Article

The Fast and the Furriest: Investigating the Rate of Selection on Mammalian Toxins

Leah Lucy Joscelyne Fitzpatrick [1,2,*], **Vincent Nijman** [1,2], **Rodrigo Ligabue-Braun** [3]
and K. Anne-Isola Nekaris [1,2]

1   Nocturnal Primate Research Group, Department of Social Sciences, Oxford Brookes University, Oxford OX3 0BP, UK
2   Centre for Functional Genomics, Department of Health and Life Sciences, Oxford Brookes University, Oxford OX3 0BP, UK
3   Department of Pharmacosciences, Federal University of Health Sciences of Porto Alegre (UFCSPA), Avenida Sarmento Leite 245, Porto Alegre 90050-130, Brazil
*   Correspondence: 19148055@brookes.ac.uk

Abstract: The evolution of venom and the selection pressures that act on toxins have been increasingly researched within toxinology in the last two decades, in part due to the exceptionally high rates of diversifying selection observed in animal toxins. In 2015, Sungar and Moran proposed the 'two-speed' model of toxin evolution linking evolutionary age of a group to the rates of selection acting on toxins but due to a lack of data, mammals were not included as less than 30 species of venomous mammal have been recorded, represented by elusive species which produce small amounts of venom. Due to advances in genomics and transcriptomics, the availability of toxin sequences from venomous mammals has been increasing. Using branch- and site-specific selection models, we present the rates of both episodic and pervasive selection acting upon venomous mammal toxins as a group for the first time. We identified seven toxin groups present within venomous mammals, representing Chiroptera, Eulipotyphla and Monotremata: KLK1, Plasminogen Activator, Desmallipins, PACAP, CRiSP, Kunitz Domain One and Kunitz Domain Two. All but one group (KLK1) was identified by our results to be evolving under both episodic and pervasive diversifying selection with four toxin groups having sites that were implicated in the fitness of the animal by TreeSAAP (Selection on Amino Acid Properties). Our results suggest that venomous mammal ecology, behaviour or genomic evolution are the main drivers of selection, although evolutionary age may still be a factor. Our conclusion from these results indicates that mammalian toxins are following the two-speed model of selection, evolving predominately under diversifying selection, fitting in with other younger venomous taxa like snakes and cone snails—with high amounts of accumulating mutations, leading to more novel adaptions in their toxins.

Keywords: mammals; selection rates; dN/dS; venom evolution; Primates; Eulipotyphla; Chiroptera; Monotremata

Key Contribution: For the first time, we compared the selection rates of all known mammal toxins, to provide new perspectives on what drives toxin evolution within venomous mammals. We show that all but one group of toxins are evolving under intense episodic and pervasive diversifying selection, indicating a larger proportion of mutations that can lead to unique adaptions, such as new targets or greater potency of venom. This places mammals with snakes and cone snails together due to the rate of evolution occurring on their toxins, creating opportunities to compare mammals to other venomous groups.

1. Introduction

The advent of venomous animal transcriptomes and genomes, advances in proteomics and the reduction in costs for next-generation sequence technology has allowed advances

in the study of venom evolution [1,2]. Venom is an extraordinary trait to document the mechanics of evolution, having been identified within >250,000 animal species, with over 11 different functional roles and posited to have evolved independently in 101 different animal lineages [3]. In addition, the thousands of identified toxins within animal venoms and different venom delivery systems provide examples of parallel and convergent recruitment across and within clades [4].

Random mutations in coding genes that lead to variation within traits are what enables species to adapt to their environments. Novel mutations may become fixed within an organism or population if they provide an advantageous adaption against various selection pressures encountered throughout an organism's life (i.e., predation, sexual selection or availability of resources) [5]. Measuring the rate of synonymous mutations (known as dS) in comparison to neutral mutations (known as non-synonymous; dN) found within a gene or protein of interest can allow researchers to infer if adaption is occurring and which selection pressures may be responsible. Toxins from venomous animals are particularly interesting to evolutionary biologists as they often contain a high proportion of synonymous mutations and are observed to be some of the fastest evolving genes on the planet due to their rapid accumulation of synonymous mutations over a relatively short evolutionary time [6–8]. In addition, toxins are often represented with multiple isomers or homologous sequences within a venom—in part due to the neo- or sub-functionalisation of toxins recruited into the venom from duplication of non-toxin genes. These homologous sequences can be well conserved between species, families or even orders [9]. It also leads to a greater observed diversity of functions and prey targets due to these homologous sequences.

In 2015, Sungar and Moran examined the rate of selection acting on toxin groups available and found a link between the evolutionary age of a venomous taxon and the rate of selection acting on their toxins [9]. Older groups (e.g., Chilopoda, Cnidaria and Araneae) had a higher percentage of purifying (dS < dN) selection on their toxins, in comparison to younger groups (e.g., Conidae and Serpentes), which had greater diversifying (dS > dN) selection acting on their toxins. The authors presented this as a 'two-speed' mode of evolution within toxins: that a rapid and diversifying recruitment of toxins is then followed by a retention and stabilising of toxins under purifying selection. A drawback of the study was that the data had come exclusively from predatory venoms—the study used no data from either mammals or fish.

In mammals, venom represents a unique model to investigate evolution, as it has evolved independently in eight lineages in four different clades with five different functional uses, the most diverse of any other venomous group [10,11] (Figure 1, Table 1).

These uses include intraspecific, both sexual (*Ornithorhynchus anatinus*) [31] and territorial (*Nycticebus* spp. and *Xanthonycticebus pygmaeus*) [32], haematophagy (Desmodontinae) [33] and food storage (Eulipotyphla) [13]. Currently, there are over 100 different molecules (both toxic and non-toxic) identified from mammal venoms. The exact purpose of many of these molecules to their target organisms remains unclear, due to the small amount of venom produced by many of these animals, difficulty in collecting samples from rare and endangered taxa, or difficulty in replicating the exact function of these venoms (particularly for intraspecific usage).

The venom of the platypus has been well documented in part due to the platypus being one of the first mammals to receive a full genome assembly [14,34]. The venom of the platypus is noted for being highly painful although no human deaths are recorded. Platypus venom is exclusively found within males during the breeding season (August to October, January to March) [35]. The most recognisable toxins found within platypus venom are the OvDLPs (defensins), CRiSPs (known as Cysteine-rich secretory proteins) and Kunitzs, all thought to play a role in producing pain via neurotoxicity [36].
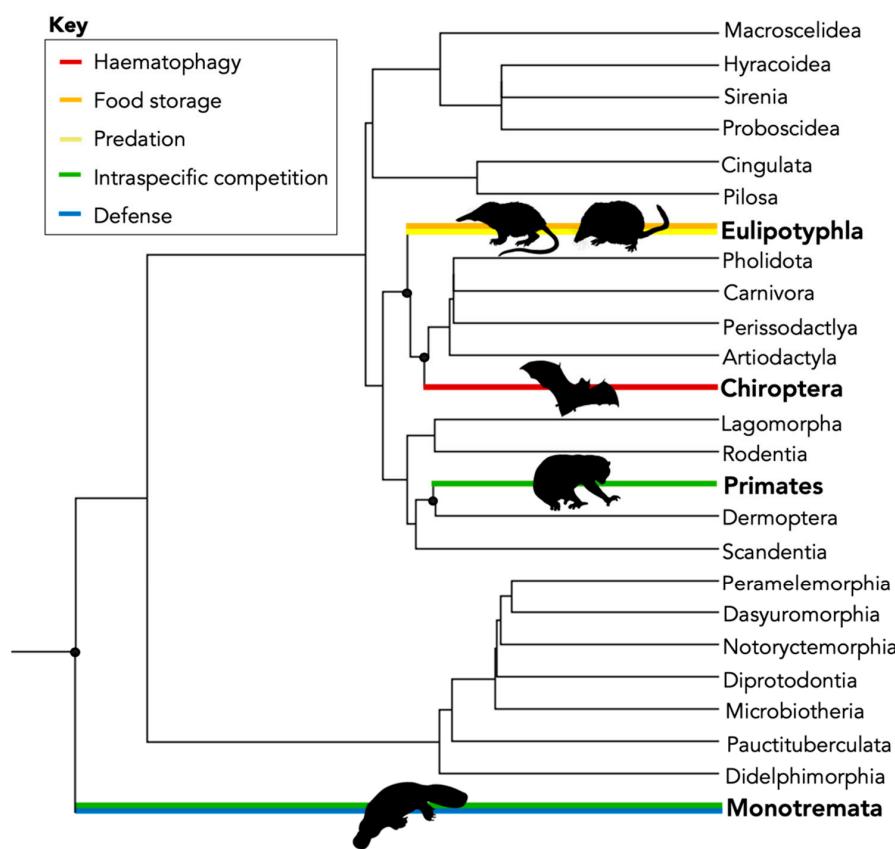
Figure 1. Current phylogeny of the class Mammalia, branches with orders of venomous mammals are coloured with corresponding functional usages of venom. Estimated divergent times for venomous mammals are approximately 178 MYA for Monotremes, 73 MYA for Eulipotyphla, 5 MYA for Chiroptera (Desmodontinae) and 25 MYA for Primates (Lorisidae). Divergent times and proportionally representative branches were obtained from TimeTree [12].

Table 1. The currently confirmed or highly likely venomous mammal species—based on either observation or relation to confirmed venomous species. We have chosen not to include several Eulipotyphla that have been suggested to be venomous as this has been highly speculative for many years with no studies quantifying potential venoms or observing behaviours that confirm likely venom use. For detailed information on specimens of Eulipotyphla that may potentially be venomous, please refer to Kowalski and Rychlik [13].

Common Name	Order	Family	Genus	Species	Venom Status	Reference
Platypus	Monotremata	Ornithorhynchidae	*Ornithorhynchus*	*anatinus*	Confirmed [1]	[14]
Short-beaked echidna	Monotremata	Tachyglossidae	*Tachyglossus*	*aculeatus*	Vestigial [2]	[15]
Cuban solenodon	Eulipotyphla	Solenodontidae	*Atopogale*	*cubana*	Highly likely (evolutionary)	[16]
Hispaniolan solenodon	Eulipotyphla	Solenodontidae	*Solenodon*	*paradoxus*	Confirmed	
Northern short-tailed shrew	Eulipotyphla	Soricidae	*Blarina*	*brevicauda*	Confirmed	
Southern short-tailed shrew	Eulipotyphla	Soricidae	*Blarina*	*carolinensis*	Highly likely (evolutionary)	
Elliot's short-tailed shrew	Eulipotyphla	Soricidae	*Blarina*	*hylophaga*	Highly likely (evolutionary)	[17]
Everglades short-tailed shrew	Eulipotyphla	Soricidae	*Blarina*	*peninsulae*	Highly likely (evolutionary)	

Table 1. *Cont.*

Common Name	Order	Family	Genus	Species	Venom Status	Reference
Asiatic short-tailed shrew	Eulipotyphla	Soricidae	*Blarinella*	*quadraticauda*	Confirmed	[18]
Indochinese short-tailed shrew	Eulipotyphla	Soricidae	*Blarinella*	*griselda*	Highly likely (evolutionary)	
Burmese short-tailed shrew	Eulipotyphla	Soricidae	*Blarinella*	*wardi*	Highly likely (evolutionary)	
Mediterranean water shrew	Eulipotyphla	Soricidae	*Neomys*	*anomalus*	Confirmed	[19]
Eurasian water shrew	Eulipotyphla	Soricidae	*Neomys*	*fodiens*	Confirmed	[20]
Transcaucasian water shrew	Eulipotyphla	Soricidae	*Neomys*	*teres*	Highly likely (evolutionary)	
Crawford's grey shrew	Eulipotyphla	Soricidae	*Notiosorex*	*crawfordi*	Likely (behavioural observation)	[21]
Canarian shrew	Eulipotyphla	Soricidae	*Crocidura*	*canariensis*	Likely (behavioural observation)	[22]
Common vampire bat	Chiroptera	Phyllostomidae	*Desmodus*	*rotundus*	Confirmed	[23]
Hairy-legged vampire bat	Chiroptera	Phyllostomidae	*Diphylla*	*ecaudata*	Confirmed	[24]
White-winged vampire bat	Chiroptera	Phyllostomidae	*Diaemus*	*youngi*	Confirmed	
Pygmy slow loris	Primates	Lorisidae	*Xanthonycticebus*	*pygamaeus*	Confirmed	[25]
Greater slow loris	Primates	Lorisidae	*Nycticebus*	*coucang*	Confirmed	[25]
Javan slow loris	Primates	Lorisidae	*Nycticebus*	*javanicus*	Confirmed	[26]
Kayan slow loris	Primates	Lorisidae	*Nycticebus*	*kayan*	Confirmed	[27]
Bengal slow loris	Primates	Lorisidae	*Nycticebus*	*bengalensis*	Confirmed	[28,29]
Bornean slow loris	Primates	Lorisidae	*Nycticebus*	*borneanus*	Highly likely (evolutionary)	[26,27,30]
Philippine slow loris	Primates	Lorisidae	*Nycticebus*	*menagensis*	Highly likely (evolutionary)	
Sumatran slow loris	Primates	Lorisidae	*Nycticebus*	*hilleri*	Highly likely (evolutionary)	
Bangka slow loris	Primates	Lorisidae	*Nycticebus*	*bancanus*	Highly likely (evolutionary)	

[1]. Male only, active during breeding season (approx. June–October) [2]. Vestigial gland present, adaptation now possibly for olfactory communication, supports theory that venom is an ancestral trait of Monotremata. See Wong, et al., [15] for further details.

Within Chiroptera, all molecules so far identified within the venom are assumed to help facilitate blood feeding: plasminogen activator prevents blood from clotting to allow a constant flow of blood when feeding [24]. Other toxins found within Chiroptera venom include CRiSP (which are proposed to either have a peripheral vasodilatory effect or be potentially neurotoxic), Kunitz domains (suggested to have anticoagulation properties), Kallikrein-1 (known as KLK1, a serine protease, suggested to aid in cleaving fibrogen which prevents blood clotting and vasodilation), Pituitary Adenylate Cyclase-Activating Polypeptide (known as PACAP, proposed to aid in vasodilation). In addition, *Desmodus* allergen-related lipocalins (called Desmallipins) have been identified within the venom although it is not clear what purpose they provide [23].

Within Eulipotyphla, the most prominent and well-documented venom is a Kallikrein-1 derivative (KLK1). Eulipotyphla paralyse their prey and store them in food larders. Due to the exceptionally high metabolism found within members of Eulipotyphla, they must eat every few hours [13]. KLK1 in shrews is thought to cause a drop in blood pressure and in turn result in paralysis within prey—both invertebrate and vertebrate [37].

The only protein identified within slow loris venom so far is a secretoglobin called Brachial Gland Secretion (BGE). This protein must be mixed with slow loris saliva to

become toxic, being the only known two-step venom observed in the animal kingdom so far [26]. The BGE is a secretoglobin that shares a homology with the Fel-d-1 protein produced by domestic cats (*Felis catus*) that causes allergies in humans—humans bitten by slow lorises often display anaphylactic shock. Slow lorises are highly territorial, with venom used in intraspecific competition between females or between males, resulting in necrotic wounds that can lead to intraspecific injury or death [32].

Despite this convergent origin of venom in mammals, previous work has identified shared toxins and molecules found within mammal venoms including kallikrein-1 (Eulipotyphla and Chiroptera) [16,17,23], secretoglobins (Primates, Eulipotyphla and Chiroptera) [17,23,25] and CRiSPs (Platypus and Chiroptera) [23,38]. Venom evolving convergently within and across the class, as well as having multiple different functional usages but still with shared toxins is distinct in mammals amongst venomous animals. Several mammal toxins are represented by multiple paralogs or isomers, which allows an opportunity to investigate if selection has acted equally across sequences within a toxin group or if there is evidence of episodic selection promoting variation within similar sequences.

Several researchers have investigated selection of mammalian toxins, but these papers have been published intermittently over two decades, using different methodology and different criteria to test hypotheses. The availability of mammalian toxin data has also grown, allowing an opportunity to not just test species evolution but the evolution of toxins across the whole Class, affording a direct comparison. This presents the opportunity to investigate if there are patterns of selection across the Class due different factors such as the functional use of venom, the evolutionary age of the group or the type of toxin. This direct comparison between different venomous mammal groups is a novel contribution to the literature.

By using multiple representatives of mammalian toxins, we can begin to answer three questions pertaining to mammal toxin evolution:

1. What is the dN/dS ratio of mammal toxins and where does the Class fit within the 'two-speed' model of evolution proposed by Sunagar and Moran (2015) [9]?
2. Are sites found under selection subject to pervasive selection (found across the whole phylogeny) or episodic selection (found within a subset of branches)? Is there evidence that different branches (toxin sequences) are under different rates of selection?
3. Are any sites identified to be under diversifying selection within a toxin having a possible structural or chemical impact on the toxin?

To answer these questions, we are using a combination of selection tests from the Datamonkey Adaptive Evolution Server FUBAR (Fast Unconstrained Bayesian Approximation, MEME (Mixed Effects Model of Evolution) and aBSREL (Adaptive Branch-Site Random Effects Likelihood), CodeML from PAML and TreeSAAP (Tree Selection on Amino Acid Properties). Using these in combination will enable us not only to describe the selection rates in mammals but begin to discuss the possible function of a toxin or selection pressures acting upon it due to the different layers of selection we are looking at. For question one, CodeML can provide the overall dN/dS ratio—comparing if the models are significant will identify the overall rate of selection acting upon that toxin.

For question two, using the DataMonkey Adaptive Evolutionary Server, each toxin group alignment will generate a phylogenetic tree to estimate the rate of dN/dS along a branch or at a specific site; FUBAR identifies pervasive site selection (found at a site across the whole phylogeny generated), MEME identifies episodic site selection (found at a site across a subsection of the phylogeny generated) and aBSREL identifies if sites along a branch within the phylogeny has undergone episodic selection. These methods in combination can help identify how much variation there is within sequences and where in these toxins groups there exists variation (especially if highlighted by episodic method like MEME and aBSREL), which may allow toxins to work on different targets. For example, the venom of shrews is predominately comprised of KLK1 but shrews consume both invertebrate and vertebrate prey [17].

Finally, to investigate question three, we will use TreeSAAP that calculates the likelihood of diversifying selection at a site and verifies if that selection could impact an amino acid property (such as hydropathy, the measure of how hydrophobic or hydrophilic a molecule is). Results from TreeSAAP can implicate if selection pressures are having an impact on the toxicity of the molecule and in turn, contribute significantly to the fitness of the mammal.

Although sequence data have improved for mammal toxins, there is still limited availability of the groups we can investigate as well as uncertain identification of non-toxin precursors. Due to this, the toxin groups we are using do not have non-toxins sequences for comparison. Therefore, we cannot state *categorically* if selection identified within this study has been selected for toxicity, but we do know that sequences used in this research have already identified to be recruited as toxins and therefore the possibility of selected sites being implicated in toxicity is still plausible. The methods chosen not only allow a complete comparison of all toxin groups identified within venomous mammals, but enable mammals to be compared to other venomous taxa: three of the above programmes (FUBAR, MEME and CodeML) are used in Sungar and Moran's paper [9], three are used by Barua, et al., [39] (aBSREL, MEME and CodeML) and four by Low et al., [23] (FUBAR, MEME, CodeML and TreeSAAP). For further justification and explanation of each method used, see Section 5 (Materials and Methods).

Based on previous research, we hypothesize that the overall rate of evolution within mammals will be acting under diversifying selection [16,17,23], and be linked to the evolutionary age of venomous mammals and the different functional usages of venom. We expect to see a proportional amount of diversifying selected sites found across all toxins, although we expect to see more episodic diversifying selection than pervasive selection in KLK1 due to the diverse diet that Eulipotyphla consume requiring more targets to be hit when compared to other toxins which are more focused on smaller targets (i.e., intraspecific in platypuses or blood feeding in vampire bats). Finally, we expect to see some selected sites found on active sites or other structurally important areas within the toxins' protein, particularly in 'younger' groups of mammals (i.e., Desmodontinae) when compared to 'older' groups of mammals (i.e., Eulipotyphla).

2. Results

For a toxin group to be analysed, at least four sequences representing that toxin had to be identified to minimise errors and skew. We identified eight venomous mammal species with a total of 139 toxin (or non-toxin but found within the venom) sequences that have been described. Due to the criteria, two species that had sequence data (*Blarinella quadraticauda* and *Nycticebus javanicus*) were excluded as each had less than four sequences. Of the 139 sequences identified, only 82 representing seven different toxin groups were able to be used: Plasminogen Activator (Chiroptera), KLK1 (*Blarina brevicauda*, *Solenodon paradoxus*, and *Desmodus rotundus*)), CRiSP (Cysteine-rich secretory proteins, *Desmodus* only and *Desmodus* and *Ornithorhynchus anatinus*), PACAP (Pituitary Adenylate Cyclase-Activating Polypeptide Pituitary Adenylate Cyclase-Activating Polypeptide, *Desmodus*), Kunitz Domain One (*Desmodus*), Kunitz Domain Two (*Desmodus*) and Desmallipins (*Desmodus* allergen-related lipocalins, *Desmodus*.) We also chose to conduct species or order specific alignments (in the case of CRiSP and KLK1), which brought the total number of toxin alignments up to eleven (Table S1). The majority of alignments (seven out of eleven) are only represented by single species, evolutionary studies on mammals [23] have shown that the methodology still provides informative answers although caution should be taken regarding inferring toxicity and false positives. No sequences could be obtained from slow loris species. This is due to both a lack of sequenced venom and genomes assembled to chromosome level, either of which would have enabled sequences to be included.

2.1. HyPhy DataMonkey

2.1.1. FUBAR and MEME Results

Overall trends reported by FUBAR and MEME were consistent across both analyses on the number of sites evolving under purifying or diversifying selection for each alignment, with equivalent numbers of MEME diversifying sites found to as many FUBAR diversifying sites. (Table 2). All but a few sites identified by FUBAR and MEME are different from each other (as they are testing different aspects) although there are some sites identified on alignments that are shared (see Figure 2, point 179). Six alignments have less than or exactly ten sites per method identified to be evolving under either diversifying or purifying selection: PACAP, Desmodus CRiSP, and Kunitz Domain Two. Two alignments have more than ten sites estimated to be evolving under diversifying selection by one or both methods: Plasminogen Activator (FUBAR identified 15 sites evolving under diversifying selection) and Desmallipins (FUBAR identified 22 sites and MEME identified 15 sites evolving under diversifying selection). The four KLK1 alignments have a significant number of sites identified to be acting under purifying selection by FUBAR (Blarina KLK1 *n* = 23, Blarina KLK1 *n* = 16, Eulipotyphla KLK1 *n* = 35 and All KLK1 = 47). However, MEME identifies a large amount of diversifying episodic selected sites in Eulipotyphla KLK1 (*n* = 16) and All KLK1 (*n* = 21) despite also reporting that both alignments are evolving under purifying selection (Eulipotyphla KLK1 ω = 0.593 and All KLK1 ω = 0.513) (Figure 2). This result is likely due to the convergent origin of KLK1 within Eulipotyphla and Chiroptera. As well, in both CRiSP alignments, MEME did not identify any episodic diversifying selection.
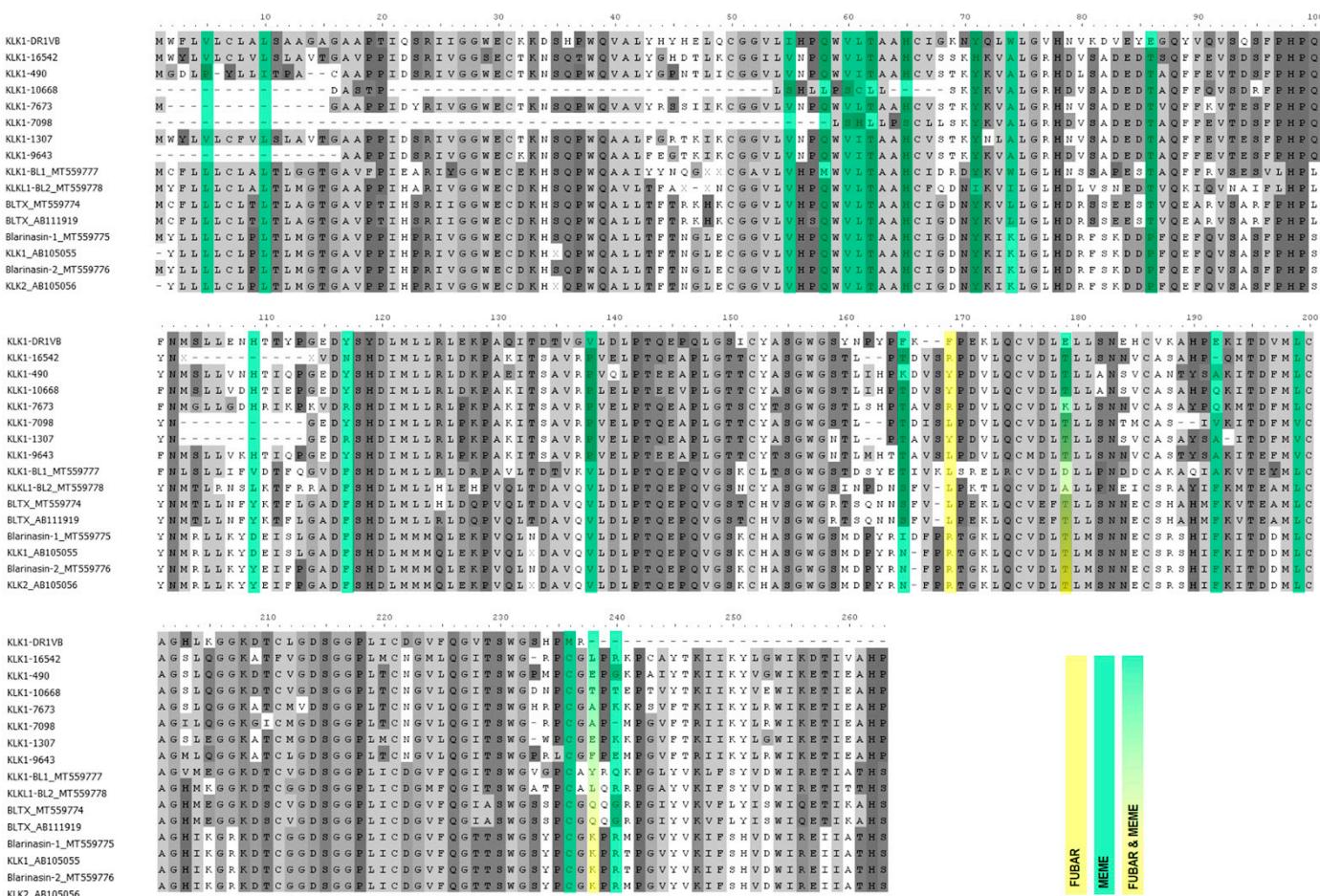

Figure 2. Amino acid alignment of KLK1 All, with diversifying selection highlighted according to each programme. Yellow represents FUBAR (pervasive selection) and Green represents MEME (episodic selection). Overlapping sites with partial or complete programme agreement are coloured as mixed hues. Please see Figures S1–S10 for all other alignments.

Table 2. Summary of results for eleven alignments representing the seven toxin groups (bolded). Results include sites identified by FUBAR evolving under diversifying ($\omega > 1$) or purifying ($\omega < 1$) selection, sites identified to be evolving under diversifying selection by MEME, total dN/dS for Model 0, Model 2a, Model 8 and Model 8a and significant diversifying sites (<0.95–0.99) for Model M2a and M8. Sequence alignments reflecting these results are shown as Figures S1–S10.

Species	Toxin	FUBAR		MEME	CODEML						
		$\omega > 1$	$\omega < 1$	$\omega > 1$	Model 0 Weighted Average	Model 2a Weighted Average	M2a Positive Sites >0.95	M2a Positive Sites >0.99	Model 8 Weighted Average	M8 Positive Sites >0.95	M8 Positive Sites $p > 0.99$
Blarina brevicauda	Kallikren-1 (KLK1)	5	23	4	0.57984	0.8139	N/A	N/A	0.8075	13	3
Solenodon paradoxus	Kallikren-1 (KLK1)	8	16	6	0.58254	0.7298	1	2	0.7073	2	4
Eupotiphyla (*Blarina* and *Solenodon*)	Kallikren-1 (KLK1)	5	35	16	0.49379	0.6748	1	2	0.6260	3	2
Desmodus rotundus, *Blarina* **and** *Solenodon*	**Kallikren-1 (KLK1)**	**3**	**47**	**21**	**0.43905**	**0.6339**	**1**	**N/A**	**0.5592**	**3**	**1**
Diphylla ecaudata, *Desmodus* **and** *Diameus youngi*	**Plasminogen Activator**	**15**	**8**	**4**	**1.0549**	**1.3046**	**10**	**4**	**1.3082**	**14**	**10**
Desmodus	**Desmallipin**	**22**	**8**	**15**	**1.09464**	**1.3216**	**3**	**7**	**1.3393**	**17**	**12**
Desmodus	**PACAP**	**3**	**5**	**2**	**0.90099**	**1.3239**	**4**	**1**	**1.3208**	**14**	**N/A**
Desmodus	CRiSP	5	2	0	1.48836	2.0592	4	N/A	2.0806	4	2
Desmodus **and** *Ornithorhynchus anatinus*	**CRiSP**	**7**	**2**	**0**	**0.34629**	**1.3843**	**3**	**N/A**	**1.2923**	**4**	**3**
Desmodus	**Kunitz Domain One**	**1**	**7**	**1**	**0.29323**	**1.4170** [NS]	**1**	**N/A**	**1.3985** [NS]	**2**	**2**
Desmodus	**Kunitz Domain Two**	**2**	**1**	**2**	**2.28040** [NS]	**2.6283**	**3**	**1**	**2.6283**	**4**	**3**

2.1.2. aBSREL Results

Diversifying episodic selection was identified on nine of the alignments. Four alignments had one branch identified as evolving under diversifying selection (Blarina KLK1, Desmallipins and Kunitz Domain One). Five alignments had multiple branches identified to evolve under diversifying selection including Plasminogen Activator (three branches), Solenodon KLK1 (five branches), All KLK1 (five branches), PACAP (five branches) and Eulipotyphla KLK1 (six branches). No evidence of episodic selection was found by aBSREL in three of the alignments: Desmodus CRiSP, CRiSP All and Kunitz Domain Two (Spreadsheet S4).

2.2. PAML Selection Results

2.2.1. Diversifying Selection Detected by Alternative Model

Reporting the result taken from the M8 model, most mammal toxins are evolving under diversifying selection, with high levels of significance ($p < 0.001$) between all comparisons of null and alternative models (M0 vs. M3, M1a vs. M2a, M7 vs. M8 and M8 vs. M8a) confirming that the alternative model (those that allow for $\omega > 1$) should be accepted. We accept significance as $p < 0.01$ for PAML results to reduce the likelihood of a Type 1 error. Toxins confirmed to have high levels of diversifying selection included Plasminogen Activator (Model 8 $\omega = 1.3082$) (Tables S12 and S13), Desmallipins (Model 8 $\omega = 1.3393$) (Tables S14 and S15), PACAP (Model 8 $\omega = 1.3208$) (Tables S16 and S17) and CRiSP Desmodus only (Model 8 $\omega = 2.0806$) (Tables S18 and S19). CRiSP All received a lower but still significant ($p < 0.01$) level of support for M1a vs. M2a and M7 vs. M8 but had significance $p < 0.001$ for an acceptance of the alternative model for M8a vs. M8, indicating diversifying selection (Model 8 $\omega = 1.2923$) (Tables S20 and S21).

2.2.2. Purifying Selection Detected by Alternative Model

All four of the KLK1 alignments are evolving under purifying selection—even with the acceptance of the alternative model with high level of significance ($p < 0.001$) across all four sets of models: KLK1 Blarina (Model 8 $\omega = 0.8075$), KLK1 Solenodon (Model 8 $\omega = 0.7073$) and Kallikren-1 Eulipotyphla (Model 8 $\omega = 0.6260$). Kallikren-1 All also is highly significant ($p < 0.001$) for three of the four sets of models, except for M1a vs. M2a, which has a high level of significance ($p < 0.01$), still indicating purifying selection (Model 8 $\omega = 0.5592$) (Tables S4–S11).

2.2.3. Alternative Models Rejected

The two Kunitz Domain groups provided mixed results of significance—Kunitz Domain Two is highly significant for the acceptance of alternative model M8a vs. M8 (Model 8 $\omega = 2.6283$); with a significant ($p < 0.01$), result for the acceptance of M1a vs. M2a and M7 vs. M8. Although the M0 vs. M3 acceptance rate rejects the alternative model ($p < 0.05$), the null model still proposes that the group is evolving under a diversifying selection rate (Model 0 $\omega = 2.8040$) (Tables S24 and S25). Kunitz Domain One has a very high level of acceptance for the M0 vs. M3 alternative model ($\omega M3 = 1.4033$) but poor levels of acceptance ($p < 0.05$) for the other model matches, therefore, rejecting the alternative model (Tables S22 and S23).

2.3. Diversifying Sites Detected by Model 8 BEB

The Bayes Empirical Bayes was implemented in Model 8, which detected the sites in each alignment that are evolving under a very high (0.95–0.99) level of diversifying selection. A total of seven alignments had below 10% of total sites (taken from trimmed result in CodeML) evolving under detected diversifying selection: Solenodon KLK1 (3%, 6/187 total sites), Eulipotyphla KLK1 (3%, 5/179 total sites), All KLK1 (3%, 4/151 total sites), PACAP (3%, 6/182 total sites), All CRiSP (6%, 7/122 total sites), Blarina KLK1 (6%, 16/269 total sites) and Plasminogen Activator (8%, 24/303 total sites). Three of the alignments had above 10% of all sites evolving under diversifying selection: CRiSP Desmodus (11%, 14/124

total sites), Kunitz Domain Two (11%, 7/61 total sites) and Desmallipins (17%, 29/167 total sites). Due to the rejection of the alternative Model 8, the number of positive sites reported for Kunitz Domain One (9%, 4/47 total sites) are not included.

2.4. TreeSAAP Results

2.4.1. Properties Detected in Alignments by TreeSAAP

All but one alignment (PACAP) was identified to have sites evolving under diversifying selection that impacted different amino acid properties. The properties identified to have been affected across the alignments are long range non-bonded energy (detected in one alignment, CRiSP Desmodus), normalised partial specific volume (detected in one alignment, Kunitz Domain Two), consensus hydrophobicity (detected in one alignments, Kunitz Domain Two), power to be at the middle of the alpha-helix (detected in two alignments, Solenodon KLK1 and KLK1 All), polar requirement (detected in two alignments, CRiSP Desmodus and All CRiSP), hydropathy (detected in two alignments, KLK1 Eulipotyphla and KLK1 ALL), power to be at the C-terminal (detected in four alignments, KLK1 Shrew, KLK1 Eulipotyphla, KLK1 All and CRiSP All) and isoelectric point (detected in five alignments, Plasminogen Activator, KLK1 Shrew, Desmallipins, CRiSP All, Kunitz Domain One) (Table S26).

2.4.2. Alignments Which Had Significant Impact on Properties Identified by TreeSAAP

Only eight of the eleven alignments had specific sites identified by TreeSAAP: KLK1 Shrew, KLK1 Solenodon, KLK1 Eulipotyphla, KLK1 All, Desmallipins, CRiSP Desmodus, CRiSP All, and Kunitz Domain 2. Further, the four KLK1 alignments were identified to have more than 1 site present for each property identified. Some sites were also recorded to be evolving at diversifying selected rates that affected amino acid properties, however they were recorded at a different magnitude than observed.

2.4.3. Location of Selection on the Protein Structure

When mapped to the protein structures, the positions under diversifying selection identified by multiple methods were located on the proteins' surfaces. Based on functional annotation, no active site residue (or region) is affected. In the plasminogen activator, one glycosylation position is under diversifying selection, as is an initial position in the signal peptide region of PACAP (Figure 3).
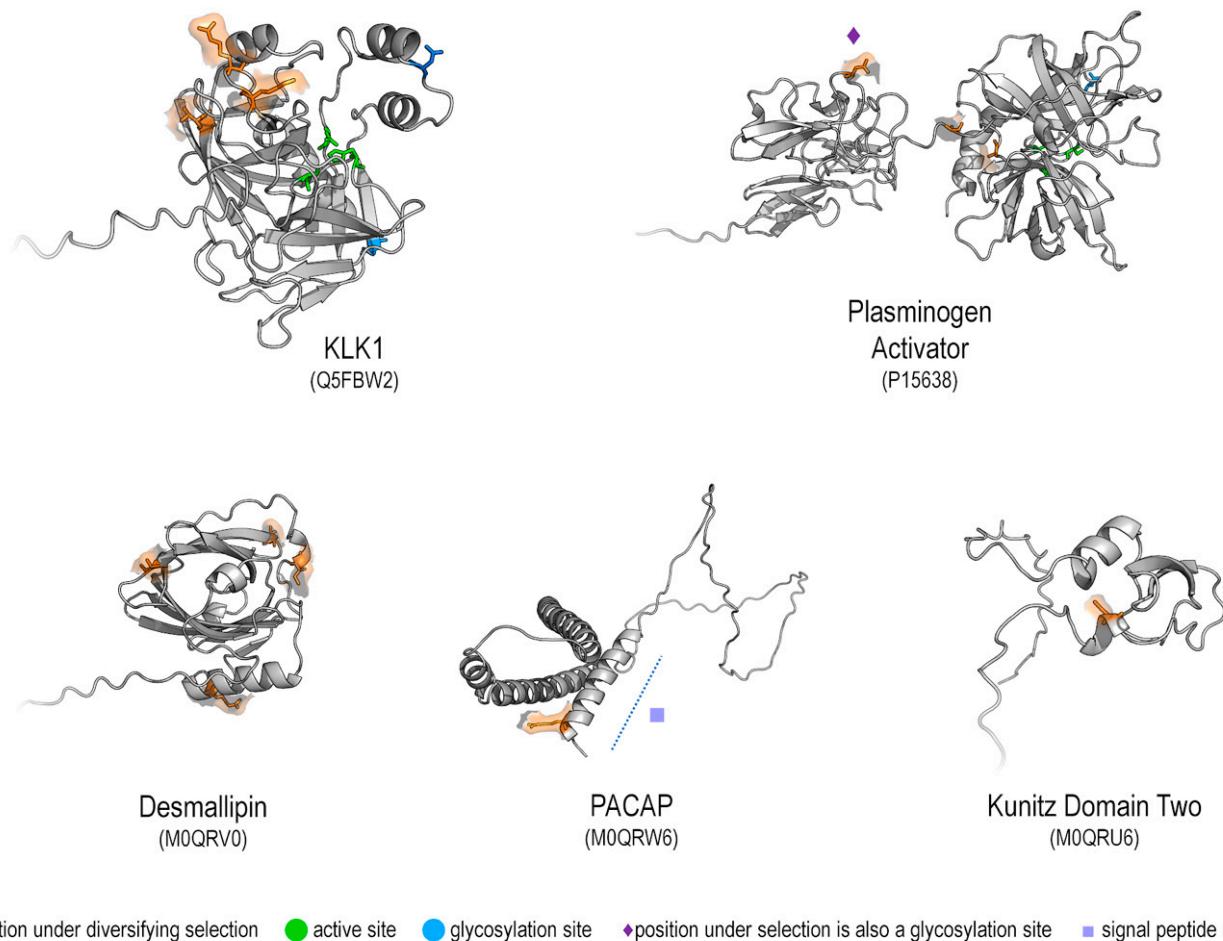
Figure 3. Protein structures highlighting positions under diversifying selection detected in this paper. Overlapping sites with partial (yellow) or complete (orange) FUBAR and MEME agreement are shown as sticks with molecular surface contour on overall cartoon representation. Structure IDs correspond to AlphaFold Protein Structure Database.

3. Discussion

Using different branch and site-specific methodologies, as well as TreeSAAP, we analysed the rate of selection acting on toxins found across venomous mammals for the first time. Overall, mammal toxins are evolving under diversifying selection (Figure 4). Six are evolving under diversifying selection (Plasminogen Activator, CRiSP, Desmallipins, PACAP, Kunitz Domain Two and Kunitz Domain One (although Kunitz Domain One is not a significant result). In contrast, the KLK1 toxin group (including all species-specific alignments) is under purifying selection. These results indicate that mammalian toxins are overall evolving under diversifying selection, following the two-speed model. There are 89 diversifying sites (>0.95) identified by PAML Model 8 across the toxin groups, with the majority from Plasminogen Activator ($n = 24$) and Desmallipins ($n = 29$). The total dN/dS ratio of mammal toxins (12.3% dN, 85.7% dS) and large number of diversifying sites identified would group mammal toxins with the other high synonymous groups found by Sungar and Moran [9], the advanced snakes (diversifying sites = 531, diverged ~54 MYA) and cone snails (diversifying sites = 166, diverged ~33–50 MYA).

Figure 4. A visual summary of the main findings of our research showing that mammal toxins are evolving overall under diversifying selection.

Throughout all the toxins, there is evidence of similar amounts of pervasive and episodic selection found in each alignment. While this is observed in a number of other venomous orders [9], the very high amount of episodic selection identified in KLK1 Eulipotyphla and KLK1 All (by both MEME and aBSREL) and the lack of episodic selection identified in both CRiSP alignments are both. This episodic selection identified in KLK1 could be to do with the diverse diet that venomous Eulipotyphla species are observed to eat [16,17,20]—with multiple KLK1s able to target both physiologically different invertebrate and vertebrate systems at the same time [16,17,20]. When observed, *Neomys* spp., show a preference for taking large prey items (such as frogs or rodents) over smaller prey items [19] and venom of *B. brevicauda* has been shown to be potent on vertebrates as large as domestic cats (*Felis catus*) [13] despite the shrew's size—it is possible that this episodic selection has slowly developed as shrews expanded their niche from to include vertebrates with selection pressure and competition from other invertebrate-eating organisms pushing them to target larger, risky but more rewarding prey.

Structural models of mammalian toxins revealed that positions under diversifying selection are located on protein surfaces (Figure 3). This is in accordance with the observation that changes in exposed residues are less detrimental to protein stability than those occurring in buried residues [40–42]. These changes in surface-exposed residues are considered facilitators for neofunctionalisation of toxins by modification of interactions between protein and its target(s) [2,43]. For KLK1, the active and glycosylation sites are not under selection. Positions under selection are located in regions close to regulatory loops identified as critical for toxicity in KLKs from reptiles and mammals [44]. For plasminogen activator, the active site is not under selection, but one glycosylation site is. Glycosylation is relevant for protein secretion and localisation [45], something that has been shown to be critical for some vampire bat toxins [46]. In PACAP, the identified residue under selection is

located in the N-terminus of its signal peptide segment [47]. Signal peptides are ubiquitous in animal venom toxins [43], being essential for protein secretion [47]. For Desmallipins and Kunitz Domain Two, there are no pinpointed functional positions, with all sites under selection being found in the protein surface. The selection pressure on positions related to protein secretion and localisation may point to a progressive transition of these proteins to be present in venom glands (or in alternative glands/sites). Likewise, the pressure on surface coils on KLK1 may indicate further target alteration (either broadening prey classes or becoming more active towards specific groups). We hypothesised that we see more significant sites under selection on evolutionary younger groups; while the results do concur with that, this is more likely because sequence data are overwhelming from *Desmodus rotundus*.

Research has already indicated that toxins within mammal venom are experiencing mixed selection rates [9,16,17,23,24]. In Chiroptera, the majority of toxins are evolving under weak to strong diversifying selection [23] with an unusually high proportion of sites evolving under diversifying selection (from 20 to 36% of total sites). Within Eulipotyphla, the *Solenodon* KLK1 paralogs are evolving under purifying selection compared to other mammal non-toxic KLK1 despite a number of diversifying sites identified within the alignments [16] but in *Blarina*, KLK1 toxins are identified to be evolving under diversifying selection compared to non-toxin KLK1 paralogs [17]. Although these results are highly credible and provide support for the two-speed model, there has been little consistency with the methodology used due to the advancements made in selection tests since 2001 and the varied areas of research by each study. This mixed pattern was also detected by Harris and Arbuckle [48] when examining loss/gain of venom (and poison) within tetrapod orders, with mammals displaying a higher rate of losing venom as a trait despite only venom gains being recorded by ancestral shifts. The authors explained that this could be a result of venom in mammals evolving as an ancestral trait but being lost with no trace. Alternatively, due to the unusual spread of the trait and the multiple specific usages of venom recorded in the mammals, venom as a trait may not have evolved at the same time as the divergence of venomous mammals. Therefore, toxins may still be undergoing high levels of diversifying selection due to the more recent recruitment of venom within venomous mammals. Our results do corroborate previous work, with the general trend of whether a toxin is acting under diversifying or purifying selection. There are differences between the proportion of sites under significant selection and the site numbers identified. This could likely be due to differences in alignment, using a different methodology or the inclusion of different sequences compared to past research on mammal toxins.

3.1. Mammal Toxins Are Acting under Diversifying Selection

We confirm previous suggestions [16,17,23] that the majority of mammal toxins are evolving under diversifying selection. Taking into account the evolutionary ages of each group of mammal analysed (~178 MYA for Monotremes, ~73 MYA for Eulipotyphla and ~5 MYA for Chiroptera [12,16,34,49]), mammal toxins fit the two-speed model as a 'young' evolutionary group.

The specific molecular mechanism responsible for toxins may contribute to the diversifying rate of selection observed in this study; while duplication were thought to be most responsible for toxins in venom [7,50], there are other mechanisms that also identified in the neofunctionalisation of toxins [51]. For example, alternative splicing (where exons from the same gene are spliced in different mRNA strings to produce differing proteins) is responsible for the different plasminogen activators identified in all three species of vampire bat, a mechanism that can accelerate selection rates [52].

Our findings support previous research that suggested KLK1 occurs under purifying selection [16]. As this toxin predominantly has a predatory function (See Figure 1), more research is required to investigate why this does not have higher selection pressure as would be expected in the predator–prey arms race [17,18,53]. Our results for venom of *Blarina* indicated that other molecules, such as HYLAP (Hyaluronidase acid, PH-20), may

play a role as they facilitate the 'spread' of venom by degradation of target cells. These molecules are under diversifying selection [17] and may act to maintain KLK1 under purifying selection by minimising mutations. Given, HYLAP was not identified in all KLK1 venoms (such as Solenodon) further research is required to identify equivalent molecules in those venoms.

Another intriguing aspect of the purifying selection KLK1 is under is the number of diversifying sites identified by PAML Model 8 in *B. brevicauda*—a total of 16 sites (Table S4). Although this may appear to be a false positive (indeed, FUBAR identify significantly more pervasive purifying than diversifying sites across the *B. brevicauda* KLK1 alignment, Tables S1 and S2), in [44], *B. brevicauda* toxic KLK1 was identified to be under purifying selection but inserted loops were evolving under episodic diversifying selection, raising the suggestion that these sites are responsible for the toxic effect by these molecules. Indeed, all KLK1 alignments had diversifying sites that TreeSAAP identified as having affected a structural protein property (Table S26). Research on KLK1 proposed it as a toxipotent molecule [39]—that even prior to neofunctionalisation of KLK1 in venom, it possessed properties that made it suitable for venom usage. Due to the parallel origins of KLK1 toxins identified in mammal and reptile venom, sharing structural and functional similarities despite the divergence of the groups 300 MYA, KLK1 was suggested to be more susceptible to becoming a toxin than other related groups [54]. This would explain the intense pervasive purifying selection acting on KLK1, with its ancient origins but with specific episodically selected sites that are toxipotent. Although *S. paradoxus* and the grouped KLK1 that included all identified KLK1 in the mammals had significantly less positively sites (only six and four respectively), this could be a result of the convergence origin of KLK1.

3.2. Different Functions Have Different Rates of Selection on Mammal Toxins

In the Class Mammalia, venom does appear to follow the two-speed model but when examining the separate mammalian orders in this study (Monotremata, Eulipotyphla and Chiroptera), we found no clear pattern, even when considering site or branch selection rates. Beyond the evolutionary divergence times of each order, it is the functional usages (even when broadly grouped under predatory, defensive, or intraspecific), ecology and behaviour of the animal that appear to have a stronger influence on the rate of selection of a particular toxin.

Only three of the toxin groups identified (KLK1, Desmallipins and CRiSP) that have diversifying sites identified by TreeSAAP as significantly affecting the toxins' protein properties and therefore have a significant impact on the fitness of the animal. These results differ from those reported by Low et al., [23] who reported that Kunitz Domain One, Kunitz Domain Two and Plasminogen Activator had significant sites reported by TreeSAAP. This is likely due to the differences in methodology, as we used nucleotide sequences whereas they used amino acid sequences.

In other toxin groups (except for PACAP), diversifying sites that had influenced nucleotide sequences were identified although they were not significant. A possible explanation for why only these three groups had significant sites in TreeSAAP could be due to pluripotency of toxins—Desmallipins are allergen-related lipocalin proteins, a candidate for olfactory communication in *D. rotundus* [55], a social species that partakes frequently in grooming. This sociality could also explain the high number of diversifying sites identified with both FUBAR and MEME—in particularly the episodic selection indicating different levels of diversifying selection acting on 22 sites within Desmallipins. The data used within this study come from two separate studies on *D. rotundus* [23,56]—if sociality is a selection pressure on Desmallipins, this could be why there are so many sites, as two different colonies of bats will have distinct communication between each other. This is similar to slow lorises, which use brachial gland exudate (BGE), a toxin with high structural and functional homology with Fel-d-1 but which is still involved in olfactory communication [25]. For the CRiSPs, both in *D. rotundus* and *Ornithorhynchus anatinus*, they might preform a role in immunity—although vasodilation (as proposed in [23]) may be a stronger factor

for *D. rotundus*. Further expanding on the CRiSP pleiotropy, the PAML Model 8 result for CRiSP All (M8 = 1.2923) while still weakly diversifying, is significantly less than just CRiSP *Desmodus* (M8 = 2.0806). Running the different CRiSP sequences through GenBank BLAST [57] shows that the CRiSPs in *O. anatinus* and *D. rotundus* evolved from different paralogs, with *O. anatinus* evolving from CRiSP 2 and *D. rotundus* evolving from CRiSP 3. The non-toxic mammalian CRiSPs are expanding under purifying selection (although both [58,59] identify that CRiSP 2 is not-significant) and while both are associated with immunity, CRiSP 2 is recognised for its role in sperm gamete fusion. Further evidence to support this is the fact that no episodic selection was identified in either CRiSP alignment, this could imply that the function for each is still somewhat retained as there is little to no non-synonymous mutation between each sequence identified. The *O. anatinus* CRiSP is also expressed at the same level within the venom gland both during and out of breeding season, further providing evidence of a possible pleiotropic nature [38,60].

The higher rate of selection on Chiroptera toxins can also be interpreted as genomic evolution. The jump from the ancestral diet of insectivory to blood feeding in the vampire bats occurred approx. 5 MYA [61]. Haematophagy requires numerous adaptions due to the potential dangers of viral infection, iron poisoning and poor quality of nutrients [52,62]. In *D. rotundus*, adaptions to aid in this diet include a heavily adapted microbiome that has evolved parallel to the bats genome and a unique immune system amongst mammals, with loss of immune genes that result in an inflammatory response to viruses (include defensins and natural killer) but had an expansion in major histone complex genes that result in less cell death [63]. The genome of *D. rotundus* also showed upwards of 1.6–2.6 times higher number of transposon mutagenic changes compared to other Chiroptera, with these transposons found in regions related to immunity amongst other adaptions related to haematophagy [62]. This could have resulted in rapid selection rate in vampire bats on homeostasis genes and immune genes, as well as genes associated with secretion [64], in comparison to other Chiroptera species. Vampire bats regurgitate blood to other bats within a colony if they were unsuccessful in feeding, with gut microbiomes closer in similarity between bats that had engaged in food sharing compared to kin or social groups [65]. Whilst many of the toxins found in vampire bat venom are posited to aid in blood feeding at the host (via anticoagulation or vasodilation [56]) and could explain the higher rate of evolution due to predator–prey interactions, the constant threat of viral infection could also justify why some of the toxins in vampire bats (i.e., CRiSP or Desmallipins) are under intense diversifying selection pressures.

Whilst PACAP is evolving under moderate diversifying selection (M8 = 1.3208) and has 14 diversifying sites identified, it was the only toxin that had no properties identified by TreeSAAP. Previous work suggested that this indicates the rate of selection and positive sites therefore had minimal impact on the fitness of *D. rotundus*. The discovery of vCGRP (vampire bat-derived form of calcitonin gene-related peptide) [33] which causes intense vasodilatory properties indicates that multiple toxins found within venom of *D. rotundus* are responsible for similar functions. Like the mixed result from the two Kunitz Domains (with a strong diversifying selection acting on Kunitz Domain One and weaker diversifying but not significant selection acting on Kunitz Domain Two), which have a proposed anticoagulation effect similar to plasminogen activator, having multiple toxins with similar functions can lead to relaxed selection.

A final point worth considering is the co-evolution of toxins with their target and the toxin families that represent several groups within this study. As observed by Zhu, et al., [66], some homologous scorpion alpha-toxins in *Mesobuthus* spp. contain insertions that reduce the efficiency of an alpha-toxin on mammalian voltage gated Na+ channels. These insertions did increase the insecticidal effects of the alpha-toxins, an adaption that allows the scorpions to take a greater variety of prey more efficiently and that prey selection had been a selection pressure on the alpha-toxins [67]. Toxins groups used in this study came from species-specific toxin families (i.e., KLK1 Solenodon), with aBS (Spreadsheet S4) indicating that some branches have had diversifying selection acting on them at some point

(i.e., KLK1 9643 and KLK1 7098)—these different rates of selection and multigene copies of the same toxin may allow venomous mammals to take different prey (such as the case of the solenodon that consumes both vertebrate and invertebrate prey) or keep the ancestral function of a toxin (as discussed above with CRiSP and Desmallipins). This would require further testing on a molecular, protein or genomic level of venomous mammals and their targets to verify this.

3.3. Future Research

The results are overwhelmingly represented by one species (*Desmodus rotundus*), which accounts for six of the seven toxin groups examined here. Additionally, the toxins analysed in this study have a small sample size (ranging from four to 16); smaller sample sizes are more likely to be skewed by sequences that are acting under very strong selection rates. To mitigate this, steps were taken to maximise the stability of the results with high levels of bootstrapping or taking only very significant results for CodeML (defined as $p < 0.01$). We cannot exclude the possibility of skewed or insufficient data within this paper. Therefore, caution should be taken when interpreting these results—although numerous results concurring with previous work do aid in providing support to our results.

The lack of primate toxin sequences also limits this research. Currently, only two amino acid and one nucleotide BGE sequence have been identified; all three sequences were identified by mass-spectrometry [26]. The slow lorises are the only example of a two-step venom system (with the BGE gland needing to be mixed with saliva via licking to induce potency) [11]. Identifying any additional molecules found in the gland for comparison and collecting a full sequence of BGE from multiple slow loris species or specimens would significantly increase our understanding of mammalian venoms. As well, intraspecific behaviour between slow lorises is territorial, between male-male and female-female [30,32]. Venom composition, toxicity and amount differs between the sexes in some venomous species [1]. Identifying if the rate of selection acts similar on venom within each sex in slow lorises would be an excellent contribution to our understanding of venom systems.

Identifying non-toxin precursors would allow researchers an opportunity to infer which sites or branches have been selected for in the transition to toxicity. Using the gene-wide methods such as BUSTED (Branch-Site Unrestricted Statistical Test for Episodic Diversification) to seek out where episodic positive selection has occurred across a gene (or a subset of branches) or RELAX which identifies if rates of purifying selection differ across a tree (this method requires a test set of branches, which would be non-toxin homologs). Identifying which specific non-toxic precursor (as well as any transitional non-toxin, such as venom like beta-defensin (DEFB-VL) found in the platypus [14]) has led to a mammalian toxin would provide a fascinating opportunity to identify what has driven parallel recruitment of similar toxins as mentioned in the Discussion about CRiSP recruitment between Chiroptera and Monotremes. Further expanding upon this point would provide excellent evolutionary insight.

There are additional shared toxins in mammal venoms that are widely shared by other venomous animals including phospholipase A2, secretoglobin and acid hyaluronidase. If further sequences are discovered in species not yet sequenced (i.e., *Crocidura canariensis*) [22] or species are revised for better quality sequences, they would provide an excellent comparison within the class and also across different animal venoms.

4. Conclusions

For the first time, the selection rates of available mammal toxins have been analysed together. We have shed light on the evolution of venom within mammals, reviewing cutting edge ecological, behavioural, and genomic research, as well as previously discovered toxins, to compare mammals to more well studied taxa. Overall, mammal toxins are acting under a diversifying selection of evolution and have a significantly large number of sites across eight toxin groups that are particularly evolving under both pervasive and episodic diversifying

selection which does follow the two-speed evolutionary model. These results can be indicative of different selection pressures acting on venomous mammals including sexual selection, immune responses, pluripotent proteins balancing toxic and non-toxic roles and predisposition of certain proteins for toxicity. Several toxins identified (KLK1, Desmallipins and CRiSP) have positively selected sites identified by TreeSAAP and visualised as protein models implying that these are significant to the fitness of the animal. These sample sizes contain few sequences: obtaining more samples of described toxins from numerous mammals (such as the platypus specific OvDLPs) and obtaining high quality sequences from missing toxins (e.g., BGE from *Nycticebus* or toxins from *Crocidura canariensis*) would provide more conclusive results and better support for understanding which factors are greater selection pressures on toxins in different mammal groups. It is only the beginning of furthering our understanding of this fascinating venomous group and the applications for its venom.

5. Materials and Methods

5.1. Data Collection

Data collection methodology and criteria is identical to the methodology as used by Sunagar and Moran, 2015 [9]. There is speculation on which species of mammal are venomous; we decided to be conservative in our definition and only included species that had published, accessible complete nucleotide sequences available for use. Thus, data from *Sorex araneus* and *Neomys foidens* were not included as venom has only been identified as fractions by Mass Spectrometry. Very few molecules from mammal toxins have been isolated and their toxicity verified—we assume unless specified that molecules described in text by researchers as toxic based on homology or experiments are toxins (including Plasminogen Activator [23,24] and Kunitz [36]). Identifying selection on mammal toxins within an alignment requires that there are at least four sequences represented by that toxin available—these criteria excluded the majority of available toxins identified in mammals as only one sequence has been published. Data were predominately collected from published literature using either review papers to identify recently published papers, then using key word searches on Google Scholar ('venomous mammal', 'poisonous mammal' or 'toxins mammals') and then finally cross referencing the found toxin sequences using UniProt, GenBank and EMBL-EBI to confirm best available sequences for analysis. To ensure no sequences where missed, a final search on GenBank and UniProt using either the venomous species or toxin name were conducted. Sequences were also obtained via individual repositories or via emailing researchers (pers. comm. Nick Casewell). For a full list of where sequences were obtained, see Spreadsheet S1.

5.2. Alignments and Trimming

Toxins that had multiple representatives were identified and their nucleotide sequences where grouped together in Geneious Prime v2022.1.1 (https://www.geneious.com/; Biomatters, Auckland, New Zealand) (Table S1), aligned with MUSCLE 3.8.425 [68] and finished with manual editing of alignment including necessary removal of STOP codons for use in analysis. Genus or Order specific alignments (such as CRiSP *Desmodus*) were done as well as Class wide alignments to confirm significance as well as identify site-specific changes.

Aligned sequences were run through TrimAI v1.3 (accessed via Phylemon 2.0 web server) [69] on automated1 mode (chosen as this mode is optimised for alignments used in maximum likelihood analysis and identifies if gappyout or strict will be more suitable to use on the alignment based on average identify score, sequence number and maximum identity score). Spurious sequences were not removed. Trimmed output was downloaded as FASTA files.

5.3. Datamonkey Selection Analysis

Selection analysis was conducted on Datamonkey Adaptive Evolution Server [70–72] looking at both branch-site and site-specific selection levels on each alignment. Site-specific models examine which amino acid positions in an alignment are evolving under selection, with level of significance given at 0.95 or 0.99. Site-specific models cannot indicate which sequences the site-specific selection is found in. Branch site level selection levels can indicate which branch or node selection has occurred in an alignment and specifically which sites.

5.4. Site-Specific Selection
5.4.1. Fast Unconstrained Bayesian AppRoximation

FUBAR (Fast Unconstrained Bayesian AppRoximation [73], a Bayesian selection model) was used to examine pervasive site selection. FUBAR assumed a constantly level of selection across the whole alignment and significance is taken at posterior probability < 0.9. FUBAR reports if diversifying or purifying selection is present at a site.

5.4.2. Mixed Effects of Model Evolution

MEME (Mixed Effects of Model Evolution [74], a maximum likelihood model) was used to examine episodic site selection with significance taken at <0.05. MEME will examine a proportion of branches at a time and then infers two different ω classes with different weighting, Null and Alternative. Both classes measure alpha (dS) and two different beta (dN) values, Beta+ and Beta-. Null allows values in alpha to be unrestricted but both beta values must be smaller or the same as alpha. Alt allows values in alpha and beta+ to be unrestricted but beta- must be the same or smaller than alpha. Sites are examined for the probability that a site evolves at each ω class on a given branch. If beta+ is greater than alpha (and adjusted with LRT) then a site is assumed to be under diversifying selection. MEME will report if you have diversifying selection present at a site only.

5.5. Branch Specific Selection
Adaptive Branch-Site Random Effective Likelihood

aBSREL (Adaptive Branch-Site Random Effective Likelihood [75]) was used to examine episodic branch-site selection on each alignment with significance set at <0.05 and all branches selected. aBSREL measures both site and branch ω values but will only report branch results for tests. aBSREL examines each branch in the phylogenetic tree it constructs from the alignment to measure which proportion has evolved under diversifying selection. The Akaike Information Criterion corrected (AICc) is used to calculate the optimal rate present on each branch and LRT is used to fit the best model on each branch. These optimal models are then compared to a null mode to identify diversifying selection. aBSERL will report if you have diversifying selection present on a branch.

5.6. PAML Selection Tests
5.6.1. Tree Data

Maximum likelihood trees were generated for each trimmed alignment, with an appropriate model selected using MEGAX [76] model selection test (Table S27), ran with 1000 bootstrap replicates and unrooted. Each tree was then saved as two separate Newick files (one with branch lengths and support, one clean tree with no labels).

5.6.2. Selection Tests Using CodeML

Trimmed FASTA files were first converted into PAML files using the file converter function available in EasycodeML [77]. For the selection tests, CodeML (from PAML v4.8 [78]) was opened via the command line. To confirm if the dN/dS site-specific results are significant, the alternative model must be compared to a null model and compared via an LRT. Therefore, six models were chosen for site-specific comparison: three null models, M0 (assumes constant ω value across alignment), M1a (assumes only neutral or purifying

evolution across the alignment) and M7 (assumes neutral evolution as well but using beta) and three alternative models M3 (allows some level of variation amongst the alignment but very limited) M2a (assumes diversifying, neutral or purifying evolution across the alignment) and M8 (assumes neutral or diversifying evolution as well but using beta)). Additionally, M8a (a null model that assumes neutral evolution using the beta model) was used for comparison with M8 model, as it provides the best supported result for each alignment. The comparisons were M0 vs. M3, M1a vs. M2a, M7 vs. M8 and M8a vs. M8. The clean Newick file and PAML file were placed into a folder with a codeml.ctl file with the above models coded (M8a was coded separately with a new codeml.ctl file as ω had to be changed to equal 1 from default of 0.4). All options remained default excluding the option to clean the data which was selected.

5.7. Comparison of Toxin Sequences to Ancestral Group

5.7.1. Tree Selection on Amino Acid Properties

To investigate the impact sites evolving under natural selection have on the protein structure of these alignments and to confer with sites identified by selection tests used in this meta-analysis, TreeSAAP v1.4 [79] was used on each alignment's FASTA file and cleaned Newick files where converted into NEXUS format and then ran through on an initial 'length' sliding window. The Z-score statistics were then used to identify which of the 32 properties were evolving under the highest amount of selection (ranked from 8 (most) to 1 (least)) with only those properties at +3.092 significance and ranked 8–6 taken as the properties worth investigating. The alignments were run again with the list of interesting properties and a sliding window of one (1) was used to examine results.

5.7.2. Visualisation of Results

Structural models for proteins with amino acid positions under diversifying selection were obtained from AlphaFold Protein Structure Database [80,81]. They were represented as cartoons following Richardson conventions [82] and the images were rendered with PyMol 1.3 (Schrodinger, LLC., New York, NY, USA). Functional annotation was obtained from UniProt [83].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/toxins14120842/s1, Figure S1: Amino acid alignment of KLK1 *Blarina*, with diversifying selection highlighted according to each programme, Figure S2: Amino acid alignment of KLK1 *Solenodon*, with diversifying selection highlighted according to each programme, Figure S3: Amino acid alignment of KLK1 Eulipotyphla, with diversifying selection highlighted according to each programme, Figure S4: Amino acid alignment of Plasminogen Activator, with diversifying selection highlighted according to each programme, Figure S5: Amino acid alignment of Desmallipins, with diversifying selection highlighted according to each programme, Figure S6: Amino acid alignment of PACAP, with diversifying selection highlighted according to each programme, Figure S7: Amino acid alignment of CRiSP *Desmodus*, with diversifying selection highlighted according to each programme, Figure S8: Amino acid alignment of CRiSP *Desmodus* and *Ornithorhynchus*, with diversifying selection highlighted according to each programme, Figure S9: Amino acid alignment of Kunitz Domain One, with diversifying selection highlighted according to each programme, Figure S10: Amino acid alignment of Kunitz Domain Two, with diversifying selection highlighted according to each programme, Table S1: The number of sequences and length of each alignment, Table S2: FEL results for each alignment including number of sites identified to be under significant diversifying or purifying selection and at which site in each alignment, Table S3: FUBAR results for each alignment including number of sites identified to be under significant diversifying or purifying selection and at which site in each alignment, Table S4: MEME results for each alignment including number of sites identified to be under significant diversifying selection and at which site in each alignment, Table S5: CodeML results for the alignment of KLK1 Blarina brevicauda, including results of all models identified in this study, Table S6: Significance of CodeML results for the alignment of KLK1 Blarina brevicauda, Table S7: CodeML results for the alignment of KLK1 Solenodon paradoxus, including results of all models identified in this study, Table S8: Significance of CodeML results

for the alignment of KLK1 Solenodon paradoxus, Table S9: CodeML results for the alignment of KLK1 Eulipotyphla, including results of all models identified in this study, Table S10: Significance of CodeML results for the alignment of KLK1 Eulipotyphla, Table S11: CodeML results for the alignment of KLK1 All, including results of all models identified in this study, Table S12: Significance of CodeML results for the alignment of KLK1 All, Table S13: CodeML results for the alignment of Plasminogen Activator, including results of all models identified in this study, Table S14: Significance of CodeML results for the alignment of Plasminogen Activator, Table S15: CodeML results for the alignment of Desmallipins, including results of all models identified in this study, Table S16: Significance of CodeML results for the alignment of Desmallipins, Table S17: CodeML results for the alignment of PACAP, including results of all models identified in this study, Table S18: Significance of CodeML results for the alignment of PACAP, Table S19: CodeML results for the alignment of CRiSP Desmodus, including results of all models identified in this study, Table S20: Significance of CodeML results for the alignment of CRiSP Desmodus, Table S21: CodeML results for the alignment of CRiSP All, including results of all models identified in this study, Table S22: Significance of CodeML results for the alignment of CRiSP All, Table S23: CodeML results for the alignment of Kunitz Domain One, including results of all models identified in this study, Table S24: Significance of CodeML results for the alignment of Kunitz Domain One, Table S25: CodeML results for the alignment of Kunitz Domain Two, including results of all models identified in this study, Table S26: Significance of CodeML results for the alignment of Kunitz Domain Two, Table S27: TreeSAAP values for each alignment, including magnitude, the value and the site number corresponding to each alignment, Table S28: Model selection result from MEGA for maximum likelihood tree for each alignment.

Author Contributions: Conceptualization, L.L.J.F., K.A.-I.N. and V.N.; methodology, L.L.J.F.; analysis, L.L.J.F. and R.L.-B.; investigation, L.L.J.F.; writing—original draft preparation, L.L.J.F.; writing—review and editing, K.A.-I.N., V.N. and R.L.-B.; supervision, K.A.-I.N. and V.N. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All eleven alignments used for analysis are available to download here. For any additional files or questions from analysis' used in this study, please contact the first author (L.L.J.F.).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Surm, J.M.; Moran, Y. Insights into How Development and Life-History Dynamics Shape the Evolution of Venom. *EvoDevo* **2021**, *12*, 1. [CrossRef] [PubMed]
2. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex Cocktails: The Evolutionary Novelty of Venoms. *Trends Ecol. Evol.* **2013**, *28*, 219–229. [CrossRef] [PubMed]
3. Schendel, V.; Rash, L.D.; Jenner, R.A.; Undheim, E.A.B. The Diversity of Venom: The Importance of Behavior and Venom System Morphology in Understanding Its Ecology and Evolution. *Toxins* **2019**, *11*, 666. [CrossRef] [PubMed]
4. Walker, A.A. The Evolutionary Dynamics of Venom Toxins Made by Insects and Other Animals. *Biochem. Soc. Trans.* **2020**, *48*, 1353–1365. [CrossRef] [PubMed]
5. Kryazhimskiy, S.; Plotkin, J.B. The Population Genetics of DN/DS. *PLOS Genet.* **2008**, *4*, e1000304. [CrossRef]
6. Chang, D.; Duda, T.F., Jr. Extensive and Continuous Duplication Facilitates Rapid Evolution and Diversification of Gene Families. *Mol. Biol. Evol.* **2012**, *29*, 2019–2029. [CrossRef]

7. Hargreaves, A.D.; Swain, M.T.; Hegarty, M.J.; Logan, D.W.; Mulley, J.F. Restriction and Recruitment—Gene Duplication and the Origin and Evolution of Snake Venom Toxins. *Genome Biol. Evol.* **2014**, *6*, 2088–2095. [CrossRef]

8. Aird, S.D.; Aggarwal, S.; Villar-Briones, A.; Tin, M.M.-Y.; Terada, K.; Mikheyev, A.S. Snake Venoms Are Integrated Systems, but Abundant Venom Proteins Evolve More Rapidly. *BMC Genom.* **2015**, *16*, 647. [CrossRef]

9. Sunagar, K.; Moran, Y. The Rise and Fall of an Evolutionary Innovation: Contrasting Strategies of Venom Evolution in Ancient and Young Animals. *PLOS Genet.* **2015**, *11*, e1005596. [CrossRef]

10. Ligabue-Braun, R. Venom Use in Mammals: Evolutionary Aspects. In *Evolution of Venomous Animals and Their Toxins*; Gopalakrishnakone, P., Malhotra, A., Eds.; Springer: Dordrecht, The Netherlands, 2016; pp. 1–23. ISBN 978-94-007-6727-0.

11. Rode-Margono, J.E.; Nekaris, K.A.-I. Cabinet of Curiosities: Venom Systems and Their Ecological Function in Mammals, with a Focus on Primates. *Toxins* **2015**, *7*, 2639. [CrossRef] [PubMed]

12. Kumar, S.; Stecher, G.; Suleski, M.; Hedges, S.B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol. Biol. Evol.* **2017**, *34*, 1812–1819. [CrossRef] [PubMed]

13. Kowalski, K.; Rychlik, L. Venom Use in Eulipotyphlans: An Evolutionary and Ecological Approach. *Toxins* **2021**, *13*, 231. [CrossRef]

14. Warren, W.C.; Hillier, L.W.; Marshall Graves, J.A.; Birney, E.; Ponting, C.P.; Grützner, F.; Belov, K.; Miller, W.; Clarke, L.; Chinwalla, A.T.; et al. Genome Analysis of the Platypus Reveals Unique Signatures of Evolution. *Nature* **2008**, *453*, 175–183. [CrossRef] [PubMed]

15. Wong, E.S.W.; Nicol, S.; Warren, W.C.; Belov, K. Echidna Venom Gland Transcriptome Provides Insights into the Evolution of Monotreme Venom. *PLOS ONE* **2013**, *8*, e79092. [CrossRef] [PubMed]

16. Casewell, N.R.; Petras, D.; Card, D.C.; Suranse, V.; Mychajliw, A.M.; Richards, D.; Koludarov, I.; Albulescu, L.-O.; Slagboom, J.; Hempel, B.-F.; et al. Solenodon Genome Reveals Convergent Evolution of Venom in Eulipotyphlan Mammals. *Proc. Natl. Acad. Sci.* **2019**, *116*, 25745–25755. [CrossRef]

17. Hanf, Z.R.; Chavez, A.S. A Comprehensive Multi-Omic Approach Reveals a Relatively Simple Venom in a Diet Generalist, the Northern Short-Tailed Shrew, Blarina Brevicauda. *Genome Biol. Evol.* **2020**, *12*, 1148–1166. [CrossRef]

18. Liao, Z.; Tang, X.; Chen, W.; Jiang, X.; Chen, Z.; He, K.; Li, Q.; Duan, Z.; He, X.; Kamau, P.M.; et al. Shrew's Venom Quickly Causes Circulation Disorder, Analgesia and Hypokinesia. *Cell. Mol. Life Sci.* **2022**, *79*, 35. [CrossRef]

19. Rychlik, L. CHANGES IN PREY SIZE PREFERENCES DURING SUCCESSIVE STAGES OF FORAGING IN THE MEDITERRANEAN WATER SHREW NEOMYS ANOMALUS. *Behaviour* **1999**, *136*, 345–365. [CrossRef]

20. Kowalski, K.; Marciniak, P.; Rosiński, G.; Rychlik, L. Evaluation of the Physiological Activity of Venom from the Eurasian Water Shrew Neomys Fodiens. *Front. Zool.* **2017**, *14*, 46. [CrossRef] [PubMed]

21. Camargo, I.; Alvarez-Castañeda, S.T. Analyses of Predation Behavior of the Desert Shrew Notiosorex Crawfordi. *Mammalia* **2018**, *83*, 276–280. [CrossRef]

22. Lopez-Jurado, L.F.; Mateo, J.A. Evidence of Venom in the Canarian Shrew (*Crocidura Canariensis*): Immobilizing Effects on the Atlantic Lizard (*Gallotia Atlantica*). *J. Zool.* **1996**, *239*, 394–395. [CrossRef]

23. Low, D.H.W.; Sunagar, K.; Undheim, E.A.B.; Ali, S.A.; Alagon, A.C.; Ruder, T.; Jackson, T.N.W.; Pineda Gonzalez, S.; King, G.F.; Jones, A.; et al. Dracula's Children: Molecular Evolution of Vampire Bat Venom. *J. Proteom.* **2013**, *89*, 95–111. [CrossRef] [PubMed]

24. Tellgren-Roth, Å.; Dittmar, K.; Massey, S.E.; Kemi, C.; Tellgren-Roth, C.; Savolainen, P.; Lyons, L.A.; Liberles, D.A. Keeping the Blood Flowing—Plasminogen Activator Genes and Feeding Behavior in Vampire Bats. *Naturwissenschaften* **2009**, *96*, 39–47. [CrossRef]

25. Hagey, L.R.; Fry, B.G.; Fitch-Snyder, H. Talking Defensively, a Dual Use for the Brachial Gland Exudate of Slow and Pygmy Lorises. In *Primate Anti-Predator Strategies*; Gursky, S.L., Nekaris, K.A.I., Eds.; Developments in Primatology: Progress and Prospects; Springer: Boston, MA, USA, 2007; pp. 253–272. ISBN 978-0-387-34807-0.

26. Scheib, H.; Nekaris, K.A.-I.; Rode-Margono, J.; Ragnarsson, L.; Baumann, K.; Dobson, J.S.; Wirdateti, W.; Nouwens, A.; Nijman, V.; Martelli, P.; et al. The Toxicological Intersection between Allergen and Toxin: A Structural Comparison of the Cat Dander Allergenic Protein Fel D1 and the Slow Loris Brachial Gland Secretion Protein. *Toxins* **2020**, *12*, 86. [CrossRef]

27. Madani, G.; Nekaris, K.A.-I. Anaphylactic Shock Following the Bite of a Wild Kayan Slow Loris (Nycticebus Kayan): Implications for Slow Loris Conservation. *J. Venom. Anim. Toxins Trop. Dis.* **2014**, *20*, 43. [CrossRef]

28. Gardiner, M.; Weldon, A.; Poindexter, S.A.; Gibson, N.; Nekaris, K.A.I. Survey of Practitioners Handling Slow Lorises (Primates: *Nycticebus*): An Assessment of the Harmful Effects of Slow Loris Bites. *J. Venom Res.* **2018**, *9*, 1–7.

29. Bal, A.K.; Giordano, A.J.; Gouda, S. Effects of a Bengal Slow Loris Nycticebus Bengalensis (Primates: Lorisidae) Bite: A Case Study from Murlen National Park, Mizoram, India. *J. Threat. Taxa* **2022**, *14*, 21449–21452. [CrossRef]

30. Nekaris, K.A.-I.; Moore, R.S.; Rode, E.J.; Fry, B.G. Mad, Bad and Dangerous to Know: The Biochemistry, Ecology and Evolution of Slow Loris Venom. *J. Venom. Anim. Toxins Trop. Dis.* **2013**, *19*, 1–10. [CrossRef] [PubMed]

31. Thomas, J.L.; Parrott, M.L.; Handasyde, K.A.; Temple-Smith, P. Female Control of Reproductive Behaviour in the Platypus (*Ornithorhynchusa anatinus*), with Notes on Female Competition for Mating. *Behaviour* **2018**, *155*, 27–53. [CrossRef]

32. Nekaris, K.A.I.; Campera, M.; Nijman, V.; Birot, H.; Rode-Margono, E.J.; Fry, B.G.; Weldon, A.; Wirdateti, W.; Imron, M.A. Slow Lorises Use Venom as a Weapon in Intraspecific Competition. *Curr. Biol.* **2020**, *30*, R1252–R1253. [CrossRef] [PubMed]

33. Kakumanu, R.; Hodgson, W.C.; Ravi, R.; Alagon, A.; Harris, R.J.; Brust, A.; Alewood, P.F.; Kemp-Harper, B.K.; Fry, B.G. Vampire Venom: Vasodilatory Mechanisms of Vampire Bat (*Desmodus rotundus*) Blood Feeding. *Toxins* **2019**, *11*, 26. [CrossRef]

34. Zhou, Y.; Shearwin-Whyatt, L.; Li, J.; Song, Z.; Hayakawa, T.; Stevens, D.; Fenelon, J.C.; Peel, E.; Cheng, Y.; Pajpach, F.; et al. Platypus and Echidna Genomes Reveal Mammalian Biology and Evolution. *Nature* **2021**, *592*, 756–762. [CrossRef]

35. Hawkins, M.; Battaglia, A. Breeding Behaviour of the Platypus (*Ornithorhynchus anatinus*) in Captivity. *Aust. J. Zool.* **2009**, *57*, 283. [CrossRef]

36. Whittington, C.M.; Papenfuss, A.T.; Locke, D.P.; Mardis, E.R.; Wilson, R.K.; Abubucker, S.; Mitreva, M.; Wong, E.S.; Hsu, A.L.; Kuchel, P.W.; et al. Novel Venom Gene Discovery in the Platypus. *Genome Biol.* **2010**, *11*, R95. [CrossRef]

37. Kowalski, K.; Marciniak, P.; Rychlik, L. A New, Widespread Venomous Mammal Species: Hemolytic Activity of *Sorex araneus* Venom Is Similar to That of *Neomys fodiens* Venom. *Zool. Lett.* **2022**, *8*, 7. [CrossRef]

38. Wong, E.S.W.; Morgenstern, D.; Mofiz, E.; Gombert, S.; Morris, K.M.; Temple-Smith, P.; Renfree, M.B.; Whittington, C.M.; King, G.F.; Warren, W.C.; et al. Proteomics and Deep Sequencing Comparison of Seasonally Active Venom Glands in the Platypus Reveals Novel Venom Peptides and Distinct Expression Profiles. *Mol. Cell. Proteom.* **2012**, *11*, 1354–1364. [CrossRef]

39. Barua, A.; Koludarov, I.; Mikheyev, A.S. Co-Option of the Same Ancestral Gene Family Gave Rise to Mammalian and Reptilian Toxins. *BMC Biol.* **2021**, *19*, 268. [CrossRef]

40. Tokuriki, N.; Stricher, F.; Schymkowitz, J.; Serrano, L.; Tawfik, D.S. The Stability Effects of Protein Mutations Appear to Be Universally Distributed. *J. Mol. Biol.* **2007**, *369*, 1318–1332. [CrossRef]

41. Tokuriki, N.; Tawfik, D.S. Stability Effects of Mutations and Protein Evolvability. *Curr. Opin. Struct. Biol.* **2009**, *19*, 596–604. [CrossRef]

42. Studer, R.A.; Dessailly, B.H.; Orengo, C.A. Residue Mutations and Their Impact on Protein Structure and Function: Detecting Beneficial and Pathogenic Changes. *Biochem. J.* **2013**, *449*, 581–594. [CrossRef]

43. Roelants, K.; Champagne, D.; Scheib, H.; Tyndall, J.; King, G.; Nevalainen, T.; Norman, J.; Lewis, R.; Norton, R.; Renjifo, C.; et al. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms. *Annu. Rev. Genom. Hum. Genet.* **2009**, *10*, 483–511. [CrossRef]

44. Aminetzach, Y.T.; Srouji, J.R.; Kong, C.Y.; Hoekstra, H.E. Convergent Evolution of Novel Protein Function in Shrew and Lizard Venom. *Curr. Biol.* **2009**, *19*, 1925–1931. [CrossRef]

45. Reily, C.; Stewart, T.J.; Renfrow, M.B.; Novak, J. Glycosylation in Health and Disease. *Nat. Rev. Nephrol.* **2019**, *15*, 346–366. [CrossRef] [PubMed]

46. Fernandez, A.Z.; Tablante, A.; Bartoli, F.; Beguin, S.; Hemker, H.C.; Apitz-Castro, R. Expression of Biological Activity of Draculin, the Anticoagulant Factor from Vampire Bat Saliva, Is Strictly Dependent on the Appropriate Glycosylation of the Native Molecule. *Biochim. Biophys. Acta BBA—Gen. Subj.* **1998**, *1425*, 291–299. [CrossRef] [PubMed]

47. Owji, H.; Nezafat, N.; Negahdaripour, M.; Hajiebrahimi, A.; Ghasemi, Y. A Comprehensive Review of Signal Peptides: Structure, Roles, and Applications. *Eur. J. Cell Biol.* **2018**, *97*, 422–441. [CrossRef] [PubMed]

48. Harris, R.J.; Arbuckle, K. Tempo and Mode of the Evolution of Venom and Poison in Tetrapods. *Toxins* **2016**, *8*, 193. [CrossRef]

49. Botero-Castro, F.; Tilak, M.-K.; Justy, F.; Catzeflis, F.; Delsuc, F.; Douzery, E.J.P. In Cold Blood: Compositional Bias and Positive Selection Drive the High Evolutionary Rate of Vampire Bats Mitochondrial Genomes. *Genome Biol. Evol.* **2018**, *10*, 2218–2239. [CrossRef]

50. Wong, E.S.W.; Belov, K. Venom Evolution through Gene Duplications. *Gene* **2012**, *496*, 1–7. [CrossRef]

51. Casewell, N.R.; Jackson, T.N.W.; Laustsen, A.H.; Sunagar, K. Causes and Consequences of Snake Venom Variation. *Trends Pharmacol. Sci.* **2020**, *41*, 570–581. [CrossRef] [PubMed]

52. Phillips, C.D.; Baker, R.J. Secretory Gene Recruitments in Vampire Bat Salivary Adaptation and Potential Convergences With Sanguivorous Leeches. *Front. Ecol. Evol.* **2015**, *3*, 122. [CrossRef]

53. Folinsbee, K.E. Evolution of Venom across Extant and Extinct Eulipotyphlans. *Comptes Rendus Palevol.* **2013**, *12*, 531–542. [CrossRef]

54. Barua, A.; Mikheyev, A.S. An Ancient, Conserved Gene Regulatory Network Led to the Rise of Oral Venom Systems. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2021311118. [CrossRef] [PubMed]

55. Chaverri, G.; Ancillotto, L.; Russo, D. Social Communication in Bats. *Biol. Rev.* **2018**, *93*, 1938–1954. [CrossRef] [PubMed]

56. Francischetti, I.M.B.; Assumpção, T.C.F.; Ma, D.; Li, Y.; Vicente, E.C.; Uieda, W.; Ribeiro, J.M.C. The "Vampirome": Transcriptome and Proteome Analysis of the Principal and Accessory Submaxillary Glands of the Vampire Bat Desmodus Rotundus, a Vector of Human Rabies. *J. Proteom.* **2013**, *82*, 288–319. [CrossRef]

57. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

58. Arévalo, L.; Brukman, N.G.; Cuasnicú, P.S.; Roldan, E.R.S. Evolutionary Analysis of Genes Coding for Cysteine-RIch Secretory Proteins (CRISPs) in Mammals. *BMC Evol. Biol.* **2020**, *20*, 67. [CrossRef]

59. Sunagar, K.; Johnson, W.E.; O'Brien, S.J.; Vasconcelos, V.; Antunes, A. Evolution of CRISPs Associated with Toxicoferan-Reptilian Venom and Mammalian Reproduction. *Mol. Biol. Evol.* **2012**, *29*, 1807–1822. [CrossRef]

60. Jackson, T.N.W.; Koludarov, I. How the Toxin Got Its Toxicity. *Front. Pharmacol.* **2020**, *11*, 574925. [CrossRef]

61. Potter, J.H.T.; Davies, K.T.J.; Yohe, L.R.; Sanchez, M.K.R.; Rengifo, E.M.; Struebig, M.; Warren, K.; Tsagkogeorga, G.; Lim, B.K.; dos Reis, M.; et al. Dietary Diversification and Specialization in Neotropical Bats Facilitated by Early Molecular Evolution. *Mol. Biol. Evol.* **2021**, *38*, 3864–3883. [CrossRef]

62. Zepeda Mendoza, M.L.; Xiong, Z.; Escalera-Zamudio, M.; Runge, A.K.; Thézé, J.; Streicker, D.; Frank, H.K.; Loza-Rubio, E.; Liu, S.; Ryder, O.A.; et al. Hologenomic Adaptations Underlying the Evolution of Sanguivory in the Common Vampire Bat. *Nat. Ecol. Evol.* **2018**, *2*, 659–668. [CrossRef]

63. Moreno Santillán, D.D.; Lama, T.M.; Gutierrez Guerrero, Y.T.; Brown, A.M.; Donat, P.; Zhao, H.; Rossiter, S.J.; Yohe, L.R.; Potter, J.H.; Teeling, E.C.; et al. Large-Scale Genome Sampling Reveals Unique Immunity and Metabolic Adaptations in Bats. *Mol. Ecol.* **2021**, *30*, 6449–6467. [CrossRef] [PubMed]

64. Vandewege, M.W.; Sotero-Caio, C.G.; Phillips, C.D. Positive Selection and Gene Expression Analyses from Salivary Glands Reveal Discrete Adaptations within the Ecologically Diverse Bat Family Phyllostomidae. *Genome Biol. Evol.* **2020**, *12*, 1419–1428. [CrossRef] [PubMed]

65. Yarlagadda, K.; Razik, I.; Malhi, R.S.; Carter, G.G. Social Convergence of Gut Microbiomes in Vampire Bats. *Biol. Lett.* **2021**, *17*, 20210389. [CrossRef] [PubMed]

66. Zhu, L.; Peigneur, S.; Gao, B.; Zhang, S.; Tytgat, J.; Zhu, S. Target-Driven Positive Selection at Hot Spots of Scorpion Toxins Uncovers Their Potential in Design of Insecticides. *Mol. Biol. Evol.* **2016**, *33*, 1907–1920. [CrossRef]

67. Zhang, S.; Gao, B.; Zhu, S. Target-Driven Evolution of Scorpion Toxins. *Sci. Rep.* **2015**, *5*, 14973. [CrossRef]

68. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef]

69. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]

70. Delport, W.; Poon, A.F.Y.; Frost, S.D.W.; Kosakovsky Pond, S.L. Datamonkey 2010: A Suite of Phylogenetic Analysis Tools for Evolutionary Biology. *Bioinformatics* **2010**, *26*, 2455–2457. [CrossRef]

71. Pond, S.L.K.; Frost, S.D.W. Datamonkey: Rapid Detection of Selective Pressure on Individual Sites of Codon Alignments. *Bioinformatics* **2005**, *21*, 2531–2533. [CrossRef]

72. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [CrossRef]

73. Murrell, B.; Moola, S.; Mabona, A.; Weighill, T.; Sheward, D.; Kosakovsky Pond, S.L.; Scheffler, K. FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. *Mol. Biol. Evol.* **2013**, *30*, 1196–1205. [CrossRef]

74. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Pond, S.L.K. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLOS Genet.* **2012**, *8*, e1002764. [CrossRef]

75. Kosakovsky Pond, S.L.; Murrell, B.; Fourment, M.; Frost, S.D.W.; Delport, W.; Scheffler, K. A Random Effects Branch-Site Model for Detecting Episodic Diversifying Selection. *Mol. Biol. Evol.* **2011**, *28*, 3033–3043. [CrossRef]

76. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]

77. Gao, F.; Chen, C.; Arab, D.A.; Du, Z.; He, Y.; Ho, S.Y.W. EasyCodeML: A Visual Tool for Analysis of Selection Using CodeML. *Ecol. Evol.* **2019**, *9*, 3891–3898. [CrossRef]

78. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [CrossRef]

79. Woolley, S.; Johnson, J.; Smith, M.J.; Crandall, K.A.; McClellan, D.A. TreeSAAP: Selection on Amino Acid Properties Using Phylogenetic Trees. *Bioinformatics* **2003**, *19*, 671–672. [CrossRef]

80. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef]

81. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Žídek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; et al. Highly Accurate Protein Structure Prediction for the Human Proteome. *Nature* **2021**, *596*, 590–596. [CrossRef]

82. Richardson, J.S. The Anatomy and Taxonomy of Protein Structure. In *Advances in Protein Chemistry*; Anfinsen, C.B., Edsall, J.T., Richards, F.M., Eds.; Academic Press: Cambridge, MA, USA, 1981; Volume 34, pp. 167–339.

83. The UniProt Consortium UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]