**Title:** Solenodon genome reveals convergent evolution of venom in eulipotyphlan mammals

**Short Title:** Venom evolution in eulipotyphlans

**Authors:** Nicholas R. Casewell1,\*, Daniel Petras2,3, Daren C. Card4,5,6, Vivek Suranse7, Alexis M. Mychajliw8,9,10, David Richards11,12, Ivan Koludarov13, Laura-Oana Albulescu1, Julien Slagboom14, Benjamin-Florian Hempel2, Neville M. Ngum11, Rosalind J. Kennerley15, Jorge L. Brocca16, Gareth Whiteley1, Robert A. Harrison1, Fiona M. S. Bolton1, Jordan Debono17, Freek J. Vonk18, Jessica Alföldi19, Jeremy Johnson19, Elinor K. Karlsson19,20, Kerstin Lindblad-Toh19,21, Ian R. Mellor11, Roderich D. Süssmuth2, Bryan G. Fry17, Sanjaya Kuruppu22,23, Wayne C. Hodgson22, Jeroen Kool14, Todd A. Castoe4, Ian Barnes24, Kartik Sunagar7, Eivind A. B. Undheim25,26,27 and Samuel T. Turvey28.

**Affiliations:**

1 Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK

2 Technische Universität Berlin, Institut für Chemie, Strasse des 17. Juni 124, 10623 Berlin, Germany

3 University of California San Diego, Collaborative Mass Spectrometry Innovation Center, 9500 Gilman Drive, La Jolla, California 92093, USA

4 Department of Biology, University of Texas at Arlington, 501 S. Nedderman Dr., Arlington, Texas 76010, USA

5 Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford St., Cambridge, MA, 02138, USA

6 Museum of Comparative Zoology, Harvard University, 26 Oxford St., Cambridge, MA, 02138, USA

7 Evolutionary Venomics Lab, Centre for Ecological Sciences, Indian Institute of Science, Bangalore, 560012, India

8 Department of Biology, Stanford University, 371 Serra Mall, Stanford, California 94305, USA.

9 Department of Rancho La Brea, Natural History Museum of Los Angeles County, 5801 Wilshire Blvd, Los Angeles, California 90036, USA.

10 Institute of Low Temperature Science, Hokkaido University, Kita-19 Nishi-8, Kita-ku, Sapporo 060-0819, Japan.

11 School of Life Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

12 Biomedical Research Centre, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

13 Ecology and Evolution Unit, Okinawa Institute of Science and Technology, Onna, Kunigami-gun, Okinawa, Japan

14 Division of BioAnalytical Chemistry, Amsterdam Institute of Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

15 Durrell Wildlife Conservation Trust, Les Augrès Manor, Trinity, Jersey JE3 5BP, British Channel Islands

16 SOH Conservación, Apto. 401 Residencial Las Galerías, Calle Gustavo Mejia Ricart No. 119 B, Santo Domingo, Dominican Republic

17 Venom Evolution Lab, School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4067, Australia

18 Naturalis Biodiversity Center, 2333 CR, Leiden, The Netherlands

19 Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA

20 Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01655, USA

21 Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

22 Monash Venom Group, Department of Pharmacology, Biomedicine Discovery Institute, Monash University, VIC 3800, Australia

23 Department of Biochemistry & Molecular Biology, Biomedicine Discovery Institute, Monash University, VIC 3800, Australia

24 Department of Earth Sciences, Natural History Museum, London, SW7 5BD, UK

25 Centre for Advanced Imaging, the University of Queensland, Brisbane QLD 4072, Australia

26 Institute for Molecular Bioscience, the University of Queensland, Brisbane QLD 4072, Australia

27 Centre for Ecology and Evolutionary Synthesis, Department of Biosciences, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway

28 Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK

**\* Corresponding author:** Nicholas R. Casewell, Centre for Snakebite Research & Interventions, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK. Email: Nicholas.casewell@lstmed.ac.uk

**Key words:** convergent molecular evolution; genotype phenotype; gene duplication; venom systems; kallikrein toxin

**Classification:** Biological Science: Evolution

**Abstract**

Venom systems are key adaptations that have evolved throughout the tree of life and typically facilitate predation or defence. Despite venoms being model systems for studying a variety of evolutionary and physiological processes, many taxonomic groups remain understudied, including venomous mammals. Within the Order Eulipotyphla, multiple shrew species and solenodons have oral venom systems. Despite morphological variation of their delivery systems, it remains unclear whether venom represents the ancestral state in this group or is the result of multiple independent origins. We investigated the origin and evolution of venom in eulipotyphlans by characterising the venom system of the endangered Hispaniolan solenodon (*Solenodon paradoxus*). We constructed a genome to underpin proteomic identifications of solenodon venom toxins, before undertaking evolutionary analyses of those constituents, and functional assessments of the secreted venom. Our findings show that solenodon venom consists of multiple paralogous kallikrein 1 (*KLK1*) serine proteases, which cause hypotensive effects *in vivo*, and seem likely to have evolved to facilitate vertebrate prey capture. Comparative analyses provide convincing evidence that the oral venom systems of solenodons and shrews have evolved convergently, with the four independent origins of venom in eulipotyphlans outnumbering all other venom origins in mammals. We find that *KLK1*s have been independently co-opted into the venom of shrews and solenodons following their divergence during the late Cretaceous, suggesting that evolutionary constraints may be acting on these genes. Consequently, our findings represent a striking example of convergent molecular evolution, and demonstrate that distinct structural backgrounds can yield equivalent functions.

**Significance Statement**

Multiple representatives of eulipotyphlan mammals (shrews, hedgehogs, moles and solenodons) are venomous, but little is known about the evolutionary history and composition of their oral venom systems. Herein we characterised venom from the endangered Hispaniolan solenodon (*Solenodon paradoxus*) and find that it consists of hypotensive proteins likely used to facilitate vertebrate prey capture. We demonstrate that venom has evolved independently on at least four occasions in eulipotyphlans, and that molecular components of these venoms have also evolved convergently, with kallikrein-1 proteins co-opted as toxins in both solenodons and shrews following their divergence over 70 million years ago. Our findings present an elegant example of convergent molecular evolution, and highlight that mammalian venom systems may be subjected to evolutionary constraints.

**Introduction**

Venom systems are key ecological innovations that have evolved independently on numerous occasions across the tree of life (1). They consist of mixtures of proteinaceous components (commonly referred to as toxins), and can be defined as secretions produced in specialised tissues that cause physiological perturbations when delivered into other animals through a wound caused by a venom delivery apparatus (2). Venoms have proven to be valuable systems for understanding a variety of different evolutionary processes, including those relating to convergence (1, 2), accelerated molecular evolution (3), gene duplication (4) and protein neofunctionalization (5). Venoms are also of great medical importance, both due to the harm they can cause to people (e.g. >100,000 people die annually as a result of snake envenoming (6)), and for the value of their highly selective toxins for understanding physiological processes and the development of new pharmaceuticals (7).

Ecologically, venoms are primarily used for prey capture and/or to defend the producing animal from aggressors or predators, though in some instances venom is utilised for intraspecific competition or to facilitate offspring survival (1). Despite extensive research focus on a number of venomous lineages, many remain almost completely unstudied, including venomous mammals. Mammalian venom systems are rare and, based on the definition above, are restricted to members of four extant orders: the monotremes, chiropterans, primates, and eulipotyphlans (8). Their venoms are utilised for distinct ecological purposes, such as male-male combat to facilitate breeding (platypus; *Ornithorhynchus anatinus*), aiding haematophagy (vampire bats; *Desmodus rotundus*, *Diaemus youngi* and *Diphylla ecaudata*), predation (shrews; *Blarina brevicauda*, *Crocidura canariensis*, *Neomys fodiens* and *N. anomalus*), and potentially defensive or antagonistic purposes (slow lorises; *Nycticebus* spp.) (8).

The Eulipotyphla, a group historically referred to as the insectivores, consists of the hedgehogs, moles, shrews and solenodons. Within this group, species from three separate genera of shrews (*Blarina*, *Neomys* and *Crocidura*) and the solenodons (*Solenodon paradoxus* and *Atopogale cubana*) exhibit convincing evidence of an oral venom system (9) (Fig. 1). Shrews utilise their venom for overpowering vertebrate prey much larger than they would otherwise be able to feed upon (e.g. similar mass to themselves), and for paralysing invertebrate prey for long-term storage purposes (‘prey caching’), presumably to provide a continual resource to help offset the extreme metabolic demands of these small animals (10–13). While convincing evidence of a venom system is lacking for hedgehogs (14) and moles, it has been proposed that the nesophontids (*Nesophontes* spp.), a recently extinct family of eulipotyphlans that were the sister group to the solenodons (15), may also have been venomous based on morphological evidence (16). All of the extant venomous eulipotyphlans produce venom in submaxillary glands, but their venom delivery apparatus varies, with solenodons using elaborate tubular lower incisors (Fig. 1) and shrews having little modified but pointed lower incisors and canines (9, 17). While this morphological variation might point towards independent origins of these venom systems, with the more extensive morphological adaptation in solenodons potentially indicating a longer evolutionary history or tighter ecological integration of venom use, it is worth noting that the venom delivering dentition of snakes also varies extensively, despite the common origin of their venom secretions and toxins (18). Thus, it remains unclear whether eulipotyphlan venom systems share a single early evolutionary origin, or whether multiple groups of shrews, the solenodons, and possibly other eulipotyphlans have each evolved venom independently following their divergence during the late Cretaceous period.

To address this fundamental question, we characterised the venom system of the Hispaniolan solenodon (*Solenodon paradoxus*). Solenodons, which are slot-toothed distant relatives of shrews, are relatively large (~1 kilogram) nocturnal eulipotyphlans found on the Caribbean islands of Hispaniola (*S. paradoxus*) and Cuba (*A. cubana*), and molecular and fossil evidence suggest that they diverged from all other mammals over 70 million years ago (MYA) (19) (Fig. 1). Both species have long been considered rare and threatened, and have experienced range declines associated with habitat loss and predation by invasive dogs and cats (20). Despite these enigmatic animals likely being the largest extant venomous terrestrial mammals, little is known about the composition, function and ecological role of their venom, other than its relatively weak toxicity to mice (17). Consequently, we sequenced the genome of the Hispaniolan solenodon and used this information to underpin identifications of the proteins present in its venom. We then characterised the function of solenodon venom via a range of *in vitro* and *in vivo* assays to determine the likely role of this adaptation. Our findings reveal that eulipotyphlan venom systems and their constitutive toxins have evolved on multiple independent occasions via the process of convergent evolution.

**Results and Discussion**

We constructed a genome for *S. paradoxus* from DNA isolated from blood collected from an adult male Hispaniolan solenodon from the northern Dominican Republic (*S. p. paradoxus*), housed in captivity in the Dominican Republic National Zoo (ZOODOM). DNA was sequenced using Illumina paired end short read technology, and the genome was assembled using DISCOVAR *de novo*. The resulting assembly had a scaffold N50 of 407.7 kb and performed well on BUSCO (Benchmarking Universal Single-Copy Orthologs (21)), with 92.9% complete and 4.7% partial BUSCOs recovered. The assembly is thus relatively higher quality than a recently published ‘consensus’ genome for *S. p. woodi* constructed using DNA from multiple individuals (22) (SI Appendix Table S1). Next, we annotated the repetitive and protein-coding portions of the genome using MAKER (23). Because RNAseq data was not available for this endangered species, our annotations were based on homology searches alone, which may be less effective for identifying highly divergent genes. Nonetheless, homology searches with existing protein databases and genome and RNAseq data from related eulipotyphlan species (hedgehog, mole, shrew) identified a comparably high number of protein encoding genes (18,112 *vs* 19,372-20,798), of which the vast majority exhibited orthology with those previously detected from other eulipotyphlans (97.4-98.0%), indicating that our approach was broadly effective.

Venom was collected from two wild male adult Hispaniolan solenodons (*S. p. woodi*) that were caught near Pedernales, southwestern Dominican Republic, and we also collected saliva from one of these individuals. Saliva was collected via direct pipetting from the back of the mouth prior to venom stimulation, while venom was collected by encouraging solenodons to chew onto soft plastic tubing and collecting the resulting secretions. Thus, saliva is unlikely to contain venom proteins, but venom may, perhaps, contain small amounts of salivary proteins. However, initial 1D SDS-PAGE gel electrophoretic analysis of these samples validated the collection approach, as distinct protein profiles were observed between the collected venom and saliva (Fig. 2A). We also observed highly similar venom compositions between the two sampled individuals (Fig. 2A), suggesting venom conservation. However, the small sample size and possibility of high genetic relatedness of these individuals means that future work is required to robustly explore venom variation in solenodons. For in-depth comparisons between venom and saliva, we applied three different mass spectrometry-based proteomics workflows; shotgun analyses of digested crude samples, bottom-up proteomic analyses of prefractionated (decomplexed) samples, and top-down proteomic analyses of reduced and non-reduced samples. In addition to orthogonal confirmation of the main venom components, the application of these three different approaches offer complementary merits such as higher sensitivity, optimal quantitative estimation of toxin abundance, and proteoform-resolved compositional information, respectively. For all approaches, venom proteins were identified by peptide/protein spectrum matching against the protein database derived from the assembled *S. paradoxus* genome.

Initial analysis via shotgun experiments revealed solenodon venom is primarily composed of proteins that exhibit high-scoring annotations to kallikrein-1-like serine proteases (*KLK1*-like; 7 of 17 total venom proteins identified), although various other protein types were also detected (Fig. 2B and 2C, SI Appendix Table S2). None of the venom proteins directly identified here show similarity to those recently predicted by others, who used genomic data alone to predict venom toxin identity based on sequence similarity to previously described, yet distinct, animal venom toxins (22). These findings highlight the importance of direct sampling (e.g. gene expression or protein) to robustly characterise proteins associated with venom secretions (24). The majority (ten) of the solenodon venom proteins detected were also identified in saliva, although solenodon saliva contained an additional 48 proteins with diverse functional annotations (Fig. 2B and 2C, SI Appendix Table S2). Next, we applied a validated venom decomplexation strategy that utilised HPLC fractionation followed by SDS-PAGE, in-gel trypsin digestion, and LC-MS/MS analysis (25, 26). This approach yielded 29 venom peaks (Fig. 2D), with peaks 1-19 containing molecules with masses below 3 kDa, peaks 20-25 showing masses of 10-15 kDa according to reductive SDS-PAGE, and peaks 27-29 showing two masses around 14 and 28 kDa. From these excised bands (14 and 28 kDa), we identified three distinct *KLK1*-like proteins (Fig. 2G, SI Appendix Table S2), and no other proteins, with the exception of keratin contaminants. We complemented these data with top-down analyses of crude reduced and non-reduced venom and saliva. In this experiment, the venom and saliva were not digested and were instead directly analysed by LC-MS/MS, which allows better comparison of homologous proteins and proteoforms that would otherwise be indistinguishable after trypsin digestion (27). According to UV peak area, the main protein observed in the native (non-reduced) venom was found in peak 28 with the monoisotopic mass 27589.64 Da and a retention time (RT) of 92.3 min, and which again corresponded to *KLK1* (Fig. 2F). This protein was not detected in the saliva from the in-gel digest (Fig. 2E and 2H). Several proteins in the same mass range (26952.59 Da, 76.4 min; 26486.43 Da, 77.9 min; 25430.46 Da, 81.6 min; 27589.48 Da, 91.83 min; 25130.54 Da, 107.5 min) were also mainly detected in the venom. Another high abundance *KLK1*-like isoform with a mass of 25130.40 Da and RT of 80.5 min was detected in both saliva and venom, although its relative abundance (normalised peak area) was around five-fold higher in the venom (Fig. 2E, 2F and 2H).

These proteomic data demonstrate that: (i) solenodon venom is relatively compositionally streamlined in comparison with saliva, (ii) venom consists predominately of *KLK1*-like proteins, and (iii) while some of these *KLK1*-like proteins are also found in solenodon saliva, they are of much higher abundance in venom. Kallikreins are members of the S1 group of serine proteases and likely originated in early tetrapods (28, 29). They are diverse in placental mammals, consisting of up to 15 paralogs, and they act by enzymatically cleaving peptide bonds (28, 29). Kallikreins can have diverse functions, includingcleaving kininogens and plasminogen, resulting in the liberation of kinins and plasmin, respectively (29). Here we demonstrate that solenodon venom exhibits activities consistent with the presence of secretions rich in kallikreins. Using substrate-specific kinetic biochemical assays, we find that solenodon venom exhibits serine protease activity and potently activates plasminogen (Fig. 3, SI Appendix Fig. S1). In both cases, solenodon venom showed significant increases in activity when compared with solenodon saliva, and also when compared with snake venoms known to exert serine protease and plasminogen activating activities (Fig. 3A and 3B) (30, 31). We demonstrated that multiple *KLK1*-likeproteins are responsible for the activation of plasminogen observed with solenodon venom via the use of a nanofractionation approach consisting of LC-MS, undertaken in parallel with a specific bioassay (Fig. 3C, SI Appendix Table S3). Both venom and saliva also demonstrated cleavage of high molecular weight kininogen (HMWK), with the venom being most potent, as it rapidly cleaved this substrate in the absence of pre-incubation, unlike saliva (Fig. 3D). While both venom and saliva were also found to cleave other substrates known to be targeted by serine proteases (e.g. fibrinogen) (SI Appendix Fig. S1), their higher potency to HMWK is consistent with the identification of *KLK1* in these samples. In combination, these *in vitro* bioactivity studies reveal that solenodon venom exhibits functional specificities consistent with the identification of kallikrein serine proteases as the most numerous and abundant proteins found in the venom.

Physiologically, the cleavage of kininogens by kallikreins results in the liberation of the kinins bradykinin and kallidin, which in turn stimulate hypotensive responses in vertebrates, via the kinin-kallikrein system (29). To test whether solenodon venom causes hypotension *in vivo*, we intravenously administered a sub-lethal dose of venom in PBS to mice (25 mg/kg; n=3) and compared their physiological responses with those of a control group receiving PBS only (n=3). Using a MouseOx pulse-oximeter cuff, we periodically monitored the pulse rate, respiration rate, and percentage oxygen content of the envenomed andcontrol animals but found no significant differences between the two groups (SI Appendix Fig. S2). However, measures of pulse distension - defined as local blood flow at the sensor location - showed a substantial, transient, reduction in envenomed animals compared to controls (47.5% maximal decrease from baseline), with recovery towards baseline levels occurring 30 mins after venom administration (Fig. 3E). These results suggest that solenodon venom exerts a hypotensive effect. To directly test this hypothesis, we assessed the bioactivity of solenodon venom in an *in vivo* cardiovascular assay. We found that solenodon venom (1 mg/kg; n=5) caused a marked depressor effect on the mean arterial pressure of anaesthetized rats, consisting of a transient depressor response and resulting in a maximal decrease of 22% (+/-6%) from baseline readings (Fig. 3F, SI Appendix Fig. S3).

Our findings demonstrate that *KLK1*-like proteins are the major functional components of solenodon venom. S1 serine proteases are common constituents of animal venoms, with diverse venomous taxa such as snakes, lizards, cephalopods and lepidopterans all utilising representatives of this large multilocus gene family as toxins via the process of convergent evolution (2). Reconstructing the molecular evolutionary history of tetrapod kallikreins (*KLK1*-*KLK15*) (Fig. 4A, SI Appendix Fig. S4) revealed that all of the annotated *KLK1*-like genes identified in the solenodon genome are indeed found nested within a strongly supported clade containing *KLK1*s from other mammals. Fascinatingly, this clade also includes proteins previously identified in the venom of the shrew *B. brevicauda* (Blarina toxin and blarinasin 1 and 2 (11, 32)) (Fig. 4A). However, the seven *KLK1*s we identified in solenodon venom (SI Appendix Table S2) formed a strongly supported monophyletic sub-cluster (Bayesian posterior probability: 1.00; bootstrap: 100), and included an additionalsolenodon *KLK1* isoform not identified by our proteomic analyses of venom (Fig. 4A). These findings strongly suggest that solenodon *KLK1* venom genes have arisen as the result of lineage-specific gene duplication events, rather than duplications occurring prior to the diversification of eulipotyphlans, thereby indicating independent venom-related diversifications in solenodons and shrews. To investigate this further, we performed sequence analyses of representative eulipotyphlan *KLK1*s. Prior work has suggested that a combination of multiple small insertions and alterations to the physicochemical patterns (hydropathicity and charge) of the five regulatory loops present in *KLK1*s are responsible for the increases in toxicity observed between Blarina toxin and the blarinasins (33). Here, we find small novel insertions in the regulatory loops of solenodonvenom *KLK1*s, although we find no consistent patterns of changes to the mean hydropathicity or charge of these regions when broadly comparing eulipotyphlan venom *KLK1*s with those identified from non-venomous taxa (hydropathicity, *p* = 0.18; charge, *p* = 0.20) (SI Appendix Fig. S5). However, comparisons of the locations of the regulatory loop insertions reveals a differential pattern between *Blarina* and the solenodon, with the former exhibiting insertions predominately in loop 1 and 2, and the latter in loop 3 (SI Appendix Fig. S5), thereby confirming that these toxins have evolved independently for a role in venom.

Next, we employed site-, branch-, and branch-site specific maximum likelihood and Bayesian models to assess the regime of natural selection influencing the evolution of the kallikrein gene family in tetrapods. Site-specific selection analyses (model 8, PAML) (34) revealed a significant influence of purifying selection on the evolution of all *KLK* paralogs. Each paralog was characterised by a very small omega (ω) value (mean *KLK* ω = 0.29), which represents the ratio of non-synonymous to synonymous substitutions, with the exception of *KLK1* (ω= 0.55) (SI Appendix Table S4 and Fig. S6). Our analyses identified 18 positively selected amino acid sites in *KLK1*, only two in *KLK10*, and none in any of the other kallikreins. These results suggest that while the majority of amino acid sites in *KLK1* remain extremely well-conserved, a number have experienced positive selection for amino acid replacements. When overlaying these positively selected sites onto the *KLK1* sequence alignment, we find that 14 of these 18 sites are found within the five regulatory loops (SI Appendix Fig. S5), which is consistent with the prior suggestion that modification of these regions may be important for venom toxin function (33). The findings of the site-specific selection analyses are further supported by Fast Unconstrained Bayesian AppRoximation (FUBAR) and Mixed Effects Model Evolution (MEME) analyses, which identified numerous *KLK1* sites evolving under the pervasive influence of purifying selection, but with only a small number evolving under pervasive or episodic positive selection (SI Appendix Table S4). To identify whether positive selection has shaped the evolution of venom *KLK1*s detected in eulipotyphlans, we employed branch- and branch-site specific maximum likelihood and Bayesian models. Together, these analyses revealed an increased influence of positive selection on the *KLK1* clade, in comparison with the other *KLK* paralogs (SI Appendix Tables S5-S7). The branch-site specific model identified 39 positively selected sites (pp ≥ 0.95) and computed a ω of 1.3 for this clade (SI Appendix Table S5). Interestingly, 4 out of the 13 foreground branches that were identified to have undergone episodic positive selection (p ≤ 0.05) were *KLK1* genes identified from the genome of *S. paradoxus* (SI Appendix Table S7). In combination, our findings suggest that at least four of the eight solenodon *KLK1* genes (three of the seven *KLK1*s detected in venom) exhibit evidence of evolving under the influence of episodic positive selection. Thus, solenodon venom genes have evolved via the process of gene duplication coupled, in some cases, with episodic positive selection – a phenomenon that is consistent with the evolutionary histories of a number of toxin families found in other venomous animal lineages (1, 35).

Analysis of the genomic organisation of kallikreins provides additional support for multiple independent origins of venom in eulipotyphlans. Mammalian kallikreins are found in a tandem array of linked genes, and while most exist as single copy orthologs, *KLK1* and its upstream flanking gene *KLK15* show a more variable pattern (Fig. 4B). We found between one and three intact paralogs of these genes present in different mammals, and they are often interspersed by *KLK* pseudogenes or exon fragments (e.g. *Homo sapiens* has three *KLK1*-like genes, annotated as *KLK1*, *KLK2* and *KLK3*, and one *KLK15* gene, with remnants of at least one other *KLK15*). Within Eulipotyphla, the representative mole (*Condylura cristata*), hedgehog (*Erinaceus europaeus*) and shrew (*Sorex araneus*) species analysed, all of which are non-venomous, were found to have two or three *KLK1* paralogs and a single *KLK15* gene, though the organisation and orientation of these genes varied among species, suggesting independent evolutionary histories (Fig. 4B). Contrastingly, the solenodon genome revealed the presence of at least eight *KLK15* and eight *KLK1* paralogs, of which we detected seven of the *KLK1*s proteomically in venom. While the contiguity of the solenodon genome is insufficient to perform synteny analysis, we note that multiple scaffolds contain *KLK1* and *KLK15* genes adjacent to one another (Fig. 4B), suggesting that the process giving rise to the extensive number of paralogs uniquely observed in this species may involve a duplication consisting of at least one of each of these genes. The combined findings from our molecular evolution and synteny analyses provide convincing evidence that solenodons have evolved multiple *KLK1* genes for use in their venom system, and that both solenodons and shrews have independently utilised *KLK1*s for a role in venom. Future work is required to assess whether *KLK1*s show similar evolutionary trajectories in other venomous eulipotyphlans (e.g. *Neomys* and *Crocidura* shrews), as comparative molecular data are currently unavailable for those species.

Next, we sought to infer the timing of the origin of venom in the Order Eulipotyphla. To do so, we used ancestral trait reconstructions to reconstruct the character state for venom across this group. The resulting posterior probabilities (all >0.85) provided strong support for four independent origins of venom in this group: in solenodons, *Blarina* shrews, *Neomys* shrews and *Crocidura* shrews (Fig. 4C, SI Appendix Fig. S7). The unlikely alternative hypothesis of an early evolution of venom followed by the loss of this character state in multiple taxa required a single gain and at least nine loss events, as *Blarina*, *Neomys* and *Crocidura* are not closely related to one another within the Soricidae (last common ancestor ~16-20 MYA; (36)) (Fig. 4C, SI Appendix Fig. S7). While further research effort could change this interpretation - for example the future identification of additional shrew species or other eulipotyphlan families as venomous - the combination of diverse data types described above strongly suggest that both solenodons and shrews, which diverged from one another over 70 MYA (19), have independently evolved oral venom systems. Moreover, both these groups have independently recruited *KLK1*s for a role in venom, and thus provides a fascinating example of convergent molecular evolution. In this instance, molecular convergence seems likely to be underpinned by pre-adaptations, as bothsolenodons and shrews have evolved an oral venom system via the modification of submaxillary salivary glands (37, 38), and *KLK1* has previously been demonstrated to be an abundant component found in the salivary glands of a variety of mammals (SI Appendix Fig. S8). Thus, *KLK1* likely existed as an abundant starting substrate in the oral secretions of ancestral eulipotyphlans before being independently selected for increased expression and diversification for use in the venom systems of multiple different eulipotyphlan groups. Therefore, solenodons and at least some shrews have achieved the same molecular solution for the composition of their venom, despite employing different morphological strategies to deliver those molecules (e.g. elaborate tooth grooves *vs* rapid biting with pointed incisors and canines). While the venom delivery systems of many other venomous mammals (e.g. platypus and slow lorises) are distinct from the solely oral systems of eulipotyphlans, haematophagus vampire bats (e.g. *Desmodus rotundus*) also deliver venom produced in submaxillary glands via sharp incisors. Notably, *KLK1*-like proteins have previously been detected in their venom (39), alongside other serine proteases that activate plasminogen (40), thereby representing an intriguing example of molecular and functional venom convergence with their distant eulipotyphlan relatives (last common ancestor ~87 MYA) (41).

Prior research has been unable to determine the ecological role of solenodon venom. It has previously been speculated that venom might facilitate prey capture, be a relictual trait, or be used for intra-specific competition or anti-predator defence (8, 9, 17). The use of a hypotensive venom for defence would be unusual (although not unique (42)), as most defensive venoms cause acute pain to act as both an immediate deterrent and to invoke learned avoidance behaviour (43). However, solenodon bites inflicted on humans do not tend to result in such extensive pain, with inflammatory responses and secondary infections likely responsible for much of the resulting pathology (17, 44). Importantly, the insular Caribbean contained no native terrestrial mammalian predators before the Mid-Holocene arrival of humans, who first introduced dogs, and then later cats and mongooses (45), suggesting that the evolution of a defensive venom is unlikely to be related to defence against predators. Although solenodons are known to be predated by owls and possibly other raptors (46), and co-existed pre-historically with giant Caribbean raptors that are now extinct (47), orally-delivered venom seems unlikely to protect them from the talons of such avian predators.

There is also little evidence supporting the premise that solenodons use their venom for intraspecific purposes, such as for competition during breeding seasons (as in the platypus), or for resolving territorial disputes. Although some captivity case reports suggest that solenodons may have been killed following bites by other solenodon individuals (17), most captive accounts describe antagonistic encounters among solenodons being resolved without biting (48). Moreover, solenodons are relatively social animals; both species live in family groups comprising adults, subadults, and young, with multiple family groups of Cuban solenodons sharing the same den (44, 46, 49, 50). Although a lack of natural history reports documenting the behaviour of these poorly-known mammals limits our interpretation, we find no convincing evidence supporting the hypothesis for venom having evolved for an intra-specific purpose.

It appears most likely that the solenodon venom system evolved for capturing prey, in a manner analogous to, and in parallel with, venomous shrews. This hypothesis is supported by the convergent evolution of similar venom components (*KLK1s*) found in the solenodon and *Blarina* venom systems. However, *Blarina* shrews have a bipartite venom, consisting of both *KLK1*-like proteins that act on small vertebrates (11) and potent neuropeptides for the immobilisation of invertebrates for long-term prey storage (51, 52). Although their feeding and hunting behaviour is poorly understood, solenodons do not appear to ‘cache’ their prey in this manner (48). Nonetheless, to test for the potential presence of neurotoxic venom activity, we assessed the activity of solenodon venom on nicotinic acetylcholine receptors (nAChR) and voltage gated sodium channels (Nav), both of which are ion channels commonly targeted by venoms to cause immobilisation via neuromuscular paralysis (1, 2). Solenodon venom exhibited no activity on either human muscle-type or locust nAChR at concentrations up to 50 µg/mL (Fig. 3G), but did display subtle, but significant, inhibitory activity at mammalian voltage gated sodium channels in whole-cell patch-clamp assays (SI Appendix Fig. S9 andTable S8). While the observed reduction in peak current amplitude, as well as the inhibitory shifts in threshold for activation and inactivation of NaV currents, could lead to paralysis, this was not apparent in *in vivo*toxicity studies using locusts (*Schistocerca gregaria*, n=4) and centipedes (*Ethmostigmus rubripes*, n=5); invertebrates representative of general groups that solenodons may consume in the wild (50). All animals survived 24 hr after the intrathoracic injection of venom (up to 50 µg/g and 100 µg/g, respectively) with no signs of immobilisation or incapacitation; it is therefore possible that the observed activity on Nav channels may be specific to vertebrates. Irrespectively, we conclude from these data that solenodon venom is unlikely to function to paralyse invertebrates for prey storage, unlike that of *Blarina* shrews.

Kallikrein toxins present in *Blarina* shrew venom likely facilitate the capture of vertebrate prey, specifically by enabling relatively large-bodied prey to be subdued, which would otherwise be capable of escaping or defending themselves (11). To investigate the extent to which *S. paradoxus* feeds on vertebrate prey items, we assessed their dietary composition using a vertebrate DNA barcoding approach on 64 faecal samples collected during the dry and rainy seasons in Pedernales Province, Dominican Republic. Our conservative ‘frequency of occurrence’ approach detected evidence of vertebrate prey in 12.3% of the faecal samples analysed (Fig. 1), demonstrating that while solenodons feed predominately on invertebrates (53), vertebrate prey make up a considerable proportion of their diet. These findings are in line with a number of natural history reports describing solenodons feeding on amphibians, reptiles, and occasionally nestling and ground-dwelling birds (44, 46, 49, 50). Although many vertebrate species remain as potential solenodon prey items in Cuba and Hispaniola, it is also possible that the solenodon venom system initially evolved for capturing vertebrate prey, and is now at least partly relictual due to the extinction of considerable amounts of regionally endemic vertebrate biodiversity (e.g. lizards, birds and mammals) from their habitats. In particular, all of the smallest-bodied native Caribbean land mammals (nesophontids, spiny rats, small hutias), which might have ancestrally constituted solenodon prey species, became extinct following European arrival in the insular Caribbean ~500 years ago (45). Indeed, our sample site exhibits considerable evidence of anthropogenic impacts (e.g. mosaic agriculture-forest habitat), and thus the diet reported here may heavily reflect flexible responses to anthropogenic resources and the prey types currently available to solenodons. Based on venom composition, dietary data and limited available natural history reports, we conclude that the solenodon venom system likely evolved for overpowering and subduing occasional (and perhaps previously more abundant) vertebrate prey items.

Our findings collectively highlight the unique significance of eulipotyphlans within mammals for having multiple origins of venom and venom delivery systems. Venom has evolved more times in this group than in all other mammals combined, and on more instances than those found within any other class of vertebrates, other than bony fish. The reason for the frequency of venom evolution in eulipotyphlans remains unclear, but considering the majority of these venom origins relate to shrews, a group of animals well known for having high metabolic rates that require frequent feeding (10), venom may be a valuable adaptation that facilitates their near-continual foraging lifestyle. In solenodons, venom also appears to facilitate prey capture, but additional work is required to fully elucidate the nature of this venom, such as (i) compiling extensive natural history observations of foraging behaviour, (ii) testing solenodon venom on natural prey items, (iii) comparisons of venom composition and function between the two solenodon species, and between male and female individuals, and (iv) further investigation of solenodon dental and mandibular morphology to understand their prey handling capability. Nonetheless, our findings highlight the evolutionary novelty of the solenodon venom system, and stress the importance of studying and conserving endangered species in order to protect both ecological diversity and their utilitarian value, which in this case is most relevant when considering the bioactive compounds found in their toxic secretions (7, 54). Ultimately, our work reveals a surprising case of convergent molecular evolution, whereby *KLK1*s have been independently recruited for use in the non-homologous venom systems of shrews and solenodons. These findings highlight that the molecular constituents of eulipotyphlan venom systems may be subjected to constraints that limit the options available for the evolution of venom, as these lineages have highly divergent phylogenetic backgrounds and different morphological adaptations for delivering these molecules. Our findings therefore emphasise that distinct structural phenotypes, encapsulated by variation in venom delivery systems, can yield equivalent functions, and more generally, they reinforce the broad value of studying natural toxin systems to elucidate fundamental evolutionary processes.

**Materials and Methods**

Detailed materials and methods can be found in the SI Appendix.

**Genomics.** The *S. paradoxus* genome was constructed usinghigh molecular weight DNA isolated from the blood of a captive male adult individual. Paired-end library sequencing was performed on an Illumina HiSeq2500 instrument with 250 bp reads. The assembly was undertaken using DISCOVAR *de novo*, before annotation with MAKER (23), with Augustus (55) implemented to facilitate gene prediction. We then used BUSCO (21) to individually assess the quality of the genome assembly and annotation.

**Proteomics.** Venom and saliva samples were collected from two wild-caught, adult, male Hispaniolan solenodons. We used reduced SDS-PAGE gel electrophoresis for initial visualisation of venom and saliva proteins (10 µg). Shotgun proteomics was performed by digesting 5 µg of each sample with trypsin, before analysis by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Decomplexed bottom-up proteomics was performed as previously described (26). Samples (1 mg) were separated via reverse-phase HPLC, then reduced and analysed by SDS-PAGE, and protein bands subjected to in-gel tryptic digestion and analysed by LC-MS/MS using an Orbitrap XL (Agilent, Germany). For top-down proteomics we used 0.2 mg of venom and saliva for reduced and non-reduced HPLC-high-resolution (HR) MS/MS measurements. Top-down LC-electrospray ionization-HR-MS experiments were performed on an LTQ Orbitrap XL (Agilent, Germany) in data-dependent acquisition mode. For all proteomic experiments, resulting MS2 spectra were matched against translations of the protein encoding genes predicted from thegenome. Full details are displayed in SI Appendix File S1.

**Evolutionary analyses.** The 26 *KLK* genes identified in the solenodon genome were used as queries for BLAST searching the NCBI non-redundant and Ensembl tetrapod genomes databases. Resulting nucleotide and amino acid sequences were aligned (SI Appendix Files S2 and S3) and subjected to Bayesian inference (2x108 generations, parallel runs=4, MCMC simulations=6) and the nucleotides also to maximum likelihood (subtree pruning and regrafting method, 100 bootstrapping replicates) analyses. To investigate the influence of selection on tetrapod *KLK*s we employed (i) site- branch-, and branch-site maximum likelihood models implemented in CodeML of the PAML package (56), (ii) the Mixed Effect Model of Evolution (MEME) (57), (iii) Fast Unconstrained Bayesian AppRoximation (FUBAR) analyses (58), and (iv) the Adaptive Branch-Site Random Effects Likelihood (aBSREL) approach (59). Three-dimensional homology models were generated for the various tetrapod *KLK*s using the Phyre2 server (60) and PyMOL (PyMOL Molecular Graphics System, Schrödinger, LLC) was used for visualisation. Analyses of hydropathicity and charge of *KLK1* regulatory loops (33) were calculated using the ProtParam tool of the ExPASy Bioinformatics Resource Portal, with statistical comparisons performed using unpaired two-tailed *t*-tests in Graphpad Prism (La Jolla, USA). *De novo* annotation of *KLK* exons and synteny comparisons of mammalian genomic data were undertaken as recently described (61). Ancestral trait reconstructions were performed with Ape (62) and Phytools (63) in R, and the marginal ancestral states (empirical Bayesian posterior probabilities) estimated for each node in a eulipotyphlan species tree derived from prior studies (19, 36). A stochastic character mapping analysis (64) was performed for 1,000 simulations, and a trait density map generated to depict the posterior probabilities of states across the tree.

***In vitro* venom function.** Degradation gel electrophoresis experiments were performed using 5 µg of substrate (HMW kininogen or fibrinogen) and 5 µg of venom or saliva. Samples were either incubated at 37°C for 60 mins or loaded directly onto SDS-PAGE gels for electrophoretic separation under reducing conditions. For serine protease activity, we used a chromogenic assay (n=3 independent repeats) with the specific substrate S-2288 (Cambridge Biosciences). Samples (1 µg) were plated in triplicate into 384-well plates, overlaid with Tris buffer (100 mM Tris, 100 mM NaCl, pH 8.5), and 6 mM of S-2288 and absorbances measured at 405 nm kinetically. Control (PBS) readings were subtracted and the rate of substrate consumption calculated by measuring the slope between 0 and 5 mins. To monitor plasminogen cleavage we used a modified kinetic assay (n=3 independent repeats) (65) to detect the resulting plasmin activity via cleavage of the H-D-Val-Leu-Lys-AMC fluorescent substrate (I-1390, Bachem). One microgram of venom/saliva was prepared in 10 µl of assay buffer (100 mM Tris-HCl pH 7.5, 0.1% BSA), plated in triplicate into 384-well plates, and overlaid with 50 µl/well of 200ng/ml plasminogen (Sigma-Aldrich) and 5 µM of substrate. Fluorescence were measured kinetically for 45 mins (excitation 355 nm, emission 460 nm) and areas under the curves calculated for the 0-30 min interval. Statistical analyses (unpaired two-tailed *t-*tests) for the kinetic assays were performed in GraphPad Prism (La Jolla, USA). To identify the toxins responsible for plasminogen activating activity, we followed the previously described approach (65), whereby venom (250 µg) was fractionated by LC in parallel with at-line nanofractionation and subsequent identification of bioactives from the plasmin assay identified via nanoLC-MS/MS analysis of tryptic digests of the corresponding fractions. Details of the resulting peptide spectrum matching are displayed in SI Appendix File S4. For patch-clamp electrophysiology experiments, we used: (i) TE671 human rhabdomyosarcoma cells endogenously expressing embryonic muscle-type nicotinic acetylcholine receptors (nAChR) and Nav1.7 voltage gated sodium channels (VGSC), and (ii) locust (*Schistocerca gregaria*) primary neurons (natively expressing insect neuronal nAChR), the latter of which were dissected from the mushroom bodies of 6th instar locusts. Patch pipettes (resistance 5-7 MΩ) were filled with a caesium pipette solution (140 mM CsCl, 10 mM NaCl 1 mM MgCl2, 11 mM EGTA and 5 mM HEPES, pH 7.2 with CsOH) and the bath solution for TE671 cells was 135 mM NaCl, 5.4 mM KCl, 1 mM CaCl2, 1 mM MgCl2, 5 mM HEPES and 10 mM D-glucose (pH 7.4 with NaOH), and for locust neurons was 180 mM NaCl, 10 mM KCl, 2 mM CaCl2, 10 mM HEPES, pH 7.2. Whole-cell currents were monitored using an Axopatch 200A (Axon instruments, USA) patch-clamp amplifier and venom and agonist were applied to cells using a DAD-12 Superfusion system (Adams and List Associates, USA). The series resistance was compensated by 75% to minimize any voltage errors, and data were filtered at 10 kHz.

***In vivo* venom function.** Groups offifth instar desert locusts (*Schistocerca gregaria*, n=4) and juvenile giant centipedes (*Ethmostigmus rubripes*, n=5) were injected intrathoracically with various solenodon venom doses (0.1-100 µg/g) alongside controls (insect Ringer’s saline) and their status (alive, incapacitated, dead) was monitored for at least 24 hrs. Comparisons of the physiological responses of mice (20g, male CD1, Charles River) dosed intravenously with solenodon venom (25 µg/g, n=3) and PBS (control, pH 7.2, n=3) were undertaken using a MouseOx pulse-oximeter monitoring system (MouseOx, Harvard Apparatus). Measurements were collected at five different times for each animal (baseline, 1 min, 15 mins, 30 mins and 45 mins post-administration). We examined the effect of solenodon venom on the blood pressure of anaesthetised rats (male Sprague-Dawley, 250-320 g) by connecting a carotid artery cannula to a PowerLab/400 system via a Gould Statham P23 pressure transducer. Blood pressure was allowed to stabilize for at least 10 min before venom (1 mg/kg; n=5) was administered through the jugular vein and flushed with saline (0.2 ml).

**Dietary analyses.** DNA was isolated from 64 samples collected opportunistically from fresh *S. paradoxus* faecal samples (Pedernales Province, Dominican Republic; dry season, 40; wet season, 24) and probed for the presence of vertebrate prey DNA using primers specific to 12S and 16S ribosomal genes. Resulting DNA was sequenced on an Ilumina MiSeq instrument and operational taxonomic units identified via BLAST searches of the GenBank non-redundant nucleotide and the Barcode of Life Data Systems (BOLD) databases, using 85% identity thresholds. To calculate the frequency of occurrence of vertebrate prey, we summed the presence of each identified food item across all 64 samples and divided that figure by the total number of faecal samples.

**Permissions.** Ethical permission, collection permits and export permits for solenodon samples were granted by the Dominican Republic National Zoo (ZOODOM) and the Ministerio de Medio Ambiente y Recursos Naturales (#2577 and #VAPB-02368). The murine *in vivo* study was conducted using protocols approved by the Animal Welfare and Ethical Review Boards of the Liverpool School of Tropical Medicine and the University of Liverpool, and performed under licenced approval of the UK Home Office, in accordance with the Animal [Scientific Procedures] Act 1986 (UK) and institutional guidance on animal care. The rat *in vivo* study was approved by the Monash Animal Research platform (MARP) Animal Ethics Committee, Monash University, Australia (#MARP/2017/147).

**Data deposition.** The genome assembly has been deposited to NCBI (<https://www.ncbi.nlm.nih.gov>) with the accession number RJWH00000000, the BioProject identifier PRJNA399371 and the BioSample identifier SAMN07678062. A dedicated website for the *Solenodon paradoxus* assembly and genome annotation has also been established (https://darencard.github.io/solPar), and repeat and gene annotation files have also been deposited to Figshare (https://doi.org/10.6084/m9.figshare.7640456).The proteomic data has been deposited to the ProteomeXchange database (<http://proteomecentral.proteomexchange.org>) with the accession number PXD009593 via the Mass Spectrometry Interactive Virtual Environment (MassIVE, <https://massive.ucsd.edu/>) with the accession number MSV000082307.

**Acknowledgments**

The authors thank Nicolas Corona, Yimell Corona, Luis Freites, Lisa Casewell, Darren Cook, Diane Genereux, Eva Murén, Voichita Marinescu, Ann-Maree, Kristine Bohmann, Martin Nielsen, Elizabeth Hadly, and the Ministerio de Estado de Medio Ambiente y Recursos Naturales (Dominican Republic). This study was supported by the Liverpool School of Tropical Medicine’s Jean Clayton Fund (to N.R.C.), a Sir Henry Dale Fellowship (200517/Z/16/Z to N.R.C.) jointly funded by the Wellcome Trust and the Royal Society, the Deutsche Forschungsgemeinschaft through the Cluster of Excellence Unifying Concepts in [Catalysis](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/catalysis) (UniCat) (to R.D.S.) and Grant PE 2600/1 (to D.P.), a Stanford Center for Computational, Evolutionary, and Human Genomics trainee grant (to A.M.M.), a Distinguished Professor grant from the Swedish Research Council (to K.L-T.), the DBT-IISc Partnership Program (to K.S.), a Department of Science and Technology (DST) INSPIRE Faculty Award (DST/INSPIRE/04/2017/000071 to K.S.), the Australian Research Council (DECRA Fellowship DE160101142 to E.A.B.U.) and the Royal Society (grants RG100902 and UF130573 to S.T.T.).

**Author Contributions**

N.R.C., I.B. and S.T.T conceived the project and designed the research; N.R.C., A.M.M., R.J.K., J.L.B and S.T.T. collected samples; N.R.C., G.W., J.A., J.J., E.K.K. and K.L-T. constructed the genome; D.C.C. and T.A.C. annotated the genome; D.P., J.S., B-F.H., R.D.S., J.K. and E.A.B.U. performed proteomics; N.R.C., V.S., I.K., F.J.V., I.B. and K.S. performed evolutionary and sequence analyses; N.R.C., D.R., L-O.A., J.S., N.M.N., J.D., I.R.M., B.G.F., and J.K. performed *in vitro* functional assays; N.R.C., D.R., R.A.H., F.M.S.B., S.K., W.C.H. and E.A.B.U. performed *in vivo* functional assays; A.M.M. performed dietary analyses. N.R.C. wrote the paper with assistance from D.P., D.C.C., A.M.M., K.S., E.A.B.U. and S.T.T. and input from all other authors.

**References**

1. Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG (2013) Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol* 28(4):219–29.

2. Fry BG, et al. (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet* 10:483–511.

3. Duda TF, Palumbi SR (1999) Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc Natl Acad Sci U S A* 96(12):6820–3.

4. Dowell NL, et al. (2016) The deep origin and recent loss of venom toxin genes in rattlesnakes. *Curr Biol* 26(18):2434–45.

5. Brust A, et al. (2013) Differential evolution and neofunctionalization of snake venom metalloprotease domains. *Mol Cell Proteomics* 12(3):651-63.

6. Gutiérrez JM, et al. (2017) Snakebite envenoming. *Nat Rev Dis Prim* 3:17063.

7. Clark GC, et al. (2019) Friends or foes? Emerging impacts of biological toxins. *Trends Biochem Sci* 44(4):365-79.

8. Rode-Margono JE, Nekaris KA-I (2015) Cabinet of curiosities: venom systems and their ecological function in mammals, with a focus on primates. *Toxins (Basel)* 7(7):2639–58.

9. Folinsbee KE (2013) Evolution of venom across extant and extinct eulipotyphlans. *Comptes Rendus Palevol* 12(7–8):531–42.

10. Deavers DR, Hudson JW (1979) Water metabolism and estimated field water Budgets in two rodents (*Clethrionomys gapperi* and *Peromyscus leucopus*) and an insectivore (*Blarina brevicauda*) inhabiting the same mesic environment. *Physiol Zool* 52(2):137–52.

11. Kita M, et al. (2004) Blarina toxin, a mammalian lethal venom from the short-tailed shrew *Blarina brevicauda*: Isolation and characterization. *Proc Natl Acad Sci U S A* 101(20):7542–7.

12. Kowalski K, Rychlik L (2018) The role of venom in the hunting and hoarding of prey differing in body size by the Eurasian water shrew, *Neomys fodiens*. *J Mammal* 99(2):351–62.

13. Lopez-Jurado LF, Mateo JA (1996) Evidence of venom in the Canarian shrew (*Crocidura canariensis*): immobilizing effects on the Atlantic lizard (*Gallotia atlantica*). *J Zool* 239(2):394–5.

14. Mebs D (1999) Studies on biological and enzymatic activities of salivary glands from the European hedgehog (*Erinaceus europaeus*). *Toxicon* 37(11):1635–8.

15. Brace S, et al. (2016) Evolutionary history of the Nesophontidae, the last unplaced recent mammal family. *Mol Biol Evol* 33(12):3095–103.

16. Turvey ST (2010) Evolution of non-homologous venom delivery systems in West Indian insectivores? *J Vertebr Paleontol* 30(4):1294–9.

17. Rabb G (1959) Toxic salivary glands in the primitive insectivore Solenodon. *Nat Hist Misc* 170:1–3.

18. Fry BG, et al. (2008) Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol Cell Proteomics* 7(2):215–46.

19. Springer MS, Murphy WJ, Roca AL (2018) Appropriate fossil calibrations and tree constraints uphold the Mesozoic divergence of solenodons from other extant mammals. *Mol Phylogenet Evol* 121:158–65.

20. Turvey ST, Kennerley RJ, Nuñez-Miño JM, Young RP (2017) The Last Survivors: current status and conservation of the non-volant land mammals of the insular Caribbean. *J Mammal* 98(4):918–36.

21. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31(19):3210–12.

22. Grigorev K, et al. (2018) Innovative assembly strategy contributes to understanding the evolution and conservation genetics of the endangered *Solenodon paradoxus* from the island of Hispaniola. *Gigascience* 7(6).

23. Campbell MS, Holt C, Moore B, Yandell M (2014) Genome Annotation and Curation Using MAKER and MAKER-P. *Current Protocols in Bioinformatics* (John Wiley & Sons, Inc., Hoboken, NJ, USA), pp 4.11.1-4.11.39.

24. Smith JJ, Undheim EAB (2018) True lies: using proteomics to assess the accuracy of transcriptome-based venomics in centipedes uncovers false positives and reveals startling intraspecific variation in *Scolopendra subspinipes*. *Toxins (Basel)* 10(3):96.

25. Göçmen B, Heiss P, Petras D, Nalbantsoy A, Süssmuth RD (2015) Mass spectrometry guided venom profiling and bioactivity screening of the Anatolian Meadow Viper, *Vipera anatolica*. *Toxicon* 107:163–74.

26. Nalbantsoy A, et al. (2017) Combined venom profiling and cytotoxicity screening of the Radde’s mountain viper (*Montivipera raddei*) and Mount Bulgar Viper (*Montivipera bulgardaghica*) with potent cytotoxicity against human A549 lung carcinoma cells. *Toxicon* 135:71–83.

27. Petras D, Heiss P, Süssmuth RD, Calvete JJ (2015) Venom proteomics of Indonesian king cobra, *Ophiophagus hannah*: integrating top-down and bottom-up approaches. *J Proteome Res* 14(6):2539–56.

28. Koumandou VL, Scorilas A (2013) Evolution of the plasma and tissue kallikreins, and their alternative splicing isoforms. *PLoS One* 8(7):e68074.

29. Lawrence MG, Lai J, Clements JA (2010) Kallikreins on steroids: structure, function, and hormonal regulation of prostate-specific antigen and the extended kallikrein locus. *Endocr Rev* 31(4):407–46.

30. Zhang Y, Wisner A, Xiong Y, Bon C (1995) A novel plasminogen activator from snake venom. Purification, characterization, and molecular cloning. *J Biol Chem* 270(17):10246–55.

31. Sanchez EF, et al. (2000) Isolation of a proteinase with plasminogen-activating activity from *Lachesis muta muta* (Bushmaster) snake venom. *Arch Biochem Biophys* 378(1):131–41.

32. Kita M, et al. (2005) Purification and characterisation of blarinasin, a new tissue kallikrein-like protease from the short-tailed shrew *Blarina brevicauda*: comparative studies with blarina toxin. *Biol Chem* 386(2):177–82.

33. Aminetzach YT, Srouji JR, Kong CY, Hoekstra HE (2009) Convergent evolution of novel protein function in shrew and lizard venom. *Curr Biol* 19(22):1925–31.

34. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–91.

35. Sunagar K, et al. (2015) The rise and fall of an evolutionary innovation: contrasting strategies of venom evolution in ancient and young animals. *PLOS Genet* 11(10):e1005596.

36. Dubey S, Salamin N, Ohdachi SD, Barrière P, Vogel P (2007) Molecular phylogenetics of shrews (Mammalia: Soricidae) reveal timing of transcontinental colonizations. *Mol Phylogenet Evol* 44(1):126–37.

37. Pournelle G (1968) Venomous animals and their venoms. *Venomous Animals and Their Venoms*, eds Bucherl W, Buckley E, Deulofen V (Academic Press, New York), pp 31–42.

38. Dufton MJ (1992) Venomous mammals. *Pharmacol Ther* 53(2):199–215.

39. Low DHW, et al. (2013) Dracula’s children: molecular evolution of vampire bat venom. *J Proteomics* 89:95–111.

40. Gardell SJ, et al. (1989) Isolation, characterization, and cDNA cloning of a vampire bat salivary plasminogen activator. *J Biol Chem* 264(30):17947–52.

41. Murphy WJ, Pringle TH, Crider TA, Springer MS, Miller W (2007) Using genomic data to unravel the root of the placental mammal phylogeny. *Genome Res* 17(4):413–21.

42. Casewell NR, et al. (2017) The evolution of fangs, venom, and mimicry systems in blenny fishes. *Curr Biol* 27(8):1184–91.

43. Trim SA, Trim CM (2013) Venom: the sharp end of pain therapeutics. *Br J Pain* 7(4):179–88.

44. Silva Taboada G, Suárez Duque W, Díaz Franco S (2007) *Compendio de los mamíferos terrestres autóctonos de Cuba vivientes y extinguidos.* (Museo Nacional de Historia Natural, La Habana, Cuba).

45. Cooke SB, Dávalos LM, Mychajliw AM, Turvey ST, Upham NS (2017) Anthropogenic extinction dominates Holocene declines of West Indian mammals. *Annu Rev Ecol Evol Syst* 48(1):301–27.

46. Borroto-Páez R, Mancina CA (2011) *Mamíferos en Cuba* (UPC Print, Vaasa, Finland).

47. Suárez W, Olson SL (2015) Systematics and distribution of the giant fossil barn owls of the West Indies (Aves: Strigiformes: Tytonidae). *Zootaxa* 4020(3):533.

48. Eisenberg J, Gould E (1966) The behaviour of *Solenodon paradoxus* in captivity with comments on the behaviour of other Insectivora. *Zool NY* 51:49–58.

49. Varona LS (1983) Remarks on the biology and zoogeography of the *Solenodon* (*Atopogale*) *cubanus*. *Bijdr Dierk* 53:93–8.

50. Ottenwalder JA (1991) The systematics, biology and conservation of solenodon. Dissertation (Univeristy of Florida, Gainsville, Florida, USA).

51. Martin IG (1981) Venom of the short-tailed shrew (*Blarina brevicauda*) as an insect immobilizing agent. *J Mammal* 62(1):189–92.

52. Bowen C V, et al. (2013) In vivo detection of human TRPV6-rich tumors with anti-cancer peptides derived from soricidin. *PLoS One* 8(3):e58866.

53. Mychajliw AM (2017) From the Pleistocene to the Anthropocene: fossils, genes, and the future of Caribbean mammals. Ph.D Dissertation (Stanford University, USA).

54. Pearson RG (2016) Reasons to conserve nature. *Trends Ecol Evol* 31(5):366–71.

55. Stanke M, Diekhans M, Baertsch R, Haussler D (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24(5):637–44.

56. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24(8):1586–91.

57. Murrell B, et al. (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8(7):e1002764.

58. Murrell B, et al. (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30(5):1196–205.

59. Smith MD, et al. (2015) Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* 32(5):1342–53.

60. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10(6):845–58.

61. Koludarov I, Aird SD (2019) Snake venom NAD glycohydrolases: Primary structures, genomic location, and gene structure. *PeerJ* 7:e6154.

62. Paradis E, Schliep K (2018) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35(3):526-8.

63. Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3(2):217–23.

64. Huelsenbeck JP, Nielsen R, Bollback JP (2003) Stochastic mapping of morphological characters. *Syst Biol* 52(2):131–58.

65. Zietek BM, et al. (2018) Liquid chromatographic nanofractionation with parallel mass spectrometric detection for the screening of plasmin inhibitors and (metallo)proteinases in snake venoms. *Anal Bioanal Chem* 410(23):5751–63.

**Figure Legends**

**Fig. 1.** **The two competing hypotheses relating to the origin of venom in eulipotyphlans and key characteristics of the Hispaniolan solenodon (*Solenodon paradoxus*).** The schematic phylogeny (grey lines) highlights the estimated divergence times of eulipotyphlan families and solenodon species, and the two competing hypotheses relating to the origin of venom in this group. Purple lines indicate the early origin of venom hypothesis followed by losses in moles, hedgehogs and some shrews (narrowing line), whereas red lines indicate the alternative hypothesis of multiple independent origins of venom in shrews (three times), solenodons, and possibly nesophontids. Shaded areas of the map indicate the modern distribution of the two solenodon species (*Atopogale cubana* and *S. paradoxus*) on the islands of Cuba and Hispaniola, respectively. Boxes on the right show (from top to bottom): a wild specimen of *S. paradoxus*, its lower jaw morphology, its enlarged tubular lower second incisor used for venom delivery visualised via stereo microscopy, and the composition and frequency of occurrence (in percentages) of vertebrates detected in their diet determined by DNA barcoding analyses of faecal samples. Divergence times displayed on the phylogeny are from (15, 19). The photograph of the wild solenodon is courtesy of Rocio Pozo.

**Fig. 2.** **Proteomic analyses of Hispaniolan solenodon (*S. paradoxus*) venom and saliva reveal *KLK1* proteins as major venom components.** **A)** Reduced SDS-PAGE gel electrophoretic profiles of venom and saliva samples. **B)** Gene ontology (GO) term analysis of proteins identified via shotgun proteomic-based annotation to the genome. GO term categories are only displayed for those with at least two matches. **C)** A Venn diagram displaying the number of proteins in the venom and saliva, and those identified in both samples via shotgun proteomic-based annotation to the genome. **D)** Reverse-phase chromatographic separation of venom. Venom was separated by semi-preparative reversed-phase HPLC (UV214nm) and manually collected. Peptides were directly submitted to LC-MS/MS, whereas protein fractions were analysed by SDS-PAGE (inset) under reducing conditions. Afterwards protein bands were subjected to in-gel trypsin digestion and identified by spectrum peptide matching against the translated *S. paradoxus* genome database. **E)** and **F)** LC-top-down MS analysis of saliva (**E**) and venom (**F**). The peak nomenclature is based on the chromatogram fractions, shown in panel **D**. **E)** Total ion current (TIC) profile of native saliva separated by high performance liquid chromatography (HPLC). **F)** Total ion current (TIC) profile of native venom separated by high performance liquid chromatography (HPLC). **G)** Summary table of the proteins identified via top-down and bottom-up proteomic analyses of solenodon venom, including their mass, corresponding identification in the genome (genome ID), and protein annotation. All identified proteins are annotated as *KLK1*, with the exception of keratins, which are human contaminants. **H)** Comparison of the relative abundance of main proteins present in chromatographic fractions of venom and saliva from top down MS experiments. SI Appendix Table S2 presents a summary of the protein matches identified in solenodon venom and saliva via the various proteomic approaches.

**Fig. 3. Functional assessments of Hispaniolan solenodon (*S. paradoxus*) venom reveals kallikrein serine protease activity and hypotensive effects.** Solenodon venom has extensive **A)** serine protease activity, as measured by chromogenic enzyme assay, and **B)** plasminogen activating activity, as measured by fluorescent enzyme assay. The data displayed is the mean rate of substrate consumption (**A**) or area under the kinetic curve (**B**) for mean measurements (±SEM) taken from three independent experiments; \*\*\*\* *p*<0.0001; \*\* *p*<0.01; unpaired two-tailed *t-*tests. **C)** Nanofractionation bioassaying reveals that *KLK1* proteins are responsible for plasminogen activating activity. *i* Bioactivity chromatogram at 5 mg/ml (blue line) and 1 mg/ml (red line) venom show the activity of each fraction, where positive peaks represent bioactive compounds. Bioactive wells selected for tryptic digestion are indicated by green arrows and well numbers, and those identified by mass spectrometry as *KLK1*s are labelled in red. *ii* UV trace at 254 nm collected during the LC-MS run with a UV-VIS detector. *iii* The total ion current (TIC) shown by the LC-MS chromatogram. *iiii* Extracted ion currents (XICs) of the *m/z*-values from the LC-MS data corresponding to the bioactives detected in the plasminogen assay. **D)** Solenodon venom degrades high molecular weight kininogen more potently than saliva without incubation. SDS-PAGE gel electrophoresis profiles demonstrate that both venom and saliva completely degrade kininogen (arrows) when preincubated for 60 mins, but that venom also degrades kininogen in the absence of preincubation. **E)** Solenodon venom causes substantial reductions in the pulse distension of envenomed mice (25 mg/kg; n=3) when compared to baseline measurements and controls (saline; n=3). The data displayed represents mean measurements, and the error bars represent standard deviations. **F)** Solenodon venom causes a transient depressor effect on the mean arterial blood pressure of the anaesthetised rat. The data displayed is a representative trace from one of five experimental animals that received 1 mg/kg venom (see also SI Appendix Fig. S3). **G)** Solenodon venom has no effect on nicotinic acetylcholine receptors. Representative whole-cell patch-clamp traces showing human muscle-type TE671 (left) and locust neuron nAChR (right) responses to 10 µM acetylcholine and the co-application of acetylcholine with 5 µg/ml solenodon venom. VH = -75 mV.

**Fig. 4. Molecular analyses reveal that eulipotyphlan venom systems and their toxin constituents have evolved independently by convergent evolution. A)** Molecular phylogeny of amino acid translations of tetrapod *KLK*s demonstrate that solenodon *KLK1* venom genes form a strongly supported monophyly and are polyphyletic to *Blarina* shrew venom genes. The phylogeny was derived by Bayesian inference analysis (n=106; 2x108 generations, four parallel runs with six simultaneous MCMC simulations). Genes encoding for proteins detected in solenodon venom (SI Appendix Table S2) or *Blarina* venom(11, 32)are highlighted by red-coloured branches and tip labels. Support values represent Bayesian posterior probabilities (BPP), where black circles represent BPP = 1.00 and grey circles BPP > 0.95. See also SI Appendix Figure S4 for the nucleotide-derived phylogeny. **B)** Analysis of the genomic organisation of mammalian *KLK*s demonstrates that *KLK1*s are atypically numerous in the solenodon. Distinct patterns of *KLK1* orientation across eulipotyphlans suggest that venom genes have arisen independently in the solenodon, and evidence of multiple solenodon genome scaffolds containing *KLK1* and *KLK15* adjacent to one another suggests that these may form the basis of a duplication cassette. **C)** Ancestral state reconstruction of the origin of venom in eulipotyphlans reveals that venom most likely evolved independently on four occasions (red vertical lines). Genera containing venomous species (or the species themselves) are highlighted by red tip labels. The computed ancestral traits for each node are depicted by pie charts, where the proportion of red colour represents the posterior probability of the most recent common ancestor being venomous, and blue represents non-venomous. In all cases, ancestral nodes support the non-venomous character state with a posterior probability of 1.00, except for the *Suncus* and *Crocidura* node, where the support value was greater than 0.85. Divergence times are indicated by the scale, and these, along with the tree topology, are derived from prior studies (15, 19, 36). The specific timing of the origin of venom should not to be inferred from the placement of the vertical red bars on the tree - these are placed arbitrarily at the mid-point of each relevant branch.