Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Multiple bursts of pancreatic ribonuclease gene duplication in insect-eating bats

Huihui Xu ^a, Yang Liu ^a, Fanxing Meng ^a, Beibei He ^b, Naijian Han ^c, Gang Li ^c, Stephen J. Rossiter ^{d,*}, Shuyi Zhang ^{a,*}

^a Institute of Molecular Ecology and Evolution, Institutes for Advanced Interdisciplinary Research, East China Normal University, Shanghai 200062, China

^b School of Life Sciences, East China Normal University, Shanghai 200062, China

^c Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

^d School of Biological and Chemical Sciences, Queen Mary, University of London, London E1 4NS, UK

ARTICLE INFO

Article history: Accepted 1 April 2013 Available online 1 May 2013

Keywords: RNASE1 Chiroptera Gene duplication Positive selection

ABSTRACT

Pancreatic ribonuclease gene (RNASE1) was previously shown to have undergone duplication and adaptive evolution related to digestive efficiency in several mammalian groups that have evolved foregut fermentation, including ruminants and some primates. RNASE1 gene duplications thought to be linked to diet have also been recorded in some carnivores. Of all mammals, bats have evolved the most diverse dietary specializations, mainly including frugivory and insectivory. Here we cloned, sequenced and analyzed RNASE1 gene sequences from a range of bat species to determine whether their dietary adaptation is mirrored by molecular adaptation. We found that seven insect-eating members of the families Vespertilionidae and Molossidae possessed two or more duplicates, and we also detected three pseudogenes. Reconstructed RNASE1 gene trees based on both Bayesian and maximum likelihood methods supported independent duplication events in these two families. Selection tests revealed that RNASE1 gene duplicates have undergone episodes of positive selection indicative of functional modification, and lineage-specific tests revealed strong adaptive evolution in the *Tadarida* β clade. However, unlike the *RNASE1* duplicates that function in digestion in some mammals, the bat RNASE1 sequences were found to be characterized by relatively high isoelectric points, a feature previously suggested to promote defense against viruses via the breakdown of double-stranded RNA. Taken together, our findings point to an adaptive diversification of RNASE1 in these two bat families, although we find no clear evidence that this was driven by diet. Future experimental assays are needed to resolve the functions of these enzymes in bats.

© 2013 Published by Elsevier B.V.

1. Introduction

Vertebrate pancreatic ribonuclease (RNase1) is a secretory enzyme that belongs to the ribonuclease A superfamily (Beintema and Kleineidam, 1998). Its biological function in vertebrates is to degrade pathogenic RNA and thus protect organisms as part of the immune system (Sorrentino, 2010; Sorrentino and Libonati, 1994). Previous studies have uncovered high expression levels of RNase in the pancreas of ruminants and Old World colobine monkeys, both of which have evolved foregut fermentation (Barnard, 1969; Beintema, 1990; Beintema et al., 1973). Accordingly it has been proposed that RNase in these taxa might serve an enzymatic function to digest symbiotic bacteria in their foreguts so allowing bacterial derived nitrogen to be utilized efficiently (Barnard, 1969; Beintema, 1990).

Subsequent work has revealed that colobine monkeys possess two or more RNASE1 genes and that the gene duplicate products are functionally suited to a low pH (Zhang, 2006; Zhang et al., 2002). Specifically, amino acid replacements in pancreatic RNASE1 have resulted in a greater overall negative charge, and the consequent reduction in isoelectric point (pI) is thought to be an adaptation to the acidified environment of the foregut (Zhang, 2006). The origin of digestive RNASE1 genes in Asian and African colobines appears to be independent, providing compelling evidence that natural selection has shaped RNase1 evolution for its ability to break down the RNA of cellulolytic bacteria (Yu et al., 2010; Zhang, 2006). Similar RNASE1 duplications have been reported in ruminants (Breukelman et al., 1998, 2001; Kleineidam et al., 1999), as well as other herbivorous species such as the guinea pig (Van den Berg et al., 1977). Yet RNASE1 gene duplications are not restricted to herbivores; multiple copies have also been documented in some carnivores, which have also been attributed to dietary efficiency (Yu and Zhang, 2006).





Abbreviations: RNASE1, pancreatic ribonuclease gene; RNase1, pancreatic ribonuclease; RNase, ribonuclease; AGT, alanine-glyoxylate aminotransferase 1; GLUT4, glucose transporter 4.

^{*} Corresponding authors.

E-mail addresses: s.j.rossiter@qmul.ac.uk (S.J. Rossiter), syzhang@bio.ecnu.edu.cn (S. Zhang).

Bats (order Chiroptera) are divided into two suborders: the Yinpterochiroptera and the Yangochiroptera (Teeling et al., 2005). Both suborders show dietary specializations that include frugivory or nectarivory, insectivory and carnivory (Kunz and Fenton, 2003) and so offer a good opportunity for testing for a link between diet and RNASE1 duplication events. Recent studies have shown that the enzymes alanine-glyoxylate aminotransferase 1 (AGT) and glucose transporter 4 (GLUT4) have both undergone molecular adaptation in bats related to dietary specialization, probably promoting digestive efficiencies in fruit bats (Liu et al., 2012; Shen et al., 2012). The enzyme AGT was also reported to show adaptive evolution in some insect-eating bats (Liu et al., 2012). An early study reported that the expression level of RNase was very low in bat pancreatic tissue (Beintema et al., 1973), however, to our knowledge no previous study has examined the molecular evolution of ribonucleases in bats. Here we tested for molecular adaptation in the RNASE1 gene of 24 bat species that collectively exhibit a variety of feeding habits. Based on the results from other mammal groups (Yu and Zhang, 2006), we speculated that bats are good candidates for RNASE1 duplications, and that RNASE1 gene evolution in bats would show evidence of different selective regimes associated with divergent trajectories in diet.

2. Materials and methods

2.1. Sample collection and species coverage

Tissue samples of bat wing membrane were collected in the field using biopsy punches, and transferred to ethanol for DNA preservation. Bats were released *in situ* soon after capture. We sequenced *RNASE1* genes from 22 bat species representing a range of dietary specialists. These species covered eight families: Pteropodidae, Hipposideridae, Rhinopomatidae, Emballonuridae, Vespertilionidae, Molossidae, Phyllostomidae and Mormoopidae, five belonging to the suborder Yinpterochiroptera, and the others from the suborder Yangochiroptera (Table S1).

2.2. Experiments and data collection

Genomic DNA was extracted from wing membrane biopsies using DNeasy kits (Qiagen). One primer pair (F1 and R1; see Table S2) was designed based on aligned genome data from *Myotis lucifugus*, *Pteropus vampyrus* and other mammals. We successfully amplified the *RNASE1* gene from 16 bat species using this pair of primers. For amplifying bats from the families Phyllostomidae and Mormoopidae, we designed another pair of primers (F2 and R2) based on the sequence of *Taphozous melanopogon*. To ensure only one *RNASE1* copy was present in these two families, we also designed a set of degenerate primers (F3 and R3) and tested these in the species *Leptonycteris curasoae*, *Phyllostomus latifolius* and *Micronycteris megalotis* (Table S2).

Two independent PCRs were carried out per species using the following protocol: 95 °C for 5 min, 32 cycles of 95 °C for 30 s, 57 °C for 30 s and 72 °C for 30-40 s, and a final extension of 72 °C for 10 min. At least 44 positive clones from each species were sequenced using an ABI 3730 sequencer. To verify putative duplicates of the RNASE1 gene detected for the families Vespertilionidae and Molossidae, we conducted additional independent PCRs and sequencing reactions for these groups. Bat sequences that contained at least four amino acid differences, and which were obtained from at least two separate PCRs, were considered to be putative duplicate gene copies (Yu and Zhang, 2006; Yu et al., 2010). RNASE1 gene sequences of P. vampyrus and M. lucifugus were retrieved from the Ensembl using BLAT search (www.ensembl.org), and those of Artibeus jamaicensis and the douc langur Pygathrix nemaeus (AJ535682 and AF449642 respectively) along with an RNASE1B sequence from P. nemaeus (AF449643) were obtained from NCBI (www.ncbi.nlm. nih.gov). GenBank accession numbers of the bat RNASE1 genes sequenced in this study are JX624171–JX624211 and JX998072–JX998077 (Table S1).

2.3. Gene tree reconstruction

Bat *RNASE1* sequences (including noncoding regions) were aligned with Clustal W in the MEGA 5 software (Tamura et al., 2011; Thompson et al., 1994) and checked by eye. Based on the Akaike information criterion, TVMef+ Γ was the best model chosen by jModelTest 2 (Darriba et al., 2012). Phylogenetic trees were then reconstructed using both Bayesian (Ronquist and Huelsenbeck, 2003) and maximum likelihood (ML) methods (Stamatakis, 2006). Five million generations were set in the Bayesian inference with the first two million discarded as burn-in. We searched 200 tree topologies using the ML method with 1000 bootstrap replicates performed. Cases of duplications were identified from inspecting the tree and duplicates were sorted into groups (suffixed by Greek letters) based on their phylogenetic distinctiveness.

2.4. Molecular evolution analysis

We tested for gene conversion using the software GENECONV under different gscale parameters (Sawyer, 1989), and tested for positive selection using CODEML in PAML 4 (Yang, 2007). For selection tests we used the Bayesian tree topology. To determine the extent to which duplications have led to heterogeneous selection pressures, we first applied two nested branch-wise models: the free-ratio model that assumes independent ω (ratio of the rate of nonsynonymous substitutions $[d_N]$ to the rate of synonymous substitution $[d_S]$) along each branch and a one-ratio that fixes the same ω across the tree (Yang, 1998). In addition we implemented modified branch-site model A to test for positive selection in each of the detected groups of duplicated genes, in each case setting the ancestral branch to the duplicates as the foreground branch. We conducted test 2 in which ω of the foreground focal branch was fixed at one in the null model (Zhang et al., 2005). Finally, to test further for changes in selection pressure in two groups of duplicates (the Myotis γ group and *Tadarida* β group) we applied separate clade models (model C) in which a proportion of sites was allowed to differ in their ω values between the focal clade of duplicates and all of the remaining background sequences (Bielawski and Yang, 2004). Finally, since the parallel evolution of dietary RNASE1 genes in Asian and African leaf-eating monkeys has been shown to have arisen via parallel amino acid changes (Zhang, 2006), whereas ruminant RNASE1 genes underwent divergent changes (Zhang, 2003) we also compared the published sequences of leaf-eating monkey to those obtained from bats using the software MEGA 5 (Tamura et al., 2011).

3. Results and discussion

3.1. Gene sequences and duplications of Bat RNASE1

We obtained new sequences from 22 bat species that collectively exhibit a range of dietary habits (listed in Table S1) and supplemented these with published sequences from *P. vampyrus*, *M. lucifugus* and *A. jamaicensis*. In total we identified 52 *RNASE1* gene sequences from 24 bat species, each spanning the entire coding region as well as partial sequences of intron 1 and the 3' untranslated region. Six of the gene sequences were from six bat species from the suborder Yinpterochiroptera and the other 46 sequences were from 18 species from the suborder Yangochiroptera bats (see full details in Table S1). In most species—including all focal members of the New World family Phyllostomidae—just one *RNASE1* gene copy was detected, despite extensive cloning and sequencing. The absence of any *RNASE1* gene duplicates in the phyllostomids was particularly surprising given that these taxa encompass frugivory, nectarivory and insectivory (Monteiro and Nogueira, 2011; Schondube et al., 2001). This finding also agrees with published data from the phyllostomid *A. jamaicensis* (Dubois et al., 2003). Similarly, our inability to recover more than one copy of the *RNASE1* gene by cloning PCR products from *Pteropus faunulus* is strongly supported by the recent report of a single *RNASE1* copy in the *P. alecto* genome (Zhang et al., 2013), as well as our detection of just one copy in the *P. vampyrus* genome.

Two or more candidate *RNASE1* genes were discovered in five insect-eating members of the family Vespertilionidae (*M. lucifugus*, *M. altarium*, *M. ricketti*, *Ia io and Murina leucogaster*) and in two

insect-eating members of the Molossidae (*Tadarida brasiliensis* and *T. insignis*), all of which belong to the suborder Yangochiroptera. Cloning of PCR products from *M. lucifugus* yielded six candidate *RNASE1* genes, four of which corresponded to hits in the *Myotis* genome, and three further *RNASE1* loci were discovered in the *Myotis* genome but were not recovered from PCR and cloning. It is interesting that three copies of *RNASE1* were also very recently reported in the *Myotis* davidii genome (Zhang et al., 2013), although these were not analyzed here. According to our results, most bat *RNASE1* genes appear to be functional (Fig. 1), although pseudogenes were also cloned from

	E spelaea MAC P, raptor P, faunulus H prati H prati R. hardwickii T. melanopogon T. nachypus L io 0.1 L io 0.2	QEKSLILFP-MLLLV G L. A. S-L. A. S-L. V. W. - L.V.M - L.V. - L.V. - L.V. - L.V.	L L A L GWV E P L L G 	R E S QA KK F I R QHM	1DPVSYPNISSSY N	' C N QMM R R R NM' 	Г K GWC K PMNT F V H E P I	.VNVQAVCFQGKVTCKDG - (.D	Q S
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	P. vanpyrus P. vanpyrus H. pratti R. hardwickii T. melanopogon T. pachypus Lio α1 Lio α2 Lio β1	A S L A S L A S L A S L A S L A S L A S L A S C A S S L A S S L S S S S S S S S S S S S		K. R.M. Q		D .	R V	. D	• •
P_ministi A. S. L. V. PEE M. D. C. R. V. S. D. E. Amadexisti I. V. M. V. V. S. A. N.	P. fauniulus	. A S - L L . V . M L . V . M F L . V L . V L . V V 	FRE V Q . S . V P . S . V Q . S	K R . M Q K RTM Q	D G G T NDA N . 		R V	. D	
Interlington I. V. N. V. V. S. K. R. M. O. G. N. D. N. I. N. D. N. I. N. D. I. N. N. D. N. N. D. N. N. D. N. N. N. D. N. N.<	H. pratti R. hardwickii T. melanopogon T. pachypus I. io α1 L io β1	V . M V . M F V 	V Q . S V P . S V Q . S	K R . M Q K RTM Q K RS	GTNDAN. GN GNN N	E .	V I S V		· .
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	$\begin{array}{cccccccccccccccccccccccccccccccccccc$		V Q . S	K RS O		E K .	0 V V S	AD I L N -	. P
$ \begin{array}{c} protection for the form the for$	<i>I. pachypus I. io</i> α1 <i>I. io</i> α2 <i>I. io</i> β1 <i>I. io</i> α2 <i>I. io</i> β1			K K Q	D . R . S S N . T .	K .	. È . R V	. D I N I N	. P
$ \begin{array}{c} Lio 22 \\ Lio 11 \\ Lio 11 \\ Lio 11 \\ Lio 12 \\ Lio $	<i>L</i> io α2 <i>L</i> io β1	L . V	V 0 S	K.R.M.Q KRMO	DG SNN.N.		E.R. V	D.I.L.NINI	K P P
$ \begin{array}{c} L_{D} \\ L_{D} \\ metry $	L io B1	T 17	v Q . s	K WVM Q		HQ	. ERQ V M	. D I . L N I N	P
M. Lecogester B T. T. V. V. V. V. S. K. R. M. G. DGS SNN H. E. R. V. D L. N. N. H. R. N. N. H. R. N. N. H. R. V. D L. N. N. N. H. R. V. D L. N. N. N. H. H. N. N. N. N. N. H. H. N. N. <td< th=""><th>M leucogaster a</th><th></th><th> V Q . S</th><th>K R . M Q K R . M O</th><th> D S NN . N . DGDSSNN . N .</th><th>. T W</th><th> R Y . L V . I</th><th>. D I . L N I N D</th><th>. Р . Р</th></td<>	M leucogaster a		V Q . S	K R . M Q K R . M O	D S NN . N . DGDSSNN . N .	. T W	R Y . L V . I	. D I . L N I N D	. Р . Р
M diarring of the set of the se	M. leucogaster β	T L . V	V Q . S	KR.MR.Q	DGDSSNN.N.	н.	. E . R V	. D L . EN I N	. P
M. diatrium*12 P. L. V. V. O. S. K. R. M. O. GG. STN. N. L. E. R. V. I. D. I. L. NTI. N. M. diatrium*12 P. L. V. V. O. S. K. R. M. O. G. STN. N. H. E. R. V. D. L. NTI. N. M. diatrium*12 P. L. V. V. O. S. K. R. M. O. G. STN. N. L. E. R. V. D. L. NTI. N. M. diatrium*12 L. V. V. O. S. K. R. M. O. G. STN. N. L. E. R. V. D. L. NTI. N. M. lactingues 13 L. V. V. O. S. K. R. M. O. G. STN. N. H. E. R. V. D. L. NTI. N. M. lactingues 14 L. V. V. O. S. K. R. M. O. G. STN. N. H. E. R. V. D. L. NTI. N. M. lactingues 14 L. V. V. O. S. K. R. M. O. G. STN. N. H. E. R. V. D. I. L. NI. N. M. lactingues 14 L. V. V. O. S. K. R. M. O. G. STN. N. L. E. R. V. D. I. L. NI. N. M. lactingues 14 L. V. V. O. S. K. R. M. O. G. STN. N.	$M. altarium \alpha$ $M. altarium \beta$	L . V	V Q . S	K.R.M.Q KRMO	DGSNN.N. DG-SNN N		. E . R V	. D I . L NI N	. P
	$M. altarium \gamma 1$	P L . V	V Q . S	K R . MM . QQ	DG STN . N .	L .	. E . R V I	. D I . L NT I N I	ΚP
	M. altarium γ_2 M. lucifuans α_1	P L . V	V Q . S	K.R.M.Q	GG SNN . N . DG SNN N	H .	. E . R V I	. D I . L NI N	. P
	M. lucifugus $\alpha 2$	P L . V	V Q . S	K R . M Q	DG SNN . N .	L .	. E . R V	. D L N I N	P
	M. lucifugus B1 M. lucifugus B2	L . V	V Q . S	КР.МQ	DG SNN . N . GG SNN N		. E	. D I . L N I N I	KP
	M. lucifugus $\gamma 1$	L . V	V Q . S	K R . M Q	DG - LSNN . N .	H .	. E . R V	. D I . L N I N I	ΚΡ
M. Implement Description Description <thdescription< th=""> <thdescription< th=""></thdescription<></thdescription<>	M. lucifugus γ_2	L . V	V Q . S	K R . M Q	DG SNN . N .	H .	. E . R V	. D L N I N	. P
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. lucifugus $\gamma 3$ M. lucifugus $\gamma 4$	L . V	V Q . S	K W. MM. Q	DG S SN . N .	QL.	. E R Q V I	. D I . L NI N I	. г КР
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. lucifugus γ5	L . V	V Q . S . V	K P. M Q. R.	GG SNN . N .		. E . R V I . K	ADI.L.NTPH	. P
$ \begin{array}{llllllllllllllllllllllllllllllllllll$	M. ricketti γ2	L . V	V Q . S	K W. MM . Q	DG SND . N .	L .	. E . L V I	. D I . L NI N I	. г КР
millionsin Thransitiensis ct S. V. A. V. L. O. S. K. G. R. M. O. D. O. S. S. N. N. n. n. E. R. V. I. E. I. L. N. N. Thransitiensis f1 S. V. A. V. L. O. S. O. T. F. E. O. D. O. S. S. N. N. R. E. R. V. D. N. I. L. N. N. Thransitiensis f2 S. V. A. V. L. O. S. O. T. F. E. O. S. DON. N. N. N. R. E. R. V. D. R. I. I. P. N. N. Thransitiensis f2 S. V. A. V. L. O. S. O. G. G. R. KQ. S. N. N. N. N. K. K. K. V. D. V. D. V. D. V. D. N. R. I. N. N. N. Tinsignis f2 S. V. A. V. L. O. S. O. T. G. M. K. S. N. N. N. N. N. D. K. K. K. V. D. V. D. I. I. N. N. N. Tinsignis f2 S. L. A. V. L. O. S. O. T. G. M. K. S. N. N. N. N. N. D. E. R. V. E. D. E. I. I. N. N. Tinsignis f3 S. L. A. V. L. O. S. O. T. G. M. K. S. N. E. D. E. I. I. N. N. N. Tinsignis f4 S. L. A. V. L. O. S. M. G. G. M. M. S. S. N. R. N. N. N. N. N. N. M. N. N. N. N. N. N. N. N. N. N. N.	M. ricketti ý3	L . V	V Q . S	K R Q	DG SNN . N .	H .	. E . R V I	. D I . L N I N	. P
$ \begin{array}{c} Transitions & 0 & & S + V & A & & V & L & Q & S & K & G & R & Q & D & D & N & S N & T & & K & E & R & & V & & D & R & L & N & N & N & T & T & T & S & S & V & A & & V & L & Q & S & Q & Q & Q & G & R & Q & S & S & S & N & N & D & & K & E & R & & V & D & R & L & P & N & R & R & N & T & T & T & T & T & T & T & T & T$	T. plicata		V Q . S	K R . M Q	D . Q S . NN . N .	К.	. E . R V I	. D I . L NI N	. r
$ \begin{array}{c} 1 \\ product (1, 1, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,$	T. brasiliensis α	S - V . A	V L . Q . S	K G. R Q	D . Ñ . S NN . T .	K .	. E . R V	. D I . L N I N I	Н.
$ \begin{array}{c} T. \ breasiliens \ bre$	T. brasiliensis B2		V L . Q . S	Q. 1 P. EK. Q QGPG. M Q	. SDGN NNPN . . SNGN . QPN . N .		. E . K V	KD.E.I.I.PNIR.N	. N
$ \begin{array}{c} 1, msrgmis \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	T. brasiliensis β_3	S - V . A	V L . Q . S	Q G. R KQ	. SN. RHLPN.N.	K K K . K .	. VDA V	TD.K.I.N.RNIR.N	. N
$ \begin{array}{c} T_{1} nsignis \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	T. insignis β_1		V Q . S	Q.TG.MK	D . N . S N N . N	. D K . . D E G K .	. E. K V	ED.K.IDRILNKRI	н. DT
$ \begin{array}{c} L \text{ instgins } \text{B} \\ T \text{ instgins } \text{B} \\ S \\ $	T. insignis β_2	S - L . A	V L . Q . S	Q.TG.M.K	. SN.N NNPN.	. D E K .	. Q. R V	ED.E.I.I.DNIR.SNKI	DT
$ \begin{array}{c} T. insignis \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$	T. insignis B4		V L . Q . S V Q . SM .	Q. 1 G. M W	. SK.R. PN.T.	. D	. Q. к к	KD. EDI. IEPNIL. N	. N
$ \begin{array}{c} Lnsgms \ po \\ Mn \ megalotis \\ C - L \ I \\ V \\$	T. insignis B5	S - L . A	V L . Q . SM .	QGPG.M. Q	. S N . R P N . N .	. D KK . K .	. V . A L	KD.EDI.I.RNIR.N	. N
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. megalophylla		V L . Q . SM . V O . S	QGPG.MQ КR.МО	. SN.K. PN.N. . DGHSG-NPN.	. D C. R.	. V . A L	KD. EDI. I. KNIR N . D I. L NI N	. N . P
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. megalotis	C L . I	V A Q . S	K R . M Q	DGH . S - N	<u>K</u> .	. E . R V	. D I . L N I N	. P
$ \begin{array}{c} C. perspicillata \\ C. willosim \\ M. F. C L. I V. AQ. S. K. R. M. Q. DG. T-N V. D. I. L. NI. N. A. jamaicensis \\ M. F. C L. I. V. AQ. S. K. W. M. Q. DG. T-N. L V. D. I. L. NI. N. N. A. jamaicensis \\ \hline V. M. Q. S. S.$	L. curasoae	L . G L . T	V AQ. S	кк.мQ КW.МQ	DGH . S - N	K .	. E . K V	. D I . L N I N	. Р . Т
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	C. perspicillata	C L . I	V A Q . S	K R. M Q	DGH . S - N	K .	. E . R V	. D I . L N I N	. P
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	A. jamaicensis			K W. M Q	DG T - N	L K .	. E . R I	. D I . L N I N	
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $						120		1.60	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	E and an NON	VOCCCCNCLTDODIT	TCCKNDKCANDT	CO KKDULINACI		130		169	pl 0.51
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	P. raptor H	HK.N.R	NN	M	.T	. L			9.50
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	P. vampyrus D		NN						8.95
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	H. pratti H	Н	SN.V	E I I		I			8.67
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	R. hardwickii H	Н Н N	AN.E						8.83
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	T. pachypus H	HKV.	S	E	L . O . 				9.10
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$L io \alpha 1^{-1}$. H	HKV.	SN	EI					8.98
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	<i>L</i> io β1	HKK.MV.	S						9.16
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. leucogaster O H	HKV.	SN	El		A			8.67
M. altarium β HK. N. K. HE.S. N.D.Q. H-E. E. H. IEDQSSEMPHESSVSPSSLLPLREMTQGRAAMQHKSVSL M. altarium γ1 HK. K. VKSS. D.D.Q. H-E. E.	$M. altarium \alpha$ H	HKV.	S						9.11
M_{L} and M_{L}	$M. altarium \beta$ H	HK.NHE.	SN.D.Q.	. H E	H	I EDQ S S I	EMPHESSVSPSSLLPI	REMTQGRAAMQHKSVSLD	8.17
<i>M. altarium γ</i> 2 HK K V. S N E E	$M. altarium \gamma^2$ H	HKV.	SN	E	8				9.13
M. lucifugus al	M. lucifugus 0.1 H M. lucifugus 0.2 H	НКV. нк к v	SN.D SN.D.	EI	3	A . S			8.99
M lucifugus β1 HK N N HK S N D Q E V E IGDQSSEMPHESSVSPSSLLPLKEMTQGRAAMOHKSVSLA	M. lucifugus β1 H	HK.NHK.	SN.D.Q.	E V I		I GDQSS	EMPHESSVSPSSLLPI	K EMT Q G R A AM Q H K S V S L A	7.74
$m. uccjuggs p_ = nk, n. n. n = e.s. n. D. Q. n = e.e. = e.e. = 10DQ3 empless stars stelet kemi QGKAAMQRKS v Le$	M. lucifugus γ1 H	HR	S	E		S	EMFLESSVSFSSLLFI	.REMIQURAAMQHKSVSLA	8.45
<u>M.lucifugus 12</u>	M. lucifugus γ_2 H	HKV.	RN.D	EI		A . S			8.99
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	M. lucifugus γ4 H	HKVK	SS.E.F.D.E.	. H E V . O	3 H				8.18
$M_{\rm tric/fugus}\gamma 5$, HK, R, K,, E, N, Q, N, D, N, H - E,, A,, A	M. lucifugus $\gamma 5$ H	HKE.	NQN.DN	. H E	L	A			9.26
$M \rightarrow M$	M. ricketti γ2 H	HKVK	SSD.D.H.	. H E V . O	Б				8.45
<i>M. ΠCKetti γ2</i> HK K VKSS D. D. H H E V. G A H	M. ricketti y3 H	HKV.	SN	El		Ι			9.10
M. ricketti γ3	T. plicata	НКНЕ	N R N H .	EKQ					8.99
М. ricketti Y2 HK K. VKSS D. D. H. H E. V. G. А. H	T. brasiliensis α H	HKHE HK R RV K	NRNH. NRNH	NEKQI P-A O T I) - R				8.99
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	T. brasiliensis β_2 H	H R R V	N R N . K	. P - T Q . T V	/				9.48
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	T. brasiliensis β3 H	HK.RRV	NRH.N.K	. P - A Q I	OK R L	N			10.11
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$T. insignis \beta 1$ H	HK.KRV	N R N . K . N .	. LTTN.Q.TI	. D				9.22
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	T. insignis B2 H T insignis B3	HK.KRV	N.R.N.K.N. NRNKN	GLTTQ.TI	D				8.98
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	T. insignis B4 H	HK.KFL.NV.A	NRK.H.	. P - TM . Q . T I					9.62
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	1. insignis 35 H T. insignis 36 H	HK.KRV	N.R.N.K.H. N.R.NKH	. P - T Q . T V	(L.G-				9.63
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. megalophylla H	НКК	N N	E.QI					9.24
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	M. megalotis H P. latifolius	HKN	NN	E.QI	Δ	I			8.99
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	L curasoae	ΗΚΝ	N N	V . Q 1		A			8.99
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	C. perspicillata H C. villosum H	HKN	NN	TE.QI		S			8.99
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	A. jamaicensis	HKN	NN.V	E.QI	8				9.11

Fig. 1. RNase1 sequence alignment. Only functional bat RNase1 sequences are listed with predicted pl values. The pl values were calculated using Compute pl/Mw tool from the ExPASy server (http://web.expasy.org/compute_pi). Groups of duplicates were identified based on distinct clades in Fig. 3, and suffixed with Greek letters.

a few taxa: *Myotis ricketti, I. io* and *Nyctalus plancyi* (Fig. 2). In fact for the latter taxon, only a pseudogene was cloned, possibly due to PCR amplification bias. It follows that among all of the focal species studied here, there is a combination of functional and nonfunctional gene copies, with some taxa possessing both forms.

3.2. Multiple RNASE1 duplication events in bats

Bayesian and ML phylogenetic tree reconstructions both indicated that bat *RNASE1* gene duplications have occurred independently after the divergence of the families Vespertilionidae and Molossidae (Figs. 3A and S1), previously dated to be around 47 million years ago (Teeling et al., 2005). These two tree topologies were concordant with the exception of the position of *L. curasoae*. In the former family, *RNASE1* gene duplications appear to have occurred separately in the genera *Ia*, *Murina* and *Myotis*, with at least three main groups of *RNASE1* genes seen in the genus *Myotis*. In the *Myotis* β group, three *RNASE1* genes were identified with coding sequences that were longer, with an additional stretch of 39 amino acids at the C-terminal end. The resulting novel sequence of the *Myotis* RNase1 protein might thus confer additional function. In the Molossidae, multiple *RNASE1* copies were detected in *T. insignis* and *T. brasiliensis* but only one copy was seen in *T. plicata*. Thus it is likely that duplication events occurred prior to the divergence of these taxa, with a possible subsequent loss in the latter species. Tests of gene conversion revealed just one possible event involving *T. brasiliensis* β 3 and *T. insignis* β 3 genes, suggesting that this phenomenon has not had a major impact on gene tree reconstruction (Table S3).

3.3. Adaptive evolution of bat RNASE1

Most of the bat species that were found to show gains of functional *RNASE1* genes are predominantly insectivorous; however, *I. io* and *M. ricketti* are also carnivorous and piscivorous, respectively. To assess whether gene duplicates show evidence of molecular adaptation, we undertook tests of positive selection. The estimated ω value showed considerable variation among different bat lineages, pointing to variable selective pressures (see Fig. 3B and Table S4). The free-ratio model fitted the data significantly better than the one-ratio model, suggesting that this variation is real. Inspection of tree-wide ω values revealed multiple episodes of positive selection in *RNASE1* genes following duplications, implying functional adaptation of the enzyme in these lineages.

Further tests for positive selection were conducted with either branch-site models, in which the ancestral branches leading to different *RNASE1* groups were considered as focal lineages, or clade models

	Signal peptide region — 60						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	A T G G C C C A G G A G A A G T C C C T C A T C C T G T T C C C A C T G T T A G T C C T G G T G C T G T T A G T G C T G T A						
	120						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	GGGT G G GT C C A G C C T T C C C T G G G C A A A GA A T C A C G G G C C A T G A A G T T C C A G A G G C A A C A C						
	180						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	A T G G A C C C A G A C A G C T A C C C C A G C A A C A A T T C C A A C T A C T G C A C C C A A A T G A T G A G G C G C 						
	240						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	T G G A A T A T G A C A A G G G A C G G T A C C						
	300						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	T A GA T GT C C A G G C C A T T T G C C T C C A G G G A A C A T C A C T G C A A G A A T G G G C A G C C C A A C T						
	360						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	G C C A C - A A G A G C A G C T C C A G C A T G A A G A T C A T G G A C T G T C G T G T G A C G A G T G G C T C C A A G 						
	420						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	TACCCCCAACTGTGCCTACCGGACCAGCCAGCCAGAAGGAGGCACATTATTGTGCCCTGTGAG						
	*						
I. io β1 I. io β2 N. plancyi M. ricketti γ1	G G A A A C C C G T A C G T G C C T G T C C A C T T T G A T G C T T C A G T A G A G G T C T C C A C C T G A						

Fig. 2. Sequence alignment for three nonfunctional *RNASE1* genes. The pseudogenes are aligned with the functional β_1 gene from *Ia io*. An in-frame stop codon and two insertions are shown in boxes. The 1 bp insertion in *I. io* β_2 will cause two in-frame stop codons and the 2 bp insertion in *Myotis ricketti* γ_1 will lead to five in-frame stop codons. The asterisk indicates the stop codon in the functional *RNASE1* gene reading frame.



Fig. 3. Gene tree and molecular evolution tests. (A) Bayesian phylogenetic tree reconstructed from bat *RNASE1* genes with posterior probabilities shown. Bats from the suborder Yinpterochiroptera were used to root the tree (shown in a gray block). The two dotted lines denote the ancestral branches of families Vespertilionidae and Molossidae, respectively. The duplicated *RNASE1* genes in the two families are mainly grouped as α , β and γ , and three detected pseudogenes are marked by the symbol Ψ . (B) Branches shown as a thick line indicate that ω was estimated to be >1 based on branch-wise test of selection (free-ratio model). Branch numbers are shown in the tree and listed in Table S4, in which detailed parameters are also given.

applied to two of these groups (Fig. 3A). Of the nine branch-site models implemented, each corresponding to the ancestral branch of a different group of duplicates, we found strong positive selection only in the *Tadarida* β group (Table 1) with four sites (amino acid

positions: 1, 15, 64, 104) seen to have Bayes empirical Bayes values > 0.5. On the other hand, no significant positive selection was detected using these models in members of the Vespertilionidae. Evidence of divergent selection from both Clade models was

Table 1

Tests for positive selection on bat RNASE1 genes showing significant results.

Model	Log-likelihood	Parameters of foreground lineage	P value
Alternative hypothesis 1: Branch-site modified model A (branch leading to <i>Tadarida</i> β group) Null hypothesis 1: Fixed ω_2 in branch leading to <i>Tadarida</i> β group at one	- 4550.132 - 4553.422		0.01
Alternative hypothesis 2: Clade model C (<i>Myotis</i> γ group vs. other bats) Null hypothesis 2: M1a (nearly neutral)	- 4541.951 - 4553.688		<0.001
Alternative hypothesis 3: Clade model C (<i>Tadarida</i> β group vs. other bats) Null hypothesis 3: M1a (nearly neutral)	- 4511.797 - 4553.688		<0.001
Alternative hypothesis 4: Clade model C (<i>Tadarida</i> β group vs. other bats) Null hypothesis 4: Fixed ω_3 in <i>Tadarida</i> β group at one	- 4511.797 - 4519.559		<0.001

consistent with neofunctionalization with 37.6% and 28.1% sites in the *Tadarida* β and *Myotis* γ clades under divergent selection. Moreover, in the case of *Tadarida* β the estimated ω in the foreground clade was >2, signifying positive selection (Table 1).

Our comparisons of RNASE1 duplicate sequences in bats revealed divergent amino acid changes, which also differed from the changes reported in duplicate sequences of the douc langur (Table S5). In colobine monkeys, the protein product of the new RNASE1 gene duplicate has a lower pI value than its parental copy, and thus shows molecular adaptation for a low pH environment (Zhang, 2006). Here we predicted the pI values of bat RNase1 proteins, but found no such straightforward relationship (Fig. 1). Instead, all of the bat gene products were predicted to have relatively high isoelectric points. Previously it was shown that RNase1 proteins with higher isoelectric points are suited to the breakdown of double-stranded RNA (Zhang, 2006), supporting the theory that ribonuclease activity evolved as a means of protection against viruses (Sorrentino, 2010; Sorrentino and Libonati, 1994). Indeed, host defense might be the original evolutionary function of ribonucleases: Cho and Zhang (2007) found that RNases in zebrafish and other divergent lineages of vertebrates also possess host defense properties (Cho and Zhang, 2007). Therefore, given the lack of evidence that bat RNASE1 genes are adapted to an acidified environment, we cannot rule out the possibility that the duplication events and changes in selection pressure detected in our study are related to roles other than diet, such as pathogen defense. In this context it is noteworthy that while the Vespertilionidae has undergone dietary diversification, the Molossidae is almost exclusively insectivorous (Nowak, 1994), again presenting no obvious link between the proliferation of RNASE1 genes and dietary specialization in bats.

3.4. Conclusions

In summary, we screened the *RNASE1* gene in more than 20 bat species, covering both main clades and diverse feeding types. PCR, cloning and sequencing revealed more than two gene copies from seven bat species from two families, Vespertilionidae and Molossidae. Duplications in these families appear to have arisen independently. All species that were found to have *RNASE1* duplicates are mainly insect-eating. Evidence of positive selection following duplication events suggests these are adaptive, however, high predicted pl values do not indicate a clear link with diet. Instead, bat *RNASE1* genes may have other roles, including defense against viruses although function assays are needed to investigate whether this is indeed the case.

Acknowledgments

We thank Junpeng Zhang for help with sample collection. This study was supported by the Chinese National Science Foundation (Grant No. 31172077) to SZ and a Royal Society Fellowship (UK) to SJR.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2013.04.035.

References

- Barnard, E.A., 1969. Biological function of pancreatic ribonuclease. Nature 221, 340–344. Beintema, J.J., 1990. The primary structure of langur (*Presbytis entellus*) pancreatic ribonuclease: adaptive features in digestive enzymes in mammals. Mol. Biol. Evol. 7, 470–477.
- Beintema, J.J., Kleineidam, R.G., 1998. The ribonuclease A superfamily: general discussion. Cell. Mol. Life Sci. 54, 825–832.

- Beintema, J.J., Scheffer, A.J., van Duk, H., Welling, G.W., Zwiers, H., 1973. Pancreatic ribonuclease distribution and comparisons in mammals. Nat. New Biol. 241, 76–78.
- Bielawski, J.P., Yang, Z., 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. J. Mol. Evol. 59, 121–132.
- Breukelman, H.J., van der Munnik, N., Kleineidam, R.G., Furia, A., Beintema, J.J., 1998. Secretory ribonuclease genes and pseudogenes in true ruminants. Gene 212, 259–268.
- Breukelman, H.J., Jekel, P.A., Dubois, J.F., Mulder, P.P., Warmels, H.W., Beintema, J.J., 2001. Secretory ribonucleases in the primitive ruminant chevrotain (*Tragulus javanicus*). Eur. J. Biochem. 268, 3890–3897.
- Cho, S., Zhang, J., 2007. Zebrafish ribonucleases are bactericidal: implications for the origin of the vertebrate RNase A superfamily. Mol. Biol. Evol. 24, 1259–1268.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.
- Dubois, J.Y., Ursing, B.M., Kolkman, J.A., Beintema, J.J., 2003. Molecular evolution of mammalian ribonucleases 1. Mol. Phylogenet. Evol. 27, 453–463.
- Kleineidam, R.G., Pesole, G., Breukelman, H.J., Beintema, J.J., Kastelein, R.A., 1999. Inculsion of cetaceans within the order Artiodactyla based on phylogenetic analysis of pancreatic ribonuclease genes. J. Mol. Evol. 48, 360–368.
- Kunz, T.H., Fenton, M.B., 2003. Bat Ecology. The University of Chicago Press, Chicago.
- Liu, Y., Xu, H., Yuan, X., Rossiter, S.J., Zhang, S., 2012. Multiple adaptive losses of alanine-glyoxylate aminotransferase mitochondrial targeting in fruit-eating bats. Mol. Biol. Evol. 29, 1507–1511.
- Monteiro, L.R., Nogueira, M.R., 2011. Evolutionary patterns and processes in the radiation of phyllostomid bats. BMC Evol. Biol. 11, 137.
- Nowak, R.M., 1994. Walker's Bats of the World. The Johns Hopkins University Press, Baltimore.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
- Sawyer, S., 1989. Statistical tests for detecting gene conversion. Mol. Biol. Evol. 6, 526–538.
- Schondube, J.E., Herrera-M, L.G., Martínez del Rio, C., 2001. Diet and the evolution of digestion and renal function in phyllostomid bats. Zoology 104, 59–73.
- Shen, B., Han, X., Zhang, J., Rossiter, S.J., Zhang, S., 2012. Adaptive evolution in the glucose transporter 4 gene Slc2a4 in Old World fruit bats (family: Pteropodidae). PLoS One 7, e33197.
- Sorrentino, S., 2010. The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. FEBS Lett. 584, 2194–2200.
- Sorrentino, S., Libonati, M., 1994. Human pancreatic-type and nonpancreatic-type ribonucleases: a direct side-by-side comparison of their catalytic properties. Arch. Biochem. Biophys. 312, 340–348.
- Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
- Teeling, E.C., Springer, M.S., Madsen, O., Bates, P., O'Brien, S.J., Murphy, W.J., 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307, 580–584.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
- van den Berg, A., van den Hende-Timmer, L., Hofsteenge, J., Gaastra, W., Beintema, J.J., 1977. Guinea-pig pancreatic ribonucleases. Isolation, properties, primary structure and glycosidation. Eur. J. Biochem. 75, 91–100.
- Yang, Z., 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. 15, 568–573.
- Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.
- Yu, L., Zhang, Y.P., 2006. The unusual adaptive expansion of pancreatic ribonuclease gene in carnivora. Mol. Biol. Evol. 23, 2326–2335.
- Yu, L., Wang, X.Y., Jin, W., Luan, P.T., Ting, N., Zhang, Y.P., 2010. Adaptive evolution of digestive RNASE1 genes in leaf-eating monkeys revisited: new insights from ten additional colobines. Mol. Biol. Evol. 27, 121–131.
- Zhang, J., 2003. Parallel functional changes in the digestive RNases of ruminants and colobines by divergent amino acid substitutions. Mol. Biol. Evol. 20, 1310–1317.
- Zhang, J., 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat. Genet. 38, 819–823.
- Zhang, J., Zhang, Y.P., Rosenberg, H.F., 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat. Genet. 30, 411–415.
- Zhang, J., Nielsen, R., Yang, Z., 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22, 2472–2479.
- Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K.A., Fang, X., Wynne, J.W., Xiong, Z., Baker, M.L., Zhao, W., Tachedjian, M., Zhu, Y., Zhou, P., Jiang, X., Ng, J., Yang, L., Wu, L., Xiao, J., Feng, Y., Chen, Y., Sun, X., Zhang, Y., Marsh, G.A., Crameri, G., Broder, C.C., Frey, K.G., Wang, L.F., Wang, J., 2013. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science 339, 456–460.