Phylogeography and venom composition of the rinkhals, *Hemachatus haemachatus* (Squamata: Elapidae)

Elmé Brand

Submitted in partial fulfilment of the requirements for the degree
Master of Science in Zoology
in the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria
16 April 2020

Supervisor: **Prof. Catherine L. Sole**[1]
Co-supervisor: **Dr. Ian A. Engelbrecht**[1,2]
Co-supervisor: **Prof. Graham J. Alexander**[3]

Collaborators:
Prof. Juan J. Calvete[4]
Dr. Davinia Pla[4]

[1]Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa

[2]South African National Biodiversity Institute, Pretoria, South Africa

[3]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

[4]Evolutionary and Translational Venomics Unit, Instituto de Biomedicina de Valencia, Valencia, Spain

Declaration

I, Elmé Brand, declare that this thesis/dissertation, which I hereby submit for the degree Master of Science in Zoology at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: 

DATE: 14/04/2020

Abstract

Phylogeographic patterns as well as divergence date estimates and the past population demography of *Hemachatus haemachatus* were analysed using nuclear and mitochondrial genetic markers. My results suggest that *H. haemachatus* forms a continuous population that diverged into two broad lineages due to glacial and interglacial cycles during the Plio-Pleistocene interval. During this interval, stable populations existed in the southern Cape and eastern inland regions of southern Africa. These cycles lead to further isolation and genetic differentiation of populations within these lineages during the Tarantian stage of the Pleistocene epoch. Currently, limited genetic variation exists between the lineages, suggesting that gene flow was restricted for a limited period only, and subsequently re-established. This most likely occurred as a result of a range expansion during the Last Glacial Maximum (LGM) which led to increased open grassland habitats. In addition to genetic analyses, I examined the venom composition of this species across its distribution using bottom-up approaches, which incorporated rp-HPLC, SDS-PAGE and MS/MS analysis. A total of 20 protein families were identified from six venom pools. I found that the major components of the venom were shared between all populations sampled, although slight variations were observed. The venom was rich in cytotoxic and neurotoxic elements. Cardiotoxic and hemorrhagic elements were also present. I did not observe a definitive link between phylogeographic patterns and the proteomic composition at the protein family level.

Acknowledgements

I would like to thank my father, Louis Brand, for always supporting me in my crazy adventures doing research on venomous snakes and making it possible for me to further my studies in a field of study which I absolutely love. I would also like to thank my partner, André Coetzer, who took part in my adventures looking for venomous snakes and assisting me with transporting and handling the specimens collected for this project. I am also very grateful to my mother who taught me to never give up and never to let anything stand in the way of my goals. I would also like to thank my in laws, Bennie and Irene Coetzer, for supporting me in my studies and for always being willing to assist me.

I would like to thank my supervisors, Prof. Catherine Sole, Dr. Ian Engelbrecht and Prof. Graham Alexander for their continuous guidance and support during this project. This was a learning curve for all of us and I would not have been able to start or complete this project without their support (and patience). I am especially grateful to Prof. Sole for providing me with the opportunity to start this project in the first place, for all the opportunities that came from the project, and for being such an awesome listener. I am also grateful to Dr. Engelbrecht for the role he played in kickstarting the project with getting in touch with potential sample collectors and nature conservation permit offices, as well as all the time and effort he put into helping me with R coding and other aspects of the project. I would also like to thank Prof. Alexander for his time and valuable inputs, as an expert on *Hemachatus haemachatus*, which made a big difference in the interpretation of the patterns I observed in this species.

I would also like to thank Prof. Juan Calvete, Dr. Davinia Pla, Mrs. Yania Rodriguez, Miss Sarai Quesada Bernat and the rest of the team from the Evolutionary and Translational Venomics lab at the Instituto de Biomedicina de Valencia for teaching me the techniques implemented in venom proteomics and for making me feel right at home in Spain. I am grateful to Prof. Calvete for providing me with the opportunity to collaborate with his specialised team on this project. It was an experience that I will never forget, and I look forward to using the valuable knowledge and skills gained from this going forward.

Also, thank you to Johan Marais for putting me in contact with Prof. Calvete, and a big thank you to Mike Perry for allowing us to use his facilities to house the live specimens and for doing the venom extractions.

I also want to thank the past and current members of the Invertebrate Biosystematics and Conservation group. Thank you to Carmen Jacobs, Shannon Mitchell and Ishtiag Abdalla for assisting me with my lab work at times and to Gimo Daniels for showing me how to do some of the distribution modelling in R.

I would also like to thank the Animal Demography Unit of the University of Cape Town for providing me with distribution data for my study species.

I would like to thank the National Research Foundation for providing funding for accessing equipment abroad, which allowed me to travel to the Instituto de Biomedicina de Valencia in Spain for six weeks to learn techniques implemented in studying venom proteomics from specialists in this field.

I also want to thank the Mpumalanga Tourism and Parks Agency, Eastern Cape Department of Economic Development, Environmental Affairs and Tourism, Northern Cape Department of Rural, Environmental and Nature Conservation, North West Department of Rural, Environmental and Nature Conservation, CapeNature, Ezemvelo KZN Wildlife, Gauteng Department of Agriculture and Rural Development, Free State Province Department of Economic, Small Business Development, Tourism and Environmental Affairs and the Department of Agriculture, Forestry and Fisheries for issuing the required permits. Also, thank you to the Animal Ethics Committee of the University of Pretoria for providing ethical clearance to continue with the project.

I owe a big thank you to everyone involved in collecting tissue samples and live specimens for this project. So, thank you to André Coetzer, Andrew MacLeod, Armin du Preez, Arné Hobbs, Arno Naude, Ashley Dodkins, Blake Rawlins, Bo Gilbert, Bryan Maritz, Carl Jackson, Chad Keates, Chris Hobkirk, Chris Wilken, Craig van Rensburg, Dawie van Eeden, Debbie van den Berg, Deon Jonck, Emiel van Zweel, Francois Oosthuisen, Frans Awie Vermaak, Frans de Coning, Gary Nicolau, Gerhard Bruwer, Grant Smith, Henery Robert Budgen, Hennie Haasbroek, Hennie van der Walt, Hester Obermeyer, Jean-Pierre Lees, Jo Balmer, Johan Hefer, Johan Pretorius, Juan

Marillier, Krystal Tolley, Leon Goosen, Lizette Oosthuizen, Louis Brand, Luke Kemp, Mark Marshall, Max Bastard, Mike Perry, Neels Bothma, Nico van der Merwe, Paul Rabiega, Paul Swanepoel, Petrus Bothma, Philip Jordaan, Pieter Kruger, Rian Viljoen, Rosa van Zyl, Sakkie Fischer, Schubert Smith, Shelley Edwards, Stefan Gouws, Steve de Klerk, Victor Boshoff, Werner Conradie, Zam Paulik and Zane Barnard. A special shout out to Zane Barnard who passed away in 2019 doing what he loved – rescuing and relocating snakes, RIP Zane.

Thank you to Luke Verburgt, Ursula Verburgt and André Coetzer for assisting with capturing amazing photographs of the snakes used for the venom component of this study.

Contents

List of Figures

Figures in main text ***Page***

Supplementary figures

List of Tables

Tables in main text **Page**

Supplementary tables

List of Acronyms

5’-NT	5'-Nucleotidase
16S	16S Ribosomal RNA
3FTx	Three Finger Toxin
AChE	Acetylcholinesterase
ADU	Animal Demography Unit
AIC	Akaike Information Criterion
AUC	Area Under Curve
BI	Bayesian Inference
BIC	Bayesian Information Criterion
BS	Bootstrap
CHELSA	Climatologies At High Resolution for the Earth's Land Surface Areas
CHIT1	Chitotriosidase
CID	Collision-Induced Dissociation
CPD	Central Posterior Densities
CRISP	Cystein-Rich Secretory Protein
CVF	Cobra Venom Factor
Cyt-b	Cytochrome B
DC domain	Disintergin-Like/Cysteine-Rich Domain
DNA	Deoxyribonucleic Acid
EBSP	Extended Bayesian Skyline Plot
EC	Eastern Cape Province
ESI-QUAD-TOF	Electrospray Ionization Quadrupole Time-Of-Flight
ESM	Earth System Model
ESS	Effective Sample Size
FS	Free State Province
GCM	Global Climate Model
GP	Gauteng Province
GPx	Glutathione Peroxidase
GTR	General Time Reversible
IGF	Insulin Growth Factor
KI	Kunitz Inhibitor
KZN	Kwazulu-Natal Province
LAO	L-Amino Acid Oxidase
LC	Liquid Chromatography
LGM	Last Glacial Maximum
LIG	Last Interglacial
MIROC-ESM	Model for Interdisciplinary Research On Climate, Earth System Model
MJ	Median-Joining
ML	Maximum Likelihood
MP	Mpumalanga Province
MPI-ESM-P	Max-Planck-Institute Earth System Model, Paleoclimate
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
mtDNA	Mitochondrial DNA

MW	Molecular Weight
NA	Not Available
NC	Northern Cape Province
nDNA	Nuclear DNA
NW	North West Province
PCR	Polymerase Chain Reaction
PDE	Phosphodiesterase
PLA2	Phospholipase A2
PLB	Phospholipase B
PP	Posterior Probability
PRLR	Prolactin Receptor
RAG1	Recombination Activating Gene 1
RAxML	Randomized Axelerated Maximum Likelihood
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
rp-HPLC	Reversed-Phase High-Performance Liquid Chromatography
SA	Serum Albumin
SANBI	South African National Biodiversity Unit
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
SVMP	Snake Venom Metalloprotease
SVSP	Snake Venom Serine Protease
TFA	Trifluoroacetic Acid
VEGF	Vascular Endothelial Growth Factor
vNGF	Venom Nerve Growth Factor
WC	Western Cape Province
WHO	World Health Organisation

Chapter One

General introduction

1.1 Consequences of snakebite

Venomous snakes occur in 169 countries worldwide (Kasturiratne et al., 2008). This results in snakebites being a common and widespread threat for outdoor recreationists, farmers, plantation workers and people living in rural communities (Chippaux 2011; Longbottom et al. 2018; WHO 2019). The burden of snakebite is most prominent in developing regions, specifically countries in sub-Saharan Africa, South and Southeast Asia and South and Central America (Kasturiratne et al. 2008; WHO 2019). Victims from developing regions do not have ready access to effective treatment due to scarcity and high cost of antivenom, which is the only effective method of treatment at this stage (Kasturiratne et al. 2008; Brown 2012; WHO 2016). Healthcare is also inaccessible and suboptimal in such regions and victims, thus often seek ineffective traditional remedies (Kasturiratne et al. 2008; Brown 2012). Many victims who do survive are left with permanent disabilities, which have major economic impacts (Kasturiratne et al. 2008; WHO 2019). Despite these issues, snakebite is considered a neglected public health problem in developing countries and results in substantial morbidity and mortality worldwide (Ahmed et al., 2008; Kasturiratne et al., 2008; Chippaux, 2011).

1.2 Venom composition, phylogenetics and phylogeography

Reptiles represent the largest group of vertebrates that are able to produce venom (Mackessy 2010). Members of this group are considered excellent model systems for evolutionary studies, especially with regards to the origins of body form, asexuality, viviparity and venom (Wiens et al. 2012). Snake families Lamprophiidae, Viperidae, Colubridae, Elapidae, as well as the lizard family Helodermatidae represent the main squamate lineages that contain venom-producing species (Mackessy 2010).

Phylogenetics can be described as the study of evolutionary relationships of organisms (Dunn et al. 2014). A phylogeny groups organisms according to their evolutionary history by comparing their shared derived characteristics, i.e. synapomorphies (Dowell 2008; Dunn et al. 2014). Phylogeography, on the other hand,

can be described as the discipline in which the relationship between phylogenetics and geographic distribution are examined (Avise 1989). The historical demography of a species can be inferred from phylogeographic trees (Avise 1989). Investigating the phylogeographic and demographic history of a widespread species allows researchers to understand the biogeographic history of a specific region, investigate barriers to gene flow and place the effects of such barriers within a specific time period (Burbrink et al. 2008). Populations display the accumulated effects of processes that affect individuals within the population, which ultimately impact speciation events (Sunnucks 2000). Recent and ancient population changes can be estimated from differences in sequences and the shape of trees (Burbrink and Castoe 2009). Therefore, information related to almost any evolutionary process within a population can be inferred from genetic markers that display appropriate signals, i.e. rates of change (Sunnucks 2000; Hurst and Jiggins 2005).

Venom composition tends to vary between snake lineages (Mackessy 2010). Therefore the severity of snakebite envenomation largely depends on the local venom variant and the potency of the variant, as antibodies found in antivenoms may be ineffective against different toxin varieties (Casewell et al. 2014). However, few studies have linked variation in venom composition to differences in phylogeography and it would therefore be beneficial to investigate these aspects by studying species that contain diverse lineages (Mackessy 2010; Lourenço et al. 2013; Casewell et al. 2014).

1.3 The study species

Hemachatus haemachatus (Bonnaterre 1790), is a venomous elapid endemic to southern Africa (Branch 1998; Marais 2004). It is monotypic and is a sister taxon to true cobras of the genus *Naja* (Branch 1998; Wüster et al. 2007). This species occurs predominantly in grassland habitats, but has also been found to occur in moist savanna, fynbos and lowland forest habitats (Branch 1998; Marais 2004). It occurs from the coast up to elevations of 2500 m, throughout the eastern parts of South Africa, including the provinces of Gauteng, Free State, Kwa-Zulu Natal, Eastern Cape, Western Cape, Northern Cape, North West and Mpumalanga (Branch 1998; Marais 2004). An allopatric population also occurs in eastern Zimbabwe, although it is believed to have gone locally extinct in that region (Alexander 2018).

Hemachatus haemachatus is common throughout most of its range, with relict populations in the Western, Eastern and Northern Cape (Branch 1998). It varies greatly in size and colouration across its distribution (Branch 1998; Marais 2004). Specimens from Kwa-Zulu Natal, the Eastern Cape, Northern Cape and the Western Cape are conspicuously banded (Branch 1998; Marais 2004). Individuals from these areas are usually olive to dark brown or black in colouration alternating with orange, yellowish or pale grey asymmetrical bands across the body (Fig. 1) (Branch 1998; Marais 2004). Individuals further inland are generally not banded and are olive to dull black or dark brown and occasionally speckled with lighter grey or brown (Fig. 1) (Branch 1998; Marais 2004). The common name, rinkhals/ringhals cobra, is derived from the white cross-bands on the throat area, which are present in both banded and plain individuals (Branch 1998; Marais 2004). There is limited information as to whether the variation in colouration observed in *H. haemachatus* is due to evolutionary factors or phenotypic plasticity.

In addition to variation in colouration and size of *H. haemachatus* across its range, its venom composition is also suspected to vary between populations. The venom of *H. haemachatus* is described as predominantly cytotoxic with mild neurotoxic effects (Müller et al. 2012) or neurotoxic (Branch 1998; Marais 2004). Predominantly cytotoxic symptoms are observed in human patients, however, mild neurotoxic symptoms have also been recorded (Blaylock 2005; Müller et al. 2012). Anecdotal observations of *H. haemachatus* bites in dogs suggest that dogs respond differently to envenomation, with varying degrees of cytotoxic and neurotoxic symptoms and this appears to be dependent in part on the geographic origin of the snake. However, the relationship between the variant of venom and geography has not been studied.

Given the known variation in colouration of this species, and the reported differences in the symptoms caused by envenomation, this species is an ideal candidate for assessing phylogeographic patterns, differences in venom composition and whether such differences can be attributed to different evolutionary histories of the populations.

Figure 1: General patterns of colour variation observed across the geographic range of *Hemachatus haemachatus*. This image shows colour variations of adult and sub-adult snakes only and it should be noted that juveniles may deviate from this general trend, especially in inland regions.

1.4 Aims and objectives

I aimed to assess the phylogeographic history and genetic structure of *H. haemachatus* across its range, the demographic history of the species and the divergence times of the intraspecific lineages. I also aimed to investigate whether differences in venom composition are related to phylogeographic patterns.

1.5 Expectations

Based on a pilot study conducted in 2017, *H. haemachatus* appears to have undergone a relatively recent range expansion. This study indicated that *H. haemachatus* is divided into two broadly supported mitochondrial lineages with a third lineage, indicated by a single specimen from Mpumalanga, falling sister to both these

lineages. The two broad lineages appear to have diverged from one another near the border of the Eastern Cape and KwaZulu-Natal provinces. The first lineage includes specimens with Northern, Eastern and Western Cape, Gauteng, Mpumalanga and Kwa-Zulu Natal. The distribution of the second lineage includes specimens from Kwa-Zulu Natal. The first lineage was found to form two distinct geographic groups, while the second lineage appeared to be confined to a single geographic group.

Based on these results, a similar divergence was expected at the border of the Kwa-Zulu Natal and Mpumalanga provinces and a distinct sister lineage was expected to occur within the Mpumalanga Province. However, the results of the pilot study were based on a limited number of samples (29) and mitochondrial gene regions only, thus more extensive sampling, the addition of nuclear and ribosomal gene regions to the molecular dataset and the incorporation of more integrative molecular techniques were implemented in this study to corroborate the results from 2017 and improve the information of the population structure for *H. haemachatus*.

In addition to this, I expected to observe differences in the venom composition of *H. haemachatus* and to find a link between variations in the venom composition and phylogeographic lineages, based on the assumption that venom composition is genetically determined.

1.6 Ethical clearance and permits

The experimental procedures for this study have been cleared by the Animal Ethics Committee of the University of Pretoria: clearance number ec071-16 (Revised) (A1). Scientific collection permits were obtained from the Mpumalanga Tourism and Parks Agency (MPB. 5603), Eastern Cape Department of Economic Development, Environmental Affairs and Tourism (CRO 2/18, CRO 3/18), Northern Cape Department of Rural, Environmental and Nature Conservation (FAUNA 0063/2018, FAUNA 0064/2018), North West Department of Rural, Environmental and Nature Conservation (HO 02/03/18-056 NW), CapeNature (CN44-30-3687), Ezemvelo KZN Wildlife (OP 1761/2018), Gauteng Department of Agriculture and Rural Development (CPF 000200, CPC5 3640, CPB6 0297, CPB6 001822, CPB6 002203, CPB6 002361, CPB6 0300, CPB6 0298, CPB6 0299), and the Free State Province Department of Economic, Small Business Development, Tourism and Environmental Affairs (JM

4081/2018). I also obtained an international export permit and section 20 permit from the Department of Agriculture, Forestry and Fisheries.

Chapter Two

The phylogeography of rinkhals, *Hemachatus haemachatus* (Squamata: Elapidae), across South Africa

2.1 Introduction

2.1.1 Phylogeography

Organisms differ in the DNA sequences that comprise their genomes at an individual level. These differences lead to genetic diversity which plays a critical role in key evolutionary processes involved in adaptation, survival and speciation (Sunnucks 2000; Guicking et al. 2004; Cox and Chippindale 2014). The characteristics of a population change over time as a result of evolution, thus estimating the level of diversity within and between populations can provide insight into changes observed in population dynamics and structure (Prior et al. 1997; Fu and Li 1999).

The geography of a region can affect patterns of gene flow and genetic diversity as geographic barriers and corridors may affect movement and geneflow between populations (Burbrink and Castoe 2009; Cox and Chippindale 2014; Card et al. 2016). These patterns as well as the geographic and temporal origins of lineage diversity can be investigated using phylogeographic methods (Burbrink and Castoe 2009; Cox and Chippindale 2014; Card et al. 2016). To effectively examine the diversification of populations over a geographic area, a detailed understanding of the histories of colonisation and gene flow between the different populations is required (Hurst and Jiggins 2005).

Mitochondrial DNA (mtDNA) markers have formed the basis of phylogeographic and phylogenetic studies due to their high rates of nucleotide evolution, which are approximately five to 10 times greater than protein-coding nuclear genes. They also have single-copy status, lack of recombination and multiple copies within a cell. These attributes facilitate easy amplification as previously published primers and complete mitochondrial sequences are readily available (Burbrink and Castoe 2009; Burbrink and Guiher 2015; Hurst and Jiggins 2005; Sunnucks 2000). Mitochondrial DNA can be used to recover patterns and measures of recent historical events without extensive sequencing being necessary (Hurst and Jiggins 2005; Burbrink and Castoe 2009; Fu and Li 1999). It also represents an effective marker of

population history and gene flow, even in species that show little genetic variation (Sunnucks 2000). These characteristics result in an increased probability of mtDNA gene trees representing the species tree correctly (Burbrink and Castoe 2009).

However, there are problems associated with using only mitochondrial genes to infer the evolutionary histories of populations (Hewitt 2004; Hurst and Jiggins 2005; Ursenbacher et al. 2006; Jiang et al. 2007). Mitochondrial DNA only represents historical processes of the maternal lineage and may therefore not reflect the phylogeographic history of the species as a whole (Giannasi et al. 2001; Hurst and Jiggins 2005; Burbrink and Castoe 2009). Different mitochondrial genes do not provide independent evolutionary information regarding phylogeographic and phylogenetic patterns and only reflect the maternal lineage (Burbrink and Castoe 2009). Phylogeographic studies should therefore make conclusions based on both mtDNA and nuclear DNA (nDNA) sequences (Giannasi et al. 2001; Burbrink and Castoe 2009). However, nDNA is generally not considered useful for studying recent divergences, as limited phylogeographic information is provided by nDNA. However, sequences that evolve more slowly allow for a better understanding of the deeper phylogeographic history of a species (Hewitt 2004; Burbrink and Castoe 2009; Burbrink and Guiher 2015).

Maximum likelihood, Bayesian inference and Markov chain Monte Carlo simulation techniques are used for population genetics as well as phylogenetic/phylogeographic analyses (Sunnucks 2000; Boussau and Daubin 2010). Spatial and historical hypotheses, relative timing and direction of events, such as range expansions, and the differentiation of current restricted gene flow from historical gene flow can be tested using these techniques (Sunnucks 2000). The probability of tree topology and demographic parameters that determine whether population sizes have declined, experienced growth or remained constant are assessed using these models (Burbrink and Castoe 2009). The structure of haplotypes across a geographic range may be representative of range expansion, current or historical barriers to gene flow or past population fragmentation (Burbrink and Castoe 2009).

2.1.2 Effect of past climate

Genetic diversity and distributions of many species have been influenced significantly by past climatic events, such as glacial climate fluctuations (Hofmann 2012;

DeMenocal 2004). Climatic conditions associated with such periods impact ecological relationships, which in turn caused fauna to experience high selection pressure, leading to changes in genetic diversity (DeMenocal 2004; Montgelard and Matthee 2012; Hofmann 2012). Glacial climate fluctuations have therefore affected the genetic structure as well as current distribution of modern species through various isolation and recolonisation events (Inoue et al. 2015; Mokhatla et al. 2015).

The African continent has been described as being highly vulnerable in relation to climate change (Sintayehu 2018). It contains a large variety of climates and ecosystems, such as grasslands, savannahs, woodlands, shrublands, forests (dry and wet), steppes and deserts (Kissling et al. 2016). Evidence, based on Pan-African mammals, suggests that Africa contained open-habitat refugia during the Pleistocene, which were isolated by tropical forests which expanded during interglacial periods characterised by warm and humid conditions (Montgelard and Matthee 2012; Barlow et al. 2013). Trauth et al. (2009) proposed that a significant aridification period occurred approximately 3.4 mya. In addition to this, it has been proposed that three Plio-Pleistocene aridification events occurred in southern Africa at around 2.8, 1.7 and 1.0 mya, which were interspersed by warm and humid periods (DeMenocal 2004; Montgelard and Matthee 2012). These fluctuations resulted in the contraction and expansion of suitable habitat, leading to migration, diversification, adaptation or extinction of species occurring there (Montgelard and Matthee 2012). Regions that experienced relatively stable climatic conditions are considered to have acted as refugia during climatic fluctuations (Mokhatla et al. 2015). Paleo-vegetation and fossil data suggest that scattered refugia existed in the eastern parts and a long-standing refugium was present in the southern parts of southern Africa (Lorenzen et al. 2010).

2.1.3 The study species

Hemachatus haemachatus was selected as the study species as it is phenotypically variable and occurs across a range of habitats (Branch 1998; Marais 2004; Alexander et al. 2012). It is widespread and relatively common throughout most of its range, although only relict populations occur in the Northern and Western Cape provinces and Zimbabwe (Branch 1998). Generalist species that are relatively widespread are considered powerful model systems for investigating and understanding how various

ecological factors can drive patterns of species diversification and divergence on a regional scale (Card et al. 2016).

A clear pattern is evident in the distribution of the different colour morphs. Specimens from Kwa-Zulu Natal, the Eastern Cape, Northern Cape and the Western Cape are conspicuously banded (Branch 1998; Marais 2004). Individuals from these areas are usually olive to dark brown or black in colouration alternating with orange, yellowish or pale grey asymmetrical bands across the body (Fig. 1) (Branch 1998; Marais 2004). Individuals further inland are generally not banded and are olive to dull black or dark brown and occasionally speckled with lighter grey or brown (Fig. 1) (Branch 1998; Marais 2004). However, there are limited reports of banded and plain, or rather melanistic forms of banded individuals occurring sympatrically in Port Elizabeth, Hermanus and around Grahamstown. This species also has a wide thermal tolerance, allowing it to occur in areas with altitudes of up to 2500 m and survive being exposed to sub-zero temperatures on a frequent basis (Branch 1998; Alexander et al. 1999,2012). *Hemachatus haemachatus* is also of taxonomic interest as it is monotypic (Wüster et al. 2007) and is most closely related to cobras of the genus *Naja* (Wüster et al. 2007; Head et al. 2016).

Important information on the evolution, taxonomy and biodiversity of snakes has been revealed from phylogeographic studies (Burbrink and Castoe 2009). Phylogeographic studies on snakes may also prove vital for validating and clarifying broad phylogeographic patterns that have played a role in shaping various components of the biota within a specific region (Burbrink and Castoe 2009). Investigating the phylogeographic and demographic history of a widespread species, such as *H. haemachatus*, can improve the understanding of the biogeographic history of a specific region, as well as provide an opportunity to investigate barriers to gene flow and place the effects of such barriers within a specific time period (Burbrink et al. 2008).

2.1.4 Aims and objectives

The aim of this component of the project was to assess the population structure of *H. haemachatus* across South Africa by:

- investigating phylogeographic patterns of this species,
- estimating dates of divergence of different lineages and

- modelling past distribution using paleoclimatic data.

2.2 Methods

2.2.1 Taxon and tissue sampling

Samples and specimens were collected from the Gauteng, KwaZulu-Natal, Mpumalanga, Eastern Cape, Western Cape, Northern Cape, North West and Free State provinces in South Africa. Additional samples, including pieces of shed skin, were sourced from the South African National Biodiversity Institute's (SANBI) Herpetology Tissue Bank as well as the Port Elizabeth Museum herpetology collection (Table 1).

Table 1: Accession numbers for samples sourced from the South African National Biodiversity Institute's Herp Tissue Bank and the Port Elizabeth Museum herpetology collection.

Sample name	Source	Accession number
WC_5072	Port Elizabeth Museum	PEM R22655
WC_6676	Port Elizabeth Museum	PEM R24871
GKN_120	Port Elizabeth Museum	PEM R24997
2019_Rinkhals_Shamwari	Port Elizabeth Museum	To be confirmed
WC_DNA_1215	Port Elizabeth Museum	PEM R20666
KAT17-01	SANBI	6186

Ventral scale clippings that contain a thin layer of connective tissue at the base of the scales were collected using small sterilised scissors and stored in a vial of absolute ethanol at -20 °C. This served as a non-lethal method for collecting tissue samples for molecular analyses (Burbrink and Castoe 2009).

2.2.2 DNA extraction, amplification, sequencing and alignment

Whole genomic DNA was extracted from each ventral scale clipping, or fragment of shed skin. Extractions were performed using a Macherey-Nagel NucleoSpin® Tissue kit (Düren, Germany) according to the protocol indicated by the manufacturer.

One mitochondrial gene region, cytochrome b (Cyt-b), one ribosomal gene region, 16S ribosomal RNA (16S) and two nuclear gene regions, recombination-activating gene 1 (RAG-1) and Prolactin Receptor (PRLR), were amplified by Polymerase Chain Reaction (PCR) and sequenced (Palumbi 1996; Groth and

Barrowclough 1999; Lawson et al. 2004; Metzger et al. 2010). The Cyt-b gene region was selected as most suitable for comparing species within the same genus due to its sequence variability; the evolutionary patterns of Cyt-b in Elapidae have been found to be very similar to those of intensively studied groups such as mammals (Castresana 2001; Leavitt and McClellan 2014). The 16S gene evolves at a slower rate compared to Cyt-b (Burbrink and Castoe 2009), however it is also frequently used in snake phylogeographic studies and was therefore included in this study. The PRLR gene was included in this study as it evolves at a faster rate compared to other nuclear genes (Townsend et al. 2008). Protein-coding genes, such as Cyt-b, PRLR and RAG-1, are required to compare nonsynonymous and synonymous rates of evolution when assessing population expansions (Burbrink and Castoe 2009). These genes are easy to align and generally do not contain gaps when aligned in phylogeographic studies (Burbrink and Castoe 2009). These four gene regions were selected to maximise the amount of phylogenetically informative data at a fine genetic scale, specifically aimed at phylogeographic studies and allowed testing of whether the species tree is accurately reflected by the gene trees through congruence of these respective trees (Doyle and Hortorium 1992; Maddison 1997).

Previously published primer sequences were selected to amplify these gene regions (Table 2). The mixtures for PCR amplification constituted the following: 20 pmol of each primer, a single unit of Emerald Amp®MAX HS PCR Mastermix (TAKARA BIO INC., Otsu, Shiga, Japan), 100-500 ng of extracted mitochondrial DNA template; the mixture was made up to a total volume of 25 µl by adding distilled water. The thermocycling conditions are indicated in Table 3. Successful amplifications were purified using a Macherey-Nagel NucleoSpin® PCR and Gel Clean-up kit (Düren, Germany).

The BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA) was used to perform cycle sequencing of reactions in the forward and reverse directions. The products of the cycle sequencing reactions were precipitated by following the standard protocol for sodium acetate and ethanol precipitation. CLC Main Workbench version 7.9.2 (QIAGEN Bioinformatics) was used to generate and edit the sequences. Alignment and further editing of the sequences was done in MEGA 7 (Kumar et al. 2016).

Table 2: Primer details for the selected gene regions amplified.

Gene	Primer	Sequence	Publication
Cyt-b	H16064	5’ CTT TGG TTT ACA AGA ACA ATG CTT TA 3’	Burbrink et al. (2000)
Cyt-b	L14910	5’ GAC CTG TGA TMT GAA AAA CCA YCG TTG T 3’	De Queiroz et al. (2002)
Cyt-b	H15149	5’ AAA CTG CAG CCC CTC AGA ATG ATA TTT GTC CTC A 3’	Kocher et al. (1989)
Cyt-b	L15584	5’ TCC CAT TYC ACC CAT ACC A 3’	De Queiroz et al. (2002)
16S	L2510	5’ CGC CTG TTT ATC AAA AAC AT 3’	Palumbi (1996)
16S	H3056	5’ CTC CGG TCT GAA CTC AGA TCA CGT AGG 3’	Hedges (1994)
RAG-1	G396 (R13)	5’ TCT GAA TGG AAA TTC AAG CTG TT 3’	Groth & Barrowclough (1999)
RAG-1	G397 (R18)	5’ GAT GCT GCC TCG GTC GGC CAC CTT T 3’	Groth & Barrowclough (1999)
PRLR	PRLR_f1	5’ GAC ARY GAR GAC CAG CAA CTR ATG CC 3’	Townsend et al. (2008)
PRLR	PRLR_r3	5’ GAC YTT GTG RAC TTC YAC RTA ATC CAT 3’	Townsend et al. (2008)

Table 3: Thermocycling conditions for amplification of the selected gene regions. The grey shaded area highlights the steps that were repeated for the number of cycles specified.

Gene region	Thermocycling conditions	
Cyt-b	7 min at 94 °C (denaturation)	
	40 sec at 94 °C (denaturation)	X 40 cycles
	30 sec at 46 °C (annealing)	
	1 min at 72 °C (elongation)	
	7 min at 72 °C (elongation)	
16S	3 min at 94 °C (denaturation)	
	3 min at 94 °C (denaturation)	X 40 cycles
	40 sec at 50 °C (annealing)	
	1 min at 72 °C (elongation)	
	10 min at 72 °C (elongation)	
RAG-1	3 min at 95 °C (denaturation)	
	30 sec at 95 °C (denaturation)	X 30 cycles
	30 sec at 57 °C (annealing)	
	45 sec at 72 °C (elongation)	
	10 min at 72 °C (elongation)	
PRLR	2 min at 94 °C (denaturation)	
	30 sec at 94 °C (denaturation)	X 30 cycles
	1 min at 52 °C (annealing)	
	45 sec at 72 °C (elongation)	
	5 min at 72 °C (elongation)	

2.2.3 Phylogenetic and phylogeographic estimation

Evolutionary relationships were inferred from molecular data generated from sequences using Maximum Likelihood (ML) and Bayesian Inference (BI) approaches. Various elapid, viperid and colubrid sequences were downloaded from GenBank® as outgroups (Table 4).

Table 4: Genbank accession numbers for outgroup taxa for the selected gene regions. Accession numbers shown as NA refer to gene regions for which no data were available on Genbank.

Family	Species/Sample name	16S	Cyt-b	RAG-1	PRLR
Colubridae	*Diadophis_punctatus*_1	KX694636.1	EU193858.1	NA	NA
Colubridae	*Diadophis_punctatus*_2	AF544793.1	FJ358789.1	NA	NA
Elapidae	*Micrurus_fulvius*_1	GU045453.1	GU045453.1	NA	JN880856.1
Elapidae	*Micrurus_fulvius_2*	NC_013481.1	NC_013481.1	NA	JN880856.1
Viperidae	*Vipera_aspis*_1	AY611814.1	AY611997.1	KX357730.1	MG875590.1
Viperidae	*Vipera_berus*_1	NC_036956.1	NC_036956.1	KX357727.1	LT221008.1
Viperidae	*Vipera_berus_2*	MF945570.1	MF945570.1	KX357727.1	LT221007.1
Viperidae	*Gloydius_ussuriensis*_1	MK194042.1	JQ687503.1	NA	NA
Viperidae	*Gloydius_ussuriensis_2*	MK194043.1	JQ687501.1	NA	NA
Viperidae	*Agkistrodon_piscivorus*_1	MN135488.1	EU483487.1	NA	JN880801.1
Viperidae	*Agkistrodon_piscivorus_2*	KX694680.1	EU483486.1	NA	JN880801.1
Elapidae	*Laticauda_laticaudata*_1	FJ587203.1	AB701327.1	FJ587083.1	FJ587127.1
Elapidae	*Laticauda_laticaudata_2*	KX239663.1	AB701328.1	FJ587082.1	FJ587126.1
Elapidae	*Laticauda_colubrina*_1	MK193892.1	AF217834.1	NA	JQ910415.1
Elapidae	*Laticauda_colubrina_2*	KX239659.1	KX239643.1	NA	JN880849.1
Elapidae	*Sinomicrurus_kelloggi*_1	MK194268.1	AF220408.1	NA	NA
Elapidae	*Sinomicrurus_kelloggi_2*	MK193898.1	EF137417.1	NA	NA
Viperidae	*Macrovipera_lebetina*_1	EU624294.1	KJ415301.1	NA	NA
Viperidae	*Macrovipera_lebetina_2*	KX694652.1	MF446029.1	NA	NA

Elapidae	*Oxyuranus_scutellatus*_1	KT026532.1	EU547051.1	EU546877.1	NA
Elapidae	*Oxyuranus_scutellatus*_2	EU547149.1	DQ897772.1	EU546877.1	NA
Elapidae	*Oxyuranus_microlepidotus*	EU547148.1	EU547050.1	NA	NA
Viperidae	*Vipera_ammodytes*_1	DQ186102.1	DQ186508.1	KX357729.1	NA
Viperidae	*Vipera_ammodytes*_2	EU624297	DQ186520.1	KX357729.1	NA
Elapidae	*Micrurus_tener*_1	KR814656.1	KR814692.1	NA	NA
Elapidae	*Micrurus_alleni*	KX660164.1	KX660439.1	NA	NA
Elapidae	*Sinomicrurus_macclellandi*_1	MK194273.1	MG653604.1	NA	NA
Elapidae	*Sinomicrurus_macclellandi_2*	MK194267.1	MG653605.1	NA	NA
Colubridae	*Lycodon_rufozonatum*_1	KJ179950.1	KJ179950.1	KC733181.1	NA
Colubridae	*Lycodon_ rufozonatum*_2	KF148622.1	KF148622.1	KC733179.1	NA
Colubridae	*Boiga_multomaculata*_1	MK194200.1	FJ710798.1	NA	NA
Colubridae	*Boiga_beddomei*_1	KC347346.1	KC347466.1	NA	NA
Colubridae	*Boiga_ceylonensis*_1	KC347347.1	KC347467.1	NA	NA
Colubridae	*Lycodon_subcinctus*_1	MK194178.1	KC010385.1	KC733188.1	NA
Colubridae	*Lycodon_subcinctus*_2	MK193962.1	KC010384.1	KC733188.1	NA
Elapidae	*Naja_atra*_1	EU913475.1	EU913475.1	NA	NA
Elapidae	*Naja_atra*_2	KX694689.1	JN160676.1	NA	NA
Elapidae	*Naja_naja*_1	DQ343648.1	EU547039.1	EU366432.1	NA
Elapidae	*Naja_naja*_2	EU547137.1	GQ225657.1	EU366432.1	NA
Elapidae	*Naja_kaouthia*_1	KX277260.1	AF217835.1	JF412633.1	JN880858.1
Elapidae	*Naja_kaouthia*_2	JN687925.1	JF357939.1	JF412633.1	JN880858.1

Elapidae	*Naja_melanoleuca*_1	MF497427.1	AY611995.1	JF357955.1	MH337500.1
Elapidae	*Naja_melanoleuca*_2	AY611812.1	FR693726.1	JF357955.1	MH337498.1
Elapidae	*Naja_haje*_1	AY611811.1	AY611994.1	NA	NA
Elapidae	*Naja_mossambica*_1	GQ359744.1	AY611996.1	NA	NA
Elapidae	*Naja_nivea*_1	GQ359755.1	AF217827.1	NA	NA
Elapidae	*Naja_pallida*_1	GQ359745.1	FR693723.1	NA	NA
Elapidae	*Walterinnesia_aegyptia*_1	HQ267785.1	AF217838.1	NA	NA
Elapidae	*Aspidelaps_scutatus*_1	KX694617.1	AF217828.1	NA	NA
Elapidae	*Aspidelaps_scutatus*_2	AY188046.1	AY188007.1	NA	NA
Elapidae	*Bungarus_candidus*_1	JN687933.1	AJ749335.1	NA	NA
Elapidae	*Bungarus_multicinctus*_1	EU579522.1	AJ749345.1	NA	NA
Elapidae	*Ophiophagus_hannah*_1	KX694688.1	AF217842.1	NA	NA
Elapidae	*Ophiophagus_hannah*_2	JN687931.1	NC_011394.1	NA	NA
Elapidae	*Dendroaspis_jamesoni*_1	KX660192.1	MF407272.1	NA	NA
Elapidae	*Dendroaspis_jamesoni*_2	KX660191.1	KX660464.1	NA	NA
Elapidae	*Dendroaspis_angusticeps*_1	JF357945.1	JF357936.1	NA	NA
Elapidae	*Dendroaspis_angusticeps*_2	FJ404194.1	MF407270.1	NA	NA

Randomized Axelerated Maximum Likelihood v8.2 (RAxML v.8.2; Stamatakis, 2014) was used to estimate the ML tree under the default general time-reversible (GTR) model (Tavaré and Miura 1986) of sequence evolution and clade support was estimated using 1000 Bootstrap (Felsenstein 1985) replicates. The analysis was repeated three times using the rapid hill-climbing algorithm (Stamatakis et al. 2007) to ensure that tree space is effectively sampled – a different starting seed was used for each repeat.

The best model of sequence evolution was assessed using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) by implementing jModelTest (Posada 2008). To assess nodal support and infer trees Bayesian inference (BI) was run three times for 20 million generations using MrBayes v3.1.2 (Ronquist and Huelsenbeck 2003). The parameters were specified as “unlinked” for each partition to ensure that separate estimates were generated for each gene region. The rate-prior was specified as “variable” and flat Dirichlet-priors were employed. The analysis ran two parallel independent runs with one cold and three heated Markov chains per run. Tree sampling occurred every 500th generation. The resulting BI trees were compared to the ML trees and inferences of evolutionary relationships were considered reliable for nodes with posterior probability (PP) values greater or equal to 95% and nonparametric bootstrap values greater or equal to 70% (Hillis and Bull 1993; Burbrink et al. 2008). The resulting phylogenetic trees were viewed and edited in FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/Figtree). The locality data for each sample were plotted onto a South African map using QGIS v3.12 (https://qgis.org/en/site/) to assess phylogeographic patterns.

SplitsTree v4 (Huson and Bryant 2006) was used to also generate unrooted phylogenetic networks for the abovementioned gene regions and a bootstrap analysis was run in the program. This allowed for the comparison of the groupings of taxa between the BI and ML trees generated with the combined molecular dataset and the phylogenetic networks generated with the Cyt-b and 16S gene regions. The resulting networks were viewed and edited in FigTree.

Haplotype networks were generated in Network v4.6.1.6 (2004-2018 Fluxus Technology Ltd.) using the Median Joining (MJ) algorithm (Bandelt et al. 1999).

2.2.4 Divergence dating

The timing of the origin of each evolutionary lineage was inferred using a relaxed molecular clock. A phylogenetic estimate of the lineages with outgroups for the combined Cyt-b and 16S dataset was made using the appropriate model of sequence evolution, as determined in jModeltTest, in BEAST2 (Bouckaert et al. 2019). A Yule branching process was implemented and the timing of lineage divergence was estimated by using an uncorrelated lognormal tree prior with fossil calibration dates (shown in Table 5) (Drummond et al. 2006). The analysis was run for 40 million generations, with the first 25% being discarded as burnin. The program Tracer v1.7.1 (Rambaut et al. 2018) was used to estimate effective sample sizes (ESS) for the parameters. LogCombiner v2.6.0 (Bouckaert et al. 2019) was used to combine the log and tree output files. The resulting phylogeny was viewed in FigTree. The dated phylogeny was produced in R v3.0.3 (R Core Team 2014) using the geoscalePhylo function. The phyloc (Cusimano and Heibl 2013), XML (Lang and The CRAN Team 2019), ape (Paradis et al. 2019), ggplot2 (Wickham et al. 2019), strap (Bell and Lloyd 2015), OutbreakTools (Aanensen et al. 2015) and network (Butts et al. 2019) packages were required. The R script can be viewed under supplementary data (Script S1).

Table 5: Fossil calibration references used for divergence dating of lineages.

Fossil calibrations	Calibration priors	References and description
Micrurus_Sinomicrurus	Hard min: 18 Soft max: 32	Priors used in article by Lee et al. (2016) Hard Min: *Micrurus gallicus* • known from MN4, middle Orleanians (approximately 18 mya) (Ivanov 2002) • shares similarities with extant *Micrurus* species (Szyndlar and Schleich 1993) Soft Max: • rich fauna of colubrids found in Europe and North America up to early Oligocene (approximately 32 mya) • deposits yielded no taxa similar to *Micrurus* (Rage 1988)
Naja_Hemachatus	Hard min: 18 Soft max: 32	Priors used in article by Lee et al. (2016) Hard Min: *Naja romani* • known from MN4, middle Orleanians (approximately 18 mya) (Ivanov 2002) • shares some synapomorphies with Asiatic *Naja*, thus suggested to form part of crown *Naja*

		• retains primitive traits that are either rare or absent in extant *Naja*, thus suggested to form part of stem *Naja* (Szyndlar and Rage 1990) • despite contradictions, assumed to form part of *Naja* group as a whole (stem or crown) Soft max: • large diversity of Colubridae as well as unidentified members of Elapidae found in North America and Europe up to early Oligocene (approximately 32 mya) • deposits yielded nothing similar to *Naja* (Rage 1988) • used to calibrate divergence between *Naja* and *Hemachatus* (closest extant relative of *Naja*)
Laticauda_Oxyuranus	Lognormal mean: 1 Standard deviation: 0.8 Zero offset: 20	Priors used in article by Lukoschek et al. (2012) *Laticauda* fossil vertebra: • from late Oligocene/early Miocene (approximately 20-23 mya) • similar to *L. colubrina*, differs from *L. laticaudata* and other elapids (Scanlon and Lee 2003) • used to calibrate *Laticauda* and *Oxyuranus* divergence

Viperidae_Colubroidea	Lognormal mean: 2 Standard deviation: 1.2 Zero offset: 40	Priors used in articles by Wüster et al. (2008) and Lukoschek et al. (2012) Divergence between Viperidae and Colubroidea: • youngest unambiguous colubroid fossil from Eocene in Asia (approximately 40 mya) (Head et al. 2005) • fossils, presumed to be colubroids, described from Cenomanian (approximately 95 mya) (Rage and Werner 1999; Rage et al. 2003) • lognormal prior applied to span complete range of the potential dates of divergence

2.2.5 Historical demography

Mismatch distributions were used to provide estimates of changes in population size in the program DnaSP v6 (Rozas et al. 2017). A recent range expansion is indicated by a unimodal mismatch distribution, whereas migration, subdivision, diminishing structure of population size or historical contraction of a population is indicated by a multimodal mismatch distribution and a ragged mismatch distribution is indicative of a lineage being widespread in the past (Burbrink et al. 2008; Nuñez et al. 2011). Fu's *Fs* (Fu 1997) was also calculated in the program DnaSP v6 to assess whether *Hemachatus* populations are at equilibrium (Nuñez et al. 2011). Positive values are indicative of a genetic bottleneck and negative values are indicative of recent range expansion under the assumption of neutrality (Fu 1997; Nuñez et al. 2011). This is because positive values are representative of the elimination of rare alleles within a population due to bottlenecks and negative values are representative of increased occurrence of rare alleles as a result of sudden expansion (Nuñez et al. 2011).

Fu's *Fs* and mismatch distributions can be used to provide information on past population dynamics, but cannot be used to provide information about the shape of population growth or decline over time (Nuñez et al. 2011). Thus, Extended Bayesian Skyline Plots (EBSP) were used to examine past population dynamics by estimating the shape of population growth of the phylogeographic lineages through time (Drummond et al. 2005; Burbrink et al. 2008; Guiher and Burbrink 2008; Nuñez et al. 2011). This technique was used to estimate the effective population size (N_e) of the lineages through time without the need to specify a demographic model prior to the analysis (Burbrink et al. 2008; Nuñez et al. 2011). Bayesian skyline plots were estimated for each lineage in BEAST2 by making use of a relaxed molecular clock and the appropriate model of sequence evolution. Clock rates were estimated from the sequence data and the divergence date estimates of the most recent common ancestor of all the haplotypes. The analyses were sampled every 1000^{th} repetition for 80 million generations; 10% of the initial samples were discarded as burnin. Tracer v1.7.1 (Rambaut et al. 2018) was used to estimate the ESS for the parameters. LogCombiner was used to combine the log output files. The demographic plots were visualised in R v3.6.0 (R Core Team 2019) using the plotEBSP.R script referred to in the EBSP tutorial available online (Heled and Drummond 2008; Heled 2016). The R

script can be downloaded from GitHub using the following link: https://github.com/CompEvol/beast2/tree/master/doc/tutorials/EBSP/scripts.

2.2.6 Distribution modelling

Distribution modelling was done in R v3.6.0 (R Core Team 2019) using current and past bioclimatic variables downloaded from the PaleoClim database (Brown et al. 2018). Models were created using Maxent v3.4.1 (Phillips et al. 2017) in R v3.6.0. The model was based on current distribution data for *H. haemachatus* that were collected during this study along with data supplied by the ADU. To control for observer bias, the spThin package was used to thin out data points in densely sampled areas (Aiello-Lammens et al. 2019). A Principal Components Analysis was run in R v3.6.0 to identify correlated variables using the prcomp function. Correlated variables were selected from a biplot, thus only uncorrelated variables were included in the model. The model was trained using climatic data for the years 1979-2013, at a resolution of 2.5 arc minutes (ca. 5 km), supplied by CHELSA (http://chelsa-climate.org/bioclim/), which incorporated the average data for bioclimatic variables. The past distribution for this species was then modelled using bioclimatic variables from the Pleistocene epoch. The Global Climate Models (GCMs) used to run the simulations are shown in Table 6. The R script can be viewed under supplementary data (script S2).

Table 6: Global Climate Models (GCM) source for data used for distribution modelling based on past climatic variables.

Time period	GCM and/or data source
Current (1979-2013)	CHELSA
Pleistocene: late Holocene, Meghalayan (4.2-0.3 ka)	Fordham et al. (2017)
Pleistocene: mid-Holocene, Northgrippian (8.326-4.2 ka)	Fordham et al. (2017)
Pleistocene: early Holocene, Greenlandian (11.7-8.326 ka)	Fordham et al. (2017)
Pleistocene: Younger Dryas Stadial (12.9-11.7 ka)	Fordham et al. (2017)
Pleistocene: Bølling-Allerød (14.7-12.9 ka)	Fordham et al. (2017)
Pleistocene: Heinrich Stadial 1 (17.0-14.7 ka)	Fordham et al. (2017)
Pleistocene: Last Glacial Maximum (ca. 22 ka)	Fick & Hijmans (2017)
Pleistocene: Last Interglacial Period (ca. 130 ka)	Otto-Bliesner et al. (2006)

2.3 Results

2.3.1 Sequence data

The aligned sequence data were analysed using four different datasets. These included a full mitochondrial dataset containing 91 ingroup sequences (1652 bp; Cyt-b 1127 bp, 16S 525 bp), a pruned mitochondrial dataset containing 48 unique ingroup sequences (1640 bp; Cyt-b 1115 bp, 16S 525 bp) (Table 7), as well as a nuclear dataset (1391 bp; PRLR 450 bp, RAG-1 941 bp) and a combined nuclear and mitochondrial dataset (3021 bp; Cyt-b 1127 bp, 16S 503bp, PRLR 450 bp, RAG-1 941 bp) both containing a subset of 20 ingroup sequences. The generated sequences will be submitted to GenBank.

Table 7: Sequences included in the pruned dataset as representative samples of the identical sequences listed. Identical sequences were identified from the full mitochondrial dataset. A total of 47 sequences were found to be identical to other sequences in the alignment.

Identical sequences	Representative sequence
MP01VR, GP01AdP, GP01DvE, GP01AN, GP02FM, GP01DJ, GP02PR, GP01NvdM, MP01FV, GP02NvdM, GP02DJ	MP01VR
KZN03ZB, KZN04ZB, KZN01MP	KZN03ZB
EC01CK, GKN_120, EC02CK	EC01CK
MP05FV, MP01SG, MP04FV, MP01CH, WC_DNA_1215, FS02JL, FS08JL, MP02VC, FS04JL, FS05JL, MP01NB, MP02FV, MP02NB, FS06JL, GP01PR	MP05FV
WC01FO, WC02FO	WC01FO
GP03AN, GP01FM	GP03AN
EC02MM, EC07MM, EC01MM, EC04MM, EC08MM, EC05MM	EC01MM
KZN02LB, KZN01LB, KZN01EB	KZN01LB
MP06FV, MP07FV, MP03FV	MP06FV
WC04HA, WC05HA, WC02PB, WC02HA	WC04HA
WC03HA, WC01HA, WC01PB	WC01HA

The mitochondrial alignments provided 48 variable sites and 33 parsimony informative sites. The nuclear subset and the combined nuclear and mitochondrial subset alignments provided 7 and 43 variable and 2 and 22 parsimony informative sites, respectively. The optimal substitution models for the generated datasets are shown in Table 8.

Table 8: Optimal substitution models for the combined and individual sequence datasets generated.

Dataset	Optimal substitution model
Full mitochondrial	HKY+I
Pruned mitochondrial	HKY+I
Nuclear subset	TPM3uf
Combined nuclear and mitochondrial subset	HKY+G
16S full	F81
Cyt-b full	HKY+I
16S pruned	F81
Cyt-b pruned	HKY+I
16S nuclear subset	HKY
Cyt-b nuclear subset	HKY
PRLR	HKY
RAG-1	F81

2.3.2 Phylogeography

The results for the phylogenetic and phylogeography analyses are focused on phylogenetic trees produced using the pruned mitochondrial dataset (Fig. 2, 3, supplementary material Fig. S1, S2). Trees produced using the full mitochondrial, combined nuclear and mitochondrial, and nuclear only datasets can be viewed under supplementary material (Fig. S3-S8).

Similar topologies were produced by the BI and ML analyses of the combined datasets, with no contradictory nodes noted (Fig. 2, 3, supplementary material Fig. S1-S8). The degree of support is defined in terms of posterior probability (PP) and Bootstrap (BS) values. High support is defined as PP>0.95 and BS>70, intermediate

support is defined as 0.95>PP>0.50 and 70>BS>50 and low/no support is defined as PP<0.50 and BS<50.

The pruned, combined nuclear and mitochondrial, and full datasets indicate that *Hemachatus haemachatus* appears to form a well-supported monophyletic lineage, which is further divided into two broad mitochondrial lineages (Fig. 2, 3, supplementary material Fig. S1-S6). The nuclear only dataset shows very little phylogenetic and phylogeographic resolution within *H. haemachatus* (supplementary material Fig. S7, S8).

The split between lineage 1 and lineage 2 is well-supported (PP = 1.00, BS = 100) (Fig. 2, 3). Lineage 1 consists of two further lineages, a Cape lineage and an inland lineage, with high support (PP = 1.00, BS = 88) (Fig. 2, 3). The Cape lineage contains coastal as well as inland samples from areas around Hermanus, Gansbaai and George in the Western Cape, Port Elizabeth, Grahamstown, Maclear and Mtatha in the Eastern Cape and Sutherland in the Northern Cape provinces (Fig. 4). The Cape coastal lineage shows high PP and intermediate BS support (PP = 0.99, BS = 59). The inland Eastern Cape lineage within the Cape lineage shows high PP and intermediate BS support (PP = 0.99, BS = 60), and the inland Northern Cape lineage shows high overall PP and BS support (PP = 1.00, BS = 87) (Fig. 2, 3). The second lineage in lineage 1 consists of inland samples from areas around Underberg, Kokstad and New Castle in KwaZulu-Natal, Standerton, Secunda and Kriel in Mpumalanga, Fouresburg, Bethlehem and Clarens in the Free State, and Centurion, Midrand, Randburg, Vanderbijlpark and Fourways in Gauteng (Fig. 2-4). This lineage shows intermediate BS, but low PP support (BS = 68, PP < 0.50) (Fig. 2, 3). However, within this lineage, the MP09FV, MP01VR (represents Mpumalanga and Gauteng samples, Table 7), MP06FV (represents Mpumalanga samples, Table 7) and MP01PJ group shows high PP and intermediate BS support (PP = 0.94, BS = 64) (Fig. 2, 3).

Lineage 2 contains inland samples from areas around Volksrust, Standerton, Lydenburg, Middelburg, Carolina, Wakkerstroom and Secunda in Mpumalanga, New Castle, Vryheid, Mooi River and Kokstad in KwaZulu-Natal, Bethlehem and Clarens in the Free State, Centurion and Vanderbijlpark in Gauteng, Potchefstroom in the North West province, Paterson and Hogsback in the Eastern Cape province and Mbabane

in Eswatini (Fig. 4). Within this lineage, the MP08FV and NW01AH group as well as the KZN01BG and SL01AN group show intermediate BS support (BS = 54 and BS = 68, respectively) (Fig. 2-3). In addition to this, the KZN02EB and KZN03LB group shows high overall support (PP = 0.99, BS = 72), and the KZN02BG and KZN01RV group shows high PP and intermediate BS support (PP = 0.96, BS = 52) (Fig. 2, 3).

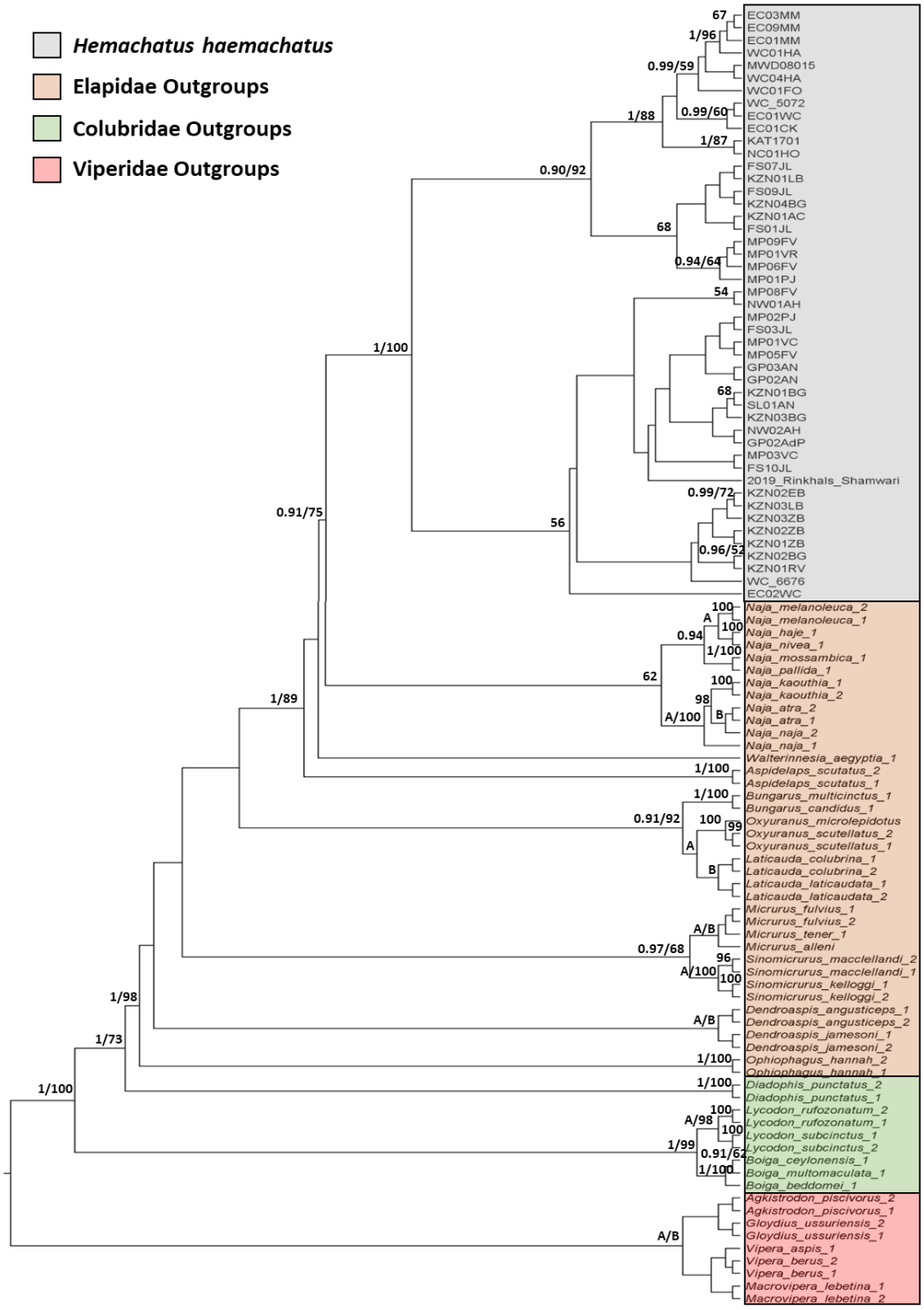

Figure 2: Bayesian phylogram for *Hemachatus haemachatus* overlaid with Bootstrap (BS) and/or posterior probability (PP) values showing nodes with intermediate to high support (BS>50, PP>90) for the pruned dataset with outgroups. Nodes shown with an "A" indicate that PP = 1 for all terminal nodes in the branch and nodes indicated with a "B" indicate that BS = 100 for all terminal nodes in the branch.

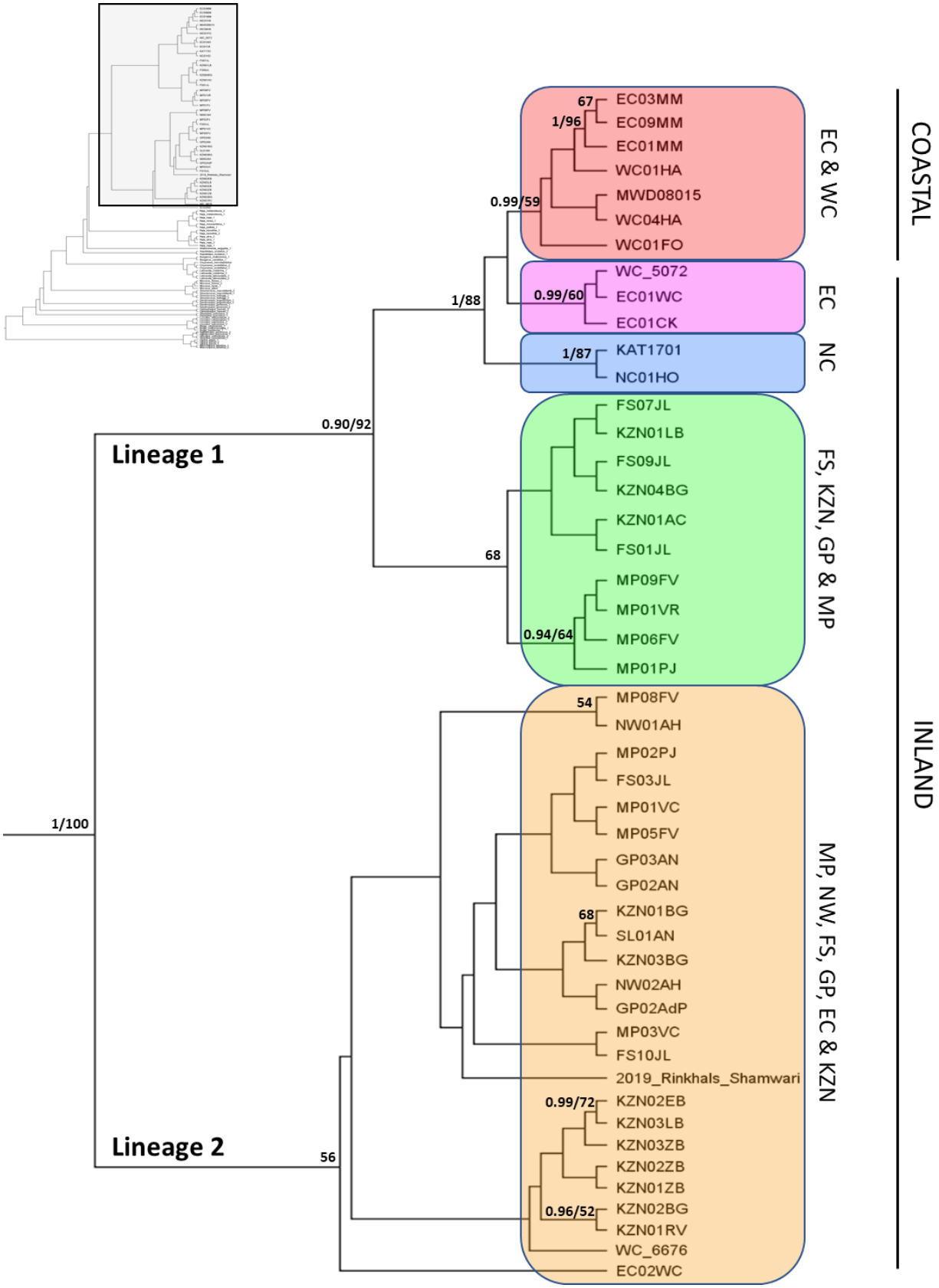

Figure 3: Bayesian phylogram for *Hemachatus haemachatus* overlaid with Bootstrap (BS) and/or posterior probability (PP) values showing nodes with intermediate to high support (BS>50, PP>90) for the pruned dataset. MP01VR represents Gauteng and Mpumalanga samples in lineage 1.

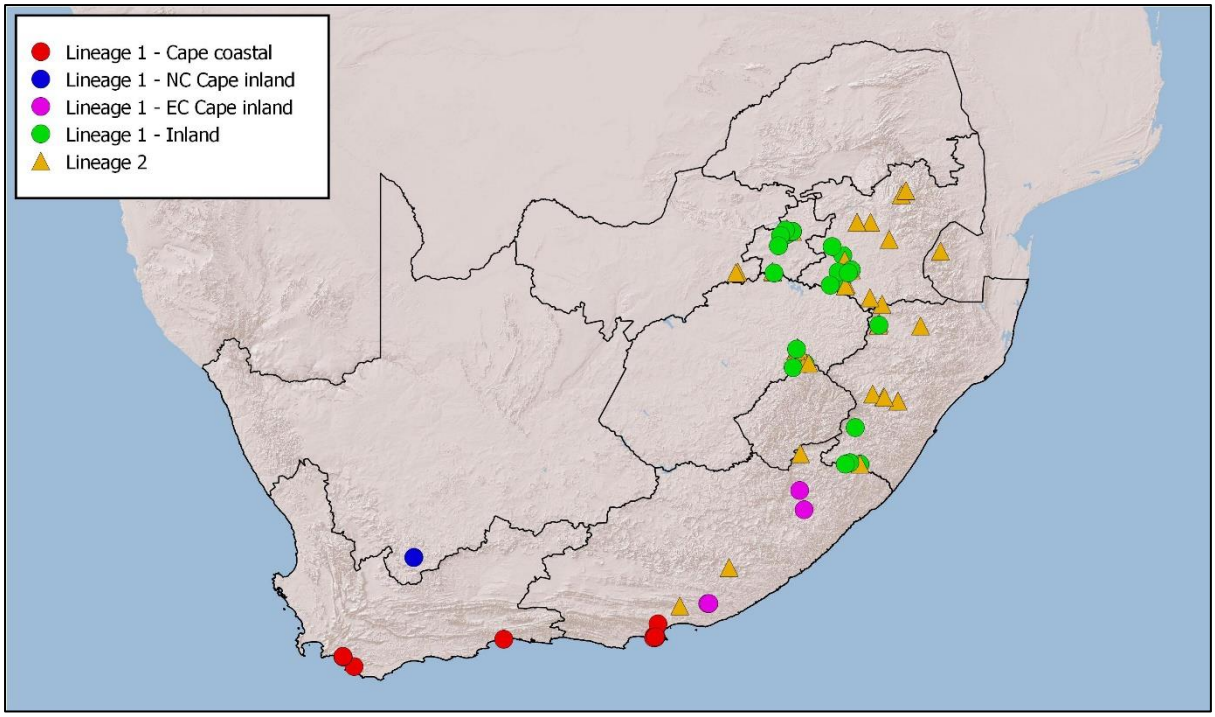

Figure 4: Map showing locality data for samples belonging to the lineages described for *Hemachatus haemachatus*. Samples shown in red, blue and purple belong to the Cape lineage and samples shown in green belong to the inland lineage in lineage 1. Samples shown in orange belong to lineage 2.

The topologies discussed above were further supported by the unrooted phylogenetic networks generated in SplitsTree for the pruned Cyt-b and 16S gene regions and the RAG-1 and PRLR subset gene regions (supplementary material Fig. S9, S12). Despite some of the branches showing intermediate to high BS support (50 < BS < 85), the two nuclear gene regions show limited phylogenetic and phylogeographic structure, as samples from the two lineages are intermixed within the topologies (supplementary material Fig. S11, S12). The 16S gene region shows phylogenetic and phylogeographic structure up to the lineage level, as samples belonging to lineages 1 and 2 are grouped together in a similar pattern as observed in the phylogenetic trees produced, however the grouping pattern differs from the tree topology at the individual sample level with samples from the inland and coastal lineages being intermixed (Fig. 2, 3 and supplementary material Fig. S10). The Cyt-b gene regions has the highest degree of phylogenetic and phylogeographic resolution producing a similar topology to the phylogenetic tree described above (Fig. 2, 3 and supplementary material Fig. S9). The split between lineages 1 and 2 is strongly supported in this case (BS = 100) and the individual samples group together in a

similar fashion described earlier with most branches showing intermediate to high support (50 < BS < 100) (Fig. 2, 3 and supplementary material Fig. S9).

Haplotype networks were generated using the full mitochondrial and nuclear only datasets. For the full dataset 30 haplotypes were identified with a haplotype diversity index of Hd = 0.905 (Fig. 5). The haplotypes grouped together in a similar pattern observed in the phylogenetic tree (Fig. 2, 3). Lineage 1 and lineage 2 are separated by 12 mutational steps between median vector 1 and 2 (mv1 and mv2). Most of the haplotypes are separated from one another by one mutational step, however there are some exceptions. Within lineage 2, haplotype 5 is separated from haplotype 1 by seven mutational steps. Within lineage 1, the Cape lineage is separated from the inland lineage by three mutational steps between mv2 and mv3 and by one mutational step between mv2 and haplotype 6. In the Northern Cape inland lineage of lineage 1, haplotype 12 is separated from mv3 and haplotype 26 by two mutational steps, respectively. In the Cape coastal lineage of lineage 1, haplotype 29 is separated from haplotype 30 by two mutational steps, and haplotype 3 is separated from haplotype 30 by three mutational steps. For the nuclear only dataset 10 haplotypes were identified with a haplotype diversity index Hd = 0.884 (Fig. S13). All haplotypes were separated by a single mutational step. Samples belonging to lineages 1 and 2 were intermixed between the haplotypes (supplementary material Fig. S13).

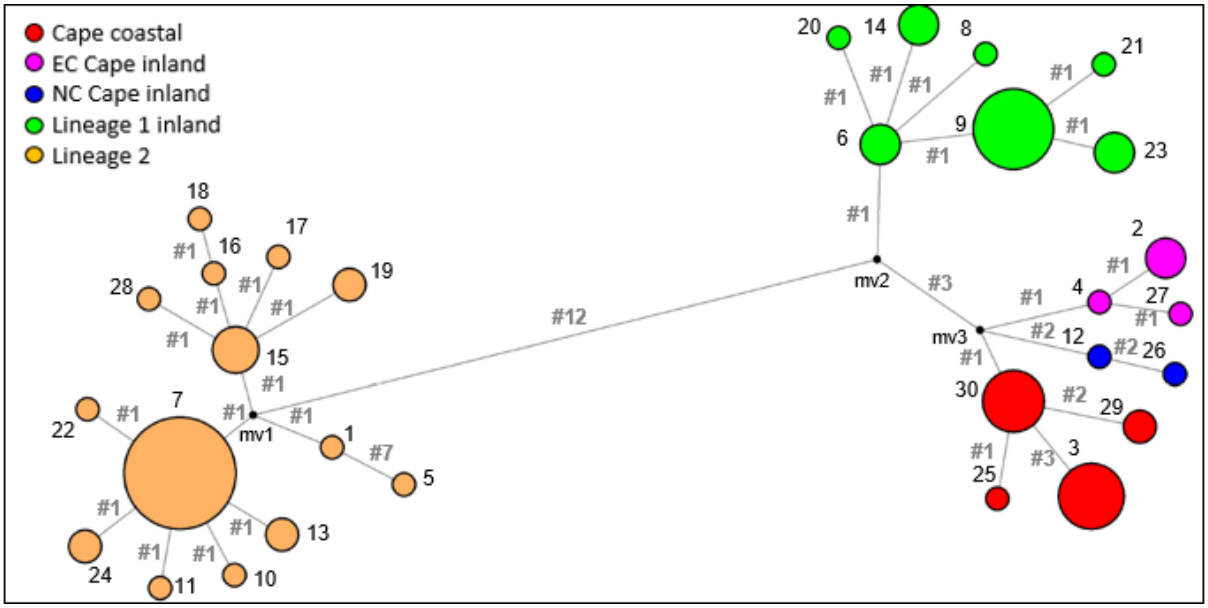

Figure 5: Haplotype network of *Hemachatus haemachatus* generated from Cyt-b and 16S gene regions for 91 samples. Black circles indicate median vectors 1-3. The coloured circles indicate the haplotypes and the haplotype number is shown next to each circle. The area of the coloured circles is proportional to the number of samples present in each haplotype. Lines connecting the haplotypes are not proportional to the number of mutations separating each haplotype. The "#" followed by a number indicates the number of mutations separating each haplotype. Samples included in each haplotype are as follows: (1) 2019_Rinkhals_Shamwari; (2) EC01CK, EC02CK, GKN_120; (3) EC01MM, EC02MM, EC03MM, EC04MM, EC05MM, EC07MM, EC08MM, EC09MM; (4) EC01WC; (5) EC02WC; (6) FS01JL, FS07JL, KZN01AC; (7) FS02JL, FS03JL, FS04JL, FS05JL, FS06JL, FS08JL, FS10JL, GP01FM, GP01PR, GP03AN, KZN03BG, MP01CH, MP01NB, MP01SG, MP02FV, MP02NB, MP02PJ, MP02VC, MP03VC, MP04FV, MP05FV, NW02AH, WC_DNA_1215; (8) FS09JL; (9) GP01AdP, GP01AN, GP01DJ, GP01DvE, GP01NvdM, GP02DJ, GP02FM, GP02NvdM, GP02PR, MP01FV, MP01VR, MP09FV; (10) GP02AdP; (11) GP02AN; (12) KAT1701; (13) KZN01BG, SL01AN; (14) KZN01EB, KZN01LB, KZN02LB; (15) KZN01MP, KZN02ZB, KZN03ZB, KZN04ZB (16) KZN01RV; (17) KZN01ZB; (18) KZN02BG; (19) KZN02EB, KZN03LB; (20) KZN04BG; (21) MP01PJ; (22) MP01VC; (23) MP03FV, MP06FV, MP07FV; (24) MP08FV, NW01AH; (25) MWD08015; (26) NC01HO; (27) WC_5072; (28) WC_6676; (29) WC01FO, WC02FO; (30) WC01HA, WC01PB, WC02HA, WC02PB, WC03HA, WC04HA, WC05HA.

2.3.3 Divergence dating

The divergence dating analysis was run for the pruned mitochondrial, combined nuclear and mitochondrial subset, and full mitochondrial datasets. The nuclear dataset was not included due to a lack of sequence data available for species included for calibration purposes. The ESS > 200 for all parameters for the combined runs for the pruned mitochondrial dataset after four runs. This point was reached after six runs for the combined nuclear and mitochondrial dataset. The full mitochondrial dataset had

low overall ESS values (100 < ESS < 200) for a number of the parameters, most likely due to excessive branch swapping caused by the presence of a high number of duplicate sequences. The pruned mitochondrial and combined nuclear and mitochondrial datasets resulted in similar, relatively recent divergence date estimates (Fig. 6-8 and supplementary material Table S1, Fig. S14). The full mitochondrial dataset resulted in overestimated divergence date estimates compared to the abovementioned datasets (supplementary material Table S1, Fig. S15).

Lineages 1 and 2 diverged from one another during the Quaternary period, within the Pleistocene epoch approximately 2.08 mya (Fig. 7, 8). The 95% confidence limits show that the timing of this event ranged between the Pleistocene and Pliocene epochs. Lineage 1 diverged approximately 1.15 mya during the Calabrian stage in the Pleistocene epoch. The 95% confidence limits place this event between the Gelasian and Ionian stages. Lineage 2 diverged more recently, during the Calabrian stage, approximately 0.84 mya. The confidence limits place this event between the Calabrian and Ionian stages. The Cape lineage (inland and coastal) and the inland lineage of lineage 1 diverged from one another approximately 0.71 mya, during the Ionian stage, with confidence limits ranging between this stage and the Calabrian stage. The Cape coastal lineage diverged during the Ionian stage, approximately 0.43 mya, with confidence limits between the Calabrian and Tarantian stages. The Eastern Cape group of the Cape inland lineage diverged approximately 0.27 mya, and the Northern Cape group of the Cape inland lineage diverged approximately 0.22 mya. Confidence limits for both Cape inland groups ranged between the Tarantian and Ionian stages. The inland group of Lineage 1 containing samples from the north-eastern parts of South Africa diverged during the Ionian stage, approximately 0.64 mya, with confidence limits between the Calabrian and Tarantian stages. The KwaZulu-Natal group within lineage 2 diverged during the Ionian stage, approximately 0.51 mya, with confidence limits falling within this stage and the Tarantian stage. The inland group within lineage 2 containing samples from the north-eastern parts of South Africa also diverged during the Ionian stage, approximately 0.58 mya, with confidence limits falling within this stage and the Calabrian stage.

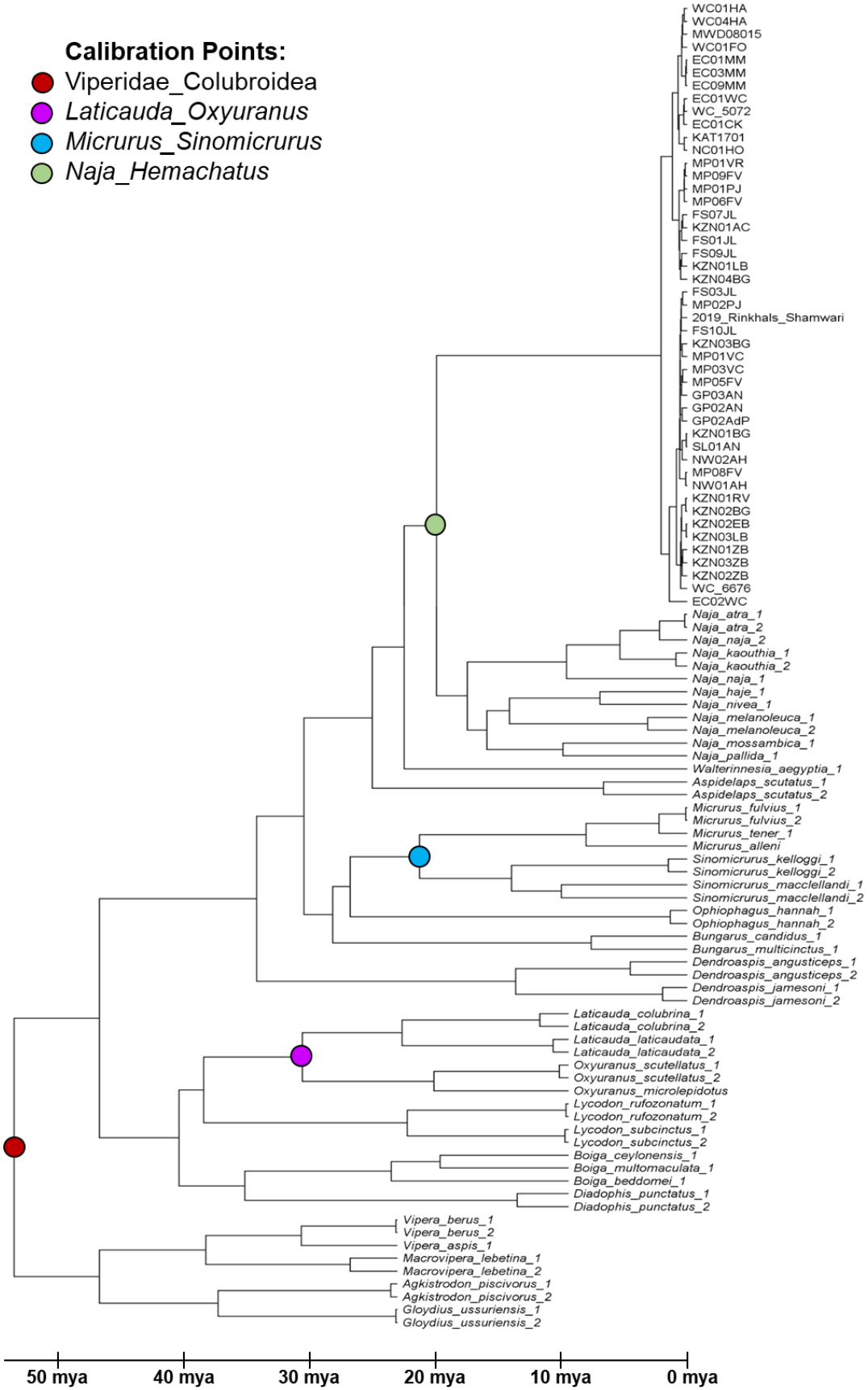

Figure 6: Dated phylogeny of *Hemachatus haemachatus* generated from the pruned mitochondrial dataset showing calibration points.

Node	Age (mya)	95% HPD (mya)
	2.08	0.79, 5.03
	1.15	0.34, 2.39
	0.84	0.18, 1.45
	0.71	0.20, 1.65
	0.43	0.09, 0.8
	0.27	0.02, 0.53
	0.22	0.01, 0.44
	0.64	0.11, 1.05
	0.24	0.02, 0.47
	0.58	0.13, 1.03
	0.51	0.09,0.91
	0.07	1×10^{-6}, 0.18
	0.12	9.97×10^{-4}, 0.27
	0.06	3×10^{-7}, 0.17

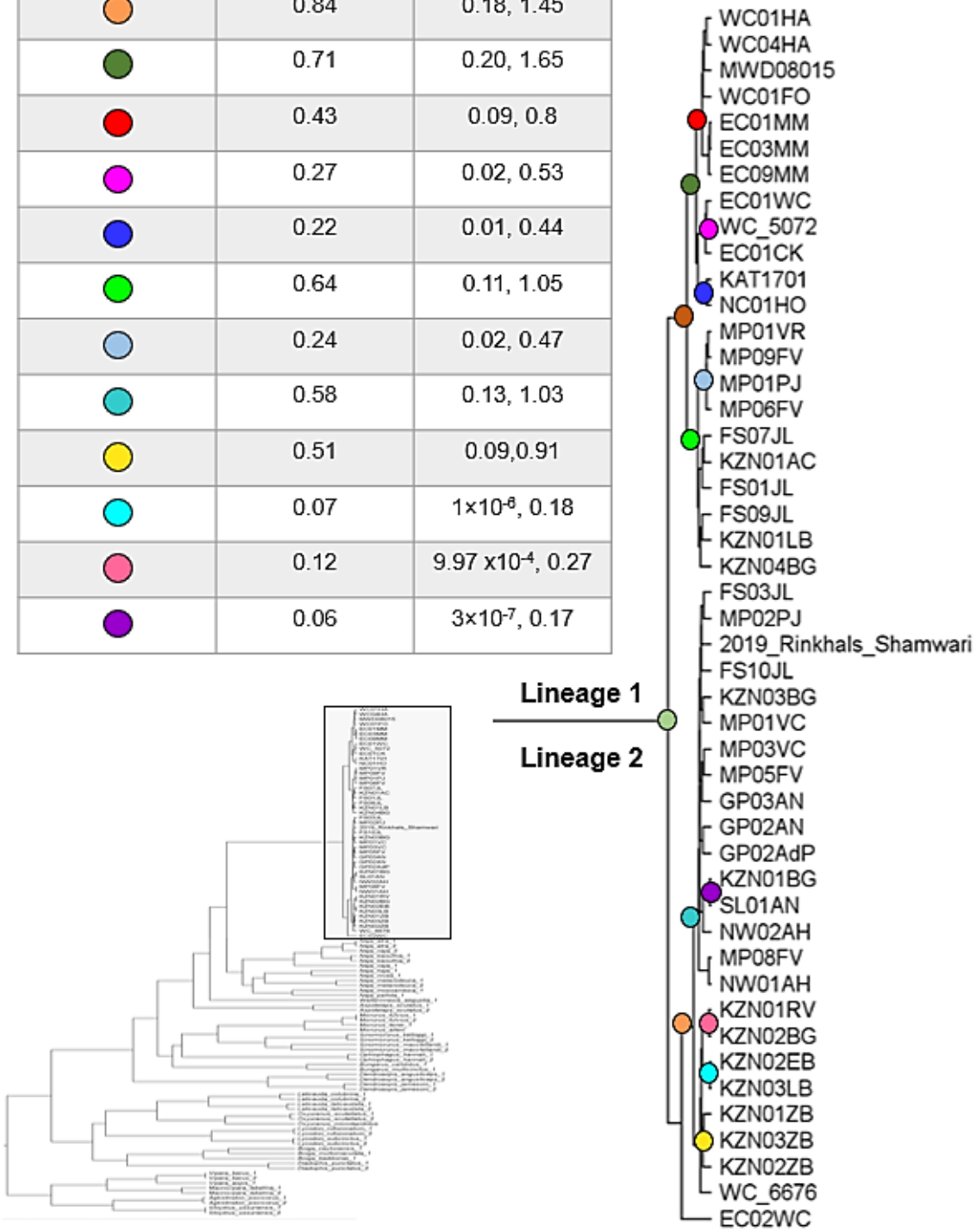

Figure 7: Subset of the dated phylogeny of *Hemachatus haemachatus* generated from the pruned mitochondrial dataset showing divergence date estimates of supported nodes.

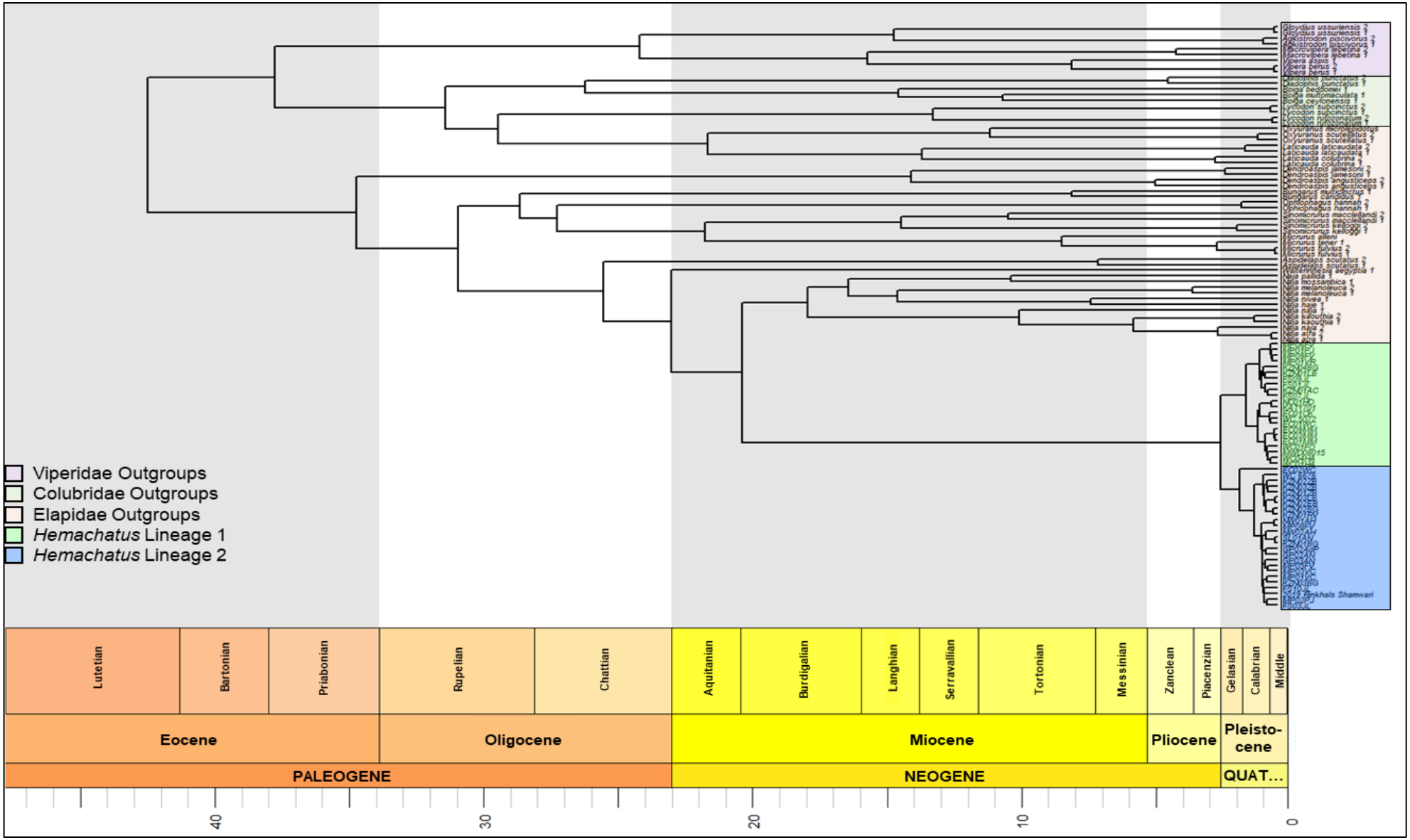

Figure 8: Dated phylogeny of *Hemachatus haemachatus* generated from the pruned mitochondrial dataset indicating in which time period and epoch the lineages are estimated to have diverged from one another.

2.3.4 Historical demography

Mismatch distributions were run using the population growth-decline model for the full mitochondrial, pruned mitochondrial, nuclear only and combined nuclear and mitochondrial datasets. The full mitochondrial, pruned mitochondrial and combined nuclear and mitochondrial datasets produced ragged mismatch distributions, with a raggedness index of r = 0.024 for the full and combined nuclear and mitochondrial datasets and r = 0.012 for the pruned mitochondrial dataset (Fig. 9 and supplementary material Fig. S16, S17). The nuclear only dataset produced a unimodal mismatch distribution with r = 0.079 (Fig. 10).

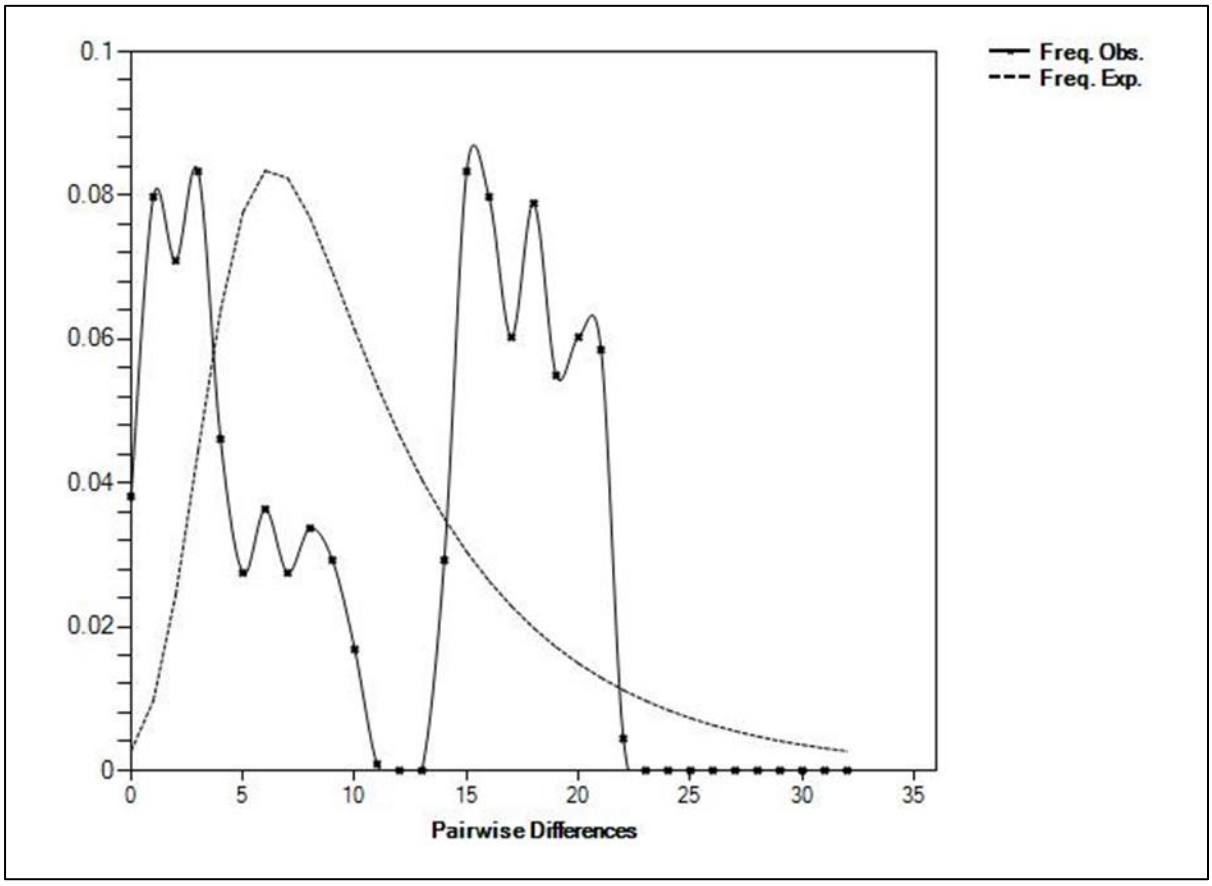

Figure 9: A ragged mismatch distribution produced for *Hemachatus haemachatus* under the population growth-decline model for the pruned mitochondrial dataset. The number of pairwise differences are shown on the x-axis and the frequency of the pairwise comparisons are shown on the y-axis. The dotted line indicates the frequency expected under the population growth-decline model and the solid line indicates the observed frequencies.

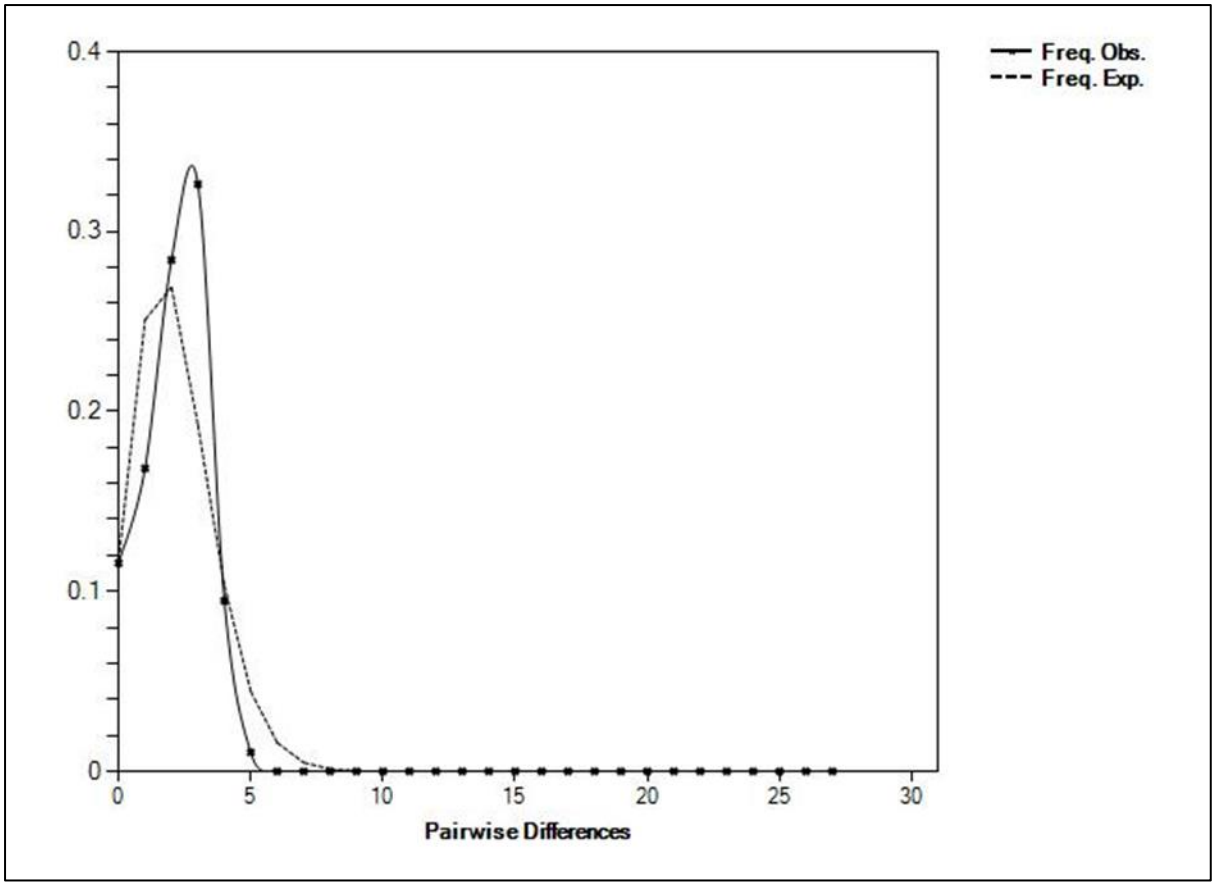

Figure 10: A unimodal mismatch distribution produced for *Hemachatus haemachatus* under the population growth-decline model for the nuclear dataset. The number of pairwise differences are shown on the x-axis and the frequency of the pairwise comparisons are shown on the y-axis. The dotted line indicates the frequency expected under the population growth-decline model and the solid line indicates the observed frequencies.

Fu's Fs was calculated for the full mitochondrial, combined nuclear and mitochondrial, nuclear only and pruned mitochondrial datasets. All four datasets produced negative values for Fu's Fs. The full mitochondrial and combined nuclear and mitochondrial datasets had a value of Fu's Fs = -2.427, whereas the pruned mitochondrial dataset had a value of Fu's Fs = -8.039, and the nuclear dataset had a value of Fu's Fs = -4.250.

Only the pruned mitochondrial dataset was included in the EBSP analysis. The nuclear only and combined nuclear and mitochondrial datasets were excluded due to small sample sizes and a lack of calibration information for the nuclear gene regions, as mentioned earlier, preventing the estimation of a relatively accurate clock rate for these gene regions. The full mitochondrial dataset was excluded as it contained a

large number of duplicate sequences and resulted in low overall ESS values. The ESS > 200 for all parameters of the pruned mitochondrial dataset after eight runs combined. The EBSP shows a general increasing trend in the effective population size of *H. haemachatus* which started approximately 100 000 years ago (Fig. 11A, 11B). However, a decreasing trend is observed in the slope of the median line as well as in the lower 95% Central Posterior Densities (CPD) moving towards time zero (Fig. 11A, 11B). When considering the histogram of tree events, this decrease in the effective population size and density starts at around 40 000 years ago (Fig. 11C). The histogram shows a spike at time zero, which is caused by the populations having a point of change at this time and will not be considered in the discussion (Heled and Drummond 2008; Heled 2016) (Fig. 11C).

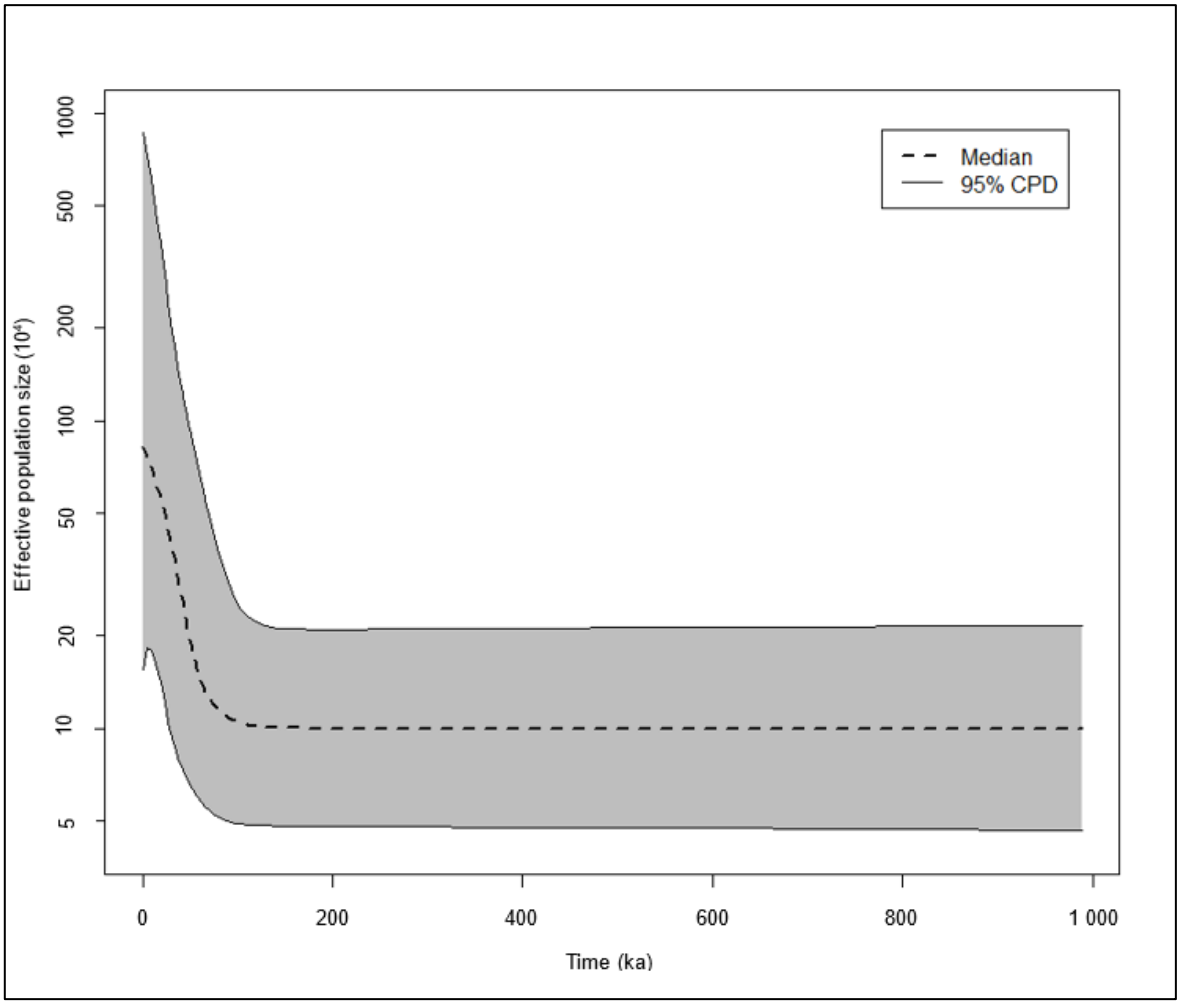

Figure 11A: Extended Bayesian Skyline Plot showing the change in effective population size of *Hemachatus haemachatus* populations. An interval of 95% of the Central Posterior Densities (CPD) of the population histories sampled in the posterior is represented by the grey area. The plot was generated using the pruned mitochondrial dataset.

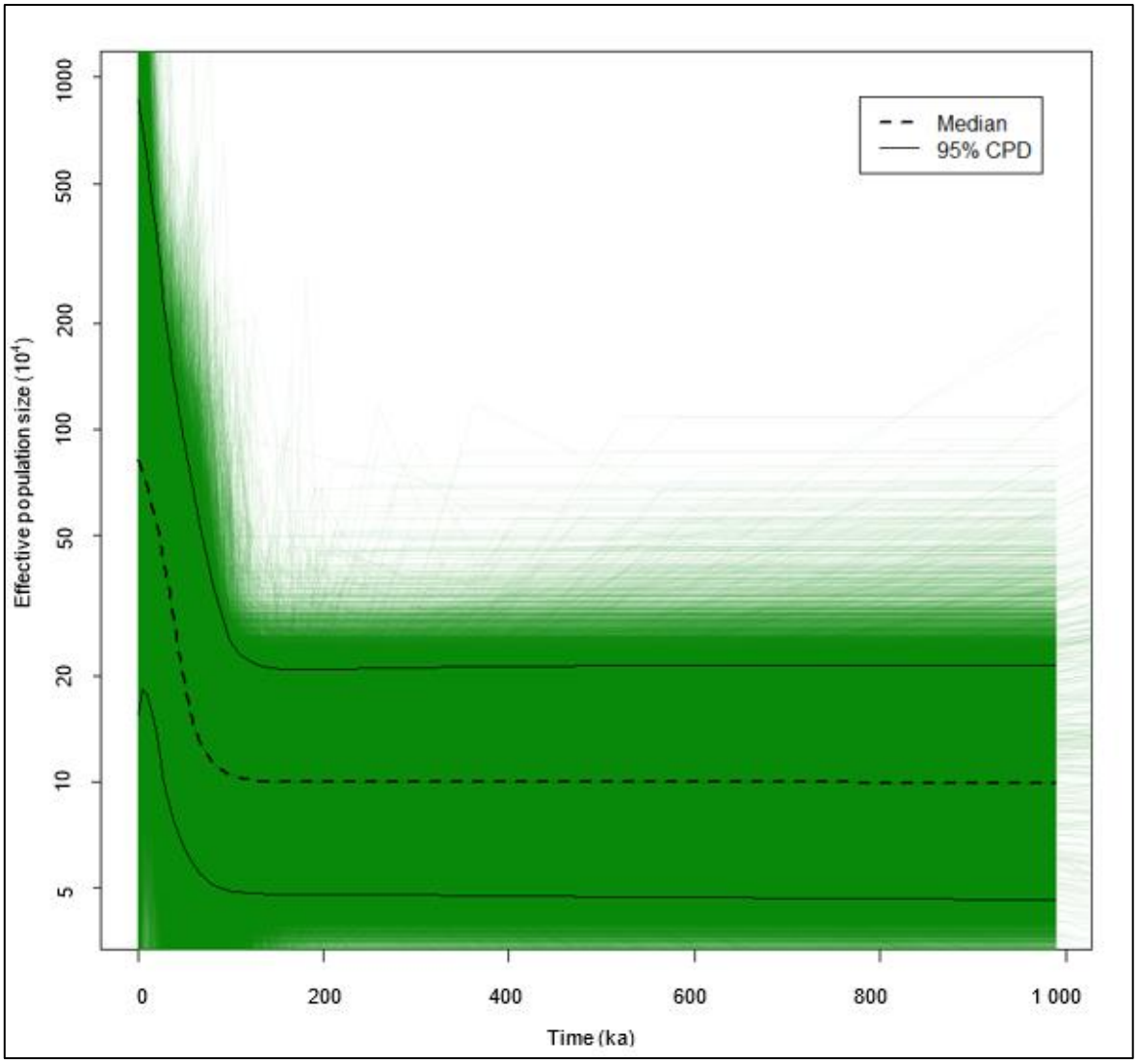

Figure 12: Extended Bayesian Skyline Plot showing the change in effective population size of *Hemachatus haemachatus* populations for individual population trajectories. The green lines represent individual population trajectories, showing the full view of the samples sampled in the posterior. An interval of 95% of the Central Posterior Densities (CPD) of the population histories sampled in the posterior is represented by the area between the two solid lines. The plot was generated using the pruned mitochondrial dataset.

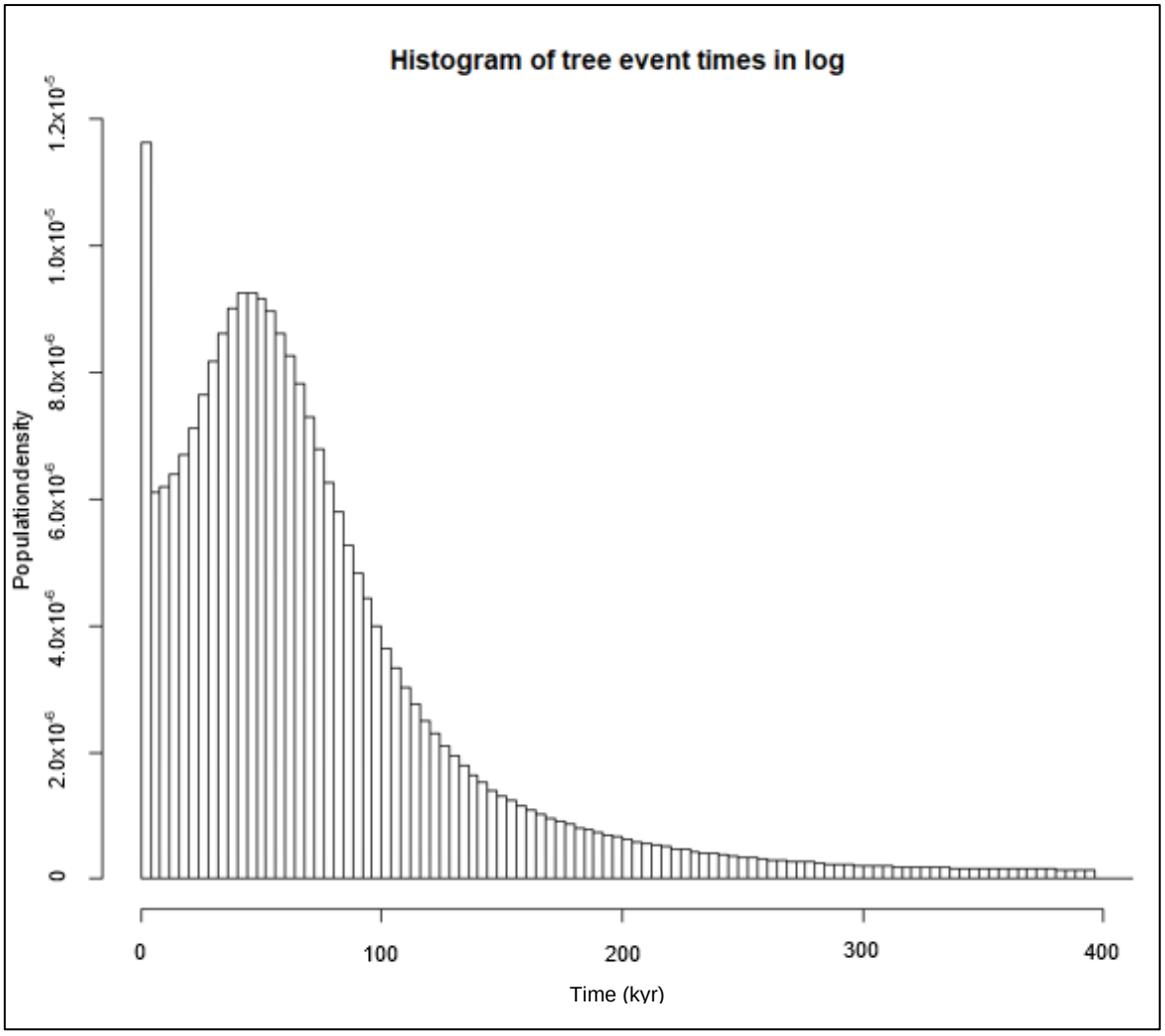

Figure 13: Histogram showing the change in population density of *Hemachatus haemachatus* populations for 95% of tree event times. The plot was generated using the pruned mitochondrial dataset. The spike at time zero is caused by the populations having a point of change at this time.

2.3.5 Distribution modelling

The Maxent model was trained using 116 presence records and determined using 10032 background and presence points. Bioclimatic variables BIO1, BIO3, BIO5, BIO7, BIO10, BIO12, BIO14, BIO15, BIO16 and BIO18 were selected to run the model. These variables represented the annual mean temperature, isothermality, maximum temperature of the warmest month, the temperature annual range, the mean temperature of the warmest quarter, annual precipitation, precipitation of driest month, precipitation seasonality, precipitation of the wettest quarter and precipitation of the warmest quarter, respectively. BIO3 was estimated to provide the highest relative contribution of the environmental variables to the model (45.1%). This was followed

by BIO1 (27.9%), BIO7 (10.1%), BIO14 (7.0%), BIO18 (4.5%), BIO12 (2.8%), BIO5 (1.0%), BIO15 (0.9%), BIO16 (0.8%) and BIO10 (0.0%). Areas with a high probability of *H. haemachatus* occurring there are shown in green, medium probabilities are shown in yellow, and low probabilities are shown in red on the modelled distribution maps.

The predicted distribution in the model covers the presence records provided (Fig. 12). The omission rate is close to the predicted omission and the model has high AUC (area under curve) scores for the receiver operating characteristic (ROC) (Fig. 13A, 13B). Thus, the model can be considered a good representation of *H. haemachatus* distribution based on climatic variables.

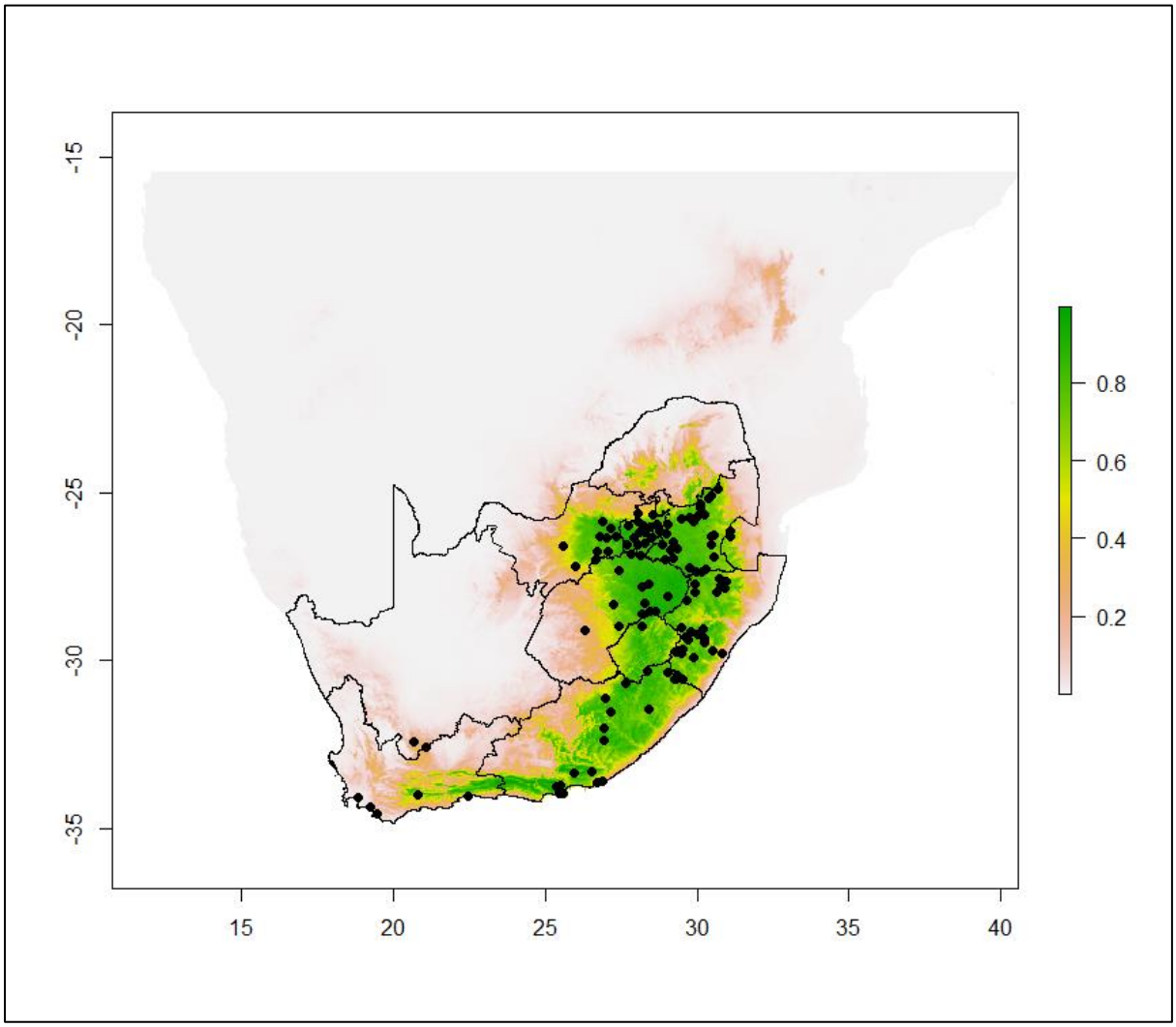

Figure 12: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the current GCM. Distribution points are shown as black dots. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

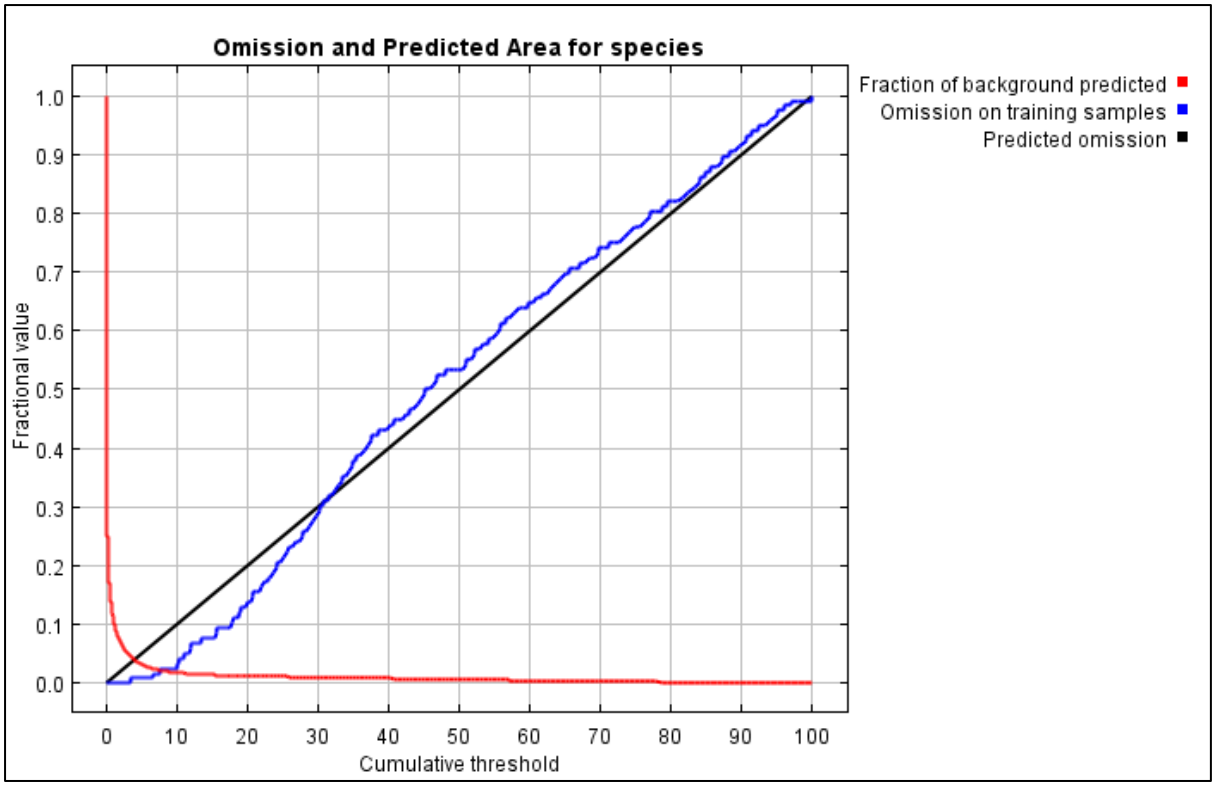

Figure 13A: Omission rate and the predicted area of training samples as a function of the cumulative threshold. The omission rate is calculated on the *Hemachatus haemachatus* presence records used for training the Maxent model.

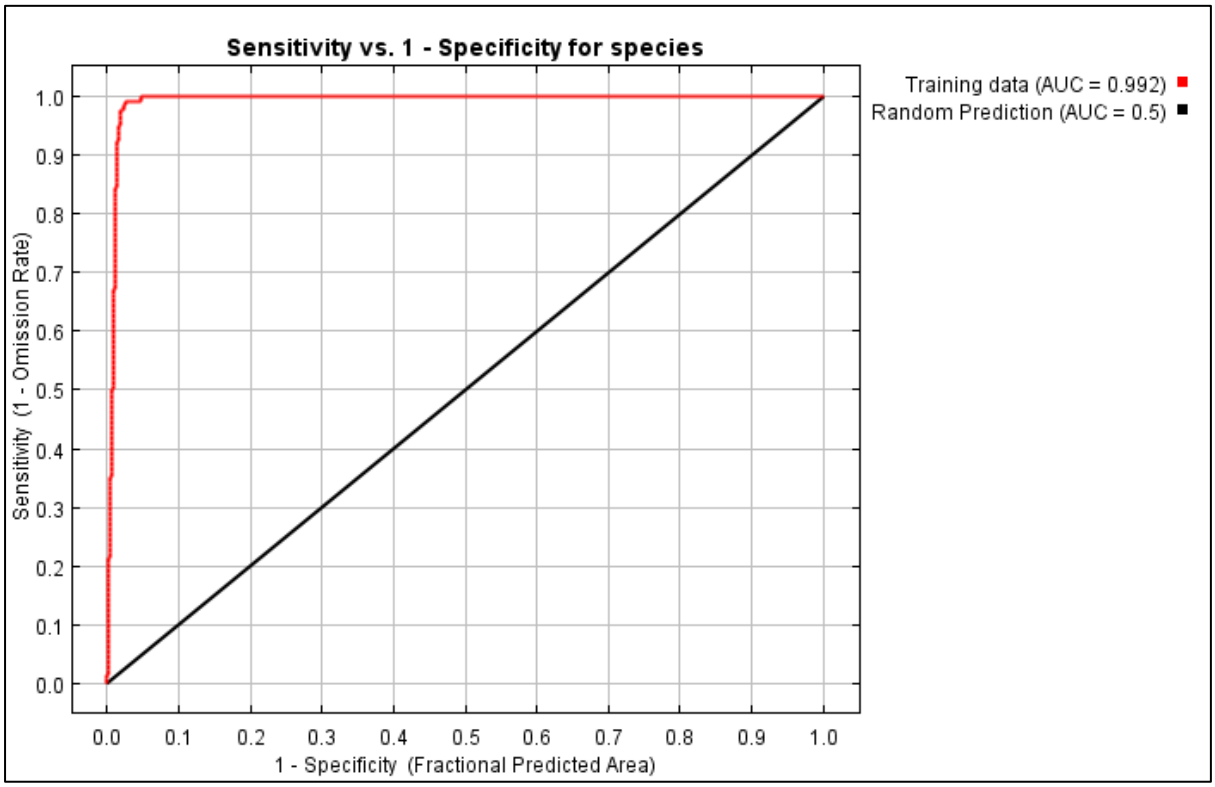

Figure 13B: Receiver operating characteristic (ROC) curve for the *Hemachatus haemachatus* presence records used for training the Maxent model.

During the Last Interglacial Period (LIG), approximately 130 000 years ago, *H. haemachatus* appeared to be relatively widespread throughout the eastern regions of South Africa. There also appeared to be two to three potentially isolated populations along the southern coast of South Africa (Fig. 14).

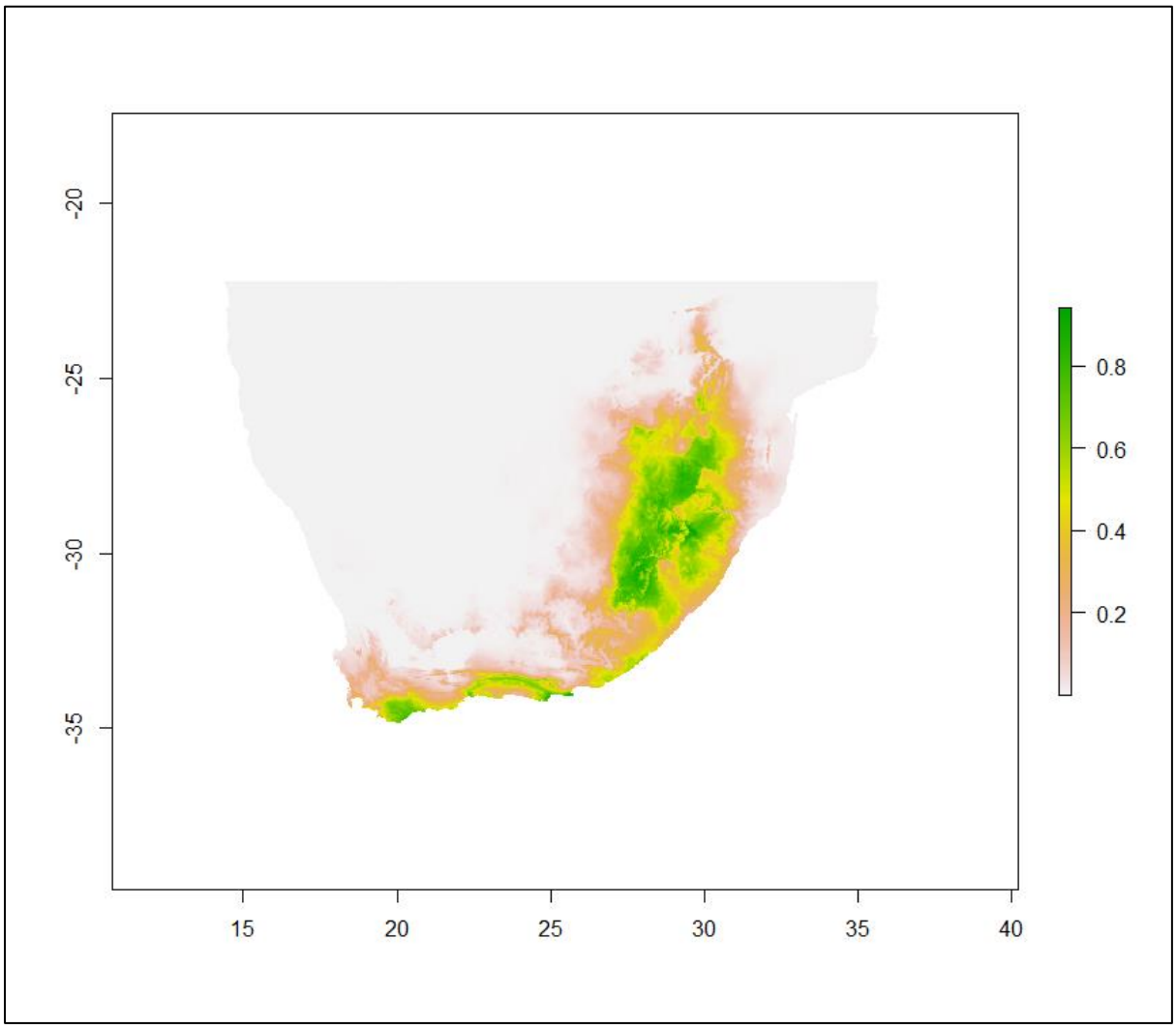

Figure 14: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the Last Interglacial Period. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

During the Last Glacial Max (LGM), approximately 22 000 years ago, the CCSM4, MIROC-ESM and MPI-ESM-P GCMs indicate that the populations of *H. haemachatus* expanded inland towards the Northern Cape region of South Africa (Fig. 15A, 15B, 15C). The MPI-ESM-P GCM predicts a larger distribution in the Northern Cape compared to the CCSM4 and MIROC-ESM GCM's, although all three models

indicate that suitable climatic conditions were present in the same general area in the Northern Cape for these snakes to survive.

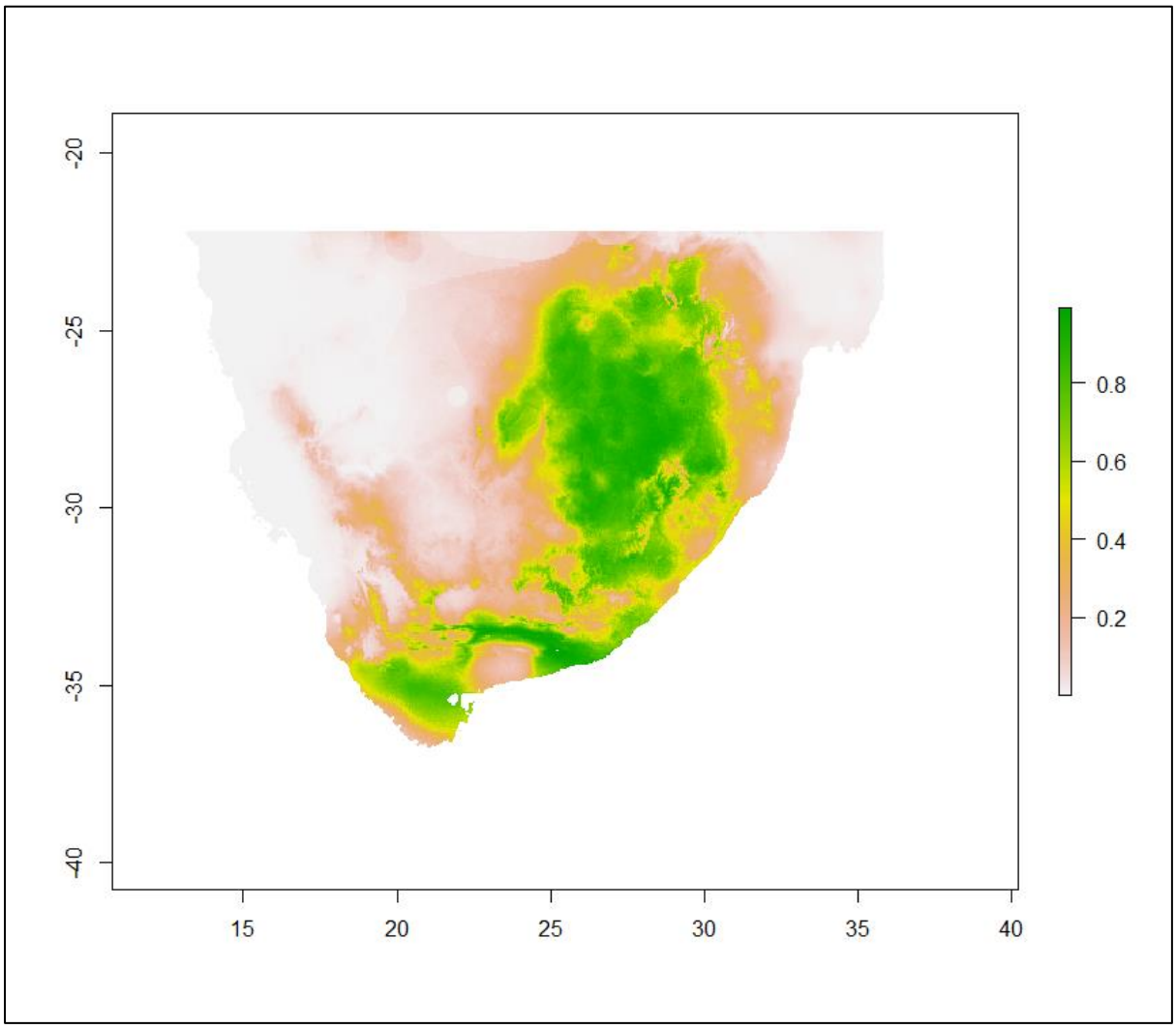

Figure 14A: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the CCSM4 Global Climate Model of the Last Glacial Maximum. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

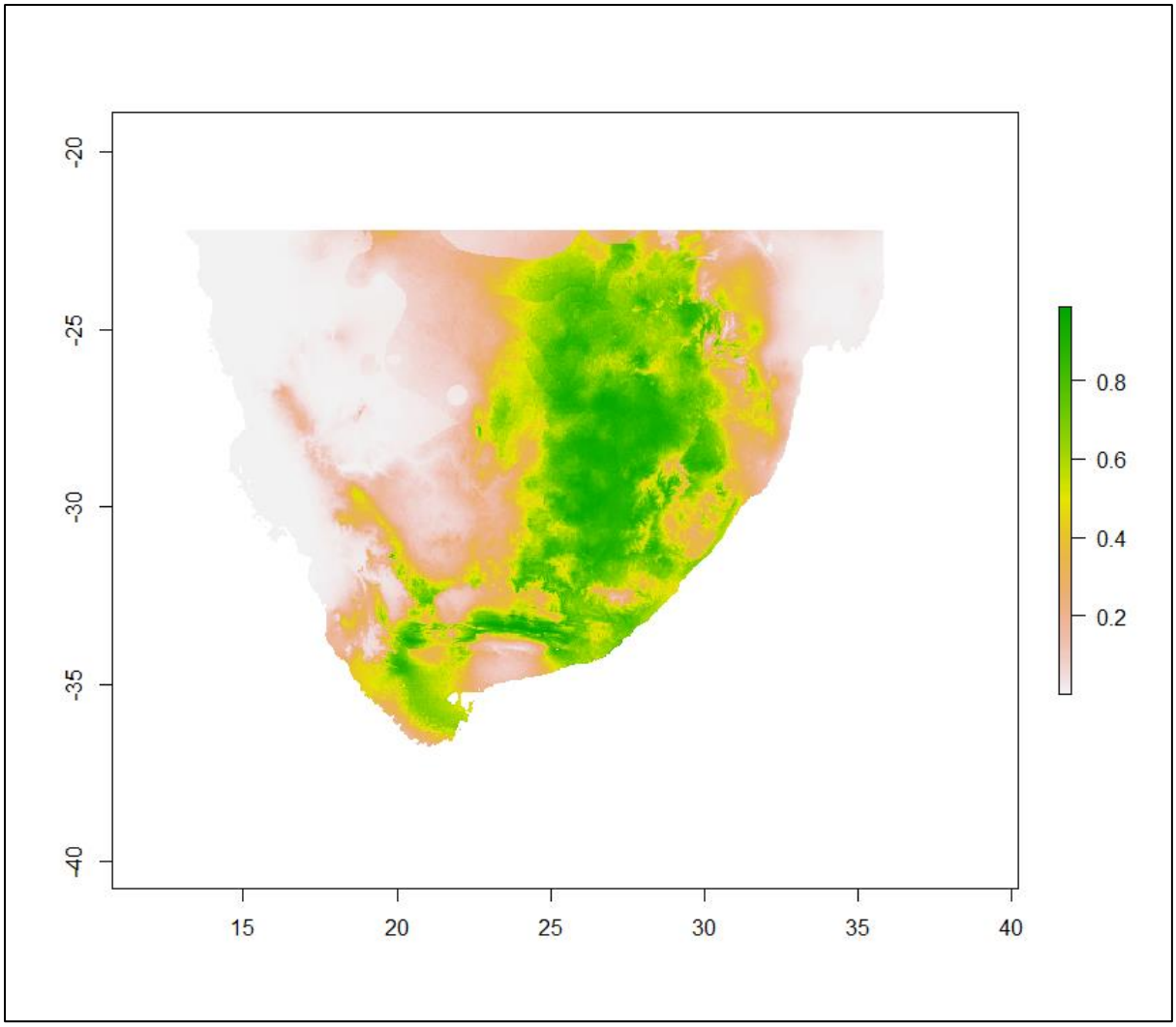

Figure 15B: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the MPI-ESM-P Global Climate Model of the Last Glacial Maximum. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

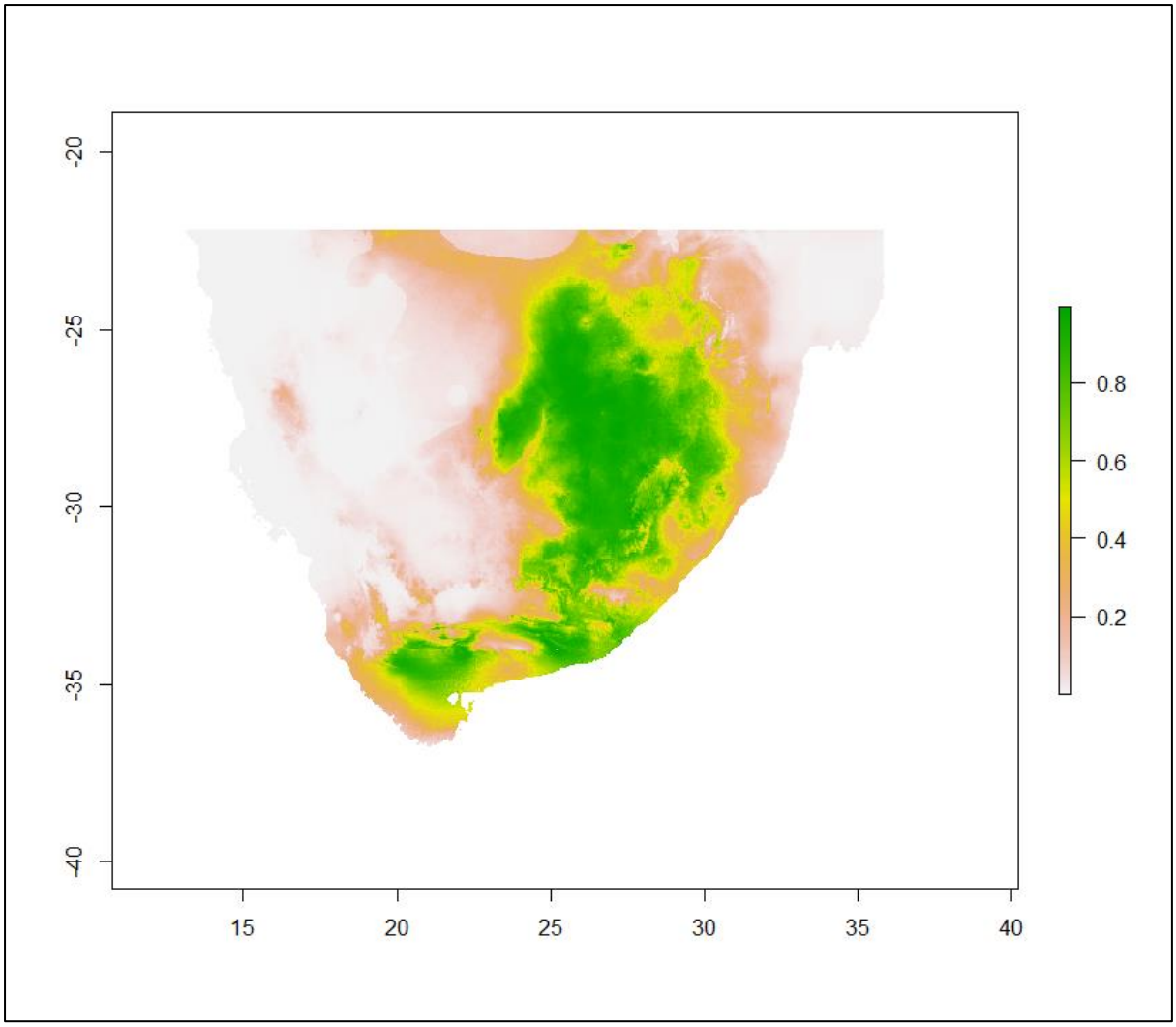

Figure 15C: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the MIROC-ESM Global Climate Model of the Last Glacial Maximum. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

During the Heinrich Stadial 1 (HS1) event, approximately 17 000 – 14 700 years ago in the Tarantian Stage of the Pleistocene epoch, *H. haemachatus* populations appeared to contract in their distribution (Fig. 16). The Northern Cape population appeared isolated from the populations in the southern and Eastern Cape. Populations from the southern and eastern parts of the Cape appeared more widespread, while populations from the eastern inland regions appeared to have contracted in their distribution.

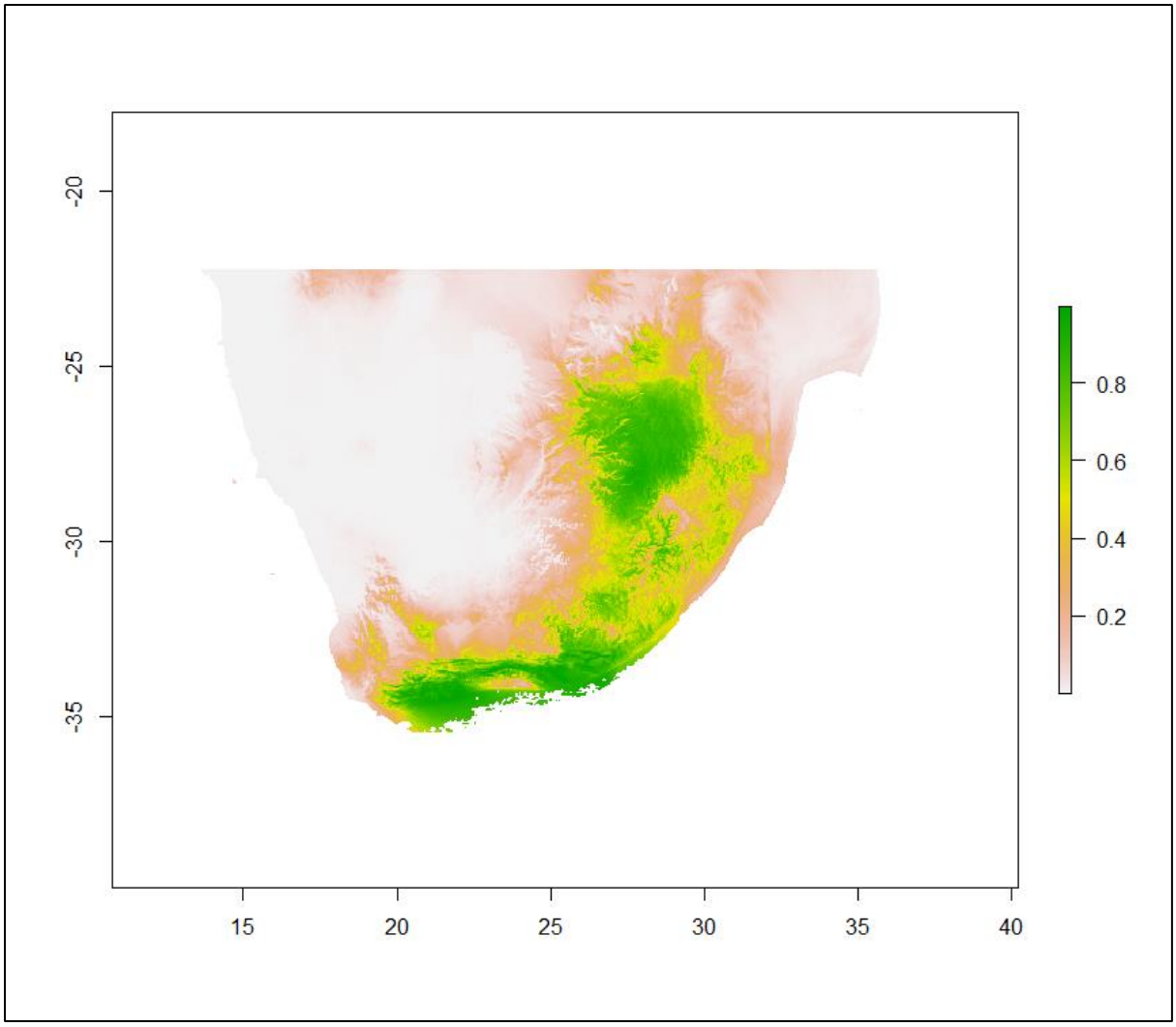

Figure 16: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the Heinrich Stadial 1 event during the Tarantian stage in the Pleistocene epoch. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

Climatic conditions during the Bølling-Allerød event, approximately 14 700 – 12 900 years ago in the Tarantian stage of the Pleistocene epoch, areas with suitable climatic conditions for *H. haemachatus* populations appeared to decrease in size, resulting in further contraction of the populations (Fig. 17). The model indicates that most of the populations were linked through a continuous distribution and that the southern Cape provided the largest continuous area with suitable climatic conditions for this species to occur there, while such areas in the eastern regions of South Africa were more fragmented.

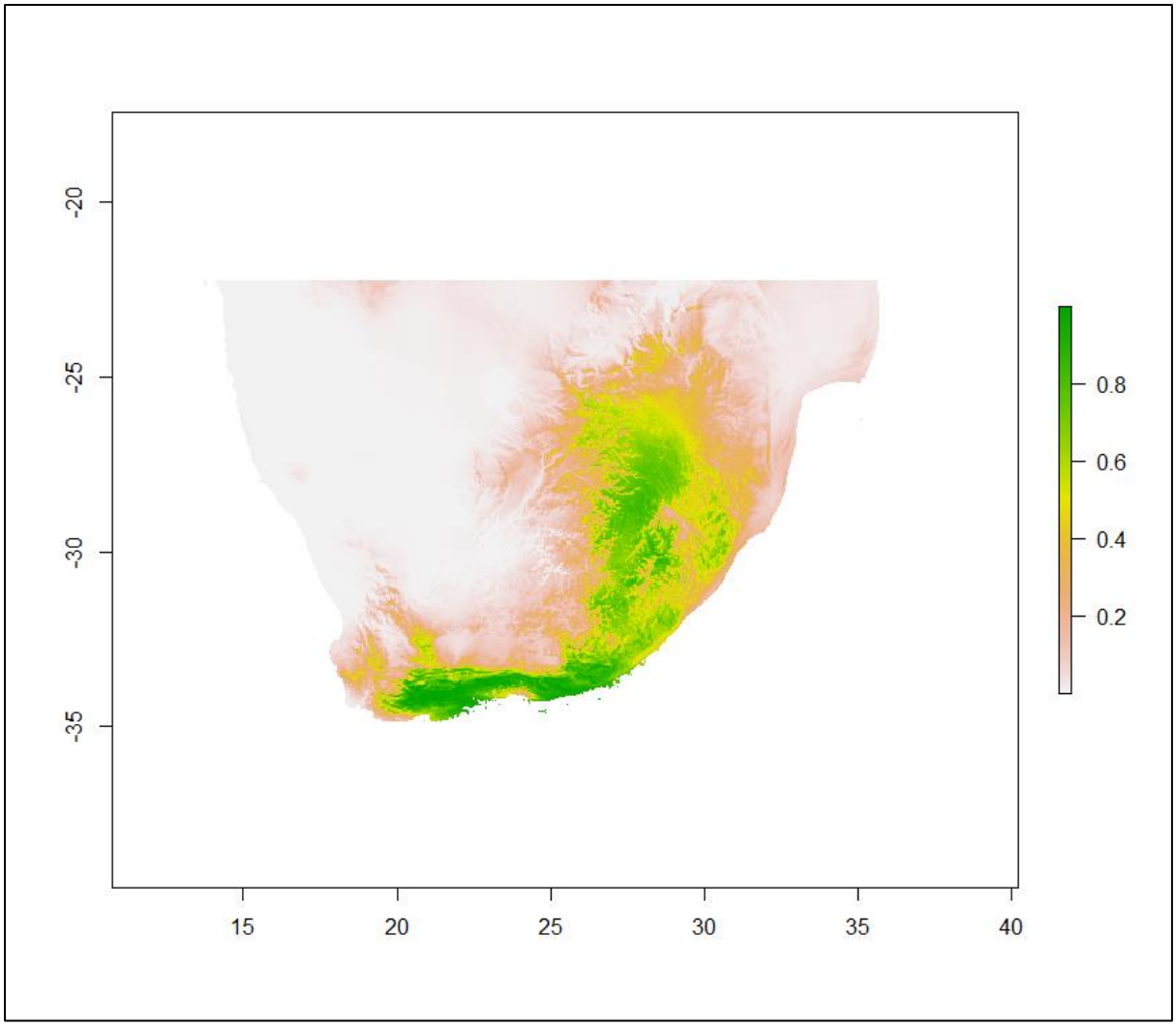

Figure 17: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the Bølling-Allerød event during the Tarantian stage in the Pleistocene epoch. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

The model from the Younger Dryas Stadial, approximately 12 900 – 11 700 years ago in the Tarantian stage, indicated a similar distribution to the one modelled using data from the Bølling-Allerød event (Fig. 18). Although, in this case, the distribution of the eastern inland populations appeared to have expanded slightly in size, especially within the KwaZulu-Natal region.

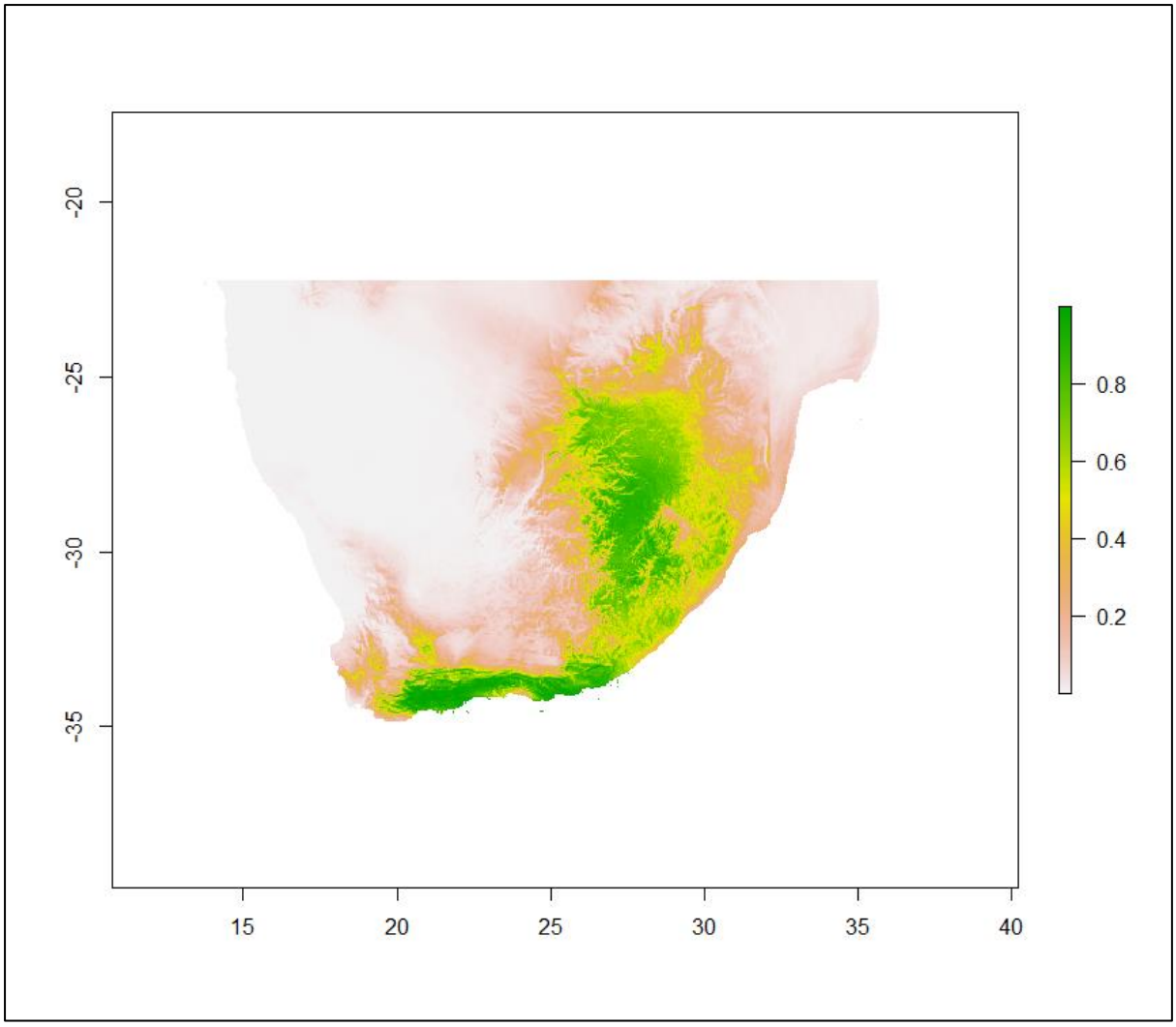

Figure 18: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the Younger-Dryas Stadial event during the Tarantian stage in the Pleistocene epoch. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

During the early-Holocene (11 700 – 8 326 years ago), the climatic conditions resulted in a contraction of suitable habitat in the southern Cape region of South Africa, indicated through larger areas shown in orange and/or red on the modelled distribution map (Fig. 19). Along with this, an increase in suitable habitat was observed in the eastern inland regions of South Africa moving towards the coast, indicated through larger areas shown in green on the modelled distribution map.

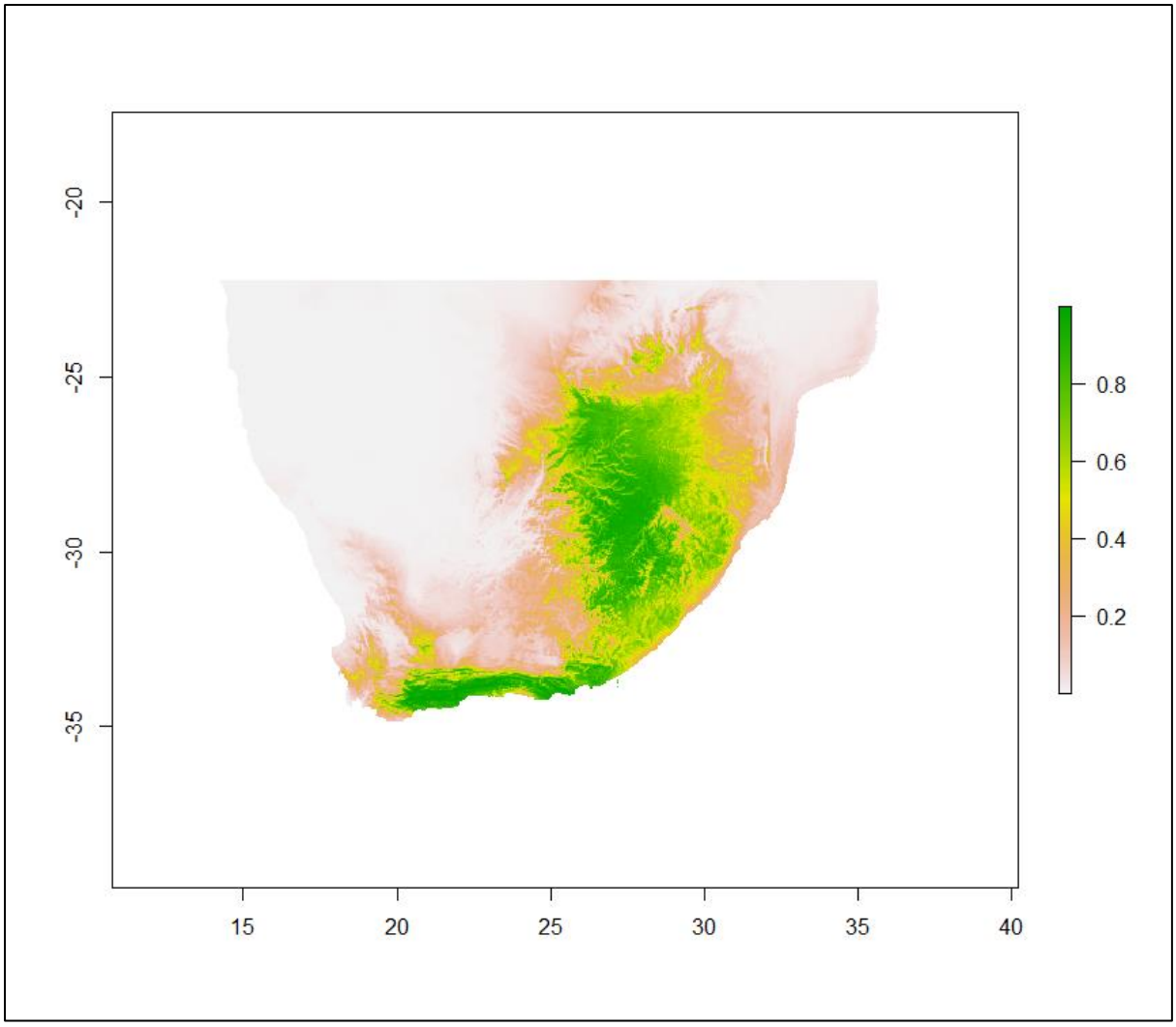

Figure 19: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the early-Holocene. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

This increasing trend of the modelled distribution continues into the mid-Holocene (8 326 – 4 200 years ago). Further expansion was observed in the eastern inland regions of South Africa, specifically along the south-eastern parts of South Africa within the Eastern Cape region (Fig. 20). Despite the possible contraction in the southern Cape region, a relatively continuous distribution was observed along the southern and eastern parts of South Africa during the mid-Holocene.

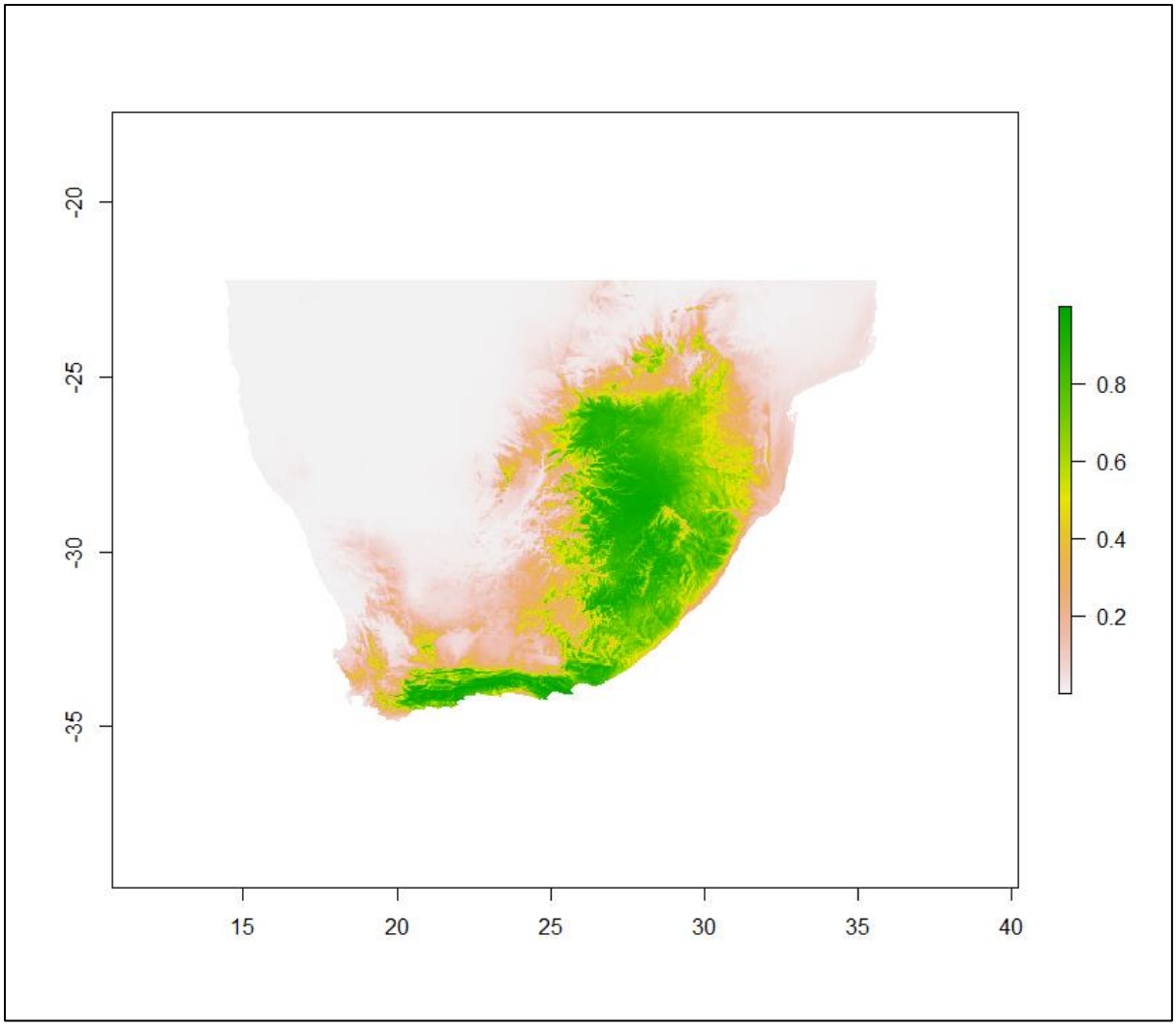

Figure 20: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the mid-Holocene. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

During the late-Holocene (4 200 – 300 years ago) the modelled distribution of *H. haemachatus* shows further contraction within the southern Cape region (Fig. 21). Along with this, climatic conditions around the edges of the predicted inland distribution in the eastern and south-eastern parts resulted in possible expansion of populations in these areas, indicated through the yellow to yellow-green areas shown on the modelled distribution map.

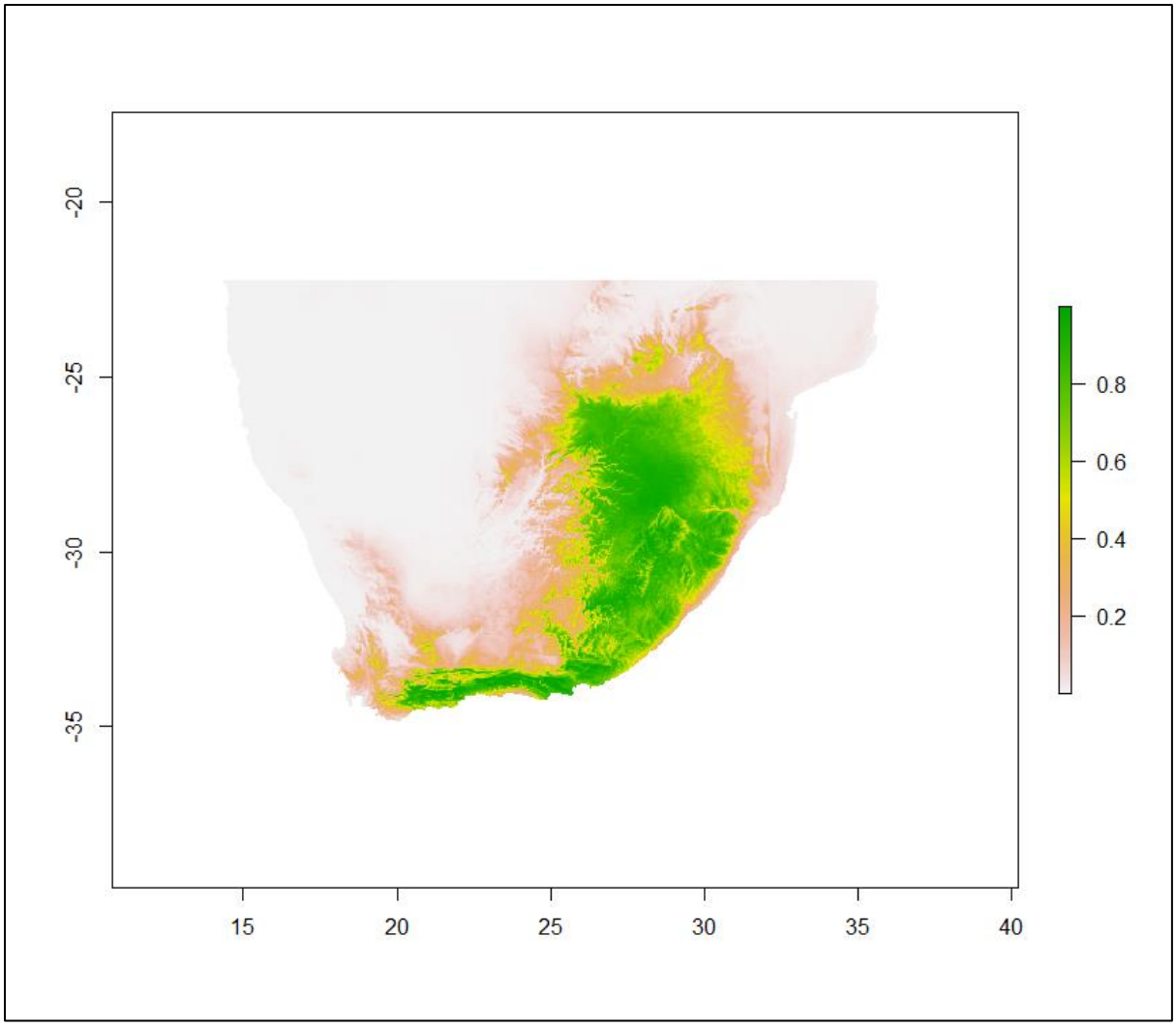

Figure 21: Predicted distribution of *Hemachatus haemachatus* modelled using average bioclimatic variables from the late-Holocene. The scale bar on the right shows the probability of the species occurring in specific areas based on climatic data for those areas.

2.3 Discussion

2.3.1 *Hemachatus haemachatus* population structure and phylogeography

Hemachatus is divided into two well-supported lineages. Lineage 1 includes populations from the southern and eastern Cape coastal and inland regions, as well as eastern inland regions of South Africa. Within Lineage 1, the Cape Lineage is divided into coastal and inland populations. The Eastern and Western Cape coastal populations are most closely related to each other and occur close to one another in terms of distribution as well. The most closely related population to these is a genetically distinct inland population, which occurs within the Eastern Cape region,

around Grahamstown, Hogsback and Maclear. The Northern Cape inland population is a relict population and falls sister to the abovementioned populations. Lineage 1 also contains populations from the eastern inland regions of South Africa, from the Gauteng, Free State, Mpumalanga and KwaZulu-Natal provinces, which are genetically distinct from the Cape populations. Lineage 2 consists of a similar inland lineage, with populations stretching from the eastern regions of South Africa, throughout the Gauteng, North West, Mpumalanga, Free State, and KwaZulu-Natal provinces, and further south into the Eastern Cape Province. However, this lineage shows limited genetic structuring between the populations as most nodes have low support. The inland populations of Lineages 1 and 2 overlap in their distribution. This may explain why, despite the clear genetic split between Lineages 1 and 2, very little variation is observed between the sequences generated and analysed. The patterns observed are most likely due to range expansions and contractions that occurred during glaciation events, although this will be discussed further in section 2.3.2 below.

One of the key elements that is responsible for the development of distinct genetic lineages among populations are historical barriers across landscapes (Krysko et al. 2016). Typically, reptiles have a higher degree of genetic structure over large spatial scales, compared to birds and mammals (Ettling and Parker 2017). The Cape populations occur in Fynbos and Karoo habitats. The Western Cape coastal populations occur in the Cape Floristic Region, consisting of fynbos habitats (Verboom et al. 2015). The Northern Cape inland population from Sutherland occurs in a region characterised by fynbos vegetation (van der Merwe et al. 2008). The Eastern Cape coastal populations occur in areas containing fynbos as well as grassland habitats, which overlap to an extent (Taylor and Morris 1981). The inland populations of the Eastern Cape, KwaZulu-Natal, Mpumalanga, Gauteng, Free State and North West provinces occur in grassland habitats (Mucina and Rutherford 2006). Thus, topographic heterogeneity may have resulted in the Cape populations showing increased differentiation compared to the populations occurring in grassland habitats (Verboom et al. 2015). These factors may explain the current population and phylogeographic structure, however the split between Lineages 1 and 2 was most likely caused by glaciation events in the past, which lead to prevention and subsequent

re-establishment of gene flow between certain populations (Scott 2002). This will be discussed in further detail in section 2.3.2 below.

2.3.2 The effect of past climate on *H. haemachatus* divergence and population demography

Grasslands and fynbos were well established in southern Africa by the Miocene (23.0 to 5.3 mya) (Scott 2002; Neumann 2015). Changes in the African climate and shifts in climate variability during the Pliocene-Pleistocene interval mediated important evolutionary changes in African fauna (DeMenocal 2004). Climatic shifts that occur over large scales alter the ecological composition of landscapes, which present speciation pressures on populations and thereby lead to genetic innovation and selection (DeMenocal 2004). Molecular dating suggests that the radiation of spitting cobras, and *H. haemachatus*, occurred before the end-Miocene/Pliocene grassland transitions at around 15 million years ago, if not earlier (Wüster et al. 2007). Our divergence date estimates fall within the 95% confidence limits of this specified timeframe. I further estimated that *H. haemachatus* differentiated into two separate lineages within the late Pliocene to early Pleistocene interval. In Africa, this stage was characterised by a shift from a warm Pliocene climate to a cooler, more arid climate with alternating glacial cycles (DeMenocal 2004; Barlow et al. 2013). Cycles between peaks of aridity, which coincided with glacial maxima, and interglacial periods of warmer and more humid conditions occurred (Barlow et al. 2013). These events resulted in population isolations, subsequent divergence of lineages, and admixture among diverged lineages in areas of secondary sympatry (Krysko et al. 2016). This pattern is clearly evident in *H. haemachatus*.

Hemachatus haemachatus has a wide thermal tolerance (Alexander et al. 1999) and thus it is unlikely that thermal tolerance limits alone play a major role as a range limiting factor. However, despite its wide thermal tolerance, this species favours low body temperatures and regularly experiences hypothermy (Alexander 1996). It possesses no distinctive physiological traits that would exclude it from surviving in warm environments, however individuals express discomfort at body temperatures of 36 °C or higher, resulting in physiological distress at high temperatures (Alexander 1996). Furthermore, the species does not occur in warmer areas such as savannas or bushveld environments which, along with their preference for fynbos and grassland

environments that generally experience cooler temperatures, lends further support to their preference for cooler climates (Alexander 1996; Branch 1998; Marais 2004; Rutherford et al. 2006). I also found that bioclimatic variables related to temperature made the highest contribution to the distribution model. As *H. haemachatus* is a temperate species adapted for cooler environments (Alexander and Brooks 1999; Alexander 1996), I suspect that warm climatic conditions during the Pliocene resulted in the contraction and isolation of populations, which ultimately lead to genetic differentiation and the formation of two lineages during the Pliocene-Pleistocene interval.

Limited genetic divergence was observed, which suggests that populations were isolated for relatively short periods of time. The presence of identical or very similar haplotypes in localities separated by large geographic distances are indicative of a recent range expansion (Guicking et al. 2009). The mismatch distributions indicate a recent range expansion, based on nuclear gene regions, and that the population was widespread in the past, based on mitochondrial genes (Burbrink et al. 2008; Nuñez et al. 2011). As mitochondrial DNA is a useful marker for examining more recent events, and nuclear DNA is more useful for deep population histories (Hewitt 2004; Hurst and Jiggins 2005; Ursenbacher et al. 2006; Jiang et al. 2007), I can conclude that *H. haemachatus* underwent a range expansion at some point and has been relatively widespread since. The negative Fu's Fs values further support the event of a recent range expansion (Fu 1997; Nuñez et al. 2011).

An increase in the effective population size is illustrated by the EBSP from approximately 100 000 years ago, during the Tarantian stage of the Pleistocene epoch. Several population histories also indicate a downward trajectory in size towards approximately 10 000 years ago. This is coupled to an increase in population density from approximately 100 000 years ago, and a decrease in the population density from around 40 000 years ago towards more recent times. Further population growth was most likely limited due to limiting environmental factors (Edwards and Edwards 2011; Tarsi and Tuff 2012).

The Tarantian stage ranges from 126 000 years to 11 700 years ago (Gibbard and Head 2009) and the distribution models described hereafter all fall within this

period. The modelled distribution, based on bioclimatic variables from the LIG, illustrates three to four potentially isolated populations. When comparing the distribution to the divergence date estimates, I can infer that the Cape population and inland population of Lineage 1 became isolated from one another during the LIG. I can also infer that the KwaZulu-Natal Midlands and Vryheid populations, and the eastern inland populations of Lineage 2 became isolated from one another during this period. The lower 95% confidence limit estimates fall within this specified timeframe. The contraction of populations during this period is most likely due to the contraction of open grassland habitats caused by the expansion of forested habitats during warmer, more humid interglacial periods (Barlow et al. 2013; Lima-Rezende et al. 2019). Open grassland/savannah habitats were found in the western Cape region as well as the eastern regions of South Africa, and *H. haemachatus* populations were restricted to these regions, resulting in genetic differentiation (Barlow et al. 2013; Ecker et al. 2018).

During the LGM, the Cape populations and inland populations of Lineage 1 appear to have expanded and as a result re-established gene flow. These populations appear to have expanded northward, with the northern and inland eastern Cape populations of Lineage 1 potentially became isolated during this period. The formation of a genetically distinct inland population, now represented by populations from Mpumalanga and Gauteng, in lineage 1 also occurred at this time. Populations of lineage 2 underwent further expansion inland, resulting in increased gene flow between these populations. The lower 95% confidence limit estimates fall within this specified timeframe. The expansion of these populations was most likely due to an increase in open grassland/savannah habitats caused by increased aridification leading to decreased forest habitats (Barlow et al. 2013; Lima-Rezende et al. 2019).

A slight contraction was observed in the distribution of *H. haemachatus* populations during the Heinrich Stadial 1, Bølling-Allerød and Younger Dryas Stadial events. The Heinrich Stadial 1 event was characterised as a period with extreme climatic conditions with cold, arid and relatively unstable climate (Camuera et al. 2019). As *H. haemachatus* has a relatively wide thermal tolerance and is well-adapted for cooler climatic conditions (Alexander et al. 1999), it would have been somewhat resilient to such conditions and thus only experienced a slight range contraction. The Bølling-Allerød event was characterised by abrupt warming conditions (Zhan and

Ingersoll 2016), which lead to further range contraction in *H. haemachatus.* The Younger-Dryas Stadial event, on the other hand, was characterised by abrupt cooling (Liu et al. 2012), which again lead to a slight range expansion in *H. haemachatus.* Despite these range contraction and expansion events, *H. haemachatus* populations still appeared to be connected through corridors with suitable climatic conditions.

This range expansion continued throughout the early-, mid- and late-Holocene. This was most likely due to the warm, dry climatic conditions in the early-Holocene, and a shift to cooler, dry conditions towards the late-Holocene which allowed open grassland/savannah habitats to expand (Lee-Thorp et al. 2001; Zhao et al. 2016). It should be noted that the largest expansion occurred during the late-Holocene, again showing the preference that this species has for cooler climatic conditions. The distribution was continuous during the Holocene epoch, further supporting that gene-flow was not restricted between populations.

2.4 Conclusion

The current *H. haemachatus* lineages most likely arose during the Pleistocene epoch as a result of glacial and interglacial cycles leading to cycles of aridification and forestation (Barlow et al. 2013). The interglacial cycles resulted in *H. haemachatus* retreating into open-habitat refugia which were located in the eastern and southern regions of southern Africa (Barlow et al. 2013). A large, stable refugium persisted over a long period of time in the south, which was most likely the refugium populated by Lineage 1 (Barlow et al. 2013; Ecker et al. 2018). Lineage 2 most likely populated the open grassland habitat which was located in the eastern parts of southern Africa (Barlow et al. 2013; Ecker et al. 2018). These lineages appear to have re-established gene flow, as a result of a range expansion during the LGM and the Holocene epoch due to the increasing size of open grasslands/savannah habitats, leading to a continuous distribution (Mokhatla et al. 2015; Bond and Zaloumis 2016).

The past population demography analyses, divergence date estimates and distribution models based on past climatic conditions provide support that the limited genetic variation observed in *H. haemachatus* is most likely due to a recent range expansion (Guicking et al. 2009). The species has remained relatively widespread since its initial range expansion. I expect the distribution of *H. haemachatus* to become

more contracted due to warmer climatic conditions and anthropogenic activities leading to contraction of grassland habitats (Mokhatla et al. 2015; Bond and Zaloumis 2016); this is already evident in certain regions of the Western Cape and Northern Cape regions as only relict populations occur there today (Branch 1998). The distribution model based on data from the Anthropocene period provides further support for this. This species has a clear preference for cooler climatic conditions and are generally associated with open grassland habitats.

Chapter Three

The venom composition of rinkhals, *Hemachatus haemachatus* (Squamata: Elapidae), populations across South Africa

3.1 Introduction

3.1.1 Overview of venomous snakes and snakebite

Snakes represent a highly specialised and diverse group of organisms that are thought to have evolved approximately 100-150 mya (Graham et al. 2008; Caldwell et al. 2015). The presence of snakes has been recorded in all major ecosystems, excluding the polar regions (Graham et al. 2008). Snakes are currently grouped into three superfamilies, namely the Caenophidia, Scolephidia and Henophidia (Keogh 1998; Graham et al. 2008). The Caenophidia contains the families Elapidae, Colubridae and Viperidae (Keogh 1998; Graham et al. 2008). The Scolephidia represent a group of burrowing snakes that still exhibit some ancestral traits such as the presence of pelvic vestiges, although many derived traits, such as skull structure, are present as well (Graham et al. 2008; Lock and Wellehan 2015). The Henophidia contain families of snakes that range from Scolephidia-like species to species that are similar to the Caenophidia (Graham et al. 2008).

All venomous snake species fall within the Caenophidia (Fry et al. 2008; Graham et al. 2008). Venom is a trophic adaptation that plays a vital role in the foraging success of snakes (Gibbs et al. 2011; Margres et al. 2013; Vieira et al. 2013). Snake venom proteins arise through duplications of genes that have physiological functions in other parts fo the body (Fry et al. 2008; Kohlhoff et al. 2012). These genes are selectively expressed in venom glands (Fry et al. 2008; Kohlhoff et al. 2012). The proteins and peptides undergo significant structural and sequence changes, but retain the ancestral protein's molecular scaffold (Kohlhoff et al. 2012). Snakes have evolved the ability to store large amounts of different proteins and peptides in venom glands to immobilise prey, aid in digestion and act as a defence mechanism (Kohlhoff et al. 2012; Margres et al. 2013). Venom constitutes a complex mixture of proteins and peptides that are contained in specialised venom glands located in the upper jaw (Li

et al. 2004; Graham et al. 2008; Gibbs et al. 2011; Margres et al. 2013; Vieira et al. 2013).

The World Health Organisation (WHO) has recently recognised snakebite as a neglected public health issue that causes significant economic and social problems, specifically in developing countries (WHO 2016, 2017). It has been estimated that approximately 5.8 billion people are potentially at risk of coming across a venomous snake (Longbottom et al. 2018; WHO 2019). It is therefore not surprising that around 5.4 million snakebites occur annually, of which approximately 2.7 million result in envenomation, causing 81 000-138 000 deaths and up to 400 000 permanent physical disabilities (WHO 2016, 2019). Although many victims survive, such physical disabilities could potentially have a major economic impact as many victims are relatively young (Gutiérrez et al. 2006; Chippaux 2011).

The most severe cases of snakebite are caused by snakes from the families Elapidae and Viperidae (Kasturiratne et al. 2008). In southern Africa, venomous snakes can broadly be classified as those that have mainly neurotoxic, cytotoxic or haemotoxic venoms (Müller et al. 2012). Most elapids have predominantly neurotoxic venoms, whereas viperids have predominantly cytotoxic venoms (Graham et al. 2008; Müller et al. 2012). However, this classification oversimplifies reality as a single snake's venom can consist of several components, components can vary within a species in relation to location and the age of the individual snake (Williams et al. 1988; Chippaux et al. 1991), and there may be significant overlap of venom types between taxa (Graham et al. 2008; Müller et al. 2012). For example, elapid and viperid venom generally contain components such as proteolytic enzymes, phospholipases, metalloproteinases, serine proteinases, disintegrins, potassium channel-binding neurotoxins, pre- and postsynaptic neurotoxins, cardiotoxins and cytotoxins (Graham et al. 2008; Mackessy and Saviola 2016). These components may have unique activities and can be highly selective (Graham et al. 2008). Toxins found in cytotoxic venom are generally phospholipases, proteolytic enzymes and polypeptides that destroy tissues, skeletal muscle and cell membranes (Gutiérrez et al. 2006; Müller et al. 2012; Mackessy and Saviola 2016). These effects lead to blistering, local swelling and oedema as a result of the toxins increasing the permeability of the vascular endothelium (Gutiérrez et al. 2006; Müller et al. 2012). Irreversible necrosis may also

occur (Müller et al. 2012). Pre- and postsynaptically acting neurotoxins target neuromuscular junctions and cause progressive descending paralysis, which may lead to death (Gutiérrez et al. 2006; Müller et al. 2012). In southern Africa, cytotoxic symptoms are mainly attributed to envenomation by *Bitis arietans*, *Naja mossambica*, *Naja nigricollis*, *Naja nigricincta* and *H. haemachatus* (Müller et al. 2012; WHO 2016). Neurotoxic effects are observed in envenomation caused by non-spitting *Naja*, *Dendroaspis*, *Aspidelaps lubricus*, *Aspidelaps scutatus*, *Elapsoidea* and *Pelamis platurus* (Müller et al. 2012; WHO 2016). However, mild neurotoxic effects have also been reported from *H. hemachatus* envenomation in humans (Müller et al. 2012).

Antivenom is currently the only specific treatment available for the treatment of serious cases of envenomation. However, this treatment remains largely inaccessible to victims from developing countries (Gutiérrez et al. 2006; Brown 2012; WHO 2019). Furthermore, the efficacy of antivenom is restricted to a limited biological spectrum and geographic range as a result of the large immunochemical diversity of snake venoms and may carry significant risk of anaphylaxis (Gutiérrez et al. 2006; Lourenço et al. 2013). Antivenom thus offers only a partial solution to snakebite.

3.1.2 Variation in venom composition

Variation in the venom composition of snakes is a universal phenomenon that occurs at intra- and interspecific levels (Chippaux et al. 1991; Casewell et al. 2014; Amazonas et al. 2019). Venom proteins are under strong selection and may therefore be subjected to rapid rates of evolution leading to variations in venom composition from the family through to the individual level (Gibbs et al. 2011; Casewell et al. 2014; Amazonas et al. 2019). Evidence suggests that this variation can be attributed to some degree of genetic control as a result of the presence or absence of alleles that code for specific venom proteins or due to amino acid substitutions in venom-coding genes (Williams et al. 1988; Gibbs et al. 2011; Casewell et al. 2014). Further evidence suggests that gene regulation can lead to individual venom composition being plastic through time, allowing snakes to adapt to different or changing ecological conditions or in response to changes in diet with age (Gibbs et al. 2011; Amazonas et al. 2019).

Snakes can differ in terms of their venom composition and toxicity over a geographic range (Aguilar et al. 2007; Lourenço et al. 2013). Assessing the

biochemical makeup of snake venom is crucial as treatment is dependent on the composition and activity of the venom (Mendoza et al. 1992; Li et al. 2004; Lourenço et al. 2013). Significant differences have been recorded in the biological and pharmacological activities of venom depending on the origin of the venom (Lourenço et al. 2013; Mackessy and Saviola 2016). Therefore, providing a regional map of the different biological and pharmacological patterns of venom activities could not only play a vital role in characterising the venom of a specific species, but could also aid in selecting region-specific venom for antivenom production (Schenberg 1963; Lourenço et al. 2013).

3.1.3 The study species

Hemachatus haemachatus was selected as the study species as it is morphologically variable and a generalist, occurring in a range of habitats (Branch 1998; Marais 2004; Alexander et al. 2012). It is widespread and relatively common throughout most of its range, apart from in the Northern and Western Cape provinces and the eastern highlands of Zimbabwe where only relict populations occur (Branch 1998). Generalist species that are relatively widespread are considered powerful model systems for investigating and understanding how various ecological factors can drive patterns of species diversification and divergence at a regional scale (Card et al. 2016).

A recent study by Sánchez et al. (2018) investigated the proteomic composition of *H. haemachatus* venom. These authors found that the venom of this species consisted predominantly of three-finger toxins, followed by phospholipase A2, snake venom metalloproteases, cysteine-rich secretory proteins and Kunitz type serine protease inhibitors (Sánchez et al. 2018). Other proteins included cystatin, nerve growth factors and vespryn (Sánchez et al. 2018). There are contradictory views in the literature regarding the venom activities of *H. haemachatus*, with the venom being classified as predominantly neurotoxic, cytotoxic, or both (Strydom and Botes 1971; Joubert 1975a; Condrea et al. 1981; Müller et al. 2012). To add to this confusion, *H. haemachatus* venom appears to have predominantly neurotoxic effects when tested on mice, whereas human envenomation results in predominantly cytotoxic symptoms (Müller et al. 2012; Sánchez et al. 2018). Furthermore, anecdotal evidence, based on dog envenomation, suggests that *H. haemachatus* populations differ in their venom composition with the degree of cytotoxicity and neurotoxicity potentially varying at the

population level. The WHO has listed *H. haemachatus* under Category 2 of medical importance in South Africa, Lesotho, Zimbabwe and Swaziland (WHO 2016). Category 2 refers to snakes that are highly venomous and capable of causing disability, morbidity or death but are not frequently implicated in snakebite cases and/or limited information is available on the exact clinical or epidemiological activity of the venom (WHO 2016). The widespread distribution of *H. haemachatus* as well as disagreements regarding its venom composition makes this species a good candidate for investigating variations in venom composition across a geographical range.

3.1.4 Aims and objectives

To evaluate whether differences in the venom composition of *H. haemachatus* exist across the geographic range and assess whether such differences are genetically linked by:

- analysing the venom composition of *H. haemachatus* specimens collected,
- comparing venom profiles of specimens sampled from different localities and
- linking venom profiles to phylogeographic patterns of this species.

3.2 Methods

3.2.1 Taxon sampling and venom extraction

A total of 16 live specimens were collected from different localities within the Gauteng, Kwa-Zulu Natal, Mpumalanga, Eastern Cape, Western Cape, North West and Free State provinces (Fig. 22). Samples were collected from across the geographic range of this species.

Venom was extracted by inducing each individual snake to bite into a beaker covered by parafilm, which ensured that the venom was injected into the beaker. The venom collected from each individual specimen was lyophilised and stored in a -80 °C freezer.

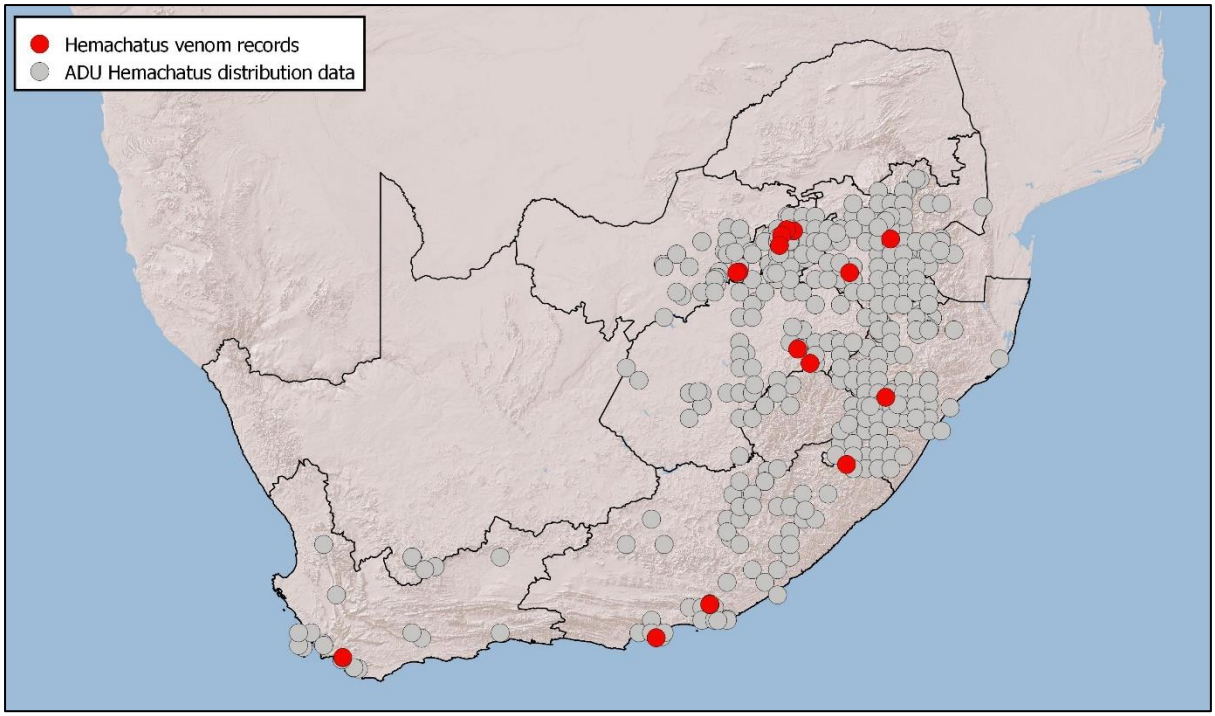

Figure 22: Map of South Africa showing distribution records of the *Hemachatus haemachatus* and localities of venom samples. The red circles represent the 16 venom samples and the grey circles represent *Hemachatus haemachatus* distribution records as supplied by the ADU, Department of Biological Sciences, University of Cape Town.

3.2.2 Proteomic characterisation of the venom peptidome and proteome

A total of 5 mg of each crude venom sample was dissolved in 200 µL of 5% acetonitrile in ultrapure water (MilliQ®; Millipore Co.) containing 0.1% trifluoroacetic acid (TFA). These were centrifuged to remove debris and separated by reverse-phase HPLC (rp-HPLC) using a Teknokroma Europa Protein 300 C18 (4 mm × 250 mm, 5 µm particle size, 300 Å pore size) column and an LC 1100 high pressure gradient system (Agilent Technologies, Santa Clara, CA, USA) which is equipped with a DAD detector and micro-auto-sampler. The flow-rate was set to 1 mL/min. The column was developed with a linear gradient of 0.1% TFA in water (solution A) and acetonitrile (solution B) under the following column elution conditions: isocratically (5% B) for 5 min, followed by 5-25% B for 10 min, 25-45% B for 60 min, and 45-70% for 10 min. Protein detection along the chrmnatogram was carried out at 215 nm with a reference wavelength of 400 nm. The resulting rp-HPLC elution profiles were then compared and samples with similar profiles were pooled.

Fractions with absorbance intensities of 50 mAU or higher were collected manually across the entire elution range, dried in a vacuum centrifuge (Savant™, ThermoFisher Scientific), and redissolved in MilliQ® water. Molecular masses of the purified proteins were estimated by non-reduced and reduced SDS-PAGE (on 15% polyacrylamide gels) or determined by electrospray ionization (ESI) mass spectrometry (MS). For SDS-PAGE analysis sample aliquots were mixed with ¼ volume of 4x sample buffer (0.25M Tris-HCl pH 6.8, 8% SDS, 30% glycerol, 0.02% bromophenol blue, with or without 10% 2-mercaptoethanol) and heated at 85 ºC for 15 min, run under reducing conditions, and the gels were stained with Coomassie Brilliant Blue G-250. For ESI-MS mass profiling, the proteins eluted in the different RP-HPLC fractions were separated by nano-Acquity UltraPerformance LC® (UPLC®) using BEH130 C18 (100µm x 100mm, 1.7 µm particle size) column in-line with a Waters SYNAPT G2 High Definition Mass Spectrometry System. The flow rate was set to 0.6 µL/min and the column was developed with a linear gradient of 0.1% formic acid in water (solution A) and 0.1% formic acid in ACN (solution B), isocratically 1% B for 1 min, followed by 1-12% B for 1min, 12-40% B for 15min, 40-85% B for 2min. Monoisotopic and isotope-averaged molecular masses were calculated by manually deconvolution of the isotope-resolved multiply-charged MS1 mass spectra.

Electrophoretic protein bands were excised from Coomassie Brilliant Blue-stained SDS-PAGE gels and thereafter subjected to in-gel reduction (10 mM dithiothreitol) and alkylation (50 mM iodoacetamide). Thereafter protein bands were digested overnight using sequencing-grade trypsin (66 ng/µL in 25 mM ammonium bicarbonate, 10% acetonitrile; 0.25 µg/sample) in an automated ProGest Protein Digestion Workstation (Genomic Solution Ltd., Cambridgeshire, UK) following the manufacturer's instructions. Tryptic digests were dried in a SpeedVac (Savant™, Thermo Scientific Inc., West Palm Beach, FL, USA), re-dissolved in 15 µL of 0.1% formic acid in water, and then submitted to LC-MS/MS. Tryptic peptides were separated by nano-Acquity UltraPerformance LC® (UPLC®) using BEH130 C18 (100 µm × 100 mm, 1.7 µm particle size) column in-line with a SYNAPT® G2 High Definition Mass Spectrometry System (Waters Corp. Milford Massachusetts, USA). The flow rate was set to 0.6 µL/min. The column was developed with a linear gradient of 0.1% formic acid in water (solution A) and 0.1% formic acid in acetonitrile (solution B) under the

following column elution conditions: isocratically 1% B for 1 min, followed by 1%-12% B for 1 min, 12%-40% B for 15 min, 40%-85% B for 2 min. Ions with a charge of 2^+, 3^+ and 4^+ were selected for collision-induced dissociation (CID) MS/MS. Fragmentation spectra were then interpreted manually (*de novo* sequencing), and by using the on-line form of the MASCOT program at http://www.matrixscience.com against the NCBI non-redundant database. MS/MS mass tolerance was set to ± 0.6 Da. Carbamidomethyl cysteine and oxidation of methionine were selected as fixed and variable modifications, respectively. Amino acid sequence similarity searches were performed against the available databanks using the BLAST program implemented in the WU-BLAST2 search engine (http://www.bork.embl-heidelberg.de).

The relative abundances, i.e. the percentage of the total venom proteins, of the different protein families were calculated as the ratio of the sum of the areas of the reverse-phase chromatographic peaks containing proteins from the same family to the total area of venom protein peaks in the reverse-phase chromatogram (Eichberg et al. 2015; Calvete 2014). If more than one protein band was present in a reverse-phase fraction, their proportions were estimated by densitometry of Coomassie-stained SDS-polyacrylamide gels using ImageJ v1.47 (http://rsbweb.nih.gov/ij). Conversely, the relative abundances of different proteins contained in the same SDS-PAGE band were estimated based on the relative ion intensities of the three more abundant peptide ions associated with each protein by MS/MS analysis. Finally, protein family abundances were estimated as the percentages of the total venom proteome.

To determine similarity between pools, a cluster analysis was performed in R v3.2.3 (R Core Team 2015) using the Aitchison distance measure and the complete linkage clustering method in the Compositions package (Van den Boogaart et al. 2019) to indicate which pools were more similar in terms of composition. The R script can be viewed under supplementary data (Script S3).

3.3 Results

3.3.1 Venom pools

The venom samples were grouped into six pools based on similarities in elution profiles. Each pool consisted of the following samples:

- Pool 1: GP01DJ, GP02DJ and GP01DvE,
- Pool 2: GP02NvdM, MP01CH, MP01FV and FS06JL,
- Pool 3: EC02CK and KZN02LB,
- Pool 4: FS07JL, WC01PB and WC02PB,
- Pool 5: EC08MM and KZN01MP, and
- Pool 6: NW01AH and NW02AH.

The elution profiles of the individual and pooled venom samples can be viewed under supplementary material (Fig. S18A-S18F). SDS-PAGE gel images for each pool can also be viewed under supplementary material (Fig. S19A-S19F).

3.3.2 Proteomic composition

The proteomic composition of each venom pool is shown in Table 9 and Fig. 23. Only Pool 1 and Pool 4 were sampled and analysed extensively. Protein families present in the remaining pools were inferred from these two pools based on similarities in the chromatograms and band patterns in the SDS-PAGE gels. Selected bands were sampled and analysed from these pools to confirm which protein families were present in bands that appeared to differ from those found in Pools 1 and 4.

A total of 20 protein families were identified, of which 17 were identified in Pool 1, 16 in Pool 5, 15 in Pools 3 and 4, 13 in Pool 6 and 11 in Pool 2 (Table 9). However, only eight were present in all six pools (Table 9).

Table 9: The proteomic composition of the six venom pools indicated as percentages. Zero-values are indicated with a dash (-).

Protein Family	Pool 1	Pool 2	Pool 3	Pool 4	Pool 5	Pool 6
5'-Nucleotidase	0.231	0.066	0.213	0.022	0.182	0.178
Acetylcholinesterase	-	-	-	0.011	-	-
Chitotriosidase	0.003	0.024	0.007	-	0.010	0.009
Cobra Venom Factor	-	-	-	0.047	-	-
Cysteine-Rich Secretory Protein	7.093	4.226	3.057	6.370	5.080	3.503
Disintegrin-Like/Cysteine-Rich Domain	0.037	-	-	0.191	0.064	-
Glutathione Peroxidase	0.023	-	0.144	0.115	0.036	-
Insulin-Like Growth Factor	0.061	-	0.078	-	0.026	-
Kunitz-Type Serine Protease Inhibitor	2.032	0.912	5.899	1.643	4.058	1.761
L-Amino Acid Oxidase	1.426	1.281	1.736	1.057	1.236	0.945
Phosphodiesterase	0.430	-	0.186	0.433	0.326	0.394
Phospholipase A2	14.027	19.256	13.453	12.625	9.526	14.052
Phospholipase B	0.002	-	0.009	0.004	-	-
Snake Venom Metalloprotease	17.011	11.111	16.048	17.961	15.989	7.861
Snake Venom Serine Protease	0.016	-	-	-	0.054	0.009
Three Finger Toxin	53.567	61.647	53.990	59.264	61.234	70.192
Vascular Endothelial Growth Factor	0.359	0.575	0.188	-	0.182	0.480
Snake venom nerve growth factor	1.848	0.147	3.511	0.255	1.797	0.061
Vespryn	0.650	0.754	1.481	-	0.201	0.551
Serum Albumin	-	-	-	0.002	-	-
Unknown	1.184	-	-	-	-	0.004

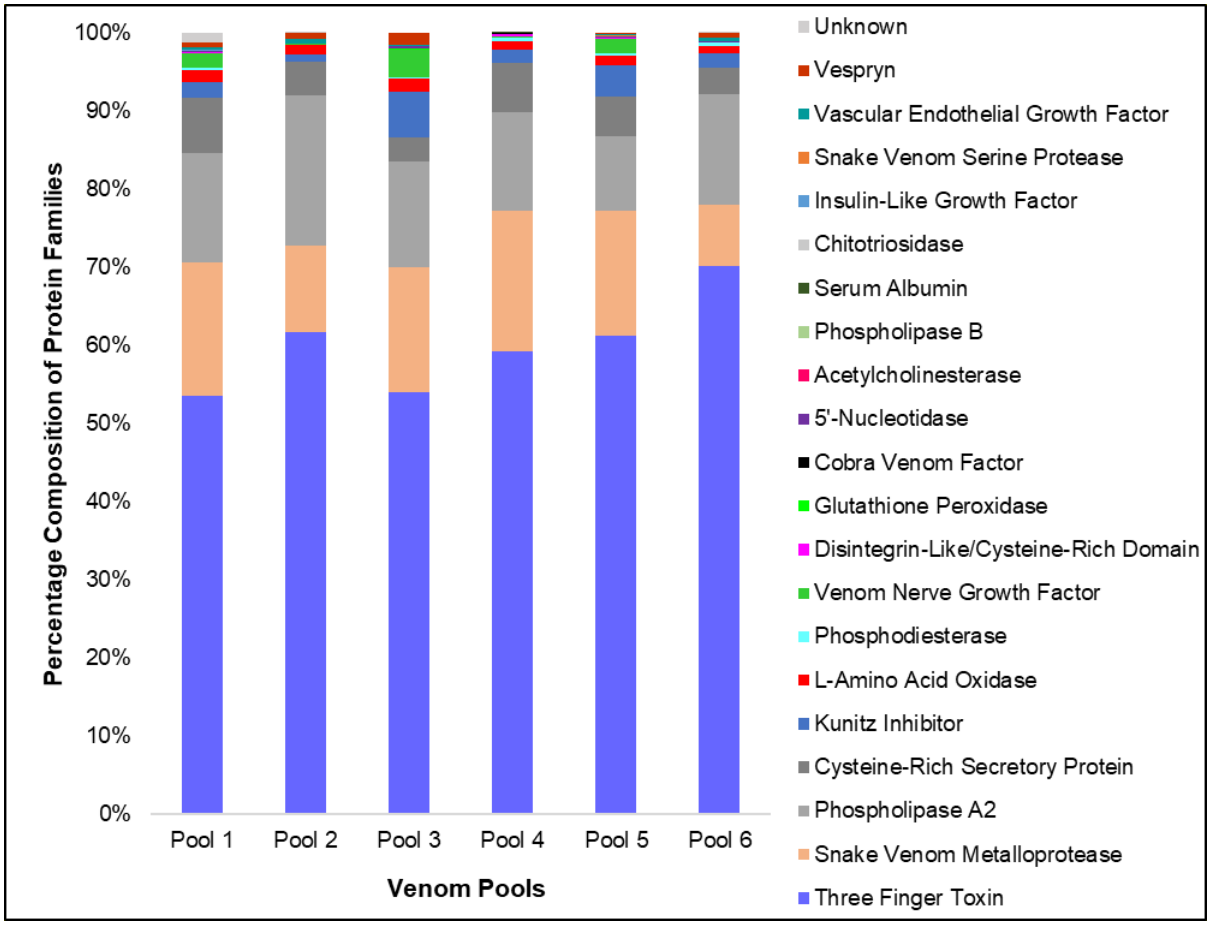

Figure 23: Bar chart indicating percentage composition of protein families present in venom Pools 1 to 6.

The following sections (3.3.2.1-3.3.2.2) provide an overview of the protein families and proteins identified in this study. GenBank accession numbers for the proteins listed are shown in brackets.

3.3.2.1 Protein families present in all pools

Protein families present in all six pools included 5'-nucleotidase (5'-NT), cysteine-rich secretory protein (CRISP), Kunitz-type serine protease inhibitor (KI), L-amino acid oxidase (LAO), phospholipase A2 (PLA2), snake venom metalloprotease (SVMP), three-finger toxin (3FTx), and snake venom nerve growth factor (vNGF) families. These protein families made up between 97% (Pool 1) to 99% (Pool 4) of the overall venom composition in each pool and thus contributed the largest portion to the overall venom composition.

Three-finger toxins made up the largest component of the overall venom composition in all six pools (Table 9 and Fig. 23). These toxins were found between

peaks 1 and 43 in Pool 1, peaks 1 and 35 in Pool 2, peaks 1 and 38 in Pool 3, peaks 1 and 41 in Pools 4 and 5, and peaks 1 and 36 in Pool 6 (supplementary material Tables S2A-S2F, fig. S18A-S18F, fig. S19A-S19F). The largest variety of protein fragments were identified in this protein family. The venom consisted of cytotoxins, cardiotoxins and neurotoxins (supplementary material Tables S2A-S2F). A number of proteins were identified, which included cytotoxin homolog toxin 9B/9BB (P24778), cytotoxin homolog 3 (P01473), cytotoxin 1 (P01471), cytotoxin 2 toxin 12A (P24776), cytotoxin V(II)2 (P01469), cytotoxin V(II)3 (P01470), cytotoxin 3 toxin 11/11A (P24777), cytotoxin 13 (BAU24670), short neurotoxin 1 (P01431, P01425 and P68417), short neurotoxin 1 toxin II (P01425), short neurotoxin 2 toxin IV (P01433), short neurotoxin 3 (P01420), short neurotoxin A (P25495), short neurotoxin SNTX-1 (A6MFK6), long neurotoxin (AHZ08824), long neurotoxin 1 (P01383), long neurotoxin homolog (O93422), long-chain neurotoxin isoform 2 (ACY68696), long neurotoxin OH-57 (Q53B56), neurotoxin homolog NL1 (Q9DEQ3), probable weak neurotoxin NNAM3 (Q9YGI1), weak neurotoxin 9 (Q9W7I3), weak toxin CM-11 (P01401), weak toxin CM-1c (P25676), three-finger hemachatoxin (B3EWH9), putative three-finger toxin precursor (AED89572 and AED89561), exactin (P0DQH2), ringhalexin (C0HJT5), toxin Lc b (P0C8R8), actiflagelin (C0HKZ8), alpha-elapitoxin-Lh2a (Q8UW28), toxin S6C6 (P25682), cobrotoxin homolog (Q9PTT0), cobrotoxin-b (P80958), beta-cardiotoxin CTX23 (Q2VBN4), muscarinic toxin 7 (Q8QGR0), and a muscarinin toxin-like protein (Q9W727).

The SVMP protein family was the second most abundant proteins present in Pools 1, 3, 4 and 5, and the third most abundant in Pools 2 and 6 (Table 9 and Fig. 23). These proteins were found between peaks 43 and 52 in Pool 1, peaks 37 and 46 in Pool 2, peaks 38 and 46 in Pool 3, peaks 40 and 53 in Pool 4, peaks 44 and 57 in Pool 5, and peaks 36 and 46 in Pool 6 (supplementary material Tables S2A-S2F, fig. S18A-S18F, fig. S19A-S19F). The venom was very rich in zinc-metalloproteinase-disintegrin like proteins, as well as metalloproteinase precursors (supplementary material Tables S2A-S2F). Majority of the elements belonged to P-III SVMP's. Fragments identified were similar to P-III snake venom metalloprotease (AHZ08819), snake venom metalloproteinase HT-2 (P20897), snake venom metalloproteinase-disintegrin-like mocarhagin (Q10749), nigresceaese-1 (ABQ01139), carinatease-1

(ABQ01132), australease-1 (ABQ01134), predicted zinc metalloproteinase-disintegrin-like VLAIP-B (XP_013929562), zinc metalloproteinase-disintegrin-like BfMP (A8QL48), zinc metalloproteinase-disintegrin-like atragin (D3TTC2), zinc metalloproteinase-disintegrin-like kaouthiagin-like (D3TTC1), zinc metalloproteinase-disintegrin-like cobrin (Q9PVK7), zinc metalloproteinase-disintegrin-like MTP4 isoform X9 (XP_026570659), zinc metalloproteinase-disintegrin-like atrase-A (D5LMJ3), metalloproteinase precursor (ABK63559 and AAM27042), predicted hemorrhagic metalloproteinase-disintegrin-like kaouthiagin (XP_013922279), and asrin (ABH10621).

The PLA2 protein family was the second most abundant protein family in Pools 2 and 6, and the third most abundant in Pools 1, 3, 4 and 5 (Table 9 and Fig. 23). This protein family was found between peaks 11 and 33 in Pool 1, peaks 10 and 28 in Pool 2, peaks 18 and 34 in Pool 3, peaks 2 and 36 in Pool 4, peaks 13 and 27 in Pool 5, and peaks 12 and 27 in Pool 6 (supplementary material Tables S2A-S2F, fig. S18A-S18F, fig. S19A-S19F). Within this protein family, basic phospholipase A2 DE-1 (P00595) was the most common protein fragment identified (supplementary material Tables S2A-S2F). Other fragments identified were similar to phospholipase A2 (AHZ08815), basic phospholipase A2 3 (Q6SLM0), acidic phospholipase A2 1 (Q6SLM2, P00596 and P00598), acidic phospholipase A2 2 (P15445), neutral phospholipase A2 muscarinic inhibitor (Q92084), and neutral phospholipase A2 homolog taipoxin beta chain 1 (P00615).

The CRISP protein family formed part of the fourth most abundant family in Pools 1, 2, 4, 5 and 6 (Table 9 and Fig. 23). This protein family was the sixth most abundant in Pool 3 (Table 9 and Fig. 23). It was found between peaks 37 and 48 in Pool 1, peaks 31 and 42 in Pool 2, peaks 34 and 42 in Pool 3, peaks 32 and 44 in Pool 4, peaks 40 and 46 in Pool 5, and peaks 33 and 38 in Pool 6 (supplementary material Tables S2A-S2F, fig. S18A-S18F, fig. S19A-S19F). A number of cysteine-rich venom and secretory proteins were identified, along with hypothetical proteins and a CRISP isoform (supplementary material Tables S2A-S2F). The fragments identified were similar to CRISP isoform 2 (ACY68723), najanajin (P86543), cysteine-rich secretory protein Bc-CRPb (ACE73577), cysteine-rich venom protein annuliferin-a (P0DL14), cysteine-rich venom protein annuliferin-b (P0DL15), cysteine-rich venom protein hematin (P0DL17),

cysteine-rich venom protein kaouthin-2 (P84808), cysteine-rich venom protein latisemin (Q8JI38), cysteine-rich venom protein LEI1 (Q2XXQ1), cysteine-rich venom protein natrin-2 (Q7ZZN8), cysteine-rich venom protein ophanin (Q7ZT98), cysteine-rich venom protein tigrin (Q8JGT9), and hypothetical protein L345_12110 (ETE62137).

The KI family formed part of the fourth most abundant protein family in Pool 3, the fifth most abundant protein family in Pools 1, 4, 5 and 6, and the sixth most abundant protein family in Pool 2 (Table 9 and Fig. 23). This protein family was found between peaks 14 and 20 in Pool 1, peaks 3 and 18 in Pool 2, peaks 6 and 20 in Pool 3, peaks 3 and 19 in Pool 4, peaks 4 and 25 in Pool 5, and peaks 3 and 21 in Pool 6 (supplementary material Tables S2A-S2F, fig. S18A-S18F, fig. S19A-S19F). Serine protease inhibitors, venom basic protease inhibitors and protease inhibitors were identified (supplementary material Tables S2A-S2F). The most prominent fragment was similar to the protein identified as Kunitz-type serine protease inhibitor 2 (P00985, P00986 and P20229). Fragments similar to Kunitz-type serine protease inhibitor bungaruskunin (B2KTG1), Kunitz-type serine protease inhibitor carinatin-1 (Q6ITB0), Kunitz-type serine protease inhibitor IX (P25660), Kunitz-type serine protease inhibitor NACI (Q5ZPJ7), Kunitz-type serine protease inhibitor vestiginin-4 (A6MFL4), protease inhibitor 4 (C1IC53), and venom basic protease inhibitor II (P00986) were also identified.

The LAO protein family was the fifth most abundant protein family in Pool 2, the sixth most abundant in Pools 4 and 6, and the seventh most abundant in Pools 1, 3 and 5 (Table 9 and Fig. 23). This protein family was found between peaks 51 and 52 in Pool 1, peaks 45 and 46 in Pools 2, 3 and 6, peaks 50 and 53 in Pool 4, peaks 56 and 57 in Pool 5 (supplementary material Tables S2A-S2F, Fig. S18A-S18F, Fig. S19A-S19F). Four fragments of LAO were identified (A8QL51, A8QL58, AVX27607 and P81382), of which the fragment most similar to AVX27607 was the most common (supplementary material Tables S2A-S2F)

The vNGF protein family was the fifth most abundant in Pool 3, the sixth most abundant family in Pools 1 and 5, the eighth most abundant in Pool 4, the ninth most abundant in Pool 2 and the 11th most abundant in Pool 6 (Table 9 and Fig. 23). This

protein family was found between peaks 22 and 41 in Pool 1, peaks 19 and 31 in Pool 2, peaks 20 and 35 in Pool 3, peak 19 in Pool 4, peaks 23 and 27 in Pool 5, and peak 33 in Pool 6 (supplementary material Tables S2A-S2F, Fig. S18A-S18F, Fig. S19A-S19F). Three fragments of vNGF were identified (P61898, BAN82142 and P01140), of which the P61898 fragment was the most common (supplementary material Tables S2A-S2F).

The 5'-NT protein family was the ninth most abundant in Pool 3, the 10th most abundant in Pools 2, 5 and 6, the 11th most abundant in Pool 1 and the 12th most abundant in Pool 4 (Table 9 and Fig. 23). This protein family was found in peak 51 in Pool 1, peak 45 in Pools 2, 3 and 6, peak 53 in Pool 4, and peak 56 in Pool 5 (supplementary material Tables S2A-S2F, Fig. S18A-S18F, Fig. S19A-S19F). Three 5'-NT fragments were identified (B6EWW8, BAN82018 and AHJ80886) (supplementary material Tables S2A-S2F).

3.3.2.2 Other protein families

The remaining 12 protein families identified were chitotriosidae (CHIT1), disintegrin-like/cysteine-rich domain (DC domain), glutathione peroxidae (GPx), insulin-like growth factor (IGF), phosphodiesterase (PDE), phospholipase B (PLB), snake venom serine protease (SVSP), vascular endothelial growth factor (VEGF), vespryn, acetylcholinesterase (AChE), cobra venom factor (CVF), and serum albumin (SA). Three of these protein families, namely AChE, CVF and SA, were unique to Pool 4. Majority of these protein families contributed less than 1% to the overall venom composition in all six pools (Table 9).

The PDE protein family was present in all pools, except Pool 2 (Table 9). It was the seventh most abundant protein family in Pool 4, the eighth most abundant in Pool 5, the ninth most abundant in Pools 1 and 6, and the eleventh most abundant in Pool 3 (Table 9 and Fig. 23). This protein family was found between peaks 48 and 49 in Pool 1, in peak 42 in Pools 3 and 6, peak 48 in Pool 4, and in peak 52 in Pool 5 (supplementary material Tables S2A, S2C-S2F, Fig. S18A-S18F, Fig. S19A-S19F). Phosphodiesterase family member 3 (AHJ80885, ALA20853, XP_026561286, XP_013908397 and XP_007430581) was the main PDE protein present (supplementary material Tables S2A, S2C-S2F).

Some of the SVMP proteins listed in section 3.3.2.1 were situated in a disintegrin or cysteine rich region, and thus form part of the DC domain protein family. These included fragments identified as zinc metalloproteinase-disintegrin-like cobrin, zinc metalloproteinase-disintegrin-like VLAIP-B and australease-1 (supplementary material Tables S2A, S2D, S2E). The DC domain protein family was absent from pools 2, 3 and 6. It was the ninth most abundant protein family in Pool 4, the 12th most abundant in Pool 5 and the 13th most abundant in Pool 1 (Table 9 and Fig. 23). This protein family was found in peak 9 in Pool 1, between peaks 12 and 14 in Pool 4, and between peaks 14 and 19 in Pool 5 (supplementary material Tables S2A, S2D, S2E, Fig. S18A, S18D, S18E, Fig, S19A, S19D, S19E).

The GPx protein family was absent from Pools 2 and 6. It was the 10th most abundant protein family in Pool 4, the 12th most abundant in Pool 3 and the 14th most abundant in Pools 1 and 5 (Table 9 and Fig. 23). This protein family was found in peak 49 in Pool 1, peak 43 in Pool 3, between peaks 48 and 50 in Pool 4, and in peak 52 in Pool 5 (supplementary material Tables S2A, S2C-S2E, Fig. S18A, S18C-S18E, Fig. S19A, S19C-S19E). The only GPx protein identified was glutathione peroxidase 3, of which two fragments were identified (ETE68810 and AFS63888) (supplementary material Tables S2A, S2C-S2E).

Phospholipase B was absent from Pools 2, 5 and 6. It is the 14th most abundant protein family in Pools 3 and 4, and the least abundant protein family in Pool 1 (Table 9 and Fig. 23). This protein family was found in peak 51 in Pool 1, peak 45 in Pool 3, and peak 50 in Pool 4 (supplementary material Tables S2A, S2C, S2D, Fig. S18A, S18C, S18D, Fig. S19A, S19C, S19D). Phospholipase B (F8S101), phospholipase-B 81 (F8J2D3) and phospholipase B-like 1 (ETE59578) proteins were identified in these pools (supplementary material Tables S2A, S2C, S2D).

Chitotriosidase was present in all pools except Pool 4. It was the least abundant protein family in Pools 2, 3, 5 and 6 and the 16th most abundant protein family in Pool 1 (Table 9 and Fig. 23). This protein family was found in peak 50 in Pool 1, peak 44 in Pools 2 and 3, peak 54 in Pool 5, and peak 43 in Pool 6 (supplementary material Tables S2A-S2C, S2E-S2F, Fig. S18A-S18C, S18E-S18F, Fig. S19A-S19C, S19E-

S19F). A single protein, chitotriosidase-1 (XP_015678645), was identified in this family (supplementary material Tables S2A-S2C, S2E-S2F).

The IGF protein family was absent from Pools 2, 4 and 6. It was the 12th most abundant protein family in Pool 1, the 13th most abundant in Pool 3 and the 15th most abundant in Pool 5 (Table 9 and Fig. 23). This protein family was found between peaks 24 and 28 in Pool 1, peaks 22 and 24 in Pool 3, and in peak 26 in Pool 5 (supplementary material Tables S2A, S2C, S2E, Fig. S18A, S18C, S18E, Fig. S19A, S19C, S19E). Three proteins were identified in this protein family. These included insulin-like growth factor I (ETE61012), insulin-like growth factor-binding protein 3 (ETE63541) and insulin-like growth factor-binding protein 3 isoform X1 (XP_026558105) (supplementary material Tables S2A, S2C, S2E).

The SVSP protein family was only present in Pools 1, 5 and 6. It was the 12th most abundant protein family in Pool 6, the 13th most abundant in Pool 5 and the 15th most abundant in Pool 1 (Table 9 and Fig. 23). This protein family was found in peak 50 in Pool 1, between peaks 20 and 57 in Pool 5, and between peaks 36 and 46 in Pool 6 (supplementary material Tables S2A, S2E, S2F, Fig. S18A, S18E, S18F, Fig. S19A, S19E, S19F). The only protein identified in this family was beta-fibrinogenase brevinase (Q9PT51) (supplementary material Tables S2A, S2E, S2F).

The VEGF protein family was present in all pools, excluding Pool 4. It was the eighth most abundant protein family in Pools 2 and 6, the 10th most abundant in Pools 1 and 3, and the 11th most abundant in Pool 5 (Table 9 and Fig. 23). This protein family was found in peak 35 in Pools 1 and 5, peak 29 in Pool 2, and peak 30 in Pools 3 and 6 (supplementary material Tables S2A-S2C, S2E-S2F, Fig. S18A-S18C, S18E-S18F, Fig. S19A-S19C, S19E-S19F). Snake venom vascular endothelial growth factor toxin barietin (C0K3N1) was the only protein identified in this protein family (supplementary material Tables S2A-S2C, S2E-S2F).

Vespryn was only absent from Pool 4. It was the seventh most abundant protein family in Pools 2 and 6, the eighth most abundant in Pools 1 and 3, and the ninth most abundant in Pool 5 (Table 9 and Fig. 23). This protein family was found between peaks 32 and 37 in Pool 1, peaks 28 and 29 in Pool 2, peaks 27 and 34 in Pool 3, peaks 35 and 37 in Pool 5, and peaks 30 and 33 in Pool 6 (supplementary material Tables S2A,

S2B, S2C, S2E, S2F, Fig. S18A, S18B, S18C, S18E, S18F, Fig. S19A, S19B, S19C, S19E, S19F). This family was represented by a protein identified as thaicobrin (P82885).

In Pool 4, the CVF protein family was the 11th most abundant protein family, the AChE protein family was the 13th most abundant, and the SA protein family was the least abundant (Table 9 and Fig. 23). The CVF protein family was found between peaks 52 and 53, while the AChE and SA protein families were found in peak 50 (supplementary material Table S2D, Fig. S18D, Fig. S19D). Only a single CVF protein (Q91132), a serum albumin precursor (S59517) and two AChE proteins (AAC59905 and XP_015668657) were identified (supplementary material Table S2D).

Approximately 1.184% and 0.004% of the proteins in Pools 1 and 6, respectively, could not be identified and thus could not be assigned to a specific protein family (Table 9 and Fig. 23). This could be due to incomplete collection of fractions during rp-HPLC, incomplete excision of bands from SDS-PAGE gels, or an inability to interpret fragmentation spectra using de novo sequencing.

Although I mentioned specific proteins that were identified in this study, the focus remained at the protein family level when comparing the venom composition between pools in section 3.3.3 below. Inferences made were based on patterns identified at the protein family level. This is because the methods used in this study are not appropriate for fine-scale analyses at the individual protein level. This is discussed in further detail in section 3.4.1.

3.3.3 Comparison of venom profiles

When comparing only the shared protein families, similar profiles were observed in Pools 1, 3 and 5, and Pools 2 and 6, respectively (Fig. 24). The order of abundance of these protein families in Pools 1 and 5 were 3FTx > SVMP > PLA2 > CRISP > KI > vNGF > LAO (Fig. 24). Pool 3 only differed in terms of the percentage composition of CRISP being smaller than the percentage composition of the KI and vNGF protein families (Fig. 24). However, when including all the protein families, the cluster analysis indicated that Pools 1, 3 and 5 cluster together, with Pools 1 and 3 being more similar than Pools 1 and 5, despite the profiles of the major components being similar (Fig. 25). Pools 1 and 3 merged at a height of approximately 5.1, and these merged with

Pool 5 at a height of approximately 6.0 (Fig. 25). Protein families present in all three these pools were 5'-NT, vNGF, CHIT1, CRISP, GPx, IGF, KI, LAO, PDE, PLA2, SVMP, 3FTx, VEGF and vespryn. The DC domain and SVMP protein families were found in Pools 1 and 5, but not in Pool 3, and PLB was found in Pools 1 and 3, but not in Pool 5 (Table 9). No protein families were unique to Pools 3 and 5 (Table 9). Pools 1, 3 and 5 were the only pools in which the percentage composition of vNGF was higher than LAO (Fig. 24).

Pools 2 and 6 showed similar profiles, however Pool 2 had a higher percentage composition for LAO compared to KI, and Pool 6 had a higher percentage composition of KI compared to LAO (Fig. 24). These were the only pools in which the percentage composition of PLA2 was higher than SVMP (Fig. 24). These two pools also clustered together in the cluster analysis and merged at a height of approximately 4.9. This cluster then merged with Pools 1, 3 and 5 at a height of approximately 9.3 (Fig. 25).

Pool 4 contained a total of six unique features. This was the only pool that did not contain the CHIT1, VEGF and vespryn protein families, and the only pool that contained the AChE, CVF and SA protein families (Table 9). This pool did not cluster with any of the other pools and merged at a height of approximately 10.6 (Fig. 25).

Figure 26 shows the localities for the samples from each pool. Pool 1 contained samples from the West Rand of Johannesburg in the Gauteng province of South Africa. Pool 2 contained samples from Centurion in Gauteng, Carolina and Standerton in Mpumalanga and Clarens in the Free State. Pool 3 contained samples from Grahamstown in the Eastern Cape and Kokstad from KwaZulu-Natal. Pool 4 contained samples from Hermanus in the Western Cape and Bloemfontein in the Free State; however, I did not consider the Free State sample as it did not reflect the unique features of the venom composition of the Western Cape population. Pool 5 contained samples from Port Elizabeth in the Eastern Cape and Mooi River in KwaZulu Natal, and Pool 6 contained samples from Potchefstroom in the North West province of South Africa. Pools 1, 3, and 5 and Pool 2 and 6 were most similar to each other in terms of venom composition (Fig. 24, 25, 26). Pool 4 was the most unique (Fig. 25, 26). I observed an overlap in the distribution of all pools, except Pool 4 (Fig. 26).

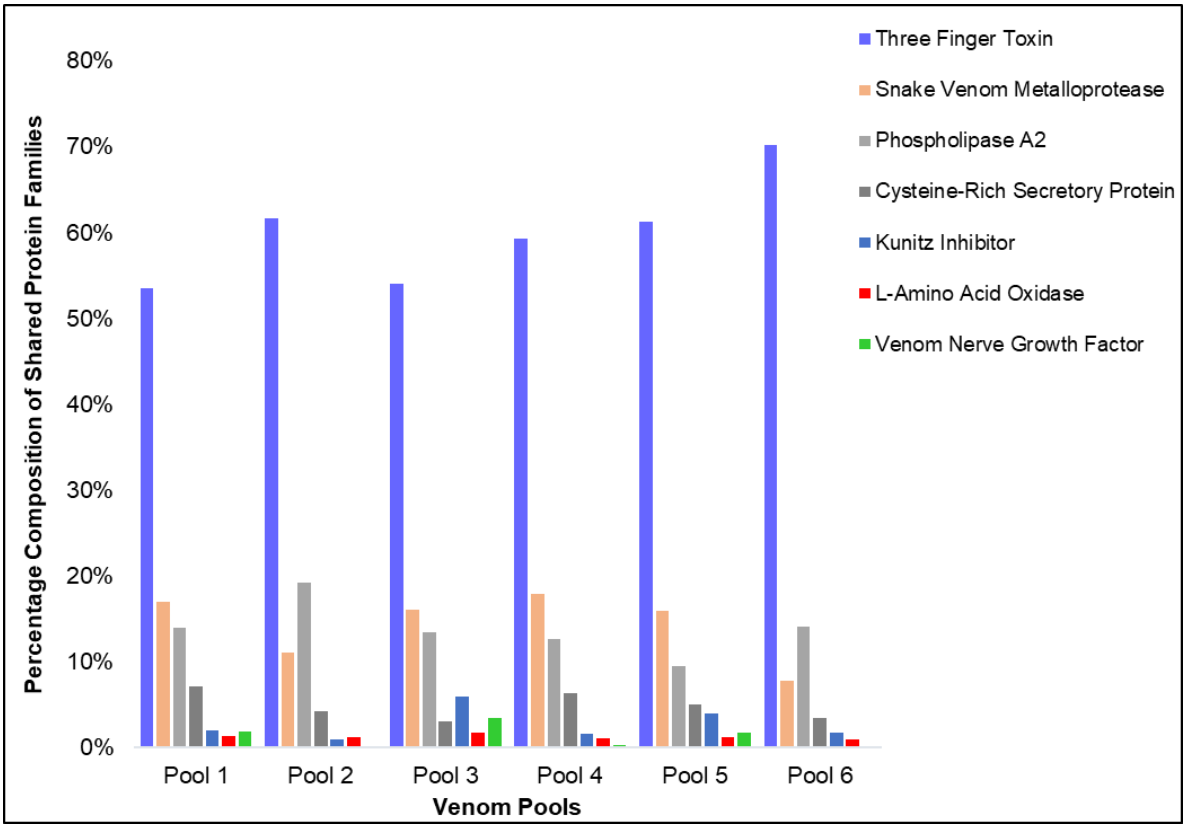

Figure 24: Bar chart indicating percentage composition of protein families present in all venom pools. The 5'-nucleotidase protein family is not shown in this graph as percentages are smaller than 0.01%.

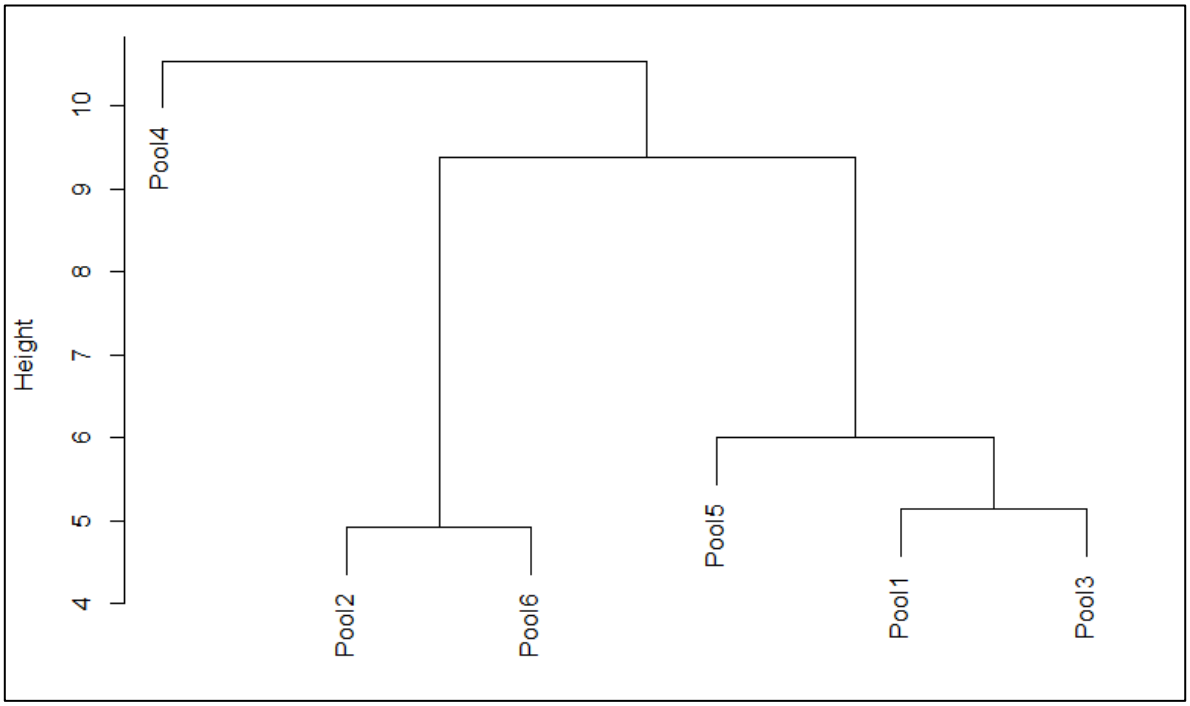

Figure 25: Cluster dendogram indicating which venom pools were more similar in terms of composition. The y-axis displays the height, which represents the distance between clusters. The point at which the clusters are merged is indicated by the horizontal bars.

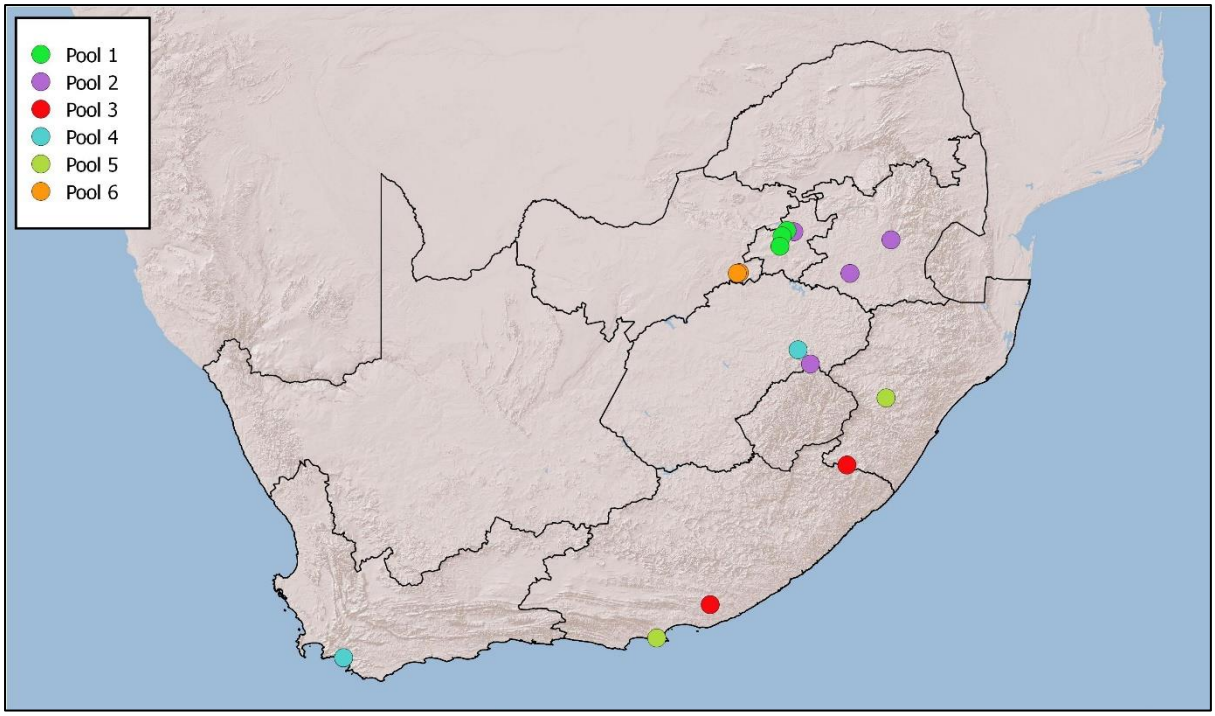

Figure 26: Localities of samples included in the six pools. The samples are colour coded according to pools.

3.4 Discussion

3.4.1 Venom proteomics of *Hemachatus haemachatus*

The venom composition of *H. haemachatus* has been investigated by a number of researchers (Joubert 1975a; Hokama et al. 1976; Strydom 1976; Joubert and Taljaard 1980; Girish et al. 2012; Barnwal et al. 2016; Girish and Kini 2016; Sánchez et al. 2018). Most of the research was conducted during the late 1900's and focused on isolating specific components of the venom. More recently, Sánchez et al. (2018) provided a more complete overview of the proteomic composition of *H. haemachatus* venom. My results are similar to those described by Sánchez et al. (2018). The protein families described were divided into three groups, based on the following criteria:

- Major protein families made up 97%-99% of the overall venom composition and were present in all six pools.
- Minor protein families made up 1%-2% of the overall venom composition and were present in two to five pools.
- Unique protein families were present in only a single pool.

3.4.1.1 Major protein families

Eight protein families were identified in this group. These included 3FTx, PLA2, SVMP, CRISP, KI, vNGF, LAO and 5'-NT. These protein families made up majority of the venom composition in all six pools. The most prominent protein family was 3FTx, followed by PLA2 and SVMP, CRISP, KI, vNGF and LAO. The 5'-NT protein family was the least abundant protein family in this group.

In all pools, more than 50% of the overall venom composition consisted of 3FTx. These toxins are non-enxymatic proteins that are commonly found in elapid venoms (Girish et al. 2012). Members of this toxin family have a similar structure, containing four conserved disulfide bonds located in the core that are responsible for maintaining the characteristic three-finger fold structure, but differ in their biological activities (Girish et al. 2012; Sánchez et al. 2018). The biological activities of the cytotoxins identified in the 3FTx family have been previously investigated by a number of researchers (Fryklund and Eaker 1973; Louw 1974; Shipolini et al. 1975; Joubert 1977; Otting et al. 1987; Chien et al. 1994; Girish et al. 2012; Girish and Kini 2016). The biological activities of these toxins include cytotoxic (Shipolini et al. 1975), cytolytic (Louw 1974; Otting et al. 1987; Chien et al. 1994), cardiotoxic (Fryklund and Eaker 1973; Joubert 1977; Girish et al. 2012; Vonk et al. 2013), hemolytic (Fryklund and Eaker 1973; Joubert 1977; Girish et al. 2012), hypotensive (Joubert 1977; Fryklund and Eaker 1973; Girish et al. 2012) and hemostasis impairing (Girish and Kini 2016) effects. Previous research indicates that the neurotoxins identified in this study target the muscular or nicotinic acetylcholine receptors leading to inhibited binding of acetylcholine to the receptors, and thereby impair neuronal and/or neuromuscular transmission (Botes and Strydom 1969; Strydom and Botes 1971; Shipolini et al. 1974; Joubert 1975b; Kim and Tamiya 1982; Tamiya et al. 1983; Lin et al. 1998; Qian et al. 1998; Poh et al. 2002; He et al. 2004; St Pierre et al. 2007).

The PLA2 proteins found in snake venom are acidic or basic enzymes with a stable structure consisting of seven disulfide bonds (Rafhaella et al. 2018). Neurotoxic, cytotoxic, myotoxic, hemolytic, proinflammatory and edematogenic effects have been reported (Rafhaella et al. 2018). However, only anticoagulant and neurotoxic effects have been reported for the proteins identified in this study (Joubert 1975a; Fohlman et al. 1976; Pan et al. 1994; Miyoshi and Tu 1996; Singh et al. 2005a,b).

The SVMP proteins vary in the structure of their domains as some consist only of a metalloproteinase domain, while others contain DC domains (Gutiérrez and Rucavado 2000). The primary function of these proteins is to induce hemorrhage (Takeda et al. 2012). Proteins identified in the SVMP protein family were zinc metalloproteinases, belonging to the P-III subfamily, that inhibit platelet aggregation and cell migration, cleave specific proteins and induce hemorrhage (Takeya et al. 1990; Guo et al. 2007; Guan et al. 2010).

Members of the CRISP protein family share a common secondary structure, consisting of two main domains and 16 conserved cysteine residues (Tadokoro et al. 2020). Similar to 3FTx, diverse biological activities have been recorded in this protein family (Sunagar et al. 2012; Tadokoro et al. 2020). The CRISP proteins in snake venom have been found to inhibit ion channels, smooth muscle contraction and angiogenesis, and promote vascular permeability and inflammation (Sunagar et al. 2012; Tadokoro et al. 2020). The toxins identified in this study have been found to inhibit smooth muscle contraction through interactions with voltage-gated potassium channels, calcium-activated potassium channels, and calcium release channel receptors (Yamazaki et al. 2002; Chang et al. 2005; Matsunaga et al. 2009).

Kunitz-type serine protease inhibitors are venom peptides with a low molecular mass, characterised by disulfide-rich alpha/beta fold structures and a conserved active site characterized by a low molecular mass (Mukherjee et al. 2014). As with 3FTx and CRISP, proteins belonging to the KI family have been shown to exhibit diverse biological activities, despite structural similarities (Mukherjee et al. 2014). These proteins function by inhibiting serine proteases and blocking specific ion-channels (Mukherjee et al. 2014). This may lead to inflammation, fibrinolysis, disruption of blood pressure regulation and blood coagolutation, causing blood clotting, hemorrhage, hypotension, immobilization and, potentially, death (Mukherjee et al. 2014; Inagaki 2017). All proteins identified in this protein family were serine protease inhibitors (Hokama et al. 1976; Shafqat et al. 1990; St Pierre et al. 2007; Lu et al. 2008; Tsai et al. 2008; Lin et al. 2013; Yang et al. 2014). Some proteins have been found to only affect a single serine protease, such as trypsin (Shafqat et al. 1990) or chymotrypsin (Lin et al. 2013), while others affected a number of serine proteases, such as trypsin, chymotrypsin and elastase (Lu et al. 2008). One of the proteins has been found to

have a dual activity, inhibiting the chymotrypsin and elastase serine proteases as well as voltage-gated potassium channels (Yang et al. 2014).

Nerve growth factors are dimeric protein molecules that play a role in survival and production of nervous cells (Tong et al. 2012). Members of this family are also thought to affect cells in the endocrine and immune systems (Tong et al. 2012). The main function of vNGF is to maintain sensory and sympathetic nervous systems, and the role it plays in snake venom remains unclear at this stage (Trummal et al. 2011; Tong et al. 2012). Some evidence suggests that it may play a role in inhibiting metalloproteinases, possibly preventing autodigestion, or providing protection against proteases found in prey (Wijeyewickrema et al. 2010).

Proteins belonging to the LAO protein family are homodimers with a conserved structure (Costa et al. 2014; Bhattacharjee et al. 2017). These proteins have been shown to induce toxicity and cause predominantly cytotoxic effects, including hemorrhage, apoptosis, edema, hemolysis and inhibition of platelet aggregation (Costa et al. 2014; Bhattacharjee et al. 2017; Jin et al. 2007).

Proteins belonging to the 5'-NT protein family are hydrolytic enzymes that are thought to aid in prey immobilization by generating adenosine, a multitoxin (Dhananjaya and D'Souza 2010). Although the biological activities of 5'-NT are not well-studied, some evidence suggests that these enzymes have anticoagulant effects and inhibit platelet aggregation (Ogawa et al. 2009; Dhananjaya and D'Souza 2010).

3.4.1.2 Minor protein families

Nine of the 20 protein families identified belonged to this group. These included vespryn, PDE, VEGF, IGF, GPx, SVSP, CHIT1 and PLB. The DC domain also falls under this group, but as proteins belonging to this family are SVMP's, its properties were described in section 3.4.1.1 above.

The biological activities of a number of the proteins present in this group have not yet been fully established. Only a single protein, thaicobrin, was identified in the vespryn protein family. This protein is thought to function in a similar manner to ohanin, which aids in subduing prey through neurotoxic effects (Tan et al. 2017). The IGF proteins are considered putative toxins based on the presence of adaptive evolution and selective expression in venom proteomes (Sparkman et al. 2012; Vonk et al.

2013). Glutathione peroxidase is an antioxidant enzyme that is rarely found in snake venom and always occurs in low amounts (Leonardi et al. 2019). The role of GPx in venom is unclear (Viala et al. 2015; Leonardi et al. 2019), but it is thought to play a role in protecting proteins and lipids from oxidative damage caused by hydrogen peroxide (Leonardi et al. 2019). The role of CHIT1 in snake venom also remains unclear, although a study on varanid venom suggests that it may play a role in aiding the digestion of chitinous prey (Koludarov et al. 2017). Limited information is available related to the biological activities of PLB proteins, although studies have found that these proteins are generally involved in hemolytic and cytotoxic activities (Gay et al. 2016; Kunalan et al. 2018).

Other protein families in this group have been found to exhibit diverse effects upon envenomation. The PDE family falls within this group. Proteins within this family have been found to cause hypotension, edema, inhibition of platelet aggregation, and paralysis (Uzair et al. 2018; Ullah et al. 2019). The VEGF protein family has also been shown to exhibit great intraspecific variation in terms of biological activities, with the protein identified in this study playing a role in promoting angiogenesis and permeability of capillaries (Yamazaki et al. 2009). The SVSP protein family plays a role in cell differentiation, system activation, homeostasis and prey digestion (Zaqueo et al. 2014). It has been found to regulate the activation and inactivation of coagulation factors that lead to coagulation, platelet aggregation and fibrinolysis, causing hemotoxic effects (Lee et al. 1999; Zaqueo et al. 2014; Roldán-Padrón et al. 2019). The SVSP protein identified in this study has been found to cause fibrinogenolytic activity, and may also play a role in thrombolytic and antithrombotic reactions (Lee et al. 1999).

3.4.1.3 Unique protein families

Only three protein families belonged to this group. These families were all found only in Pool 4 and made up less than 1% of the overall venom composition of this pool. The families identified were AChE, CVF and SA.

Members of the AChE family play a role in cholinergic transmission and have been found to produce neurotoxic effects upon enevenomation (Cousin and Bon 1997; Runakwa et al. 2013). Interestingly, CVF is not a toxin, but functions by facilitating venom entry into the bloodstream (Kock et al. 2004; Vogel and Fritzinger 2017).

Furthermore, SA proteins also do not appear to play a direct role in venom toxicity, but potentially function in venom neutralization, suggesting a role in defense (Clark and Voris 1969; Straight et al. 1976).

3.4.2 Venom variation across the geographic range

The exact protein composition and biological activities of the venom would need to be confirmed using more intensive methods such as top-down proteomics and in-vivo testing (Melani et al. 2017; Sánchez et al. 2018). Bottom-up approaches to venom proteomics, as used in this study, are not sufficient to quantify or identify toxin proteoforms due to the protein inference problem (Melani et al. 2017; Sánchez et al. 2018). Proteoforms refer to different variants of proteins from the same toxin (Melani et al. 2017). Bottom-up approaches may lead to ambiguities in determining the identity of a specific toxin protein, as the same peptide is often present in multiple proteoforms, leading to what is termed the protein inference problem (Melani et al. 2017). As my approach did not allow for accurate estimation of the number of proteoforms present in the venom samples, variations in the venom are only discussed at the protein family level.

No major variation was observed in the proteomic composition of the six pools. The major protein families, discussed above, made up the largest component of the venom in all six pools. The minor protein families, along with the unique protein families only made up a small proportion of the overall venom composition. Despite the overall similarity, I observed slight variations in the minor and unique protein families along with variations in the proportions of the protein families present in each pool. Samples collected from populations from the eastern regions of the Gauteng, Mpumalanga and North West provinces varied from those collected from the more southern regions of Gauteng and the KwaZulu-Natal and Eastern Cape, although a degree of distributional overlap was observed. The Western Cape samples had the most unique venom composition. Interpopulation venom variation over a geographic range is a common occurrence in snakes and has been described in a number of species, including *Crotalus scutulatus scutulatus* (Wilkinson et al. 1991), *Daboia russelii* (Sharma et al. 2014), *Naja naja* (Dissanayake et al. 2018), *Naja atra* (Huang et al. 2015), and *Naja kaouthia* (Tan et al. 2017). Geographic variation has been observed in response to seasonal variation, habitat variability, and prey variability (Chippaux et al. 1991; Chang

et al. 2015; Aird et al. 2017; Strickland et al. 2018). Changes in venom composition occur due to the strong selection pressure experienced as a result of the importance of venom in prey capture and defense (Aird et al. 2017; Strickland et al. 2018). Habitat heterogeneity may affect venom composition as heterogenous environments cause species to be subjected to a range of selection pressures (Chang et al. 2015). Seasonal and habitat variability are known to affect prey availability, leading to diet and activity shifts in snakes (Carfagno and Weatherhead 2006; Christian et al. 2007; Willson et al. 2010; Michael et al. 2014). Local adaptation in response to prey has been recorded in a number of snake species, including *Crotalus oreganus* (Holding et al. 2016)*, Crotalus scutulatus* (Strickland et al. 2018), and *Crotalus adamanteus* (Margres et al. 2015)*.* These factors may explain why the Western Cape population contained the most unique features. This population occurs in fynbos whereas the eastern populations occur in grassland habitats (Branch 1998; Marais 2004). These populations are subjected to seasonal variation as the Western Cape population occurs in a different climatic region with winter rainfall, whereas the eastern populations occur in summer rainfall regions (Stock and Lewis 1986; Carbutt et al. 2011). Both fynbos and grassland habitats are characterized by high degrees of endemism, therefore it will not be surprising if differences are observed in the diet of *H. haemachatus* between these areas. Unfortunately, little information is available on specific prey species targeted by *H. haemachatus.* This species appears to have a preference for frogs and toads but is mostly considered opportunistic (Marais 2004; Alexander et al. 2012).

Seasonal, habitat and prey variability, or a combination of these factors, may be driving venom variability within this species. However, the variations observed may also be attributed to intrapopulation/individual variation (Amazonas et al. 2019). Intrapopulation variation has been observed in a number of snake species, including *Vipera berus berus* (Malina et al. 2017), *Bothrops atrox* (Amazonas et al. 2019), *Bothrops moojeni* (da Silva Aguiar et al. 2019), and *Micrurus mipartitus* (Henao Duque and Nuñez Rangel 2016). As my sample size is limited, more extensive sampling is required to confirm this possibility.

3.5 Conclusion

The venom composition of *H. haemachatus* is relatively consistent across its distribution. Protein families that make up the majority of the venom were found in all samples tested. The largest degree of variation, although limited, is observed between the Western Cape samples and the samples collected from eastern regions. The interpopulation variation observed may be due to seasonal, prey and habitat variability, or a combination of these, across the distribution of this species (Chippaux et al. 1991; Aird et al. 2017; Strickland et al. 2018). However, the possibility of intrapopulation variability should also be investigated.

Neurotoxic and cytotoxic elements were prevalent in the venom of *H. haemachatus.* A strong hemotoxic element was also present. Cardiotoxic elements were also identified, although these were less prevalent. These elements are responsible for diverse biological activities, and the reported activities fall in line with symptoms observed upon envenomation (Müller et al. 2012; Sánchez et al. 2018). As snake venom is a rich assortment of toxins and immune system reactions may vary between snakebite victims (León et al. 2011; Stone et al. 2013), it is not surprising that variable symptoms have been observed upon envenomation. The differences in the proportions of the protein families and the presence or absence of specific toxins in the venom may explain why different symptoms are observed in dog envenomations from different regions, however a range of other factors should also be considered. Examples of such factors include the site of envenomation, the amount of venom injected, victim age and health status, and treatment administered (Hayes et al. 2002; Sankar et al. 2013; Stone et al. 2013; Firth et al. 2016; Parker-Cote and Meggs 2018).

As bottom-up approaches did not allow precise identification of proteins present in the venom composition, top-down approaches will need to be implemented in future to get a more accurate measure of which specific proteins are present in the venom of *H. haemachatus* (Melani et al. 2017). I also suggest that the venom composition be determined for samples individually, as subtle differences may be lost by pooling of venom samples (Schenberg 1963; Chippaux et al. 1991; Lourenço et al. 2013). Specific biological activities of the venom proteins will need to be tested and verified in-vivo. Despite the abovementioned shortcomings, I believe that the samples collected and analysed were sufficient in providing an overview of the venom

composition of *H. haemachatus* across its range, as a large portion of the distribution was covered in our sampling effort. Therefore, this project can be considered a baseline for further studies on variation in venom composition of venomous snakes in South Africa.

Chapter Four

General summary and conclusion

Hemachatus haemachatus populations experienced cycles of vicariance and expansion during glacial and interglacial cycles in the Pleistocene epoch, which gave rise to two broad lineages. The range of this species expanded during the LGM, in response to increased aridification leading to the expansion of open grassland habitats (Barlow et al. 2013; Lima-Rezende et al. 2019). As this species has a wide thermal tolerance, it was somewhat resilient to slight changes in climatic conditions during the late Pleistocene, described by abrupt warming and cooling events, allowing it to remain relatively widespread (Alexander 1996; Lee-Thorp et al. 2001; Liu et al. 2012; Zhao et al. 2016; Camuera et al. 2019). The range of this species continued to expand during the Holocene epoch, and it is currently widespread throughout the eastern regions of southern Africa, although only fragmented populations now occur in the western and northern Cape regions and eastern Zimbabwe (Branch 1998; Alexander 2018).

I did not observe a definitive link between the venom composition and phylogeographic patterns of *H. haemachatus* lineages. Only four of the venom samples collected belonged to Lineage 2, whereas the remaining samples belonged to Lineage 1. Pools 1, 3, and 4 contained samples belonging to Lineage 1, and Pool 6 contained samples belonging to Lineage 2. Discrepancies were found in Pools 2 and 5, which contained one sample from Lineage 1 and two samples from Lineage 2, and one sample from Lineage 1 and one from Lineage 2, respectively. If patterns of variation were genetically linked, the dendogram produced using the venom data would have mirrored the phylograms produced using the sequence data. However, as mentioned earlier, definitive relationships can only be inferred if the venom composition for each venom sample is investigated individually (Chippaux et al. 1991).

Variation observed in the venom pools appear to mirror the current distribution of the species, indicating that the variations are likely related to factors associated with the distribution of this species. Venom variation is common between geographically distinct populations and has been observed in other species of venomous snakes as well, including *Crotalus lepidus lepidus* (Forstner et al. 1997), *Naja naja* (Dissanayake

et al. 2018) and *Vipera berus berus* (Malina et al. 2013). However, the possibility of intrapopulation variation should also be investigated.

To conclude, *H. haemachatus* is a widespread species with a preference for cooler climatic conditions and open grassland environments. It has a wide thermal tolerance, which allowed it to remain relatively widespread since its range expansion during the LGM. This species shows limited genetic and venom variability across its distribution, although increased genetic structuring and subtle variations in venom composition are observed within the Cape provinces of South Africa. As expected, the venom was rich in both neurotoxic and cytotoxic elements; a strong hemotoxic and cardiotoxic elements were also identified. There is no definitive link between the venom composition at the family level and phylogeographic patterns of this species. Thus, patterns observed in venom variability may not reflect phylogeographic patterns observed and may be a result of distribution-related factors, such as geographic isolation, seasonal variation, habitat and prey variability, or intrapopulation variation. The possibility of a link between the phylogeographic patterns and the venom composition of populations at the individual protein level will need to be confirmed. Increased sampling and individual analysis of venom samples is required to provide a more detailed representation of inter- and intrapopulation venom variability.

References

Aanensen, D., M. Baguelin, P. Birrell, S. Cauchemez, A. Camacho, C. Colijn, A. Cori, X. Didelot, K. Eames, C. Fraser, S. Frost, N. Hens, J. Hugues, T. Jombart, L. Opatowski, O. Ratmann, S. Soubeyrand, M. Suchard, J. Wallinga, and R. Ypma. 2015. Package 'OutbreakTools'. Available at http://sites.google.com/site/therepiproject/r-pac/about.

Aguilar, I., B. Guerrero, A.M. Salazar, M.E. Giro, E.E. Sanchez, J.C. Pérez, and A. Rodriguez-Acosta. 2007. Individual venom variability in the South American rattlesnake *Crotalus durissus cumanensis*. *Toxicon* 50:214–224.

Ahmed, S.M., M. Ahmed, A. Nadeem, J. Mahajan, A. Choudhary, and J. Pal. 2008. Emergency treatment of a snake bite: pearls from literature. *J. Emerg. Trauma. Shock* 1:97–105.

Aiello-Lammens, M.E., R.A. Boria, A. Radosavljevic, B. Vilela, and R.P. Anderson. 2019. spThin: functions for spatial thinning of species occurrence records for use in ecological models. Available at https://cran.r-project.org/package=spThin.

Aird, S.D., J. Arora, A. Barua, L. Qiu, K. Terada, and A.S. Mikheyev. 2017. Population genomic analysis of a pitviper reveals microevolutionary forces underlying venom chemistry. *Genome Biol. Evol.* 9:2640–2649.

Alexander, G.J. 1996. Thermal physiology of *Hemachatus haemachatus* and its implications to range limitation. University of the Witwatersrand, Johannesburg, South Africa.

Alexander, G.J. 2018. *Hemachatus haemachatus*. *The IUCN Red List of Threatened Species* 2018: e.T177556A115659071. Available at https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T177556A115659071.en.

Alexander, G.J., and R. Brooks. 1999. Circannual rhythms of appetite and ecdysis in the Elapid snake, *Hemachatus haemachatus*, appear to be endogenous. *Copeia* 1:146–152.

Alexander, G.J., S.A. Hanrahan, and D. Mitchell. 2012. Assimilation efficiency and gut passage time in an african elapid snake, *Hemachatus haemachatus*. *African J. Herpetol*. 61:3–13.

Alexander, G.J., D. Mitchell, and S.A. Hanrahan. 1999. Wide thermal tolerance in the African Elapid, *Hemachatus haemachatus. J. Herpetol*. 33:164–167.

Amazonas, D.R., L.A. Freitas de Sousa, D.P. Orefice, L.F. de Sousa, M.G. Martinez, R.H.V. Mourão, H.M. Chalkidis, P.B. Camargo, and A.M. Moura da Silva. 2019. Evidence for snake venom plasticity in a long-term study with individual captive *Bothrops atrox. Toxins (Basel).* 11:294.

Avise, J.C. 1989. Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution* 43:1192.

Bandelt, H.J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16:37–48.

Barlow, A., K. Baker, C.R. Hendry, L. Peppin, T. Phelps, K.A. Tolley, C.E. Wüster, and W. Wüster. 2013. Phylogeography of the widespread African puff adder (*Bitis arietans*) reveals multiple Pleistocene refugia in southern Africa. *Mol. Ecol.* 22:1134–1157.

Barnwal, B., C. Jobichen, V.M. Girish, C.S. Foo, J. Sivaraman, and R.M. Kini. 2016. Ringhalexin from *Hemachatus haemachatus*: a novel inhibitor of extrinsic tenase complex. *Sci. Rep.* 6:25935.

Bell, M.A., and G.T. Lloyd. 2015. Strap: an R package for plotting phylogenies against stratigraphy and assessing their stratigraphic congruence. *Palaeontology* 58:379–389.

Bhattacharjee, P., J. Mitra, and D. Bhattacharyya. 2017. L-amino acid oxidase from venoms. Pp. 295–320 in *Toxins and Drug Discovery* (P. Gopalakrishnakone, L.J. Cruz, and S. Luo, eds.). Springer, Dordrecht, Dordrecht.

Blaylock, R.S. 2005. The identification and syndromic management of snakebite in South Africa. *South African Fam. Pract.* 47:48–53.

Bond, W., and N.P. Zaloumis. 2016. The deforestation story: testing for anthropogenic origins of Africa's flammable grassy biomes. *Philos. Trans. R. Soc. B* 371:20150170.

Bonnaterre, A. 1790. Tableau encyclopédique et méthodique des trois règnes de la nature. Chez Panckoucke, libraire, hôtel de Thou, rue des Poitevins, Paris.

Botes, D.P., and D.J. Strydom. 1969. A neurotoxin, toxin alpha, from Egyptian cobra (*Naja haje haje*) venom. I. Purification, properties, and complete amino acid sequence. *J. Biol. Chem.* 244:4147–4157.

Bouckaert, R., T.G. Vaughan, J. Barido-Sottani, S. Duchêne, M. Fourment, A. Gavryushkina, J. Heled, G. Jones, D. Kühnert, N. de Maio, M. Matschiner, F.K. Mendes, N.F. Müller, H.A. Ogilvie, L. du Plessis, A. Popinga, A. Rambaut, D. Rasmussen, I. Siveroni, M.A. Suchard, C.H. Wu, D. Xie, C. Zhang, T. Stadler, and A.J. Drummond. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS *Comput. Biol.* 15:e1006650.

Boussau, B., and V. Daubin. 2010. Genomes as documents of evolutionary history. *Trends Ecol. Evol.* 25:224–232.

Branch, B. 1998. *Field guide to snakes and other reptiles of southern Africa*. 3rd ed. (P. Parker and H. Reid, eds.). Struik Publishers, Cape Town. 109–110 pp.

Brown, J.L., D.J. Hill, A.M. Dolan, A.C. Carnaval, and A.M. Haywood. 2018. PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Nature* 5:180254.

Brown, N.I. 2012. Consequences of neglect: analysis of the sub-Saharan African snake antivenom market and the global context. *PLoS Negl. Trop. Dis.* 6:e1670.

Burbrink, F.T., and T.A. Castoe. 2009. Molecular phylogeography of snakes. Pp. 38–77 in *Snakes: Ecology and Conservation* (S.J. Mullin and R.A. Seigel, eds.). Cornell University Press, New York.

Burbrink, F.T., F. Fontanella, R. Alexander Pyron, T.J. Guiher, and C. Jimenez. 2008. Phylogeography across a continent: the evolutionary and demographic history of the North American racer (Serpentes: Colubridae: *Coluber constrictor*). *Mol. Phylogenet. Evol.* 47:274–288.

Burbrink, F.T., and T.J. Guiher. 2015. Considering gene flow when using coalescent methods to delimit lineages of North American pitvipers of the genus *Agkistrodon*. *Zool. J. Linn. Soc.* 173:505–526.

Burbrink, F.T., R. Lawson, and J.B. Slowinski. 2000. Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. *Evolution* 54:2107–2118.

Butts, C.T., D. Hunter, M. Handcock, S. Bender de Moll, J. Horner, and L. Wang. 2019. Package 'network'. Available at http://statnet.org/.

Caldwell, M.W., R.L. Nydam, A. Palci, and S. Apesteguía. 2015. The oldest known snakes from the middle Jurassic-lower Cretaceous provide insights on snake evolution. *Nat. Commun.* 6:1–11.

Calvete, J.J. 2014. Next-generation snake venomics: protein-locus resolution through venom proteome decomplexation. *Expert Rev. Proteomics* 11:315–329.

Camuera, J., G. Jiménez-Moreno, M.J. Ramos-Román, A. García-Alix, F. Jiménez-Espejo, J.L. Toney, R.S. Anderson, and C. Webster. 2019. Climatic subdivision of Heinrich Stadial 1 based on centennial-scale paleoenvironmental changes observed in the western Mediterranean area. *Clim. Past.* Accepted for Review.

Carbutt, C., M. Tau, A. Stephens, and B. Escott. 2011. The conservation status of temperate grasslands in southern Africa. *Grassroots* 11:17–23.

Card, D.C., D.R. Schield, R.H. Adams, A.B. Corbin, B.W. Perry, A.L. Andrew, G.I.M. Pasquesi, E.N. Smith, T. Jezkova, S.M. Boback, W. Booth, and T.A. Castoe. 2016. Phylogeographic and population genetic analyses reveal multiple species of *Boa* and independent origins of insular dwarfism. *Mol. Phylogenet. Evol.* 102:104–116.

Carfagno, G.L.F., and P.J. Weatherhead. 2006. Intraspecific and interspecific variation in use of forest-edge habitat by snakes. *Can. J. Zool.* 84:1440–1452.

Casewell, N.R., S.C. Wagstaff, W. Wüster, D.A.N. Cook, F.M.S. Bolton, S.I. King, D. Pla, L. Sanz, J.J. Calvete, and R.A. Harrison. 2014. Medically important differences in

snake venom composition are dictated by distinct postgenomic mechanisms. *Proc. Natl. Acad. Sci.* 111:9205–9210.

Castresana, J. 2001. Cytochrome b phylogeny and the taxonomy of great apes and mammals. *Mol. Biol. Evol.* 18:465–471.

Chang, D., A.M. Olenzek, and T.F. Duda. 2015. Effects of geographical heterogeneity in species interactions on the evolution of venom genes. *Proc. R. Soc. B* 282:20141984.

Chang, L.S., J.C. Liou, S.R. Lin, and Y.C. Cheng. 2005. Purification and characterization of Taiwan cobra venom proteins with weak toxicity. *Toxicon* 45:21–45.

Chien, K.Y., C.M. Chiang, Y.C. Hseu, A.A. Vyas, G.S. Rule, and W. Wu. 1994. Two distinct types of cardiotoxin as revealed by the structure and activity relationship of their interaction with zwitterionic phospholipid dispersions. *J. Biol. Chem.* 269:14473–14483.

Chippaux, J.P. 2011. Estimate of the burden of snakebites in sub-Saharan Africa: a meta-analytic approach. *Toxicon* 57:586–599.

Chippaux, J.P., V. Williams, and J. White. 1991. Snake venom variability: methods of study, results and interpretation. *Toxicon* 29:1279–1303.

Christian, K., J.K. Webb, T. Schultz, and B. Green. 2007. Effects of seasonal variation in prey abundance on field metabolism, water flux, and activity of a tropical ambush foraging snake. *Physiol. Biochem. Zool.* 80:522–533.

Clark, W.C., and H.K. Voris. 1969. Venom neutralization by rattlesnake serum albumin. *Science* 164:1402–1404.

Condrea, E., J.E. Fletcher, B.E. Rapuano, C.C. Yang, and P. Rosenberg. 1981. Effect of modification of one histidine residue on the enzymatic and pharmacological properties of a toxic phospholipase A2 from *Naja nigricollis* snake venom and less toxic phospholipases A2 from *Hemachatus haemachatus* and *Naja atra* snake venoms. *Toxicon* 19:61–71.

Costa, T.R., S.M. Burin, D.L. Menaldo, F.A. de Castro, and S. V. Sampaio. 2014. Snake venom L-amino acid oxidases: an overview on their antitumor effects. *J. Venom. Anim. Toxins Incl. Trop. Dis.* 20:23.

Cousin, X., and C. Bon. 1997. Acetylcholinesterase from snake venoms. *C. R. Séances Soc. Biol. Fil.* 191:381–400.

Cox, C.L., and P.T. Chippindale. 2014. Patterns of genetic diversity in the polymorphic ground snake (*Sonora semiannulata*). *Genetica* 142:361–370.

Cusimano, N., and C. Heibl. 2013. Phyloch-package: phylogenetic data analysis and display. Available at https://rdrr.io/github/fmichonneau/phyloch/man/phyloch-package.html.

Da Silva Aguiar, W., N. da Costa Galizio, C. Serino-Silva, S.S. Sant'Anna, K. Fernandes Grego, A.K. Tashima, E.S. Nishiduka, K. de Morais-Zani, and A.M. Tanaka-Azevedo. 2019. Comparative compositional and functional analyses of *Bothrops moojeni* specimens reveal several individual variations. *PLoS One* 14:e0222206.

DeMenocal, P.B. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet. Sci. Lett.* 220:3–24.

Dhananjaya, B.L., and C.J. D'Souza. 2010. The pharmacological role of nucleotidases in snake venoms. *Cell Biochem. Funct.* 28:171–177.

Dissanayake, D.S.B., L.D. Thewarage, R.N. Waduge, J.G.S. Ranasinghe, S.A.M. Kularatne, and R.P. V. Jayanthe Rajapakse. 2018. The venom of spectacled cobra (Elapidae: *Naja naja*): in vitro study from distinct geographical origins in Sri Lanka. *J. Toxicol.* 2018:7358472.

Dowell, K. 2008. Molecular phylogenetics: an introduction to computational methods and tools for analyzing evolutionary relationships. Available at http://www.math.umaine.edu/~khalil/courses/MAT500/papers/MAT500_Paper_Phylogenetics.pdf.

Doyle, J.J., and L.H.B. Hortorium. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.* 17:144–163.

Drummond, A.J., S.Y.W. Ho, M.J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.

Drummond, A.J., A. Rambaut, B. Shapiro, and O.G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22:1185–1192.

Dunn, C.W., G. Giribet, G.D. Edgecombe, and A. Hejnol. 2014. Animal phylogeny and its evolutionary implications. *Annu. Rev. Ecol. Evol. Syst.* 45:371–395.

Ecker, M., J.S. Brink, L. Rossouw, M. Chazan, L.K. Horwitz, and J.A. Lee-Thorp. 2018. The palaeoecological context of the Oldowan-Acheulean in southern Africa. *Nat. Ecol. Evol.* 2:1080–1086.

Edwards, W.J., and C.T. Edwards. 2011. Population limiting factors. *Nat. Educ. Knowl.* 3:1.

Eichberg, S., L. Sanz, J.J. Calvete, and D. Pla. 2015. Constructing comprehensive venom proteome reference maps for integrative venomics. *Expert Rev. Proteomics* 12:557–573.

Ettling, J.A., and P.G. Parker. 2017. Genetic diversity and population structure of Armenian vipers, *Montivipera raddei*, in two landscapes: implications for conservation. *Herpetol. Conserv. Biol.* 12:435–446.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the Bootstrap. *Evolution* 39:783–791.

Fick, S.E., and R.J. Hijmans. 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37:4302–4315.

Firth, G.B., M. Street, Y. Ramguthy, and L. Doedens. 2016. Mortality following snake bite envenomation by *Bitis arietans* in an HIV positive child. *Medicine (Balt.)* 95:e4001.

Fohlman, J., D. Eaker, E. Karlsoon, and S. Thesleff. 1976. Taipoxin, an extremely potent presynaptic neurotoxin from the venom of the Australian snake taipan

(*Oxyuranus s. scutellatus*). Isolation, characterization, quaternary structure and pharmacological properties. *Eur. J. Biochem.* 68:457–469.

Fordham, D.A., F. Saltré, S. Haythorne, T.M.L. Wigley, B.L. Otto-Bliesner, K. Ching Chan, and B.W. Brook. 2017. PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography (Cop.)* 40:1348–1358.

Forstner, M.R.J., R.A. Hilsenbeck, and J.F. Scudday. 1997. Geographic variation in whole venom profiles from the Mottled Rock rattlesnake (*Crotalus lepidus lepidus*) in Texas. *J. Herpetol.* 31:277–287.

Fry, B.G., H. Scheib, L. van der Weerd, B. Young, J. McNaughtan, S.F. Ryan Ramjan, N. Vidal, R.E. Poelmann, and J.A. Norman. 2008. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol. Cell. Proteomics* 7:215–246.

Fryklund, L., and D. Eaker. 1973. Complete amino acid sequence of a nonneurotoxic hemolytic protein from the venom of *Haemachatus haemachates* (African ringhals cobra). *Biochemistry* 12:661–667.

Fu, Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.

Fu, Y.X., and W.H. Li. 1999. Coalescing into the 21st century: An overview and prospects of coalescent theory. *Theor. Popul. Biol.* 56:1–10.

Gay, C., L. Sanz, J.J. Calvete, and D. Pla. 2016. Snake venomics and antivenomics of *Bothrops diporus*, a medically important pitviper in northeastern Argentina. *Toxins* 8:9.

Giannasi, N., R.S. Thorpe, and A. Malhotra. 2001. The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, *Trimeresurus albolabris*. *Mol. Ecol.* 10:419–426.

Gibbard, P., and M.J. Head. 2009. The definition of the quaternary system/era and the Pleistocene series/epoch. *Quaternaire* 20:125–133.

Gibbs, H.L., L. Sanz, J.E. Chiucchi, T.M. Farrell, and J.J. Calvete. 2011. Proteomic analysis of ontogenetic and diet-related changes in venom composition of juvenile and adult Dusky Pigmy rattlesnakes (*Sistrurus miliarius barbouri*). *J. Proteomics* 74:2169–2179.

Girish, V.M., and R.M. Kini. 2016. Exactin: a specific inhibitor of Factor X activation by extrinsic tenase complex from the venom of *Hemachatus haemachatus. Sci. Rep.* 6:32036.

Girish, V.M., S. Kumar, L. Joseph, C. Jobichen, R.M. Kini, and J. Sivaraman. 2012. Identification and structural characterization of a new three-finger toxin hemachatoxin from *Hemachatus haemachatus* venom. *PLoS One* 7:e48112.

Graham, L.R.J., C. Graham, D. Theakston, G. Mcmullan, and C. Shaw. 2008. Elucidation of trends within venom components from the snake families Elapidae and Viperidae using gel filtration chromatography. *Toxicon* 51:121–129.

Groth, J.G., and G.F. Barrowclough. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol. Phylogenet. Evol.* 12:115–123.

Guan, H.H., K.S. Goh, F. Davamani, P.L. Wu, Y.W. Huang, J. Jeyakanthan, W.G. Wu, and C.J. Chen. 2010. Structures of two elapid snake venom metalloproteases with distinct activities highlight the disulfide patterns in the D domain of ADAMalysin family proteins. *J. Struct. Biol.* 169:294–303.

Guicking, D., A. Herzberg, and M. Wink. 2004. Population genetics of the Dice snake (*Natrix tessellata*) in Germany: implications for conservation. *Salamandra* 4:217–234.

Guicking, D., U. Joger, and M. Wink. 2009. Cryptic diversity in a Eurasian water snake (*Natrix tessellata*, Serpentes: Colubridae): evidence from mitochondrial sequence data and nuclear ISSR-PCR fingerprinting. *Org. Divers. Evol.* 9:201–214.

Guiher, T.J., and F.T. Burbrink. 2008. Demographic and phylogeographic histories of two venomous North American snakes of the genus *Agkistrodon*. *Mol. Phylogenet. Evol.* 48:543–553.

Guo, X.X., L. Zeng, W.H. Lee, Y. Zhang, and Y. Jin. 2007. Isolation and cloning of a metalloproteinase from king cobra snake venom. *Toxicon* 49:954–965.

Gutiérrez, J.M., and A. Rucavado. 2000. Snake venom metalloproteinases: their role in the pathogenesis of local tissue damage. *Biochimie* 82:841–850.

Gutiérrez, J.M., R.D.G. Theakston, and D.A. Warrell. 2006. Confronting the neglected problem of snake bite envenoming: the need for a global partnership. *PLoS Med.* 3:0727–0731.

Hayes, W., S.S. Herbert, G.C. Rehling, and J.F. Gennaro. 2002. Factors that influence venom expenditure in viperids and other snake species during predatory and defensive contexts. Pp. 207–233 in *Biology of the Vipers* (G.W. Schuett, M. Hoggren, M.E. Douglas, and H.W. Greene, eds.). Eagle Mountain Publishing, Eagle Mountain, Utah.

He, Y.Y., W.H. Lee, and Y. Zhang. 2004. Cloning and purification of alpha-neurotoxins from king cobra (*Ophiophagus hannah*). *Toxicon* 44:295–303.

Head, J.J., J.H. Hutchison, and R.I. Ciochon. 2005. First report of snakes (Serpentes) from the late middle Eocene Pondaung formation, Myanmar. *J. Vertebr. Paleontol.* 25:246–250.

Head, J.J., K. Mahlow, and J. Müller. 2016. Fossil calibration dates for molecular phylogenetic analysis of snakes 2: Caenophidia, Colubroidea, Elapoidea, Colubridae. *Palaeontol. Electron.* 19.2.2FC:1–21.

Hedges, S.B. 1994. Molecular evidence for the origin of birds. *Proc. Natl. Acad. Sci. U.S.A.* 91:2621–2624.

Heled, J. 2016. Extended Bayesian Skyline Plot tutorial for BEAST2. Available at www.beast2.org/tutorials/.

Heled, J., and A.J. Drummond. 2008. Bayesian inference of population size history from multiple loci. *BMC Evol. Biol.* 8:289.

Henao Duque, A.M., and V. Nuñez Rangel. 2016. Maintenance of red-tail coral snake (*Micrurus mipartitus*) in captivity and evaluation of individual venom variability. *Acta Biológica Colomb.* 21:593–600.

Hewitt, G.M. 2004. The structure of biodiversity – insights from molecular phylogeography. *Front. Zool.* 1:4.

Hillis, D.M., and J.J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42:182–192.

Hofmann, S. 2012. Population genetic structure and geographic differentiation in the Hot Spring snake *Thermophis baileyi* (Serpentes, Colubridae): indications for glacial refuges in southern-central Tibet. *Mol. Phylogenet. Evol.* 63:396–406.

Hokama, Y., S. Iwanaga, T. Tatsuki, and T. Suzuki. 1976. Snake venom proteinase inhibitors. III. Isolation of five polypeptide inhibitors from the venoms of *Hemachatus haemachatus* (Ringhal's corbra) and *Naja nivea* (Cape cobra) and the complete amino acid sequences of two of them. *J. Biochem.* 79:559–578.

Holding, M.L., J.E. Biardi, and H.L. Gibbs. 2016. Coevolution of venom function and venom resistance in a rattlesnake predator and its squirrel prey. *Proc. R. Soc. B Biol. Sci.* 283:20152841.

Huang, H.W., B.S. Liu, K.Y. Chien, L.C. Chiang, S.Y. Huang, W.C. Sung, and W.G. Wu. 2015. Cobra venom proteome and glycome determined from individual snakes of *Naja atra* reveal medically important dynamic range and systematic geographic variation. *J. Proteomics* 128:92–104.

Hurst, G.D., and F.M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. B Biol. Sci.* 272:1525–1534.

Huson, D.H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23:254–267.

Inagaki, H. 2017. Snake venom protease inhibitors: enhanced identification, expanding biological function, and promising future. Pp. 161–186 in *Snake Venoms*

(P. Gopalakrishnakone, H. Inagaki, C.W. Vogel, A.K. Mukherjee, and T.R. Rahmy, eds.). Springer, Dordrecht, Dordrecht.

Inoue, K., B.K. Lang, and D.J. Berg. 2015. Past climate change drives current genetic structure of an endangered freshwater mussel species. *Mol. Ecol.* 24:1910–1926.

Ivanov, M. 2002. The oldest known Miocene snake fauna from Central Europe: Merkur-North locality, Czech Republic. *Acta Palaeontol. Pol.* 47:513–534.

Jiang, Z.J., T.A. Castoe, C.C. Austin, F.T. Burbrink, M.D. Herron, J.A. McGuire, C.L. Parkinson, and D.D. Pollock. 2007. Comparative mitochondrial genomics of snakes: extraordinary substitution rate dynamics and functionality of the duplicate control region. *BMC Evol. Biol.* 7:123.

Jin, Y., W.H. Lee, L. Zeng, and Y. Zhang. 2007. Molecular characterization of L-amino acid oxidase from king cobra venom. *Toxicon* 50:479–489.

Joubert, F.J. 1975a. *Hemachatus haemachatus* (ringhals) venom. Purification, some properties and amino-acid sequence of Phospholipase A (fraction DE-I). *Eur. J. Biochem.* 52:539–554.

Joubert, F.J. 1975b. The amino acid sequences of three toxins (CM-10, CM-12 and CM-14) from *Naja haje annulifera* (Egyptian cobra) venom. *Hoppe. Seylers. Z. Physiol. Chem.* 356:53–72.

Joubert, F.J. 1977. Snake venom toxins. The amino-acid sequences of three toxins (9B, 11 and 12A) from *Hemachatus haemachatus* (ringhals) venom. *Eur. J. Biochem.* 74:387–397.

Joubert, F.J., and N. Taljaard. 1980. The complete primary structure of toxin CM-1b from *Hemachatus haemachatus* (ringhals) snake venom. *Toxicon* 18:191–198.

Kasturiratne, A., A.R. Wickremasinghe, N. De Silva, N.K. Gunawardena, A. Pathmeswaran, R. Premaratna, L. Savioli, D.G. Lalloo, and H.J. De Silva. 2008. The global burden of snakebite: a literature analysis and modelling based on regional estimates of envenoming and deaths. *PLoS Med.* 5:1591–1604.

Keogh, J.S. 1998. Molecular phylogeny of elapid snakes and a consideration of their biogeographic history. *Biol. J. Linn.* Soc. 63:177–203.

Kim, H.S., and N. Tamiya. 1982. Amino acid sequences of two novel long-chain neurotoxins from the venom of the sea snake *Laticauda colubrina*. *Biochem. J.* 207:215–223.

Kissling, W.D., A. Blach-Overgaard, R.E. Zwaan, and P. Wagner. 2016. Historical colonization and dispersal limitation supplement climate and topography in shaping species richness of African lizards (Reptilia: Agaminae). *Sci. Rep.* 6:1–14.

Kocher, T.D., W.K. Thomas, A. Meyer, S. V. Edwards, S. Pääbo, F.X. Villablanca, and A.C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. U.S.A.* 86:6196–6200.

Kock, M.A., B.E. Hew, H. Bammert, D.C. Fritzinger, and C.W. Vogel. 2004. Structure and function of recombinant cobra venom factor. *J. Biol. Chem.* 279:30836–30843.

Kohlhoff, M., M.H. Borges, A. Yarleque, C. Cabezas, M. Richardson, and E.F. Sanchez. 2012. Exploring the proteomes of the venoms of the Peruvian pit vipers *Bothrops atrox*, *B. barnetti* and *B. pictus*. *J. Proteomics* 75:2181–2195.

Koludarov, I., T.N. Jackson, B. op den Brouw, J. Dobson, D. Dashevsky, K. Arbuckle, C.J. Clemente, E.J. Stockdale, C. Cochran, J. Debono, C. Stephens, N. Panagides, B. Li, M.L.R. Manchadi, A. Violette, R. Fourmy, I. Hendrikx, A. Nouwens, J. Clements, P. Martelli, H.F. Kwok, and B.G. Fry. 2017. Enter the dragon: the dynamic and multifunctional evolution of Anguimorpha lizard venoms. *Toxins* 9:242.

Krysko, K.L., L.P. Nuñez, C.A. Lippi, D.J. Smith, and M.C. Granatosky. 2016. Pliocene-Pleistocene lineage diversifications in the Eastern Indigo snake (*Drymarchon couperi*) in the southeastern United States. *Mol. Phylogenet. Evol.* 98:111–122.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA 7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33:1870–1874.

Kunalan, S., I. Othman, S. Syed Hassan, and W.C. Hodgson. 2018. Proteomic characterization of two medically important Malaysian snake venoms, *Calloselasma rhodostoma* (Malayan pit viper) and *Ophiophagus hannah* (king cobra). *Toxins* 10:434.

Lang, D.T., and The CRAN Team. 2019. Tools for parsing and generating XML within R and S-Plus. Available at http://www.omegahat.net/RSXML.

Lawson, R., J.B. Slowinski, and F.T. Burbrink. 2004. A molecular approach to discerning the phylogenetic placement of the Enigmatic snake *Xenophidion schaeferi* among the Alethinophidia. *J. Zool.* 263:285–294.

Leavitt, D.H., and D.A. McClellan. 2014. Evolution of cytochrome b in Elapidae (Squamata: Serpentes). *J. Undergrad. Res.* Available at http://jur.byu.edu/?p=9491.

Lee-Thorp, J.A., K. Holmgren, S.-E. Lauritzen, H. Linge, A. Moberg, T.C. Partridge, C. Stevenson, and P.D. Tyson. 2001. Rapid climate shifts in the southern African interior throughout the mid to late Holocene. *Geophys. Res. Lett.* 28:4507–4510.

Lee, J.W., J.H. Seu, I.K. Rhee, I. Jin, Y. Kawamura, and W. Park. 1999. Purification and characterization of brevinase, a heterogeneous two-chain fibrinolytic enzyme from the venom of Korean snake, *Agkistrodon blomhoffii brevicaudus*. *Biochem. Biophys. Res. Commun.* 260:665–670.

Lee, M.S.Y., K.L. Sanders, B. King, and A. Palci. 2016. Diversification rates and phenotypic evolution in venomous snakes (Elapidae). *R. Soc. Open Sci.* 3:150277.

León, G., L. Sánchez, A. Hernández, M. Villalta, M. Herrera, Á. Segura, R. Estrada, and J.M. Gutiérrez. 2011. Immune response towards snake venoms. *Inflamm. Allergy - Drug Targets* 10:000–000.

Leonardi, A., T. Sajevic, J. Pungerčar, and I. Križaj. 2019. Comprehensive study of the proteome and transcriptome of the venom of the most venomous European viper: discovery of a new subclass of ancestral snake venom metalloproteinase precursor-derived proteins. *J. Proteome Res.* 18:2287–2309.

Li, S., J. Wang, X. Zhang, Y. Ren, N. Wang, K. Zhao, X. Chen, and C. Zhao. 2004. Proteomic characterization of two snake venoms: *Naja naja atra* and *Agkistrodon halys. Biochem. J.* 127:119–127.

Lima-Rezende, C.A., A. V. Rocha, A.F.J. Couto, E. de Souza Martins, V. Vasconcelos, and R. Caparroz. 2019. Late Pleistocene climatic changes promoted demographic expansion and population reconnection of a Neotropical savanna-adapted bird, *Neothraupis fasciata* (Aves: Thraupidae). *PLoS One* 14:e0212876.

Lin, S.R., H.B. Huang, B.N. Wu, and L.S. Chang. 1998. Characterization and cloning of long neurotoxin homolog from *Naja naja atra. Biochem. Mol. Biol. Int. J.* 46:1211–1217.

Lin, Y.J., T. Ikeya, P. Guntert, and L.S. Chang. 2013. NMR solution structure of a chymotrypsin inhibitor from the Taiwan cobra *Naja naja atra. Molecules* 18:8906–8918.

Liu, Z., A.E. Carlson, F. He, E.C. Brady, B.L. Otto-Bliesner, B.P. Briegleb, M. Wehrenberg, P.U. Clark, S. Wu, J. Cheng, J. Zhang, D. Noone, and J. Zhu. 2012. Younger Dryas cooling and the Greenland climate response to CO2. *Proc. Natl. Acad. Sci. U.S.A.* 109:11101–11104.

Lock, B.A., and J. Wellehan. 2015. Ophidia (snakes). Pp. 60–74 in *Fowler's Zoo and Wild Animal Medicine* (E.R. Miller and M.E. Fowler, eds.). W.B. Saunders Ltd., Amsterdam.

Longbottom, J., F.M. Shearer, M. Devine, G. Alcoba, F. Chappuis, D.J. Weiss, S.E. Ray, N. Ray, D.A. Warrell, R. Ruiz de Castañeda, D.J. Williams, S. Hay, and D.M. Pigott. 2018. Vulnerability to snakebite envenoming: a global mapping of hotspots. *Lancet* 392:673–684.

Lorenzen, E.D., C. Masembe, P. Arctander, H.R. Siegismund, and D. Hafner. 2010. A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. *J. Biogeogr.* 37:571–581.

Lourenço, A., C.F.Z. Creste, L.C. de Barros, L.D. dos Santos, D.C. Pimenta, B. Barraviera, and R.S. Ferreira. 2013. Individual venom profiling of *Crotalus durissus terrificus* specimens from a geographically limited region: crotamine assessment and captivity evaluation on the biological activities. *Toxicon* 69:75–81.

Louw, A.I. 1974. Snake venom toxins. The amino acid sequences of three cytotoxin homologues from *Naja mossambica mossambica* venom. *Biochim. Biophys. Acta* 336:481–495.

Lu, J., H. Yang, H. Yu, W. Gao, R. Lai, J. Liu, and X. Liang. 2008. A novel serine protease inhibitor from *Bungarus fasciatus* venom. *Peptides* 29:369–374.

Lukoschek, V., J. Scott Keogh, and J.C. Avise. 2012. Evaluating fossil calibrations for dating phylogenies in light of rates of molecular evolution: A comparison of three approaches. *Syst. Biol.* 61:22–43.

Mackessy, S.P. (ed.). 2010. *Handbook of venoms and toxins of reptiles*. CRC Press, Boca Raton. 3–8 pp.

Mackessy, S.P., and A.J. Saviola. 2016. Understanding biological roles of venoms among the Caenophidia: the importance of rear-fanged snakes. *Integr. Comp. Biol.* 56:1004–1021.

Maddison, W.P. 1997. Gene trees in species trees. *Syst. Biol.* 46:523–536.

Malina, T., G. Babocsa, L. Krecsák, and C. Erdész. 2013. Further clinical evidence for the existence of neurotoxicity in a population of the European adder (*Vipera berus berus*) in eastern Hungary: second authenticated case. *Wilderness Environ. Med.* 24:378–383.

Malina, T., L. Krecsák, A. Westerström, G. Szemán-Nagy, G. Gyémánt, M. M-Hamvas, E.G. Rowan, A.L. Harvey, D.A. Warrell, B. Pál, Z. Rusznák, and G. Vasas. 2017. Individual variability of venom from the European adder (*Vipera berus berus*) from one locality in Eastern Hungary. *Toxicon* 135:59–70.

Marais, J. 2004. *A complete guide to the snakes of southern Africa*. 2nd ed. (E. Bowles, ed.). StruikNature, Cape Town. 113–115 pp.

Margres, M.J., J.J. McGivern, M. Seavy, K.P. Wray, J. Facente, and D.R. Rokyta. 2015. Contrasting modes and tempos of venom expression evolution in two snake species. *Genetics* 199:165–176.

Margres, M.J., J.J. Mcgivern, K.P. Wray, M. Seavy, K. Calvin, and D.R. Rokyta. 2013. Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J. Proteomics* 96:145–158.

Matsunaga, Y., Y. Yamazaki, F. Hyodo, Y. Sugiyama, M. Nozaki, and T. Morita. 2009. Structural divergence of cysteine-rich secretory proteins in snake venoms. *J. Biochem.* 145:365–375.

Melani, R.D., F.C.S. Nogueira, and G.B. Domont. 2017. It is time for top-down venomics. *J. Venom. Anim. Toxins Incl. Trop. Dis.* 23:44.

Mendoza, C.E.C., T. Bhatti, and A.R. Bhatti. 1992. Electrophoretic analysis of snake venoms. *J. Chromatogr. B Biomed. Sci. Appl.* 580:355–363.

Metzger, G.A., F. Kraus, A. Allison, and C.L. Parkinson. 2010. Uncovering cryptic diversity in *Aspidomorphus* (Serpentes: Elapidae): evidence from mitochondrial and nuclear markers. *Mol. Phylogenet. Evol.* 54:405–416.

Michael, D.R., R.B. Cunningham, C. Macgregor, D. Brown, and D.B. Lindenmayer. 2014. The effects of prey, habitat heterogeneity and fire on the spatial ecology of peninsular Diamond Pythons (*Morelia spilota spilota*: Pythonidae). *Austral Ecol.* 39:181–189.

Miyoshi, S., and A.T. Tu. 1996. Phospholipase A2 from *Naja naja sputatrix* venom is a muscarinic acetylcholine receptor inhibitor. *Arch. Biochem. Biophys*. 328:17–25.

Mokhatla, M.M., G. Kemp, D. Rödder, and G.J. Measey. 2015. Assessing the effects of climate change on distributions of Cape Floristic Region amphibians. *S. Afr. J. Sci.* 111:1–7.

Montgelard, C., and C.A. Matthee. 2012. Tempo of genetic diversification in southern African rodents: the role of Plio-Pleistocene climatic oscillations as drivers for speciation. *Acta Oecologica* 42:50–57.

Mucina, L., and M.C. Rutherford (eds.). 2006. *The vegetation of South Africa, Lesotho and Swaziland*. South African National Biodiversity Institute, Pretoria.

Mukherjee, A.K., S.P. Mackessy, and S. Dutta. 2014. Characterization of a Kunitz-type protease inhibitor peptide (Rusvikunin) purified from *Daboia russelii russelii* venom. *Int. J. Biol. Macromol.* 67:154–162.

Müller, G.J., H. Modler, C.A. Wium, D.J.H. Veale, C.J. Marks, H. Modler, C.A. Wium, D.J.H. Veale, and C.J. Marks. 2012. Snake bite in southern Africa: diagnosis and management. *Contin. Med. Educ.* 30:362–380.

Neumann, F.H. 2015. Shaping of modern southern African biomes: Neogene vegetation and climate changes. *Trans. R. Soc. South Africa* 70:195–212.

Nuñez, J.J., N.K. Wood, F.E. Rabanal, F.M. Fontanella, and J.W. Sites. 2011. Amphibian phylogeography in the Antipodes: refugia and postglacial colonization explain mitochondrial haplotype distribution in the Patagonian frog *Eupsophus calcaratus* (Cycloramphidae). *Mol. Phylogenet. Evol.* 58:343–352.

Ogawa, Y., N. Murayama, and R. Yanoshita. 2009. Molecular cloning and characterization of ecto-5'-nucleotidase from the venoms of *Gloydius blomhoffi*. *Toxicon* 54:408–412.

Otting, G., W.E. Steinmetz, P.E. Bougis, H. Rochat, and K. Wuthrich. 1987. Sequence-specific 1H-NMR assignments and determination of the secondary structure in aqueous solution of the cardiotoxins CTXIIa and CTXIIb from *Naja mossambica mossambica. Eur. J. Biochem*. 168:609–620.

Otto-Bliesner, B.L., S.J. Marshall, J.T. Overpeck, G.H. Miller, and A. Hu. 2006. Simulating arctic climate warmth and icefield retreat in the Last Interglaciation. *Science* 311:1751–1753.

Palumbi, S.R. 1996. Nucleic acids II: the polymerase chain reaction. Pp. 205–247 in *Molecular Systematics* (D.M. Hillis, C. Moritz, and B.K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

Pan, F.M., M.S. Yeh, W.C. Chang, C.C. Hung, and S.H. Chiou. 1994. Sequence analysis and expression of phospholipase A2 from Taiwan cobra. *Biochem. Biophys. Res. Commun.* 199:969–976.

Paradis, E., S. Blomberg, B. Bolker, J. Brown, J. Claude, H.S. Cuong, R. Desper, G. Didier, B. Durand, J. Dutheil, R.J. Ewing, O. Gascuel, T. Guillerme, C. Heibl, A. Ives, B. Jones, F. Krah, D. Lawson, V. Lefort, P. Legendre, J. Lemon, E. Marcon, R. McCloskey, J. Nylander, R. Opgen-Rhein, A.A. Popescu, M. Royer-Carenzi, K. Schliep, K. Strimmer, and D. de Vienne. 2019. Analyses of Phylogenetics and Evolution. Available at https://cran.r-project.org/web/packages/ape/ape.pdf.

Parker-Cote, J., and W.J. Meggs. 2018. First aid and pre-hospital management of venomous snakebites. *Trop. Med. Infect. Dis.* 3:45.

Phillips, S.J., M. Dudík, and R.E. Schapire. 2017. Maxent software for modeling species niches and distributions. Available at http://biodiversityinformatics.amnh.org/open_source/maxent/.

Poh, S.L., G. Mourier, R. Thai, A. Armugam, J. Molgo, D. Servent, K. Jeyaseelan, and A. Menez. 2002. A synthetic weak neurotoxin binds with low affinity to Torpedo and chicken alpha7 nicotinic acetylcholine receptors. *Eur. J. Biochem.* 269:4247–4256.

Posada, D. 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25:1253–1256.

Prior, K.A., H.L. Gibbs, and P.J. Weatherhead. 1997. Population genetic structure in the black rat snake: implications for management. *Conserv. Biol.* 11:1147–1158.

QIAGEN. 2017. CLC Main Workbench version 7.9.2. Available at https://www.qiagenbioinformatics.com/products/clc-main-workbench/.

Qian, Y.C., C.Y. Fan, Y. Gong, and S.L. Yang. 1998. cDNA sequence analysis and expression of four long neurotoxin homologues from *Naja naja atra. Biochim. Biophys. Acta* 1443:233–238.

De Queiroz, A., R. Lawson, and J.A. Lemos-espinal. 2002. Phylogenetic relationships of North American garter snakes (*Thamnophis*) based on four mitochondrial genes: how much DNA sequence is enough? *Mol. Phylogenet. Evol.* 22:315–329.

R Core Team. 2014. R: A language and environment for statistical computing version 3.0.3. Available at http://www.r-project.org/.

R Core Team. 2015. R: A language and environment for statistical computing version 3.2.3. Available at http://www.r-project.org/.

R Core Team. 2019. R: A language and environment for statistical computing version 3.6.0. Available at http://www.r-project.org/.

Rafhaella, C.A., D.L. Menaldo, T.R. Costa, K.F. Zoccal, M.A. Sartim, N.A. Santos-Filho, L.H. Faccioli, and S. V. Sampaio. 2018. Cytotoxic and inflammatory potential of a phospholipase A2 from *Bothrops jararaca* snake venom. *J. Venom. Anim. Toxins Incl. Trop. Dis.* 24:33.

Rage, J.C. 1988. The oldest known colubrid snakes: the state of the art. *Acta Zool. Cracoviensia* 31:457–474.

Rage, J.C., S. Bajpai, J.G.M. Thewissen, and B.N. Tiwari. 2003. Early Eocene snakes from Kutch, Western India, with a review of the Palaeophiidae. *Geodiversitas* 25:695–716.

Rage, J.C., and C. Werner. 1999. Mid-Cretaceous (Cenomanian) snakes from Wadi Abu Hashim, Sudan: the earliest snake assemblage. *Palaeontol. Africana* 35:85–110.

Rambaut, A., A.J. Drummond, D. Xie, G. Baele, and M.A. Suchard. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67:901–904.

Roldán-Padrón, O., J.L. Castro-Guillén, J.A. García-Arredondo, M.S. Cruz-Pérez, L.F. Díaz-Peña, C. Saldaña, A. Blanco-Labra, and T. García-Gasca. 2019. Snake venom hemotoxic enzymes: biochemical comparison between *Crotalus* species from central Mexico. *Molecules* 24:1489.

Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Rozas, J., A. Ferrer-Mata, J. Carlos Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S.E. Ramos-Onsins, and A. Sánchez-Gracia. 2017. DnaSP v6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol.* 34:3299–3302.

Runakwa, U.K., D.G. Lalloo, and H.J. de Silva. 2013. Neurotoxicity in snakebite-the limits of our knowledge. *PLoS Negl. Trop. Dis.* 7:e2302.

Rutherford, M.C., L. Mucina, and L.W. Powrie. 2006. Biomes and bioregions of southern Africa. *Strelitzia* 19:46–50.

Sánchez, A., M. Herrera, M. Villalta, D. Solano, Á. Segura, B. Lomonte, J.M. Gutiérrez, G. León, and M. Vargas. 2018. Proteomic and toxinological characterization of the venom of the South African Ringhals cobra *Hemachatus haemachatus. J. Proteomics* 181:104–117.

Sankar, J., R. Nabeel, M.J. Sankar, L. Priyambaba, and S. Mahadevan. 2013. Factors affecting outcome in children with snake envenomation: a prospective observational study. *Arch. Dis. Child.* 98:596–601.

Scanlon, J.D., and M.S.Y. Lee. 2003. Mid-Tertiary elapid snakes (Squamata, Colubroidea) from Riversleigh, northern Australia: early steps in a continent-wide adaptive radiation. *Geobios* 36:573–601.

Schenberg, S. 1963. Immunological (ouchterolony method) identification on intrasubspecies qualitative diffferences in snake venom composition. *Toxicon* 1:67–75.

Scott, L. 2002. Grassland development under glacial and interglacial conditions in southern Africa: review of pollen, phytolith and isotope evidence. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 177:47–57.

Shafqat, J., O.U. Beg, S.J. Yin, Z.H. Zaidi, and H. Jornvall. 1990. Primary structure and functional properties of cobra (*Naja naja naja*) venom Kunitz-type trypsin inhibitor. *Eur. J. Biochem.* 194:337–341.

Sharma, M., N. Gogoi, B.L. Dhananjaya, J.C. Menon, and R. Doley. 2014. Geographical variation of Indian Russell's viper venom and neutralization of its coagulopathy by polyvalent antivenom. *Toxin Rev.* 33:7–15.

Shipolini, R.A., G.S. Bailey, and B.E. Banks. 1974. The separation of neurotoxin from the venom of *Naja melanoleuca* and the primary sequence determination. *Eur. J. Biochem.* 42:203–211.

Shipolini, R.A., M. Kissonerghis, and B.E. Banks. 1975. The primary structure of a major polypeptide component from the venom of *Naja melanoleuca. Eur. J. Biochem.* 56:449–454.

Singh, G., S. Gourinath, K. Saravanan, S. Sharma, S. Bhanumathi, C. Betzel, A. Srinivasan, and T.P. Singh. 2005a. Sequence-induced trimerization of phospholipase A2: structure of a trimeric isoform of PLA2 from common krait (*Bungarus caeruleus*) at 2.5 A resolution. *Acta Crystallogr.* F61:8–13.

Singh, G., J. Jasti, K. Saravanan, S. Sharma, P. Kaur, A. Srinivasan, and T.P. Singh. 2005b. Crystal structure of the complex formed between a group I phospholipase A2 and a naturally occurring fatty acid at 2.7 A resolution. *Protein Sci.* 14:395–400.

Sintayehu, D.W. 2018. Impact of climate change on biodiversity and associated key ecosystem services in Africa: a systematic review. *Ecosyst. Heal. Sustain*. 4:225–239.

Sparkman, A.M., T.S. Schwartz, J.A. Madden, S.E. Boyken, N.B. Ford, J.M. Serb, and A.M. Bronikowski. 2012. Rates of molecular evolution vary in vertebrates for insulin-like growth factor-1 (IGF-1), a pleiotropic locus that regulates life history traits. *Gen. Comp. Endocrinol*. 178:164–173.

St Pierre, L., G.W. Birrell, S.T. Earl, T.P. Wallis, J.J. Gorman, J. De Jersey, P.P. Masci, and M.F. Lavin. 2007. Diversity of toxic components from the venom of the evolutionarily distinct black whip snake, *Demansia vestigiata. J. Proteome Res.* 6:3093–3107.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.

Stamatakis, A., F. Blagojevic, D.S. Nikolopoulos, and C.D. Antonopoulos. 2007. Exploring new search algorithms and hardware for phylogenetics: RAxML meets the IBM cell. *J. VLSI Signal Process. Syst. Signal Image. Video Technol.* 48:271–286.

Stock, W.D., and O.A.M. Lewis. 1986. Atmospheric input of nitrogen to a coastal fynbos ecosystem of the south-western Cape Province, South Africa. *South African J. Bot.* 52:273–276.

Stone, S.F., G.K. Isbister, S. Shahmy, F. Mohamed, C. Abeysinghe, H. Karunathilake, A. Ariaratnam, T.L. Jacoby-Alner, C.L. Cotterell, and S.G.A. Brown. 2013. Immune response to snake envenoming and treatment with antivenom; complement activation, cytokine production and mast cell degranulation. *PLoS Negl. Trop. Dis*. 7:e2326.

Straight, R., J.L. Glenn, and C.C. Snyder. 1976. Antivenom activity of rattlesnake blood plasma. *Nature* 261:259–260.

Strickland, J.L., C.F. Smith, A.J. Mason, D.R. Schield, M. Borja, G. Castañeda-Gaytán, C.L. Spencer, L.L. Smith, A. Trápaga, N.M. Bouzid, G. Campillo-García, O.A. Flores-Villela, D. Antonio-Rangel, S.P. Mackessy, T.A. Castoe, D.R. Rokyta, and C.L. Parkinson. 2018. Evidence for divergent patterns of local selection driving venom variation in Mojave Rattlesnakes (*Crotalus scutulatus*). *Sci. Rep.* 8:17622.

Strydom, A.J., and D.P. Botes. 1971. Snake venom toxins. Purification, properties, and complete amino acid sequence of two toxins from Ringhals (*Hemachatus haemachatus*) venom. *J. Biol. Chem.* 246:1341–1349.

Strydom, D.J. 1976. Snake venom toxins. Purification and properties of low-molecular-weight polypeptides of *Dendroaspis polylepis polylepis* (Black Mamba) venom. *Eur. J. Biochem.* 69:169–176.

Sunagar, K., W.E. Johnson, S.J. O'Brien, V. Vasconcelos, and A. Antunes. 2012. Evolution of CRISPs associated with toxicoferan-reptilian venom and mammalian reproduction. *Mol. Biol. Evol.* 29:1807–1822.

Sunnucks, P. 2000. Efficient genetic markers for population biology. *Trends Ecol. Evol.* 15:199–203.

Szyndlar, Z., and J.C. Rage. 1990. West Palearctic cobras of the genus *Naja* (Serpentes: Elapidae): interrelationships among extinct and extant species. *Amphibia-Reptilia* 11:385–400.

Szyndlar, Z., and H. Schleich. 1993. Description of Miocene snakes from Petersbuch 2 with comments on the lower and middle Miocene ophidian faunas of southern Germany. *Stuttgarter Beiträge zur Naturkd. Ser. B* 192:1–47.

Tadokoro, T., C.M. Modahl, K. Maenaka, and N. Aoki-Shioi. 2020. Cysteine-rich secretory proteins (CRISPs) from venomous snakes: an overview of the functional diversity in a large and underappreciated superfamily. *Toxins* 12:175.

Takeda, S., H. Takeya, and S. Iwanaga. 2012. Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochim. Biophys. Acta* 1824:164–176.

Takeya, H., A. Onikura, T. Nikai, H. Sugihara, and S. Iwanaga. 1990. Primary structure of a hemorrhagic metalloproteinase, HT-2, isolated from the venom of *Crotalus ruber ruber. J. Biochem*. 108:711–719.

Tamiya, N., A. Sato, H.S. Kim, T. Teruuchi, C. Takasaki, Y. Ishikawa, M.L. Guinea, M. McCoy, H. Heatwole, and H.G. Cogger. 1983. Neurotoxins of sea snakes genus *Laticauda. Toxicon* 3:445–447.

Tan, K.Y., C.H. Tan, L. Chanhome, and N.H. Tan. 2017. Comparative venom gland transcriptomics of *Naja kaouthia* (Monocled cobra) from Malaysia and Thailand: elucidating geographical venom variation and insights into sequence novelty. *PeerJ* 5:e3142.

Tarsi, K., and T. Tuff. 2012. Introduction to population demographics. *Nat. Educ. Knowl.* 3:3.

Tavaré, S., and R.M. Miura. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math. Life Sci*. 17:57–86.

Taylor, H.C., and J.W. Morris. 1981. A brief account of coast vegetation near Port Elizabeth. *Bothalia* 3:519–525.

Tong, Q., F. Wang, H.-Z. Zhou, H.L. Sun, H. Song, Y.Y. Shu, Y. Gong, W.T. Zhang, T.X. Cai, F.Q. Yang, J. Tang, and T. Jiang. 2012. Structural and functional insights into lipid-bound nerve growth factors. *Res. Commun.* 26:3811–3821.

Townsend, T.M., R.E. Alegre, S.T. Kelley, J.J. Wiens, and T.W. Reeder. 2008. Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles. *Mol. Phylogenet. Evol.* 47:129–142.

Trauth, M.H., J.C. Larrasoaña, and M. Mudelsee. 2009. Trends, rhythms and events in Plio-Pleistocene African climate. *Quat. Sci. Rev*. 28:399–411.

Trummal, K., K. Tonismagi, V. Paalme, L. Jarvekulg, J. Siigur, and E. Siigur. 2011. Molecular diversity of snake venom nerve growth factors. *Toxicon* 58:363–368.

Tsai, H.Y., Y.M. Wang, and I.H. Tsai. 2008. Cloning, characterization and phylogenetic analyses of members of three major venom families from a single specimen of *Walterinnesia aegyptia. Toxicon* 51:1245–1254.

Ullah, A., K. Ullah, H. Ali, C. Betzel, and S. ur Rehman. 2019. The sequence and a three-dimensional structural analysis reveal substrate specificity among snake venom phosphodiesterases. *Toxins* 11:625.

Ursenbacher, S., M. Carlsson, V. Helfer, H. Tegelström, and L. Fumagalli. 2006. Phylogeography and Pleistocene refugia of the adder (*Vipera berus*) as inferred from mitochondrial DNA sequence data. *Mol. Ecol.* 15:3425–3437.

Uzair, B., B.A. Khan, N. Sharif, F. Shabbir, and F. Menaa. 2018. Phosphodiesterases (PDEs) from snake venoms: therapeutic applications. *Protein Pept. Lett.* 25:612–618.

Van den Boogaart, K.G., R. Tolosana-Delgado, and M. Bren. 2019. Compositional data analysis. Available at http://www.stat.boogaart.de/compositions%0D.

Van der Merwe, H., M.W. van Rooyen, and N. van Rooyen. 2008. Vegetation of the Hantam-Tanqua-Roggeveld subregion, South Africa. Part 2: Succulent Karoo biome and related vegetation. *Koedoe* 50:160–183.

Verboom, G.A., N.G. Bergh, S.A. Haiden, V. Hoffmann, and M.N. Britton. 2015. Topography as a driver of diversification in the Cape Floristic Region of South Africa. *New Phytol.* 207:368–376.

Viala, V.L., D. Hildebrand, T.M. Fucase, J.M. Sciani, J.P. Prezotto-Neto, M. Riedner, M. Sanches, P.J. Nishimura, N. Oguiura, D.C. Pimenta, H. Schlüter, C. Betzel, R.K. Arni, and P.J. Spencer. 2015. Proteomic analysis of the rare Uracoan rattlesnake *Crotalus vegrandis* venom: evidence of a broad arsenal of toxins. *Toxicon* 107:234–251.

Vieira, L.F., A.J. Magro, C.A.H. Fernandes, B.M. De Souza, W.L.G. Cavalcante, M.S. Palma, J.C. Rosa, A.L. Fuly, M.R.M. Fontes, M. Gallacci, D.S. Butzke, L.A. Calderon, R.G. Stábeli, J.R. Giglio, and A.M. Soares. 2013. Biochemical, functional, structural and phylogenetic studies on Intercro, a new isoform phospholipase A2from *Crotalus durissus terrificus* snake venom. *Biochimie* 95:2365–2375.

Vogel, C.W., and D.C. Fritzinger. 2017. Cobra venom factor: the unique component of cobra venom that activates the complement system. Pp. 345–404 in *Snake Venoms* (P. Gopalakrishnakone, H. Inagaki, C.W. Vogel, A.K. Mukherjee, and T.K. Rahmy, eds.). Springer Dordrecht, Dordrecht.

Vonk, F.J., N.R. Casewell, C. V. Henkel, A.M. Heimberg, H.J. Jansen, R.J. McCleary, H.M. Kerkkamp, R.A. Vos, I. Guerreiro, J.J. Calvete, W. Wüster, A.E. Woods, J.M. Logan, R.A. Harrison, T.A. Castoe, A.P. de Koning, D.D. Pollock, M. Yandell, D. Calderon, C. Renjifo, R.B. Currier, D. Salgado, D. Pla, L. Sanz, A.S. Hyder, J.M. Ribeiro, J.W. Arntzen, G.E. van den Thillart, M. Boetzer, W. Pirovano, R.P. Dirks, H.P. Spaink, D. Duboule, E. McGlinn, R.M. Kini, and M.K. Richardson. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. U.S.A.* 110:20651–20656.

WHO. 2016. *Guidelines for the production control and regulation of snake antivenom immunoglobulins*. Expert Comm. Biol. Stand. Post ECBS version. Available at https://www.who.int/bloodproducts/snake_antivenoms/snakeantivenomguide/en/.

WHO. 2017. *Report of the Tenth Meeting of the WHO Strategic and Technical Advisory Group for Neglected Tropical Diseases*. Geneva. 7 pp. Available at https://www.who.int/neglected_diseases/NTD_STAG_report_2017.pdf.

WHO. 2019. *Snakebite envenoming: a strategy for prevention and control*. Geneva. 50 pp. Available at https://www.who.int/snakebites/resources/9789241515641/en/.

Wickham, H., W. Chang, L. Henry, T.L. Pedersen, K. Takahashi, C. Wilke, K. Woo, and H. Yutani. 2019. Create elegant data visualisations using the grammar of graphics version 3.2.1. Available at https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf.

Wiens, J.J., C.R. Hutter, D.G. Mulcahy, B.P. Noonan, T.M. Townsend, J.W. Sites, and T.W. Reeder. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biol. Lett.* 8:1043–1046.

Wijeyewickrema, L.C., E.E. Gardiner, E.L. Gladigau, M.C. Berndt, and R.K. Andrews. 2010. Nerve growth factor inhibits metalloproteinase-disintegrins and blocks ectodomain shedding of platelet glycoprotein VI*. *J. Biol. Chem.* 285:11793–11799.

Wilkinson, J.A., J.L. Glenn, R. Straight, and J.W. Sites. 1991. Distribution and genetic variation in venom A and B populations of the Mojave rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Herpetologica* 47:54–68.

Williams, V., J. White, T.D. Schwaner, and A. Sparrow. 1988. Variation in venom proteins from isolated populations of tiger snakes (*Notechis ater niger*, *N. scutatus*) in South Australia. *Toxicon* 26:1067–1075.

Willson, J.D., C.T. Winne, M.A. Pilgrim, C.S. Romanek, and J. Whitfield Gibbons. 2010. Seasonal variation in terrestrial resource subsidies influences trophic niche width and overlap in two aquatic snake species: a stable isotope approach. *Oikos* 119:1161–1171.

Wüster, W., S. Crookes, I. Ineich, Y. Mane, C.E. Pook, and D.G. Broadley. 2007. The phylogeny of cobras inferred from mitochondrial DNA sequences: Evolution of venom spitting and the phylogeography of the African spitting cobras (Serpentes: Elapidae: *Naja nigricollis* complex). *Mol. Phylogenet. Evol.* 45:437–453.

Wüster, W., L. Peppin, C.E. Pook, and D.E. Walker. 2008. A nesting of vipers: phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes). *Mol. Phylogenet. Evol.* 49:445–459.

Yamazaki, Y., H. Koike, Y. Sugiyama, K. Motoyoshi, T. Wada, S. Hishinuma, M. Mita, and T. Morita. 2002. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. *Eur. J. Biochem*. 269:2708–2715.

Yamazaki, Y., Y. Matsunaga, Y. Tokunaga, S. Obayashi, M. Saito, and T. Morita. 2009. Snake venom vascular endothelial growth factors (VEGF-Fs) exclusively vary their structures and functions among species. *J. Biol. Chem*. 284:9885–9891.

Yang, W., J. Feng, B. Wang, Z. Cao, W. Li, Y. Wu, and Z. Chen. 2014. BF9, the first functionally characterized snake toxin peptide with Kunitz-type protease and potassium channel inhibiting properties. *J. Biochem. Mol. Toxicol.* 28:76–83.

Zaqueo, K.D., A.M. Kayano, R. Simões-Silva, L.S. Moreira-Dill, C.F.C. Fernandes, A.L. Fuly, V.G. Maltarollo, K.M. Honório, S.L. da Silva, G. Acosta, M.A.O. Caballol, E. de Oliveira, F. Albericio, L.A. Calderon, A.M. Soares, and R.G. Stábeli. 2014. Isolation and biochemical characterization of a new thrombin-like serine protease from *Bothrops pirajai* snake venom. *Anim. Toxins their Advantages Biotechnol. Pharmacol.* 595186:1–13.

Zhan, S., and A.P. Ingersoll. 2016. On the abruptness of Bølling-Allerød warming. *J. Clim.* 29:4965–4975.

Zhao, X., L. Dupont, E. Schefuß, M.E. Meadows, A. Hahn, and G. Wefer. 2016. Holocene vegetation and climate variability in the winter and summer rainfall zones of South Africa. *The Holocene* 26:843–857.

Supplementary Material

Script S1: R script specifying commands to visualise dated phylogenetic tree files produced in BEAST2.

```
# Load required libraries

library(devtools)

library(geoscale)

library(phyloch)

library(strap)

library(OutbreakTools)


# Set working directory

setwd("directory path")


# Read in the annotated nexus file produced in TreeAnnotator as part of the BEAST2 package

t <- read.annotated.nexus("NucSub_Test.nex")


# Specify estimated root time calculated in BEAST2

t$root.time <- 42.5


# Run geoscalePhylo command to produce dated phylogeny

geoscalePhylo(t, direction="rightwards", units=c("Period", "Epoch",
"Age"), boxes="none", tick.scale="myr", cex.age=0.7, cex.ts=0.7,
cex.tip=0.55, width=2, label.offset=0.2, ts.col=TRUE,
vers="ICS2013", x.lim=c(-10,50), quat.rm = FALSE)
```

Script S2: R script specifying commands for distribution modelling based on past climatic conditions using Maxent.

```
## OCCURRENCE DATA ##

# Set working directory to folder containing the R script

setwd("directory path")


# Import occurrence data csv file

RinkData<- read.csv(file.choose(), header=TRUE, sep=";")


# Double check that all required fields are present

str(RinkData)

names(RinkData)

head(RinkData)

tail(RinkData)


# Load required R packages

library(maptools)

library(rgdal)

library(raster)

library(spThin)

library(rgeos)


# Import Africa and SA shapefile with provinces indicated (edited in
GIS)

Africa <- readOGR(dsn = "directory path containing file", layer =
"Africa")

SA_Provs <- readOGR(dsn = "directory path containing file", layer =
"SOU-level_1")


# Merge shapefiles

AfricaShape <- bind(Africa, SA_Provs)
```

```
# Plot the Africa shapefile background

plot(AfricaShape)


# Thin out points in Gauteng to avoid bias in distribution data

# Create a RasterLayer with the extent of South Africa

SA <- crop(AfricaShape, extent(10, 40, -35, -18))

plot(SA, col="khaki", bg="azure2")


# Thin data using spThin package

RinkData_Thin <-

  thin(loc.data = RinkData, spec.col = "species", long.col = "x",
lat.col = "y",

        thin.par = 10, reps = 10,

        locs.thinned.list.return = FALSE,

        write.files = TRUE,

        max.files = 1,

        out.dir = "Thinned_RinkData", out.base = "Rink_thinned",

        write.log.file = TRUE,

        log.file = "Rink_Thinned_log_file.txt", verbose = TRUE)


# Read in Thinned Dataset

RinkData_Thinner<- read.csv(file.choose(), header=TRUE, sep=",")

str(RinkData_Thinner)


# Specify bounding box for Gauteng province - for further thinning
of points

box = c(26.03329,  29.69998, -27.09327, -25.49911)

boxRecordIndexes = which(RinkData_Thinner$x > box[1] &
RinkData_Thinner$x < box[2] & RinkData_Thinner$y > box[3] &
RinkData_Thinner$y < box[4])

boxRecords = RinkData_Thinner[boxRecordIndexes,]
```

```
# Plot the points to view

plot(SA_Provs, xlim = box[1:2], ylim = box[3:4])

points(boxRecords$x, boxRecords$y)


# Thin out Gauteng records further using a bounding box

RinkDataThinnest <-

thin(loc.data = boxRecords, spec.col = "species", long.col = "x",
lat.col = "y",

    thin.par = 20, reps = 10,

    locs.thinned.list.return = FALSE,

    write.files = TRUE,

    max.files = 1,

    out.dir = "Thinned_RinkData", out.base = "Rink_thinnest",

    write.log.file = TRUE,

    log.file = "Rink_Thinnest_log_file.txt", verbose = TRUE)


thinnedBoxRecords = read.csv(file.choose(), header=TRUE, sep=",")

points(RinkDataThinnest$x, RinkDataThinnest$y, pch = 16)


RinkDataThinnest = rbind(RinkData_Thinner[-boxRecordIndexes,],
thinnedBoxRecords)


# Save csv file for GIS (final file with thinned records to be used
for modeling)

write.csv(RinkDataThinnest, "rinkdatathinnest.csv")


## ENVIRONMENTAL PREDICTORS ##

# Set working directory to folder with environmental layers

fileName = "directory path"

setwd(fileName)

# Get names of tif files
```

```
PredictList <- list.files(fileName, pattern='tif$', full.names=TRUE
)

PredictList


# Create raster stack of variables

PredictorStack <- stack(PredictList)

names(PredictorStack)


# Extract and bind the environmental data for the species'
occurrence records

DataExtract <- extract(PredictorStack, RinkDataThinnest[,2:3])

# Test for correlation between environmental predictors

Env.pca <- prcomp(DataExtract, center = TRUE, scale = TRUE)

summary(Env.pca)

biplot(Env.pca, scale = 0) # Plot data


# Only include 1 of correlated predictors

# Selected these visually off the biplot

uncorrelated =
c("BIO14","BIO16","BIO1","BIO10","BIO5","BIO7","BIO15","BIO18","BIO1
2","BIO3")


# Make a new raster stack with these variables

UsePredictorList =  vector("character",0)


for (i in 1:length(PredictList)){

  for (j in 1:length((uncorrelated))){

    if(length(grep(paste(uncorrelated[j], "\\.", sep = ""),
PredictList[i])) == 1){

      UsePredictorList = c(UsePredictorList, PredictList[i])

    }

  }

}
```

```
UsePredictorStack <- stack(UsePredictorList)


modelExtent = c(10.74670, 40.58141,-35.01683, -15.93254)
ModelPredictorStack = crop(UsePredictorStack, modelExtent)
plot(subset(ModelPredictorStack, 1)) # look at one just to see if
it's okay


# Load libraries for maxent
library(dismo)
library(rJava)
maxent()


# Run maxent
model = maxent(UsePredictorStack, RinkDataThinnest[,2:3])
show(model) #show the model diagnostics


# Predict the distribution (current distribution based on model)
predictExtent = c(10.74670, 40.58141, -35.01683, -15.478768)
PredictPredictorStack = crop(UsePredictorStack, predictExtent)
plot(subset(PredictPredictorStack, 1)) # Look at one to test
prediction = predict(model, PredictPredictorStack)
plot(prediction) # the whole map
plot(SA_Provs, xlim = c(10.74670, 40.58141), ylim = c(10.74670,
40.58141), add=T) #Add provincial borders
points(RinkDataThinnest$x, RinkDataThinnest$y, pch = 16) #View
sample points
writeRaster(prediction, "rinkAfricaPredicted.grd")


# Predict for past climate
# First prepare the raster layers - too big to import all
directory = "directory path containing files"
```

```
PredictList <- list.files(directory, pattern='bil$', full.names=TRUE
)

list(PredictList) #List tif files


# Create raster stack of variables

PredictorStack <- stack(PredictList)

names(PredictorStack)


#Adjust resolution to 2.5min if necessary (only applies to cases
with resolution > 2.5min)

UsePredictorStack <- resample(PredictorStack, UsePredictorStack)


# Specify uncorrelated variables

uncorrelated =
c("BIO14","BIO16","BIO1","BIO10","BIO5","BIO7","BIO15","BIO18","BIO1
2","BIO3")


# Make a new raster stack with these variables

LICUsePredictorList =  vector("character",0)


for (i in 1:length(LICPredictList)){

  for (j in 1:length((uncorrelated))){

    if(length(grep(paste(uncorrelated[j], "\\.", sep = ""),
LICPredictList[i])) == 1){

      LICUsePredictorList = c(LICUsePredictorList,
LICPredictList[i])

    }

  }

}


# Set model extent

modelExtent = c(16.74670, 32.58141,-35.01683, -23.93254)

LICModelPredictorStack = crop(LICUsePredictorStack, modelExtent)
```

```
plot(subset(LICModelPredictorStack, 1))
```

 # Predict the distribution

```
predictExtent = c(9.614803, 42.134331, -35.423823, -11.478768)

LICPredictPredictorStack = crop(LICUsePredictorStack, predictExtent)

plot(subset(LICPredictPredictorStack, 1))


LICprediction = predict(model, LICPredictPredictorStack)

plot(LICprediction) # the whole map

plot(LICprediction, xlim = c(10.755915, 40.224664), ylim = c(-
34.848822, -22.206562)) # Crop to SA

plot(SA_Provs, xlim = c(10.755915, 40.224664), ylim = c(-34.848822,
-22.206562), add=T) # Add provincial borders
```

 # Write raster layer for viewing in GIS

```
writeRaster(prediction, "rinkAfricaPredicted.grd")
```

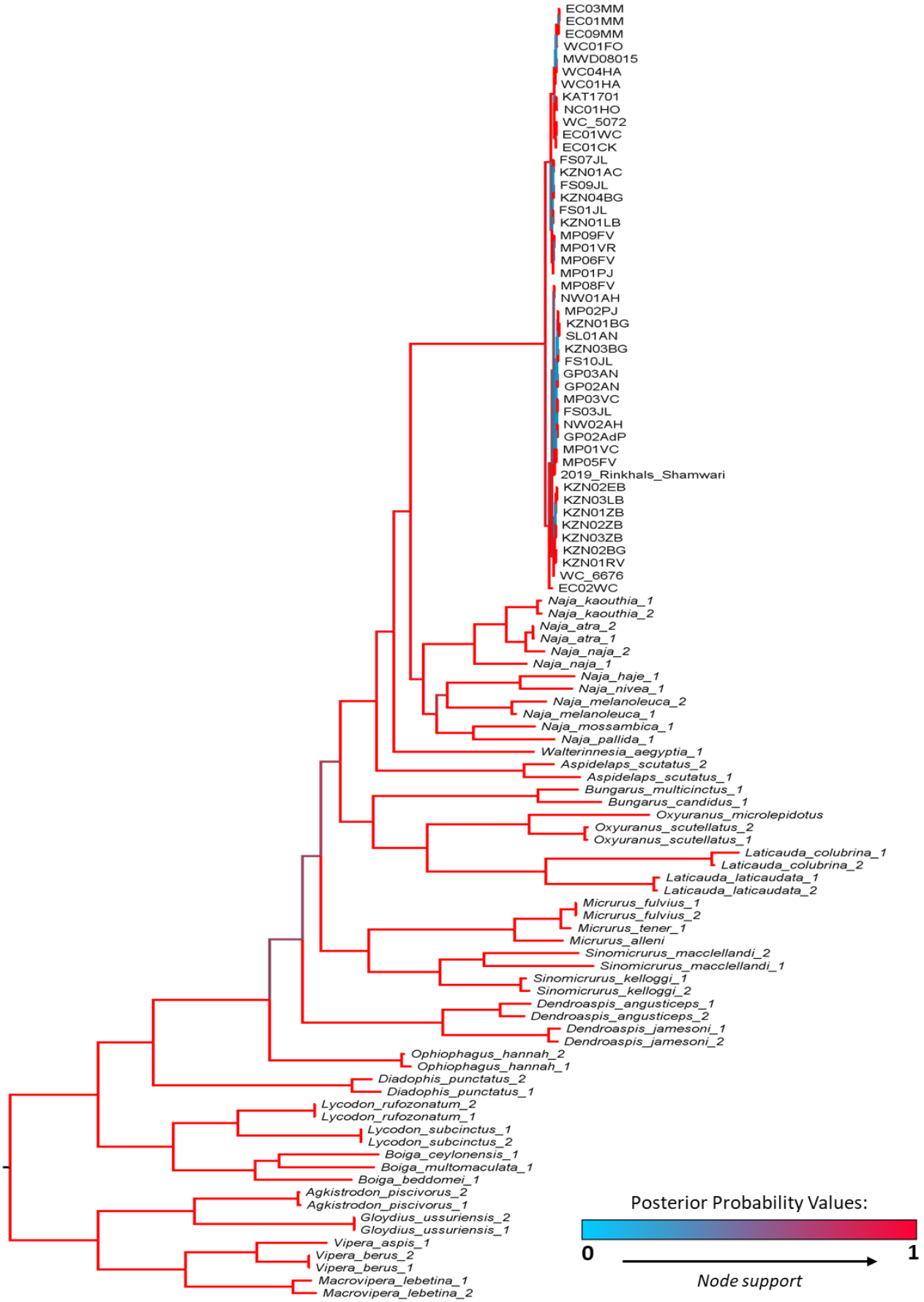

Figure S1: Bayesian phylogram for *Hemachatus haemachatus* with outgroups for the pruned mitochondrial dataset.

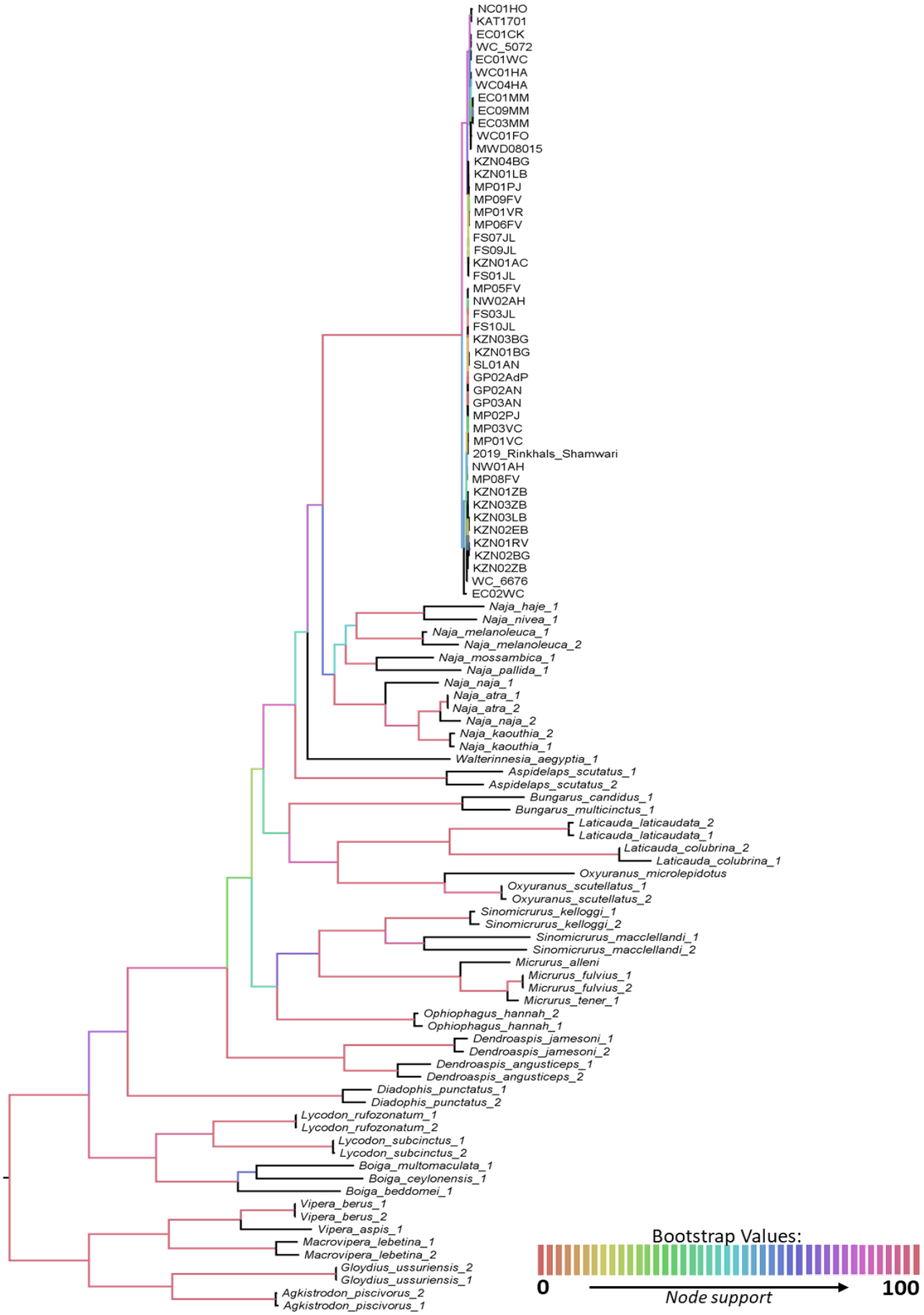

Figure S2: Maximum likelihood phylogenetic tree for *Hemachatus haemachatus* with outgroups for the pruned mitochondrial dataset.

Figure S3: Bayesian phylogram for *Hemachatus haemachatus* with outgroups for the combined nuclear and mitochondrial dataset.

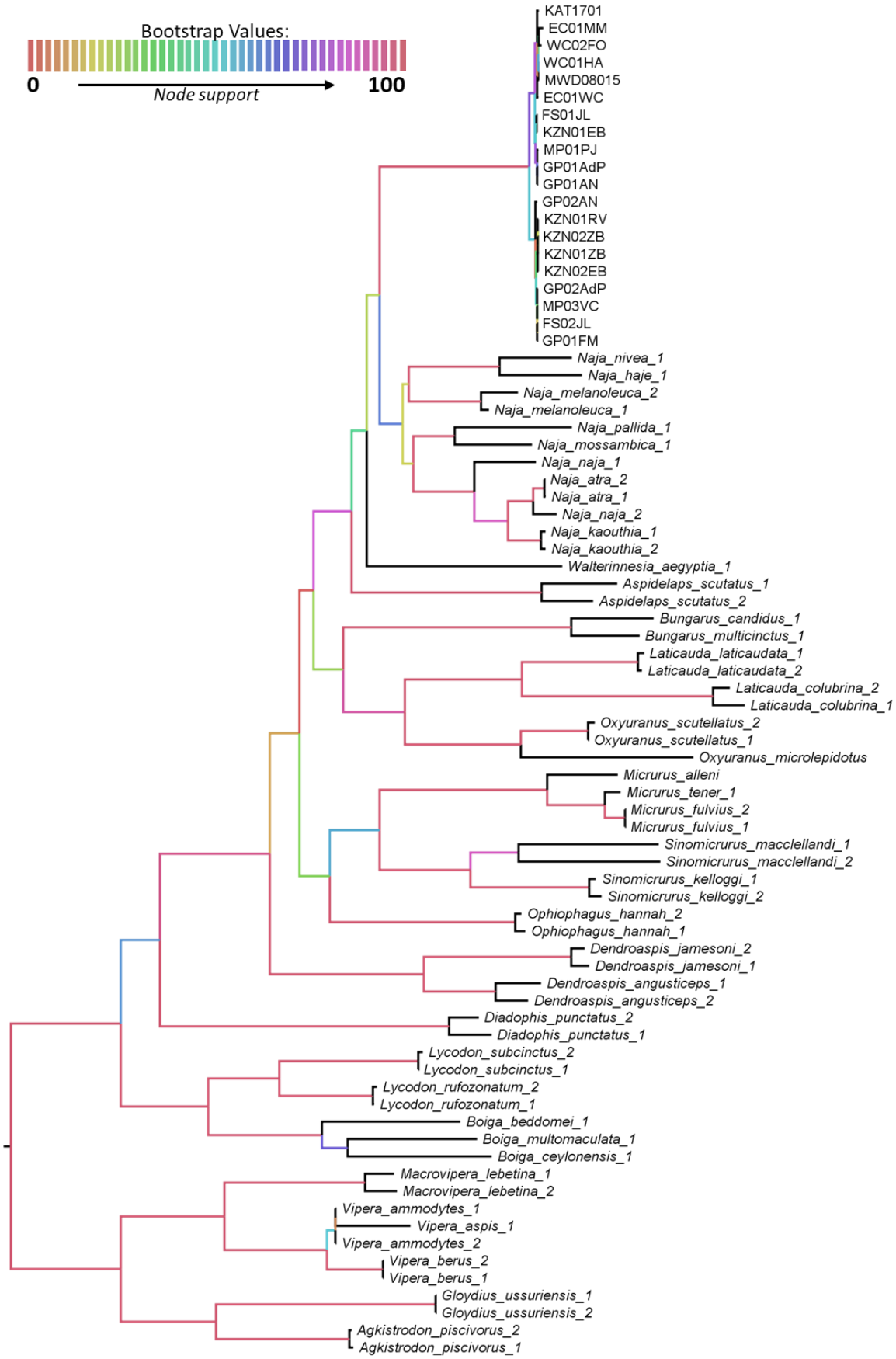

Figure S4: Maximum likelihood phylogenetic tree for *Hemachatus haemachatus* with outgroups for the combined nuclear and mitochondrial dataset.

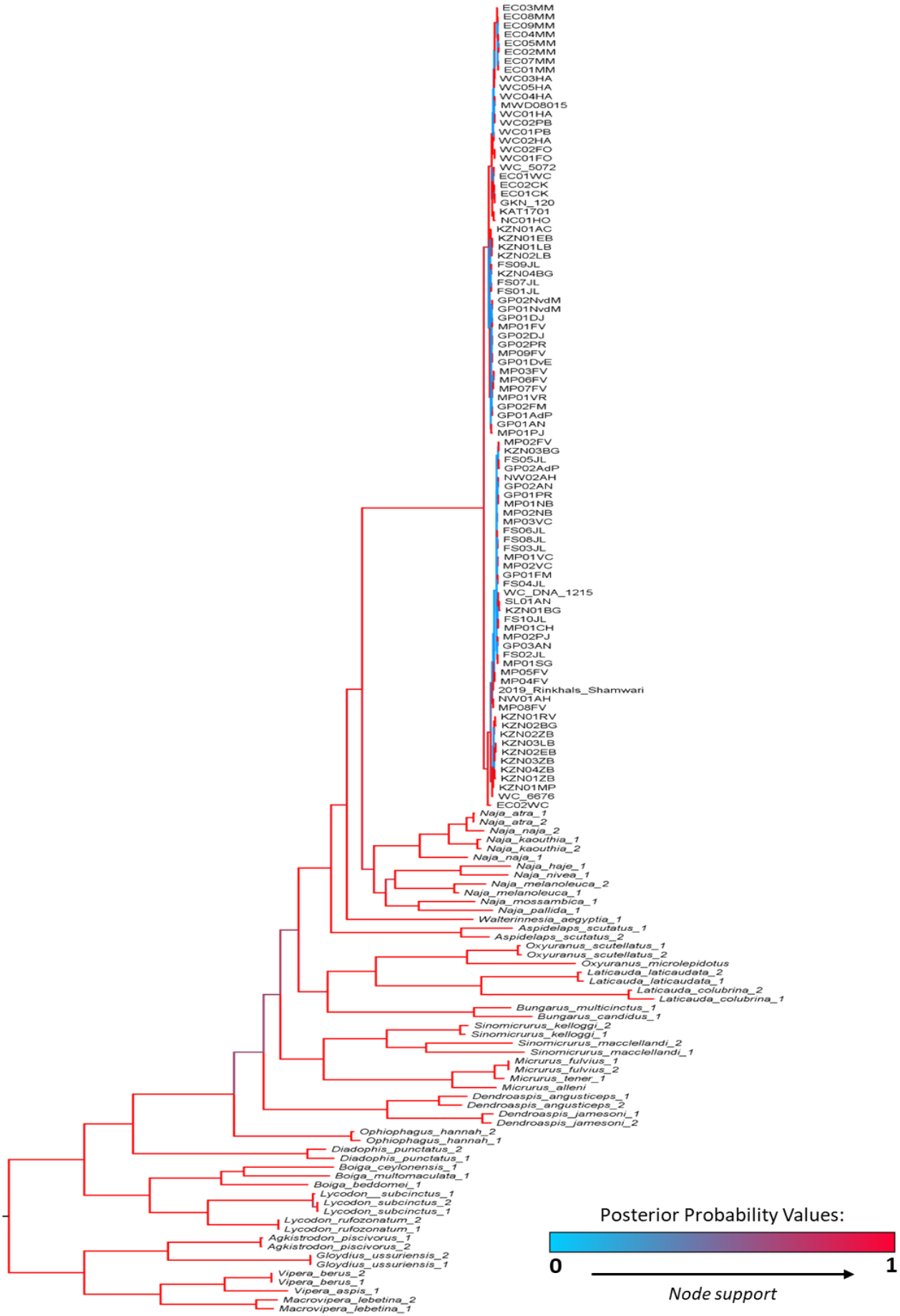

Figure S5: Bayesian phylogram for *Hemachatus haemachatus* with outgroups for the full mitochondrial dataset.

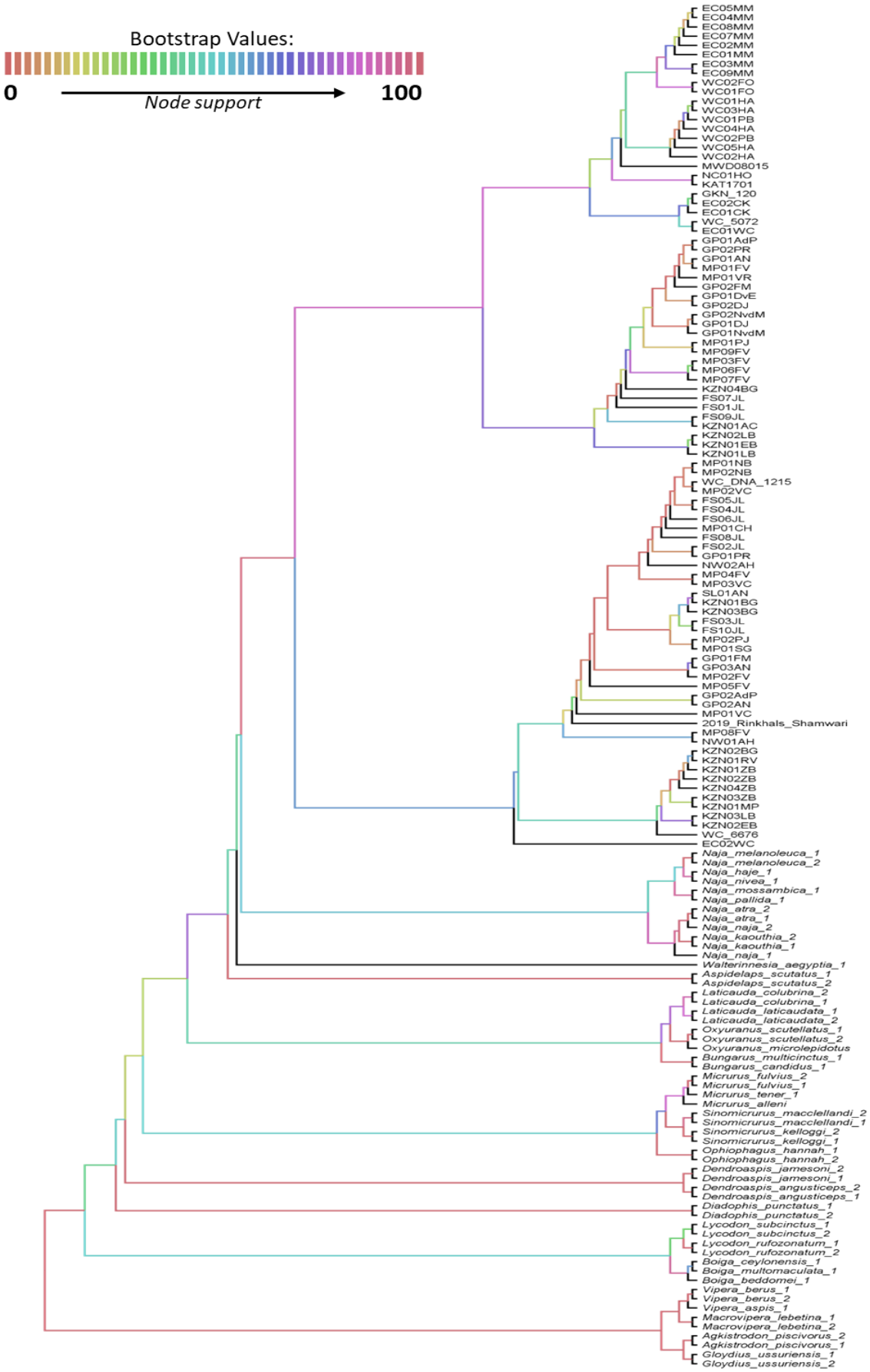

Figure S6: Maximum likelihood phylogenetic tree for *Hemachatus haemachatus* with outgroups for the full mitochondrial dataset.

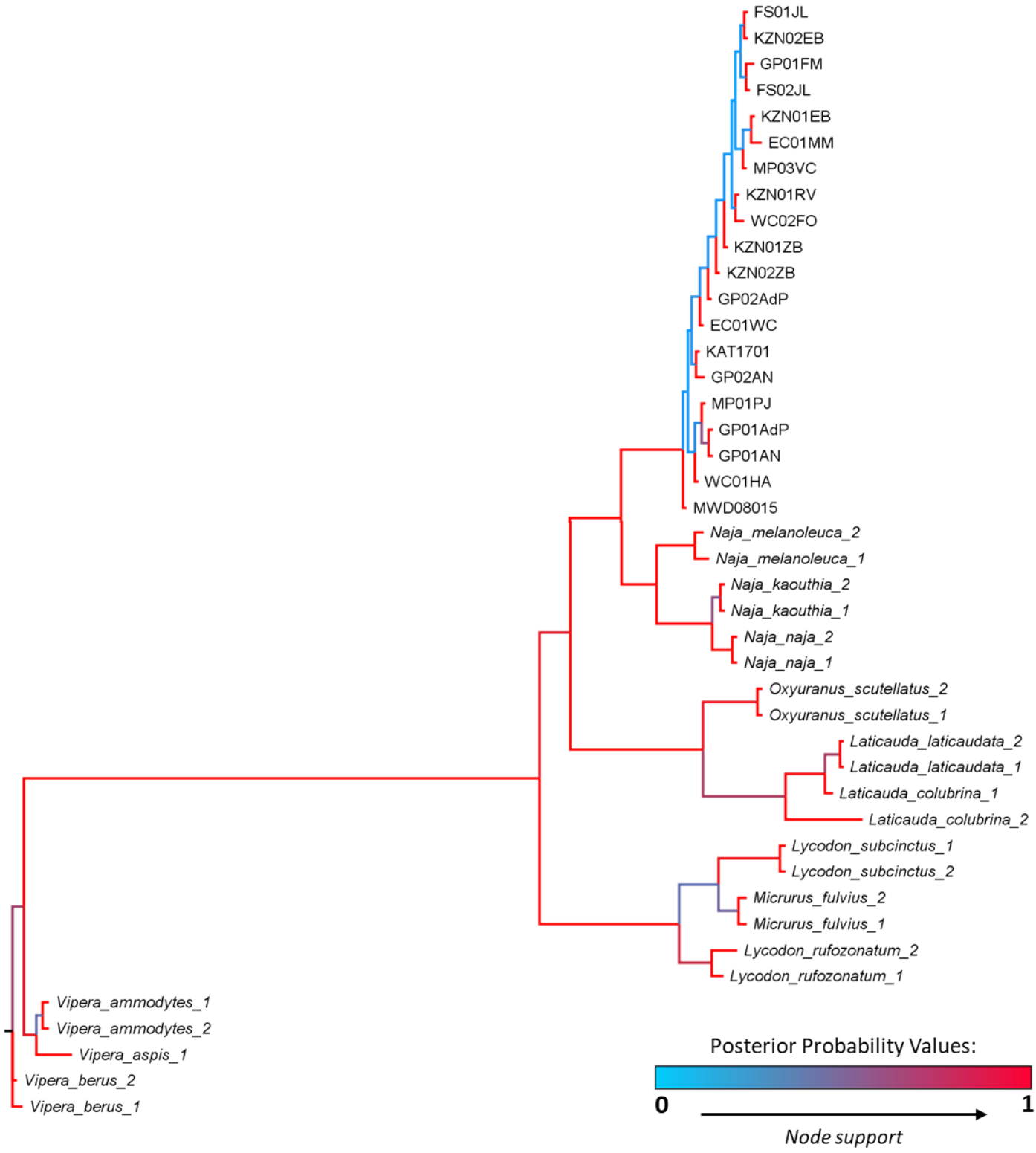

Figure S7: Bayesian phylogram for *Hemachatus haemachatus* with outgroups for the nuclear dataset.

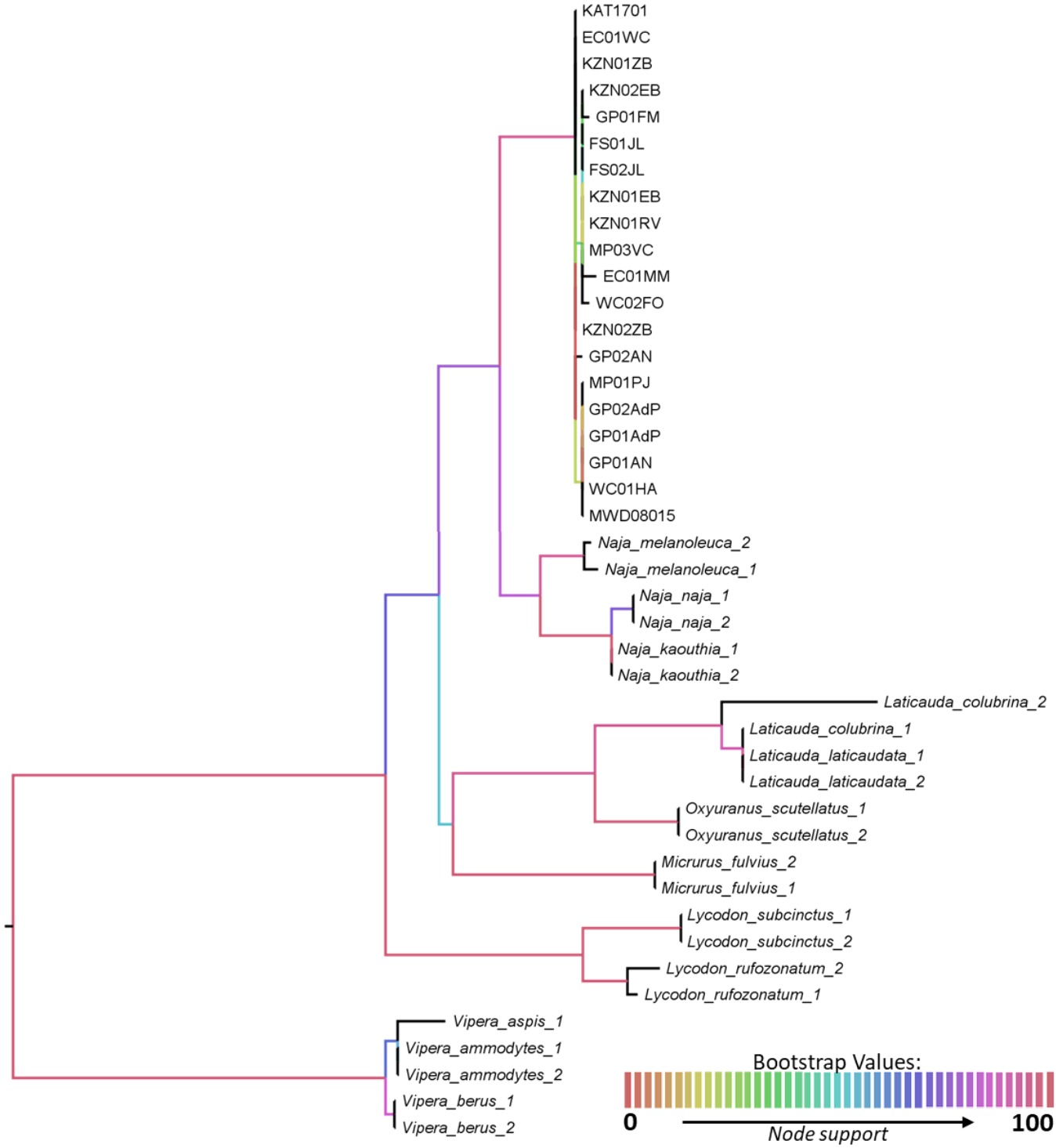

Figure S8: Maximum likelihood phylogenetic tree for *Hemachatus haemachatus* with outgroups for the nuclear dataset

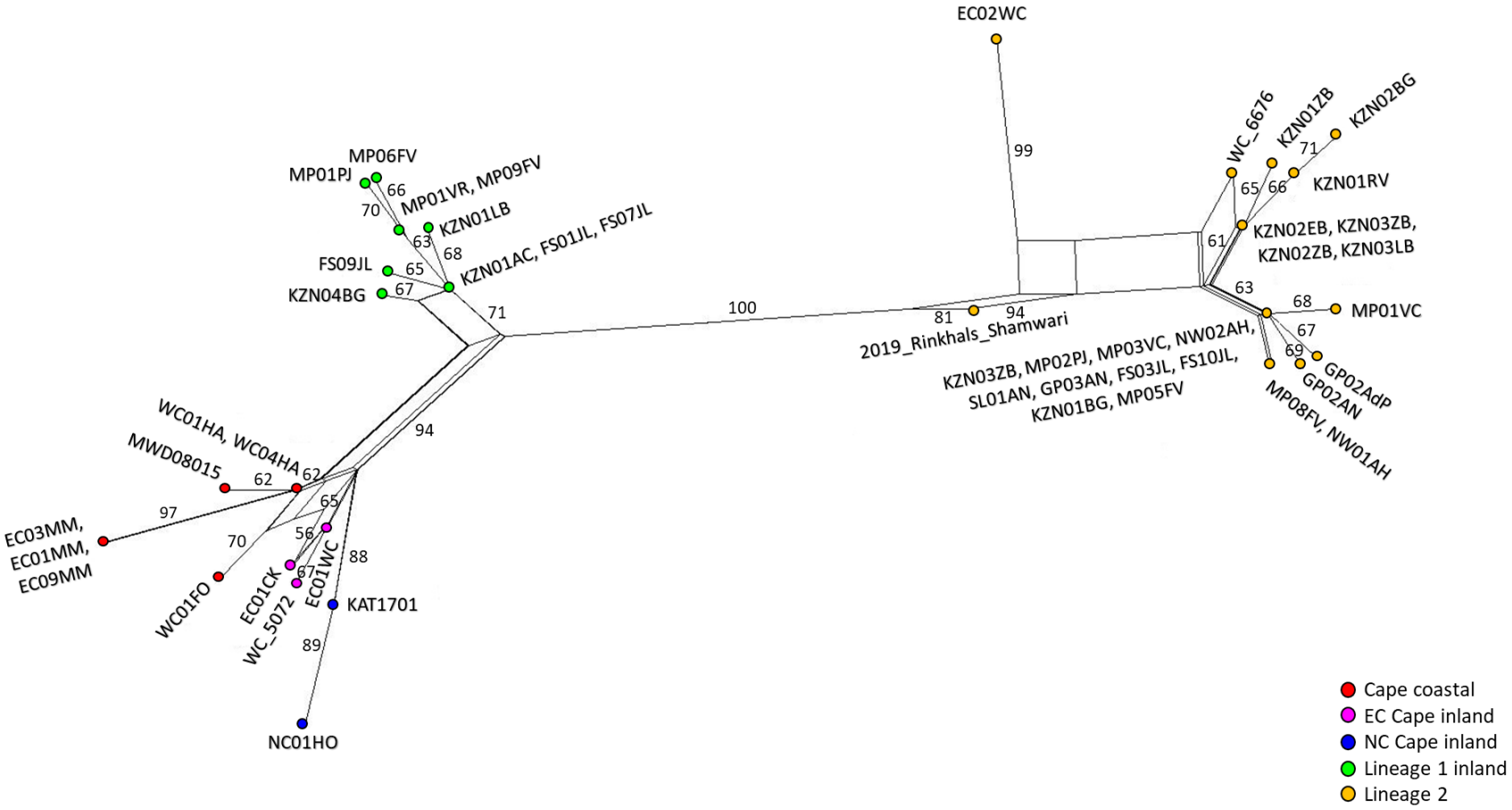

Figure S9: Phylogenetic network of the Cyt-b gene region for the pruned *Hemachatus haemachatus* dataset overlaid with Bootstrap values showing intermediate to high support (50 < BS ≤ 100).

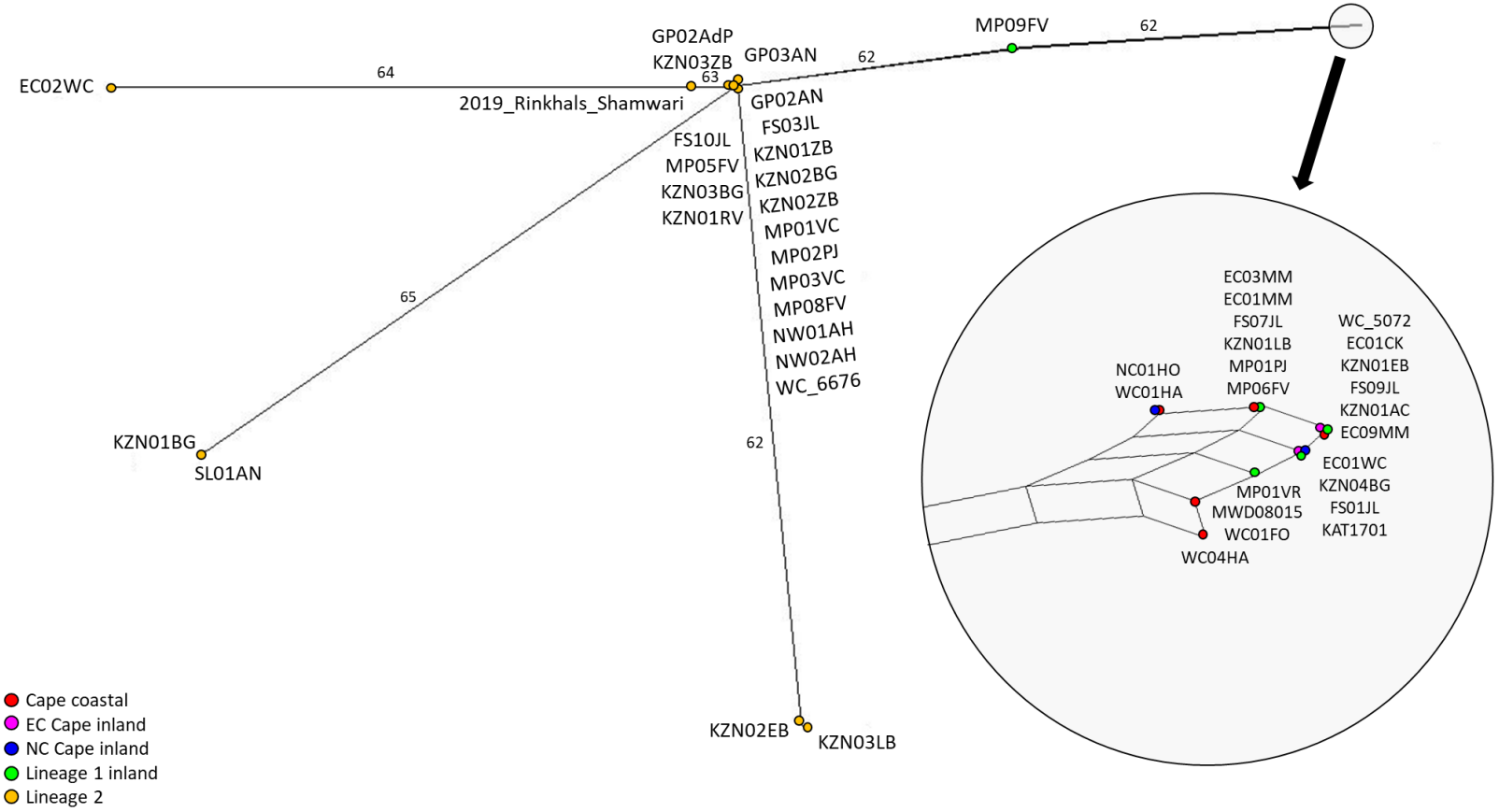

Figure S10: Phylogenetic network of the 16S gene region for the pruned *Hemachatus haemachatus* dataset overlaid with Bootstrap values showing intermediate to high support (50 < BS ≤ 100).

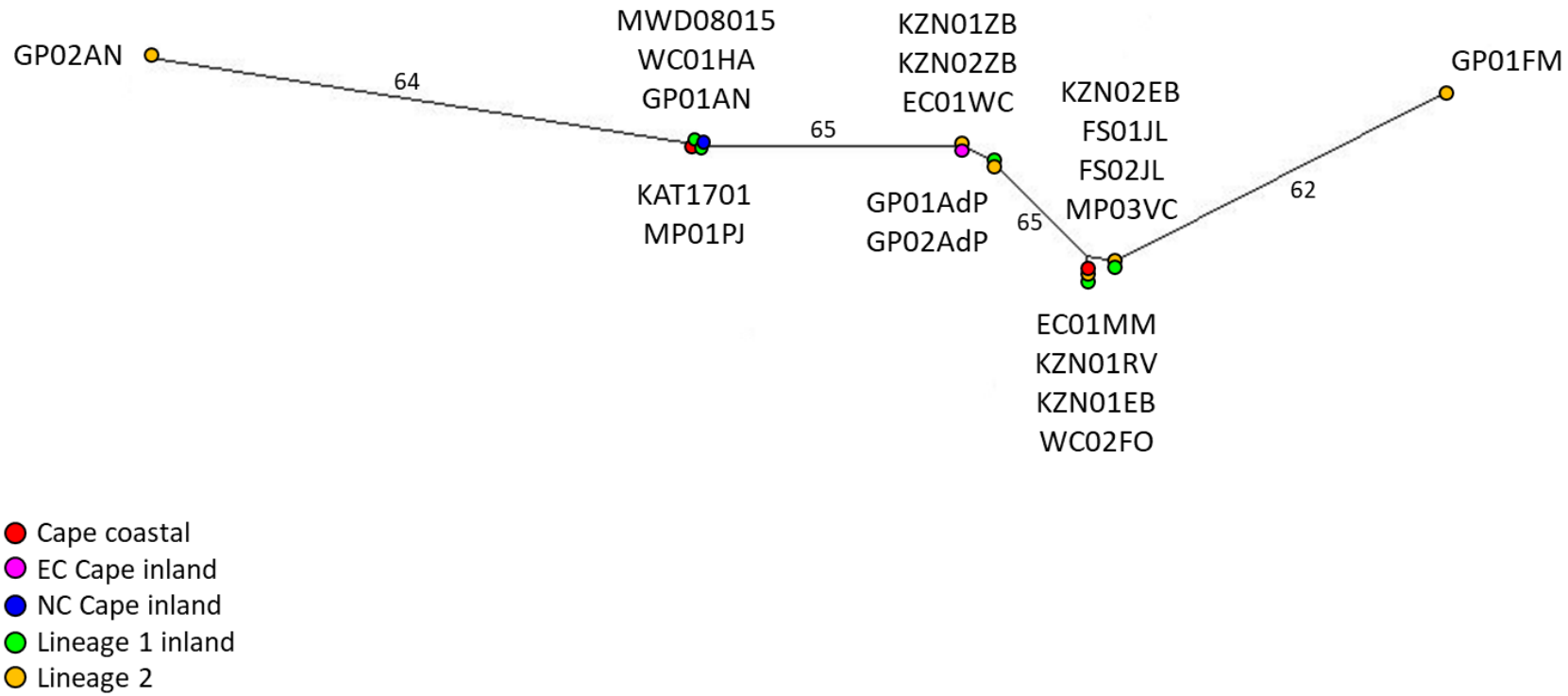

Figure S11: Phylogenetic network of the RAG-1 gene region for the pruned *Hemachatus haemachatus* dataset overlaid with Bootstrap values showing intermediate to high support (50 < BS ≤ 100).

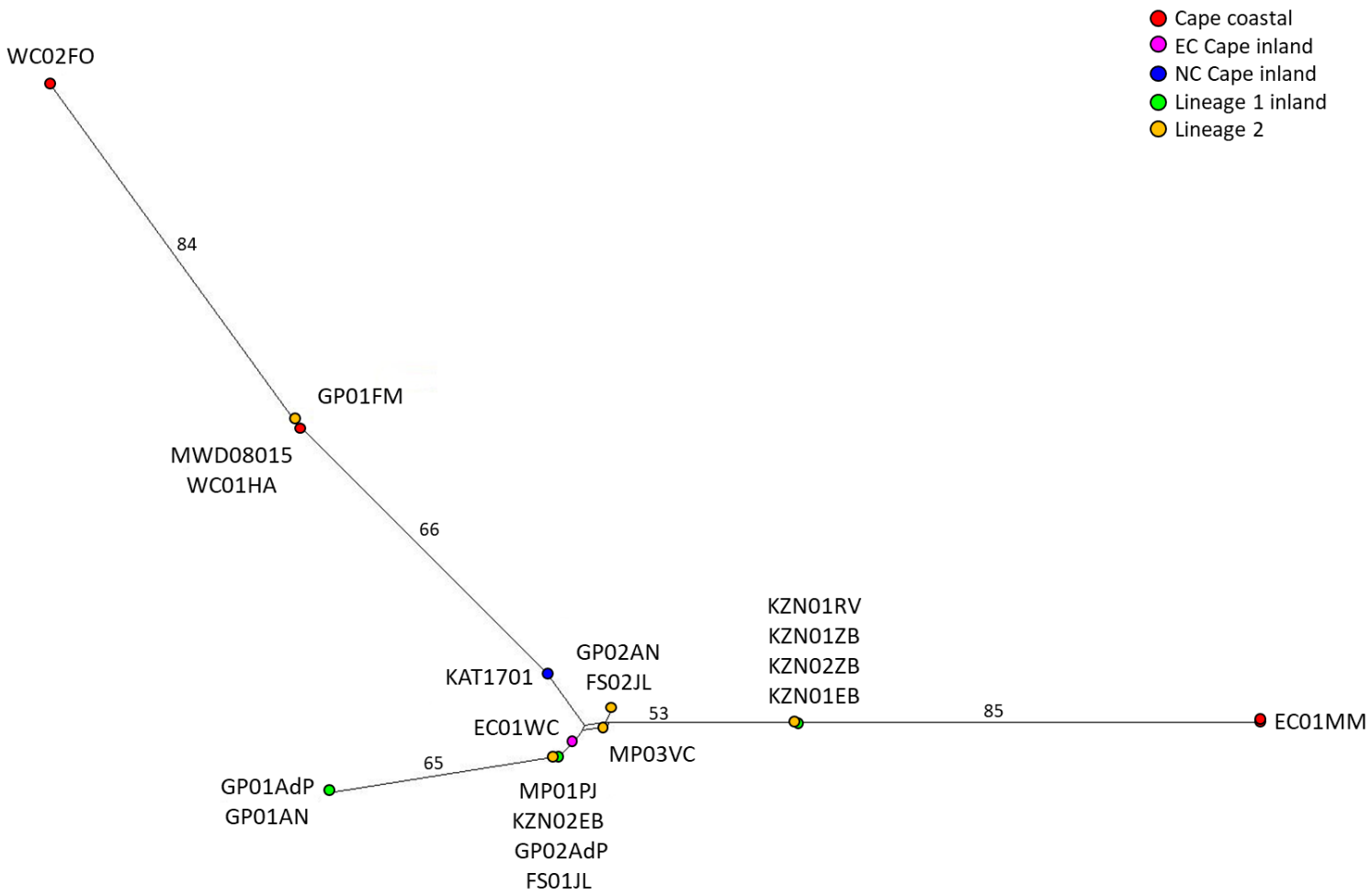

Figure S12: Phylogenetic network of the PRLR gene region for the pruned *Hemachatus haemachatus* dataset overlaid with Bootstrap values showing intermediate to high support (50 < BS ≤ 100).

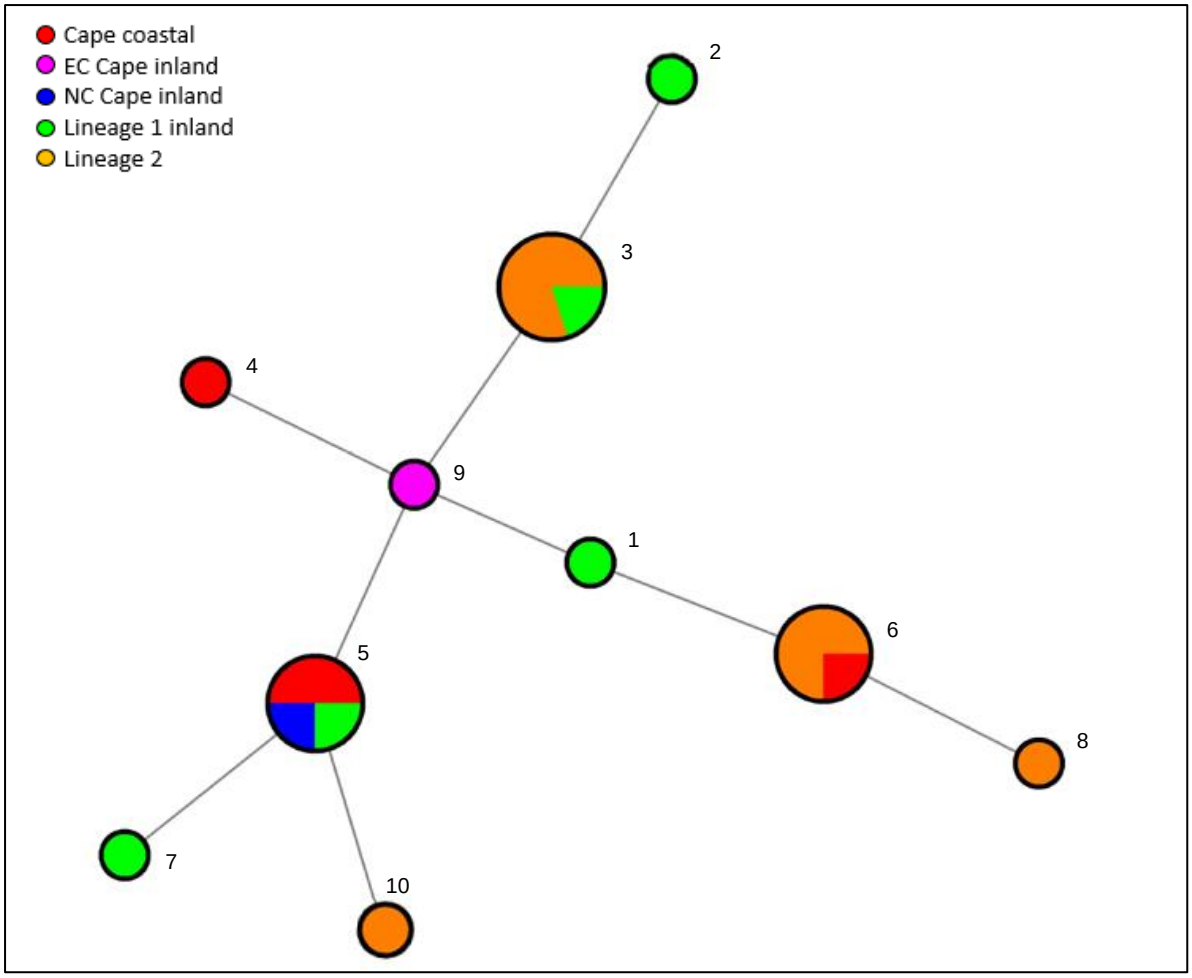

Figure S13: Haplotype network of *Hemachatus haemachatus* generated from RAG-1 and PRLR gene regions for 20 samples. The coloured circles indicate the haplotypes and the haplotype number is shown next to each circle. The area of the coloured circles is proportional to the number of samples present in each haplotype. Lines connecting the haplotypes are not proportional to the number of mutations separating each haplotype. Samples included in each haplotype are as follows: (1) KZN01EB; (2) GP01AdP; (3) GP01FM, FS02JL, GP02AdP, KZN02EB, FS01JL; (4) WC02FO; (5) MP01PJ, KAT1701, WC01HA, MWD08015; (6) KZN01ZB, EC01MM, KZN02ZB, KZN01RV; (7) GP01AN; (8) MP03VC; (9) EC01WC; (10) GP02AN.

Table S1: Divergence date estimates of *Hemachatus haemachatus* lineages for the full mitochondrial, pruned mitochondrial, and combined nuclear and mitochondrial datasets. The divergence dates were estimated from the combined results of four runs of 40 million generations per dataset. L1/L2 refers to the split between lineages 1 and 2. L1 Cape Coastal/Cape Inland refers to the split between the coastal lineages and inland lineages found in the Eastern Cape, Northern Cape and Western Cape provinces belonging to lineage 1. L1 Cape/Rest refers to the split between the Cape lineage (includes coastal and inland samples) and the inland lineage described in lineage 1. L2 KZN/Rest refers to the split between the KwaZulu-Natal lineage and the remaining samples from Mpumalanga, Gauteng, North West, the Free State and other KwaZulu-Natal samples in lineage 2.

Lineages	**Unpruned**		**Pruned**		**Nuclear Subset**	
	Age (mya)	**95% HPD (mya)**	**Age (mya)**	**95% HPD (mya)**	**Age (mya)**	**95% HPD (mya)**
L1/L2	21.3	14.17, 29.13	2.08	0.79, 5.03	1.67	0.86, 2.22
L1 Cape Coastal/Cape Inland	12.97	6.20, 20.44	0.71	0.20, 1.65	0.52	0.22, 0.73
L1 Cape/Rest	16.98	NA	1.15	0.34, 2.39	0.86	0.41, 1.20
L2 KZN/Rest	14.88	NA	0.84	0.18, 1.45	0.56	0.21, 0.88

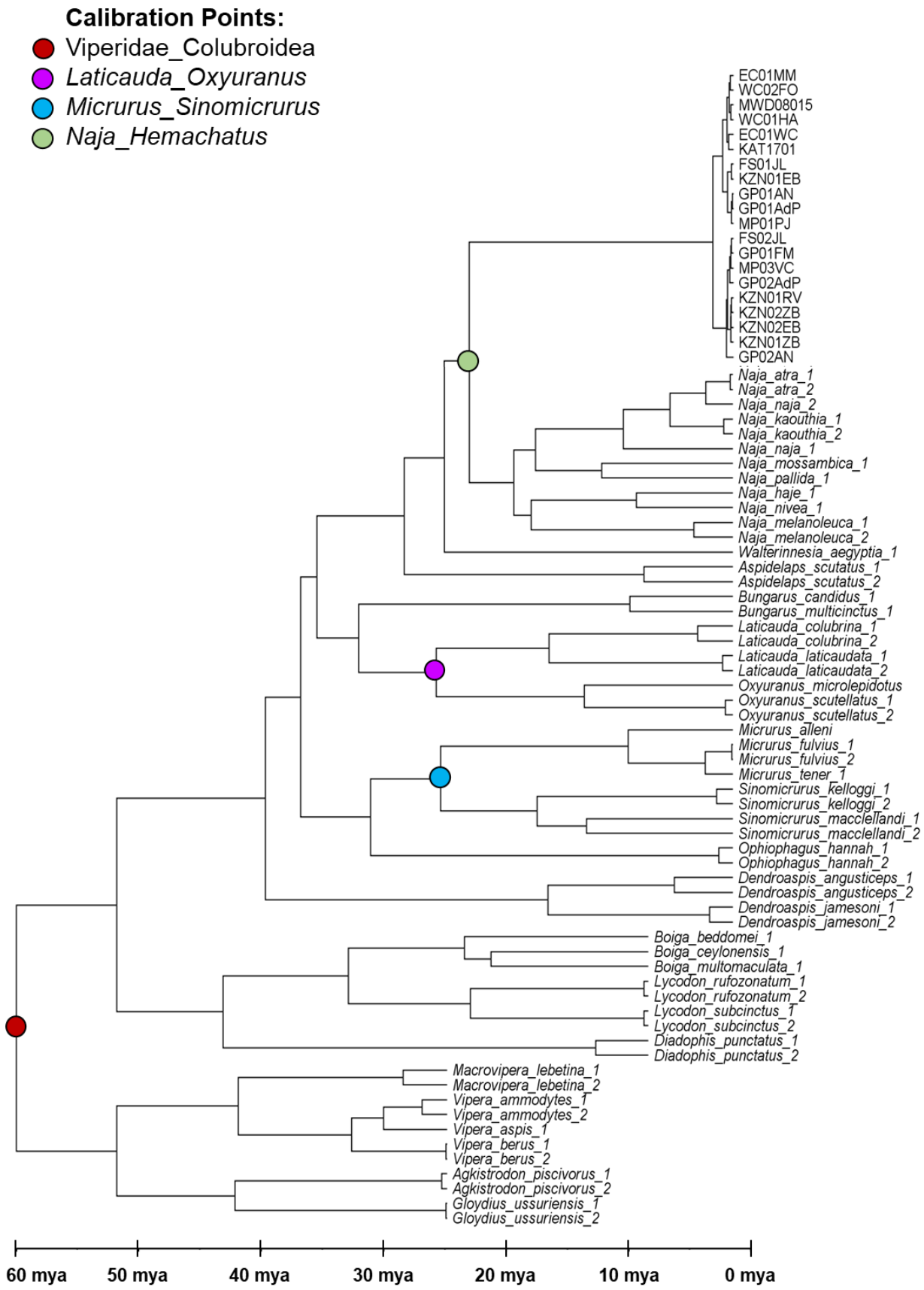

Figure S14: Dated phylogeny of *Hemachatus haemachatus* with outgroups generated from the combined nuclear and mitochondrial dataset showing calibration points.

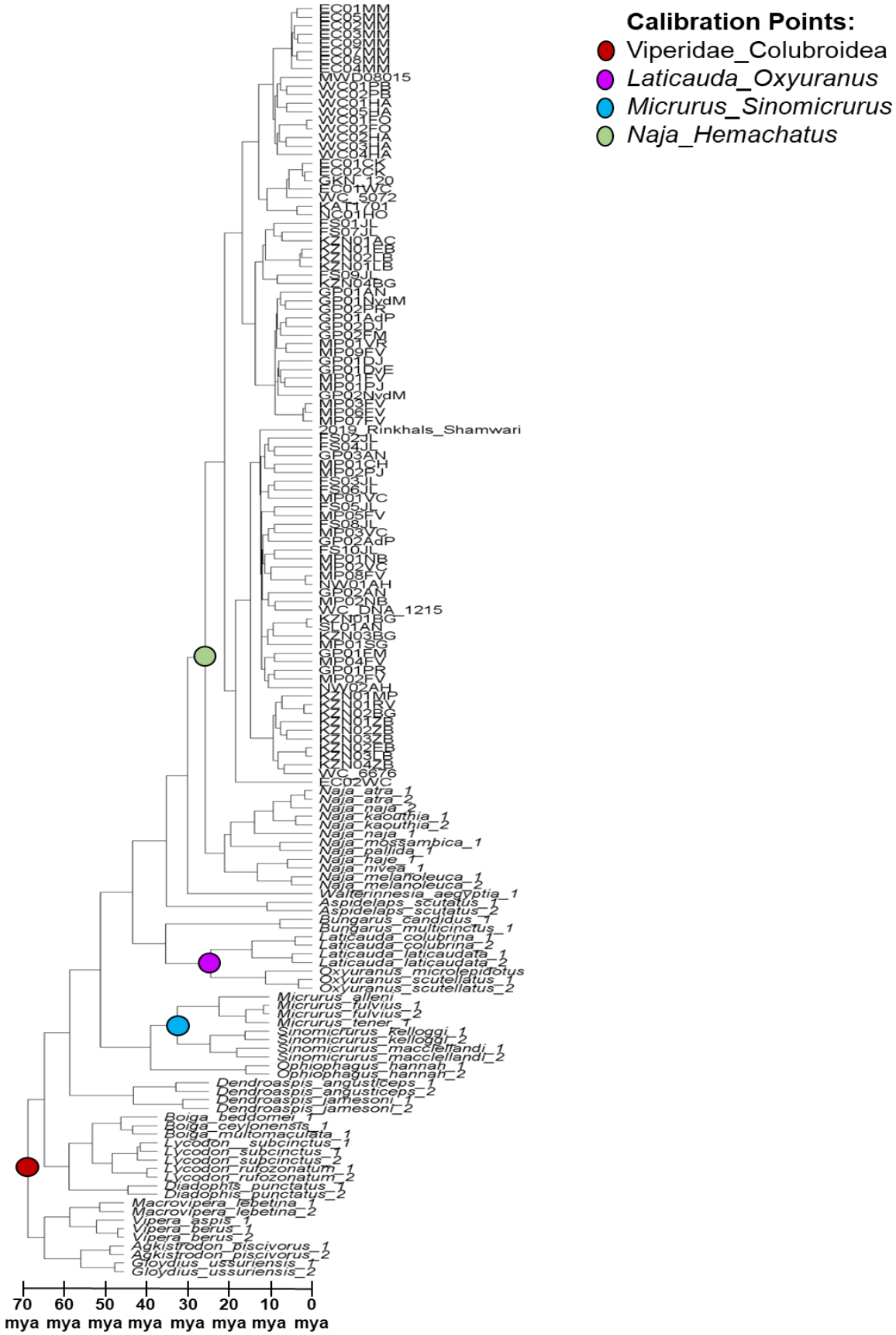

Figure S15: Dated phylogeny of *Hemachatus haemachatus* with outgroups generated from the full mitochondrial dataset showing calibration points.

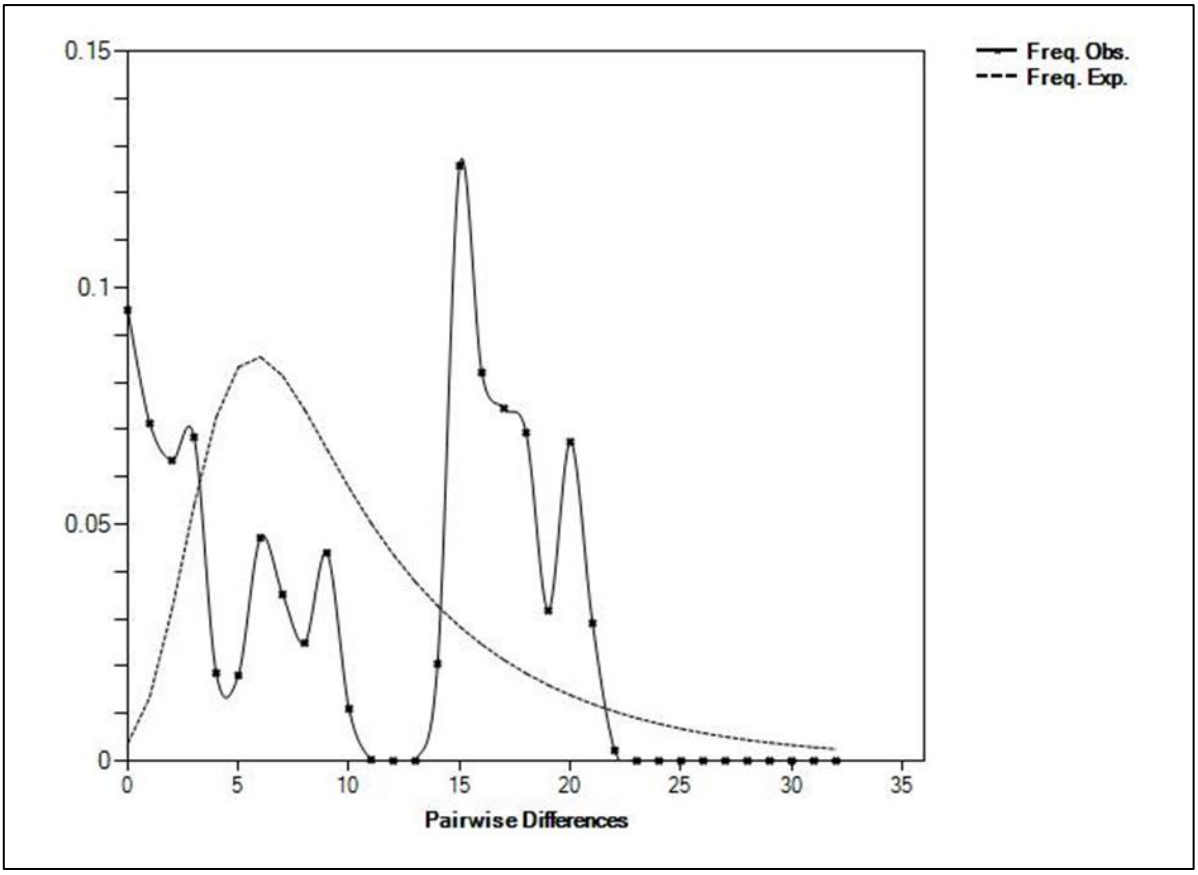

Figure S16: A ragged mismatch distribution produced for *Hemachatus haemachatus* under the population growth-decline model for the full mitochondrial dataset. The number of pairwise differences are shown on the x-axis and the frequency of the pairwise comparisons are shown on the y-axis. The dotted line indicates the frequency expected under the population growth-decline model and the solid line indicates the observed frequencies.

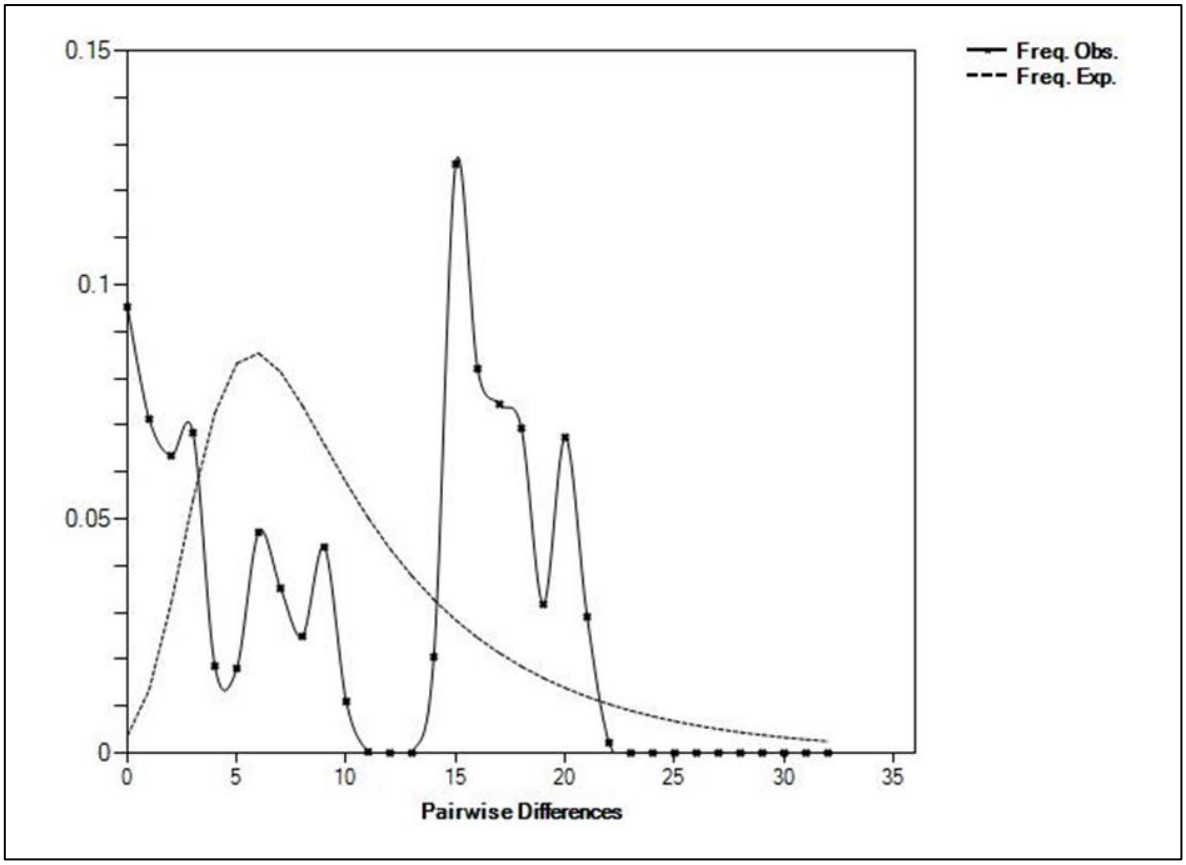

Figure S17: A ragged mismatch distribution produced for *Hemachatus haemachatus* under the population growth-decline model for the combined nuclear and mitochondrial dataset. The number of pairwise differences are shown on the x-axis and the frequency of the pairwise comparisons are shown on the y-axis. The dotted line indicates the frequency expected under the population growth-decline model and the solid line indicates the observed frequencies.

Script S3: R script specifying commands to run the cluster analysis for the venom data using the Aitchison distance measure and the complete linkage clustering method in the Compositions package.

```
# Load required packages

library(compositions)


# Load dataset

mydata = read.csv("directory path to file", row.names = 1)# load
dataset


# Mark data as Aitchison composition

cdata <- acomp(mydata)

cdata # view data


# Compute clustering

clusters <- hclust(dist(cdata), method = "ward.D")


# Plot dendogram

plot(clusters)

plot(as.dendrogram(clusters), horiz=T)
```

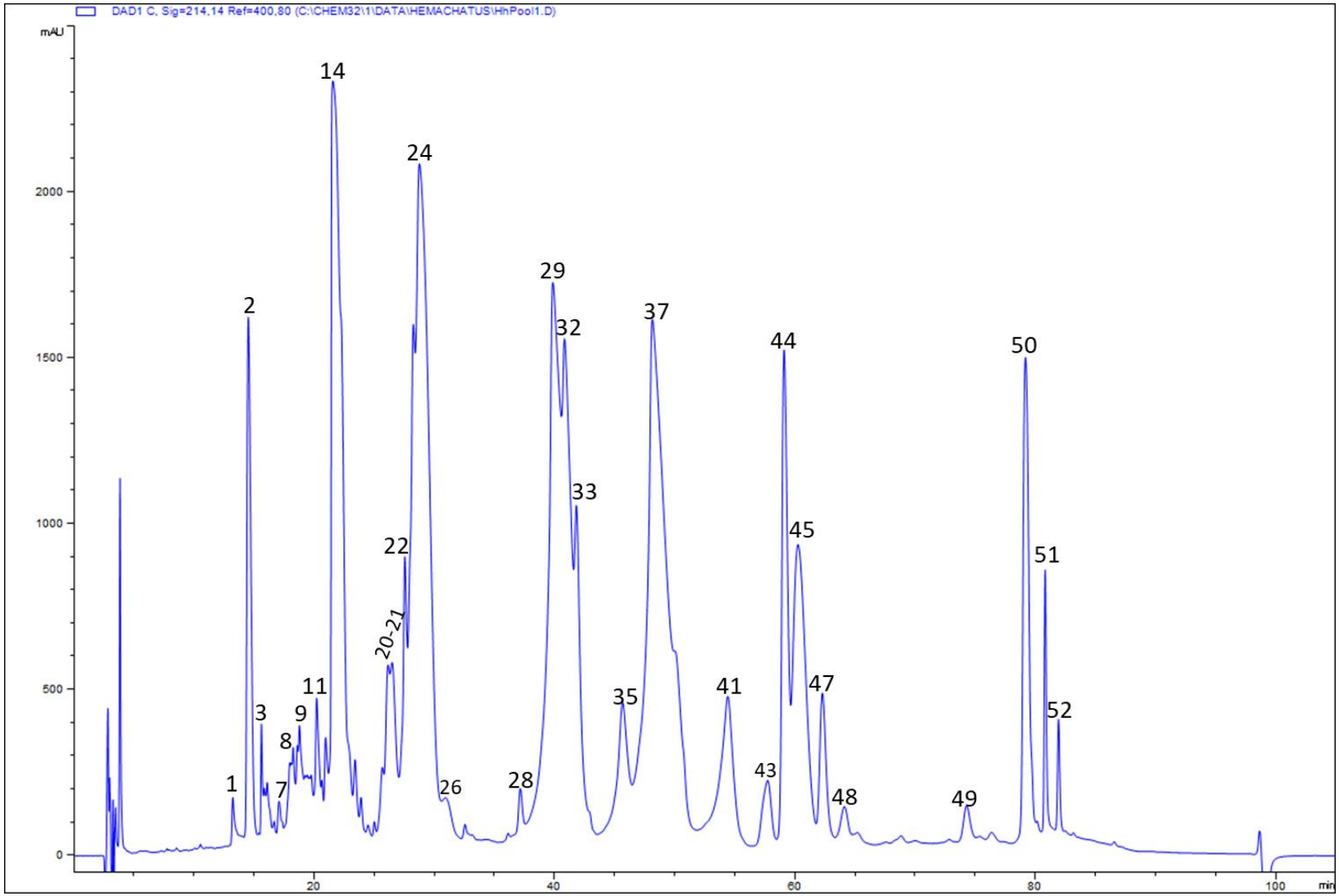

Figure S18A: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 1. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

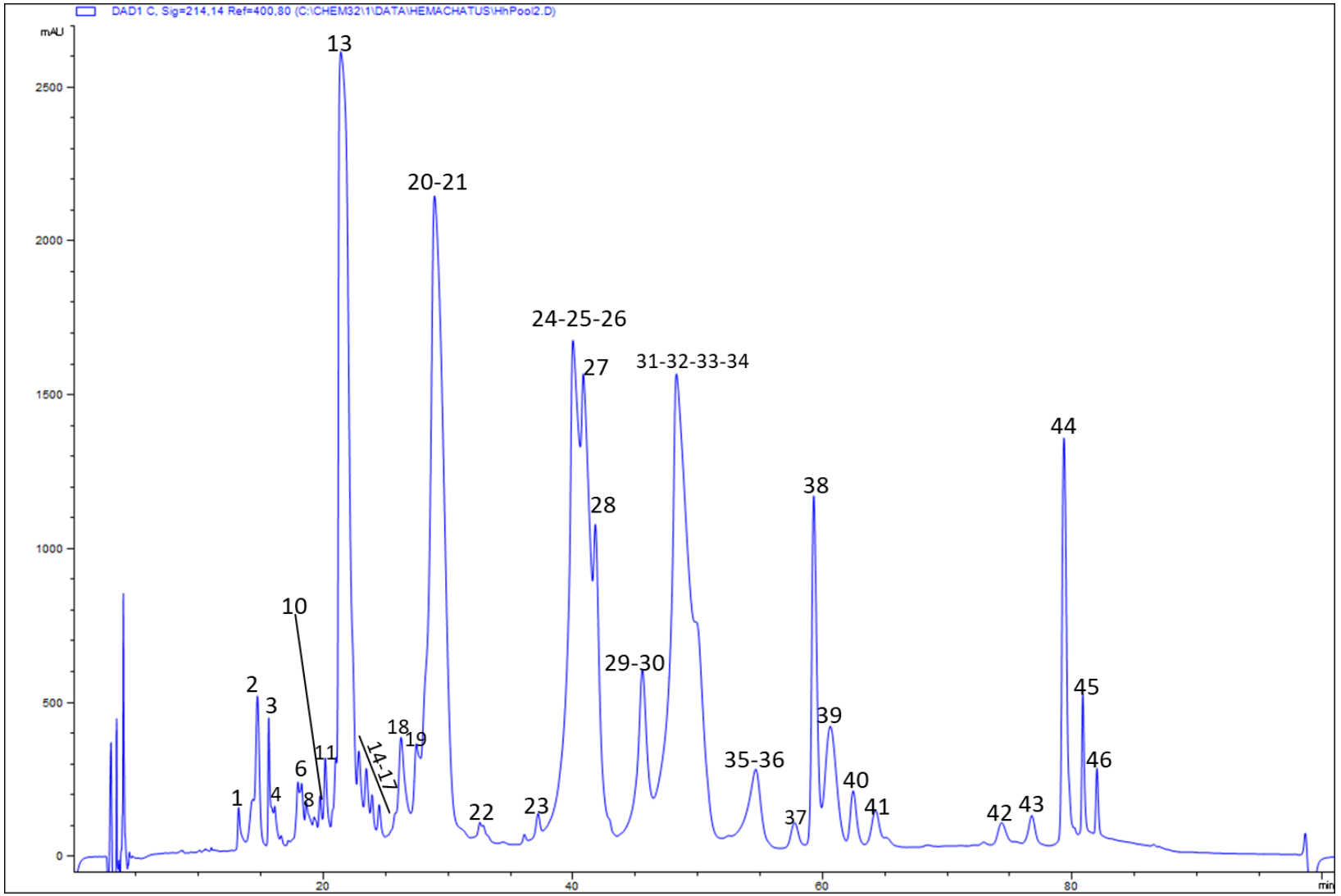

Figure S18B: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 2. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

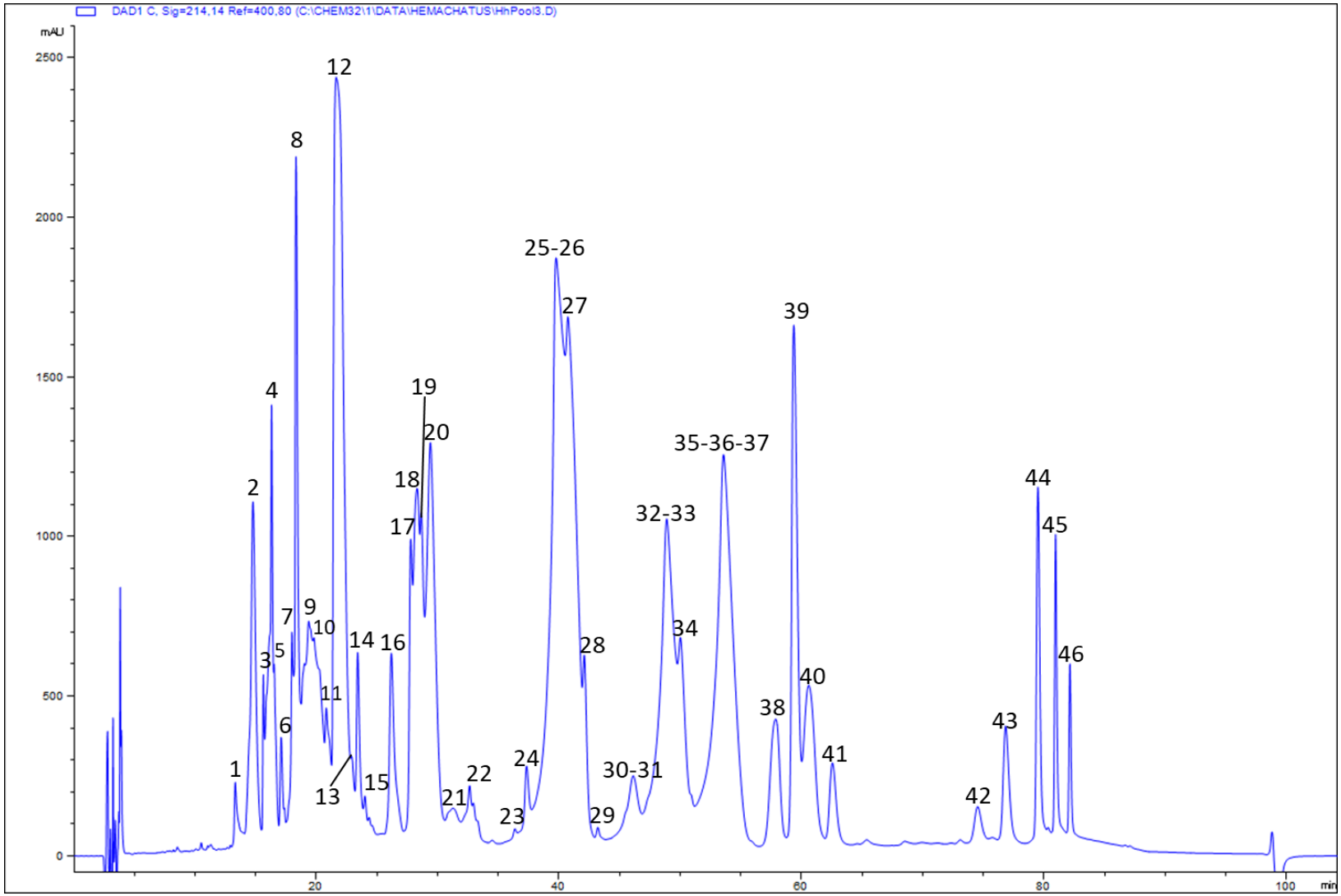

Figure S18C: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 3. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

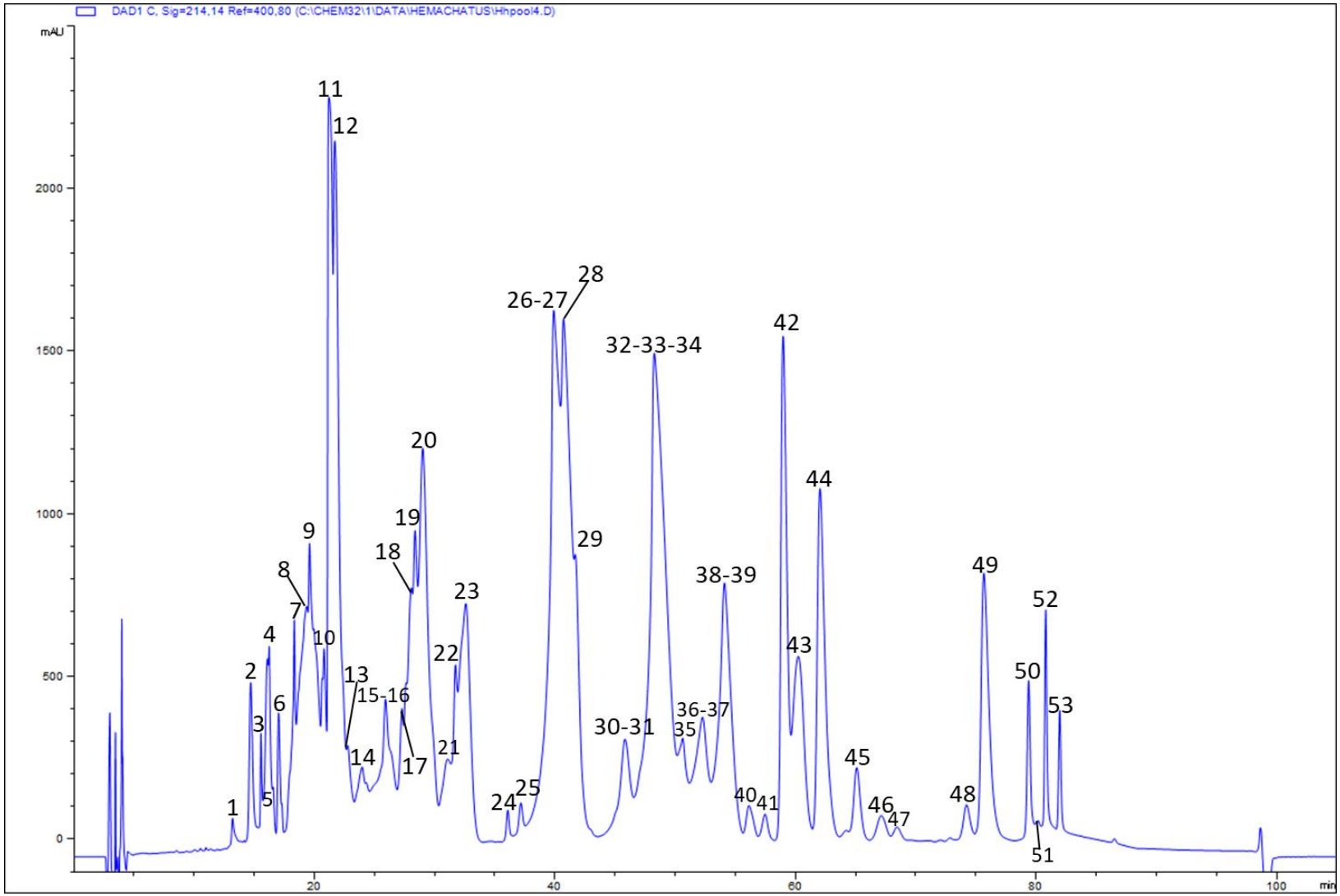

Figure S18D: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 4. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

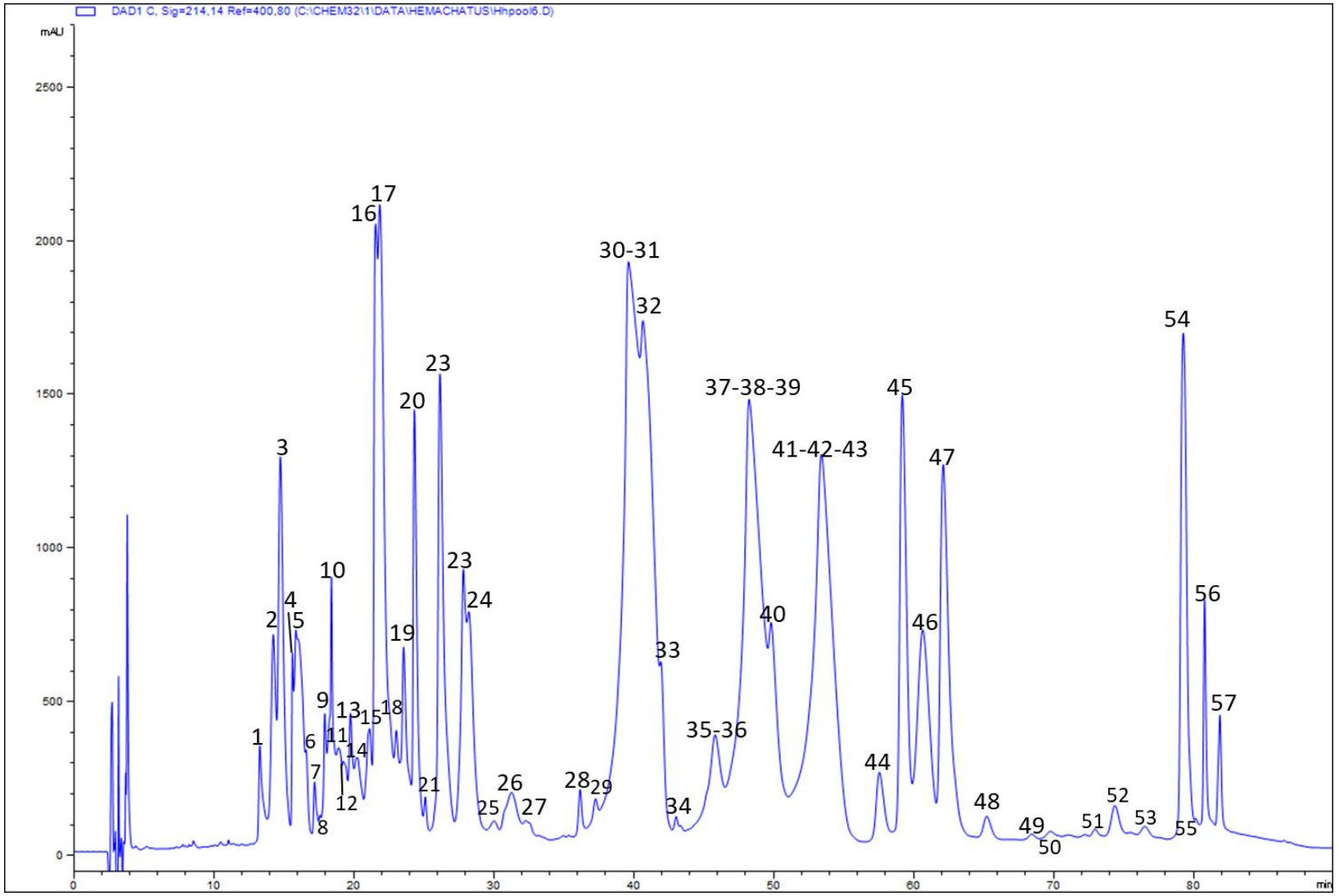

Figure D18E: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 5. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

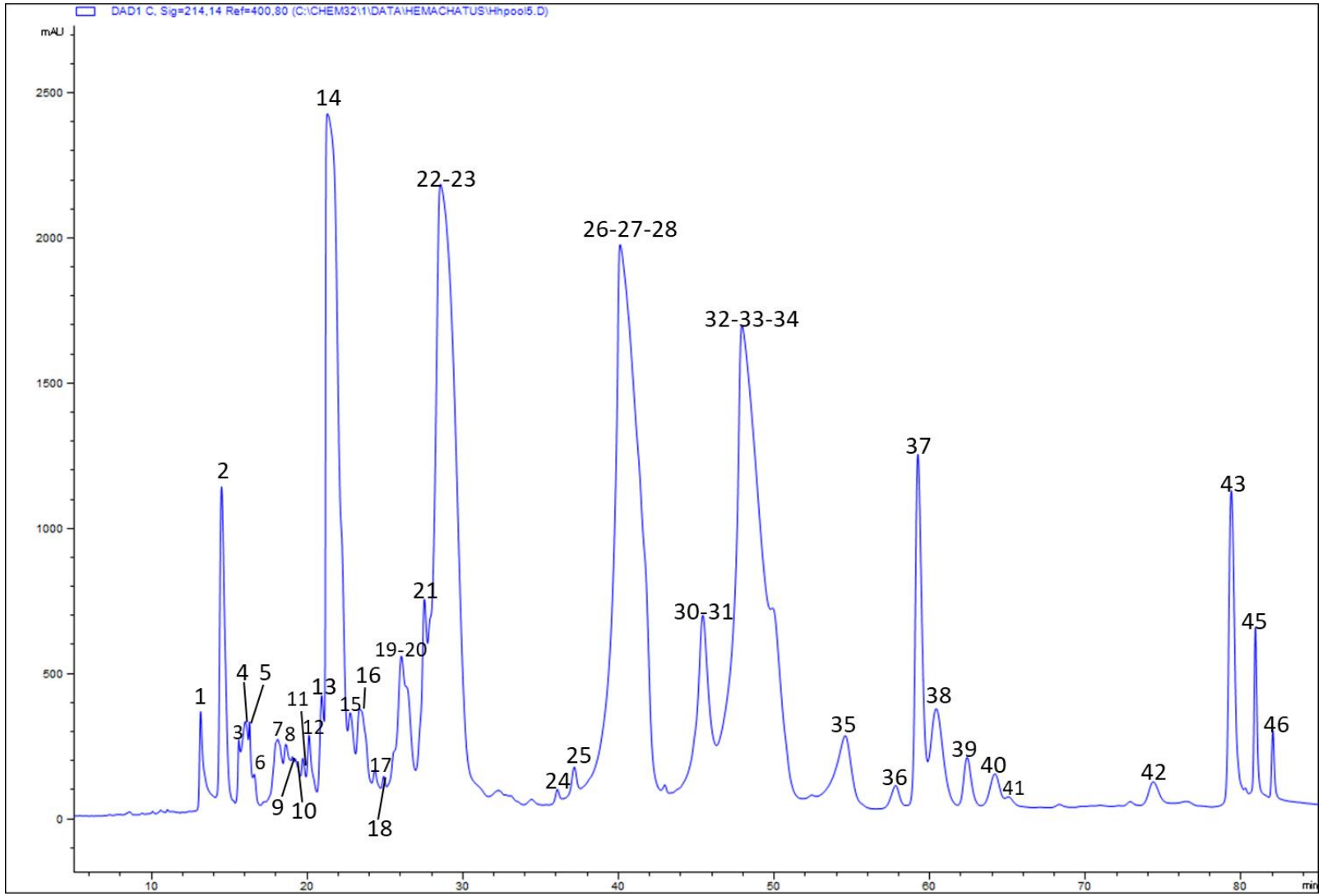

Figure S18F: Chromatogram showing the elution profile and band numbers of peaks from the rp-HPLC run for Pool 6. The x-axis shows time (min) as the run progressed and the y-axis shows absorbance intensities (mAU).

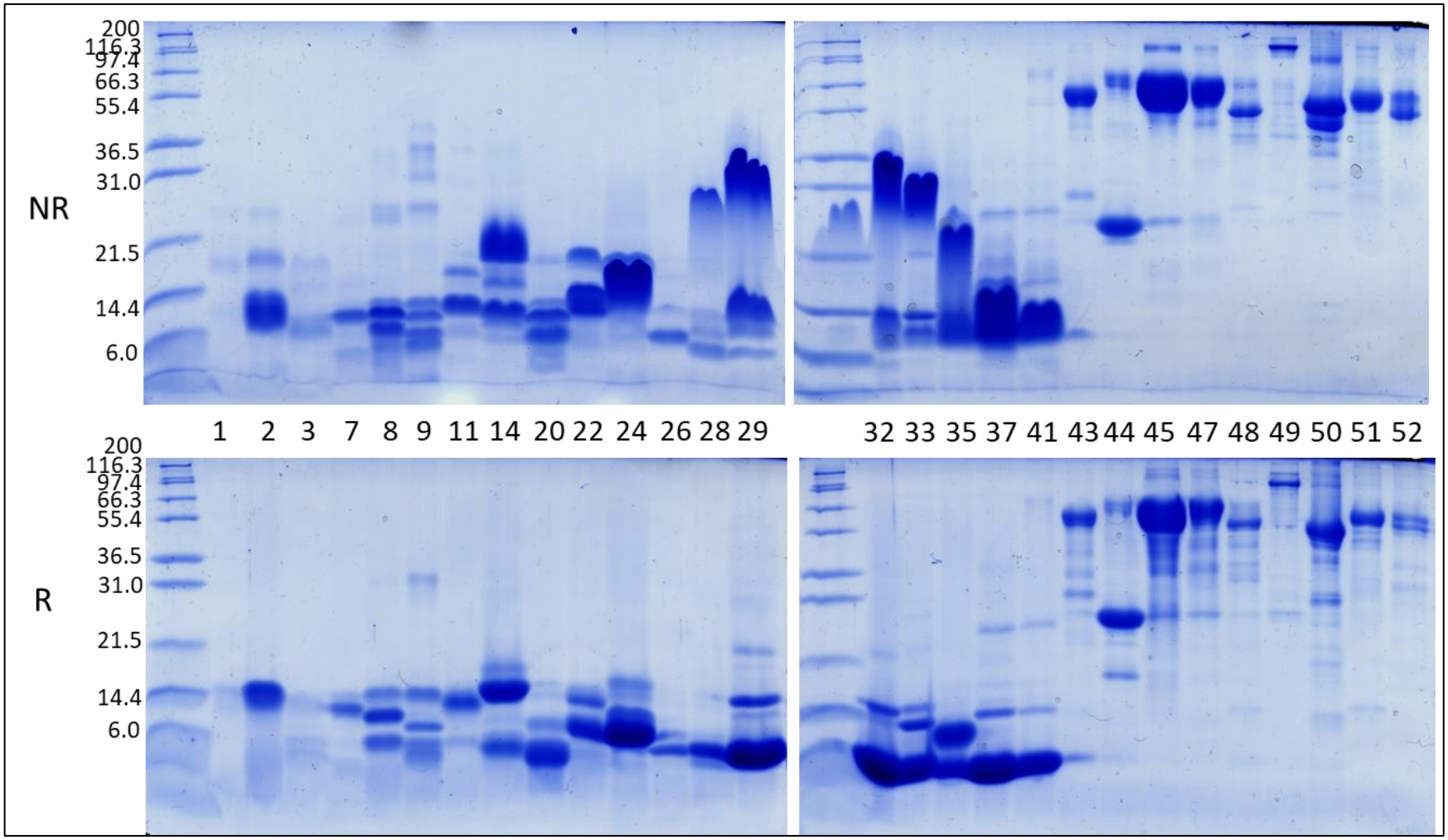

Figure S19A: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 1. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

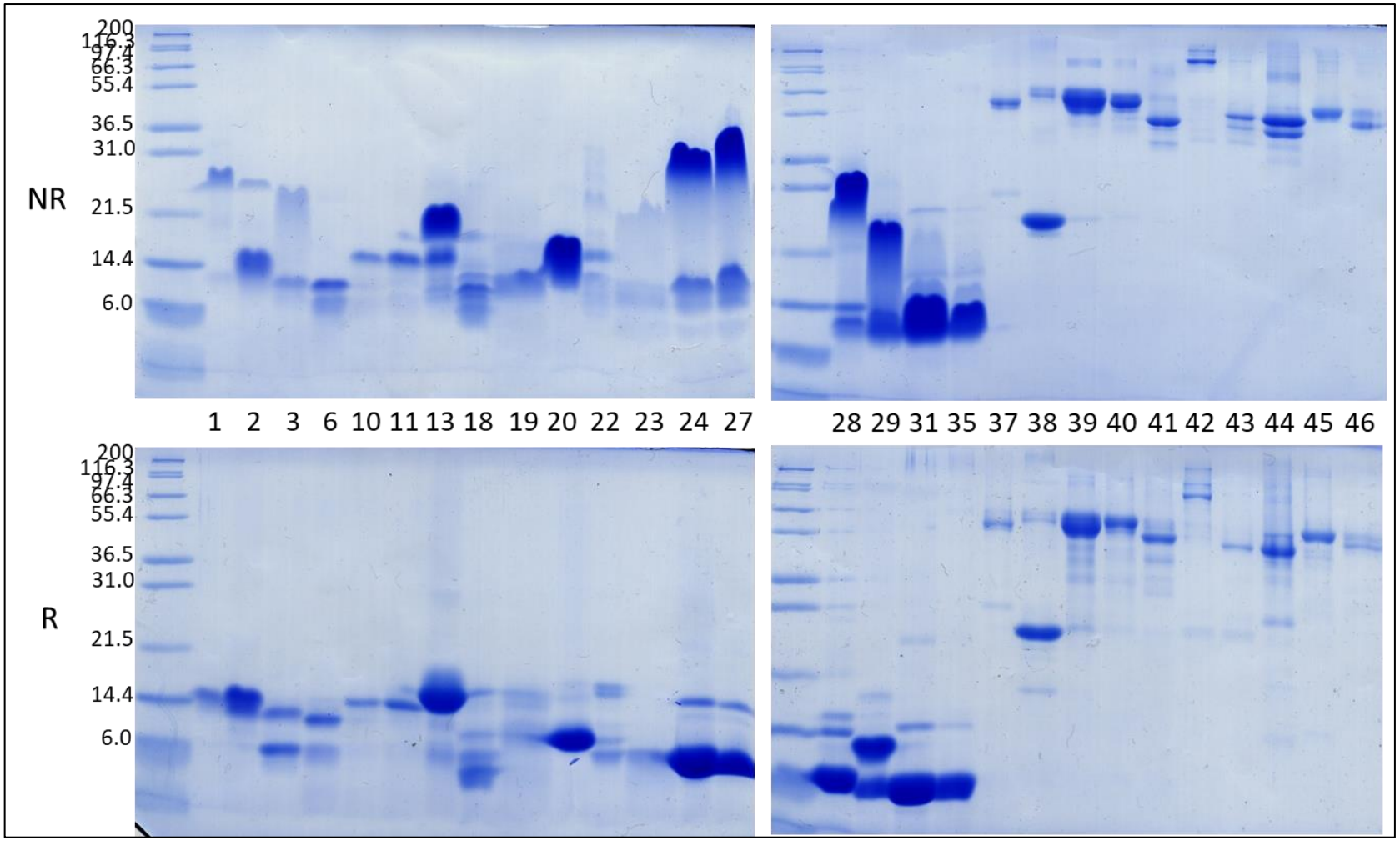

Figure S19B: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 2. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

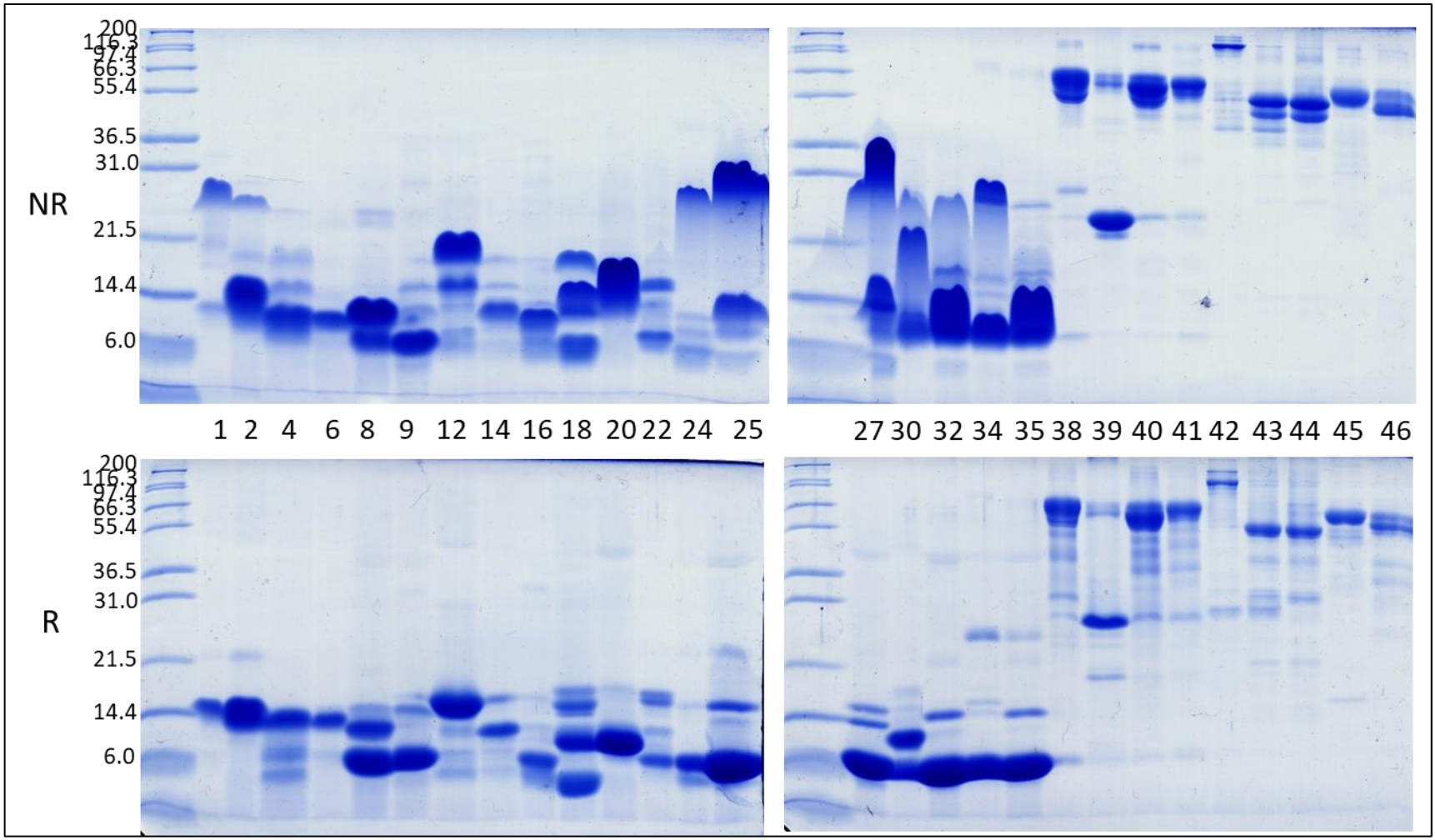

Figure S19C: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 3. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

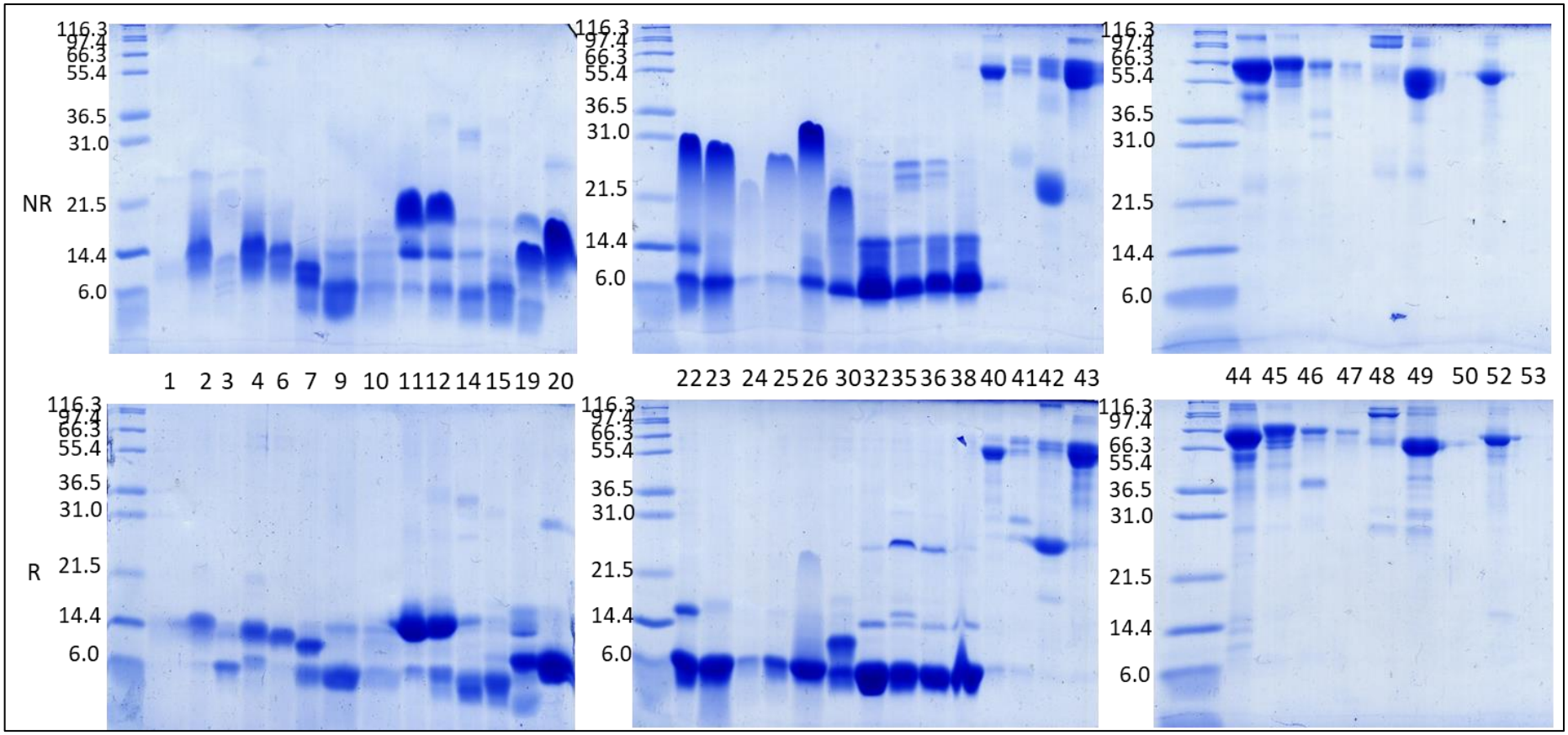

Figure S19D: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 4. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

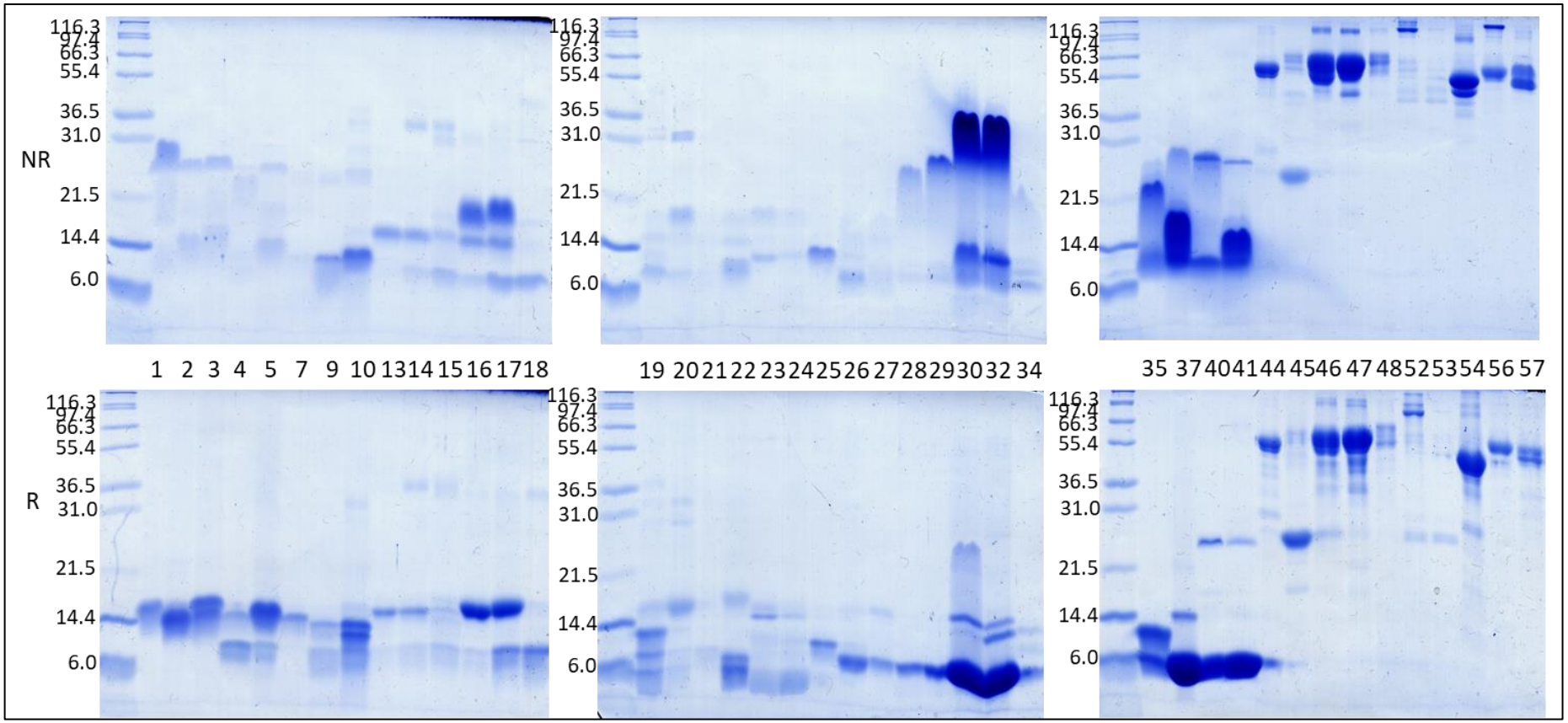

Figure S19E: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 5. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

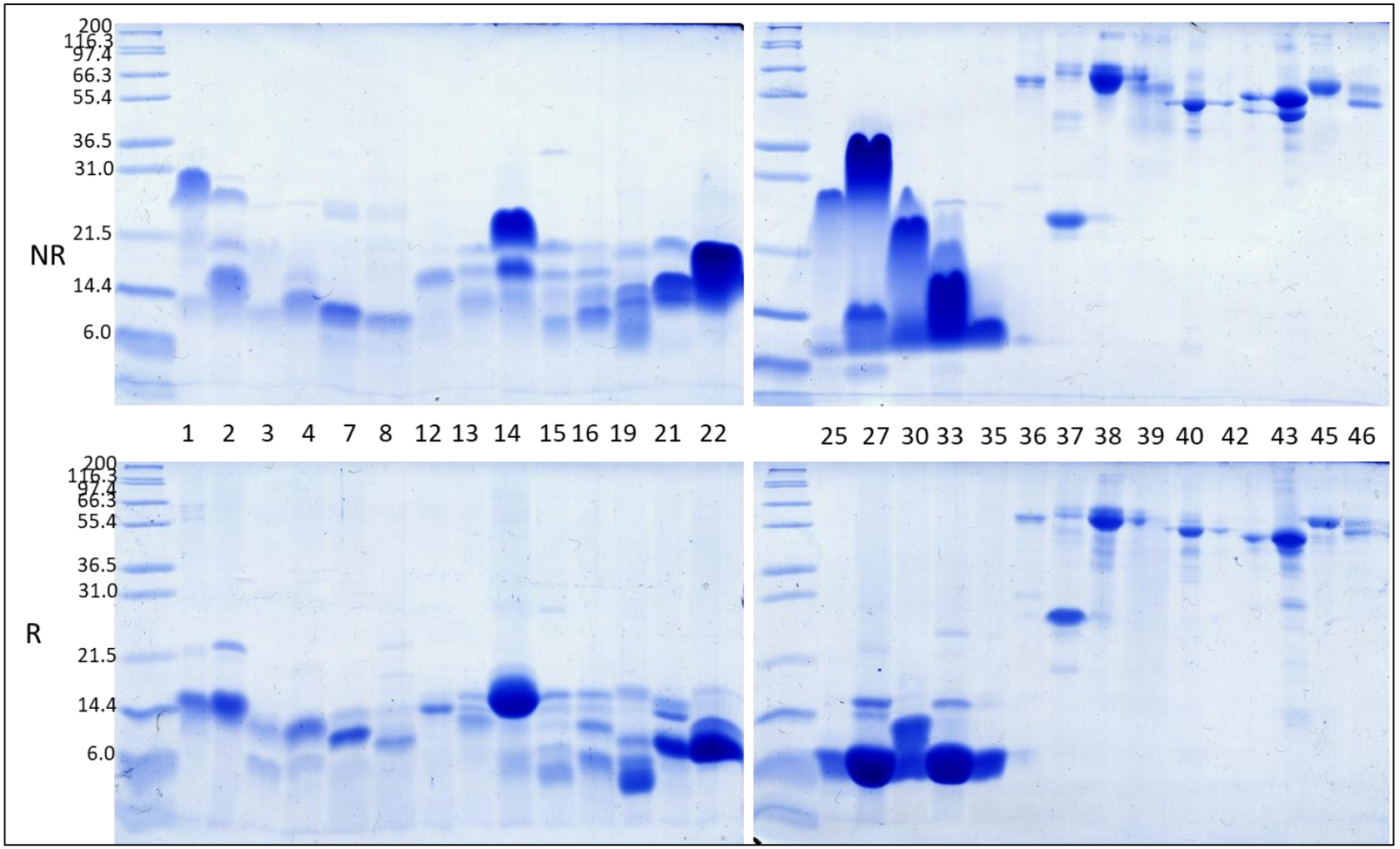

Figure S19F: SDS-PAGE gel image band numbers of peaks/fragments collected from the rp-HPLC run for Pool 6. The top two gels are were run under non-reducing conditions (NR) and the bottom two gels were run under reducing conditions (R). The vertical numbers on left indicate the size standard. The horizontal numbers in the middle indicate the band numbers corresponding to the peak numbers.

Table S2A: Proteins and protein families assigned to rp-HPLC fractions of *Hemachatus haemachatus* by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 1. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ▪ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that point. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		MS/MS-derived sequence	Best NCBI match	Protein family	Best protein match
			m/z	z				
1	0.331	18 ▪	669.8	2	GIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
			489.6	3	KGIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
2	4.187	14 ▼	863.9	2	GIELNCCTTDRCNN	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
			489.6	3	KGIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
						Hemachatus haemachatus P01433	3FTx	Short neurotoxin 2_Toxin IV
						Hemachatus haemachatus P01425	3FTx	Short neurotoxin 1_Toxin II
3	0.390	14 ▼				-	Unknown	Unknown
	0.390	6 ▼				-	Unknown	Unknown
7	0.035	14 ▼	669.8	2	GIELNCCTTDR	~ *Naja mossambica* P01431	3FTx	Short neurotoxin 1
						Hemachatus haemachatus P01433	3FTx	Short neurotoxin 2_Toxin IV
	0.140	12 ▼	649.8	2	TWCDAWCGSR	~ *Micropechis ikaheka* AHZ08824	3FTx	Long neurotoxin
8	0.088	14 ▼	649.8	2	TWCDAWCGSR	*Micropechis ikaheka* AHZ08824	3FTx	Long neurotoxin
			567.3	2	LTCLICPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11

	0.198	10 ▼	649.8	2	TWCDAWCGSR	*Micropechis ikaheka* AHZ08824	3FTx	Long neurotoxin
			567.3	2	LTCLICPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11
	0.154	5 ▼				*Naja haje haje* P01401	3FTx	Weak toxin CM-11
9	0.037	31 ▼	860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheka* AHZ08819	DC-domain	P-III snake venom metalloprotease
	0.111	14 ▼	567.3	2	LTCLICPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11
	0.093	7 ▼				-	Unknown	Unknown
	0.130	4 ▼	567.3	2	LTCLICPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11
11	0.598	13 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00596	PLA_2	Acidic phospholipase A2 1
	0.031	6 ▼	939.9	2	FTCFTTPSDTSETCPK	~ *Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
13	0.223					~ *Hemachatus haemachatus* P0DQH2	3FTx	Exactin
14	0.640	17 ▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			732.8	4	CDRLAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00596	PLA_2	Acidic phospholipase A2 1
			622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
						~ *Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
	10.243	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			732.8	4	CDRLAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFADYGCYCGR	~ *Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor

	0.075	8 ▼	622.2	3	SWWHFADYGCYCGR	~ *Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
	0.053	8 ▼	685.3	2	TCPEGKNLCYK	~ *Hemachatus haemachatus* B3EWH9	3FTx	Exactin
	1.489	4 ▼	622.2	3	SWWHFADYGCYCGR	~ *Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
	0.303	4 ▼	598.3	3	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
15	0.118					~ *Naja mossambica* P01469	3FTx	Cytotoxin V(II)2
16	0.251					~ *Naja mossambica* P01469	3FTx	Cytotoxin V(II)2
19	0.141					*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
20	0.018	16 ▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
	0.018	15 ▼	601.3	2	GGSGTPVDKLDR	*Bungarus caeruleus* Q6SLM0	PLA_2	Basic phospholipase A2 3
			594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
	0.088	7 ▼	685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
	1.641	4 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			949.4	3	SFHYNLAAQQCLQFIYGGCGGNANR	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			643.3	2	FIYGGCGGNANR	*Naja naja* P20229	KI	Kunitz-type serine protease inhibitor 2
			712.8	2	FCELPAETGLCK	*Naja nivea* P00986	KI	Venom basic protease inhibitor II
22	0.180	14 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
			690.3	2	ALTMoxEGNQASWR	*Naja atra* P61898	vNGF	Venom nerve growth factor

			754.3	3	GIDSSHWNSYCTETDTFIK	*Naja atra* P61898	vNGF	Venom nerve growth factor
						~ *Naja naja* P01140	vNGF	Venom nerve growth factor
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	0.010	14 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.445	7 ▼	754.3	3	GIDSSHWNSYCTETDTFIK	*Naja atra* P61898	vNGF	Venom nerve growth factor
24	0.903	15 ▼	921.9	2	SWWDFADYGCYCGR	*Naja naja* P15445	PLA2	Acidic phospholipase A2 2
			667.0	3	NLYQFKNMIKCTVPSR	*Naja naja* P15445	PLA2	Acidic phospholipase A2 2
	0.651	15 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
	13.989	8 ▼	922.4	2	GXVEECCFBSCDXAR	~ *Ophiophagus hannah* ETE61012	IGF	Insulin-like growth factor I
			611.3	2	XEMYCA(307)K	~ *Ophiophagus hannah* ETE61012	IGF	Insulin-like growth factor I
		8 ▼	692.3	2	TCPEGBNXCYK	~ *Micrurus altirostris* AED89572	3FTx	Putative three-finger toxin precursor
26	0.096	6 ▼	685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
						Hemachatus haemachatus C0HJT5	3FTx	Ringhalexin
	0.096	4 ▼	771.7	3	GCAATCPKPEAQVYVDCCAR	*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
27	0.170					-	Unknown	Unknown
28	0.175	30 ■	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
			688.4	2	MSMEVTPMIPIK	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
	0.029	12 ■	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor

						~ *Naja naja* P01140	vNGF	Venom nerve growth factor
	0.029	10 ▪	1028.1	3	LCLSDYSIFSETIEICPEGHNYCFK	*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
			688.4	2	MSMEVTPMIPIK	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
			527.3	3	MoxSMoxEVTPMIPIKR	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
			655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.058	6 ▪	922.4	2	GXVEECCFBSCDXAR	~ *Ophiophagus hannah* ETE61012	IGF	Insulin-like growth factor I
			611.3	2	XEMYCA(307)K	~ *Ophiophagus hannah* ETE61012	IGF	Insulin-like growth factor I
						Hemachatus haemachatus P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
29	0.355	21 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.033	21 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			402.3	2	LVPFLSK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
	1.549	14 ▼	663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.387	12 ▼	663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
	5.421	5 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			487.9	3	VVPFLSKTCPEGK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			552.3	3	SSLLVNVMCCKTDK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A

						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
32	0.184	22 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777/P01471	3FTx	Cytotoxin 3 Toxin 11/11A / Cytotoxin 1
	0.714	15 ▼	454.7	2	GCTDACPK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	~ *Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
	0.023	15 ▼	465.8	2	EWAVGLAGK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	2.762	5 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			454.7	2	GCTDACPK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			492.6	3	LVPFLSKTCPEGK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
33	0.085	15 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.010	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
	0.175	12 ▼	465.8	2	EWAVGLAGK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			647.0	3	KTVENVGVSQVAPDNPER	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			1071.5	2	ADVTFDSNTAFESLVVSPDK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			1135.6	2	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	0.010	12 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
			682.3	2	ALTMEGNQASWR	*Naja atra* P61898	vNGF	Venom nerve growth factor

						~ *Naja naja* P01140	vNGF	Venom nerve growth factor
	0.005	12 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.665	5 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			688.4	2	MSMEVTPMIPIK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
35	0.179	7 ▼	1071.5	2	ADVTFDSNTAFESLVVSPDK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	0.359	7 ▼	768.4	2	ETLVSILEEYPDK	*Bitis arietans* C0K3N1	VEGF	Snake venom vascular endothelial growth factor toxin barietin
	0.380	7 ▼	695.4	2	CHNTPLPFIYK	*Naja melanoleuca* P01473	3FTx	Cytotoxin homolog 3
	0.485	5 ▼	469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			688.9	2	MSMEVTPMIPIK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			521.9	3	MSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.126	5 ▼	757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						~ *Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
37	1.906	28▼	925.4	2	NMLQMEWNSNAAQNAK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			851.1	4	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			433.5	3	CSFAHSPSHLR	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin

	0.147	14▼	757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	3.505	14▼	516.6	3	MSMoxEVTPMIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			402.2	2	XVPFXSK	~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			410.2	2	XVPYXSK	~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
						~ *Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.159	14▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
	13.340	5▼	402.2	2	XVPFXSK	~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			688.3	2	MSMBVTPMIPIK	~ *Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						~ *Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~ *Naja annulifera* P01420	3FTx	Short neurotoxin 3
41	0.187	28▼	925.4	2	NMLQMEWNSNAAQNAK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			803.3	3	PSACVNGLCTNPCK	~*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			433.5	3	CSFAHSPSHLR	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin
	0.375	14▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
	3.183	5▼	527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
43	0.839	63▪	860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			708.9	2	YIEFYVVVDNR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1

	0.105	30▪	567.8	2	SVSPTASNMLK	*Leioheterodon madagascariensis* Q2XXQ1	CRISP	Cysteine-rich venom protein LEI1
			598.3	2	EIVDLHNSLR	*Naja naja* P86543	CRISP	Cysteine-rich venom protein (Najanajin)
	0.105	8▪	516.6	3	MSMoxEVTPMIPIKR	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.2	2	XVPFXSK	~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			688.3	2	MSMBVTPMIPIK	~ *Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
44	0.834	66▼	860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
			649.3	2	SAECPTDSFKR	*Thamnophis sirtalis* XP_013929562	SVMP	PREDICTED: Zinc metalloproteinase-disintegrin-like VLAIP-B
	4.169	26▼	918.8	3	NFVYGVGANPPGSVTGHYTQIVWYK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
	0.056	21▼	567.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.6	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			484.6	3	QKEIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
	0.500	17▼	575.8	2	SVSPTASNMoxLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.6	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			918.8	3	NFVYGVGANPPGSVTGHYTQIVWYK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin

			422.2	3	NVDFNSESTRR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			938.1	3	NFVYGVGANPPDSVTGHYTQIVWYK	*Rhabdophis tigrinus tigrinus* Q8JGT9	CRISP	Cysteine-rich venom protein tigrin
45	0.070	>116▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.070	116▼	725.8	2	DDCDLPEICTGR	~ *Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	6.320	55▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
			570.8	2	DSCFTLNQR	*Bungarus fasciatus* A8QL48	SVMP	Zinc metalloproteinase-disintegrin-like BfMP
			694.9	2	YIEFYVAVDNR	*Bungarus fasciatus* A8QL48	SVMP	Zinc metalloproteinase-disintegrin-like BfMP
	0.140	45▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	0.140	36▼	649.3	2	SAKCPTDSFQR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.140	34▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			649.3	2	SAKCPTDSFQR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.140	26▼	576.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			918.8	3	NFVYGVGANPPGSVTGHYTQIVWYK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin

47	0.033	>116▪	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	1.468	66▪	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			695.3	2	TSQLTNTPEQDR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			649.3	2	SAECPTDSFKR	*Thamnophis sirtalis* XP_013929562	SVMP	PREDICTED: Zinc metalloproteinase-disintegrin-like VLAIP-B
	0.114	50▪	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.016	24 ▪	484.6	3	QKEIVDLHNSLR	*Hemachatus haemachatus* P0DL17	CRISP	Cysteine-rich venom protein hematin
			575.8	2	SVSPTASNMoxLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			918.8	3	NFVYGVGANPPGSVTGHYTQIVWYK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
48	0.014	97▼	554.8	2	TLGMLMoxEGLK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			567.8	2	NPFYNPSPAK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			401.6	3	IDKVNLMVDR	*Python bivittatus* XP_007430581	PDE	Phosphodiesterase family member 3
			806.7	3	TEVTSFENIEVYNLMCDLLK	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
	0.610	55▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			736.4	2	TAPAFQFSSCSIR	*Naja atra* D3TTC1	SVMP	Zinc metalloproteinase-disintegrin-like kaouthiagin-like

	0.014	51▼	574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.014	36▼	574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.014	31▼	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			736.4	2	TAPAFQFSSCSIR	*Naja atra* D3TTC1	SVMP	Zinc metalloproteinase-disintegrin-like kaouthiagin-like
	0.014	26▼	567.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
49	0.046	>116▼	546.8	2	TLGMLMEGLK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			567.8	2	NPFYNPSPAK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
	0.046	116▼	567.8	2	NPFYNPSPAK	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			666.8	4	NLHNCVNLILLADHGMEAISCNR	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			431.7	2	VNLMoxVDR	*Python bivittatus* XP_007430581	PDE	Phosphodiesterase family member 3
			554.8	2	TLGMLMoxEGLK	*Python bivittatus* XP_007430581	PDE	Phosphodiesterase family member 3
			401.6	3	IDKVNLMVDR	*Python bivittatus* XP_007430581	PDE	Phosphodiesterase family member 3
	0.278	97▼	567.8	2	NPFYNPSPAK	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			888.8	3	NLHNCVNLILLADHGMEAISCNR	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			771.0	3	NPAWWGGQPIWHTATYQGLK	*Thamnophis sirtalis* XP_013908397	PDE	Phosphodiesterase family member 3
			554.8	2	TLGMoxLMEGLK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3

			502.3	3	SMQAIFLAHGPGFK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			464.2	2	BPXAETXR	*~Pseudonaja textilis* XP_026561286	PDE	Phosphodiesterase family member 3
	0.046	70▼	567.8	2	NPFYNPSPAK	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			666.8	4	NLHNCVNLILLADHGMEAISCNR	*Macropisthodon rudis* ALA20853	PDE	Phosphodiesterase family member 3
			554.8	2	TLGMoxLMEGLK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			678.3	2	AATYFWPGSEVK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			507.6	3	SMoxQAIFLAHGPGFK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase family member 3
			401.6	3	IDKVNLMVDR	*Python bivittatus* XP_007430581	PDE	Phosphodiesterase family member 3
	0.023	55▼	640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.023	29▼	746.9	2	LVILGFPCNQFGK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
50	0.117	>116▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.117	116▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.042	97▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin

	0.016	97▼	559.8	2	TLCAGILEGGK	*Gloydius blomhoffii* Q9PT51	SVSP	Beta-fibrinogenase brevinase
	5.279	56▼	570.3	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			648.8	2	DPSYGMoxVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.003	38▼	637.3	2	FTFLVQEMMK	*Protobothrops mucrosquamatus* XP_015678645	CHIT1	Chitotriosidase-1
			763.4	2	GFLAYYEICDFK	*Protobothrops mucrosquamatus* XP_015678645	CHIT1	Chitotriosidase-1
			805.9	2	ITTAEWNDEQLYK	*Protobothrops mucrosquamatus* XP_015678645	CHIT1	Chitotriosidase-1
	0.056	38▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			648.8	2	DPSYGMoxVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.176	32▼	648.8	2	DPSYGMoxVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			841.4	2	VYEIVNHLNMMYR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
	0.059	22▼	640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			710.3	2	NSPMVCXCPVSPR	~ *Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
51	1.038	64▼	575.8	2	IHFAGEYTGR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			608.8	2	FWEADGIHGGK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			661.3	2	EIQALCYPSIK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			670.3	2	EGWYVNMGPMR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			448.9	3	KFWEADGIHGGK	*Naja atra* AVX27607	LAO	L-amino acid oxidase

			721.4	2	VWEVKKDPSLLK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			559.6	4	HVVVVGAGMAGLSAAYVLAGAGHK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
	0.231	64▼	476.3	2	VGIIGYTTK	*Gloydius brevicaudus* B6EWW8	5'-NT	Snake venom 5'-nucleotidase
			725.4	2	VVSLNVLCTECR	*Gloydius brevicaudus* B6EWW8	5'-NT	Snake venom 5'-nucleotidase
	0.027	55▼	670.3	2	EGWYVNMGPMR	*Naja atra* A8QL58	LAO	L-amino acid oxidase
			559.6	4	HVVVVGAGMAGLSAAYVLAGAGHK	*Naja atra* A8QL58	LAO	L-amino acid oxidase
			834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
	0.025	50▼	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			641.3	2	SASQLYQESLR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			661.3	2	EIQALCYPSIK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			448.9	3	KFWEADGIHGGK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			470.6	3	SASQLYQESLRK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
	0.002	50▼	703.9	2	FTAYAINGPPVEK	*Drysdalia coronoides* F8J2D3	PLB	Phospholipase-B 81
			732.8	2	YGLDFSYEMoxAPR	*Drysdalia coronoides* F8J2D3	PLB	Phospholipase-B 81
			505.7	2	TWAQTFEK	*Ophiophagus hannah* ETE59578	PLB	Phospholipase B-like 1
	0.013	14▼	509.3	2	VTLLEASER	*Naja atra* A8QL58	LAO	L-amino acid oxidase
			563.6	4	HVVVVGAGMoxAGLSAAYVLAGAGHK	*Naja atra* A8QL58	LAO	L-amino acid oxidase
52	0.322	64▼	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			641.3	2	SASQLYQESLR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			481.3	3	VWEVKKDPSLLK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			559.6	4	HVVVVGAGMAGLSAAYVLAGAGHK	*Naja atra* AVX27607	LAO	L-amino acid oxidase

	0.322	60▼	742.8	2	DDCDFPELCTGR	*Thamnophis sirtalis* XP_013922279	SVMP	PREDICTED: Hemorrhagic metalloproteinase-disintegrin-like kaouthiagin
			853.9	2	ESPDVCFLLN(EGNG)R	*Pseudonaja textilis* XP_026570659	SVMP	Zinc metalloproteinase-disintegrin-like MTP4 isoform X9

Table S2B: Proteins and protein families assigned to rp-HPLC fractions of *Hemachatus haemachatus* by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 2. In this case, only a limited number of peaks were subjected to MS/MS mass determination. The results from Pools 1 and 4 were used to infer protein families where peptide ions were not subjected to MS/MS mass determination. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ▪ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that point. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		MS/MS-derived sequence	Best NCBI match	Protein family	Best protein match
			m/z	z				
1	0.397	14 ▼				~ *Naja annulifera* P68417	3FTx	Short neurotoxin 1
2	1.882	14 ▼				~ *Hemachatus haemachatus* P01433	3FTx	Short neurotoxin 2_Toxin IV
3	0.194	12 ▼					3FTx	
	0.292	6 ▼					KI	
4	0.414						3FTx	
6	0.499	12 ▼					3FTx	
	0.125	6 ▼					KI	
8	0.194						3FTx	
10	0.182	14 ▼					PLA2	
11	0.483	14 ▼					PLA2	
13	17.958	14 ▼					PLA2	

	0.181	6 ▼				~ *Hemachatus haemachatus* P0DQH2	3FTx	Exactin
14	0.259						3FTx	
15	0.296					~ *Naja mossambica* P01469	3FTx	Cytotoxin V(II)2
16	0.077						PLA2	
	0.078						3FTx	
17	0.087						PLA2	
	0.087						3FTx	
18	0.248	14 ▼	622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
	0.124	8 ▼	685.3	2	TCPEGQNICYK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
			741.3	3	GCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
			727.3	2	LEFGCAATCPTVK	*Laticauda colubrina* P0C8R8	3FTx	Toxin Lc b
	0.372	6 ▼	512.6	3	ITCSAEEKFCYK	*Walterinnesia aegyptia* C0HKZ8	3FTx	Actiflagelin
						Hemachatus haemachatus P00985	KI	Kunitz-type serine protease inhibitor 2
	0.496	3 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			949.4	3	SFHYNLAAQQCLQFIYGGCGGNANR	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			712.8	2	FCELPAETGLCK	*Naja nivea* P00986	KI	Venom basic protease inhibitor II
19	0.064	14 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
	0.078	14 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1

			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.213	8 ▼	685.3	2	TCPEGQNICYK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
			741.3	3	GCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
20	18.226	8 ▼				*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
22	0.132	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			932.9	2	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
	0.099	12 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			741.3	3	GCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
	0.027	6 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
			690.3	2	ALTMoxEGNQASWR	*Naja atra* P61898	vNGF	Venom nerve growth factor
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
	0.072	6 ▼	741.3	3	GCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
			663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
23	0.180	6 ▼				*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
24	0.757	14 ▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	0.379	12 ▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
	6.437	5 ▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
27	0.593	15 ▼					3FTx	
	3.360	5 ▼				*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1

28	0.099	15 ▼					3FTx	
	0.012	15 ▼					PLA_2	
	0.204	12 ▼					Vespryn	
	0.011	12 ▼					vNGF	
	0.006	12 ▼					3FTx	
	0.774	5 ▼				*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
29	0.134	18 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.287	7 ▼					Vespryn	
						Hemachatus haemachatus P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.575	7 ▼					VEGF	
	0.609	7 ▼				*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.848	5 ▼					3FTx	
	0.221	5 ▼					Vespryn	
31	0.217	28▼					CRISP	
	0.042	14▼					Vespryn	
	1.000	14▼					3FTx	
	0.045	14▼					vNGF	
	20.442	5▼				*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~*Naja annulifera* P01420	3FTx	Short neurotoxin 3
35	0.122	14▼					3FTx	
	2.321	5▼					3FTx	

37	0.384	63■					SVMP	
	0.043	30■					CRISP	
38	0.433	66▼					SVMP	
	3.463	26▼					CRISP	
	0.130	21▼					CRISP	
	0.303	17▼					CRISP	
39	0.031	>116▼					SVMP	
	0.031	116▼					SVMP	
	2.919	55▼					SVMP	
	0.031	45▼					SVMP	
	0.031	36▼					SVMP	
	0.031	34▼					SVMP	
	0.031	26▼					CRISP	
40	0.740	66■					SVMP	
41	0.605	55▼					SVMP	
	0.151	51▼					SVMP	
42	0.077	>116▼					SVMP	
	0.077	116▼					SVMP	
	0.194	97▼					SVMP	
	0.039	29▼					CRISP	
43	0.463	50 ▼	912.5	2	YIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin

			651.3	3	KYIEFYVVVDNIMYR			
44	0.936	55 ▼	542.2	2	QCVDVKTAY	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			912.5	2	YIEFYVVVDNIMYR			
			651.3	3	KYIEFYVVVDNIMYR			
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	2.340	50 ▼					SVMP	
	0.024	36 ▼					CHIT1	
	0.444	36 ▼					SVMP	
	0.936	28 ▼					SVMP	
45	1.024	55 ▼					LAO	
	0.066	55 ▼					5'-NT	
46	0.257	55 ▼					LAO	
	0.257	50 ▼					SVMP	

Table S2C: Proteins and protein families assigned to rp-HPLC fractions of *Hemachatus haemachatus* by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 3. In this case, only a limited number of peaks were subjected to MS/MS mass determination. The results from Pools 1 and 4 were used to infer protein families where peptide ions were not subjected to MS/MS mass determination. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ▪ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that point. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		MS/MS-derived sequence	Best NCBI match	Protein family	Best protein match
			m/z	**z**				
1	0.441	14 ▼					3FTx	
2	0.171	21 ▼	669.8	2	GIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
	3.254	14 ▼				*Hemachatus haemachatus* P01433	3FTx	Short neurotoxin 2_Toxin IV
4	0.467						3FTx	
6	1.839	14 ▼				~ *Naja annulifera* P68417	3FTx	Short neurotoxin 1
	0.920	6 ▼	695.8	2	ECEEFIYGGCK	*Bungarus fasciatus* B2KTG1	KI	Kunitz-type serine protease inhibitor bungaruskunin
	0.920	4 ▼	695.8	2	ECEEFIYGGCK	*Bungarus fasciatus* B2KTG1	KI	Kunitz-type serine protease inhibitor bungaruskunin
8	0.421	14 ▼					3FTx	
9	0.566						3FTx	
12	0.277	14 ▼					3FTx	
	0.830	10 ▼					3FTx	

	1.660	5 ▼					3FTx	
14	0.710	14 ▼					3FTx	
	1.657	5 ▼					3FTx	
16	0.239						3FTx	
18	10.639	15 ▼					PLA_2	
	0.626	12 ▼	622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA_2	Neutral phospholipase A2 muscarinic inhibitor
	0.626	6 ▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA_2	Basic phospholipase A2 DE-1
	0.626	5 ▼	557.8	2	(201.1)PAFFYNR	*Bungarus fasciatus* P25660	KI	Kunitz-type serine protease inhibitor IX
20	0.077	15 ▼					PLA2	
	0.077	15 ▼					PLA2	
	0.463	12 ▼					PLA2	
	0.077	6 ▼					PLA2	
	0.077	5 ▼					KI	
	0.079					~ *Naja mossambica* P01469	3FTx	Cytotoxin 2
	1.815	6 ▼				*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
	0.441	16 ▼					PLA2	
	0.881	9▼					vNGF	
	1.542	6▼					3FTx	
	1.542	3▼					KI	
	3.077	6▼				*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB

	0.200	16 ▼				*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
22	0.143	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			932.9	2	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.143	15▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.053	12▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.004	12▼	456.3	2	ILNVLSPR	*Ophiophagus hannah* ETE63541	IGF	Insulin-like growth factor-binding protein 3
	0.228	6▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			771.7	3	GCAATCPKPEAQVYVDCCAR	*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
24	0.221	30 ▪					3FTx	
	0.037	12 ▪					vNGF	
	0.037	10 ▪					3FTx	
	0.074	6 ▪					IGF	
						Hemachatus haemachatus P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
25	0.307	21 ▼					3FTx	
	0.028	21 ▼					PLA2	
	2.239	14 ▼					3FTx	
	0.336	12 ▼					3FTx	
	8.284	5 ▼					3FTx	
27	0.013	16 ▼	465.8	2	EWAVGLAGK	*Naja kaouthia* P82885	Vespryn	Thaicobrin

			604.3	3	TVENVGVSQVAPDNPER	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	1.071	16 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3
	1.084	14 ▼	905.9	2	TVENVGVSQVAPDNPER	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	5.060	5 ▼				*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
	0.123						PLA2	
	0.287					*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
30	0.094	7 ▼					Vespryn	
	0.188	7 ▼					VEGF	
	0.199	7 ▼					3FTx	
	0.381	5 ▼					3FTx	
	0.099	5 ▼					Vespryn	
32	0.059	14▼					Vespryn	
	1.395	14▼					3FTx	
	0.063	14▼					vNGF	
	6.066	6▼				~ *Naja annulifera* P01420	3FTx	Short neurotoxin 3
34	0.356	26▼	925.4	2	NMLQMEWNSNAAQNAK	*Ophiophagus hannah* ETE62137	CRISP	Hypothetical protein L345_12110
			884.8	3	FVYGVGANPPGSVIGHYTQIVWYK	*Ophiophagus hannah* ETE62137	CRISP	Hypothetical protein L345_12110
			1276.1	2	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			860.8	3	YLYVCQYCPTGNIIGSIATPYK	*Naja atra* Q7ZZN8	CRISP	Cysteine-rich venom protein natrin-2
	0.015	16▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1

	0.223	16▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3
	0.132	15▼	604.3	3	TVENVGVSQVAPDNPER	*Naja kaouthia* P82885	Vespryn	Thaicobrin
			1071.5	2	ADVTFDSNTAFESLVVSPDK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
	0.106	15▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
	1.544	6▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3
			469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3
			511.3	3	MSMEVTPMIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3
						~ *Naja annulifera* P01420	3FTx	Short neurotoxin 3
			695.4	2	CHNTPLPFIYK	*N. melanoleuca* P01473	3FTx	Cytotoxin homolog 3
			544.3	3	IKCHNTPLPFIYK	*N. melanoleuca* P01473	3FTx	Cytotoxin homolog 3
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
35	1.212	28▼					CRISP	
	2.424	14▼					vNGF	
	8.485	6▼					3FTx	
38	0.048	116▪	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	2.119	63▪					SVMP	
	0.120	30▪					CRISP	
	0.120	8▪					3FTx	
39	1.770	66▼					SVMP	
	2.950	26▼					SVMP	
	0.295	21▼					CRISP	
	0.885	17▼					CRISP	

40	0.035	>116▼					SVMP	
	0.035	116▼					SVMP	
	3.136	55▼					SVMP	
	0.070	45▼					SVMP	
	0.070	36▼					SVMP	
	0.070	34▼					SVMP	
	0.070	26▼					CRISP	
41	0.011	>116▪					SVMP	
	1.062	60▪					SVMP	
	0.011	45▪					SVMP	
	0.057	24 ▪					CRISP	
42	0.186	97▼					PDE	
	0.371	55▼					SVMP	
	0.062	26▼					CRISP	
43	0.029	116▼	598.8	2	TSVAVABDYSK	~ *Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.029	97▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.029	66▼	598.8	2	TSVAVABDYSK	~ *Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	1.110	55▼	912.5	2	YIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin

			976.5	2	KYIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.072	39▼	606.0	4	(821.4)AHEMGHNXGXNHDR	~ *Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			598.8	2	TSVAVABDYSK	~ *Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.072	31▼	746.9	2	LVILGFPCNQFGK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
	0.072	28▼	746.9	2	LVILGFPCNQFGK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
	0.029	21▼	583.3	2	PSYGMVAPDTK	~ *Naja mossambica* Q10749	SVMP	Snake venom metalloproteinase-disintegrin-like mocarhagin
44	2.345	56▼					SVMP	
	0.007	38▼					CHIT1	
	0.123	38▼					SVMP	
	0.130	32▼					SVMP	
45	0.960	64▼					LAO	
	0.213	64▼					5'-NT	
	0.156	55▼					LAO	
	0.147	50▼					LAO	
	0.009	50▼					PLB	
	0.078	14▼					LAO	
46	0.395	64▼					LAO	
	0.395	60▼					SVMP	

Table S2D: Proteins and protein families assigned to rp-HPLC fractions of Hemachatus haemachatus by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 4. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ■ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that base. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		Peptide sequence	Best NCBI match	Protein family	Best protein match
			m/z	z				
1	0.191	14 ▼	669.8	2	GIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
						~ *Naja naja* BAU24670	3FTx	~Cytotoxin 13
2	1.006	14 ▼	669.8	2	GIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
			425.2	3	GCGCPTVKPGIK	*Laticauda crockeri* P25495	3FTx	Short neurotoxin A
						~ *Hemachatus hemachatus* P01433	3FTx	Short neurotoxin 2_Toxin IV
	0.010	5 ▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
3	0.039	9▼	500.9	3	GCGCPTVKPGINLK	*Hemachatus haemachatus* P01425	3FTx	Short neurotoxin 1_Toxin II
			877.4	2	LECHNQQSSQPPTTK	*Hemachatus haemachatus* P01425	3FTx	Short neurotoxin 1_Toxin II
	0.154	5 ▼	638.8	2	CEEFXYGGCK	~ *Walterinnesia aegyptia* C1IC53	KI	Protease inhibitor 4
4	0.308	14 ▼	750.9	2	GCGCPTVKPGINLK	*Hemachatus haemachatus* P01425	3FTx	Short neurotoxin 1_Toxin II
	1.234	6▼	549.0	4	ECEEFIYGGCKGNKNNFK	*Bungarus fasciatus* B2KTG1	KI	Kunitz-type serine protease inhibitor bungaruskunin

			729.0	3	(513.2)NYKHAYYYNPAXR	~ *Bungarus fasciatus* B2KTG1	KI	Kunitz-type serine protease inhibitor bungaruskunin
6	0.661	14 ▼	496.9	3	GCGCPTVKP(357.3)K	~ *Demansia vestigiata* A6MFK6	3FTx	Short neurotoxin SNTX-1
7	0.736	10 ▼	567.3	2	XTCXXCPEK	~ *Naja atra* Q9YGI1	3FTx	Probable weak neurotoxin NNAM3
	0.490	5 ▼	543.6	3	GCVASCPSFESHYK	*Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
			567.3	2	LTCLICPEK	*Naja atra* Q9YGI1	3FTx	Probable weak neurotoxin NNAM3
9	0.508	14 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus P00595*	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFANYGCYCGR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	4.568	5 ▼	543.6	3	GCVASCPSFESHYK	*Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
10	0.237	14 ▼	564.3	2	ISGCRPYFK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFANYGCYCGR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.237	5 ▼	814.8	2	GCVASCPSFESHYK	*Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
11	5.317	15 ▼	519.2	2	TYSYDCTK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.6	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.280	6 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1

			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
12	0.155	36 ▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	DC domain	Zinc metalloproteinase-disintegrin-like cobrin
			649.3	2	(145.1)ECPTDSFQR	~ *Pseudechis australis* ABQ01134	DC domain	Zinc metalloproteinase-disintegrin-like cobrin
	4.959	15 ▼	564.3	2	ISGCRPYFK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.3	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.155	8 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.930	5▼	747.3	2	TTPSDTSETCPXGK	~ *Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
13	0.080						3FTx	
14	0.018	33▼	649.3	2	(145.1)ECPTDSFQR	~ *Pseudechis australis* ABQ01134	DC domain	Australease-1
			570.8	2	DSCFTLNQR	~ *Naja kaouthia* Q9PVK7	DC domain	Zinc metalloproteinase-disintegrin-like cobrin
	0.018	26▼	648.8	2	SAECPTDSFKR	*Thamnophis sirtalis* XP_013929562	DC domain	Zinc metalloproteinase-disintegrin-like VLAIP-B

			656.3	2	SAEC^{pro}PTDSFKR	~*Thamnophis sirtalis* XP_013929562	DC domain	Zinc metalloproteinase-disintegrin-like VLAIP-B
	0.072	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.108	5 ▼	747.3	2	TTPSDTSETCPXGK	~ *Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
	0.144	4▼	747.3	2	TTPSDTSETCPXGK	~ *Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
15	0.065	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.065	9▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.065	6▼	727.3	2	IEMoxGCAATCPTVK	*Hydrophis hardwickii* Q8UW28	3FTx	Alpha-elapitoxin-Lh2a
			537.3	3	RIEMoxGCAATCPTVK	*Hydrophis hardwickii* Q8UW28	3FTx	Alpha-elapitoxin-Lh2a
	1.108	5▼	948.1	3	FTCFTTPSDTSETCPIGNNICYEK	*Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
			747.3	2	TTPSDTSETCPXGK	~ *Hemachatus haemachatus* P25676	3FTx	Weak toxin CM-1c
			585.2	2	TCTEEDTWK	~ *Dendroaspis jamesoni kaimosae* P25682	3FTx	Toxin S6C6
19	0.128	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1

			611.3	2	NMoxIKCTVPSR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			921.9	2	SWWDFADYGCYCGR	*Naja atra* P00598	PLA2	Acidic phospholipase A2 1
	0.255	12▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
			682.3	2	ALTMEGNQASWR	*Naja atra* P61898	vNGF	Venom nerve growth factor
			754.3	3	GIDSSHWNSYCTETDTFIK	*Naja atra* P61898	vNGF	Venom nerve growth factor
	0.638	6▼	595.3	4	RGCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	0.255	3▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			712.8	2	FCELPAETGLCK	*Naja nivea* P00986	KI	Kunitz-type serine protease inhibitor 2
20	0.258	30 ▼	685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			566.8	2	PXVQVECCK	~ *Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	2.324	6▼	685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			566.8	2	PXVQVECCK	~ *Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
			573.8	2	PXVQVEC^{pro}CK	~ *Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
			568.8	2	SXVQVEC^{pro}CK	~ *Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
22	0.499	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	1.497	6▼	663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			457.2	3	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1

						~ *Naja mossambica* P01470	3FTx	Cytotoxin V(II)3
23	0.224	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	4.247	6▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			454.7	2	GCTDACPK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			457.5	3	TCPEGKNLCYK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			469.6	3	SSLLVKVVCCNK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~ *Naja mossambica* P01470	3FTx	Cytotoxin V(II)3
24	0.158	6▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			696.4	2	MSMEVTPMoxIPIK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			516.6	3	MoxSMEVTPMIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~ *Naja mossambica* P01470	3FTx	Cytotoxin V(II)3
25	0.002	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus P00595*	PLA2	Basic phospholipase A2 DE-1
	0.183	6▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1

26	8.820	6▼	454.7	2	GCTDACPK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			402.3	2	LVPFLSK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			658.6	3	TCPEGKNLCYKMTLMK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
			688.4	2	MSMEVTPMIPIK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						Hemachatus haemachatus P01471	3FTx	Cytotoxin 1
28	10.257					*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
30	0.059	16 ▼	685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* B3EWH9	3FTx	Three-finger hemachatoxin
	0.003	16 ▼	832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.042	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
	0.019	15 ▼	527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.552	7 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			695.4	2	CHNTPLPFIYK	*Naja melanoleuca* P01473	3FTx	Cytotoxin homolog 3
	0.552	4 ▼	469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			452.5	3	TCPDGKNLCYK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			516.6	3	MSMEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			685.3	2	TCPEGKNLCYK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
32	0.644	26 ▼	529.2	2	CPASCFCR	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin

	1.932	14 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			521.9	3	MoxSMEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	3FTx	Thaicobrin
	10.302	6▼	469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			452.2	3	TCPDGKNLCYK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			516.6	3	MSMoxEVTPMIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~*Naja annulifera* P01420	3FTx	~Short neurotoxin 3
35	0.137	26 ▼	933.4	2	NMLQMoxEWNSNAAQNAK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			803.6	3	SGPPCGDCPSACVNGLCTNPCK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			1034.2	3	CSSSKYLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			884.8	3	FVYGVGANPPGSVIGHYTQIVWYK	*Bungarus candidus* ACE73577	CRISP	Cysteine-rich secretory protein Bc-CRPb
			886.9	4	FSCGENLFMoxSSQPYAWSKVIQSWY DENKK	*Bungarus candidus* ACE73577	CRISP	Cysteine-rich secretory protein Bc-CRPb
			860.8	3	YLYVCQYCPTGNIIGSIATPYK	*Naja atra* Q7ZZN8	CRISP	Cysteine-rich venom protein natrin-2
			808.0	3	SGPPCGDCPSACVKGLCTNPCK	*Parasuta nigriceps* ACY68723	CRISP	CRISP isoform 2
	0.006	15 ▼	529.2	2	CPASCFCR	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Laticauda semifasciata* Q8JI38	CRISP	Cysteine-rich venom protein latisemin
	0.029	15 ▼	594.8	2	GGSGTPVDDLDR	*Oxyuranus scutellatus scutellatus* P00615	PLA2	Neutral phospholipase A2 homolog taipoxin beta chain 1
	0.034	14 ▼	757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	3FTx	Thaicobrin
			402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A

			521.9	3	MSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.481	6▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			521.9	3	MSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
36	0.237	26 ▼	470.6	3	VIQSWYDENKK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			925.4	2	NMLQMEWNSNAAQNAK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			803.6	3	SGPPCGDCPSACVNGLCTNPCK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			860.8	3	YLYVCQYCPTGNIIGSIATPYK	*Naja atra* Q7ZZN8	CRISP	Cysteine-rich venom protein natrin-2
			808.0	3	SGPPCGDCPSACVKGLCTNPCK	*Parasuta nigriceps* ACY68723	CRISP	CRISP isoform 2
	0.047	14 ▼	757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	3FTx	Thaicobrin
	0.032	14 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	1.264	6▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			469.7	2	GCTDTCPK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			452.5	3	TCPDGKNLCYK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
38	0.261	26 ▼	925.4	2	NMLQMEWNSNAAQNAK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
			851.1	3	YLYVCQYCPAGNIIGSIATPYK	*Naja kaouthia* P84808	CRISP	Cysteine-rich venom protein kaouthin-2
	0.522	14 ▼	516.6	3	MSMEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	4.434	6▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A

			452.5	3	TCPDGKNLCYK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			527.3	3	MoxSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
40	0.007	67 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			910.0	3	DDCDLPEICTGRSAKCPTDSFQR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.334	55 ▼	860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			944.4	2	NGLPCQNNQGYCYNGK	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			695.3	2	TSQLTNTPEQDR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			510.6	3	TKPAYQFSSCSVR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.007	46 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.007	40 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			649.3	2	SADC^{pro}PTDSFBR	~ *Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.004	36 ▼	649.3	2	SGEC^{pro}PTDSFBR	~ *Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			574.3	3	AAKDDCDXPEXCTGR	~ *Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.004	31 ▼	575.8	2	SVSPTASNMoxLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL14	CRISP	Cysteine-rich venom protein annuliferin-a

	0.004	26 ▼	575.8	2	SVSPTASNMoxLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.6	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL14	CRISP	Cysteine-rich venom protein annuliferin-a
	0.004	6 ▼	516.6	3	MSMoxEVTPMIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
41	0.066	67 ▼	758.9	2	KYIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	0.066	55 ▼	574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.066	31 ▼	585.3	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			598.3	2	EIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			484.6	3	QKEIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			575.8	2	SVSPTASNMoxLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
	0.053	26 ▼	567.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.9	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL14	CRISP	Cysteine-rich venom protein annuliferin-a
	0.013	6 ▼	402.3	2	LVPFLSK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			521.9	3	MoxSMEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			663.8	2	SSLLVNVMoxCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
42	0.604	116 ▼	451.3	3	EIVDLHNSLRR	*Naja naja* P86543	CRISP	Cysteine-rich venom protein

	0.604	66 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			860.9	2	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			695.3	2	TSQLTNTPEQDR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.604	55 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			758.9	2	KYIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	3.925	26 ▼	597.8	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.6	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			402.7	4	QKEIVDLHNSLRR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			918.8	3	NFVYGVGANPPGSVTGHYTQIVWYK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			938.1	3	NFVYGVGANPPDSVTGHYTQIVWYK	*Rhabdophis tigrinus tigrinus* Q8JGT9	CRISP	Cysteine-rich venom protein tigrin
	0.302	16 ▼	567.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			598.3	2	EIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			484.6	3	QKEIVDLHNSLR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			402.7	4	QKEIVDLHNSLRR	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
			584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			938.1	3	NFVYGVGANPPDSVTGHYTQIVWYK	*Rhabdophis tigrinus tigrinus* Q8JGT9	CRISP	Cysteine-rich venom protein tigrin
43	0.075	75 ▼	536.6	3	VHEIVNFINGFYR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			1077.5	3	HDNAQLLTAIVLDEDTLGLAYLSSMCNPR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor

			932.4	4	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			758.9	2	KYIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
			1241.6	2	VSLTDLEIWSDQDFITVQSSAK	*Crotalus ruber ruber* P20897	SVMP	Snake venom metalloproteinase HT-2
	3.387	50 ▼	1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
			758.9	2	KYIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	0.075	38 ▼	1077.5	3	HDNAQLLTAIVLDEDTLGLAYLSSMCNPR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			1241.6	2	VSLTDLEIWSDQDFITVQSSAK	*Crotalus ruber ruber* P20897	SVMP	Snake venom metalloproteinase HT-2
	0.075	36 ▼	1077.5	3	HDNAQLLTAIVLDEDTLGLAYLSSMCNPR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	0.075	33 ▼	1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor

	0.075	24 ▼	584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			484.6	3	QKEIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			598.3	2	EIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			938.1	3	NFVYGVGANPPDSVTGHYTQIVWYK	*Rhabdophis tigrinus tigrinus* Q8JGT9	CRISP	Cysteine-rich venom protein tigrin
44	0.053	>116 ▼	828.1	3	VSLTDLEIWSDQDFITVQSSAK	*Crotalus ruber ruber* P20897	SVMP	Snake venom metalloproteinase HT-2
			1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
	0.053	116 ▼	1242.9	3	VSLTDLEIWSDQDYITVQSSAENTLHSFGEWR	*Crotalus molossus molossus* AAM27042	SVMP	Metalloproteinase precursor
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.053	70 ▼	649.3	2	SGEC^{pro}PTDSFBR	*~ Demansia vestigiata ABK63559*	SVMP	Metalloproteinase precursor
			569.3	2	(187.1)AVXBDYSR	*~ Naja atra* D5LMJ3	SVMP	Zinc metalloproteinase-disintegrin-like atrase-A
	4.748	60 ▼	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			515.6	3	KYIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.158	50 ▼	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			943.9	2	NGLPCQNNQGYCYNGK	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.053	47 ▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			772.9	2	KYIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor

	0.053	37▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.053	33 ▼	694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca AHZ08819*	SVMP	P-III snake venom metalloprotease
	0.053	31 ▼	584.8	2	NVDFNSESTR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			484.6	3	QKEIVDLHNSLR	*Naja annulifera* P0DL15	CRISP	Cysteine-rich venom protein annuliferin-b
			567.8	2	SVSPTASNMLK	*Ophiophagus hannah* Q7ZT98	CRISP	Cysteine-rich venom protein ophanin
45	0.021	>116 ▼	725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
	0.021	116 ▼	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.830	70 ▼	708.9	2	YIEFYVVVDNR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
			569.3	2	WAVXBDYSR	~ *Naja atra* D5LMJ3	SVMP	Zinc metalloproteinase-disintegrin-like atrase-A
			649.3	2	SAECPTDSFBR	~ *Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			856.4	2	(185.1)VSSPXCGNYFVER	~ *Austrelaps superbus* ABH10621	SVMP	Asrin
	0.104	58 ▼	649.3	2	SAECPTDSFQR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			515.6	3	KYIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.031	48 ▼	656.3	2	VYEMVNTXNTK	~ *Pseudechis australis* ABQ01134	SVMP	Australease-1
			856.4	2	(185.1)VSSPXCGNYFVER	~ *Austrelaps superbus* ABH10621	SVMP	Asrin
	0.031	36 ▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
46	0.268	66 ▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
	0.045	58 ▼	708.9	2	YIEFYVVVDNR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
	0.134	38 ▼	695.3	2	TSQLTNTPEQDR	*Naja atra* D5LMJ3	SVMP	Zinc metalloproteinase-disintegrin-like atrase-A

			637.6	3	NGHPCQNNQGYCYNGK	*Naja atra* D5LMJ3	SVMP	Zinc metalloproteinase-disintegrin-like atrase-A
			708.9	2	YIEFYVVVDNR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
47	0.193	66▪	649.3	2	SAKCPTDSFQR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
			574.3	3	AAKDDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
48	0.047	>166▪	567.8	2	NPFYNPSPAK	*Thamnophis sirtalis* XP_013908397	PDE	PREDICTED: ectonucleotide pyrophosphatase/phosphodiesterase family member 3
			678.3	2	AATYFWPGSEVK	*Thamnophis sirtalis* XP_013908397	PDE	PREDICTED: ectonucleotide pyrophosphatase/phosphodiesterase family member 3
			666.8	4	NLHNCVNLILLADHGMEAISCNR	*Thamnophis sirtalis* XP_013908397	PDE	PREDICTED: ectonucleotide pyrophosphatase/phosphodiesterase family member 3
			546.8	2	TLGMLMEGLK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase
	0.011	>166▪	770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
	0.386	166▪	567.8	2	NPFYNPSPAK	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase
			470.2	3	TLGMoxLMoxEGLKQR	*Macrovipera lebetina* AHJ80885	PDE	Phosphodiesterase
	0.010	66▪	476.3	2	VGIIGYTTK	*Protobothrops flavoviridis* BAN82018	5'-NT	5'-nucleotidase
			513.8	2	ASGNPILLNK	*Protobothrops flavoviridis* BAN82018	5'-NT	5'-nucleotidase
			859.9	2	ETPVLSNPGPYLEFR	*Protobothrops flavoviridis* BAN82018	5'-NT	5'-nucleotidase
			807.7	3	FHECNLGNLICDAVIYNNVR	*Protobothrops flavoviridis* BAN82018	5'-NT	5'-nucleotidase
	0.010	55▪	570.8	2	DSCFTLNQR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			494.9	3	VSRDSCFTLNQR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin

	0.010	46▪	746.9	2	LVILGFPCNQFGK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
	0.010	44▪	467.9	3	FLVNPQGKPVMoxR	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
			746.9	2	LVILGFPCNQFGK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Ophiophagus hannah* ETE68810	GPx	Glutathione peroxidase 3
49	0.042	>116 ▼	542.2	2	QCVDVKTAY	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			570.8	2	DSCFTLNQR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.042	116 ▼	542.2	2	QCVDVKTAY	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			570.8	2	DSCFTLNQR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	4.000	58 ▼	790.4	2	MNDNABXXTGXDFK	~ *Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.042	40 ▼	976.5	2	KYIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
	0.042	33 ▼	746.9	2	LVILGFPCNQFGK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
	0.042	28 ▼	746.9	2	LVILGFPCNQFGK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
			770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
50	0.011	>116 ▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin

	0.011	116 ▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.011	68 ▼	453.7	2	ADFLEGVR	*Bungarus fasciatus* AAC59905	AChE	Acetylcholinesterase
			896.9	2	DEGSYFLIYGLPGFSK	*Protobothrops mucrosquamatus* XP_015668657	AChE	Acetylcholinesterase
			713.0	3	AILQSGAPNAPWATVTPAESR	*Protobothrops mucrosquamatus* XP_015668657	AChE	Acetylcholinesterase
	0.002	58 ▼	557.3	2	ESFLFTLTR	*Naja kaouthia* S59517	SA	Serum albumin precursor
			770.9	2	DSVLAQYIFELSR	*Naja kaouthia* S59517	SA	Serum albumin precursor
	0.019	58 ▼	509.3	2	VTLLEASER	*Naja atra* A8QL58	LAO	L-amino-acid oxidase
			661.3	2	EIQALCYPSIK	*Naja atra* A8QL58	LAO	L-amino-acid oxidase
	0.954	50 ▼	542.2	2	QCVDVKTAY	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			651.3	3	KYIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			841.4	2	VYEIVNHLNMMYR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
	0.004	40 ▼	703.9	2	FTAYAINGPPVEK	*Ophiophagus hannah* ETE59578	PLB	Phospholipase B-like 1
			730.9	2	QVVPESLFAWER	*Ophiophagus hannah* ETE59578	PLB	Phospholipase B-like 1
			814.1	3	TTHQGLPESYNFDFVTMoxKPVL	*Ophiophagus hannah* ETE59578	PLB	Phospholipase B-like 1
			597.8	2	TPVPAGCYDSK	*Crotalus adamanteus* F8S101	PLB	Phospholipase B
	0.017	40 ▼	542.2	2	QCVDVKTAY	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin

			570.8	2	DSCFTLNQR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			651.3	3	KYIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			732.8	2	YGLDFSYEMoxAPR	*Drysdalia coronoides* F8J2D3	PLB	Phospholipase-B 81
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.011	33 ▼	542.2	2	QCVDVKTAY	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			651.3	3	KYIEFYVVVDNIMYR	*Naja atra* D3TTC2	SVMP	Zinc metalloproteinase-disintegrin-like atragin
			841.4	2	VYEIVNHLNMMYR	*Cryptophis nigrescens* ABQ01139	SVMP	Nigrescease-1
	0.011	28▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.010	21▼	570.8	2	DSCFTLNQR	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			648.8	2	DPSYGMoxVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
			542.2	2	QCVDVKTAY	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
	0.001	21▼	770.9	2	QEPGQNSEILQGIK	*Thamnophis elegans* AFS63888	GPx	Glutathione peroxidase 3
52	0.086	>116■	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			575.8	2	IHFAGEYTGR	*Naja atra* AVX27607	LAO	L-amino acid oxidase

			641.3	2	SASQLYQESLR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			448.9	3	KFWEADGIHGGK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
	0.016	>116▪	588.8	2	INYENALLAR	*Naja kaouthia* Q91132	CVF	Cobra venom factor
			681.3	2	ACETNVDYVYK	*Naja kaouthia* Q91132	CVF	Cobra venom factor
	0.102	96▪	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			575.8	2	IHFAGEYTGR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
	0.818	55▪	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			575.8	2	IHFAGEYTGR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			641.3	2	SASQLYQESLR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			661.3	2	EIQALCYPSIK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			745.7	3	HVVVVGAGMAGLSAAYVLAGAGHK	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
			592.8	2	FDEIVDGMDK	*Calloselasma rhodostoma* P81382	LAO	L-amino acid oxidase
			670.8	2	RFDEIVDGMoxDK	*Calloselasma rhodostoma* P81382	LAO	L-amino acid oxidase
53	0.016	>116▪	834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
	0.031	>116▪	517.3	2	VAVIIYLDK	*Naja kaouthia* Q91132	CVF	Cobra venom factor
			462.9	3	QLDIFVHDFPR	*Naja kaouthia* Q91132	CVF	Cobra venom factor
	0.012	>116▪	476.3	2	VGIIGYTTK	*Macrovipera lebetina* AHJ80886	5'-NT	5'-nucleotidase
			509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase
	0.016	55▪	509.3	2	VTLLEASER	*Naja atra* AVX27607	LAO	L-amino acid oxidase

			641.3	2	SASQLYQESLR	*Naja atra* AVX27607	LAO	L-amino acid oxidase
			834.7	3	SPLEECFREADYEEFLEIAR	*Bungarus multicinctus* A8QL51	LAO	L-amino acid oxidase
	0.132	55▪	742.8	2	DDCDFPELCTGR	*Thamnophis sirtalis* XP_013922279	SVMP	Hemorrhagic metalloproteinase-disintegrin-like kaouthiagin
	0.385	50▪	742.8	2	DDCDFPELCTGR	*Thamnophis sirtalis* XP_013922279	SVMP	Hemorrhagic metalloproteinase-disintegrin-like kaouthiagin
			570.8	2	DSCFTXNQR	~ *Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin

Table S2E: Proteins and protein families assigned to rp-HPLC fractions of *Hemachatus haemachatus* by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 5. In this case, only a limited number of peaks were subjected to MS/MS mass determination. The results from Pools 1 and 4 were used to infer protein families where peptide ions were not subjected to MS/MS mass determination. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ▪ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that point. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		Peptide sequence	Best NCBI match	Protein family	Best protein match
			m/z	z				
1	0.446	15 ▼					3FTx	
2	1.262	14 ▼	669.8	2	GIEINCCTTDR	Naja naja Q9PTT0	3FTx	Cobrotoxin homolog
			567.3	2	LTCLICPEK	*Naja atra* O93422	3FTx	Long neurotoxin homolog
3	2.895	15 ▼	990.1	3	LECHNQQSSQTPTTQTCPGETNCYK	*Hemachatus haemachatus* P01433	3FTx	Short neurotoxin 2 Toxin IV
4	0.443	14 ▼				*Hemachatus haemachatus* P01425	3FTx	Short neurotoxin 1_Toxin II
		7 ▼					KI	
5	2.726	14 ▼				*Hemachatus haemachatus* P01425	3FTx	Short neurotoxin 1_Toxin II
		7 ▼					KI	
7	0.186	14 ▼					3FTx	
9	0.370	13 ▼					3FTx	
		7 ▼					3FTx	
10	1.635	13 ▼	567.3	2	LTCLICPEK	*Naja atra* O93422	3FTx	Long neurotoxin homolog

			649.8	2	TWCDAWCGSR	*Micropechis ikaheca* AHZ08824	3FTx	Long neurotoxin
		7 ▼	567.3	2	LTCLICPEK	*Naja atra* O93422	3FTx	Long neurotoxin homolog
		5 ▼	567.3	2	LTCLICPEK	*Naja atra* O93422	3FTx	Long neurotoxin homolog
13	0.336	15 ▼					PLA2	
14	0.330	35 ▼	649.3	2	SAECPTDSFKR	*Pseudechis australis* ABQ01134	DC domain	Australease-1
		15 ▼					PLA2	
15	0.284	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.3	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		7 ▼	836.9	2	(322.2)NVXVECCNT(262.1)	~ *Naja sputatrix* Q9W7I3	3FTx	Weak neurotoxin 9
16	2.901	15 ▼					PLA2	
17	5.138	15 ▼					PLA2	
		7 ▼	546.3	3	GCTFTCPEXR(415.3)	~ *Naja atra* Q9DEQ3	3FTx	Neurotoxin homolog NL1
			532.9	2	SPSBYDVXR	~ *Ophiophagus hannah* Q2VBN4	3FTx	Beta-cardiotoxin CTX23
						~ *Naja mossambica* P01469		Cytotoxin V(II)2
18	0.146	7 ▼	546.3	3	GCTFTCPEXR(415.3)	~ *Naja atra* Q9DEQ3	3FTx	Neurotoxin homolog NL1
			532.9	2	SPSBYDVXR	~ *Ophiophagus hannah* Q2VBN4	3FTx	Beta-cardiotoxin CTX23
						~ *Naja mossambica* P01469		Cytotoxin V(II)2
19	0.627	29 ▼	648.8	2	SAECPTDSFKR	*Thamnophis sirtalis* XP_013929562	DC domain	Zinc metalloproteinase-disintegrin-like VLAIP-B
		16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1

			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		13 ▼	903.9	2	(258.2)PVDCPDGBNXCFK	~ *Dendroaspis angusticeps* Q8QGR0	3FTx	Muscarinic toxin 7
		7 ▼	836.9	2	(322.2)NVXVECCNT(262.1)	~ *Naja sputatrix* Q9W7I3	3FTx	Weak neurotoxin 9
			903.9	2	(258.2)PVDCPDGBNXCFK	~ *Dendroaspis angusticeps* Q8QGR0	3FTx	Muscarinic toxin 7
						~ *Naja mossambica* P01469	3FTx	Cytotoxin V(II)2
		5 ▼	836.9	2	(322.2)NVXVECCNT(262.1)	~ *Naja sputatrix* Q9W7I3	3FTx	Weak neurotoxin 9
20	2.061	37 ▼	594.8	2	GGSGTPVDDLDR	*Bungarus caeruleus* Q6SLM2	PLA2	Acidic phospholipase A2 1
		33 ▼	648.8	2	DPSYGMoxVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
		29 ▼	640.8	2	DPSYGMVEPGTK	*Naja kaouthia* Q9PVK7	SVMP	Zinc metalloproteinase-disintegrin-like cobrin
		16 ▼	699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.3	3	SWWHFANYGCYCGR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			564.3	2	ISGCRPYFK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		7 ▼	655.8	2	SSLLVNVMCCK	~ *Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
		5 ▼	620.9	2	GCTFTCPEXR	~ *Bungarus multicinctus* Q9W727	3FTx	Muscarinic toxin-like protein
		4 ▼	667.4	2	FVYGGCFGNANK	~ *Tropidechis carinatus* Q6ITB0	KI	Kunitz-type serine protease inhibitor carinatin-1
21	0.089	7 ▼	620.9	2	GCTFTCPEXR	~ *Naja atra* Q9DEQ3	3FTx	Neurotoxin homolog NL1
22	4.029	18 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			622.3	3	SWWHFADYGCYCGR	~ *Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor

			612.4	2	SYSYDCTEGK	~ *Micropechis ikaheca* AHZ08815	PLA2	Phospholipase A2
		8 ▼	585.4	2	TCTEEDTWK	~ *Walterinnesia aegyptia* C0HKZ8	3FTx	Actiflagelin
			500.8	2	(201.1)VSGCHXK	~ *Micrurus altirostris* AED89561	3FTx	Putative three finger toxin precursor
		6 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			949.4	3	SFHYNLAAQQCLQFIYGGCGGNANR	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
			643.3	2	FIYGGCGGNANR	*Naja naja* P20229	KI	Kunitz-type serine protease inhibitor
23	2.019	16 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
		15 ▼	949.4	3	SFHYNLAAQQCLQFIYGGCGGNANR	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
		15 ▼	754.7	3	GIDSSHWNSYCTETDTFIK	*Naja atra* P61898	vNGF	Venom nerve growth factor
		15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			682.9	2	SYCTETDTFXK	~ *Naja atra* P61898	vNGF	Venom nerve growth factor
			481.8	2	BYFFETK	~ *Ovophis okinavensis* BAN82142	vNGF	Nerve growth factor
		13 ▼	1055.6	2	(338.3)VSETCAEGENXCYXK	~ *Ophiophagus hannah* Q53B56	3FTx	Long neurotoxin OH-57
		5 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
24	2.130	16 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		15 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
		13 ▼	692.8	3	(486.4)TCCSTDNCNPHPK	~ *Naja melanoleuca* P01383	3FTx	Long neurotoxin 1
			566.9	2	PLVQVECCK	~ *Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
		7 ▼	598.3	3	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
		5 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
25	0.091	12 ▼	685.3	2	TCPEGQNICYK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB

			741.3	3	GCAATCPSERPLVQVECCK	*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog Toxin 9B/9BB
		5 ▼	896.9	2	RPDFCELPAETGLCK	*Hemachatus haemachatus* P00985	KI	Kunitz-type serine protease inhibitor 2
26	0.712	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		8 ▼	772.0	3	GCAATCPKPEAQVYVDCCAR	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						Hemachatus haemachatus C0HJT5	3FTx	Ringhalexin
		8 ▼	449.2	2	VLNVLSPR	~ *Pseudonaja textilis* XP_026558105	IGF	Insulin-like growth factor-binding protein 3 isoform X1
		6 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			772.0	3	GCAATCPKPEAQVYVDCCAR	*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
			402.4	2	LVPFLSK	~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
			599.9	2	(201.2)XTCC^{pro}STDSK	~ *Parasuta nigriceps* ACY68696	3FTx	Long-chain neurotoxin isoform 2
27	0.169	15 ▼	594.8	2	GGSGTPVDDLDR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			699.8	2	EGNNECAAFVCK	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
			832.7	3	LAAICFAGAHYNDNNNYIDLAR	*Hemachatus haemachatus* P00595	PLA2	Basic phospholipase A2 DE-1
		15 ▼	647.3	2	IETACVCVITK	*Naja atra* P61898	vNGF	Venom nerve growth factor
		6 ▼	772.0	3	GCAATCPKPEAQVYVDCCAR	*Hemachatus haemachatus* C0HJT5	3FTx	Ringhalexin
28	0.221	6▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
29	0.098	6▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1

30	12.795	14 ▼					3FTx	
		5 ▼				*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
						~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
32	9.833	22 ▼					3FTx	
		15 ▼					3FTx	
		5 ▼				~ *Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
34	0.103	5 ▼	655.8	2	SSLLVNVMCCK	*Hemachatus haemachatus* P24776	3FTx	Cytotoxin 2 Toxin 12A
			521.9	3	MSMoxEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~ *Naja naja* BAU24670	3FTx	~Cytotoxin 13
35	0.932	7 ▼					Vespryn	
		7 ▼					VEGF	
		7 ▼				*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
		5 ▼					3FTx	
		5 ▼					Vespryn	
37	10.414	14▼	757.4	3	ADVTFDSNTAFESLVVSPDKK	*Naja kaouthia* P82885	Vespryn	Thaicobrin
		14▼	704.4	2	MSMoxEVTPMoxIPIK	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
			521.9	3	MoxSMEVTPMoxIPIKR	*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
		5▼				*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
						~ *Naja annulifera* P01420	3FTx	~Short neurotoxin 3
40	2.128	28▼					CRISP	
		5▼				~ *Naja annulifera* P01420	3FTx	~Short neurotoxin 3

41	10.786	28▼					CRISP	
		5▼					3FTx	
44	0.743	63■					SVMP	
45	4.592	66▼					SVMP	
		26▼					CRISP	
		17▼					CRISP	
46	3.911	>116▼					SVMP	
		55▼					SVMP	
		45▼					SVMP	
		26▼					CRISP	
47	5.217	>116■					SVMP	
		66■					SVMP	
		50■					SVMP	
48	0.261	70 ▼					SVMP	
		58 ▼					SVMP	
52	0.362	>116▼					PDE	
		116▼					PDE	
		97▼					PDE	
		57▼					PDE	
		29▼					GPx	
54	4.857	>116▼					SVMP	
		116▼					SVMP	

		97▼					SVMP	
		97▼					SVSP	
		45▼					SVMP	
		33▼					CHIT1	
		33▼					SVMP	
		28▼					SVMP	
56	1.113	50▼					LAO	
		50▼					5'-NT	
		45▼					LAO	
57	0.610	54▼					SVMP	
		45▼					LAO	

Table S2F: Proteins and protein families assigned to rp-HPLC fractions of *Hemachatus haemachatus* by ESI-QUAD-TOF of selected peptide ions collected from SDS-PAGE gels for Pool 6. In this case, only a limited number of peaks were subjected to MS/MS mass determination. The results from Pools 1 and 4 were used to infer protein families where peptide ions were not subjected to MS/MS mass determination. Peak numbers correspond to peaks collected from the chromatogram and run on the SDS-PAGE gel. The % column refers to the percentage that each band in the SDS-PAGE gel contributes to the overall composition of each protein family. MW (kDA) refers to the molecular weight of each band on the SDS-PAGE gel; in this column values shown with a ▪ refer to bands excised from the gel run under non-reducing conditions, while values shown with a ▼ refer to bands excised from the gel run under reducing conditions. In the peptide ion column, m/z refers to mass per charge and z refers to the charge of the molecule. Sequences showing an amino acid followed by the letters "ox" indicate that the amino acid is oxidised at that point. The tilde (~) sign refers to sequences identified via de novo sequencing or proteins identified from the transcriptome, based on molecule weights calculated from ESI-MS analysis of the peaks.

Peak number	%	MW (kDa)	Peptide ion		Peptide sequence	Best NCBI match	Protein family	Best protein match
			m/z	z				
1	0.710	14 ▼					3FTx	
2	0.547	21 ▼	669.8	2	GIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
			489.6	3	KGIELNCCTTDR	*Naja mossambica* P01431	3FTx	Short neurotoxin 1
						~ *Naja atra* P80958	3FTx	Cobrotoxin-b
	2.189	14 ▼					3FTx	
3	0.178	9▼					3FTx	
	0.118	5 ▼					KI	
						Hemachatus haemachatus P01425	3FTx	Short neurotoxin 1_Toxin II
4	0.986	9▼					3FTx	
	0.247	5 ▼					KI	
7	0.141	14 ▼	567.3	2	LTCLICPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11
	0.563	8 ▼					3FTx	

8	0.589	7 ▼	567.3	2	XTCXXCPEK	*Naja haje haje* P01401	3FTx	Weak toxin CM-11
	0.091						3FTx	
12	0.315	14 ▼					PLA2	
13	0.082	16 ▼	622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
14	0.084	14 ▼	622.2	3	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.082	10 ▼	594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
			699.8	2	EGNNECAAFVCK	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
15	12.366	14 ▼					PLA2	
						~ *Hemachatus haemachatus* P0DQH2	3FTx	Exactin
	0.125	6 ▼					3FTx	
16	0.063	14 ▼					PLA2	
	0.047	6 ▼	726.5	2	GCTFTCPEXR(230.2)	~ *Bungarus multicinctus* Q9W727	3FTx	Muscarinic toxin-like protein
			814.5	2	(209.1)TFTCPEXRNGPK	~ *Bungarus multicinctus* Q9W727	3FTx	Muscarinic toxin-like protein
	0.047	4 ▼	667.4	2	FVYGGCFGNANK	~ *Demansia vestigiata* A6MFL4	KI	Kunitz-type serine protease inhibitor vestiginin-4
			1030.9	3	(1104.7)SNTCHSFVYSGCG(326.14)R	~ *Naja atra* Q5ZPJ7	KI	Kunitz-type serine protease inhibitor NACI

19	0.142	15 ▼	932.9	2	SWWHFADYGCYCGR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
			594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.035	12 ▼	594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
			699.8	2	EGNNECAAFVCK	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.212	9 ▼	594.8	2	GGSGTPVDDLDR	*Naja sputatrix* Q92084	PLA2	Neutral phospholipase A2 muscarinic inhibitor
	0.319	4 ▼	512.6	2	ITCSAEEKFCYK	*Walterinnesia aegyptia* C0HKZ8	3FTx	Actiflagelin
	0.078						3FTx	
21	0.450	15 ▼					PLA2	
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	0.450	8 ▼					3FTx	
	1.349	4 ▼					KI	
22	0.204	14 ▼					PLA2	
						Hemachatus haemachatus P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	0.305	8 ▼					3FTx	
25	15.982	8 ▼				*Hemachatus haemachatus* P24778	3FTx	Cytotoxin homolog_Toxin 9B/9BB
	0.092					*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
	0.222	6 ▼				*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
27	0.205	21 ▼					3FTx	
	0.019	21 ▼					PLA2	

	2.237	14 ▼					3FTx	
	1.119	12 ▼					3FTx	
	18.793	5 ▼				*Hemachatus haemachatus* P01471	3FTx	Cytotoxin 1
30	0.240	7 ▼					Vespryn	
	0.480	7 ▼					VEGF	
	0.509	7 ▼					3FTx	
	0.975	5 ▼				*Hemachatus haemachatus* P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.254	5 ▼					Vespryn	
33	0.210	28▼					CRISP	
						Hemachatus haemachatus P24777	3FTx	Cytotoxin 3 Toxin 11/11A
	0.057	14▼					Vespryn	
	1.354	14▼					3FTx	
	0.061	14▼					vNGF	
	19.354	5▼					3FTx	
35	2.003	5▼					3FTx	
36	0.234	63▪					SVMP	
	0.029	30▪					CRISP	
	0.029	8▪					3FTx	
37	0.806	66▼					SVMP	
	2.820	26▼					CRISP	
	0.403	17▼					CRISP	
38	0.021	>116▼					SVMP	

	0.021	116▼					SVMP	
	1.851	55▼					SVMP	
	0.041	45▼					SVMP	
	0.041	36▼					SVMP	
	0.041	34▼					SVMP	
	0.041	26▼					CRISP	
39	0.062	>116■					SVMP	
	0.554	66■					SVMP	
40	0.142	60▼	736.4	2	TAPAFQFSSCSIR	*Naja atra* D3TTC1	SVMP	Zinc metalloproteinase-disintegrin-like kaouthiagin-like
			694.9	2	YIEFYVVVDNK	*Tropidechis carinatus* ABQ01132	SVMP	Carinatease-1
	0.427	55▼					SVMP	
	0.142	35▼	708.9	2	YIEFYVVVDNR	*Demansia vestigiata* ABK63559	SVMP	Metalloproteinase precursor
			736.4	2	TAPAFQFSSCSIR	*Naja atra* D3TTC1	SVMP	Zinc metalloproteinase-disintegrin-like kaouthiagin-like
			725.8	2	DDCDLPEICTGR	*Micropechis ikaheca* AHZ08819	SVMP	P-III snake venom metalloprotease
42	0.197	116▼					PDE	
	0.197	97▼					PDE	
43	0.033	>116▼					SVMP	
	0.033	116▼					SVMP	
	0.024	97▼					SVMP	
	0.009	97▼					SVSP	
	2.820	56▼					SVMP	

	0.009	38▼					CHIT1	
	0.157	38▼					SVMP	
	0.033	32▼					SVMP	
	0.100	22▼					SVMP	
	0.100	12▼	542.2	2	KCVDVKTAY	~ *Pseudechis australis* ABQ01134	SVMP	Australease-1
45	0.799	64▼					LAO	
	0.178	64▼					5'-NT	
46	0.145	64▼					LAO	
	0.178	60▼					SVMP	