilionidae) based on mitochondrial and nuclear sequence data. *J Mammal* 91:1073–1082.

Sijen T, Plasterk RH. 2003. Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* 426:310–314.

Simmons NB. 2005. Order Chiroptera. In: Wilson DE, Reeder DM, editors. Mammal species of the world: a taxonomic and geographic reference. Baltimore (MD): The Johns Hopkins University Press. p. 312–529.

Simpson GG. 1944. Tempo and mode in evolution. New York: Columbia University Press.

RepeatModeler Open-1.0. 2008-2010 [Internet]. 2008–2010. [cited 2014, April 2]. Available from: http://www.repeatmasker.org.

Smith JD, Ray DA. 2011. Expedited batch processing and analysis of transposon insertions. *BMC Res Notes.* 4:482.

Stadelmann B, Lin LK, Kunz TH, Ruedi M. 2007. Molecular phylogeny of New World Myotis (Chiroptera, Vespertilionidae) inferred from mitochondrial and nuclear DNA genes. *Mol Phylogenet Evol.* 43: 32–48.

Stanke M, Waack S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* 19(2 Suppl), ii215–ii225.

Thomas J, Sorourian M, Ray D, Baker RJ, Pritham EJ. 2011. The limited distribution of Helitrons to vesper bats supports horizontal transfer. *Gene* 474:52–58.

Vasudevan S. 2012. Functional validation of microRNA–target RNA interactions. *Methods* 58:126–134.

Zanazzi A, Kohn MJ, MacFadden BJ, Terry DO. 2007. Large temperature drop across the Eocene–Oligocene transition in central North America. *Nature* 445:639–642.

Zeh DW, Zeh JA, Ishida Y. 2009. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays* 31:715–726.