

Article

Exploring Toxin Genes of Myanmar Russell's Viper, *Daboia siamensis*, through De Novo Venom Gland Transcriptomics

Khin Than Yee ¹, Jason Macrander ², Olga Vasieva ^{3,4} and Ponlapat Rojnuckarin ^{5,*}¹ Department of Medical Research, Ministry of Health, Yangon 11191, Myanmar² Department of Biology, Florida Southern College, Lakeland, FL 33801, USA³ Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK⁴ BioSynthetic Machines, Inc., Chicago, IL 60062, USA⁵ Excellence Center in Translational Hematology, Division of Hematology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand

* Correspondence: rojnuckarinp@gmail.com or ponlapat.r@chula.ac.th

Abstract: The Russell's viper (*Daboia siamensis*) is a medically important venomous snake in Myanmar. Next-generation sequencing (NGS) shows potential to investigate the venom complexity, giving deeper insights into snakebite pathogenesis and possible drug discoveries. mRNA from venom gland tissue was extracted and sequenced on the Illumina HiSeq platform and de novo assembled by Trinity. The candidate toxin genes were identified via the Venomix pipeline. Protein sequences of identified toxin candidates were compared with the previously described venom proteins using Clustal Omega to assess the positional homology among candidates. Candidate venom transcripts were classified into 23 toxin gene families including 53 unique full-length transcripts. C-type lectins (CTLs) were the most highly expressed, followed by Kunitz-type serine protease inhibitors, disintegrins and Bradykinin potentiating peptide/C-type natriuretic peptide (BPP-CNP) precursors. Phospholipase A₂, snake venom serine proteases, metalloproteinases, vascular endothelial growth factors, L-amino acid oxidases and cysteine-rich secretory proteins were under-represented within the transcriptomes. Several isoforms of transcripts which had not been previously reported in this species were discovered and described. Myanmar Russell's viper venom glands displayed unique sex-specific transcriptome profiles which were correlated with clinical manifestation of envenoming. Our results show that NGS is a useful tool to comprehensively examine understudied venomous snakes.



Citation: Yee, K.T.; Macrander, J.; Vasieva, O.; Rojnuckarin, P. Exploring Toxin Genes of Myanmar Russell's Viper, *Daboia siamensis*, through De Novo Venom Gland Transcriptomics. *Toxins* **2023**, *15*, 309. <https://doi.org/10.3390/toxins15050309>

Received: 24 January 2023

Revised: 3 April 2023

Accepted: 20 April 2023

Published: 26 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: venom gland transcriptomes; Russell's viper; *Daboia siamensis*; next-generation sequencing; toxin genes; Venomix pipeline

Key Contribution: Myanmar Russell's viper venom glands showed unique sex-specific transcriptomes. Sequence data are correlated with clinical envenoming and being a reference for future proteomic and toxicogenomic research.

1. Introduction

Russell's viper is a widespread species found in South and Southeast Asia, as well as China. Based on previous molecular [1] and morphological [2] studies, the *Daboia* genus has been divided into two well-defined taxa: *D. russelii* (west of the Bengal Bay) and *D. siamensis* (east of the Bengal Bay). The species designation and regional distribution align with variations observed in the clinical features of Russell's viper envenomation [2,3]. The venom proteomes of Russell's viper vary across their geographical distribution and have been attributed to these differences [4–9].

Snake venoms are mixtures of proteins working synergistically to produce various toxic functions. These compounds evolved to immobilize, kill and initially digest their preys [10,11]. The myriad toxin compositions and functions are a byproduct of their interrelation between natural selection and biochemical processes shaping the protein–protein

or protein–cellular interactions. Venom variation is found not only across lineages, but also within individuals of the same species [12]. Additional factors influencing venom variation include age, sex, diet, season, frequency of extraction and a post-extraction process [13–15]. The ontogenetic change and venom gland location can also influence the venom composition in an individual snake [16].

Transcriptomic and proteomic approaches have been used to better understand venom diversity [17–19]. Transcriptomic studies using next-generation sequencing (NGS) techniques deliver a much broader view of the venom protein coding genes and a deeper understanding of toxin transcript abundance. Transcriptomic approaches permit the identification of unique isoforms that are candidates for drug discovery [20]. Using these techniques, multiple toxin gene copies have been identified within comprehensive datasets, aiding our understanding of how venom protein gene families have evolved. Each toxin family shows a unique evolutionary rate, which can be linked to a change in ecological opportunities, such as predation [21].

Proteomic approaches are often used to provide an expanded overview of venom composition and expression. Previous proteomic studies of Russell’s viper from different countries have shown that toxin gene families exhibit varying venom compositions throughout their distribution [4,6–8,22,23].

Despite the differences across their population observed in proteomic data, the Russell’s viper protein-coding toxin genes have not been previously investigated through a comprehensive transcriptomic analysis, nor have any Russell’s viper studies focused on sex-specific venom expression despite anecdotal evidence indicating size- or sex-specific envenomation symptoms in humans. The present study aims to generate a sex-specific de novo venom-gland transcriptomes of Myanmar Russell’s viper (*D. siamensis*) and provide a deeper insight into the correlation between the toxin composition and pathophysiology of Russell’s viper envenomation. The results of the current study can be used to validate diverse predicted toxins as well as identify novel toxin families in this species that are yet to be identified.

2. Results

2.1. De Novo Transcriptome Assembly

Sequencing of the cDNA libraries yielded a total of 78.64 and 74.44 million clean reads for male and female Myanmar Russell’s vipers, respectively (Table 1). De novo assembly using the Trinity [24] program created 88,325 contigs (N50 = 1870) comprising 77,668 Unigenes (N50 = 1397) for the male snakes and 50,858 contigs (N50 = 1379) comprising 46,884 Unigenes (N50 = 1183) for the female snakes.

Table 1. The sequencing and the assembly quality of the venom gland transcriptomes of Myanmar Russell’s vipers (*Daboia siamensis*).

Trinity Assembly Parameters	Male	Female
Total raw reads	82,251,232	79,578,046
Total clean reads	78,647,670	74,444,200
Total clean nucleotides (nt)	7,752,565,867	7,235,655,822
Q20 percentage	98.93%	98.55%
GC percentage	46.26%	53.43%
Contigs created	88,325	50,858
Total assembled bases	75,163,976	39,200,694
Mean length (nt)	850.99	770.79
N50	1870	1379
Unigenes/transcripts assembled (only longest isoform per gene)	77,668	46,884

2.2. Categorization of Transcripts and Gene Expression

The expression levels of toxin transcripts from both samples were compared with those of non-toxin transcripts from tissues of the same species (Bioproject PRJNA269317, Biosamples SAMN03081270) (Supplementary Tables S1 and S2).

The toxin transcripts were confirmed to be highly expressed in venom gland tissues. They were categorized according to TPM values. The individual toxin transcripts with a TPM value of more than 100 were categorized as highly expressed, a TPM value of 10 to 100 as moderately expressed and a TPM value of less than 10 as lowly expressed toxin groups. Only highly expressed groups were considered for further detailed evaluation and description in this study. Non-toxin full-length transcripts, such as glutamyl-peptide cyclotransferase (highly expressed), tomoregulin-2 (moderately expressed) and endothelin-converting enzyme (lowly expressed), were also identified in the current transcriptome. Highly expressed transcripts were no less than 150-fold higher in the venom gland when compared to samples comprising other tissues (brain, kidney, heart, lung, spleen and liver).

The highly expressed venom transcripts were categorized into 23 toxin gene families, including 53 unique full-length sequences. The most prevalent transcripts were C-type lectins (CTLs) (26%), Kunitz-type protease inhibitors (KSPIs) (19%), disintegrins (18%) and bradykinin potentiating peptide/C-type natriuretic precursors (BPP-CNPs) (16%). The remaining 21% consisted of phospholipase A₂ (PLA₂), serine proteases, metalloproteinases, vascular endothelial growth factor (VEGF), L-amino acid oxidase (LAAO) and cysteine-rich secretory proteins (CRISPs) (Figure 1).

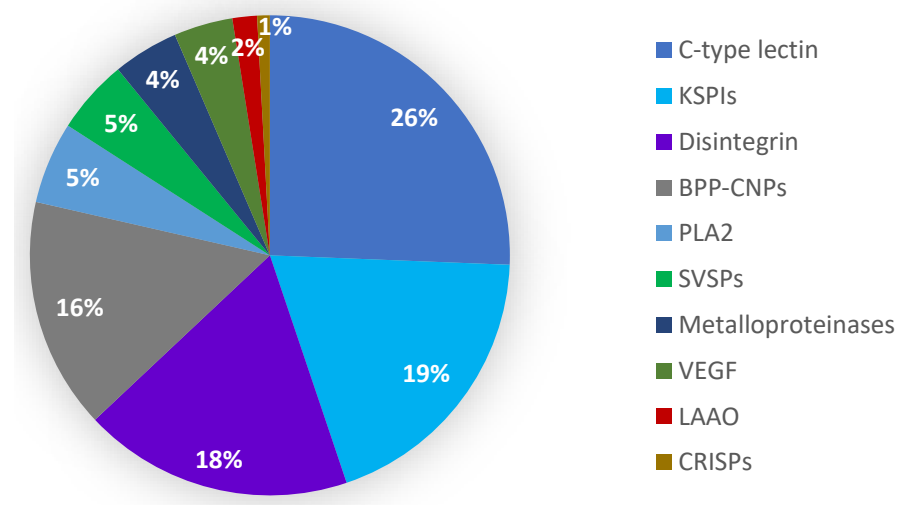


Figure 1. Overall toxin transcript (the top ten highly expressed toxin families) distribution in Myanmar Russell's viper venom gland.

When comparing males and females as shown in Figure 2, the male sample had CTLs, KSPIs, BPP-CNPs and disintegrins as the most highly expressed toxin groups, while metalloproteinases, CTLs, BPP-CNPs and KSPIs were the top highly expressed toxin groups in female snakes. The expression levels of transcripts for most male toxin groups were higher than those in the female-mapped reads, with 1.39- to 11.68-fold differences. However, the number of transcripts for serine proteases, glutamyl-peptide cyclotransferase-1 and metalloproteinases, were 1.16, 1.22 and 2.33 times higher in female than male venom glands, respectively. Some toxin transcripts, such as cathelicidin-related protein (Vipericidin) and techylectin-like protein (Veficolin), were only identified in male samples.

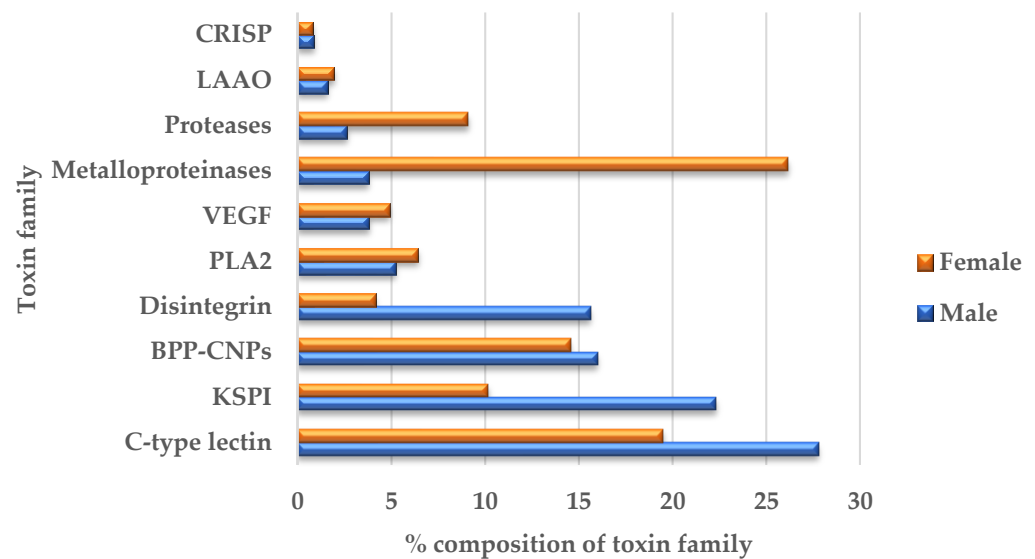


Figure 2. Comparison of toxin transcriptome compositions (the top ten highly expressed toxin families) of male and female snakes.

2.3. Snake Venom C-Type Lectins (CTLs) and Related Proteins

The present study revealed 23 CTL transcripts (4 full-length and 19 partial-length transcripts), including 7 different subtypes of CTLs. They are (I) newly identified *Bitis gabonica* C-type lectin 1 homolog, (II) *D. siamensis* p31 β subunit, (III) RVV-X LC1, (IV) *D. siamensis* p68 α subunit and *Microvipera lebetina* Snaclec A15 homolog, (V) Dabocetin α and *Microvipera lebetina* Snaclec A12 homolog, (VI) RVV-X LC2 and (VII) *D. siamensis* p68 β subunits and *D. siamensis* Snaclec-7 (Supplementary Table S3). Figure 3 shows the gene tree of these sequences. RVV-X CL1 and RVV-X CL2 linked to the metalloproteinase (RVV-X) to act on factor X of the coagulation cascade.

Our sequence alignment (Supplementary Figure S1) identified that these sequences contained six conserved cysteine residues (typically eight in total for each sequence), along with a WIGL conserved motif for CTL and the WND Gla-domain binding site.

2.4. Kunitz-Type Serine Protease Inhibitors

A total of 15 Kunitz-type serine protease inhibitor (KSPI) transcripts were identified, including seven full-length and eight partial-length transcripts. Thirteen unique isoforms of KSPIs were observed (Table 2).

Table 2. Kunitz-type protease inhibitor (KSPI) transcripts from Myanmar Russell's viper transcriptomes.

No.	Contig ID	Annotation	Accession No.	Species	Amino Acid	TPM Value
1	c17840_g2_i1_M_FL c11654_g1_i1_F_FL	Kunitz-type serine protease inhibitor 1	Q2ES50	<i>Daboia russelii</i>	84	109,755.22 16,745.60
2	c6152_g1_i1_M_FL c6152_g1_i2_M_FL	Serine protease inhibitor 1	A0A6G5ZVV7	<i>Vipera anatolica senliki</i>	148	1.59 0.76
3	c3688_g1_i1_M_FL c5062_g2_i1_F_FL	Serine protease inhibitor 4	A0A6G5ZUI5	<i>Vipera anatolica senliki</i>	514	26.46 9.09
4	c12297_g1_i1_M_FL	Tissue factor pathway inhibitor	U3TBN9	<i>Protobothrops flavoviridis</i>	303	2.45
5	c1956_g1_i1_M_PL	Serine protease inhibitor 3	A0A6G5ZW53	<i>Vipera anatolica senliki</i>	103	1.48
6	c12495_g1_i1_M_PL c12495_g1_i2_M_PL c12495_g1_i3_M_PL c2247_g1_i2_F_PL c2247_g1_i1_F_PL	Serine protease inhibitor 6	A0A6G5ZUI2	<i>Vipera anatolica senliki</i>	272	56.61 50.11 74.49 51.7 33.08
7	c15981_g1_i1_M_PL	kunitz/BPTI-like toxin	A0A6P9AX37	<i>Protobothrops guttatus</i>	139	47.25
8	c11558_g3_i1_F_PL	kunitz/BPTI-like toxin	XP_015682430.1	<i>Protobothrops mucrosquamatus</i>	139	25.45

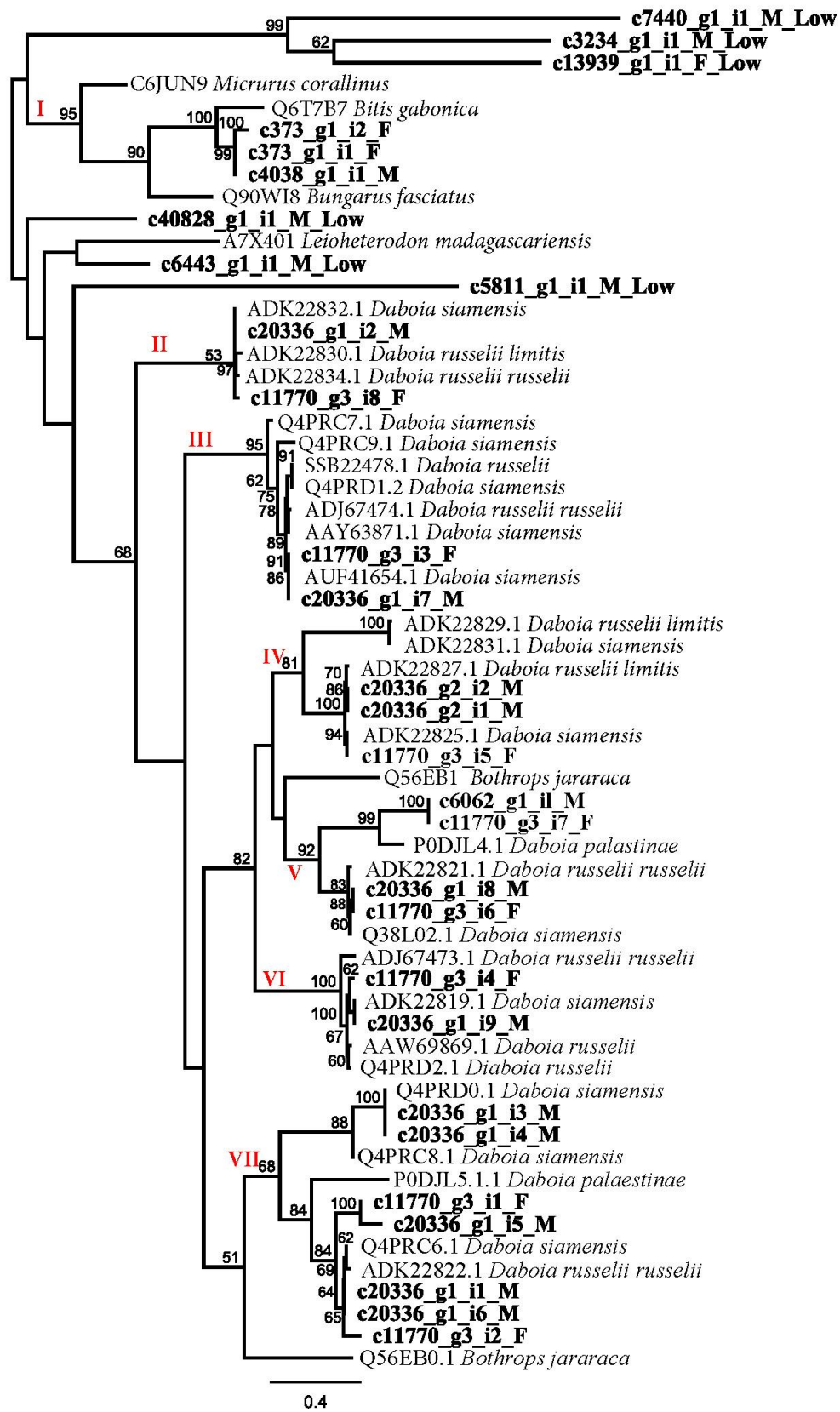


Figure 3. The maximum likelihood C-type lectins (CTLs) gene tree including *D. siamensis* with CTLs from other species. Bootstrap values greater than 50 are shown. Subtypes based on sequence alignments are indicated in Roman numerals. Refer to Supplementary Table S3 for highly expressed transcript TPM values.

2.5. RTS-Disintegrins

A total of three full-length disintegrin transcripts were identified, including two isoforms of RTS-containing disintegrins which were also described as Dis1a and Dis1b in our previous paper [25]. In the gene tree analysis, Dis1a shared a high sequence similarity with jerdostatin from *Protobothrops jerdonii* (Q7ZZM2), while Dis1b was in the same branch as acostatin- α from *D. siamensis* (AUF416558) (Supplementary Figure S2). In the sequence alignment, the N-terminal sequence of Dis1b was more similar to that of acostatin- α than jerdostatin and it possessed an RTS motif instead of an RGD motif (Figure 4).



Figure 4. The peptide sequence alignment of Myanmar Russell’s viper disintegrin with acostatin- α from *D. siamensis* and jerdostatin from *Protobothrops jerdonii*. Signal peptides are underlined, while cysteine, RGD and RTS binding motifs are highlighted in blue, yellow and green correspondingly. (*) fully conserved residues; (:) strongly similar residues; (.) weakly similar residues.

2.6. Bradykinin-Potentiating Peptides and C-Type Natriuretic Peptide (BPP-CNP) Precursors

A total of four full-length BPP-CNP transcripts were identified in this study. In Figure 5, the repeated BPP units were present in a BPP-CNP precursor. These BPP units appeared in tripeptide metalloproteinase inhibitors. CNPs were found to be variable between these sequences. The sequence alignments of the natriuretic peptide domains of Myanmar Russell’s viper BPP-CNP precursors with those of human and other snakes are shown in Figure 6.

2.7. Phospholipase A₂ (PLA₂)

In the focal transcriptomes, there were five full-length and two partial-length PLA₂ candidates. Five full-length transcripts contain three group II PLA₂ and two group III PLA₂ transcripts. Three of the group II PLA₂s (acidic and basic PLA₂) were highly expressed (TPM >10,000–14,000), followed by two group III PLA₂ transcripts (TPM 14–35), whereas two partial-length transcripts, group IIF PLA₂ homologs of *Crotalus atrox* and a crotoxin homolog from *Protobothrops guttatus*, were the lowest expressed transcripts (TPM 1–8). The highly expressed group II PLA₂ transcripts were aligned with PLA₂ from Eastern Russell’s viper PLA₂s, DsM b1and daboiatoxin A from Myanmar, PLA₂-I from China and RV-7 from Taiwan (Figure 7).

SVMI_ECHOC	-----	0
VIPAA_MPi_3	-----	0
VIPAA_ETMI	-----	0
c19711_g10_i1_M_FL	-----	0
c10992_g1_i1_F_FL	MPSVRHLQVADIRRGRRRCRCERLRPGRQRRLDEALPRQSLRLGSARLGSARRSAPSRIL	60
c19711_g10_i2_M_FL	-----	0
c10992_g1_i2_F_FL	MPSVRHLQVADIRRGRRRCRCERLRPGRQRRLDEALPRQSLRLGSARLGSARRSAPSRIL	60
SVMI_ECHOC	-----MFVSRLAASGLLLLSLLALS LDGKPLPQRQPHHIQPM <u>QKWLAPDA</u>	46
VIPAA_MPi_3	-----MFVSRLAASGLLLLSLLALS LDGKPL -----K <u>QKWL</u> GPEI	35
VIPAA_ETMI	-----MFVSRLAASGLLLLSLLALS LDGKPL -----K <u>QKWL</u> GPEI	35
c19711_g10_i1_M_FL	-----MSVARLAASGLLLLSLLALS LDGKPL -----E <u>QKWL</u> GPEI	35
c10992_g1_i1_F_FL	SLSPVALPAPAHPGMSVARLAASGLLLLSLLALS LDGKPL -----E <u>QKWL</u> GPEI	109
c19711_g10_i2_M_FL	-----MSVARLAASGLLLLSLLALS LDGKPL -----E <u>QKWL</u> GPEI	35
c10992_g1_i2_F_FL	SLSPVALPAPAHPGMSVARLAASGLLLLSLLALS LDGKPL -----E <u>QKWL</u> GPEI	109
	* :***** :*:*.*:	
SVMI_ECHOC	<u>PPLEQKWLAPDAPPEQKWLAPAAAPPEQKWLAPDAPPEQKWLAPDAPPEQKWLAPDA</u>	106
VIPAA_MPi_3	<u>PPMEQKWLGPPIPPLEQRWPGPKVPPLEQKWL</u> -----	67
VIPAA_ETMI	<u>PPMEQKWLGPPIPPLEQRWPGPKVPPLEQKWLGPPIPPLEQKWLGPPIPPLEQRWPGPKV</u>	95
c19711_g10_i1_M_FL	-----	35
c10992_g1_i1_F_FL	-----	109
c19711_g10_i2_M_FL	-----	35
c10992_g1_i2_F_FL	-----	109
SVMI_ECHOC	<u>PPMEQKWLAPDAPPEQKWLAPDAAPEQKWLAPDAPPEQKWLAPDAPPEQKWLQPIQ</u>	166
VIPAA_MPi_3	-----GPEIPPLE <u>QKWLGPPIPPLEQRWPGPKVPPLEQKWRGP</u> -----LQPEGP	111
VIPAA_ETMI	<u>PPLEQKWLGPPIPPLEQKWLGPPIPPLEQRWPGPKVPPLEQKWRGP</u> -----LQPEGP	147
c19711_g10_i1_M_FL	-----PPLE <u>QRWRGP</u> -----LQPEGP	51
c10992_g1_i1_F_FL	-----PPLE <u>QRWRGP</u> -----LQPEGP	125
c19711_g10_i2_M_FL	-----PPLE <u>QRWRGP</u> -----LQPEGP	51
c10992_g1_i2_F_FL	-----PPLE <u>QRWRGP</u> -----LQPEGP	125
	**:*:*.* **:*	
SVMI_ECHOC	SLMEQRQLSSGGTTALRQELSRAEAASGPAVVGGGGGGGSSKAALALPKPKKAGAAA	226
VIPAA_MPi_3	<u>PLMEPHEL</u> SAGGTTALREELSRAEAASA-----	140
VIPAA_ETMI	<u>PLMEPHEL</u> SAGGTTALREELSRAEAASA-----	176
c19711_g10_i1_M_FL	<u>PLMEPHEL</u> SAGGTTALREEPSRAEAAQHPG---G---GGGS-----	87
c10992_g1_i1_F_FL	<u>PLMEPHEL</u> SAGGTTALREEPSRAEAAQHPG---G---GGGS-----	161
c19711_g10_i2_M_FL	<u>PLMEPHEL</u> SAGGTTALREEPSRAEAASGPA AAAAGG ---GGSSKAALAVPKPKKAGASA	109
c10992_g1_i2_F_FL	<u>PLMEPHEL</u> SAGGTTALREEPSRAEAASGPA AAAAGG ---GGSSKAALVVPKPKKAGASA	183
	*** :*:*****.* *****.	

Figure 5. Cont.

SVMI_ECHOC	ATSRLMRDLRDPGKQASQKWGRLVDHDDHHHHHPGSSVGGGGGGGGGARRLKGLAKK	286
VIPAA_MPi_3	-----SLRTDGEQARQKWARMVHHEQQ-----RGGGGGWRADNEPPKK	179
VIPAA_ETMI	-----SLRTDGEQARQKWARMVHHEQQ-----RGGGGGWRADNEPPKK	215
c19711_g10_i1_M_FL	-----GNQPP--	92
c10992_g1_i1_F_FL	-----GNQPP--	166
c19711_g10_i2_M_FL	-ASRLMRDLRTDGKQARQKWGRMAHHDQQ-----RGGGGGWRGDNEPPNK	154
c10992_g1_i2_F_FL	-ASRLMRDLRTDGKQARQKWGRMAHHDQQ-----RGGGGGWRGDNEPPNK	228
:		
SVMI_ECHOC	<u>G</u> VAKGCFGLKLDRI <u>G</u> SMSG <u>L</u> GC-----	308
VIPAA_MPi_3	<u>V</u> PPNSCFGHKIDRI <u>G</u> SVSGLGCNKFKPGH-----	208
VIPAA_ETMI	<u>V</u> PPNSCFGHKIDRI <u>G</u> SVSGLGCNKFKPGH-----	244
c19711_g10_i1_M_FL	-NGNGCFGHKIDRVN-HNGFGCNKVDDHKQGGD*	124
c10992_g1_i1_F_FL	-NGNGCFGHKIDRVN-HNGFGCNKVDDHKQGGD*	198
c19711_g10_i2_M_FL	<u>R</u> PPDGCFGHKIDRI <u>G</u> AHSG <u>L</u> GCNKVDDRKQGGP*	188
c10992_g1_i2_F_FL	<u>R</u> PPDGCFGHKIDRVN-HNGFGCNKVDDHKQGGD*	261

Figure 5. Sequence comparison of BPP-CNP precursors from Myanmar Russell’s viper with those of snake venom metalloprotease inhibitor (A8YPR6) from *Echis ocellatus* (SVMI_ECHOC) and MPi-3 (A0A6B7FNM6) (VIPAA_MPi-3) and endogenous tripeptide metalloproteinase inhibitor (A0A1I9KNP8) from *Vipera ammodytes ammodytes* aligned by Clustal Omega. Signal peptides are underlined. BPPs are shown by wavy underlines. C-type natriuretic peptides are highlighted in gray. Metalloproteinase tripeptide inhibitors, QKW and QRW are highlighted in yellow and green, respectively. (*) fully conserved residues; (:) strongly similar residues; (.) weakly similar residues.

c19711_g10_i1_M	-----GC <u>F</u> GHKIDRVN-HNG <u>F</u> GCNKVDDHKQGGD*--	29
c10992_g1_i1_F	-----GC <u>F</u> GHKIDRVN-HNG <u>F</u> GCNKVDDHKQGGD*--	29
c10992_g1_i2_F	-----GC <u>F</u> GHKIDRVN-HNG <u>F</u> GCNKVDDHKQGGD*--	29
c19711_g10_i2_M	-----GC <u>F</u> GHKIDRI <u>G</u> AHSG <u>L</u> GCNKVDDRKQGGP*--	30
DENAN_DNP	---EVKYDPC <u>F</u> GHKIDRI <u>N</u> HVS <u>N</u> LG <u>C</u> PSLRDPRPNAPSTSA	38
hANP	---SLRRSSC <u>F</u> GGRM <u>D</u> RI <u>G</u> AQ <u>S</u> GL <u>G</u> CNSFR-----	27
hBNP	SPKMVQSGC <u>F</u> GRKM <u>D</u> RI <u>S</u> SS <u>S</u> GL <u>G</u> CKVLRH-----	32
LACMU_CNP	----RVGDGC <u>F</u> GLKL <u>D</u> RI <u>G</u> SMSG <u>L</u> GC-----	22
CROAT_CNP	-AKKRAGNGC <u>F</u> GLKL <u>D</u> RI <u>G</u> SMSG <u>L</u> GC-----	25
hCNP	----GLSKGC <u>F</u> GLKL <u>D</u> RI <u>G</u> SMSG <u>L</u> GC-----	22
BOTJA_CNP	----GAAKGC <u>F</u> GLKL <u>D</u> RI <u>G</u> TMSG <u>L</u> GC-----	22
*** ::*:. .:.*		

Figure 6. Sequence comparison of human natriuretic peptides (hANP (P01160), hBNP(P16860) and hCNP (P23582)) with those of different snake venoms (DENAN_DNP (P28374) from *Dentroaspis angusticeps*, LACMU_CNP (Q27J49) from *Lachesis muta muta*, CROAT_CNP (P0CV87) from *Crotalus atrox* and BOTJA_CNP(Q6LEM5) from *Bothrops jararaca*) aligned by Clustal Omega. Conserved cysteine residues are underlined, while relatively conserved amino acid residues known for NPR binding are highlighted. (*) fully conserved residues; (:) strongly similar residues; (.) weakly similar residues. Important residues for NP receptor binding are highlighted in yellow color with variable residues are in green and pink color.

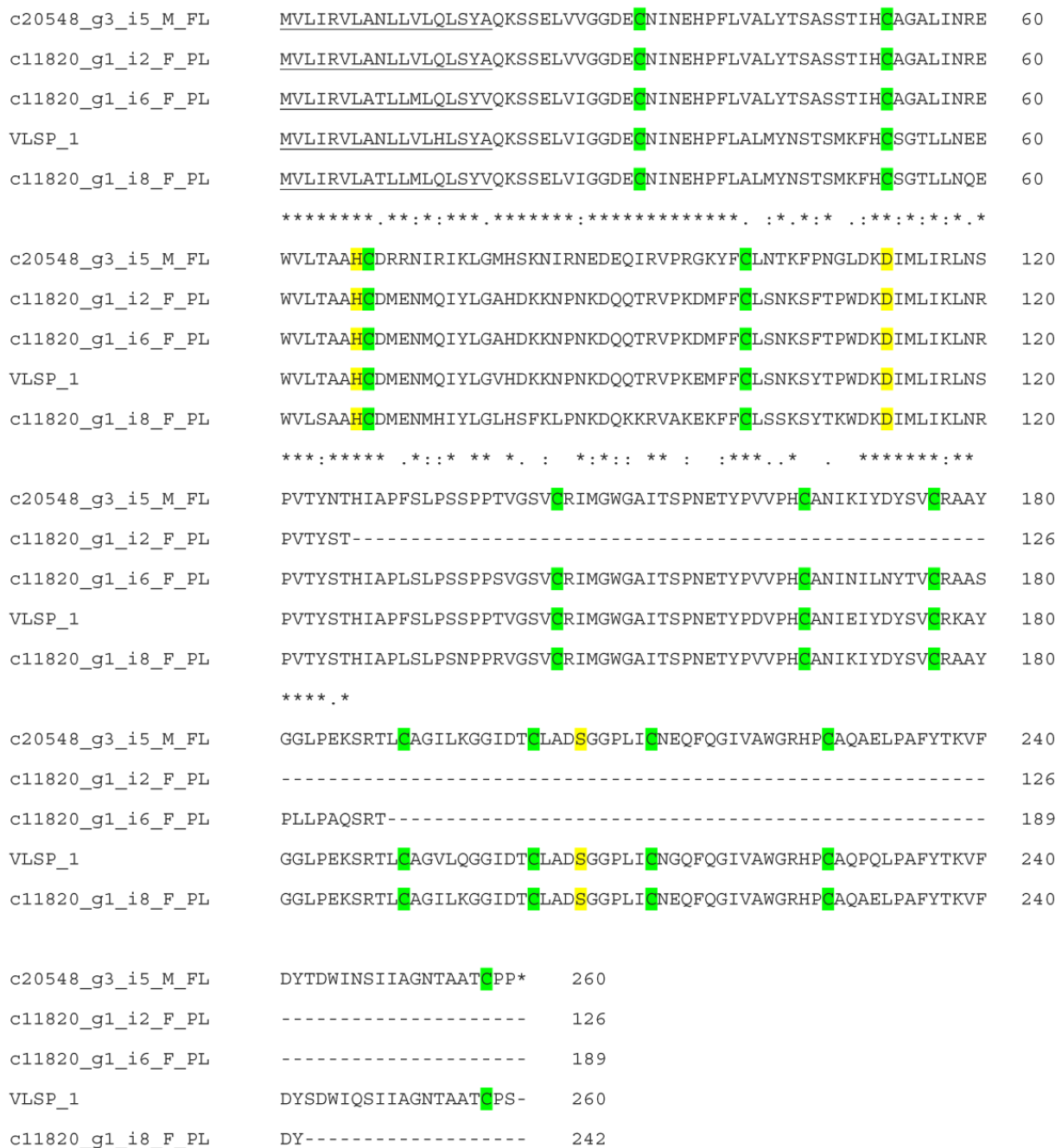


Figure 8. Sequence alignment of VLSP-1 from *Macrovipera lebetinus* and deduced amino acid sequences of transcripts from the transcriptomes. Signal peptide is single-underlined. Catalytic triad His57, Asp102 and Ser195 are highlighted in yellow. Cysteine residues in six disulfide bridges are highlighted in green. (*) fully conserved residues; (:) strongly similar residues; (.) weakly similar residues.

Table 3. Transcripts with snake venom serine protease (SVSP) annotation.

No.	Contig ID	Annotation	Accession No.	Species	Amino Acid	TPM Value
1	c20548_g3_i4_M_FL	Factor V activator RVV-V γ	P18965	<i>Daboia siamensis</i>	259	5495.65
2	c11820_g1_i3_F_PL	Thrombin-like_enzyme_TLBM (Factor V activator RVV-V γ)	P18965	<i>D. siamensis</i>	251	2205.37
3	c11820_g1_i1_F_PL	Thrombin-like_enzyme_TLBM (Factor V activator RVV-V γ)	P18965	<i>D. siamensis</i>	241	1794.30

Table 3. Cont.

No.	Contig ID	Annotation	Accession No.	Species	Amino Acid	TPM Value
4	c11820_g1_i4_F_PL	Thrombin-like_enzyme_TLBm (Factor V activator RVV-Vγ)	P18965	<i>D. siamensis</i>	251	1355.44
5	c38983_g1_i1_F_PL	α-fibrinogenase-like	E5L0E3	<i>D. siamensis</i>	105	3768.67
6	c20548_g3_i1_M_PL	α-fibrinogenase-like	E5L0E3	<i>D. siamensis</i>	131	3196.39
7	c20548_g3_i3_M_PL	α-fibrinogenase-like	E5L0E3	<i>D. siamensis</i>	155	952.41
8	c11820_g1_i9_F_PL	Thrombin-like_enzyme_catroboxin-1 (β-fibrinogenase-like)	E5L0E4	<i>D. siamensis</i>	126	981.63
9	c20548_g3_i5_M_FL	Serine protease VLSP-1	E0Y418	<i>Microvipera lebetina</i>	265	3027.55
10	c11820_g1_i8_F_PL	Thrombin-like_enzyme_TLBm (Serine protease VLSP-1)	E0Y418	<i>M. lebetina</i>	242	1400.97
11	c11820_g1_i6_F_PL	Thrombin-like_enzyme_TLBm (Serine protease VLSP-1)	E0Y418	<i>M. lebetina</i>	189	816.73
12	c11820_g1_i2_F_PL	Thrombin-like_enzyme_TLBm (Serine protease VLSP-1)	E0Y418	<i>M. lebetina</i>	126	448.44
13	c11820_g1_i5_F_PL	Thrombin-like_enzyme_TLBm (Serine protease VLSP-3)	E0Y420	<i>M. lebetina</i>	134	1318.30
14	c11820_g1_i7_F_PL	Thrombin-like_enzyme_catroboxin-1 (β-fibrinogenase-like)	E5L0E4	<i>D. siamensis</i>	104	919.53

2.9. Snake Venom Metalloproteinases (SVMPs)

In the current transcriptomes, two full-length transcripts for RVV-X heavy chain (Russell’s viper factor X activator, A0A2H4Z2W1) and two partial-length DSAIP (*Daboia siamensis* VLAIP-like SVMP, A0A2H4Z2Y9) are the most highly expressed SVMP transcripts (>1.0–1.8 × 10⁴ TPM). Of the remaining three full-length transcripts with TPM values of 15–47, c15329_g2_i1_M_FL and c10072_g2_i1_F_FL were homologous to the A disintegrin and metalloproteinase domain (ADAM)-containing protein 10 from *Vipera anatolica senliki* (A0A6G5ZVR9) (Figure 9), and c19271_g2_i1_M_FL was homologous to SVMP from *Agkistrodon contortrix contortrix* (A0A1W7RJN7) (Figure 10). All conceptually translated proteins contained the signal peptide, a metalloproteinase domain containing a Zn⁺-binding catalytic site and a disintegrin domain. There was an additional ADAM17 MPD domain containing a CxxC motif in the c19271_g2_i1_M_FL transcript.

2.10. Snake Venom Vascular Endothelial Growth Factors (svVEGFs)

In the current study, seven full-length VEGF transcripts were identified including three isoforms which comprised 136, 148 and 192 amino acids. Two transcripts were homologous to VR-1 (*D. russelii*) (P67861), as shown in Supplementary Figure S4. Two transcripts were related to VEGF1 (*Vipera anatolica senliki*) (QHR82857) and VEGFA-VIPAA (*Vipera ammodytes ammodytes*) (C0K3N5), as shown in Figure 11A,B.

VIPAN	MVLLRVLVLLCSWAAGAGGQYGNPLNKYIRHYEGLSYDVDSLHQKHQRAKRAVSHEDQFL	60
c15329_g2_i1_M_FL	MVLLRVLVLLCSWAAGAGGQYGNPLNKYIRHYEGLSYDVDSLHQKHQRAKRAVSHEDQFL	60
c10072_g2_i1_F_FL	MVLLRVLVLLCSWAAGAGGQYGNPLNKYIRHYEGLSYDVDSLHQKHQRAKRAVSHEDQFL	60

Figure 9. Cont.

VIPAN	RLDFHAHGRHFNLRLKREMSLFSDDFKLEMSDKVFDYDTSHIYTGHIYGEQGSFSGSVV	120
c15329_g2_i1_M_FL	RLDFHAHGRHFNLRLKREMSLFSDDFKLEMSDKVFDYDTSHIYTGHIYGEQGSFSGSVV	120
c10072_g2_i1_F_FL	RLDFHAHGRHFNLRLKREMSLFSDDFKLEMSDKVFDYDTSHIYTGHIYGEQGSFSGSVV *****	120
VIPAN	DGRFEGFVKTHGTGYIEPAERYIKDKALPFHSVMYHEDDIKYPHKYGPQGGCADHAVFE	180
c15329_g2_i1_M_FL	DGRFEGFVKTHGTGYIEPAERYIKDKALPFHSVMYHEDDIKYPHKYGPQGGCADHAVFE	180
c10072_g2_i1_F_FL	DGRFEGFVKTHGTGYIEPAERYIKDKALPFHSVMYHEDDIKYPHKYGPQGGCADHAVFE *****	180
VIPAN	RMRKYQMARGAEEPRHQQPSEEQIRDEPELLRKKRAAQAEKNTCQLYIQTDLHFFRYYG	240
c15329_g2_i1_M_FL	RMRKYQLARGAEEPRHQQPSEEQIRDEPELSRKKRAAQAEKNTCQLYIQTDLHFFRYYG	240
c10072_g2_i1_F_FL	RMRKYQLARGAEEPRHQQPSEEQIRDEPELSRKKRAAQAEKNTCQLYIQTDLHFFRYYG *****;*****	240
VIPAN	<u>TREAVIAQISSHVKAIDTIYQSTDFSGIRNISFMVKRIRINTTODEKDPSPNFRFPNIGV</u>	300
c15329_g2_i1_M_FL	<u>TREAVIAQISSHVKAIDTIYQSTDFSGIRNISFMVKRIRINTTODEKDPSPNFRFPNIGV</u>	300
c10072_g2_i1_F_FL	<u>TREAVIAQISSHVKAIDTIYQSTDFSGIRNISFMVKRIRINTTODEKDPSPNFRFPNIGV</u> *****	300
	ZnMc_TACE_like domain	
VIPAN	<u>EKFLELNSEQNHDDYCLAYVFTDRDFDDGVLGLAWVGAPSGSSGGICEKSKLYSDGKKKS</u>	360
c15329_g2_i1_M_FL	<u>EKFLELNSEQNHDDYCLAYVFTDRDFDDGVLGLAWVGAPSGSSGGICEKSKLYSDGKKKS</u>	360
c10072_g2_i1_F_FL	<u>EKFLELNSEQNHDDYCLAYVFTDRDFDDGVLGLAWVGAPSGSSGGICEKSKLYSDGKKKS</u> *****	360
VIPAN	<u>LNTGIITVQNYGSHVPPKVSHTFAHEVGHNFQSPHDSGTECTPGESKNLGQKENGNYIM</u>	420
c15329_g2_i1_M_FL	<u>LNTGIITVQNYGSHVPPKVSHTFAHEVGHNFQSPHDSGTECTPGESRNLGQKENGNYIM</u>	420
c10072_g2_i1_F_FL	<u>LNTGIITVQNYGSHVPPKVSHTFAHEVGHNFQSPHDSGTECTPGESRNLGQKENGNYIM</u> *****;*****	420
VIPAN	<u>YARATSGDKLNNKFSICSIRNISQVLEKKRNCFVESGQPICGNGLVEQGEQCDCGYSD</u>	480
c15329_g2_i1_M_FL	<u>YARATSGDKLNNKFSICSIRNISQVLEKKRNCFVESGQPICGNGLVEQGEQCDCGYSD</u>	480
c10072_g2_i1_F_FL	<u>YARATSGDKLNNKFSICSIRNISQVLEKKRNCFVESGQPICGNGLVEQGEQCDCGYSD</u> *****	480
	Disintegrin domain	
VIPAN	<u>QCKDECCYDANQPEDKCKLRPGKSCSPSEGPCCTQNCPEFKKEREKCRNDSEAEEGRCD</u>	540
c15329_g2_i1_M_FL	<u>QCKDECCYDANQPEDKCKLRPGKSCSPSEGPCCTQNCPEFKKEREKCRNDSEAEEGRCD</u>	540
c10072_g2_i1_F_FL	<u>QCKDECCYDANQPEDKCKLRPGKSCSPSEGPCCTQNCPEFKKEREKCRNDSEAEEGRCD</u> *****;*****	540

Figure 9. Cont.

VIPAN	<u>GNSSF</u> <u>PTSTAKKNL</u> <u>TE</u> <u>NRHTQV</u> <u>ISGQ</u> <u>AGSI</u> <u>EKFNLEE</u> <u>TT</u> <u>ASSDAKDDKEL</u> <u>CHV</u>	600
c15329_g2_i1_M_FL	<u>GNSSF</u> <u>PTSTAKKNL</u> <u>TE</u> <u>NRHTQV</u> <u>ISGQ</u> <u>AGSI</u> <u>EKFDLEE</u> <u>TT</u> <u>ASSDAKDDKEL</u> <u>CHV</u>	600
c10072_g2_i1_F_FL	<u>GNSSF</u> <u>PTSTAKKNL</u> <u>TE</u> <u>NRHTQV</u> <u>ISGQ</u> <u>AGSI</u> <u>EKFDLEE</u> <u>TT</u> <u>ASSDAKDDKEL</u> <u>CHV</u>	600
	*****;*****	
VIPAN	CMKKMHPETCASTGSEAWKAHFNYKTITLQPGSPCNDFKGYCDVFMRCRLVDADGPLARL	660
c15329_g2_i1_M_FL	CMKKMHPETCASTGSEAWKAHFNYKTITLQPGSPCNDFKGYCDVFMRCRLVDADGPLARL	660
c10072_g2_i1_F_FL	CMKKMHPETCASTGSEAWKAHFNYKTITLQPGSPCNDFKGYCDVFMRCRLVDADGPLARL	660

VIPAN	KKAI FNPELYENIAEWIVAHWWAVLLMGIALIMLMAGFIKICSVHTPSSNPKLPPPKPLP	720
c15329_g2_i1_M_FL	KKAI FNPELYENIAEWIVAHWWAVLLMGIALIMLMAGFIKICSVHTPSSNPKLPPPKPLP	720
c10072_g2_i1_F_FL	KKAI FNPELYENIAEWIVAHWWAVLLMGIALIMLMAGFIKICSVHTPSSNPKLPPPKPLP	720

VIPAN	GTLKRRRPPPTTQAPPRQRPRESYQMGHMRR-	751
c15329_g2_i1_M_FL	GTLKRRRPPPLTTQAPPRQRPRESYQMGHMRR*	751
c10072_g2_i1_F_FL	GTLKRRRPPPTTQAPPRQRPRESYQMGHMRR*	751

Figure 9. Sequence alignment of VIPAN, A disintegrin and metalloproteinase domain (ADAM)-containing protein 10 from *Vipera anatolica senliki* (A0A6G5ZVR9) and deduced amino acid sequences of c15329_g2_i1_M_FL and c10072_g2_i1_F_FL. Signal peptide is single-underlined. Zinc dependent metalloproteinase domain is double-underlined. Disintegrin domain is dotted-underlined. Cysteine residues are highlighted in green. Zinc binding site and RCD sequence are highlighted in yellow. (*) fully conserved residues; (;) strongly similar residues.

AGKCO	<u>MQPQSR</u> <u>SGRRL</u> <u>LP</u> <u>AWFS</u> <u>LLLLLLLL</u> <u>L</u> <u>ALGVS</u> <u>ALR</u> <u>PRGGH</u> <u>RES</u> <u>R</u> <u>LER</u> <u>LES</u> <u>SV</u> <u>LD</u> <u>YD</u> <u>IL</u> <u>SV</u> <u>SN</u>	60
c19271_g2_i1_M_FL	<u>MQPQSR</u> <u>SGRRL</u> <u>LP</u> <u>ALF</u> <u>LL</u> <u>-</u> <u>LLLL</u> <u>ALGVS</u> <u>ALR</u> <u>PRGGH</u> <u>RES</u> <u>R</u> <u>FER</u> <u>LES</u> <u>SV</u> <u>LD</u> <u>YD</u> <u>IL</u> <u>SV</u> <u>SN</u>	58
	***** * * *****;*****	
AGKCO	IQQHSVRRRDLPESHIEKLISFSALHRHFQLYLTSTADHFSKNFQALILDGQGNERYH	120
c19271_g2_i1_M_FL	IQQHSVRRRDLPESHIEKLISFSALHRHFQLYLRSTADHFSKNFQALILDGQGNERYEP	118
	***** ***** *	
AGKCO	VQWQEFFTGHI VGHNSKVI AHIGEGDFTVRINTDAEEYNVEPLWRF LNNTEDERLLVYR	180
c19271_g2_i1_M_FL	VQWQEFFTGHI VQNSKVI AHIGESDFTVRINTDAEEYNVEPLWRF LNNTKDERLLVYR	178
	*****;***** . *****;*****	
AGKCO	SEDIKDISRLESPKVCYVKPEE EGIWGG LKDNHVNKENI PRRRRAIPDP M <u>KNTCKMLV</u>	240
c19271_g2_i1_M_FL	SEDIKDISRLESPKVCYVKPEE KGLWGRGHKDNVNKENI PRRRRATPDPL <u>KNTCKMLV</u>	238
	*****;:* * * * . ***** * * *;*****	
	ZnMc_TACE_like domain	
AGKCO	<u>VADHRFF</u> <u>KNMGR</u> <u>GEEST</u> <u>TINY</u> <u>LIEL</u> <u>IDR</u> <u>VDDI</u> <u>YRNT</u> <u>SWD</u> <u>NAH</u> <u>FKGY</u> <u>GIQ</u> <u>IEQ</u> <u>IVH</u> <u>SEPE</u>	300
c19271_g2_i1_M_FL	<u>VADHRFF</u> <u>KNMGR</u> <u>GEEST</u> <u>TINY</u> <u>LIEL</u> <u>IDR</u> <u>VDDI</u> <u>YRNT</u> <u>SWD</u> <u>NAH</u> <u>FNGY</u> <u>GVQ</u> <u>IEQ</u> <u>IVH</u> <u>SEPE</u>	298
	*****;*****	

Figure 10. Cont.

AGKCO	<u>NVPAKQKHYNMAKSOPDEEKDAWEVKOLLEQFSDIAEKAATVCLAHLFTYQDFDMGTLG</u>	360
c19271_g2_i1_M_FL	<u>NLPPKQKHYNMAKSOPDEEKDAWEVKOLLEQFSDIAEKAATVCLAHLFTYQDFDMGTLG</u>	358
	: *****	
AGKCO	<u>LAYVASQKPNSHGGICPKAYSNSIPGKAVYLNLSGLTSTKNYKGTILTKEADLVTT</u> <u>HELGH</u>	420
c19271_g2_i1_M_FL	<u>LAYVASQKPNSHGGICPKAYSNSVPGKAVYLNLSGLTSTKNYKGTILTKEADLVTT</u> <u>HELGH</u>	418
	*****:*****	
AGKCO	<u>NFGAEHDPDSMPECAPTEDOGQKGYVMYPIAVSGDHENNKMFNSCSKASILQTIQAKAPEC</u>	480
c19271_g2_i1_M_FL	<u>NFGAEHDPDSMLECAPNEDOGQKGYVMYPIAVSGDHENNKMFNSCSKASILQTIQAKAPEC</u>	478
	***** ****.*****	
	Disintegrin domain	
AGKCO	<u>FKERTNKVCGNSRVDEG</u> <u>EECD</u> <u>PGLLNLDL</u> <u>CG</u> <u>TSD</u> <u>KLKPNAK</u> <u>SDRNS</u> <u>PC</u> <u>KR</u> <u>QFES</u>	540
c19271_g2_i1_M_FL	<u>FKERSNKVCGNSRVDEG</u> <u>EECD</u> <u>PGLLNLDL</u> <u>CG</u> <u>TSD</u> <u>KLKPNAK</u> <u>SDRNS</u> <u>PC</u> <u>KR</u> <u>QFES</u>	538
	:**	
AGKCO	<u>ADKK</u> <u>QEAINAT</u> <u>KGESY</u> <u>TGNSSD</u> <u>SPPGNALDDTV</u> <u>VDL</u> <u>GK</u> <u>KDGE</u> <u>VPFCE</u> <u>LK</u> <u>RLR</u>	599
c19271_g2_i1_M_FL	<u>ADKK</u> <u>QEAINAT</u> <u>KGESY</u> <u>TGNSSD</u> <u>PPPGNALDDTV</u> <u>VDL</u> <u>GK</u> <u>KDGE</u> <u>VPFCE</u> <u>WEK</u> <u>SLR</u>	598
	***** ***** : **	
	ADAM17_Membrane-proximal domain	
AGKCO	<u>SCACNETDNS</u> <u>KI</u> <u>CRNESGRCS</u> <u>PYRDT</u> <u>SNFLYLRK</u> <u>GK</u> <u>PCTVGFCD</u> <u>MNGKCDKRVQDVI</u>	659
c19271_g2_i1_M_FL	<u>SCACNETDNS</u> <u>KI</u> <u>CRDEFGRCS</u> <u>PYRDIN</u> <u>QNFYLRK</u> <u>GK</u> <u>PCTVGFCD</u> <u>MNGKCDKRVQDVI</u>	658
	*****:* ***** .*****	
AGKCO	<u>ERLWVFIDQLSINTFGKFLADNIVGSVLVFSLLFWIPLSILVH</u> <u>VDK</u> <u>KLDRQYE</u> <u>ENTKSL</u>	719
c19271_g2_i1_M_FL	<u>ERLWVFIDQLSINTFGKFLADNIVGSVLVFSLLFWIPLSILVH</u> <u>VDK</u> <u>KLDRQYE</u> <u>ENTKSL</u>	718

AGKCO	<u>FGPSN</u> <u>VEM</u> <u>LSS</u> <u>MDS</u> <u>SASIRI</u> <u>IKPF</u> <u>FV</u> <u>VP</u> <u>ASR</u> <u>HQ</u> <u>FLQ</u> <u>IV</u> <u>V</u> <u>SVAAG</u> <u>PK</u> <u>LDH</u> <u>QR</u> <u>MD</u> <u>TIQE</u>	779
c19271_g2_i1_M_FL	<u>FGPSN</u> <u>VEM</u> <u>LSS</u> <u>MDS</u> <u>SASIRI</u> <u>IKPF</u> <u>FV</u> <u>VP</u> <u>ASR</u> <u>HQ</u> <u>FLQ</u> <u>IV</u> <u>V</u> <u>SVAAG</u> <u>PK</u> <u>LDH</u> <u>QR</u> <u>MD</u> <u>TIQE</u>	778

AGKCO	<u>D</u> <u>STDSHMEED</u> <u>FEKDF</u> <u>FN</u> <u>NSTA</u> <u>AKS</u> <u>FED</u> <u>LD</u> <u>HD</u> <u>V</u> <u>TR</u> <u>SEK</u> <u>SS</u> <u>SF</u> <u>KL</u> <u>QR</u> <u>HR</u> <u>VD</u> <u>SKETE</u>	839
c19271_g2_i1_M_FL	<u>D</u> <u>STDSHMEED</u> <u>FEKDF</u> <u>FN</u> <u>N</u> <u>SAA</u> <u>AKS</u> <u>FED</u> <u>LD</u> <u>HD</u> <u>V</u> <u>TR</u> <u>SEK</u> <u>SS</u> <u>SF</u> <u>KL</u> <u>QR</u> <u>HR</u> <u>VD</u> <u>SKETE</u>	838
	*****:*****	
AGKCO	- 839	
c19271_g2_i1_M_FL	* 838	

Figure 10. Sequence alignment of AGKCO (snake venom metalloproteinase from *Agkistrodon contortrix contortrix*, A0A1W7RJN7) and deduced amino acid sequences of c19271_g2_i1_M_FL. Signal peptide is single-underlined. Zinc dependent metalloproteinase domain is double-underlined. The disintegrin domain is dash-underlined. ADAM17-MPD domain is wave-underlined. CxxC motif (membrane-proximal domain) is highlighted in red. PxxP motif is highlighted in pink. Zinc binding site and EECD sequence are highlighted in yellow. Cysteine residues are highlighted in green. (*) fully conserved residues; (:) strongly similar residues; (.) weakly similar residues.

VEGF1	<u>MNFLT</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
c18061_g1_i3_M_FL	<u>MNFLT</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
c11672_g4_i2_F_FL	<u>MNFLT</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60

VEGF1	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120
c18061_g1_i3_M_FL	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120
c11672_g4_i2_F_FL	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120

VEGF1	SFQQH S C ECRPKKEVRIRQ E K C DKPRR-	148
c18061_g1_i3_M_FL	SFQQH S C ECRPKKEVRIRQ E K C EKPRR*	148
c11672_g4_i2_F_FL	SFQQH S C ECRPKKEVRIRQ E K C EKPRR*	148
***** :****		

(A)

VEGFA_VIPAA	<u>MNFL</u> LSWIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
c18061_g1_i1_M_FL	<u>MNFL</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
c18061_g1_i4_M_FL	<u>MNFL</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
c11672_g4_i1_F_FL	<u>MNFL</u> TIHWGLAALLYFHNAKVLQAAPAQGDGDRQQGEVISFLT VYERSA CRPVETMVD	60
**** :*****		
VEGFA_VIPAA	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120
c18061_g1_i1_M_FL	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120
c18061_g1_i4_M_FL	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120
c11672_g4_i1_F_FL	IFQEYPDEVEYIFKPS C VALMR C GG CC NDEALE C VPTEVYNVTMEIMKLPFQSQHIHPM	120

VEGFA_VIPAA	SFQQH S C ECRPKKEVRIRQENH C EP C SERRKHLYKQDPLT C K C SKFTDSR C KSKQLEL	180
c18061_g1_i1_M_FL	SFQQH S C ECRPKKEVRIRQENH C EP C SERRKHLYKQDPLT C K C SKFTDSR C KSKQLEL	180
c18061_g1_i4_M_FL	SFQQH S C ECRPKKEVRIRQENH C EP C SERRKHLYKQDPLT C K C SKFTDSR C KSKQLEL	180
c11672_g4_i1_F_FL	SFQQH S C ECRPKKEVRIRQENH C EP C SERRKHLYKQDPLT C K C SKFTDSR C KSKQLEL	180

VEGFA_VIPAA	NERT C R C CEKPRR-	192
c18061_g1_i1_M_FL	NERT C R C CEKPRR*	192
c18061_g1_i4_M_FL	NERT C R C CEKPRR*	192
c11672_g4_i1_F_FL	NERT C R C CEKPRR*	192

(B)

Figure 11. Sequence alignment of deduced amino acid sequences of VEGF transcripts with vascular endothelial growth factor 1 (VEGF1) from *Vipera anatolica senliki* (QHR82857) (A) and vascular endothelial growth factor A precursor (VEGFA-VIPAA) from *Vipera ammodytes ammodytes* (C0K3N5) (B). Signal peptide is single-underlined. Cysteine residues are in yellow color. (*) fully conserved residues; (:) strongly similar residues.

2.11. L-Amino Acid Oxidases (LAAOs)

In this study, two full-length and one partial-length LAAO transcripts with 99.93% identity (84% query coverage, E value 0.0) to DrLAO (G8XQX1) were identified. The sequence alignment is shown in Supplementary Figure S5.

2.12. Cysteine-Rich Venom Proteins (CRISPs)

There were two full-length highly expressed cysteine-rich secretory protein (CRISP) transcripts identified in the current transcriptomes which gave the annotation of Dr-CRIPK from *D. russelii* (ACE73567), currently termed *D. siamensis* (Supplementary Figure S6). These deduced amino acid sequences were also identical to those of CRISPs (MRV CRISP) (JZ980958, JZ980965 and JZ980970), which were expressed in previous cDNA libraries [26].

3. Discussion

The de novo venom gland transcriptomic analysis provides a unique profile of the toxin composition of Myanmar *D. siamensis* venom and its suggestive toxicogenomic diversity. Our previous study on the venom gland cDNA library revealed unique gene expression patterns [26]; however, the use of NGS technology allowed us to provide a broader examination into the variation, distribution and diversity of toxin transcripts. Using Trinity as our de novo transcriptome assembly program, we were able to assemble a robust transcriptome using high-quality (>98% Q20 for both male and female) raw reads and subsequent mapping for gene expression analysis. Our initial exploration into these transcriptomes relied on RSEM and gene expression information to identify focal toxin candidate transcripts; however, future studies would benefit from exploring additional assembly programs providing a multi-faceted approach to analyzing these transcriptomes and account for chimeric transcripts [27]. The venom proteome of *D. siamensis* comprised six toxin families, including serine proteinases, metalloproteinases, PLA₂, LAAO, VEGF and CTLs [4]. The BPP-CNP transcripts containing SVMPIs were identified only in the transcriptomes due to the co-expression of SVMPI tripeptides and atrial natriuretic peptides in the same transcripts [28]. Only the processed products could be detected in the venom. Low-molecular-weight disintegrins (<10 kDa) were detected only in a transcriptome because a proteomic study could identify proteins larger than 10 kDa. The reason that CRISPs (27 kDa) were not found in the proteome was unclear.

We found sex-related differences in venom transcript expression in Russell's viper, also detected in *Bothrops jararaca* [12], *Thamnophis elegans* [17] and *Mesobuthus martensii* [29]. These differences reflected the sex-based variations of the venom proteome and were closely associated with the biochemical and biological properties of different genders in the same species [14,30–32]. In *B. jararaca* and *B. moojeni*, the variability in the venom composition was associated with animal sizes, as females are larger than males [32,33]. This coincided with the sexual dimorphism of the venom donors used in the study, as females (average weight in ounces × total length in inches × girth in inches: 49.9 × 47.8 × 6.5) were larger than males (29.8 × 47 × 5.3) and more venom was collected from females. There is also a correlation between venom yield and the length of Myanmar Russell's vipers, which is related to the severity of envenoming, time for coagulopathy development and the severity of local manifestations [34].

The transcripts corresponding to the toxin groups described below were found to be most abundant in both sexes and aligned with Myanmar Russell's viper envenomation symptoms. Snake venom C-type lectins (snaclecs) are heterodimers of two closely related α and β subunits [35]. They target factor IX, factor X, prothrombin or α-thrombin (as pro- or anticoagulant) and platelet receptors, such as glycoprotein (Gp) Ibα, Gp Ia/Ia, or Gp VI (as pro- or antithrombotic) [36,37]. Snaclecs activate platelets and thrombus formation causing thrombo-inflammation during viper envenoming [38]. Dabocetin α together with its β subunit interacts with platelet GpIb receptor, inhibiting platelet aggregation [39]. *D. siamensis* Snaclec-7 is a β subunit of a GpIb-binding protein [40], and together with dabocetin α is involved in platelet inhibition. The *M. lebetina* Snaclec A12 homolog, originally found in

D. siamensis, has a stronger sequence similarity to α than β subunits compared to other heterodimeric snakes [41].

Snake venom protease inhibitors primarily disrupt blood coagulation resulting in blood loss and the death of preys [42]. The largest and most widely distributed family is Kunitz-type protein inhibitors, which strongly inhibit serine proteases. They contain 58–60 amino acid residues, including three disulfide bridges [43]. Several Kunitz-type protease inhibitors from *Daboia* species have previously been identified.

The snake venom disintegrins are a group of small cysteine-rich proteins found in a variety of snake families. They are divided into seven groups according to disulfide patterns and sequence signatures [44]. Jerdostatin homologs (Dis1a and Dis1b) containing eight cysteines belong to group 7A. Most of these small disintegrins are synthesized from short-coding genes, instead of a proteolytic process from PII SVMPs like the majority of disintegrins [45]. The RTS-disintegrins are expressed from the *RPTLN* gene and have a high conserved signal peptide sequence with *SVMP* genes, suggesting they may share tissue-specific regulatory elements [46,47]. The disintegrin motif is responsible for binding specificity to integrin receptors [48]. Modulating disintegrin motifs is interesting for the development of antiangiogenesis, antimetastasis, antiproliferative and proapoptotic agents for cancer treatments. Lowly expressed VGD- and MLD-containing dimeric disintegrins observed in our previous study [25] were not evaluated in the current study as the candidate transcripts had TPM values lower than 100 and were therefore not analyzed.

The BPP-CNP transcripts usually carry a tripeptide repeat together with bradykinin-potentiating peptides (BPPs) and natriuretic peptide (NP) precursors [28,49]. The two tripeptides (pERW and pEKW) from Myanmar Russell's viper completely inhibited the gelatinolytic effects of RVV-X and the fibrinogenolytic activities of daborhagin [28]. These viper-restricted inhibitory peptides protected the venom gland from autodigestion and prevented hydrolysis of other proteins in venom glands [50]. Their inhibitory activity towards the SVMP Zn²⁺ binding site [51] has made them popular targets for drug design in areas of neurodegenerative and cardiovascular diseases. Natriuretic peptides (NPs) in vertebrates promote salt secretion, regulating cardiovascular and renal systems. Three mammalian NPs are atrial natriuretic peptide (ANP), B-type natriuretic peptide (BNP) and C-type natriuretic peptide (CNP). ANP and BNP are released by cardiomyocytes in response to hypertension and hypervolemia, and CNP is produced by endothelial cells. The sequence alignment of various snake venom and human natriuretic peptides is shown in Figure 6, where five residues (F, D, R, I/V and L/F) are crucial for receptor binding [50]. Among MRV CNP transcripts, one had the same conserved amino acid residues, while remaining three possessed V instead of I and F instead of L residues. These variations should be explored in relation to their roles in terms of toxicity and their potential therapeutic properties. NPs have their specific niche in medical research in the treatment of heart failure and hypertension [51,52].

Phospholipases A₂ (PLA₂) catalyze a hydrolysis of glycerophospholipids, fatty acids and lysophospholipids and release arachidonic acids to trigger inflammation. Snake venom PLA₂s (svPLA₂) are small, calcium-dependent, secreted enzymes ranging from 12 to 14 kDa. The class I PLA₂s are found in Elapidae and Hydrophiidae, while class II are in Crotalidae and class III are in lizard and bee venoms [53]. There are two major types in Russell's vipers. PLA₂s containing an N-terminal Asn (N) are found in species from China, Thailand, Myanmar and Pakistan, while the venoms of southern Indian and Sri Lankan species contain PLA₂ with an N-terminal serine (S). They confer hypotensive with neurotoxic and myonecrotic properties, respectively [26]. Furthermore, Russell's viper PLA₂ showed nephrotoxicity in mice [54]. All of the three PLA₂ transcripts from the current transcriptomes contained N-terminal Asparagine (N) and conserved Aspartic acid (D) at the 49 position. In all transcripts, 14 cysteine residues for seven conserved disulfide bonds indicated that they were GIIA PLA₂s. They all possessed Thr41 which was found to be conserved in all the neurotoxic or myotoxic PLA₂s. From Myanmar *D. siamensis* venom,

DsM b1 had neurotoxic and lethal effects [55] and dabotoxin showed neurotoxic, myotoxic, cytotoxic, edema-inducing and indirect hemolytic activities [56].

Snake venom serine proteases (SVSPs) have two six-stranded β -barrels and a catalytic triad of His57, Asp102 and Ser159. Viper venoms contain a large variety of SVSPs, mostly as monomeric glycoproteins of 26 to 67 kDa. They act on almost all components of hemostatic and fibrinolytic systems and are responsible for the hemorrhage and shock after envenomation. Based on their biological roles, they are classified as activators of the fibrinolytic, procoagulant, anticoagulant and platelet-aggregating enzymes [57–59]. SVSPs from *D. russelii* are procoagulants that (i) activate factor V (RVV-V), (ii) have thrombin-like action causing defibrinogenation (Russelobin) and (iii) are $\alpha\beta$ fibrinogenases with factor V-activation (RV-FVP α - δ). RV-FVPs showed in vivo anticoagulant effects in addition to in vitro procoagulant effects [60].

Snake venom metalloproteinases (SVMPs) are zinc endopeptidases varying from 20 to 100 kDa. They are the major components of viper venom, varying from 11% to 65% of the whole venom cocktail [61,62]. SVMPs have sequence and domain similarities to ADAM (A disintegrin and metalloproteinase) and can be grouped into three major classes (PI-PIII) according to their organization [53,63]. SVMPs are responsible for local and systemic hemorrhage, myonecrosis, the disruption of hemostasis mediated by procoagulant or anticoagulant effects or platelet aggregation, tissue necrosis and pro-inflammation [61]. These activities are mediated by metalloproteinase, disintegrin and/or disintegrin-like domains [62]. Snake venom metalloproteinases (SVMPs) isolated from Russell's viper were RVV-X (85 kDa), Daborhagin (65 kDa), VRR-73 (73 kDa), VRH-1 (23 kDa) and RVBCMP (25 kDa). RVV-X is a well-characterized class P-IIIId SVMP. It activates factor X via a single cleavage at the Arg¹⁹⁴-Ile¹⁹⁵ bond in factor X, which is responsible for disseminated intravascular coagulation (DIC) in snakebite patients [64]. RVV-X also binds to receptors on platelets and endothelial cells, resulting in clinical bleeding [60,65] and contributing to acute kidney injury in animal models [66]. The structure and biological activity of DSAIP still needs to be elucidated. RVV-X and daborhagin are found in both Myanmar and Indian RV species, while VRR-73, VRH-1 and RVBCMP are from only Indian RV [61]. Daborhagin-M is a P-IIIC hemorrhagic SVMP with high proteolytic activity toward fibrinogen, fibronectin and collagen [67]. These activities may be related to the severe bleeding in the patients.

A disintegrin and metalloproteinase (ADAM) family proteins in focal transcriptomes were expressed at moderate levels. Transcripts homologous to VIPAN (ADAM 10) from *V. a. senliki* and AGKCO (ADAM 17) from *A. c. contortrix* were found in Myanmar Russell's viper. ADAM 10 played a major role in the signaling pathway and ADAM 17 in the processing of tumor necrosis factor- α and other cell-surface molecules [68]. Although ADAM 17 is considered to be the primary sheddase for TNF- α , ADAM 10 also worked as a TNF- α sheddase in certain cell types [69]. The biological functions of ADAM 10 and 17 from snakes require further investigation.

Snake venom vascular endothelial growth factors (svVEGFs) play important roles in angiogenesis (endothelial growth and proliferation), hypotension and vascular permeability by binding to VEGF receptors. svVEGFs are 25-kDa, homodimeric heparin-binding proteins with 50% identity to VEGF-A165. They have a VEGF homology domain (VHD), with eight conserved cysteines forming a cysteine knot. The C-terminal variable region corresponds to heparin and NP-1 (coreceptor) binding sites [70,71]. VR-1, a svVEGF from *D. russelii*, increased endothelial cell proliferation, vascular permeability and hypotension through the specific activation of endothelial KDR (kinase insert domain-containing receptor) [72]. The VR-1 isolated from *D. siamensis* also increased endothelial cell proliferation [73]. The markedly enhanced vascular permeability caused by svVEGFs can lead to hypotension and shock.

4. Conclusions

The de novo venom gland transcriptomic analysis provided a unique profile of venom gene expression in *D. siamensis* from Myanmar. The expression patterns of toxin-encoding

genes showed high structural redundancy in comparison to non-toxin genes, suggesting distinct functions of the variants in each group. The expression patterns of venom genes were also sex-specific. The distribution and expression levels of the principal toxin components yielded deeper insights into the toxic syndromes of Russell's viper envenoming, which included bleeding, hypotension and renal failure. Moreover, the full-length toxin sequences from this study would be the references for future proteomic analysis. Newly identified venom components will be a source for future structure–functional relationship studies and the development of new therapeutic modalities. Taking into account that the genomes of these species were not yet completed, this study provides rich data for a fast advancement in toxicogenomic research to study evolutionary origins and diversity in snakes.

5. Materials and Methods

5.1. Sample Preparation

Four (two male and two female) adult Russell's viper, *D. siamensis*, were captured from Kyun Gyan Gon Township, Yangon, Myanmar. Venom milking was carried out prior to tissue harvesting to stimulate the toxin transcription. The snakes were allowed to rest for four days to maximize RNA transcription. The venom glands were then removed, cut into pieces smaller than 5×5 mm and stored in an RNAlater solution (Ambion, Austin, TX, USA) at 4 °C overnight before transferring to -80 °C.

5.2. RNA Extraction, cDNA Library Construction and Sequencing

Total RNA from venom gland tissues was extracted using Trizol reagent (Life Technologies, Carlsbad, CA, USA) or a Total RNA Purification Kit (Jena Bioscience GmbH, Jena, Germany). The sample were stored at -80 °C. Subsequently, mRNA from samples was isolated using PolyAT Tract mRNA Isolation System (Promega, Madison, WI, USA) or a FastTrack MAG mRNA Isolation Kit (Invitrogen, Carlsbad, CA, USA). Both systems used MagneSphere[®] technology. Subsequently, mRNA was precipitated, concentrated using 3M sodium acetate and isopropanol and stored at -80 °C. Isolated mRNA was suspended in RNase-free water and sent to Macrogen Inc. Geumchun-gu, Seoul, Republic of Korea. The mRNA sequencing libraries were prepared using the TruSeq RNA sample preparation kit (Illumina Inc., San Diego, CA, USA) with selected insert sizes of 200–400 bp. Libraries were then sequenced on the Illumina HiSeq2000 platform using (2×100 bp) paired-end reads.

5.3. De Novo Transcriptome Assembly and Expression Quantification

The quality of raw data or raw reads from the Illumina platform was assessed with the quality assessment software FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>; accessed on 18 November 2014) and cleaned up using Trimmomatic (0.32) (<http://www.usadellab.org/cms/?page=trimmomatic>; accessed on 20 November 2014) to remove adaptor sequences, nucleotide errors and low-quality sequences. Two assembly tools, Trinity [24] and CLC Genomic Workbench 7.5 (CLC bio, a QIAGEN Company, Aarhus, Denmark), were used to assemble and compare alternative pathways for de novo transcriptome assemblies. The quality assessment tool QUAST was used to evaluate the quality of the assembled transcriptomes [74], indicating that the Trinity assembly had a higher quality score than the CLC genomic assembly (Supplementary Table S4). Thus, assembled data via Trinity were used for downstream analysis.

The transcriptomes were assembled in the Trinity software (r20140717) [24]. For each transcriptome, expression values were determined as transcripts per million (TPM) and fragments per kilobase million (FKPM) were calculated in the program RSEM [75,76] as part of the Trinity package. The raw data of the assembly were uploaded to NCBI database under BioProject PRJNA545823, Biosample SAMN11938797 for males and SAMN11939170 for females. The expression levels in terms of both TPM and FKPM of toxin transcripts were mapped to transcripts from non-venom gland tissues, such as brain, kidney, heart, lung, liver and spleen (Bioproject PRJNA269317, Biosamples SAMN03081270), of *D. siamensis*.

5.4. Venom Gland Transcript Classification

Toxin transcripts from the two transcriptomic datasets were identified and characterized using Venomix [77]. The Venomix pipeline identified candidate toxin genes based on BLAST hits (E-value 1×10^{-20}) against the ToxProt dataset [78]. A transcript was considered a “candidate” if the transcript had a lower E-value associated with a toxin than with a non-toxin protein in Uniprot. The candidate transcripts were translated into their protein sequences using Transdecoder [79] and further evaluated in ToxClassifier [80]. A predicted protein sequence with a score of >1 was considered a toxin candidate. The annotation of generated transcript sequences was aligned by standalone BLASTP (v. 2.10.0) to the protein database UniProtKB/Swiss-Prot [81] (Swissprot, Swiss Institute of Bioinformatics, Lausanne, Switzerland) of *Serpentes* (taxid: 8570) organism. Sequences with no matches were then searched against non-redundant protein sequences (nr) of the same organism. Proteins with the highest ranks in the BLASTP results were referenced to determine the functional annotations.

The transcripts from this study were named as “c” (component/contig), “g” (sub-graph/trinity gene) and “i” (path sequence/trinity isoform) followed by numbers. The letter “F” stood for female, “M” for male, “PL” for partial-length and “FL” for full-length transcripts.

5.5. Multiple Sequences Alignment and Gene Tree Reconstruction

The most highly toxin transcripts (TPM values ≥ 100) were conceptually translated into protein sequences and subjected to alignment using Clustal Omega [82]. Alignments of the first-described and previously reported sequences in Russell’s viper are displayed in the manuscript and Supplementary Material, respectively. The conserved domains and functional motifs are identified and referenced in the figures. For some toxin gene families, the alignments were used to reconstruct in the program RAxML [83] using the PROTGAMMABLOSUM62 model with 1000 bootstrap replicates.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/toxins15050309/s1>, Figure S1: Clustal Omega sequence alignment of the polypeptides encoded by 23 transcript candidate C-type lectins) from Myanmar *D. siamensis*; Figure S2: The gene tree analysis of the disintegrins from Myanmar *D. siamensis*, Dis1a (c13890_g2_i1_M) and Dis1b (c11797_g1_i1_F), with disintegrins from other species; Figure S3: Sequence alignment of RVV-V γ from *Daboia siamensis* and deduced amino acid sequences of transcripts from the transcriptomes; Figure S4: Sequence alignment of deduced amino acid sequences of VEGF transcripts with VR-1 from *D. russelii* (P67861); Figure S5: Sequence alignment of DrLAO (*Daboia russelii*, G8XQX1.1) with three deduced LAAO transcripts obtained from Myanmar Russell’s viper transcriptome; Figure S6: Sequence alignment of deduced amino acid sequences of CRISP transcripts from the transcriptomes with cysteine-rich secretory protein Dr-CRPK from Taiwan Russell’s viper (ACE73567); Table S1: Comparison of expression levels of toxin transcripts from male sample with non-toxin transcripts from tissues of *D. siamensis*; Table S2: Comparison of expression levels of toxin transcripts from female sample with non-toxin transcripts from tissues of *D. siamensis*; Table S3: Transcripts with snake venom C-type lectins (CTLs) annotation; Table S4; QCAST reports of assembled transcriptomes by two different assemblers; Table S5: List of full-length sequences with toxin gene families of Myanmar Russell’s viper transcriptome. Reference [84] are cited in the supplementary materials.

Author Contributions: Conceptualization, P.R. and K.T.Y.; methodology, P.R., K.T.Y. and J.M.; software, J.M. and O.V.; validation, P.R., J.M. and O.V.; formal analysis, K.T.Y., J.M. and O.V.; investigation, K.T.Y.; resources, P.R.; data curation, J.M. and O.V.; writing—original draft preparation, K.T.Y.; writing—review and editing, K.T.Y., P.R., J.M. and O.V.; visualization, P.R., J.M. and O.V.; supervision, P.R. and O.V.; project administration, P.R.; funding acquisition, P.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Thailand Science Research and Innovation Fund Chulalongkorn University, CU_FRB65_he(17)_024_30_05.

Institutional Review Board Statement: The study was approved by the Animal Care and Use Committee, Chulalongkorn University (CU-ACUC) (Protocol code 17/2552) on 28 July 2015.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data of the transcriptome assembly were uploaded to NCBI database under BioProject PRJNA545823, Biosample SAMN11938797 for males and SAMN11939170 for females.

Acknowledgments: Administrative persons and snake handlers from the Snake Farm, Myanmar Pharmaceutical Factory, Yangon, Myanmar and Lawan Chanhome, Head of Snake Farm, QSMI, Bangkok, Thailand for help in the collection of snakes and removal of venom glands.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Thorpe, R.S.; Pook, C.E.; Malhotra, A. Symptoms of envenoming. *Herpetol. J.* **2007**, *17*, 209–218.
2. Wüster, W. The genus *Daboia* (Serpentes: Viperidae): Russell's viper. *Hamadryad* **1998**, *23*, 33–40.
3. Warrell, D.A. Snake venoms in science and clinical medicine 1. Russell's viper: Biology, venom and treatment of bites. *Trans. R. Soc. Trop. Med. Hyg.* **1989**, *83*, 732–740. [[CrossRef](#)]
4. Risch, M.; Georgieva, D.; von Bergen, M.; Jehmlich, N.; Genov, N.; Arni, R.K.; Betzel, C. Snake venomomics of the Siamese Russell's viper (*Daboia russelli siamensis*)—relation to pharmacological activities. *J. Proteom.* **2009**, *72*, 256–269. [[CrossRef](#)] [[PubMed](#)]
5. Thomas, R.; Alfred, S.; Zaw, A.; Gentili, S.; Chataway, T.; Peh, C.A.; Hurtado, P.; White, J.; Wiese, M. A demonstration of variation in venom composition of *Daboia russelii* (Russell's viper), a significantly important snake of Myanmar, by tandem mass spectrometry. *Toxicon* **2019**, *158*, S44–S45. [[CrossRef](#)]
6. Lingam, T.M.C.; Tan, K.Y.; Tan, C.H. Proteomics and antivenom immunoprofiling of Russell's viper (*Daboia siamensis*) venoms from Thailand and Indonesia. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2020**, *26*, 1–17. [[CrossRef](#)]
7. Tan, K.Y.; Tan, N.H.; Tan, C.H. Venom proteomics and antivenom neutralization for the Chinese eastern Russell's viper, *Daboia siamensis* from Guangxi and Taiwan. *Sci. Rep.* **2018**, *8*, 8545. [[CrossRef](#)]
8. Pla, D.; Sanz, L.; Quesada-Bernat, S.; Villalta, M.; Baal, J.; Chowdhury, M.A.W.; León, G.; Gutiérrez, J.M.; Kuch, U.; Calvete, J.J. Phylovenomics of *Daboia russelii* across the Indian subcontinent. Bioactivities and comparative in vivo neutralization and in vitro third-generation antivenomics of antivenoms against venoms from India, Bangladesh and Sri Lanka. *J. Proteom.* **2019**, *207*, 103443. [[CrossRef](#)] [[PubMed](#)]
9. Mukherjee, A.K.; Kalita, B.; Mackessy, S.P. A proteomic analysis of Pakistan *Daboia russelii russelii* venom and assessment of potency of Indian polyvalent and monovalent antivenom. *J. Proteom.* **2016**, *144*, 73–86. [[CrossRef](#)]
10. Casewell, N.R.; Huttley, G.A.; Wüster, W. Dynamic evolution of venom proteins in squamate reptiles. *Nat. Commun.* **2012**, *3*, 1066. [[CrossRef](#)] [[PubMed](#)]
11. Calvete, J.J.; Sanz, L.; Angulo, Y.; Lomonte, B.; Gutiérrez, J.M. Venoms, venomomics, antivenomics. *FEBS Lett.* **2009**, *583*, 1736–1743. [[CrossRef](#)]
12. Zelanis, A.; Andrade-Silva, D.; Rocha, M.M.; Furtado, M.F.; Serrano, S.M.T.; Junqueira-de-Azevedo, I.L.M.; Ho, P.L. A transcriptomic view of the proteome variability of newborn and adult *Bothrops jararaca* snake venoms. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1554. [[CrossRef](#)] [[PubMed](#)]
13. Barlow, A.; Pook, C.E.; Harrison, R.A.; Wüster, W. Coevolution of diet and prey-specific venom activity supports the role of selection in snake venom evolution. *Proc. R. Soc. B Biol. Sci.* **2009**, *276*, 2443–2449. [[CrossRef](#)] [[PubMed](#)]
14. Menezes, M.C.; Furtado, M.F.; Travaglia-Cardoso, S.R.; Camargo, A.C.M.; Serrano, S.M.T. Sex-based individual variation of snake venom proteome among eighteen *Bothrops jararaca* siblings. *Toxicon* **2006**, *47*, 304–312. [[CrossRef](#)] [[PubMed](#)]
15. Mirtschin, P.J.; Shine, R.; Nias, T.J.; Dunstan, N.L.; Hough, B.J.; Mirtschin, M. Influences on venom yield in Australian tigersnakes (*Notechis scutatus*) and brownsnakes (*Pseudonaja textilis*: Elapidae, Serpentes). *Toxicon* **2002**, *40*, 1581–1592. [[CrossRef](#)]
16. Brahma, R.K.; McCleary, R.J.R.; Kini, R.M.; Doley, R. Venom gland transcriptomics for identifying, cataloging, and characterizing venom proteins in snakes. *Toxicon* **2015**, *93*, 1–10. [[CrossRef](#)]
17. Schwartz, T.S.; Tae, H.; Yang, Y.; Mockaitis, K.; Van Hemert, J.L.; Proulx, S.R.; Choi, J.H.; Bronikowski, A.M. A garter snake transcriptome: Pyrosequencing, de novo assembly, and sex-specific differences. *BMC Genom.* **2010**, *11*, 694. [[CrossRef](#)]
18. Cardoso, K.C.; Da Silva, M.J.; Costa, G.G.L.; Torres, T.T.; Del Bem, L.E.V.; Vidal, R.O.; Menossi, M.; Hyslop, S. A transcriptomic analysis of gene expression in the venom gland of the snake *Bothrops alternatus* (urutu). *BMC Genom.* **2010**, *11*, 605. [[CrossRef](#)]
19. Tan, C.H.; Tan, K.Y.; Fung, S.Y.; Tan, N.H. Venom-gland transcriptome and venom proteome of the Malaysian king cobra (*Ophiophagus hannah*). *BMC Genom.* **2015**, *16*, 687. [[CrossRef](#)]
20. Ducancel, F.; Durban, J.; Verdenaud, M. Transcriptomics and venomomics: Implications for medicinal chemistry. *Future Med. Chem.* **2014**, *6*, 1629–1643. [[CrossRef](#)]
21. Barua, A.; Mikheyev, A.S. Toxin expression in snake venom evolves rapidly with constant shifts in evolutionary rates. *Proc. R. Soc. B Biol. Sci.* **2020**, *287*, 20200613. [[CrossRef](#)]

22. Faisal, T.; Tan, K.Y.; Sim, S.M.; Quraishi, N.; Tan, N.H.; Tan, C.H. Proteomics, functional characterization and antivenom neutralization of the venom of Pakistani Russell's viper (*Daboia russelii*) from the wild. *J. Proteom.* **2018**, *183*, 1–3. [[CrossRef](#)]
23. Saethang, T.; Somparn, P.; Payungporn, S.; Sriswasdi, S.; Yee, K.T.; Hodge, K.; Knepper, M.A.; Chanhom, L.; Khoo, O.; Chaiyabutr, N.; et al. Identification of *Daboia siamensis* venom using integrated multi-omics data. *Sci. Rep.* **2022**, *12*, 13140. [[CrossRef](#)]
24. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
25. Yee, K.T.; Tongsim, S.; Vasieva, O.; Ngamphiw, C.; Wilantho, A.; Wilkinson, M.C.; Somparn, P.; Pisitkun, T.; Rojnuckarin, P. Analysis of snake venom metalloproteinases from Myanmar Russell's viper transcriptome. *Toxicon* **2018**, *146*, 31–41. [[CrossRef](#)] [[PubMed](#)]
26. Yee, K.T.; Rojnuckarin, P. Complementary DNA library of Myanmar Russell's viper (*Daboia russelii siamensis*) venom gland. *Comp. Biochem. Physiol. Part-C Toxicol. Pharmacol.* **2020**, *227*, 108634. [[CrossRef](#)]
27. Holding, M.L.; Margres, M.J.; Mason, A.J.; Parkinson, C.L.; Rokyta, D.R. Evaluating the performance of *de novo* assembly methods for venom-gland transcriptomics. *Toxins* **2018**, *10*, 249. [[CrossRef](#)] [[PubMed](#)]
28. Yee, K.T.; Pitts, M.; Tongyoo, P.; Rojnuckarin, P.; Wilkinson, M.C. Snake venom metalloproteinases and their peptide inhibitors from Myanmar Russell's viper venom. *Toxins* **2017**, *9*, 15. [[CrossRef](#)]
29. Yang, Y.; Zhu, Y.; Luo, Y.; Liu, Q.; Hua, X.; Li, J.; Gao, F.; Hofer, J.; Gao, X.; Xiao, L.; et al. Transcriptome analysis of *Mesobuthus martensii* revealed the differences of their toxins between females and males. *Eur. Zool. J.* **2022**, *89*, 1294–1311. [[CrossRef](#)]
30. Lourenço, A.; Zorzella Creste, C.F.; Curtolo de Barros, L.; Delazari dos Santos, L.; Pimenta, D.C.; Barraviera, B.; Ferreira, R.S. Individual venom profiling of *Crotalus durissus terrificus* specimens from a geographically limited region: Crotamine assessment and captivity evaluation on the biological activities. *Toxicon* **2013**, *69*, 75–81. [[CrossRef](#)]
31. Tan, C.H.; Tan, K.Y.; Yap, M.K.K.; Tan, N.H. Venomics of *Tropidolaemus wagleri*, the sexually dimorphic temple pit viper: Unveiling a deeply conserved atypical toxin arsenal. *Sci. Rep.* **2017**, *7*, 43237. [[CrossRef](#)]
32. Amorim, F.G.; Costa, T.R.; Baiwir, D.; De Pauw, E.; Quinton, L.; Sampaio, S.V. Proteo-peptidomic, functional and immunoreactivity characterization of *Bothrops moojeni* snake venom: Influence of snake gender on venom composition. *Toxins* **2018**, *10*, 177. [[CrossRef](#)] [[PubMed](#)]
33. Freitas-de-Sousa, L.A.; Nachtigall, P.G.; Portes-Junior, J.A.; Holding, M.L.; Nystrom, G.S.; Ellsworth, S.A.; Guimarães, N.C.; Tiomyama, E.; Ortiz, F.; Silva, B.R.; et al. Size Matters: An evaluation of the molecular basis of ontogenetic modifications in the composition of *Bothrops jararacussu* snake venom. *Toxins* **2020**, *12*, 791. [[CrossRef](#)]
34. Pe, T.; Aye, B.; Myint, A.-A.; Swe, T.-N.; Warrell, D. Effect of the snake length's and recent feeding on venom antigenaemia and severity of envenoming in Russell's viper (*Daboia russelii siamensis*) bites. In *Management of Snakebite and Research*; WHO, Regional Office for South-East-Asia: Geneva, Switzerland, 2002; pp. 41–45.
35. Clemetson, K.J.; Morita, T.; Manjunatha Kini, R. Scientific and standardization committee communications: Classification and nomenclature of snake venom C-type lectins and related proteins. *J. Thromb. Haemost.* **2009**, *7*, 360. [[CrossRef](#)] [[PubMed](#)]
36. Clemetson, K.J. Snaclecs (snake C-type lectins) that inhibit or activate platelets by binding to receptors. *Toxicon* **2010**, *56*, 1236–1246. [[CrossRef](#)] [[PubMed](#)]
37. Morita, T. Structures and functions of snake venom CLPs (C-type lectin-like proteins) with anticoagulant-, procoagulant-, and platelet-modulating activities. *Toxicon* **2005**, *45*, 1099–1114. [[CrossRef](#)]
38. Teixeira, C.; Fernandes, C.M.; Leiguez, E.; Chudzinski-Tavassi, A.M. Inflammation Induced by platelet-activating Viperid snake venoms: Perspectives on thromboinflammation. *Front. Immunol.* **2019**, *10*, 2082. [[CrossRef](#)]
39. Zhong, S.R.; Jin, Y.; Wu, J.B.; Chen, R.Q.; Jia, Y.H.; Wang, W.Y.; Xiong, Y.L.; Zhang, Y. Characterization and molecular cloning of dabocetin, a potent antiplatelet C-type lectin-like protein from *Daboia russelii siamensis* venom. *Toxicon* **2006**, *47*, 104–112. [[CrossRef](#)] [[PubMed](#)]
40. Zhong, S.; Jin, Y.; Li, D.; Wang, W.; Xiong, Y. Molecular cloning and sequence analyses of cDNAs encoding seven C-type lectin-like protein subunits from *Daboia russelii siamensis*. *Zool. Res.* **2005**, *26*, 337–343.
41. Jebali, J.; Bazaa, A.; Sarray, S.; Benhaj, K.; Karboul, A.; El Ayeb, M.; Marrakchi, N.; Gargouri, A. C-type lectin protein isoforms of *Macrovipera lebetina*: cDNA cloning and genetic diversity. *Toxicon* **2009**, *53*, 228–237. [[CrossRef](#)]
42. Inagaki, H. Snake venom protease inhibitors: Enhanced identification, expanding biological function, and promising future. In *Snake Venoms*; Inagaki, H., Vogel, C.-W., Mukherjee, A.K., Rahmy, T.R., Gopalakrishnakone, P., Eds.; Springer: Dordrecht, The Netherlands, 2015; pp. 1–26. [[CrossRef](#)]
43. Thakur, R.; Mukherjee, A.K. Pathophysiological significance and therapeutic applications of snake venom protease inhibitors. *Toxicon* **2017**, *131*, 37–47. [[CrossRef](#)] [[PubMed](#)]
44. Vasconcelos, A.A.; Estrada, J.C.; David, V.; Wermelinger, L.S.; Almeida, F.C.L.; Zingali, R.B. Structure-function relationship of the disintegrin family: Sequence signature and integrin interaction. *Front. Mol. Biosci.* **2021**, *8*, 783301. [[CrossRef](#)] [[PubMed](#)]
45. Juárez, P.; Wagstaff, S.C.; Sanz, L.; Harrison, R.A.; Calvete, J.J. Molecular cloning of *Echis ocellatus* disintegrins reveals non-venom-secreted proteins and a pathway for the evolution of ocellatusin. *J. Mol. Evol.* **2005**, *63*, 183–193. [[CrossRef](#)] [[PubMed](#)]
46. Sanz-Soler, R.; Sanz, L.; Calvete, J.J. Distribution of RPTNL genes across Reptilia: Hypothesized role for RPTLN in the evolution of SVMs. *Integr. Comp. Biol.* **2016**, *56*, 989–1003. [[CrossRef](#)]

47. Sanz-Soler, R.; Lorente, C.; Company, B.; Sanz, L.; Juarez, P.; Perez, A.; Zhang, Y.; Jin, Y.; Chen, R.; Eble, J.A.; et al. Recombinant expression of mutants of the *Frankenstein* disintegrin, RTS-ocellatusin. Evidence for the independent origin of RGD and KTS/RTS disintegrins. *Toxicon* **2012**, *60*, 665–675. [[CrossRef](#)]
48. Calvete, J.J. Structure-function correlations of snake venom disintegrins. *Curr. Pharm. Des.* **2005**, *11*, 829–835. [[CrossRef](#)] [[PubMed](#)]
49. Wagstaff, S.C.; Favreau, P.; Cheneval, O.; Laing, G.D.; Wilkinson, M.C.; Miller, R.L.; Stöcklin, R.; Harrison, R.A. Molecular characterisation of endogenous snake venom metalloproteinase inhibitors. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 650–656. [[CrossRef](#)] [[PubMed](#)]
50. Munawar, A.; Ali, S.A.; Akrem, A.; Betzel, C. Snake venom peptides: Tools of biodiscovery. *Toxins* **2018**, *10*, 474. [[CrossRef](#)] [[PubMed](#)]
51. Averin, A.S.; Utkin, Y.N. Cardiovascular Effects of Snake Toxins: Cardiotoxicity and Cardioprotection. *Acta Nat.* **2021**, *13*, 4–14. [[CrossRef](#)]
52. Frangieh, J.; Rima, M.; Fajloun, Z.; Henrion, D.; Sabatier, J.M.; Legros, C.; Mattei, C. Snake venom components: Tools and cures to target cardiovascular diseases. *Molecules* **2021**, *26*, 2223. [[CrossRef](#)]
53. Coronado, M.A.; de Moraes, F.R.; Ullah, A.; Masood, R.; Mariutti, R.; Brognaro, H.; Georgieva, D.; Murakami, M.T.; Betzel, C.; Arni, R.K. Three-dimensional structures and mechanism of snake venom serine proteinases, metalloproteinases, and phospholipase A₂s. In *Venom Genomics and Proteomics*; Gopalakrishnakone, P., Calvete, J., Eds.; Springer: Dordrecht, The Netherlands, 2016; pp. 239–267. [[CrossRef](#)]
54. Chaisakul, J.; Khoo, O.; Wiwatwarayos, K.; Rusmili, M.R.A.; Prasert, W.; Othman, I.; Abidin, S.A.Z.; Charoenpitakchai, M.; Hodgson, W.C.; Chanhom, L.; et al. A biochemical and pharmacological characterization of phospholipase A₂ and metalloproteinase fractions from Eastern Russell's viper (*Daboia siamensis*) venom: Two major components associated with acute kidney injury. *Toxins* **2021**, *13*, 521. [[CrossRef](#)]
55. Tsai, I.-H.; Tasi, H.-Y.; Wang, Y.-M.; Pe, T.; Warrel, D.A. Venom phospholipases of Russell's vipers from Myanmar and eastern India—Cloning, characterization and phylogeographic analysis. *Biochim. Biophys. Acta.* **2007**, *1774*, 1020–1028. [[CrossRef](#)]
56. Thwin, M.-M.; Gopalakrishnakone, P.; Yuen, R.; Tan, C.H. A major lethal factor of the venom of Burmese Russell's viper (*Daboia russellii siamensis*): Isolation, n-terminal sequencing and biological activities of daboiatoxin. *Toxicon* **1995**, *33*, 63–76. [[CrossRef](#)]
57. Zhang, Y.; Li, W.; Gao, R.; Lu, Q.; Wang, W.; Xiong, Y. Structure-function features and effects on blood coagulation of snake venom serine proteases. *Zool. Res.* **1998**, *19*, 97–106.
58. Kang, T.S.; Georgieva, D.; Genov, N.; Murakami, M.T.; Sinha, M.; Kumar, R.P.; Kaur, P.; Kumar, S.; Dey, S.; Sharma, S.; et al. Enzymatic toxins from snake venom: Structural characterization and mechanism of catalysis. *FEBS J.* **2011**, *278*, 4544–4576. [[CrossRef](#)]
59. Serrano, S.M.T.; Maroun, R.C. Snake venom serine proteinases: Sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon* **2005**, *45*, 1115–1132. [[CrossRef](#)]
60. Thakur, R.; Mukherjee, A.K. A brief appraisal on Russell's viper venom (*Daboia russellii russellii*) proteinases. In *Snake Venoms*; Inagaki, H., Vogel, C.W., Mukherjee, A., Rahmy, T., Eds.; Springer: Dordrecht, The Netherlands, 2017; pp. 123–144. [[CrossRef](#)]
61. Gâz Florea, Ș.A.; Gâz Florea, A.; Kelemen, H.; Muntean, D.-L. Snake Venom Metalloproteinases. *Acta Medica Marisicensis* **2016**, *62*, 106–111. [[CrossRef](#)]
62. Takeda, S.; Takeya, H.; Iwanaga, S. Snake venom metalloproteinases: Structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochim. Biophys. Acta Proteins Proteom.* **2012**, *1824*, 164–176. [[CrossRef](#)] [[PubMed](#)]
63. Chellapandi, P. Structural, functional and therapeutic aspects of snake venom metalloproteinases. *Mini. Rev. Org. Chem.* **2014**, *11*, 28–44. [[CrossRef](#)]
64. Takeda, S.; Igarashi, T.; Mori, H. Crystal structure of RVV-X: An example of evolutionary gain of specificity by ADAM proteinases. *FEBS Lett.* **2007**, *581*, 5859–5864. [[CrossRef](#)] [[PubMed](#)]
65. Chen, H.S.; Chen, J.M.; Lin, C.W.; Khoo, K.H.; Tsai, I.H. New insights into the functions and N-glycan structures of factor X activator from Russell's viper venom. *FEBS J.* **2008**, *275*, 3944–3958. [[CrossRef](#)] [[PubMed](#)]
66. Suntravat, M.; Yusuksawad, M.; Sereemasapun, A.; Pérez, J.C.; Nuchprayoon, I. Effect of purified Russell's viper venom-factor X activator (RVV-X) on renal hemodynamics, renal functions, and coagulopathy in rats. *Toxicon* **2011**, *58*, 230–238. [[CrossRef](#)]
67. Chen, H.S.; Tsai, H.Y.; Wang, Y.M.; Tsai, I.H. P-III hemorrhagic metalloproteinases from Russell's viper venom: Cloning, characterization, phylogenetic and functional site analyses. *Biochimie* **2008**, *90*, 1486–1498. [[CrossRef](#)] [[PubMed](#)]
68. Takeda, S. ADAM and ADAMTS family proteins and snake venom metalloproteinases: A structural overview. *Toxins* **2016**, *8*, 155. [[CrossRef](#)]
69. Hikita, A.; Tanaka, N.; Yamane, S.; Ikeda, Y.; Furukawa, H.; Tohma, S.; Suzuki, R.; Tanaka, S.; Mitomi, H.; Fukui, N. Involvement of a disintegrin and metalloproteinase 10 and 17 in shedding of tumor necrosis factor- α . *Biochem. Cell Biol.* **2009**, *87*, 581–593. [[CrossRef](#)] [[PubMed](#)]
70. Yamazaki, Y.; Matsunaga, Y.; Tokunaga, Y.; Obayashi, S.; Saito, M.; Morita, T. Snake venom vascular endothelial growth factors (VEGF-Fs) exclusively vary their structures and functions among species. *J. Biol. Chem.* **2009**, *284*, 9885–9891. [[CrossRef](#)]
71. Ferreira, I.G.; Pucca, M.B.; de Oliveira, I.S.; Cerni, F.A.; de Cássia da Silva Jacob, B.; Arantes, E.C. Snake venom vascular endothelial growth factors (svVEGFs): Unravelling their molecular structure, functions, and research potential. *Cytokine Growth Factor Rev.* **2021**, *60*, 133–143. [[CrossRef](#)] [[PubMed](#)]

72. Yamazaki, Y.; Takani, K.; Atoda, H.; Morita, T. Snake venom vascular endothelial growth factors (VEGFs) exhibit potent activity through their specific recognition of KDR (VEGF receptor 2). *J. Biol. Chem.* **2003**, *278*, 51985–51988. [[CrossRef](#)]
73. Tokunaga, Y.; Yamazaki, Y.; Morita, T. Specific distribution of VEGF-F in Viperinae snake venoms: Isolation and characterization of a VEGF-F from the venom of *Daboia russelli siamensis*. *Arch. Biochem. Biophys.* **2005**, *439*, 241–247. [[CrossRef](#)]
74. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [[CrossRef](#)] [[PubMed](#)]
75. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)]
76. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [[CrossRef](#)] [[PubMed](#)]
77. Macrander, J.; Panda, J.; Janies, D.; Daly, M.; Reitzel, A.M. Venomix: A simple bioinformatic pipeline for identifying and characterizing toxin gene candidates from transcriptomic data. *Peer J.* **2018**, *6*, e5361. [[CrossRef](#)]
78. Jungo, F.; Bairoch, A.; Prot, T. the toxin protein annotation program of the Swiss-Prot protein knowledgebase. *Toxicon* **2005**, *45*, 293–301. [[CrossRef](#)] [[PubMed](#)]
79. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-Seq: Reference generation and analysis with Trinity. *Nat Protoc.* **2013**, *8*, 8. [[CrossRef](#)]
80. Gacesa, R.; Barlow, D.J.; Long, P.F. Machine learning can differentiate venom toxins from other proteins having non-toxic physiological functions. *PeerJ Comput. Sci.* **2016**, *2*, e90. [[CrossRef](#)]
81. UniProt Consortium. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [[CrossRef](#)] [[PubMed](#)]
82. Madeira, F.; Pearce, M.; Tivey, A.R.N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50*, W276–W279. [[CrossRef](#)]
83. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
84. Bembom, O.; Ivanek, R. seqLogo: Sequence logos for DNA sequence alignments. R package version 1.64.0. Available online: <https://bioconductor.org/packages/release/bioc/html/seqLogo.html> (accessed on 26 March 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.