Molecular Evolution, Functional Variation, and Proposed Nomenclature of the Gene Family That Includes Sphingomyelinase D in Sicariid Spider Venoms

Greta J. Binford, Melissa R. Bodner,*[1] Matthew H.J. Cordes,† Katherine L. Baldwin,*[2] Melody R. Rynerson,*[3] Scott N. Burns,‡ and Pamela A. Zobel-Thropp**

*Department of Biology, Lewis and Clark College, Portland, OR; †Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ; and ‡Cleveland High School, Portland, OR

The venom enzyme sphingomyelinase D (SMase D) in the spider family Sicariidae (brown or fiddleback spiders [*Loxosceles*] and six-eyed sand spiders [*Sicarius*]) causes dermonecrosis in mammals. SMase D is in a gene family with multiple venom-expressed members that vary in functional specificity. We analyze molecular evolution of this family and variation in SMase D activity among crude venoms using a data set that represents the phylogenetic breadth of *Loxosceles* and *Sicarius*. We isolated a total of 190 nonredundant nucleotide sequences encoding 168 nonredundant amino acid sequences of SMase D homologs from 21 species. Bayesian phylogenies support two major clades that we name α and β, within which we define seven and three subclades, respectively. Sequences in the α clade are exclusively from New World *Loxosceles* and *Loxosceles rufescens* and include published genes for which expression products have SMase D and dermonecrotic activity. The β clade includes paralogs from New World *Loxosceles* that have no, or reduced, SMase D and no dermonecrotic activity and also paralogs from *Sicarius* and African *Loxosceles* of unknown activity. Gene duplications are frequent, consistent with a birth-and-death model, and there is evidence of purifying selection with episodic positive directional selection. Despite having venom-expressed SMase D homologs, venoms from New World *Sicarius* have reduced, or no, detectable SMase D activity, and *Loxosceles* in the Southern African *spinulosa* group have low SMase D activity. Sequence conservation mapping shows >98% conservation of proposed catalytic residues of the active site and around a plug motif at the opposite end of the TIM barrel, but α and β clades differ in conservation of key residues surrounding the apparent substrate binding pocket. Based on these combined results, we propose an inclusive nomenclature for the gene family, renaming it *SicTox*, and discuss emerging patterns of functional diversification.

Introduction

Spider venoms are complex mixtures of hundreds of proteins, peptides, and low-molecular-weight components. The composition of venom varies widely across species but includes cytotoxins, neurotoxins with specific neurophysiological targets, and antimicrobial components (reviews in Schulz 1997; Rash and Hodgson 2002; Kuhn-Nentwig 2003; Adams 2004; Tedford et al. 2004; Escoubas 2006; Estrada et al. 2007; King 2007). Much recent research has focused on the rich potential in spider venoms for discovery of novel toxic activities that may have pharmacological utility (recent review in Escoubas et al. 2008), but relatively little work has empirically analyzed evolutionary mechanisms that have influenced spider venom diversity (see Kordis and Gubensek 2000; Diao et al. 2003; Sollod et al. 2005; Escoubas 2006). The general types of toxins in spider venoms are similar to those in other animals that have independently evolved venom for prey capture, such as cone snails (reviews in Duda and Palumbi 2000; Espiritu et al. 2001; Olivera 2002), snakes (Fry et al. 2008), and scorpions (Rodríguez de la Vega and Possani 2004, 2005). Work on these toxins has uncovered interesting evolutionary patterns and mechanisms including accelerated evolution and hypermutation mechanisms (e.g., Kini and Chan 1999; Duda and Palumbi 2000; Espiritu et al. 2001; Calvete et al. 2005; Lynch 2007), diversification via birth-and-death processes (Fry et al. 2003; Li et al. 2005), toxin gene recruitment from a broad range of protein families (Fry 2005), and evolution of expression patterns of toxin genes (Duda and Remigio 2008). Comparably focused work on spider venom toxins is likely to discover similarly interesting evolutionary dynamics. Furthermore, understanding evolutionary dynamics of venom toxins should help with focusing bioprospecting efforts, understanding the distribution of taxa with dangerous bites, and developing treatments that are effective for the phylogenetic breadth of taxa with clinically important toxins.

Here, we present an evolutionary analysis of the gene family that includes the venom toxin sphingomyelinase D (SMase D). SMase D is expressed in venoms of *Loxosceles* (brown or violin spiders) and their sister genus *Sicarius* (six-eyed sand spiders). These two genera are supported by morphology to be each other's closest relatives, and they are the only two taxa in the family Sicariidae (Platnick et al. 1991). *Loxosceles* are famous for bites that cause dermonecrotic lesions in mammalian tissues (recent reviews in da Silva et al. 2004; Swanson and Vetter 2005; Vetter 2008), and venoms of some *Sicarius* species also cause dermonecrosis (Newlands and Atkinson 1988; Van Aswegen et al. 1997). The venom-expressed enzyme SMase D has been demonstrated to be a sufficient causative agent for lesion formation (Kurpiewski et al. 1981; Tambourgi et al. 1998, 2004; Fernandes-Pedrosa et al. 2002; Ramos-Cerrillo et al. 2004). Multiple members of the gene family that includes SMase D are expressed in individual venoms (Tambourgi et al. 1998; de Castro et al. 2004; Ramos-Cerrillo et al. 2004; Binford et al. 2005; Machado et al. 2005; Kalapothakis et al. 2007; Fernandes-Pedrosa et al. 2008), and these paralogs differ in substrate specificity (Tambourgi et al. 1998; Ramos-Cerrillo et al. 2004; Lee and Lynch 2005; Kalapothakis et al. 2007). A structure

[1] Present address: Department of Zoology, University of British Columbia, Vancouver, BC, Canada.
[2] Present address: Department of Cellular and Molecular Biology, University of Wisconsin, Madison, WI.
[3] Present address: Department of Genome Sciences, University of Washington, Seattle, WA.

Key words: *Loxosceles*, *Sicarius*, duplication, toxin.

E-mail: binford@lclark.edu.

Mol. Biol. Evol. 26(3):547–566. 2009
doi:10.1093/molbev/msn274
Advance Access publication November 28, 2008

has been solved for a member of this gene family (PDB 1XX1) and active sites proposed (Murakami et al. 2005). The enzyme is an $(\alpha/\beta)_8$ barrel and apparently depends upon binding of a Mg^{2+} ion for catalysis. This set of homologous genes has been referred to as "SMase D" in the literature, but the fact that known venom-expressed homologs do not have SMase D activity has led Kalapothakis et al. (2007) to propose the broader name *LoxTox* for the gene family.

Many characteristics of the *LoxTox* gene family suggest it has had an interesting evolutionary history. SMase D is a highly derived member of the ubiquitous glycerophosphodiester phosphodiesterase (GDPD) protein domain family (Binford et al. 2005; Cordes and Binford 2006; Murakami et al. 2006). Comparative analyses have not detected SMase D activity in venoms outside of sicariid spiders, which is consistent with a single evolutionary origin of SMase D as a venom toxin in the most recent common ancestor (MRCA) of this lineage (Binford and Wells 2003). Homologous SMase Ds are expressed as an exotoxin in a few *Corynebacteria* (Soucek et al. 1967; Bernheimer et al. 1985; Truett and King 1993). A C-terminal structural motif that is unique to and shared by spider and bacterial SMase Ds is evidence that this disparate distribution of SMase D may be explained by a lateral transfer event (Cordes and Binford 2006). The recent discovery of a homologous gene expressed in tick saliva (Accession DQ411855) is consistent with the gene family having an ancient presence in arachnids and implies that a horizontal transfer event would have originated from an arachnid and moved into *Corynebacteria*.

Despite the interesting emerging picture of evolutionary phenomena, detailed mechanisms influencing diversification of the *LoxTox* gene family require comparative analyses with more dense taxon sampling within sicariids. To date, a detailed analysis has been hampered by limited taxon sampling and lack of understanding of species relationships in sicariids. The approximately 100 described species of *Loxosceles* are native to the Americas, Africa, and the Mediterranean (Gertsch and Ennik 1983), and the synanthropic species *L. rufescens* has colonized many other locations. Twenty-three described species of *Sicarius* are native to Africa and Central and South America (Gerschman and Schiapelli 1979; Platnick 2008). *LoxTox* gene family members have been isolated from three North American *Loxosceles* species from the *reclusa* species group and four species from South America that represent three species groups (summary and references in table 1). There are no published homologs from Old World *Loxosceles* or *Sicarius*.

Our goal was to analyze patterns of variation and mechanisms of molecular evolution of venom-expressed members of the *LoxTox* gene family in sicariid spiders. We use phylogenetic analyses, structural modeling of amino acid conservation, and analyses of positive selection of a data set of sequences of venom-expressed members of this gene family available from this and previous work. We complement phylogenetic analyses with comparative protein separations and enzyme activity assays of whole crude venoms. Our data set includes representatives from all identified species groups of *Loxosceles* and *Sicarius* with taxon sampling guided by recent species-level molecular systematic work on members of this lineage (Binford et al. 2008). Based on our results, we propose renaming this gene family *SicTox*, abbreviated from Sicariidae toxin, to accommodate the currently known breadth of species that express members of this gene family in venoms. We also propose an inclusive, phylogenetically based nomenclature with the goal of standardizing the language to facilitate efficient and unambiguous discussion of this gene family in the literature.

Materials and Methods

Taxonomic Sampling and Collection

The representation of taxa in different analyses is summarized in table 2. Our goal was to capture the breadth of diversity within Sicariidae, guided by previous phylogenetic work (Binford et al. 2008, fig. 1). To do this, we include at least one representative of every described species group in *Loxosceles*, representatives of each geographic region to which *Sicarius* is native, and the most appropriate and available outgroup for each particular analysis (table 2). Spiders in the genus *Drymusa* are putative close relatives of sicariids that have been previously shown to not express SMase D in their venom (Binford and Wells 2003).

All spiders were collected in the field by G.J.B. and colleagues. Details of collecting localities are available from G.J.B. by request. We restricted analyses to mature individuals to allow for proper species-level confirmation using morphology. We also retained legs of spiders for genomic DNA isolation to help with confirmation of species status. If animals were not mature when collected, we reared them to maturity in the laboratory at 35% humidity, 24 °C. Species-level systematics of some taxa we include is unclear, and their taxonomic nomenclature is being revised. This is particularly true of African sicariids. For this work, we follow the same nomenclatural framework used in Binford et al. (2008). Voucher specimens are maintained in the collection of G.J.B. and will be submitted to the California Academy of Sciences, and duplicates from the same populations will be sent to National Museums from the countries of origin upon completion of our work.

SMase D Homolog Sequencing and Analyses

Venom Tissue Gland Isolation and cDNA Synthesis

We used electrical stimulation to extract venom from all specimens used for cDNA analysis (table 2) as described in Binford and Wells (2003). Two to three days later, a time when we have previously isolated SMase D mRNAs (Binford et al. 2005), we anesthetized spiders in CO_2, removed, and immediately flash-froze the venom glands in liquid nitrogen. We immediately isolated RNA from homogenized venom glands using the ChargeSwitch Total RNA Cell Kit (Invitrogen, Carlsbad, CA), or we stored glands at −70 °C until RNA isolation. We synthesized first-strand cDNA using an anchored oligo-dT primer targeted for annealing at the 3′-end of mRNAs (5′-ggccacgcgtcgactagtacttttttttttttttttt-3′) and SuperScript III Reverse Transcriptase (Invitrogen). We increased all reactions 5× from the manufacturer's protocol.

Table 1
Summary of All Known SMaseD DNA and Protein Sequences Published to Date with Summaries of Enzyme Activity, Dermonecrotic Activity, and Studies That Have Been Conducted with Recombinant Forms of the Protein.

Venom Source	*SicTox* Group	Name (also known as)	DNA accession	Protein accession	RecExp	SMaseD activity	Derm. activity	Reference
Loxosceles reclusa species group								
L.arizonica –crude						X		Binford and Wells (2003)
L.arizonica	αIB2a	SMaseD derm. enz. prec.	AF512953	AAP44735				Binford et al. (2005)
	αIB2b	SMaseD-like prot. 2	AY699703	AAW22997				Binford et al. (2005)
	αIB1a	SMaseD-like prot. 3	AY699704	AAW22998				Binford et al. (2005)
L.apachea – crude						X		Binford and Wells (2003)
L.alamosa – crude						X		Binford and Wells (2003)
L.deserta – crude						X		Binford and Wells (2003)
						X	X	Barbaro et al. (2005)
							X[a]	Gomez et al. (2001)
L.reclusa – crude						X		Binford and Wells (2003)
						X	X	Barbaro et al. (2005)
						X[a]		Kurpiewski et al. (1981)
							X[a]	Merchant et al. (1998)
							X[a]	Gomez et al. (2001)
L.reclusa	αIA1	SMaseD prec.	AY862486	AAW56831	X	X		Lee and Lynch (2005)
	αIB1	Lr1	AY559846	AAT66075	X	X		Olvera et al. (2006)
		(SMaseD prot. 1)				X[a]		Ramos-Cerrillo et al. (2004)
							X[a]	Geren et al. (1976)
	αIA1	Lr2	AY559847	AAT66076		X[a]		Ramos-Cerrillo et al. (2004)
		(SMaseD prot. 2)					†[a,b]	Geren et al. (1976)
L.boneti	αIB1	SMaseD prot. 1	AY559844	AAT66073	X	X		Olvera et al. (2006)
						X[a]	X[a]	Ramos-Cerrillo et al. (2004)
		SMaseD-like prot. 2	not submitted	not submitted		X[a]low[c]	X[a]	Ramos-Cerrillo et al. (2004)
	βIA1	SMaseD-like prot. 3	AY559845	AAT66074		†[a,b]	†[a,b]	Ramos-Cerrillo et al. (2004)
Loxosceles rufescens species group								
L.rufescens – crude						X		Binford and Wells (2003)
Loxosceles laeta species group								
L.laeta – crude						X		Binford and Wells (2003)
						X	X	Barbaro et al. (2005)
							X[a]	Barbaro et al. (1996)
L.laeta	αIII1	SMaseD Ll1	DQ369999	ABD15447	X	X		Olvera et al. (2006)
	αIII2	SMaseD Ll2	DQ370000	ABD15448	X	X		Olvera et al. (2006)
	αIII1	Clone H17(SMase I)	AY093599	AAM21154	X	X	X	Fernandes-Pedrosa et al. (2002)
	αIII3	Clone H13(SMase-like prot.)	AY093600	AAM21155	X			Fernandes-Pedrosa et al. (2002)
	βIA1	Clone H10(SMase-like prot.)	AY093601	AAM21156	X			Fernandes-Pedrosa et al. (2002)
Loxosceles spadicea species group								
L.intermedia – crude						X	X	Barbaro et al. (2005)
							X[a]	Barbaro et al. (1996)
L.intermedia	αIA1b	P1[d](SMaseP1 prec.,	AY304471	AAP97091	X	X	X	Tambourgi et al. (2004)
		SMaseD1, SM phosph				X[a]	X[a]	Tambourgi et al. (1998)
		D1 prec.)				X	X	de Andrade et al. (2006)
	αIA2a	P2(SMaseP2 prec.,	AY304472	AAP97092	X	X	X	Tambourgi et al. (2004)
		SMaseD2,				X[a]	X[a]	Tambourgi et al. (1998)
		SM phosph D2 prec.)				X	X	de Andrade et al. (2006)
		P3(LiP3)	not submitted			†[a,b]	†[a,b]	Tambourgi et al. (1998)
	αIA1b	LiD1[d]	AY340702	AAQ16123	X			Kalapothakis et al. (2002)
		(recLiD1, derm. prot. 1)				Xvery low[c]	X	Felicori et al. (2006)
							X	Araujo et al. (2003)
	αIA1a	LiRecDT1	DQ218155	ABA62021	X			Chaim et al. (2006)
		(DT isoform 1)				X	X	da Silveira et al. (2006)
							X	Ribeiro et al. (2007)

Table 1
Continued

Venom Source	*SicTox* Group	Name (also known as)	DNA accession	Protein accession	RecExp	SMaseD activity	Derm. activity	Reference
	αIA2a	LiRecDT2	DQ266399	ABB69098	X	X	X	da Silveira et al. (2006)
		(DT isoform 2)					X	Ribeiro et al. (2007)
	βIA1	LiRecDT3 [c] (DT isoform 3)	DQ267927	ABB71184	X	Xvery low[c]	†[b]	da Silveira et al. (2006)
							†[b]	Ribeiro et al. (2007)
	αII1	LiRecDT4(DT isoform 4)	DQ431848	ABD91846	X	X	X	da Silveira et al. (2001)
	βID1	LiRecDT5(DT isoform 5)	DQ431849	ABD91847	X	Xlow[c]	X	da Silveira et al (2001)
	αV1	LiRecDT6(DT isoform 6)	EF474482	ABO87656	X	X	X	Appel et al. (2008)
	βIA1	derm. prot.-like I[e]	DQ388596	ABD48088				de Moura et al. (2006) unpublished
	βIA1	derm. prot.-like II[f]	DQ388597	ABD48089				de Moura et al. (2006) unpublished
	αIA2b	LoxTox i1	EF535250	ABU43329				Kalapothakis et al. (2007)
	αIA2b	LoxTox i2	EF535251	ABU43330				Kalapothakis et al. (2007)
	αIA2a	LoxTox i3	EF535252	ABU43331				Kalapothakis et al. (2007)
	αIA1b	LoxTox i4[g]	EF535253	ABU43332				Kalapothakis et al. (2007)
	αII2	LoxTox i5	EF535254	ABU43333				Kalapothakis et al. (2007)
	βIA1	LoxTox i6[e]	EF535255	ABU43334				Kalapothakis et al. (2007)
	βIA1	LoxTox i7[f]	EF535256	ABU43335				Kalapothakis et al. (2007)
Loxosceles gaucho species group								
L.gaucho – crude						X	X	Barbaro et al. (2005)
							X[a]	Cunha et al. (2003)
							X[a]	Barbaro et al. (1996)
L.gaucho	αIA1	derm. prot. 1	AY974250	AAY42401				Silvestre et al. (2005) unpublished
L.similis	αIA1	LsD1[g]	AY929305	AAX78234			X	Silvestre et al. (2005)
L. adelaida – crude						X	X	Pretel et al. (2005)
Loxosceles spinulosa species group								
L.spinulosa (Kwazulu Natal) – *crude*						X		Binford and Wells (2003)
L.speluncarum (Groenkloof) – *crude*						X		Binford and Wells (2003)
L. sp aff. speluncarum[h]*- crude*						X		Binford and Wells (2003)
Sicarius species (Africa)								
S.sp cf. hahni[i] *– crude*						X		Binford and Wells (2003)
S.sp cf. damarensis[j] *-crude*						X		Binford and Wells (2003)

NOTE.—RecExp indicates genes whose expressed proteins have been analyzed. derm = dermonecrotic, prec = precursor, prot. = protein, SM = sphingomyelin, phosph = phosphodiesterase, DT – dermonecrotic toxin. "X" indicates the protein tested positive for activity.

[a] Includes data from column chromatography fractionation experiments.

[†b] No activity as defined in original publications.

[c] "low" as defined in original publications, but generally less than half of the maximum activity levels reported.

[d, e, f, g] refer to nucleotide sequences that are identical.

[h] The same population referred to as *Lsp* Hooenoeg in Binford & Wells (2003).

[i] The same population referred to as *S. hahni* in Binford & Wells (2003).

[j] The same population referred to as *S. testaceus* in Binford & Wells (2003).

Table 2
Taxa of *Loxosceles*, *Sicarius*, and *Drymusa* for Which New Data are Presented in This Work

Species	Locality	SMaseD cDNA	SMaseD assay	SDS-PAGE
Loxosceles				
***reclusa* species group**				
Loxosceles arizonica	United States: Tucson, AZ	X	X	X
Loxosceles apachea	United States: Stein's Ghost Town, NM	X	X	X
Loxosceles sabina	United States: Bill's Cave, Vail, AZ	X	X	X
Loxosceles deserta	United States: Granite Mtn, CA	X	X	X
Loxosceles reclusa	United States: Oxford, MS		X	X
***laeta* species group**				
Loxosceles laeta	United States: Los Angeles, CA	X	X	X
L. sp. nov. Catamarca	Argentina: Catamarca	X	X	X
***spadicea* species group**				
Loxosceles intermedia	Argentina: El Palmar		X	X
Loxosceles hirsuta	Argentina: Chaco	X	X	X
Loxosceles spadicea	Argentina: Catamarca	X	X	X
***gaucho* species group**				
Loxosceles variegata	Argentina: Corrientes	X		
***amazonica* species group**				
Loxosceles amazonica	Peru: Loreto, Pevas	X	X	X
***rufescens* species group**				
Loxosceles rufescens	United States: Indianapolis, IN	X	X	X
***vonwredei* species group**				
Loxosceles vonwredei	Namibia: Uisib Farm Caves		X	X
***spinulosa* species group**				
L. sp. aff spinulosa	Namibia: Munsterland	X	X	X
Loxosceles speluncarum	South Africa: Greensleeves Cave	X	X	X
L. sp. aff. spinulosa	Namibia: Ruacana Falls	X	X	X
Loxosceles spinulosa	Namibia: Grootfontein	X	X	X
L. spinulosa	South Africa: Borakalalo	X	X	X
L. sp. aff. spinulosa	Namibia: Windhoek	X	X	X
Sicarius				
Africa				
S. sp. cf. damarensis	Namibia: Oorloogskloof	X	X	X
S. sp. cf. hahni	Namibia: Strydpoort Mtns			
Sicarius dolichocephalus	Namibia: Ruacana Falls		X	
Sicarius damarensis	Namibia: Daan Viljoen Park		X	
S. sp. aff. damarensis	Namibia: Munsterland Farm		X	
Sicarius albospinosus	Namibia: Gobabeb	X	X	X
S. albospinosus	Namibia: Wundergat		X	
South America				
Sicarius terrosus	Argentina: Catamarca		X	X
S. terrosus	Argentina: Sierra de las Quijades	X	X	X
Sicarius rupestris	Argentina: Corallito		X	X
Sicarius patagonicus	Argentina: Picun Leufo	X	X	X
Sicarius peruensis	Peru: Lima - Pisco		X	X
S. peruensis	Peru: Olmos - Lambayeque	X	X	X
Central America				
Sicarius rugosus	Costa Rica: Palo Verde		X	X
Drymusa				
Drymusa capensis	South Africa: Capetown		X	X
Drymusa serrana	Argentina: Merlo		X	X
Drymusa dinora	Costa Rica: Osa Peninsula		X	

NOTE.—Analyses that include data for a given taxon are indicated by "X."

Amplification of SMase D Homologs

To amplify a diverse set of SMase D homologs from cDNA samples, we used two different degenerate primers that were designed to match the N-terminus of published members of this gene family. One of these we have previously used to amplify SMase D genes from *Loxosceles arizonica* ("sphing1f" 5′-tggathatgggncayatggt-3′, Binford et al. 2005). To amplify more divergent homologs, we designed another primer ("Lpara" 5′-gcncayatggtnaaygayt-3′) to match published and divergent homologs of SMase D (*laeta*H10, Fernandes-Pedrosa et al. 2002) and homologs for which expression products show little or no enzyme activity when tested (LiD1, Kalapothakis et al. 2002; Lb3,

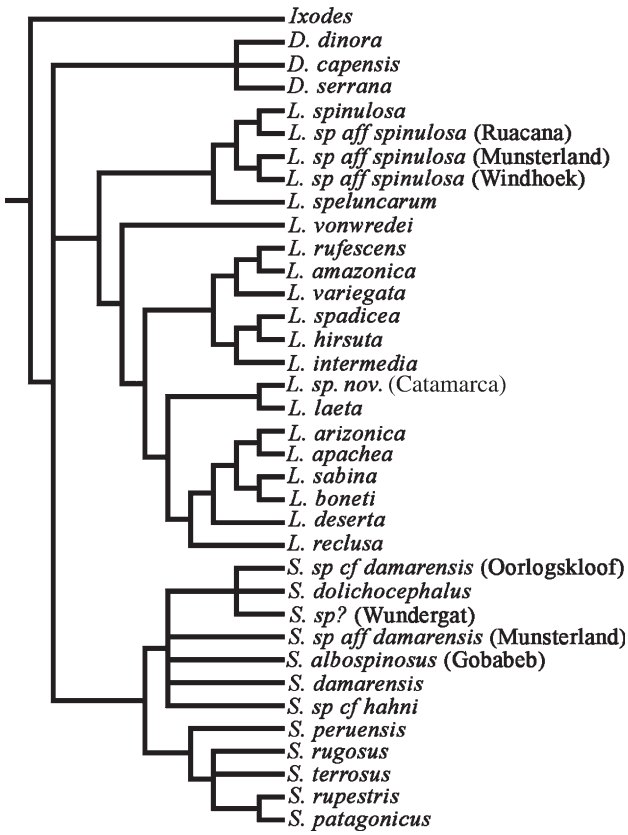

FIG. 1.—Relationships of species from which *SicTox* genes and venoms were included in our analysis. This is a summary composite from analyses of Binford et al. (2008).

Ramos-Cerrillo et al. 2004; LiRecDT3, da Silveira et al. 2006; Ribeiro et al. 2007) (table 1). We used the same primer to amplify from the C-terminus in all reactions (5′-ccacgcgtcgactagtac-3′).

For polymerase chain reaction (PCR), we used 2× MasterAmp PreMix (Epicentre Technologies, Madison, WI), NEB Taq Polymerase (New England Biolabs, Ipswich, MA), and an annealing temperature of 51 °C for all reactions except those with the primer Lpara and templates from *Loxosceles spinulosa* (Borakalalo) and *L. spinulosa* (Ruacana) that were run at 45 °C. We cloned PCR products that were ~1 kb (approximate size of the SMase D gene) into pCR4-TOPO vector using the TOPO TA Cloning Kit for Sequencing (Invitrogen). We screened bacterial transformants using whole colony PCR under the conditions used to initially amplify the inserts. From each cDNA template, we screened a minimum of 20 colonies and up to 152 for taxa that were low yielding (*L sp. aff. spinulosa* Windhoek). Colonies with products ≥700 bp were grown in selective media and purified using the QIAPrep spin mini-prep kit (Qiagen, Valencia, CA). Inserts were sequenced in both directions with T3 and T7 primers on an Applied Biosystems 3730xl DNA Analyzer (Foster City, CA) at the Genomic Analysis and Technology Core (University of Arizona).

Sequence Analysis

After trimming vector sequence, we confirmed homology to SMase D by submitting sequences to TBlastX searches in the NCBI gene database (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). We assembled all sequences that were homologous to SMase D genes using Sequencher (version 4.7, Gene Codes Corp.). We aligned all nucleotide sequences we recovered with all sequences available in GenBank, using ClustalX (Thompson et al. 1997) and refined this manually using MacClade (version 4.06; Maddison and Maddison 2005) guided by color-coding nucleotides according to their translated amino acid. We used conserved published active sites as anchors in the alignment. Sequences that ended prematurely (10) had early stop codons (19), or frame-shifting indels (8) were removed from the data set.

Molecular Phylogenetics

We analyzed phylogenetic relationships of our nonredundant amino acid data set (196 taxa in the ingroup, 349 characters) using Neighbor-Joining (NJ) in PAUP* (Swofford 1998) and Bayesian analyses in MrBayes v.3.1.2 (Ronquist and Huelsenbeck 2003). We used an SMase D homolog isolated from *Ixodes* (Arachnida:Acari) (DQ411855) as an outgroup. We assessed confidence in branches using 1,000 nonparametric bootstrap replicates for NJ analyses. All Bayesian analyses consisted of two simultaneous runs, each with four Markov Chain Monte Carlo chains. The current tree at every increment of 100 generations was saved to a file. We used default cold and heated chain parameters and compared the separate runs every 1,000 generations to facilitate convergence. We ran this analysis for 2 million generations when we considered the sampling to be adequate based on average standard deviation of split frequencies being <0.01 (Ronquist et al. 2005). We determined the burn-in period as the set of trees saved prior to log likelihood stabilization and convergence as estimated using Tracer 1.4 (Drummond and Rambaut, 2007).

Criteria for Identifying Paralogs

For many species, we recovered multiple distinct but very similar sequences that Bayesian analyses resolved as each other's closest relatives, often in a polytomy. Without further analyses, it is difficult to distinguish whether these are allelic variants or paralogs resulting from recent duplication events. We estimate the number of paralogs by three different methods. Our most liberal estimate was to define sets of sequences from a given species that were monophyletic terminal polytomies on our tree as allelic variants of a single paralog. For comparison, we estimated distinct paralogs within terminal monophyletic intraspecific groups as distinct paralogs when amino acid sequence divergence among them was 2% or greater (~5 amino acids different). For analyses that would be biased by redundancy (gene duplication analyses and models of structural conservation, see below), we analyze a subset of our data set only including intraspecific sequences that were minimally 5% divergent (~14 amino acids different).

Species Tree Gene Tree Reconciliation Analysis

We estimated the number of gene duplications and losses required to reconcile the gene tree with an

independently estimated species tree (fig. 1) using reconciliation analysis in Notung 2.5 (Chen et al. 2000; Durand et al. 2006; Vernot et al. 2007). To address biases in the analysis from overrepresentation of paralogs via allelic variants and underrepresentation of paralogs from our inexhaustive sampling methods, we did this analysis on our full data set (all terminal taxa in fig. 2), and repeated it on two reduced data sets culled to include sequences with a minimum of 2% and 5% amino acid divergence as described above. For analyses of each of these data sets, we used gene trees estimated using NJ in PAUP* that were rooted with the sequence from *Ixodes*. We removed *Ixodes* from the gene tree before reconciliation.

Tests for Positive Directional Selection

We tested for an influence of positive selection on divergence of *SicTox* nucleotide sequences using two codon-based likelihood analyses implemented in the Codeml package of PAML 3.14 (Yang 2007), branch models (Yang 1998) and site models (Yang et al. 2000). We did not test for selection at particular branches (branch-site model) (Yang and Nielsen 2002) because we had no a priori hypotheses of specific branches undergoing positive selection. To increase computational speed, we analyzed a data set that included a single representative of terminal, intraspecific, monophyletic sets of sequences that shared less than 5% sequence divergence, we analyzed the α and the β clades separately (47 and 29 sequences, respectively), and we set cleandata = 1, which only includes amino acids with no gaps. Our input trees that included only the sequences under analysis were estimated by NJ and rooted by *Ixodes*. *Ixodes* was removed from the tree before input into Codeml. For branch analyses, we estimated independent ω values for all branches (free-ratios model), we estimated a single value of ω under a constraint that ω remain constant (one-ratio model), and computed log likelihood values when ω was fixed at 1. We used likelihood ratio tests (LRTs) to determine if log likelihood estimates were significantly different when ω values were free to vary and when they were constrained to have a single value. We calculated this by comparing twice the difference between the log likelihood under the fixed model and the log likelihood of the free model to a χ^2 distribution with one degree of freedom (Yang 1998; Yang and Nielsen 1998; Yang and Bielawski 2000). For tests of selection that consider variation in selection across sites, we used the following models: M0 (one ratio), M1a (Nearly Neutral), M2a (Positive Selection), M7 (beta), and M8 (beta & ω) (Yang et al. 2005). We used LRTs comparing M1a and M2a, and M7 and M8 using df = 2.

Comparative Protein Composition and SMase D Assays

Protein Gel Electrophoresis

For protein analyses, we pooled venom collected by electrostimulation (see above) among individuals within the same population. We diluted all venoms in 1× Amplex Red buffer (5 mM $CaCl_2$, 50 mM Tris–Cl, pH 8) and quantified total venom protein using the Coomassie (Bradford) Protein Assay (Pierce, Rockford, IL).

We separated proteins from crude venom (3, 5, or 7 μg) using one-dimensional SDS-PAGE with precast Criterion 12.5% midi gels (Bio-Rad, Hercules, CA). Intensity of staining varied among lanes with 5 μg of protein loaded, so we ran select venoms with 3, 5, and 7 μg of total protein to determine the effect of staining intensity on visibility of bands. We used a broad range molecular weight marker (New England Biolabs, Ipswich, MA, P7702S) along with a low range silver stain standard (Bio-Rad, 161-0314) on each gel for size reference. For visualization, we stained the gels with silver stain using standard protocol.

Assays for SMase D Activity

We assayed SMase D activity using a modification of the Amplex Red Phospholipase D Assay Kit (Molecular Probes, Eugene, OR), as described in Binford and Wells (2003) with sphingomyelin (from chicken egg yolk, Sigma, St Louis, MO) as the substrate. We measured fluorescence emission from reactions (200 μl) in a quartz fluorimeter cell (10 mm, $z = 15$, Starna Cells, Inc., Hainault, UK) using a Perkin Elmer LS55 luminescence spectrometer (Waltham, MA).

Amino Acid Sequence Conservation

We analyzed amino acid conservation among a nonredundant set (no two members having higher than 95% amino acid sequence identity) of venom-expressed members of the *SicTox* gene family by visual inspection of alignments. Conservation levels were also mapped onto the solvent-exposed surface of *Loxosceles laeta* SMase D (PDB ID 1XX1, chain A) (Murakami et al. 2005) using the Multalign Viewer utility within the program Chimera (Pettersen et al. 2004).

Results and Discussion

We isolated homologs of SMase D from the set of taxa detailed in table 3. Although our sampling is far from exhaustive, we add sufficient phylogenetic inclusion of members for the gene family that includes SMase D to identify distinct phylogenetic groups and to detect frequent duplication events. We also detect variation in patterns of SMase D activity in whole crude venoms that correlates with phylogenetic patterns in our gene tree. Based on these data, we propose an inclusive phylogeny-based nomenclatural system for the gene family. We name the gene family *SicTox*, abbreviated from Sicariidae toxin, as an acknowledgment of the expression of members of this gene family in venoms of species across the spider family Sicariidae. We start with an overview of our naming system to create a language for reporting our results.

Rationale for Proposed Nomenclature

Our nomenclatural system follows the general guidelines established by the Human Genome Organization (HUGO) gene nomenclatural committee (HGNC) (http://

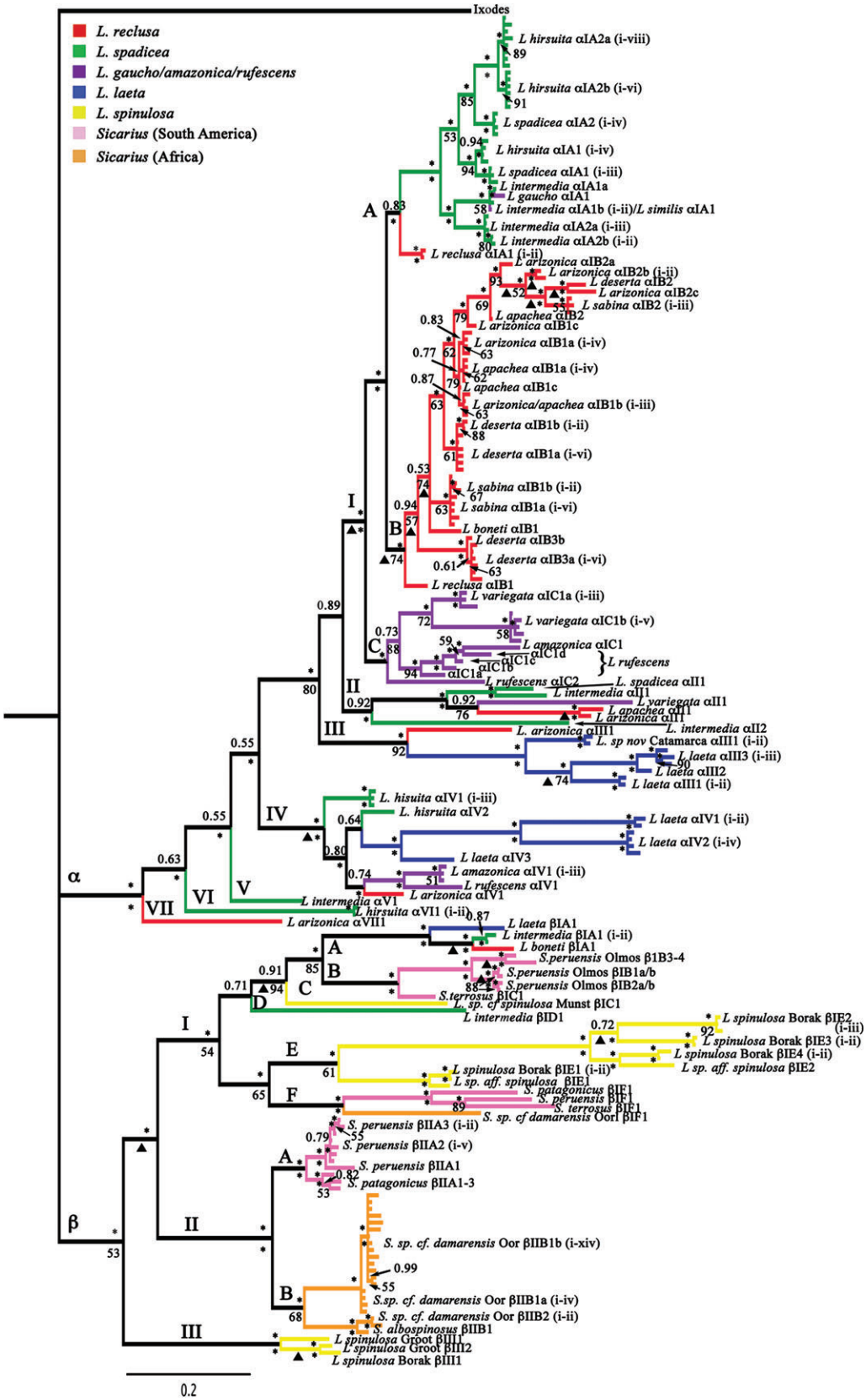

FIG. 2.—Tree topology resulting from Bayesian analyses. Posterior probabilities are labeled above branches with a * indicating support >0.95. Labels below the branch indicate bootstrap support from NJ analyses with 1,000 bootstrap replicates with a * indicating >95%. The proposed *SicTox* nomenclature delineating major clades on the tree are indicated by labels next to branches. The terminal names indicate the *SicTox* nomenclature for terminal groups (summarized in supplemental table 1, Supplementary Material online). Colored branches distinguish species groups as delineated in tables 1 and 2. Triangles below branches illustrate branches that had $\omega > 1$ in free-ratio codon likelihood analyses.

Table 3
Taxa Screened for Venom-Expressed SMase D Homologs Organized by Species Group

Species	*n* in mRNA Pool	SMase D Homologs Recovered	Nonredundant Nucleotide Sequences	Nonredundant Amino Acid Sequences
Loxosceles **reclusa species group**				
Loxosceles arizonica	2	15	11	11
Loxosceles apachea	8	22	11	10
Loxosceles sabina	4	15	13	11
Loxosceles deserta	1	23	18	16
laeta **species group**				
Loxosceles laeta	4	12	9	9
L sp. nov. Catamarca	1	4	2	2
spadicea **species group**				
Loxosceles hirsuta	2	45	31	24
Loxosceles spadicea	2	20	14	8
gaucho **species group**				
Loxosceles variegata	1	14	9	9
amazonica **species group**				
Loxosceles amazonica	4	9	4	4
rufescens **species group**				
Loxosceles rufescens	8	10	6	6
spinulosa **species group**				
L. sp. aff. spinulosa (Munsterland)	2	1	1	1
L. sp aff. spinulosa (Ruacana)	3	4	2	2
Loxosceles spinulosa (Grootfontein)	4	8	2	2
L. spinulosa (Borakolo)	2	17	10	10
Sicarius **Africa**				
S. sp. cf. damarensis (Oorlogsk)	2	27	21	21
Sicarius albospinosus (Gobabeb)	4	3	1	1
South America				
Sicarius terrosus (S. de las Quijades)	1	4	2	2
S. patagonicus	1	8	4	4
Sicarius peruensis (Olmos)	1	26	19	15
Total		291	190	168

NOTE.—The number of spiders in the mRNA pool is listed as "*n*." The distinct SMase D gene copies are distinct from known SMase D sequences. Total number of colonies screened: 1,960.

www.genenames.org/). This committee encourages using symbols to distinguish phylogenetically defined stem or root groupings and then using a hierarchical numbering system to distinguish individual members. Rather than create a rigid structure our goal is to establish a general working scaffold within which newly discovered homologs may easily be placed in a hierarchical structure representative of degree of relatedness. This will serve to standardize discussions in the literature and facilitate efficient characterization of structure and function of this gene family by targeting divergent lineages for further analyses. Recent work by Kalapothakis et al. (2007) provides an excellent starting place. The sequences used in their analysis are anchors in our phylogeny that inform hypotheses of the functional evolution within this family. A complete translation of previous published names for toxins in this family to the *SicTox* nomenclature is in table 1,and a complete list of all known *SicTox* family members is available in supplemental table 1, Supplementary Material online. We detail the nomenclature and the logic behind it as we describe the structure of diversity of the group.

Isolation and Identification of SMase D Paralogs

We isolated 329 sequences homologous to SMase D, 291 of which were full length. Of those, 190 were nonredundant at the nucleotide level (GenBank Accession numbers FJ17340–FJ17529, detailed in Supplemental table 1, Supplementary Material online), and 168 were nonredundant at the amino acid level. Of the distinct nucleotide sequences, 143 are from 16 species of *Loxosceles*, and 47 are from 5 species of *Sicarius* (table 3). Our PCR methods did not amplify any genes from our outgroup taxon, *Drymusa serrana*. There was a large range in the number of SMase D homologs discovered from within a single mRNA pool. We were particularly successful at amplifying diverse SMase D homologs from *Loxosceles hirsuta*, *Loxosceles apachea*, *L. arizonica*, *S sp cf damarensis* (Oorlogskloof), and *Sicarius peruensis* (Olmos) (table 3). The number of distinct paralogs that we isolated from within a single species ranged from 1 to a minimum of 12 (defined by 98% identity) in *L. arizonica*. Our differential recovery of sequences and paralogs among taxa is not easily explained by numbers of individual spiders whose venom glands were combined in the cDNA pool or by greater success in amplifying cDNAs from species closely related to the source species of the amino acid sequences we used to design degenerate primers (table 3). The primer sphing1f successfully amplified homologs across all taxa, whereas Lpara amplified 12 sequences nonredundant at the nucleotide and amino acid levels (1 in the α clade, *L. apachea* αIB1b, and 11 in the β clade, all from African *Loxosceles* in the *spinulosa* species group).

Phylogenetic Patterns and Delineation of Major Clades

Bayesian analyses resolved two distinct clades that we label α and β, with posterior probabilities of 1.0 and 0.99, respectively (fig. 2). NJ collapses the β clade into two clades, each with low bootstrap support, as an unresolved basal polytomy with the α clade. The average pairwise amino acid distance *p*-distance between the α and β clades is 0.537. We distinguish major lineages (groups) within the α and β clades by roman numerals (I–VII in α) and (I–III in β) (fig. 2). We delineate groups based on strong posterior probability support of monophyly (fig. 2) and high percent divergences among them (table 4). Some groups either include a large number of sequences structured into well-supported groups (e.g., αI), or include sequences that are highly divergent (e.g., βI & βII). We delineate subgroups within these groups with capital letters (A–C in αI, A–F in βI, and A–B in βII). The number of

nonredundant amino acid sequences included in groups and subgroups ranges from 1 (αVI, βIC) to 47 (αIB).

Structure within the α Clade

Our methods isolated more sequences in the α clade than the β clade. The α clade consists predominantly of genes from New World *Loxosceles* and includes published genes whose expression products have been demonstrated to cause dermonecrosis (table 1). Average amino acid differences among groups within the α clade range from 0.31 to 0.52 (table 4). Some subgroups within this clade contain sequences that are predominantly from single species groups. For example, the αIA clade includes genes isolated from South American *Loxosceles* in the *spadicea* and *gaucho* species groups. The only exception is *Loxosceles reclusa* αIA1 that Bayesian analyses resolve as a basal lineage of the αIA clade. Support for this placement is weak, and both parsimony and NJ resolve with weak support *L. reclusa* αIA1 as the basal lineage of the αIB clade. The αIB clade exclusively contains genes isolated from members of the *reclusa* species group. Although we isolated more sequences in the αIA and αIB clades than other clades, the average amino acid divergences within these clades are relatively low (table 4). The αIC clade consists exclusively of members of the *gaucho, amazonica*, and *rufescens* species groups. These lineages are well supported to share a close relationship despite the fact that *gaucho* and *amazonica* are native to South America and *rufescens* group members are native to the Old World (Binford et al. 2008, fig. 1).

The αII group contains clear paralogs of sequences in the αI group. Support for monophyly of αII is strong except for the inclusion of *LoxTox* i5 (*SicTox L. intermedia* αII2), which NJ resolves as polytomy with αI, the rest of the taxa in αII, and αIII. Conspicuously missing from the αI and αII clades are sequences from the *L. laeta* species group. This is surprising because species phylogenies tenuously support a closer relationship between members of the *laeta* and *reclusa* species groups than either of these lineages share with other American *Loxosceles* species groups (Binford et al. 2008, fig. 1). However, it is consistent with well-known patterns in the literature of *laeta* venom characteristics being divergent from those of other New World *Loxosceles* venoms that have been analyzed (see discussion below, Barbaro et al. 2005; de Roodt et al. 2007).

The rest of the groups in the α clade (IV–VII) contain paralogs of genes in αI–III. Group IV is the only one with more than one sequence. Bayesian analyses strongly support monophyly of group IV, but NJ breaks this clade into two, pulling all of the *L. laeta*, except *L. laeta* αIV3, out into a separate monophyletic group. NJ leaves these two groups in a basal polytomy within the α clade that includes the monophyetic clade of groups I–III and groups V, VI, and VII.

Structure within the β Clade

The β clade predominantly consists of *Sicarius* and African *Loxosceles* in the *spinulosa* species group within which clear paralogs are evident. This clade also includes a monophyletic group of published sequences from New World *Loxosceles* in the βIA clade (*L. laeta* H10, *Loxosceles boneti* P3, *Loxosceles intermedia* LiDT3, *L. laeta* βIA1, *L. boneti* βIA1, and *L. intermedia* βIA1 with our *SicTox* nomenclature, respectively), some of which have been demonstrated to have reduced or no SMase D or dermonecrotic activity (table 1 and references therein). Amino acid differences among groups are generally high, ranging from 0.37 to 0.52 (table 4). The only other New World sequence that falls in the β clade is LiDT5 (*L. intermedia* βID1). Bayesian and NJ analyses agree on inclusion of genes within the defined groups, but NJ weakly resolves a sister–

Table 4
Mean Amino Acid *p*-Distances within and among Groups and Subgroups of All Sequences in Our Nonredundant Data set (All Terminal Taxa in fig. 2)

	αI	αII	αIII	αIV	αV	αVI
αI						
αII	0.351					
αIII	0.421	0.489				
αIV	0.379	0.444	0.495			
αV	0.315	0.421	0.438	0.323		
αVI	0.418	0.475	0.520	0.433	0.325	
αVII	0.370	0.435	0.492	0.406	0.335	0.413

	αIA	αIB
αIA		
αIB	0.177	
αIC	0.216	0.210

	βIA	βIB	βIC	βID	βIE
βIA					
βIB	0.366				
βIC	0.379	0.384			
βID	0.461	0.446	0.439		
βIE	0.530	0.519	0.517	0.532	
βIF	0.470	0.461	0.477	0.503	0.496

	βI	βII
βI		
βII	0.507	
βIII	0.529	0.454

α Clade	Mean Distance Within	*n*
α	0.261	133
αI	0.156	98
αIA	0.093	36
αIB	0.095	48
αIC	0.161	14
αII	0.320	6
αIII	0.226	9
αIV	0.154	16
αVI	0.004	2

β Clade	Mean Distance Within	*n*
β	0.391	63
βI	0.434	28
βIA	0.129	4
βIB	0.106	7
βIE	0.317	11
βIF	0.314	4
βII	0.121	32
βIIA	0.047	11
βIIB	0.063	21
βIII	0.111	3

NOTE.—The *p*-distances between the α and β clades is 0.537. The distance between the βIIA and βIIB subgroups is 0.193.

taxon relationship between βII and βIII and collapses βID into a polytomy with a monophyletic group of βIA–C and a monophyletic group of βIE–F.

Duplication History in *SicTox* Gene Family

Although the *SicTox* gene tree has some clustering of sequences into clades consistent with species groups, duplications complicate this relationship. Recent duplications are most apparent in groups of closely related species for which we have extensive recovery of gene copies. This is particularly true for αIA and αIB members of the *spadicea* and *reclusa* groups, respectively. Our sampling includes the only three described species in the *spadicea* group, and there is solid support of *spadicea* and *hirsuta* being each other's closest relatives with *intermedia* sister to that pair (fig. 1). Within the αIA clade, there have been at least three duplication events since the MRCA of the *spadecia* group (fig. 2). Estimates of the age of the MRCA of the *spadecia* group range between 48 and 34 Ma (unpublished, methods as in Binford et al. 2008), meaning the minimally two duplications between the MRCA and *hirsuta* occurred within this time frame. The αIB clade includes only members of the North American *reclusa* group. Patterns of relationships among terminal paralog groups within αIB mirror species relationships in the *reclusa* group (fig. 1) and contain clear duplication events (fig. 2). An example is in the phylogenetic grade that includes *L arizonica* αIB2a, αIB2b, and αIB2c.

Panning out to a broad description of the history of duplications and losses in this lineage, reconciliation analyses (Notung) of a species tree (subset of fig. 1) with an SMase D gene tree (NJ) that includes all unique amino acid sequences (all taxa in fig. 1 except *Ixodes*) estimated 110 duplications and 84 losses (D/L score = 249.0). This likely overestimates duplications because of the inclusion of allelic variants in monophyletic terminal intraspecific groups. Repeating the analysis including only terminal paralogs defined by minimum amino acid divergence thresholds of 2% (105 sequences) and 5% (75 sequences) estimated 44 duplication events and 71 losses (D/L score = 137.0), and 30 duplications and 74 losses (D/L score = 119.0), respectively. The 5% threshold resulted in monophyletic, terminal intraspecific sets of sequences being reduced to single representatives and therefore eliminates the possibility of allelic variants but also will exclude very recent duplications.

It has been hypothesized that gene families encoding spider venom toxins are evolving via mechanisms similar to those driving *Conus* toxin diversification in which there is rapid duplication followed by adaptive divergence (Sollod et al. 2005; Escoubas 2006). In spiders, rapid duplication rates in toxin gene families are assumed based on large numbers of paralogs expressed in venoms (King et al. 2002; Sollod et al. 2005; Escoubas 2006; Jiang et al. 2008). Little work in any toxin lineage has attempted to quantify duplication rates, and reasonably so, given a long list of limitations including 1) lack of comprehensive knowledge of gene family size that would require a sequenced genome; 2) the complicating influences of concerted evolution and the evolution of expression patterns (Duda and Remigio 2008); and 3) few analyses with dense taxon sampling of genes among closely related species of known relationships and with known divergence times. Of these, our work only overcomes the limitation of having a species tree and broad *SicTox* gene sampling. Thus, although preliminary, estimates from reconciliation serve as a starting point for discussion of general patterns of duplication rates in venom evolution.

To translate our reconciliation results to inferred duplication rates, if duplications were evenly distributed across the 25 taxa in our analysis, the average species would have undergone 1.2 (5% threshold) to 4.4 (all data included) duplications in the evolutionary time frame encompassed. This assumption is unlikely to be true given the wide range of paralog numbers we recovered across taxa. Thus, these estimates are conservative. Molecular dating analyses and biogeographic patterns support the MRCA of Sicariidae predating the separation of South America and Africa meaning a conservative estimate of minimum age of the MRCA is 95 Ma and a maximum estimate is 157 Ma (Binford et al. 2008). This translates to 0.0126–0.0463 duplications per million years for the minimum age of sicariids and 0.0076–0.0280 duplications per million years for the maximum age.

These estimates of duplication rates are within ranges of estimates under the best of circumstances using comparisons of whole sequenced genomes, but minimizing the confounding influence of gene conversion by constraining analyses to include only pairs of duplicate genes. These include 0.028, 0.0014, and 0.024 duplications per million years in yeast, *Drosophila* and *Caenorhabditis elegans*, respectively (Gu et al. 2002 assuming a molecular clock), 0.01–0.06 duplications per billion years in yeast (not assuming molecular clock, Gao and Innan 2004), and 0.009 duplications per million years in humans (Lynch and Conery 2003). We can estimate duplication rates from a venom toxin in the four-loop conotoxin lineage. At least seven duplications have occurred since the divergence of *Conus abbreviatus* and *Conus lividus* (Duda and Palumbi 1999) whose MRCA is estimated to have lived approximately 27 Ma (Duda and Kohn 2005), which results in an average duplication rate of 0.269 duplications per million years. Within the *spadecia* group described above, estimated duplication rates of *SicTox* would be between 0.041 and 0.071 per million years, which is faster than any of our estimates from reconciliation and faster than rigorous estimates from model systems but slower than the estimates for *Conus* four-loop toxins.

A final notable pattern is that single species (*L. intermedia*, *L. boneti*, and *L. laeta*) have *SicTox* paralogs in both the α and β clades. These are descendents of a duplication event that must predate the MRCA of Sicariidae and thus the putative origin of SMase D as a venom toxin in sicariids (Binford and Wells 2003).

Evidence of Purifying Selection, Episodic Positive Directional Selection and Positively Selected Sites

LRTs indicate significant improvements in log likelihood values under the free-ratios model relative to the fixed ratios model for both the α and the β clades (table 5).

Table 5
Summary Statistics for Branch and Site Analyses

		Parameters		Estimates	−ln *L*	χ^2	df	*P*
Branch models								
α	ω fix at 1				18,339.0			
	M0 (one ratio)	ω		0.250	17,657.9			
	M1 (free ratios)	ω		Varies	17,512.2	291.4	1	<0.0001
β	ω fix at 1				14,369.0			
	M0 (one ratio)	ω		0.193	13,685.6			
	M1 (free ratios)	ω		Varies	13,607.9	155.4	1	<0.0001
Site Models			Proportions					
α	M0	ω		0.258	16,920.0			
	M1a (nearly neutral)	ω_0	0.739	0.163	16,544.9			
		ω_1	0.261	1.000				
	M2a (positive selection)	ω_0	0.727	0.165	16,538.7	12.2	2	<0.005
		ω_1	0.256	1.000				
		ω_2	0.018	**2.091**				
	M7 (beta)	p		0.576	16,364.8			
		q		1.351				
	M8 (beta & ω)	p_0	0.949		16,348.3	33.0	2	<0.001
		p		0.730				
		q		2.284				
		$\omega_s > 1$	0.051	**1.344**				
β								
	M0	ω		0.199	12,642.0			
	M1a (nearly neutral)	ω_0	0.793	0.158	12,452.0			
		ω_1	0.206	1.000				
	M2a (positive selection)	ω_0	0.793	0.158	12,452.0	0.0	2	ns
		ω_1	0.044	1.000				
		ω_2	0.162	1.000				
	M7 (beta)	p		0.672	12,280.2			
		q		2.188				
	M8 (beta & ω)	p_0	0.969		12,278.9	2.6	2	ns
		p		0.743				
		q		2.752				
		$\omega_s > 1$	0.031	1.000				

Positively selected amino acid residues based on BEB (Yang et al. 2005) ($P > 0.95$ in bold).

α	M2a	18							**151**		**204**	208					
	M8	**18**	39	40	56	58		115	**151**	176	**204**	208	215	236			268
β	M8		39				60								259		

NOTE.—There were 47 and 29 sequences in the analysis of α and β clades, respectively. Sequences included in this analysis are indicated in supplemental table 1, Supplementary Material online. χ^2 values are from LRT of the model with its nested partner. The numbering of residues identified by BEB is based on homologous sites on PDB 1xx1.

They also support significant deviation of ω from 1. Estimated ω values under the fixed model (constrained to a single value across the lineage) were considerably less than 1 (0.250 and 0.193 for the α and β clades, respectively), consistent with predominantly purifying selection in the lineage. Reconstructed estimates of ω along branches under the free-ratios model are generally less than 1; however, scattered branches within both the α and the β clades have ω values greater than 1 (fig. 2). The significantly better fit of the free-ratios model and branches with $\omega > 1$ is consistent with a pattern of episodic positive selection acting in this lineage. Branches with estimated $\omega > 1$ are particularly concentrated in the αIB and the βIA and B clades.

Site model analyses support significant improvement of log likelihood values in the α clade under the M2a and M8 models relative to their nested models. In this clade, 1.8% and 5.1% of amino acids are identified as evolving under positive selection by M2a and M8 models, respectively (table 5). Thirteen amino acids were identified by Bayes Empirical Bayes (BEB) analyses (Yang et al. 2005) (table 5) as candidates for positively selected sites, three with posterior probabilities > 0.95 (18, 151, and 204, numbering according to PDB 1xx1 *L laeta* H17, αIII1i). Some of the weakly supported amino acids are near active sites in the protein; however, they all appear to be solvent-exposed positions that are neither in the active site nor in the plug region. We are investigating possible functional roles for these residues.

Analyses of the β clade did not detect significant improvement in models that support amino acids that are under positive selection across the lineage (table 5). However, BEB analyses under the M8 model identified three sites, 39, 60, and 259 as candidates for evolving under positive selection in this clade but with low posterior probabilities (ranging from 0.55 to 0.63) (table 5). Like the residues in the α clade, these are all solvent exposed, and their functional relevance is unclear. As we fill in our data set with more thorough taxon sampling and more resolution in functional diversity, we will pursue branch-site analyses that may be more sensitive at detecting amino acids that are under selection in particular lineages.

Major Species Lineages Vary in SMase D-Sized Proteins in Venoms and in SMase D Activity

Qualitative comparative protein composition data and quantitative SMase D enzyme activity data allow a preliminarily view of how differences in expression of *SicTox* gene family members may influence differences in excreted venoms. They also add to a growing body of comparative data of whole crude venoms (table 1 and references therein).

Expressed proteins from members of this gene family are between 31 and 35 kDa (see table 1 references). Using comparative SDS-PAGE, we detected proteins in this size range from all *Loxosceles* and *Sicarius* but not in either African or Argentine *Drymusa* (fig. 3*b*). Variation among taxa that have proteins in this size region correlates to some degree with species relatedness and patterns of relationships of *SicTox* gene family members. Among venoms from North America, *L. reclusa* has a concentrated region of proteins around 32 kDa, whereas other *reclusa* species group members from the desert southwest of the United States (*L. arizonica*, *L. apachea*, *Loxosceles sabina*, and *Loxosceles deserta*), have a broader range of proteins in that region (~31–34 kDa) (fig. 3*a*). Banding patterns of *L. laeta* are strikingly different from all other New World venoms with the most dense banding at 31 kDa and a reduction in proteins larger than that (fig. 3*a* and *c*). The differences in banding between *L. intermedia* and *L. hirsuta* are notable given the putative close relationship between these species (fig. 3*a*). Banding patterns for the African species *Loxosceles vonwredei* are more similar to New World species than they are to the African *spinulosa* group (fig. 3*a*), which is consistent with patterns of species relationships (Binford et al. 2008, fig. 1). Venoms from the African *spinulosa* clade (from which we have only isolated *SicTox* genes in the β clade) tend to have higher molecular weight proteins than other *Loxosceles* venoms. Venom profiles from *Sicarius*, again with isolated *SicTox* genes only from the β clade, have notable differences between New World and Old World species (fig. 3*b* and *c*). Venoms from African *Sicarius* have a higher concentration of larger molecular weight proteins, comparable with the African *Loxosceles* venoms, whereas the South American *Sicarius* have a region of dense proteins at ~31 kDa. This apparent size variation could be influenced by posttranslational modification (glycosylation), but it likely also reflects differences in the expressed *SicTox* genes. Disentangling these relative influences on protein variation will require paralog-specific expressions and Westerns, and a better understanding of posttranslational modifications.

We assayed crude venoms from 36 taxa for SMase D activity (table 2). *Sicarius* species from Africa, all *Loxosceles* from the Americas, and *Loxosceles rufescens* have high SMase D activity using 0.5 or 1.0 μg of crude venom, whereas activity from venoms from African *Loxosceles* is generally lower (fig. 4*a*), and there is no activity in *Drymusa* venom. All of these patterns are consistent with previous work (Binford and Wells 2003). However, we detected little or no activity at these concentrations in venoms from South and Central American *Sicarius* (fig. 4*a*). When we repeated the experiment for South American *Sicarius* with concentrations of crude venom that were increased by two orders of magnitude (50 μg) enzyme activity increased but was never as high as in North American *Loxosceles* or South African *Sicarius* assayed at the original low concentrations (fig. 4*b*). The increase in SMase D activity with

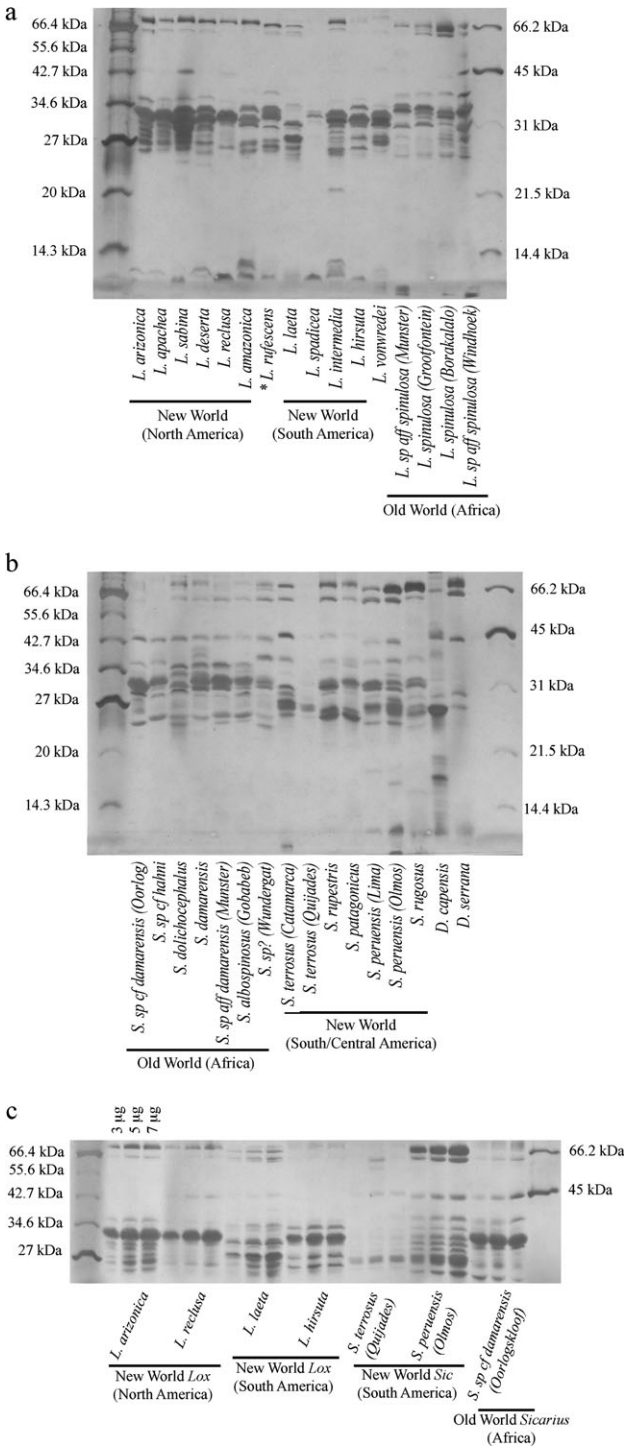

FIG. 3.—One-dimensional gel electrophoresis separations of 5 μg crude venoms from species of (*a*) *Loxosceles*; (*b*) *Sicarius* and *Drymusa*. Gels (12.5% acrylamide) were stained with silver nitrate; (c) repeated separations of crude venoms from select species loaded with 3, 5, and 7 μg of total protein to illustrate the effect of uneven loading of total venom protein on visibility of bands. The region of origin of species is indicated. **Loxosceles rufescens* is native to the Mediterranean.

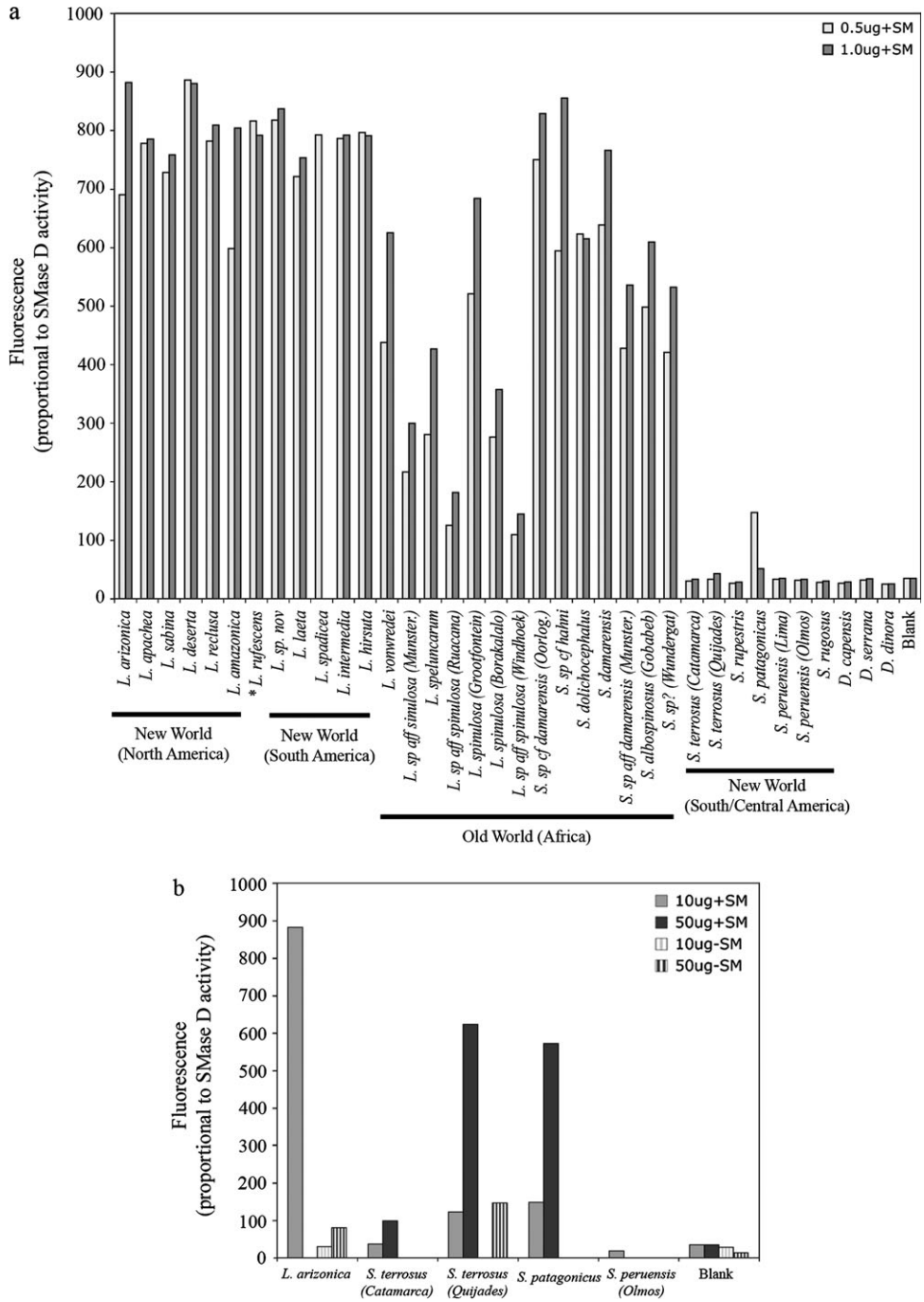

FIG. 4.—Fluorescence intensity (proportional to SMase D activity) resulting from assays of crude venom: (*a*) fluorescence measured from reactions that contained 0.5 or 1.0 μg of crude venom; (*b*) fluorescence from reactions that contained 10 or 50 μg of venom with and without sphingomyelin (SM) as the substrate. Some samples were not analyzed under all conditions because of low venom availability. Buffer was substituted for venom in blank samples. The region of origin is indicated. **Loxosceles rufescens* is native to the Mediterranean.

increasing amounts of venom from Argentine *Sicarius terrosus* (Quijades) and *Sicarius patagonicus* may reflect differences in substrate specificity of expressed *SicTox* members that results in a reduced binding efficacy for sphingomyelin. The fact that African *Sicarius* have strong SMase D signals and homologs for these species have only been recovered from the β clade suggests that at least some proteins in the β clade have strong efficacy for hydrolyzing sphingomyelin.

A growing body of work is indicating that defining the activities of venom-expressed members of this gene family as SMase D is inappropriately narrow, and there is variation in substrate specificity among homologs (Tambourgi et al. 1998; Lee and Lynch 2005; Murakami et al. 2005, da Silveira et al. 2006). Thus, our SMase D assay is a narrow assessment of the important biological activity in these venoms. Nonetheless, these assays may help identify patterns of variation in crude venom function that can focus further

work to identify features of the molecules that are associated with the evolution of functional specificity. One pattern that emerges from the integration of comparative analyses of crude venoms, the gene family analysis, and published characterizations of expression products of *SicTox* genes is evidence of evolution of functional specificity in members of the *SicTox* β clade. For identifying specific changes in the molecule responsible for functional evolution, it may be particularly fruitful to investigate functional differences among paralogs from *Sicarius*. Given the striking difference in SMase D activity between African and American *Sicarius*, it is somewhat surprising that for every paralog we recovered from African *Sicarius*, we found orthologs in American *Sicarius*. In fact, our methods recovered one paralog in American *Sicarius* (the βIB lineage) for which we did not find a candidate ortholog in African *Sicarius*. Given the support for the sister–taxon relationship with βIA, and the reduced/low SMase D activity in American *Sicarius*, *S. peruensis* and *S. terrosus* βIB genes are good candidates of genes that have a function other than SMase D. Moreover, high levels of SMase D activity in African *Sicarius* suggest that, unless our method has missed capturing mRNAs of expressed α clade genes in these species, one of the other β clade paralogs must be SMase D active.

Patterns of Sequence and Structural Conservation

The structure solved for *SicTox* gene family member *L. laeta* αIII1 (H17) (Murakami et al. 2005) allows us to analyze sequence conservation in the context of structural position and proposed active sites (fig. 5). The most conserved solvent-exposed residues are at the two openings of the barrel, either in the vicinity of the active site at the top of the barrel, or in and around a previously described "plug motif" (Cordes and Binford 2006) that caps the bottom. In the active site, all residues proposed by Murakami et al. (2005, 2006) to be directly involved in phosphate or Mg^{2+} binding (His 12, Glu 32, Asp 34, His 47, Asp 91, Lys 93, and Trp 230; fig.5*b* in green) are >98% conserved across all homologs. An additional set of surface residues near the putative phosphate binding site (Pro 50, Cys 51, Asp 52, and Asn 252; fig. 5*b* in orange) also shows >98% conservation. Two of these (Asp 52 and Asn 252) are directly hydrogen bonded to His 12. Met 250 and Tyr 228 (fig. 5*b* in yellow), which line a deep cleft beneath the phosphate site, are also >98% conserved. Many of the residues listed above have been proposed to play some significant role in chemical catalysis; the remainder could be important for general aspects of substrate binding. The role of residues in the plug motif remains unknown, but their conservation across all *SicTox* homologs as well as in *Ixodes* and *Corynebacterium* relatives (Cordes and Binford 2006) suggests some vital role either in function or maintenance of structural integrity and stability.

The distal side of the active site pocket (to the right of the Mg^{2+} and sulfate ions in fig 5*a*) contains a deep cleft whose role in substrate binding and/or catalysis is unclear at this point. Murakami et al. (2006) proposed that certain residues in this area, such as Pro 134, might contribute to substrate specificity. Interestingly, there is a contiguous set of residues in this cleft that shows >95% conservation within the α clade (Val 89, Ser 132, Pro 134, Asp 164, Ser 166, and Ser 195; see fig. 5*b* in salmon and fig. 5*c*), but subclade-specific sequence variation within the β clade. We hypothesize that these residues are involved in specific aspects of substrate binding, and consequently, we predict that members of the α clade will show relatively conserved substrate specificity profiles; by contrast, variation in these residues within the β clade may lead to more variable substrate specificity profiles among this group. It must be noted, however, that most of these residues also differ between the α clade and homologs in *Corynebacterium* (fig. 5*c*) even though both proteins are known to hydrolyze sphingomyelin substrates.

Summary of Inferred Evolutionary Mechanisms Influencing *SicTox*

Together our data indicate that evolutionary dynamics of the *SicTox* gene family are similar in some aspects and different in others from the emerging picture of evolutionary dynamics of other venom toxin lineages. Frequent duplications are consistent with a birth-and-death model of evolution (review in Nei and Rooney 2005) that has been described for other venom toxin families (e.g., Duda and Palumbi 1999; Espiritu et al. 2001; Fry et al. 2003; Lynch 2007); however, our estimates of *SicTox* duplications are lower than estimates for small peptide neurotoxins in *Conus* venoms. Once duplications occur, our data are consistent with purifying selection in the *SicTox* lineage with episodic directional selection. However, our data do not support the same level of high d_N/d_S values documented for other toxin lineages (e.g., Duda and Palumbi 1999; Kini and Chan 1999; Lynch 2007). In fact, recent estimates of average ω of 1.28 for the toxic enzyme PLA2s from snake venoms (Lynch 2007) are an order of magnitude higher than our estimates (table 5). The general pattern of *SicTox* genes having possibly lower duplication rates and lower levels of diversifying selection than other venom toxins is provocative with respect to understanding general principles that influence venom evolution. These patterns invite the consideration that some of the evolution of copy number may be best described under a model of genomic drift (Nogawa et al. 2007) in which the absolute number of paralogs is of little adaptive consequence. The little we know about the genome structure of *SicTox* is consistent with the genes lying in a region of high recombination within the genome. The coding region of *SicTox* genes in *L. arizonica* is compartmentalized into five exons, some of which are separated by large introns (Binford et al. 2005). Moreover, we have evidence that intron–exon boundaries are not conserved among paralogs (Binford GJ, unpublished data).

One limitation for understanding *SicTox* evolution is a lack of understanding of the functional role of these toxins in the complex dynamics of immobilization and/or digestion of arthropod prey. Recent expressed sequence tag analyses indicate that *SicTox* homologs are the most abundant transcripts in venoms of *L. laeta* (Fernandes-Pedrosa et al. 2008), consistent with an important functional role. Although a cytotoxic role is apparent, recent work indicates that SMase D and other phospholipase D can inhibit ion channels because of an interaction between the channel

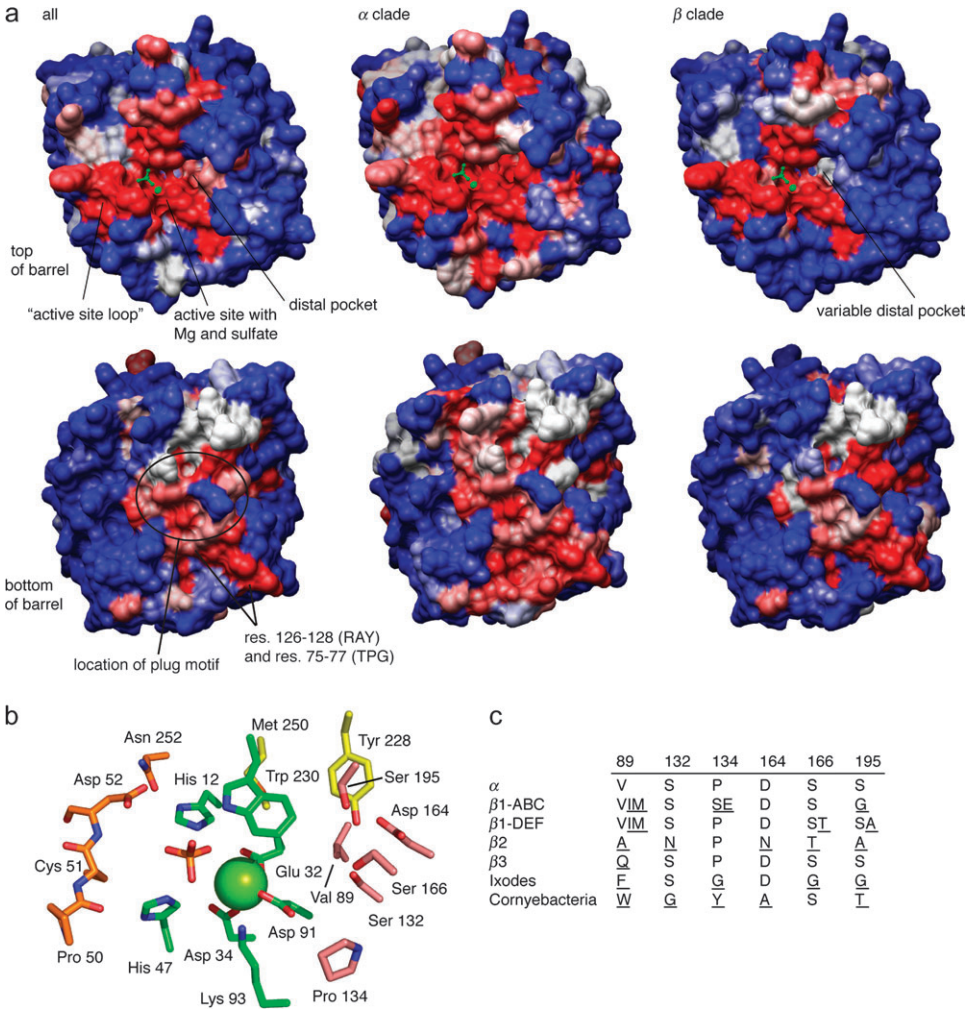

FIG. 5.—(*a*) Sequence conservation levels mapped onto the surface of *Loxosceles laeta* SMaseI (PDB ID 1XX1; chain A) for all spider sequences (left), the α clade (middle) and the β clade (right), for the top of the barrel containing the active site (top) and the bottom containing the conserved plug motif (bottom). Surface residues showing 60% conservation or less are shown in blue, whereas those showing 100% are shown in red. Intermediate levels of conservation show a range of color from blue to red, with 80% conservation appearing as white. Note that the levels of conservation observed in the α clade are higher than for the β clade. (*b*) Surface residues in and around the active site cleft showing >98% conservation (green, orange, and yellow groups) or clade-specific conservation (distal pocket residues in salmon). (*c*) Clade-specific sequence conservation patterns in the distal binding pocket. The residues shown for the α clade are >95% conserved within this group. Residue types in other clades that differ from the conserved residues in the α clade are underlined.

and the head groups of membrane phospholipids (Ramu et al. 2007; Xu et al. 2008). Thus, *SicTox* proteins may also have a neurotoxic effect on prey. As we gain insight into divergence in sequence, structure, and function more focused analyses of particular lineages and regions of the protein that have undergone directional selection will help clarify evolutionary dynamics in this lineage.

Relevance for Understanding Risks of Bites and Development of Treatments

Although the functional role of *SicTox* genes in prey capture remains unclear, the central role of SMase D activity in the pathology of human envenomation is well documented. The phylogenetic structure and diversity and insight into evolutionary dynamics of *SicTox* gene family members provides a framework for understanding the distribution of risks associated with bites from spiders in the sicariid lineage and for strategies of developing broadly effective treatments for bites.

The distribution of risks associated with bites of diverse species should be a function of the distribution of toxins in the venom of these species. The diverse set of *SicTox* homologs makes this pattern complex. It is clear that some *SicTox* gene family members are sufficient causative agents for causing dermonecrosis (table 1 and references therein), but there is much to learn about the role of the non-SMase D active members in the clinical syndrome of human envenomation. In particular, there is little understanding of the contribution of genes in the β clade. Bites from species of African *Loxosceles* from the *spinulosa* group and African *Sicarius* cause dermonecrotic lesions, and *Sicarius* envenomation can be particularly damaging (Newlands and Atkinson 1988; Van Aswegen et al. 1997). From both of these lineages, we only recovered genes in the β clade. There are no published records of the effects of

bites from New World *Sicarius*; however, injection assays of venom into mice have yielded no lesions (Alegre et al. 1977). Comparative bioassays of crude venoms and expression products using well-established rabbit models will provide better estimates of whether or not reduced SMase D activity in South American *Sicarius* venoms correlates with potential risk of bites. Results of these analyses will be particularly informative about whether or not SMase D activity is necessary for causing dermonecrosis, and the contribution of other *SicTox* gene functions to the necrotic syndrome.

Much work has focused on patterns of antigenic cross-reactivity in venoms within *Loxosceles* with the goal of developing antibody-based treatments of bites (Barbaro et al. 1996, 2005; Gomez et al. 2001; Ramos-Cerrillo et al. 2004: Olvera et al. 2006; de Roodt et al. 2007). To date, these analyses have been restricted to a subset of New World *Loxosceles*. The data presented here may help direct future work to produce treatments that are broadly, perhaps even globally, effective for treating sicariid envenomations. One important example is the phylogenetic placement of the *L. rufescens SicTox* genes. *Loxosceles rufescens* is the most cosmopolitan of all *Loxosceles* species. They are native to Mediterranean Europe and have recently dispersed to all major continents. The placement of *L. rufescens SicTox* genes in the α clade make it reasonable to predict that antivenoms developed for North and South American *Loxosceles* (Barbaro et al. 2005; Olvera et al. 2006) may be effective for treating *L. rufescens* bites. This work may also help with the possibility of designing a treatment that is effective for bites of African *Loxosceles spinulosa* group members and *Sicarius* bites. The apparent lack of α clade genes in venoms of African *spinulosa* and *Sicarius* suggests that a divergent set of genes beyond those that have been well characterized (table 1) is responsible for the clinical effects of bites of these species. For any treatment to be globally effective, antibodies raised against a diverse set of β clade genes will likely also be necessary.

Conclusion

The *SicTox* gene family is large and diverse and has undergone frequent duplications and occasional functional evolution since co-option for venomous function in the sicariid lineage. Patterns of relationships within the gene family reflect species relationships to some degree, but this is complicated by frequent duplications and likely occasional losses. We hope the framework of diversity presented here, and our nomenclature for discussing it, serve as a scaffold for more detailed analyses of mechanisms of evolution within this gene family, and for guiding particularly informative targeted analyses that will increase understanding of clinical risks associated with particular species and development of treatments for bites.

Supplementary Material

Supplemental table 1 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

For help with collecting specimens, we thank Andrew Merrell, Marjorie G. Weber, Williams Paredes, Lizette Tejada, Alfonzo Zavaleta (Peru); Facundo Labarque, Christopher Ellison, Jeremy Miller, Cristian Grismado (Argentina); Melissa Callahan, Pablo Berea Nunez, and Hartmut Kolb (Namibia); and Astri and John Leroy, Norman Larson, Wessel Pretorius, Rupert Harris, Gita Bodner (South Africa). We particularly thank Martín Ramírez and Christina Scioscia (Argentina); Tharina Bird, Helen Kolb, Joh Heschel, Leon Lotz (Namibia); Ansie Dippenaar-Schoeman, Leon Lotz (South Africa); and Gerardo Llamas and Williams Paredes (Peru) for help with obtaining collecting and export permits and with locality information. We thank Rebecca Duncan for help in the laboratory and Binford lab members, Greg Hermann and Michael Nachman, for helpful discussion and 3 anonymous reviewers for comments that improved the manuscript. Funding from Murdock Charitable Trust Life Sciences (G.J.B.), Murdock Partners in Science (S.B., G.J.B.), and NSF CAREER award IOB-0546858 (G.J.B.) supported research expenses and travel. Further support of travel was from the Student Academic Affairs Board at Lewis & Clark College (M.B., M.R.), Rogers Research Program at Lewis & Clark College (M.B., M.R.), and Instituto Bioclon S.A. de C.V. Molecular graphics images were produced using the UCSF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH P41 RR-01081).

Literature Cited

Adams ME. 2004. Agatoxins: ion channel specific toxins from the American funnel web spider, *Agelenopsis aperta*. Toxicon. 43:509–525.

Alegre AB, Meneses GO, Aguilar FPG. 1977. Peligrosidad de diez arenas communes en la costa central Peruana. Rev Per Ent. 20:63–66.

Appel MH, da Silveira RB, Chaim OM. (11 co-authors). 2008. Identification and functional characterization of a novel dermonecrotic toxin (phospholipase D) from brown spider (*Loxosceles intermedia*) venom. Biochim Biophys Acta. 1780:167–178.

Araujo SC, Castanheira P, Alvarenga LM, Mangili OC, Kalapothakis E, Chávez-Olórtegui C. 2003. Protection against dermonecrotic and lethal activities of *Loxosceles intermedia* spider venom by immunization with a fused recombinant protein. Toxicon. 41:261–267.

Barbaro KC, Knysak I, Martins R, Hogan C, Winkel K. 2005. Enzymatic characterization, antigenic cross-reactivity and neutralization of dermonecrotic activity of five *Loxosceles* spider venoms of medical importance in the Americas. Toxicon. 45:489–499.

Barbaro KC, Sousa MV, Morhy L, Eickstedt VRD, Mota I. 1996. Compared chemical properties of dermonecrotic and lethal toxins from spiders of the genus *Loxosceles* (Araneae). J Prot Chem. 15:337–343.

Bernheimer AW, Campbell BJ, Forrester LJ. 1985. Comparative toxinology of *Loxosceles reclusa* and *Corynebacterium pseudotuberculosis*. Science. 228:590–591.

Binford GJ, Callahan MS, Bodner MR, Rynerson MR, Nuñez PB, Ellison C, Duncan R. 2008. Phylogenetic relationships of *Loxosceles* and *Sicarius* spiders are consistent with western Gondwanan vicariance. Mol Phylog Evol.

Binford GJ, Cordes MHJ, Wells MA. 2005. Sphingomyelinase D from venoms of *Loxosceles* spiders: evolutionary insights from cDNA sequences and gene structure. Toxicon. 45:547–560.

Binford GJ, Wells MA. 2003. The phylogenetic distribution of sphingomyelinase D activity in venoms of Haplogyne spiders. Comp Biochem Physiol Part B. 135:25–33.

Calvete JJ, Marcinkiewicz C, Monleón D, Esteve V, Celda B, Juárez P, Sanz L. 2005. Snake venom disintegrins: evolution of structure and function. Toxicon. 45:1063–1074.

Chaim OM, Sade YB, da Silveira RB. (10 co-authors). 2006. Brown spider dermonecrotic toxin directly induces nephrotoxicity. Toxicol Appl Pharmacol. 211:64–77.

Chen K, Durand D, Farach-Colton M. 2000. NOTUNG: a program for optimizing gene family trees. J Comput Biol. 7:429–447.

Cordes MHJ, Binford GJ. 2006. Lateral gene transfer of a dermonecrotic toxin between spiders and bacteria. Bioinformatics. 22:264–268.

Cunha RB, Barbaro KC, Muramatsu D, Portaro FCV, Fontes W, de Sousa MV. 2003. Purification and characterization of loxnecrogin, a dermonecrotic toxin from *Loxosceles gaucho* brown spider venom. J Prot Chem. 22:135–146.

da Silva PH, da Silveira RB, Appel MH, Mangili OC, Gremski W, Veiga SS. 2004. Brown spiders and loxoscelism. Toxicon. 44:693–709.

da Silveira RB, Pigozzo RB, Chaim OM. (10 co-authors). 2006. Molecular cloning and functional characterization of two isoforms of dermonecrotic toxin from *Loxosceles intermedia* (Brown spider) venom gland. Biochimie. 88:1241–1253.

da Silveira RB, Pigozzo RB, Chaim OM. (10 co-authors). 2007. Two novel dermonecrotic toxins LiRecDT4 and LiRecDT5 from brown spider (*Loxosceles intermedia*) venom: from cloning to functional characterization. Biochimie. 89: 289–300.

de Andrade SA, Murakami MT, Cavalcante DP, Arni RK, Tambourgi DV. 2006. Kinetic and mechanistic characterization of the sphingomyelinases D from *Loxosceles intermedia* spider venom. Toxicon. 47:380–386.

de Castro CS, Silvestre FG, Araujo SC, Yazbeck GM, Mangili OC, Cruz I, Chávez-Olórtegui C, Kalapothakis E. 2004. Identification and molecular cloning of insecticidal toxins from the venom of the brown spider *Loxosceles intermedia*. Toxicon. 44:273–280.

de Roodt AR, Estevez-Ramirez J, Litwin S, Magana P, Olvera A, Alagon A. 2007. Toxicity of two North American *Loxosceles* (brown recluse spiders) venoms and their neutralization by antivenoms. Clin Toxicol. 45:678–687.

Diao J, Lin Y, Tang J, Liang S. 2003. cDNA sequence analysis of seven peptide toxins from the spider *Selenocosmia huwena*. Toxicon. 42:715–723.

Drummond AJ, Rambaut A. 2007. BEAST: bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 7:214.

Duda TF Jr, Kohn AJ. 2005. Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Mol Phylog Evol. 34:257–272.

Duda TF Jr, Palumbi SR. 1999. Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. Proc Natl Acad Sci. 96:6820–6823.

Duda TF Jr, Palumbi SR. 2000. Evolutionary diversification of multigene families: allelic selection of toxins in predatory cone snails. Mol Biol Evol. 17:1286–1293.

Duda TF Jr, Remigio EA. 2008. Variation and evolution of toxin gene expression patterns of six closely related venomous marine snails. Mol Ecol. 17:3018–3032.

Durand D, Halldorsson BV, Vernot B. 2006. A hybrid micro-macroevolutionary approach to gene tree reconstruction. J Comput Biol. 13:320–335.

Escoubas P. 2006. Molecular diversification in spider venoms: a web of combinatorial peptide libraries. Mol Divers. 10:545–554.

Escoubas P, Quinton L, Nicholson GM. 2008. Venomics: unraveling the complexity of animal venoms with mass spectrometry. J Mass Spec. 43:279–295.

Espiritu DJD, Watkins M, Dia-Monje V, Cartier GE, Cruz LJ, Olivera BM. 2001. Venomous cone snails: molecular phylogeny and the generation of toxin diversity. Toxicon. 39:1899–1916.

Estrada G, Villegas E, Corzo G. 2007. Spider venoms: a rich source of acylpolyamines and peptides as new leads for CNS drugs. Nat Reprod Rep. 24:145–161.

Felicori L, Araujo SC, de Avila RAM, Sachez EF, Granier C, Kalapothakis E, Chávez-Olórtegui C. 2006. Functional characterization and epitope analysis of a recombinant dermonecrotic protein from *Loxosceles intermedia* spider. Toxicon. 48:509–519.

Fernandes-Pedrosa MF, Junqueira-de Azevedo ILM, Gonçalves-de-Andrade RM, Kobashi LS, Almeida DD, Ho PL, Tambourgi DV. 2008. Transcriptome analysis of *Loxosceles laeta* (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. BMC Genomics. 9:279.

Fernandes-Pedrosa MF, Junqueira-de Azevedo ILM, Gonçalves-de-Andrade RM, van den Berg CW, Ramos CRR, Ho PL, Tambourgi DV. 2002. Molecular cloning and expression of a functional dermonecrotic and haemolytic factor from *Loxosceles laeta* venom. Biochem Biophys Res Commun. 298:638–645.

Fry BG. 2005. From genome to "venome": molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. Genome Res. 15:403–420.

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

Fry BG, Wüster W, Kini RM, Brusic V, Khan A, Venkataraman D, Rooney AP. 2003. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. J Mol Evol. 57:110–129.

Gao L-Z, Innan H. 2004. Very low gene duplication rate in the yeast genome. Science. 306:1367–1370.

Geren CR, Chan TK, Howell DE, Odell GV. 1976. Isolation and characterization of toxins from brown recluse spider venom (*Loxosceles reclusa*). Arch Biochem Biophys. 174:90–99.

de Gerschman PBS, Schiapelli RD. 1979. Caracteres morfologicos validos en la sistematica del genero Sicarius (Walckenaer 1847) Aranae: Sicariidae. Acta Zool Lilloana. 35:87–96.

Gertsch WJ, Ennik F. 1983. The spider genus *Loxosceles* in North America, Central America. and the West Indies (Aranae, Loxoscelidae). Bull Am Mus Nat Hist. 175:264–360.

Gomez HF, Miller MJ, Waggener MW, Lankford HA, Warren JS. 2001. Antigenic cross-reactivity of venoms from medically important North American *Loxosceles* spider species. Toxicon. 39:817–824.

Gu Z, Cavalcanti A, Chen F-C, Bouman P, Li W-H. 2002. Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. Mol Biol Evol. 19:256–262.

Jiang L, Peng L, Chen J, Zhang Y, Xiong X, Liang S. 2008. Molecular diversification based on analysis of expressed sequence tags from the glands of the Chinese bird spider *Ornithoctonus huwena*. Toxicon. 51:1479–1489.

Kalapothakis E, Araujo SC, de Castro CS, Mendes TM, Gomez MV, Mangili OC, Gubert IC, Chávez-Olórtegui C. 2002. Molecular cloning, expression and immunological properties of LiD1, a protein from the dermonecrotic family of *Loxosceles intermedia* spider venom. Toxicon. 40: 1691–1699.

Kalapothakis E, Chatzaki M, Gonçalves-Dornelas H. (10 co-authors). 2007. The Loxtox protein family in *Loxosceles intermedia* (Mello-Leitão) venom. Toxicon. 50:938–946.

King GF, Teford HW, Maggio F. 2002. Structure and function of insecticidal neurotoxins from Australian funnel-web spiders. Journal of Toxicology: Toxin Reviews. 21:359–389.

King GF. 2007. Modulation of insect Ca_v channels by peptide spider toxins. Toxicon. 49:513–530.

Kini RM, Chan YM. 1999. Accelerated evolution and molecular surface of venom phospholipase A_2 enzymes. J Mol Evol. 48:125–132.

Kordis D, Gubensek F. 2000. Adaptive evolution of animal toxin multigene families. Gene. 261:43–52.

Kuhn-Nentwig L. 2003. Antimicrobial and cytolytic peptides of venomous arthropods. Cell Mol Life Sci. 60:2651–2668.

Kurpiewski G, Forrester LJ, Barrett JT, Campbell BJ. 1981. Platelet aggregation and sphingomyelinase D activity of a purified toxin from the venom of *Loxosceles reclusa*. Biochim Biophys Acta. 678:467–476.

Lee S, Lynch KR. 2005. Brown recluse spider (*Loxosceles reclusa*) venom phospholipase D (PLD) generates lysophosphatidic acid. Biochem J. 391:317–323.

Li M, Fry BG, Kini RM. 2005. Putting the brakes on snake venom evolution: the unique molecular evolutionary patterns of *Aipysurus eydouxii* (marbled sea snake) phospholipase A_2 toxins. Mol Biol Evol. 22:934–941.

Lynch M, Conery JS. 2003. The origins of genome complexity. Science. 302:1401–1404.

Lynch VJ. 2007. Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A_2 genes. BMC Evol Biol. 7:2.

Machado LF, Laugesen S, Botelho ED, Ricart CAO, Fontes W, Barbaro KC, Roepstorff P, Sousa MV. 2005. Proteome analysis of brown spider venom: identification of loxnecrogin isoforms in *Loxosceles gaucho* venom. Proteomics. 5: 2167–2176.

Maddison DR, Maddison WP. 2005. MacClade: analysis of phylogenetic and character evolution. Sunderland (MA): Sinauer and Associates.

Merchant ML, Hinton JF, Geren CR. 1998. Sphingomyelinase D activity of brown recluse spider (*Loxosceles reclusa*) venom as studied by ^{31}P-NMR: effects on the time-course of sphingomyelin hydrolysis. Toxicon. 36:537–545.

Murakami MT, Fernandes-Pedrosa MF, de Andrade SA, Gabdoulkhakov A, Betzel C, Tambourgi DV, Arni RK. 2006. Structural insights into the catalytic mechanism of sphingomyelinases D and evolutionary relationship to glycerophosphodiester phosphodiesterases. Biochem Biophys Res Comm. 342:323–329.

Murakami MT, Fernandes-Pedrosa MF, Tambourgi DV, Arni RK. 2005. Structural basis for metal ion coordination and the catalytic mechanism of sphingomyelinases D. J Biol Chem. 280:13658–13664.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 39:121–152.

Newlands G, Atkinson P. 1988. Review of southern African spiders of medical importance, with notes on the signs and symptoms of envenomation. S Afr Med J. 73:235–239.

Nogawa M, Kawahara Y, Nei M. 2007. Genomic drift and copy number variation of sensory receptor genes in humans. Proc Natl Acad Sci USA. 104:20421–20426.

Olivera BM. 2002. Conus venom peptides: reflections from the biology of clades and species. Ann Rev Ecol System. 33:25–47.

Olvera A, Ramos-Cerrillo B, Estevez J. (11 co-authors). 2006. North and South American *Loxosceles* spiders: development of a polyvalent antivenom with recombinant sphingomyelinases D as antigens. Toxicon. 48(1):64–74.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera: a visualization system for exploratory research and analysis. J Comput Chem. 25:1605–1612.

Platnick NI. 2008. The world spider catalog, version 8.5. American Museum of Natural History. [Internet]. [cited July 2008]. Available from: http://research.amnh.org/entomology/spiders/catalog/index.html

Platnick NI, Coddington JA, Forster RR, Griswold CE. 1991. Spinneret morphology and the phylogeny of Haplogyne spiders (Araneae, Araneomorphae). Amer Mus Novitat. 3016:1–73.

Pretel F, Gonçalves-de-Andrade RM, Magnoli FC, da Silva MER, Ferreira JMC Jr, van den Berg CW, Tambourgi DV. 2005. Analysis of the toxic potential of venom from *Loxosceles adelaida*, a Brazilian brown spider from karstic areas. Toxicon. 45:449–458.

Ramos-Cerrillo B, Olvera A, Odell GV, Zamudio F, Paniagua-Solis J, Alagon A, Stock RP. 2004. Genetic and enzymatic characterization of sphingomyelinase D isoforms from the North American fiddleback spiders *Loxosceles boneti* and *Loxosceles reclusa*. Toxicon. 44:507–514.

Ramu Y, Xu Y, Lu Z. 2007. Inhibition of the CFTR Cl^- channel function caused by enzymatic hydrolysis of sphingomyelin. Proc Natl Acad Sci USA. 104:6448–6453.

Rash LD, Hodgson WC. 2002. Pharmacology and biochemistry of spider venoms. Toxicon. 40:225–254.

Ribeiro ROS, Chaim OM, da Silveira RB. (11 co-authors). 2007. Biological and structural comparison of recombinant phospholipase D toxins from *Loxosceles intermedia* (brown spider) venom. Toxicon. 50:1162–1174.

Rodríguez de la Vega RC, Possani LD. 2004. Current views on scorpion toxins specific for K^+-channels. Toxicon. 43:865–875.

Rodríguez de la Vega RC, Possani LD. 2005. Overview of scorpion toxins specific for Na+ channels and related peptides: biodiversity, structure-function relationships and evolution. Toxicon. 46:831–844.

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

Ronquist F, Huelsenbeck JP, van der Mark P. 2005. MrBayes 3.1 manual, draft 5/26/2005. [Internet]. [cited May 2008]. Available from: http://mrbayes.csit.fsu.edu/manual.php.

Schulz S. 1997. The chemistry of spider toxins and spider silk. Agnew Chem Int Ed Engl. 36:314–326.

Silvestre FG, de Castro CS, de Moura JF. (11 co-authors). 2005. Characterization of the venom from the Brazilian brown spider *Loxosceles similis* Moenkhaus, 1898 (Araneae, Sicariidae). Toxicon. 46:927–936.

Sollod BL, Wilson D, Zhaxybayeva O, Gogarten JP, Drinkwater R, King GF. 2005. Were spiders the first to use combinatorial peptide libraries? Peptides. 26:131–139.

Soucek A, Soucková A, Patocka F. 1967. Inhibition of the activity of alpha-toxin of *Clostridium* perfingens by toxic filtrates of corynebacteria. J Hyg Epidemiol Microbiol Immunol. 11:123–124.

Swanson DL, Vetter RS. 2005. Bites of brown recluse spiders and suspected necrotic arachnidism. N Engl J Med. 352:700–707.

Swofford DL. 1998. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer.

Tambourgi DV, Fernandes-Pedrosa MF, van den Berg CW, Gonçalves-de-Andrade RM, Ferracini M, Paixao-Cavalcante D, Morgan BP, Rushmere NK. 2004. Molecular cloning, expression, function and immunoreactivities of members of a gene family of sphingomyelinases from *Loxosceles* venom glands. Mol Immunol. 41:831–840.

Tambourgi DV, Magnoli FC, van den Berg CW, Morgan BP, de Araujo PS, Alves EW, da Silva WD. 1998. Sphingomyelinases in the venom of the spider *Loxosceles intermedia* are responsible for both dermonecrosis and complement-dependent hemolysis. Biochem Biophys Res Commun. 251: 366–373.

Tedford HW, Sollod BL, Maggio F, King GF. 2004. Australian funnel-web spiders: master insecticide chemists. Toxicon. 43:601–618.

Thompson J, Gibson T, Plewniak F, Jeanmougin F, Higgins D. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res. 25:4876–4882.

Truett AP 3rd, King LE Jr. 1993. SphingomyelinaseD: a pathogenic agent produced by bacteria and arthropods. Adv Lipid Res. 26:275–291.

Van Aswegen G, van Rooyen JM, van der Nest DG, Veldman FJ, de Villiers TH, Oberholzer G. 1997. Venom of a six-eyed crab spider, *Sicarius testaceus* (Purcell, 1908), causes necrotic and haemorrhagic lesions in the rabbit. Toxicon. 35: 1149–1152.

Vernot B, Stolzer M, Goldman A, Durand D. 2007. Computational systems bioinformatics: CSB2007. London: Imperial College Press. p. 441–452.

Vetter RS. 2008. Spiders of the genus *Loxosceles* (Araneae, Sicariidae): a review of biological, medical and psychological aspects regarding envenomations. J Arachnol. 1:150–163.

Xu Y, Ramu Y, Lu Z. 2008. Removal of phospho-head groups of membrane lipids immobilizes voltage sensors of K^+ channels. Nature. 451:826–830.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 15:568–573.

Yang Z. 2007. PAML4: a program package for phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24: 1586–1591.

Yang Z, Bielawski B. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 15:496–503.

Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J Mol Evol. 46: 409–418.

Yang Z, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol. 19:908–917.

Yang Z, Nielsen R, Goldman N, Pedersen A-MK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 155:431–449.

Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22:1107–1118.

Adriana Briscoe, Associate Editor

Accepted November 17, 2008