

**Table S1.** Key features of 72 myxomycete S1389 group I introns included in this work.

Host species	Acc No	Isolate	Size <sup>1</sup>	Insert <sup>2</sup>	P5d <sup>3</sup>	P7 <sup>4</sup>
<b>Order Physarales</b>						
<i>Diderma fallax</i> <sup>5</sup>	MZ313547	It-K52	532 bp	-	-	A1
<i>D. globosum</i>	DQ903677	AMFD110	722 bp	-	-	A1
<i>D. meyeræ</i>	HE614614	It-K61	735 bp	-	-	A2
<i>D. meyeræ</i> <sup>5</sup>	MZ313546	It-K68	921 bp	-	-	A2
<i>D. niveum</i> <sup>5</sup>	MZ313548	Fr-M26	749 bp	-	-	A1
<i>D. niveum</i>	AM231291	Fr-K10	682 bp	-	-	A1
<i>D. niveum</i>	HE614616	It-K66	683 bp	-	-	A1
<i>D. niveum</i>	HE614617	Uk-K79	683 bp	-	-	A1
<i>D. saundersii</i> <sup>5</sup>	MZ313549	Mx-K30	532 bp	-	-	A1
<i>D. testeceum</i>	AM231292	Pr3-1	529 bp	-	-	G1
<i>Didymium iridis</i>	AJ938151	CR19-1	523 bp	-	-	A1
<i>D. iridis</i>	AJ938150	CUR1-4	637 bp	-	-	A1
<i>D. nigripes</i>	AF239230	-	654 bp	-	-	A1
<i>D. squamulosum</i>	AM231293	CR10	792 bp	-	-	U1
<i>Fuligo leviderma</i>	DQ903676	AMFD130	517 bp	-	-	A1
<i>Lepidoderma carestianum</i>	HE614618	It-K71	393 bp	-	-	A2
<i>L. crustaceum</i>	HE614619	It-K62	396 bp	-	-	A2
<i>L. peyerimhoffii</i> <sup>5</sup>	MZ313552	It-K63	388 bp	-	+	A2
<i>L. tigrinum</i>	DQ903678	AMFD192	785 bp	-	-	A1
<i>Mucilago crustacea</i> <sup>5</sup>	MZ313554	No-K94	1274 bp	DR-P9	-	A1
<b>Order Stemonitales</b>						
<i>Brefeldia maxima</i>	JQ031957	MM24519	1304 bp	HEG-P1	+	A3
<i>Colloderma oculatum</i>	JQ031959	HS2885	421 bp	-	+	A3
<i>C. robustum</i>	JQ031960	AMFD270	384 bp	-	+	A1
<i>Comatricha nigra</i>	DQ903683	AMFD155	788 bp	-	+	A1
<i>C. pseudoalpina</i>	DQ903673	MM23892	815 bp	-	+	U1
<i>Diacheopsis pauxilla</i>	JQ031966	MM29883	628 bp	-	-	A1
<i>Diachea subsessilis</i>	JQ031964	MM24463	919 bp	-	-	U1
<i>Lamproderma aeneum</i>	JQ031969	MM36255	545 bp	-	-	A3
<i>L. arcyrrioides</i>	JQ031973	MM37005	1006 bp	DR-P9	+	A2
<i>L. cacographicum</i>	JQ031976	AMFD310	957 bp	DR-P9	-	A1
<i>L. columbinum</i>	HQ687204	F2	503 bp	-	-	A2
<i>L. columbinum</i>	HQ687196	106	563 bp	-	-	U2
<i>L. columbinum</i>	HQ687197	63b	474 bp	-	-	A2
<i>L. columbinum</i>	HQ687200	132	480 bp	-	+	A3
<i>L. disseminatum</i>	JQ031978	AMFD38	498 bp	-	-	A1
<i>L. echinosporum</i>	JQ031980	AMFD136	486 bp	-	+	A3
<i>L. echinosporum</i>	JQ031979	AK06016	481 bp	-	+	A3
<i>L. pseudomaculatum</i>	JQ031985	MM37354	411 bp	-	-	A3
<i>L. puncticulatum</i>	HQ687194	172	514 bp	-	+	A2
<i>L. puncticulatum</i>	HQ687202	162	511 bp	-	+	A1
<i>L. puncticulatum</i>	HQ687195	3	517 bp	-	+	A1
<i>L. sauteri</i>	DQ903674	AMFD208	543 bp	-	+	A1
<i>L. zonatum</i>	DQ903672	MM21644	434 bp	-	-	A1
<i>Meriderma carestiae</i>	JQ031999	MM35985	566 bp	-	-	A1
<i>M. carestiae</i>	DQ903671	AMFD173	566 bp	-	-	A1
<i>M. cribrarioides</i>	JQ032000	MM37106	538 bp	-	-	A1
<b>Order Liceales</b>						
<i>Licea marginata</i>	JX481296	DWM7368	455 bp	-	+	A3
<b>Order Trichiales</b>						
<i>Calomyxa metallica</i>	JX481284	AMFD483	532 bp	-	-	A3
<i>Trichia varia</i>	KM494993	sc22370	1531 bp	HEG-P1	-	A3
<i>T. varia</i>	KM494994	LE259268	586 bp	-	-	A3
<i>T. varia</i>	KM494995	LE259461	586 bp	-	-	A3
<i>T. varia</i>	KM494996	JVR848	1531 bp	HEG-P1	-	A3
<i>T. varia</i>	KM494997	sc22386	559 bp	-	-	A1
<i>T. varia</i>	KM494998	sc22408	559 bp	-	-	A1
<i>T. varia</i>	KM494999	sc22409	559 bp	-	-	A1
<i>T. varia</i>	KM495003	sc22442	559 bp	-	-	A1
<i>T. varia</i>	KM495005	sc22517	559 bp	-	-	A1
<i>T. varia</i>	KM495006	sc27697	559 bp	-	-	A1

<i>T. varia</i>	KM495009	sc27742	559 bp	-	-	A1
<i>T. varia</i>	KM495010	sc27839	559 bp	-	-	A1
<i>T. varia</i>	KM495018	KRAM M-1585	559 bp	-	-	A1
<i>T. varia</i>	KM495019	sc27686	559 bp	-	-	A1
<i>T. varia</i>	KM495020	sc27737	559 bp	-	-	A1
<i>T. varia</i>	KM495021	sc27850c1	559 bp	-	-	A1
<i>T. varia</i>	KM495022	sc27860c4	559 bp	-	-	A1
<i>T. varia</i>	KM495023	sc27507	559 bp	-	-	A1
<i>T. varia</i>	KM495024	sc27648c1	559 bp	-	-	A1
<i>T. varia</i>	KM495026	sc27667c1	559 bp	-	-	A1
<i>T. varia</i>	KM495027	sc27667c2	559 bp	-	-	A1
<i>T. varia</i>	KM495028	sc27772c2	559 bp	-	-	A1
<i>T. varia</i>	KM495029	sc27850c2	559 bp	-	-	A1
<i>T. varia</i>	KM495030	LE254838	1544 bp	HEG-P1	-	A1

Host species	Acc No	Isolate	Size	Insert	P5d	P7
--------------	--------	---------	------	--------	-----	----

Notes:

<sup>1</sup>Size in base pairs of the S1389 group I intron insertion.

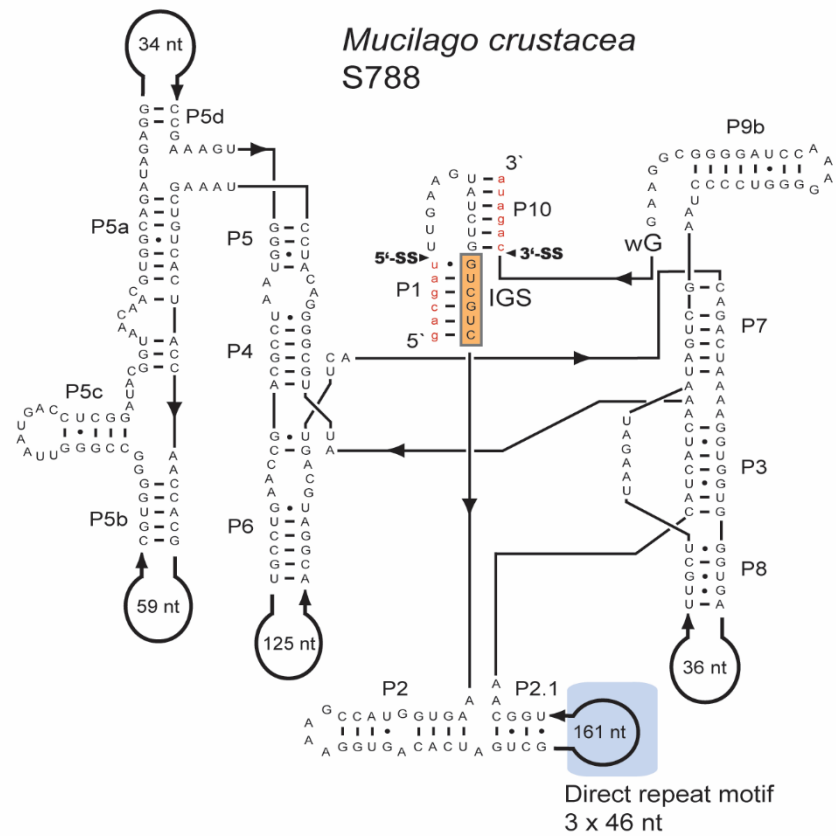
<sup>2</sup>Type of large insertions in group I ribozyme paired segment P1 or P9. HEG, homing endonuclease gene; DR, direct repeat motifs.

<sup>3</sup>Presence or absence of the optional paired segment P5d.

<sup>4</sup>G-binding sequence variants in segment P7, as indicated in Figure 3.

<sup>5</sup>This work.

a



b

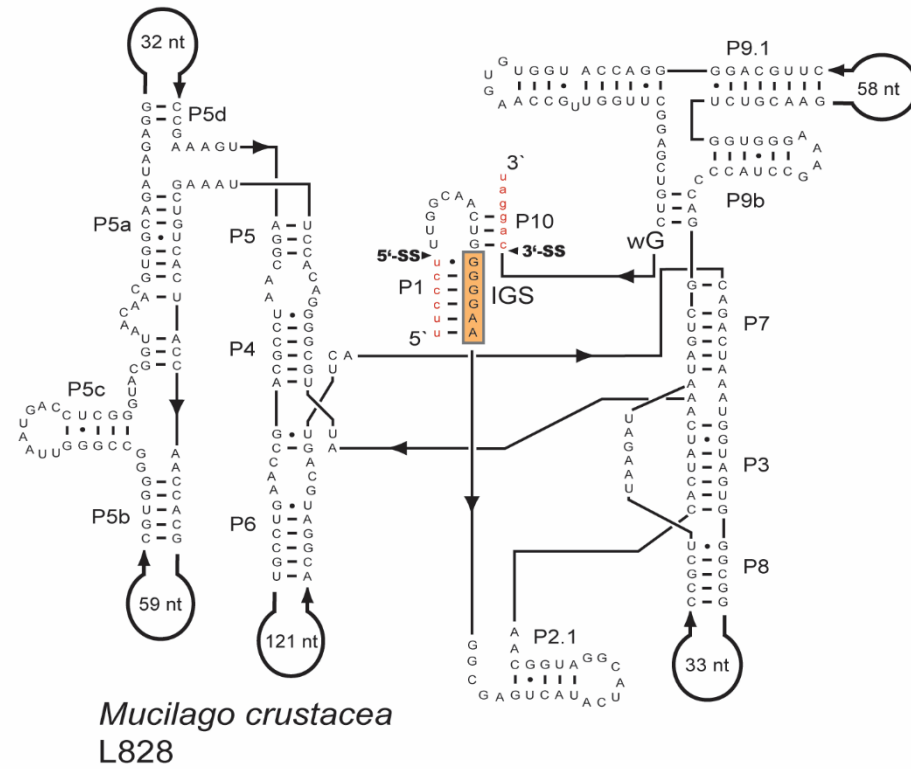
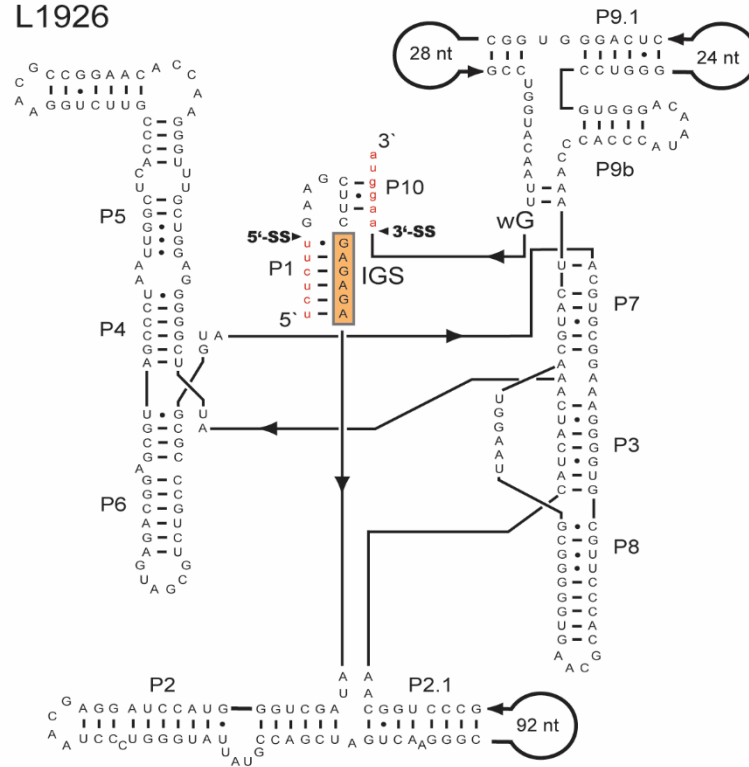


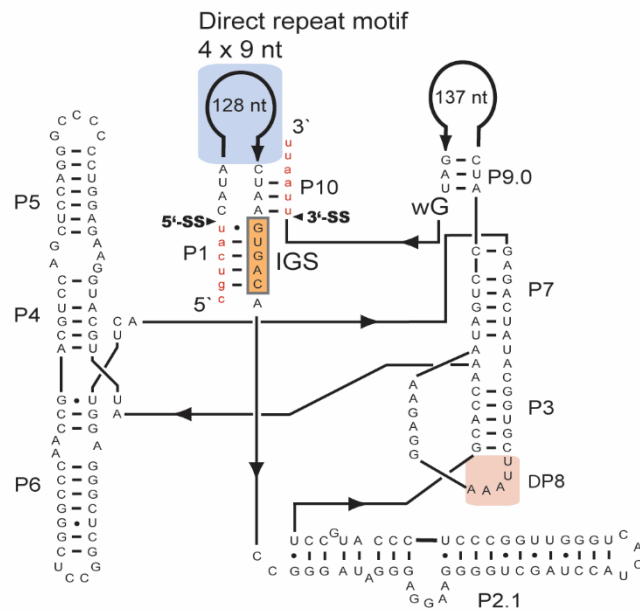
Figure S1 (a&b)

C

*Mucilago crustacea*  
L1926



d

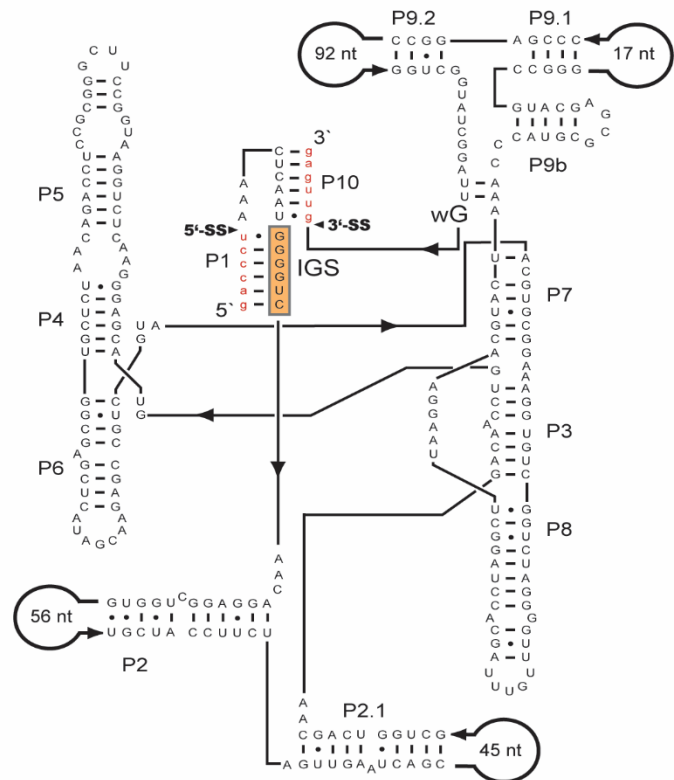


*Mucilago crustacea*  
L1949

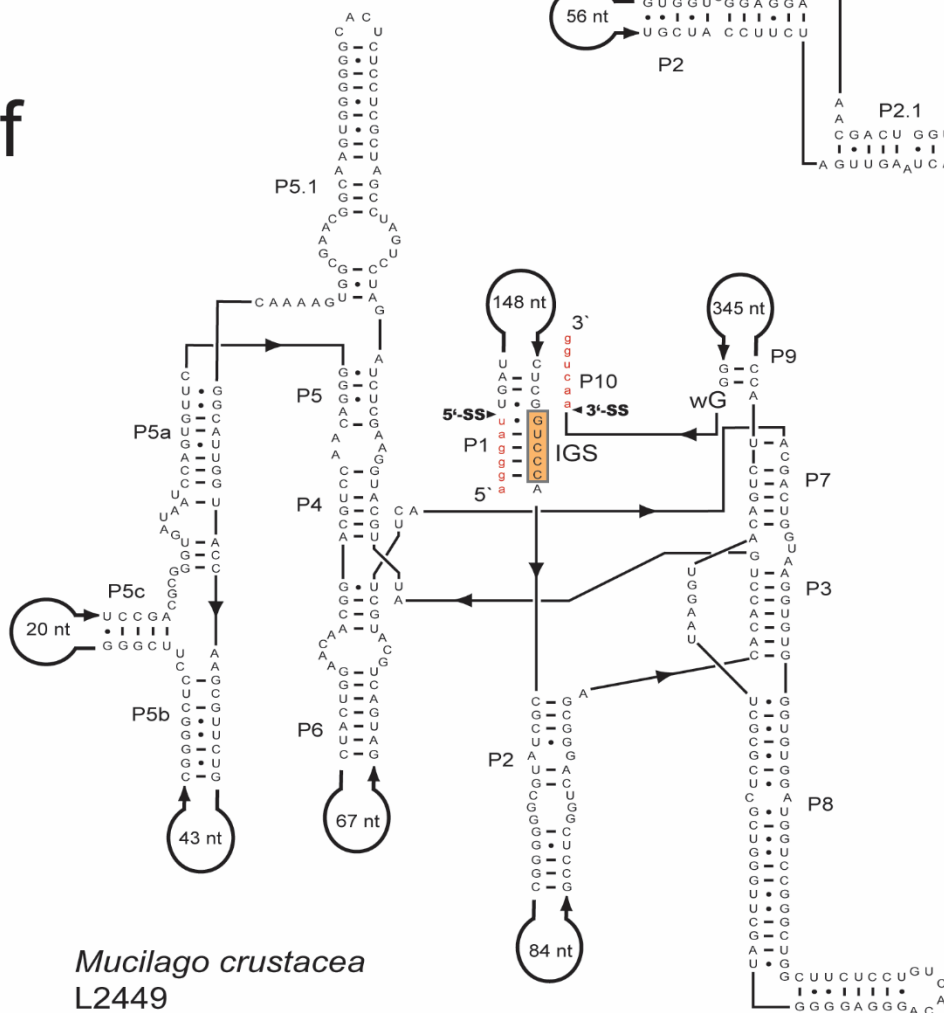
Figure S1 (c&d)

e

*Mucilago crustacea*  
L2066



f



*Mucilago crustacea*  
L2449

Figure S1 (e&f)

**Figure S1:** Secondary structure diagrams of the *Mucilago crustacea* rDNA group I introns. (a) S788 is a group IC1 intron in the SSU rRNA gene. Note that segment P2.1 (blue box) contains a 46-nt direct repeat motif. (b) L828 is a group IC1 intron in

the LSU rRNA gene. (c) L1926 is a group IE intron in the LSU rRNA gene. (d) L1949 is an unclassified group I intron in the LSU rRNA gene. Note that the P8 segment of the catalytic core is missing (pink box), and that segment P1 (blue box) contains a 9-nt direct repeat motif. (e) L2066 is a group IE intron in the LSU rRNA gene. (f) L2449 is a group IC1 intron in the LSU rRNA gene. Note that P1 and P9 contain extension sequences, but with no reading frame or direct repeat features. Common note to all intron diagrams: P1-P10, paired RNA segments; 5' SS and 3'SS, exon-intron splice sites; IGS, internal guide sequence; ωG, last nucleotide residue of intron; red lowercase letters, exon sequences.

[illegible]

**Figure S2:** Sequence alignment of core structure nucleotides of myxomycete S1389 group I intron. Dashes correspond to deleted positions. Secondary structure paired segments (P1-P8) are shown above the alignment. Intron sequences are indicated by GenBank accession numbers. V01416 and MZ313554\* correspond to the out-group *Tetrahymena* intron Tth.L1925 and *Mucilago* intron Mcr.S788, respectively.



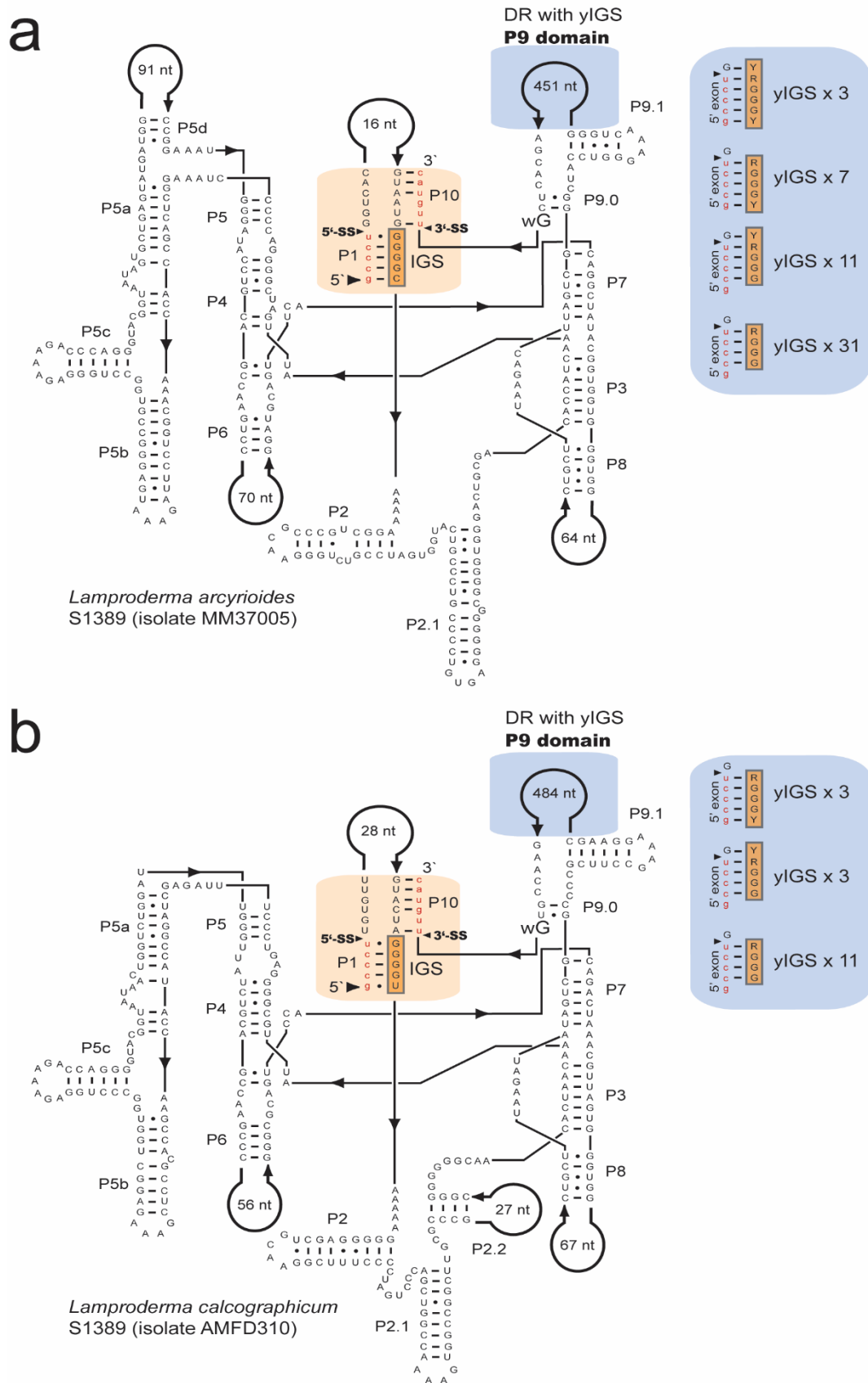
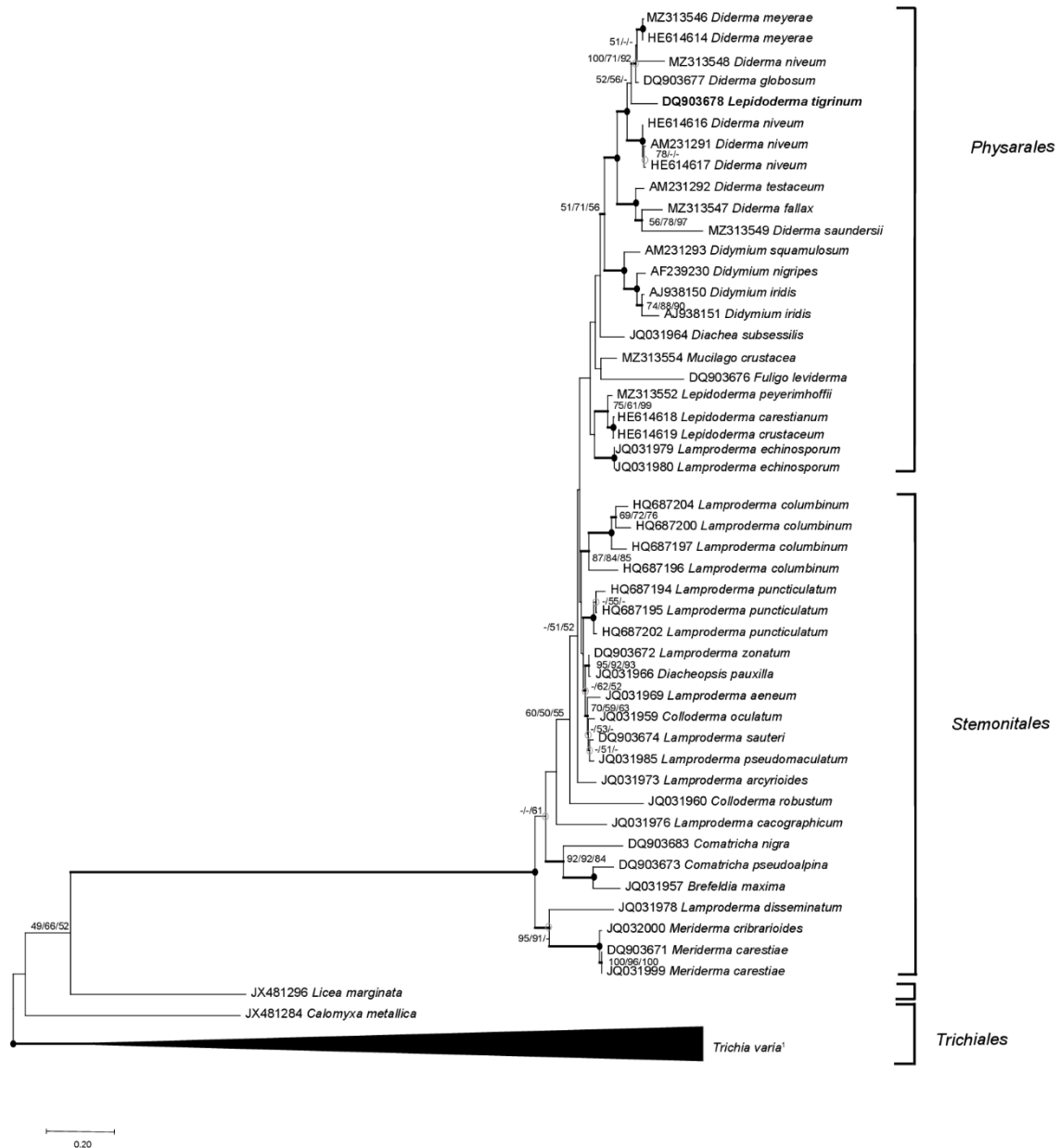


Figure S3

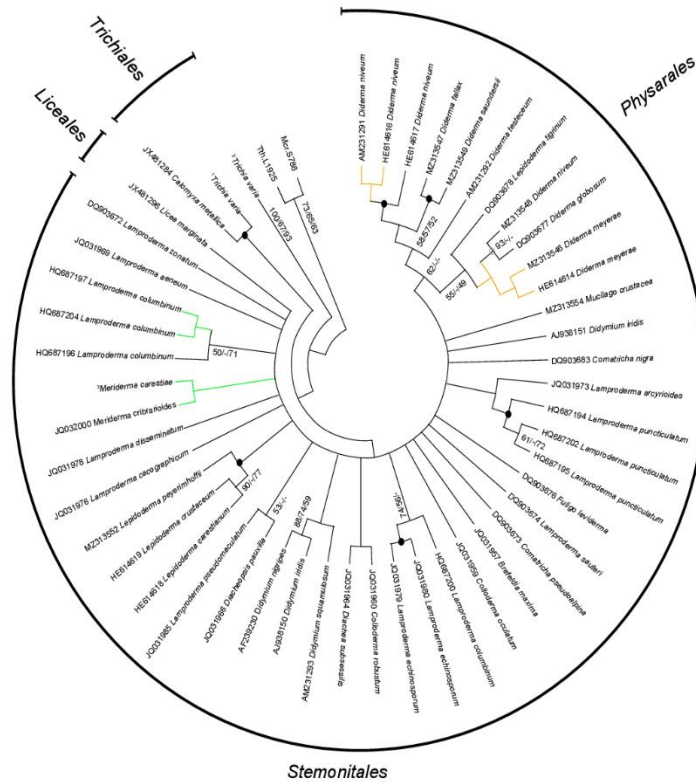
**Figure S3:** Secondary structure diagrams of (a) *Lamproderma arcyrrioides* and (b) *Lamproderma arcyrrioides* S1389 group I intron. P1-P10, paired RNA segments; 5' SS and 3' SS, exon-intron splice sites; IGS, internal guide sequence. The introns contain large extensions in P9 that contains direct repeat motifs (blue box) with ΨIGS sequences motifs.



diagram of *T. varia* group I introns (KM494996 and KM495030). Homing endonuclease gene (HEG) insertions are located in segment P1. P1-P10, paired RNA segments; 5' SS and 3'SS, exon-intron splice sites; IGS, internal guide sequence; ωG, last nucleotide residue of intron; red lowercase letters, exon sequences; black lowercase letters, nucleotide positions only present in one of the introns. R represents A or G (purine); Y represents C or U (pyrimidine); K represents G or U; M represents A or C; S represents C or G; W represents A or U.



**Figure S5:** Molecular phylogeny of myxomycete taxa based SSU rDNA sequences. The SSU topology is obtained by neighbor-joining (NJ) analysis of 72 taxa and 1,575 nt aligned positions (SSU dataset 1; Table S1). The tree is rooted with the *T. varia* SSU rDNA sequence. The NJ, maximum likelihood (ML), and maximum parsimony (MP) bootstrap replicates ( $\geq 50\%$ ) are given at each node. Bayesian posterior (BI) probabilities ( $\geq 0.95$ ) are shown in bold branches. Black dots at branch points; maximum support in NJ, ML, and MP ( $\geq 97\%$ ). The scale bar indicates the fraction of substitutions per site.



**Figure S6:** Molecular phylogeny of myxomycete S1389 group I introns. The intron topology is obtained by neighbor-joining (NJ) analysis of 74 taxa (including group introns from *Tetrahymena* (Tth.1925) and *Mucilago* (Mcr.S788), and 166 nt aligned positions (intron dataset 2; Table S1). The NJ-, maximum likelihood (ML), and maximum parsimony (MP) bootstrap replicates ( $\geq 50\%$ ) are given for each node. Bayesian posterior (BI) probabilities ( $\geq 0.95$ ) are shown in bold branches. Black dots at branch site indicate maximum support in NJ, ML, and MP ( $\geq 98\%$ ). Green line indicates maximum support in NJ, ML and MP ( $\geq 90\%$ ). Orange line indicates maximum support in NJ, ML, and MP ( $\geq 70\%$ ). Gray circles indicate alternative topologies in NJ, ML, and MP. <sup>1</sup> *Tricia varia* clade is composed of four *T. varia* isolates (KM494993-6). <sup>2</sup> *Tricia varia* clade is composed of twenty *T. varia* isolates (KM494997-9, KM495003, KM495005, KM495006, KM495009, KM495010, KM495018-30). <sup>3</sup> *Meriderma carestiae* clade is composed of two *M. carestiae* isolates (JQ031999 and DQ903671). The scale bar indicates the fraction of substitutions per site.