# Venom Proteomes of Closely Related *Sistrurus* Rattlesnakes with Divergent Diets

Libia Sanz, H. Lisle Gibbs, Stephen P. Mackessy, and Juan J. Calvete*,

Instituto de Biomedicina de Valencia, C.S.I.C., Jaume Roig 11, 46010 Valencia, Spain, Department of Evolution, Ecology and Organismal Biology, Ohio State University, 300 Aronoff Laboratory, 318 West 12th Avenue, Columbus, Ohio 43210-1293, and School of Biological Sciences, University of Northern Colorado, 501 20th Street, CB 92, Greeley, Colorado 80639-0017

[Abstract text illegible]
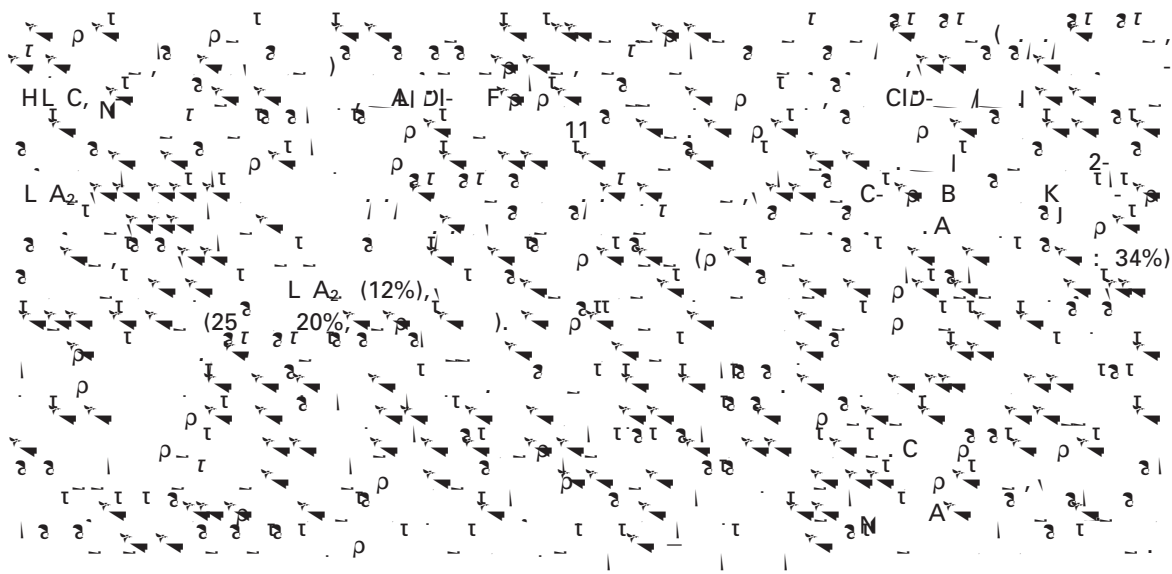
Keywords:

selection for functional diversity both within and between species levels.[15−27] Second, analysis of venom proteins using gel electrophoresis has consistently shown relatively high levels of intra- and interspecific variation.[28] Third, there is a small but increasing number of studies that strongly support the idea that this variation reflects adaptation for differential utilization of distinct prey types.[29−32] However, although the notion that evolutionary interactions between snakes and their prey may be responsible for variation in venom composition has been controversial,[33,34] the recent demonstration of taxa-specific effects of some venoms supports the hypothesis that evolutionary adjustments in venom composition have occurred.[32] The loss of a functional major toxin (α-neurotoxin) from the venom of *Aipysurus eydouxii*, which appears to have occurred following a shift from feeding on fish to fish eggs,[35] further supports a link between diet and venom composition. Of particular value in addressing this issue would be studies that adopt a comprehensive approach of assessing variation in venom proteins[4,5,36] in closely related species that show clear, significant differences in diet. The use of phylogenetically similar species would allow the strong inference that differences in venom characteristics are most likely due to differences in selection pressures alone rather than an unknown combination of the effects of selection and divergence due to phylogenetic relationships among the species being compared.

To address the need for detailed proteomic studies of

Table 1.

HPLC fraction SCC-		N-terminal sequencing	Isotope-averaged MALDI-TOF mass (± 0.2%)	peptide ion m/z	z	MS/MS-derived sequence	protein family
1		n.p.					
2		n.p.					
3		EAGEECDCGSPANPCCDAAT					disintegrin
		EECDCGSPANPCCDAATCKL					disintegrin
4		Blocked	4738 (5386*)	621.9	3	XCSPPYSDVGQXDCR	myotoxin
5		EECDCGSPANPCCDAATCKL	7069 (8340*)	743.6	3	XRPGAQCADGLCCDQCR	disintegrin
6		EAGEECDCGSPANPCCDAAT	[7069, 7187,				disintegrin
		AGEECDCGSPANPCCDAATC	7316, 7470]				disintegrin
		EECDCGSPANPCCDAATCKL					disintegrin
7		N. D.	22806				
		GCYCGTGGQGWPQDASDRCCFE		1147.6	2	(575.2)VYEAEDSCFESNQK	DC-fragment
8		[ ]	9212				2-chain PLA$_2$
		SLENCQGESQPC					
		GCYCGTGGQGWPQDASDRCCFE	[9044, 9309,				
9		[ ]	9577]				2-chain PLA$_2$
		SLENCQGESQPC					
10		N. D.	12961	744.7	3	GKPXDATDRCCFVHDCCGK	PLA$_2$
11		N. D.	15, 17 kDa§	711.9	2	AXTMEDNEASWR	nerve growth factor
				640.3	2	XDSACVCVXSR	
				556.3	2	NPNPVPTGCR	
				983.8	3	GNXVTVMVDVNXNNNVYKQY-FFETK	
12		NLLQFNKMIKIMTKK	13814				N6−PLA$_2$
13		NLLQFNKMIKIMTKK	13880				N6−PLA$_2$
14		HLLQFNKMIKFETNK	14120				N6−PLA$_2$
15		SVDFDSESPPKPEIQ	24831				CRISP
16	M)	HLIQFETLIMKIAGR	13967				PLA$_2$
	m)	Blocked	22206	753.8	2	CCFVHDCCYGK	PLA$_2$
17		HLIQFETLIMKIAGR	13952				PLA$_2$
18	M)	N. D.	13941	655.3	2	SXVQFETXXMK	PLA$_2$
				837.3	2	XTGCDPXTDVYTYR	
	m)	N. D.	28−30 kDa§	582.8	2	XMGWGTXSATK	Ser-proteinase
				595.9	2	WDKDXMXXR	
				648.3	2	NNXXDYEVCR	
19	m)	N. D.	13941	655.3	2	SXVQFETXXMK	PLA$_2$
				837.3	2	XTGCDPXTDVYTYR	
	M)	IIGGDECNINEHRFL	28−30 kDa§	582.8	2	XMGWGTXSATK	Ser-proteinase
				595.9	2	WDKDXMXXR	
				648.3	2	NNXXDYEVCR	
20		VIGGDECNINEHRSL	27090				Ser-proteinase
21		IIGGDECNINEHRFL	30918				Ser-proteinase
22		VVGGDECNINEHRFL	[26545, 28556]				Ser-proteinase
23		VVGGDECNINEHRSL	[27115, 27323]				Ser-proteinase
24		APEHQRYVELFIVVD	23053	555.8	2	TXNSFGEWR	PI-metalloprotease
				874.9	2	VAVTMTHEXGHNXGNR	
				900.9	2	YVEFVVVXDHGMYTK	
25		Blocked	33 kDa§	664.8	2	YXEXVXVADHR	metalloprotease
				840.8	2	VHEXVNFXNEFYR	
				821.4	2	HDNAQXXTAXDFQR	
				660.9	3	(233.2)EQHNPQCXXNKPXR	
			52 kDa§	627.3	2	SAGQXYEESXR	L-AMINO acid oxidase
				647.8	2	EGWYANXGPMR	
				744.8	2	EDTYEEFXEXAK	
				881.4	2	EGWYANXGPMRLPEK	
26		LTPEQQAYLDAKKYV	32 kDa§				

§2527 ([(15)-5692(SVDFDSESPPKoXR)]TJ-9)-29V. 655 49047
2NEXYXPXN-278(kDa)]TJ4.989 0 0 4.9891ND(-)Tjen5|TJ(2NEDa)]TJ4.989 0 0 4.989 267.352 310.603228.T*2527 ([(15)-5692(SVDFDSESPPKoXR)]TJ-9)-29V.
2NEXYXPXN-248(kDa)]TJ4.989 0 0 4.9891ND(-)Tjen5|T15.155aTX(640.3)(APEHQRYVELFIVVD)-7555(23053(SVDFDSESPPKoXR)]TJ-9)-29V. 655 49047

massasauga) (Q6EER2 and Q6EER3, respectively); and an N6

Table 2. [illegible]

HPLC fraction SCT-	N-terminal sequencing	isotope-averaged MALDI-TOF mass (± 0.2%)	peptide ion m/z	z	MS/MS-derived sequence	protein family
1−4	n.p.					
5	EECDCGSPANPCCDAAT	7579				disintegrin
6	EECDCGSPANPCCDAAT	7185				disintegrin tergeminin (P22828) 4−71
	M: SGMFSYSAYGCYCGWGG	11648				PLA_2
	m: Blocked	4737 (5384*)	621.9	3	XCSPPYSDVGQXDCR	myotoxin
7	GEECDCGSPANPCCDA EECDCGSPANPCCDAA	[7658, 7459]				disintegrin tergeminin (P22828)
8	GEECDCGSPANPCCDA EECDCGSPANPCCDAA	[7073, 7346, 7402]				disintegrin tergeminin (P22828) 4−70, 4−72, 3−72
9	EAGEECDCGSPANPCCDA	7438				disintegrin tergeminin (P22828) 1−71
10	SPNECGNNFVDLGEECDCGLPANP	8495				disintegrin
11	SVGEECDCGTPENDQ	11352				disintegrin-like
		23 kDa§	527.2	2	GNYYGYCR	DC-fragment
			761.2	3	(608.5)YEAEDSCFEPTXR	
12	GCYCGTGVQGWPQDASD [ ] SLENCQGESQP	[8628, 8703]				2-chain PLA_2
13	GCYCGTGVQGWPQDASD [ ] SLENCQIESQPC	9245				2-chain PLA_2
		23 kDa§	527.2	2	GNYYGYCR	DC-fragment
14	GCYCGTGGQGWPQDASD [ ] SLENCQGESQPC	9025				2-chain PLA_2 (Q6EAN6) (41−112)-S−S−(127−138)
15	N. D.	13 kDa§	744.7	3	GKPXDATDRCCFVHDCCGK	PLA_2
16	N. D.	16 kDa§				

Table 3. ... MALDI-... ...

HPLC fraction SCE-	N-terminal sequencing	isotope-averaged MALDI-TOF mass ($\pm$ 0.2%)	peptide ion m/z	z	MS/MS-derived sequence	protein family
1−6	n. p.					
7	GEECDCGSPANPCCD	[7423, 7631]				disintegrin
8	SPPVCGNKILEQGEDCDCGSP -ANCQDRCCNAATCKLTPGSQ	8960				disintegrin (bitistatin-like)
9,10	NLIQFETLILKVAKK	13832				PLA$_2$
11	N. D.	8960	534.3	2	FETPEECR	Kunitz protease inhibitor
12	GSGCFGLKLDRIGSMSGLGC	1956.0¶				C-type BPP
		27 + 30 kDa§	534.7	2	GXCCDQCR	disintegrin-like
			527.7	2	NGHPCXNNK	
			813.8	2	NQCXSFFGPSATVAK	cysteine-rich
			823.0	3	NNCNVXYTPTDEDXGmVXPGTK	
			777.3	2	VCSNGHCVDVATAY	
13−17					SVGEECDCGSPNTE	disintegrin-like
		23 kDa§	588.2	2	NFPCAQPDVK	cysteine-rich
			530.6	2	QGDNFYCR	cysteine-rich
		25 kDa§	679.3	2	VQKGQGVYYCR	cysteine-rich
			774.9	2	SPPNDPDFGFVSR	cysteine-rich
19	N. D.	33 kDa§	583.8	2	XGNYYGYCR	cysteine-rich
20,21	Blocked	16 kDa§	556.8	2	NPNPVPTGCR	nerve growth factor
			690.9	2	AXTmEGNQASWR	
			640.9	2	XDSACVCVXSR	
22	N. D.	23211	588.2	2	NFPCAQPDVK	cysteine-rich
			530.6	2	QGDNFYCR	
			626.3	2	XYNDNXNPCK	
23	NLLQFET	13856				PLA$_2$
		16 kDa§	708.2	2	(217)XPFmEVYQR	nerve growth factor
24	SVDFDSESPRKPEIQ	24829				CRISP
25	NLIQFETLILKVAKK	13842				PLA$_2$
26	IIGGEECNINEHRFL	[26202, 28742]				Ser-proteinase
27	VVGGEECNINEHRSL	29 kDa§				Ser-proteinase
28	(V+27 VVGGEECNI7260FL 28742]					Ser-protef2ase

Despite the fact that only six *Sistrurus* venom protein entries are annotated in the Swiss-Prot/TrEMBL nonredundant database, HPLC fractions that yielded unambiguous N-terminal sequences could be classified into known protein families using a BLAST amino acid similarity search (Tables 1−4), indicating that representative members of each of these families are

Table 5.

	venom			
protein family	SCC	SCT	SCE	SMB
disintegrin	2.5	4.2	0.9	7.7
myotoxin	0.4	<0.1	-	-
C-type BPP	-	—	<0.1	< 0.1
Kunitz-type inhibitor	-	—	<0.1	0.1
DC-fragment	<0.1	<0.1	<0.1	1.3
2-chain PLA$_2$	2.5	1.9	-	-
nerve growth factor	<0.1	<0.1	<0.1	<0.1
N6—PLA$_2$	12.8	14.9	—	13.9
PLA$_2$	14.6	14.8	13.7	18.6
CRISP	0.8	1.3	10.7	2.9
serine proteinase	18.2	20.4	24.4	17.1
C-type lectin	<0.1	<0.1	<0.1	<0.1
L-AMINO acid oxidase	4.2	1.6	2.5	2.1
Zn^{2+}-metalloproteinase	43.8	40.6	48.6	36.1

[a] In percentage of the total HPLC-separated proteins.

659.3^{2+} (YTXNSFGEWR) in different protein bands of fractions SCC-26, 27, 29, and 31; and 629.3^{2+} (YTXNAFGEWR), 784.3^{2+} (SAAADTXQEFGDWR), and 912.8^{2+} (VTXSADDTXQAFAEWR) in SCT-29−32, 34 and 35) further indicates that the corresponding parent molecules were isoforms. It is also noteworthy that among 93 unique Zn^{2+}-metalloprotease-derived tryptic peptide ions sequenced by MS/MS, very few are common to any pair of the *Sistrurus* snakes under study. SCC and SCT share the largest number (6 out of 45) of identical tryptic peptide ions, but this value is small (13%). The data thus indicate that Zn^{2+}-metalloproteases have evolved intra- and interspecifically in an accelerated manner such that each venom contains a distinct set of metalloprotease isoenzymes.

Table 6 shows conservation and divergence among PLA$_2$ molecules identified in the different *Sistrurus* venoms. The fact that a number of *Sistrurus* share apparently identical PLA$_2$ molecules could reflect recent common ancestry and/or convergence in aspects of their diets (Figure 1). However, the occurrence of distinct PLA$_2$ molecules in each *Sistrurus* venom is consistent with an accelerated and regional evolution of PLA$_2$ isozymes, similar to the diversity documented for PLA$_2$ isoforms of *Trimeresurus* species inhabiting the south-western islands of Japan[24,31,60] and Taiwan,[61] and the birth-and death model for the evolution of elapid three-finger toxins.[62]

The venoms of SCC and SCT also appear to be rather similar, based on the observations that (i) myotoxins and 2-chain PLA$_2$ molecules are detected only in SCC and SCT venoms and (ii) the major PIII-metalloproteases of SCC (HPLC peak 29; 48877 Da) (Table 1) and SCT (HPLC peak 33; 48861 Da) (Table 2) and the PI-metalloproteases SCC-24 (23053 Da) and SCT-27 (23082 Da) exhibit almost identical chromatographic behavior (Figure 6) and mass tryptic peptide fingerprinting. Again, this similarity could be due to either convergent selection or common ancestry—convergence appears more likely, because if common ancestry was of overriding importance, then the venom of SCT should be more similar to a member of the same clade (SCE) than to SCC, but this relationship does not hold. Alternatively, divergence of SCE metalloproteases from a shared ancestral condition could give rise to the observed similarities in metalloproteases.
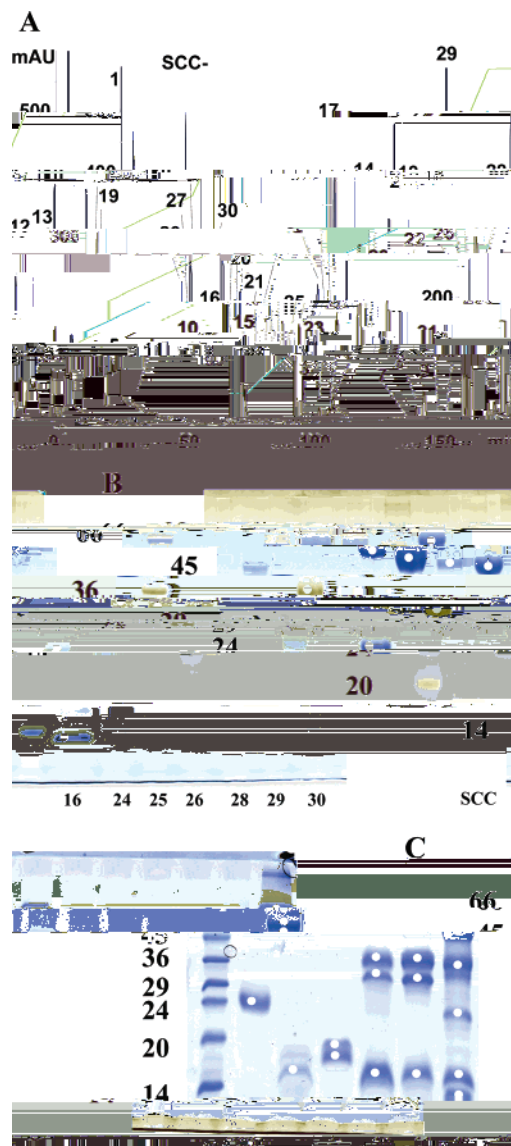
Figure 2. (A) [...] (B) [...] (C) [...]

Broad-Scale Patterns of Venom Diversity and Differentiation. The availability of detailed proteomic information on individual proteins described above make possible detailed estimates of the similarity and differentiation of the venom proteomes of different taxa that are based on comparisons of individual proteins which are then useful in revealing broad-scale evolutionary patterns. Using one measure of overall venom protein diversity (numbers of distinct proteins; see Table 7), the *Sistrurus* taxa sampled fall roughly into two groups: three species with relatively diverse venoms (SCC, 41.5; SCT,

44.0; and SCE, 35.5 distinct proteins, respectively) and SMB, which is substantially less variable[24] (Figure 7). When mapped onto a phylogeny of these taxa, this pattern points to a key event occurring early on in the evolution of the group involving the lineage connecting *S. milarius* to the other taxa (Figure 7).
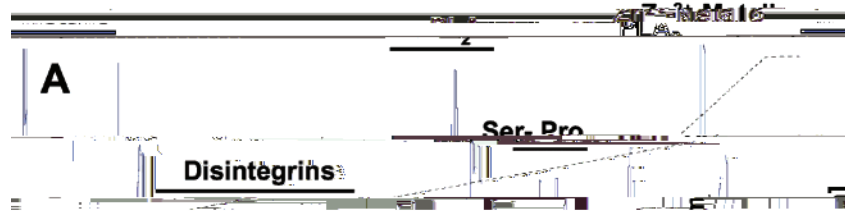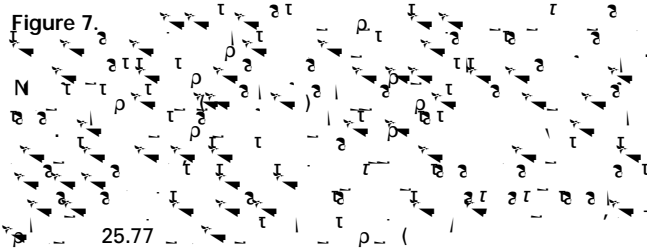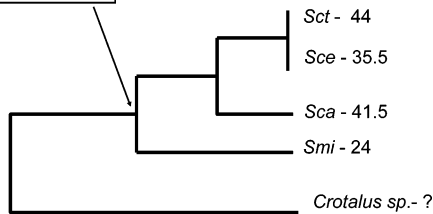
A

Disintegrins

Ser. Pro

Metal6

Figure 6. C ...(C) ... (D) ... (A) ..., (B) ...

Loss or gain of venom diversity

Sct - 44
Sce - 35.5
Sca - 41.5
Smi - 24
Crotalus sp.- ?

Figure 7.

N ... 25.77 ... (

related yet ecologically distinct venomous snakes. If we assume a link between structural and functional variation in terms of effectiveness at killing and processing different prey,[29,32,41] then our results have implications for how venom has evolved as an adaptation in these snakes.

First, the finding of the high degree of differentiation in the venom proteome among recently evolved, congeneric taxa emphasizes unique aspects of venom composition of even closely related species of venomous snakes and points to a strong role for adaptive diversification via natural selection as a cause of this distinctiveness. Moreover, the high level of within-taxon variation in almost all proteins suggests an important role for balancing selection[63] in maintaining high levels of functional variation in venom proteins within populations. The mechanism leading to this mode of selection is unclear, but we speculate that it may be related to unpredict-

ability with which a sit-and-wait predator like a rattlesnake encounters different types of prey, each of which are most efficiently subdued with different venom proteins. Thus, to deal with this uncertainty, snakes are required to have a variety of proteins "available" in their venom at all times to deal with different prey. Differential effects of venom to specific prey types (= taxa-specific toxicity) can be extreme,[32] demonstrating the clear functional link between venom composition and effects on prey. However, prey physiological responses are not static, and among mammalian prey in particular, selection for resistance mechanisms may be profound.[64,65] In a sense, the selection pressure leading to high levels of variation in venom genes, namely the capacity for the evolution of detoxifying responses by prey, may parallel the selection pressures acting

to promote high levels of variation in the genes involved in the vertebrate immune system, such as those which encode major histocompatibility complex proteins,[66] or in plant host defense genes.[67,68] Various aspects of this hypothesis could be tested by directly examining patterns of allelic variation in specific venom genes to see if they show molecular signatures of balancing selection at the DNA level,[67,69] by assaying purified components to determine if different venom components of the same protein type act more efficiently on different prey, and by assessing the predictability of the diets of individual snakes through time.

A final implication of our results is that there does not appear to be a simple relationship between levels of venom variation and diet diversity. Because species-specific effects of venom components are largely unknown, it is difficult to assign a functional role unequivocally to the variation we observed in *Sistrurus* venoms. However, if one considers relative importance of mammals in the diet of *Sistrurus*, diet trends and complexity trends are parallel and show the following order, from high to low: SCC > SCT > SCE > SMB. Both SCC and SCT include a much greater proportion of mammals in their diets than does SCE,[39] and SMB rarely feed on mammals in the wild.[40,70] It is possible that increased reliance on mammalian prey has driven selection for greater venom protein diversity, and the presence of myotoxin-a homologues and 2-chain PLA_2s *only* detected in SCC and SCT venoms is consistent with this hypothesis. The very high level of CRISP proteins in the venom of SCE, relative to the other taxa, is intriguing, as this toxin appears to be a component of virtually *all* venoms and therefore became incorporated into the venom proteome early in the evolution of venom systems;[13] unfortunately, the biological activity of most venom CRISPs are currently unknown.[71] It is clear that venom composition and diet are related, but because various species have mixed diets, including both endotherms and ectotherms, invertebrates and vertebrates, the diet/composition relationship is likely rather complex. We feel the key to understanding the relationship between diet and

(21) Deshimaru, M.; Ogawa, T.; Nakashima, K. I.; Nobuhisu, I.; Chijiwa, T.; Shimohigashi, Y.; Fukumaki, Y.; Niwa, M.; Yamashina, I.; Hattori, S.; Ohno, M. Accelerated evolution of crotaline snake venom gland serine proteases. *FEBS Lett*

(62) Fry, B. G.; Wüster, W.; Kini, R. M.; Brusic, V.; Khan, A.; Venkat-araman, D.; Rooney, A. P. Molecular evolution of elapid snake venom three-finger toxins. *J. Mol. Evol.* **2003**, *57*, 110−129.

(63) Richman, A. Evolution of balanced genetic polymorphism. *Mol. Ecol.* **2000**, *9*, 1953−1963.

(64) Perez, J. C.; Pichyangkul, S.; Garcia, V. E. The resistance of three species of warm-blooded animals to western diamondback rattlesnake (*Crotalus atrox*) venom. *Toxicon* **1979**, *17*, 601−607.

(65) Biardi, J. E.; Chien, D. C.; Coss, R. G. California ground squirrel (*Spermophilus beecheyi*) defenses against rattlesnake venom digestive and hemostatic toxins. *J. Chem. Ecol.* **2006**, *32*, 137−154.

(66) Hedrick, P. W.; Kim, T. Genetics of complex polymorphisms: parasites and maintenance of MHC variation. In *Evolutionary Genetics: From Molecules to Morphology*; S6a9d2cp275(S6R.p275(S6Krimbasp275(S6C.p2J-10.4886 0 .1333 TD[(GeEds)-2836S6Cambridge2836S6Une)-rtesy283