Supplementary Material
Part 1

Identification of species by multiplex analysis of variable-length sequences

Filipe Pereira[1], João Carneiro[1], Rune Matthiesen[1], Barbara van Asch[1,2], Nádia Pinto[1,2,3], Leonor Gusmão[1] and António Amorim[1,2]

[1] Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), R. Dr. Roberto Frias s/n, 4200-465 Porto, Portugal
[2] Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal
[3] Centro de Matemática da Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

Supplementary Table S1. Taxonomic groups used in the development of the SPInDel method.

Eukaryotes
Eukaryota; Viridiplantae
Eukaryota; Stramenopiles
Eukaryota; Fungi
Eukaryota; Alveolata
Eukaryota; Metazoa; Porifera
Eukaryota; Metazoa; Cnidaria
Eukaryota; Metazoa; Platyhelminthes
Eukaryota; Metazoa; Echinodermata
Eukaryota; Metazoa; Nematoda
Eukaryota; Metazoa; Arthropoda
Eukaryota; Metazoa; Mollusca
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Testudines
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Metatheria
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria
Prokaryotes
Archaea; Crenarchaeota
Bacteria; Tenericutes
Viruses
Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae; Lentivirus
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae
Viruses; ssRNA negative-strand viruses; Mononegavirales; Rhabdoviridae

Supplementary Table S2. General description of ribosomal RNA gene sequence alignments in eukaryotic and prokaryotic groups.

Taxonomic group	Number of reference sequences (*N*)	Species with duplicated rRNAs	Number of haplotypes[a]	Number of sites in alignments	Number of sites without gaps	Number of sites with gaps		Number of variable sites[a] (S)		Number of mutations[a] (Eta)	Nucleotide diversity per site[a] (Pi)	GC content[a]
						Nº	%	Nº	%			
Eukaryotes												
Viridiplantae (Vi)	23	3	17	11260	1856	9404	83.5	1198	64.55	1825	0.151	0.482
Stramenopiles (St)	14	1	14	3299	2224	1075	32.6	1312	58.99	1935	0.244	0.381
Fungi (Fu)	42	0	42	12332	1075	11257	91.3	909	84.56	1758	0.293	0.369
Alveolata (Al)	6	5	6	2422	2179	243	10.0	797	36.58	899	0.154	0.297
Porifera (Po)	20	0	20	5674	1582	4092	72.1	953	60.24	1447	0.192	0.402
Cnidaria (Cn)	32	0	29	3169	1341	1828	57.7	878	65.47	1334	0.218	0.377
Platyhelminthes (Pl)	29	0	29	2196	1248	948	43.2	868	69.55	1644	0.302	0.354
Echinodermata (Ec)	20	0	20	1942	814	1128	58.1	508	62.41	890	0.255	0.384
Nematoda (Ne)	29	1	29	2170	458	1712	78.9	424	92.58	801	0.329	0.250
Arthropoda (Ar)	204	0	204	4112	644	3468	84.3	513	79.66	1057	0.205	0.366
Mollusca (Mo)	50	0	50	2043	600	1443	70.6	448	74.67	1000	0.316	0.368
Actinopterygii (Ac)	539	0	536	4223	1653	2570	60.9	1217	73.62	2321	0.127	0.492
Amphibia (Am)	84	0	84	2916	1858	1058	36.3	1259	67.76	2271	0.201	0.407
Testudines (Te)	29	0	29	3128	2422	706	22.6	1120	46.24	1639	0.118	0.403
Lepidosauria (Le)	69	0	69	3194	1801	1393	43.6	1234	68.52	2378	0.239	0.447
Aves (Av)	98	0	97	3014	2335	679	22.5	1244	53.28	2181	0.149	0.466
Metatheria (Me)	31	0	31	2767	2378	389	14.1	1066	44.83	1826	0.148	0.384
Eutheria (Eu)	237	0	236	3433	2087	1346	39.2	1344	64.40	2343	0.134	0.417
Prokaryotes												
Crenarchaeota (Archaea)	48	n/a	44	1387	1243	144	10.4	588	47.30	867	0.124	0.642
Tenericutes (Bacteria)	57	n/a	55	1364	1070	294	21.6	621	58.04	996	0.171	0.473

[a] Sites with alignment gaps or missing data were excluded.

.

Supplementary Table S3. Classification performance obtained for two taxonomic groups on the species and genus level. The numbers in parenthesis represent the total number of classifications made followed by the total number of profiles in the database.

Taxonomic group	Accuracy (species)	Accuracy (genus)
Actinopterygii	0.67 (14/539)	0.99 (167/539)
Eutheria	1 (31/237)	0.96 (107/237)

Supplementary Table S4. Samples used in the development of the SPInDel profiling kit for identification of 10 eutherian species.

Species	*n*	Specimen	Origin/breed
Cat (*Felis catus*)	10	buccal swab	Persian (1), Siamese (1), calico (2), undetermined (6)
Cattle (*Bos taurus*)	10	liver	supermarket
Dog (*Canis familiaris*)	10	blood	Dogo Argentino (4), Serra de Aires (3), mixed-breed dogs (3)
Goat (*Capra hircus*)	10	blood	Serrana (2), Serpentina (2), Bravia (2), Charnequeira (2), Algarvia (2)
Horse (*Equus caballus*)	2	blood	Iberian Peninsula
Human (*Homo sapiens*)	10	blood	Iberian Peninsula
Mouse (*Mus musculus*)	10	liver	inbred strains
Pig (*Sus scrofa*)	12	blood, liver	supermarket (6), autochthonous Portuguese breeds (6)
Rabbit (*Oryctolagus cuniculus*)	10	blood, liver	domestic/supermarket (6), wild/Northern Portugal (4)
Sheep (*Ovis aries*)	10	blood, muscle	Mondegueira (2), Merino Preto (2), Saloia (2), Campaniça (2), Churra Algarvia (1), Churra da Terra Quente (1)

Supplementary Table S5. List of non-eutherian species tested with the SPInDel profiling kit.

Species	Class
Domestic chicken (*Gallus gallus domesticus*)	Aves
Domestic turkey (*Meleagris gallopavo*)	Aves
Squid (*Loligo sp*.)	Cephalopoda
Sardine (*Sardina pilchardus*)	Actinopterygii
Pouting (*Trisopterus luscus*)	Actinopterygii
Atlantic horse mackerel (*Trachurus trachurus*)	Actinopterygii
European hake (*Merluccius merluccius*)	Actinopterygii
Shrimp (Order Decapoda)	Malacostraca

Supplementary Table S6. General description of the seven SPInDel hypervariable regions used in the eutherian profiling kit.

Localization	Hypervariable region	Primer name	Primer sequence (5’-3’)	Dye label	Nucleotide position of amplicon[a]
12s rRNA	AC	A_F1	CCCCACGGGAAACAGCAG		803-1279
		A_F2	CCCCACGGGACTCAGCAG		
		A_F3	CCCCAAGGGATACAGCAG		
		A_F4	CCCCACGGGAGACAGCAG		
		C_R1	GGTTTGCTGAAGATGGCGG	TET	
	BC	B_F1	ACAATAGCTAAGACCCAAACTG		1051-1279
		B_F2	ACGATAGCTAAGGCCCAAACTG		
		C_R1	GGTTTGCTGAAGATGGCGG	TET	
	DE	D_F1	GGCAAGAAATGGGCTACATTTTC		1350-1494
		D_F2	GGGAAGAAATGGGCTACATTCTC		
		E_R1	GGTGACGGGCGGTGTGT	HEX	
12s rRNA-tRNA val-16s rRNA	FG	F_F1	GGTAAGTGTACTGGAAAGTG		1569-1921
		F_F2	GGTAAGCATACCGGAAGGTG		
		G_R1	TAGCTCGTCTGGTTTCGGG	6-FAM	
16s rRNA	HI	H_F1	GGCCTAAAAGCAGCCACCAAT		2173-2507
		H_F2	GGCCTAAAAGCAGCCATCAAT		
		I_R1	TTT TTG GTA AAC AGG CGG GG	TET	
	JK	J_F1	GACGAGAAGACCCTATGGAG	6-FAM	2716-2827
		K_R1	TCCGAGGTCGCCCCAACC		
16s rRNA-tRNA leu	LM	L_F1	GGGTTTACGACCTCGATGTTG	6-FAM	2975-3256
		M_R1	GCGATTACCGGGCTCTGC		
		M_R2	GCAGTTACCGGGCCCTG		
		M_R2	GCAATTTCCTGGCTCTGC		

[a] Numbered according to the human mtDNA reference sequence (Anderson et al. 1981. *Nature* 290:457)

Supplementary Table S7. Processed food products analyzed with the SPInDel profiling kit.

Food product	Food processing	Labeled species	Identified species
Meatball	Mechanical	Cow	Cow
Pig ham	Heat-cured	Pig	Pig
Hamburger	Mechanical	Cow	Cow
Turkey sausage	Heat-cured	Turkey and pig	Turkey[a] and pig[b]
Pig sausage	Heat-cured	Pig	Pig
Pig sausage	Slowly dried over smoke	Pig	Pig
Pig sausage	Heat-cured	Pig and chicken	Pig[c] and chicken[a]
Meat mixture sausage	Heat-cured	Pig and cow	Pig and cow
Meat mixture sausage	Heat-cured	Turkey, chicken and pig	Turkey[a], chicken[a] and pig[c]

[a] Identified by comparison with reference profiles. [b] Partial *Sus scrofa* profile with amplified products for loci JK, DE, BC, LM and HI. [c] Partial *Sus scrofa* profile with amplified products for loci JK, DE, BC, LM, HI and FG

Supplementary Table S8. List of species with equal standard SPInDel profiles.

Taxonomic group	Species with shared SPInDel profiles
Eukaryotes	
Viridiplantae (Vi)	> *Zea mays subsp. mays* (NC 007982.1) / *Zea mays subsp. parviglumis* (NC 008332.1) > *Oryza sativa Indica Group* (NC 007886.1) / *Oryza sativa Japonica Group* (NC 011033.1)
Fungi (Fu)	> *Aspergillus tubingensis* (NC 007597.1) / *Aspergillus niger* (NC 007445.1)
Cnidaria (Cn)	> *Pocillopora eydouxi* (NC 009798.1) / *Pocillopora damicornis* (NC 009797.1) > *Montastraea franksi* (NC 007225.1) / *Montastraea annularis* (NC 007224.1) / *Montastraea faveolata* (NC 007226.1)
Echinodermata (Ec)	> *Strongylocentrotus droebachiensis* (NC 009940.1) / *Strongylocentrotus pallidus* (NC 009941.1)
Nematoda (Ne)	> *Caenorhabditis elegans* (NC 001328.1) / *Caenorhabditis briggsae* (NC 009885.1)
Arthropoda (Ar)	> *Bactrocera carambolae* (NC 009772.1) / *Bactrocera papayae* (NC 009770.1) > *Bactrocera dorsalis* (NC 008748.1) / *Bactrocera philippinensis* (NC 009771.1) > *Drosophila melanogaster* (NC 001709.1) / *Drosophila simulans* (NC 005781.1) > *Eriocheir hepuensis* (NC 011598.1) / *Eriocheir sinensis* (NC 006992.1)
Mollusca (Mo)	> *Mytilus edulis* (NC 006161.1) / *Mytilus galloprovincialis* (NC 006886.1)
Actinopterygii (Ac)	> *Exocoetus volitans* (NC 003184.1) / *Cypselurus hiraii* (NC 007403.1) > *Beryx decadactylus* (NC 004393.1) / *Beryx_splendens* (NC 003188.1) > *Oncorhynchus keta* (NC 009261.1) / *Oncorhynchus masou ishikawae* (NC 008746.1) / *Oncorhynchus masou formosanus* (NC 008745.1) / *Oncorhynchus masou masou* (NC 008747.1) / *Oncorhynchus masou 'Biwa'* (NC 009262.1) > *Anguilla bicolor bicolor* (NC 006534.1) / *Anguilla bicolor pacifica* (NC 006535.1) > *Anguilla nebulosa nebulosa* (NC 006544.1) / *Anguilla bengalensis labiata* (NC 006543.1) > *Salvelinus alpinus* (NC 000861.1) / *Salvelinus fontinalis* (NC 000860.1) > *Canthigaster coronata* (NC 010978.1) / *Canthigaster rivulata* (NC 010979.1)

	> *Carassius auratus* x *Cyprinus carpio* x *Carassius cuvieri* (NC 006387.1) / *Carassius carassius* (NC 006291.1) / *Carassius auratus* (NC 002079.1) / *Carassius auratus* x *Cyprinus carpio* (NC 006136.1) / *Carassius auratus auratus* (NC 006580.1) > *Pseudaspius leptocephalus* (NC 008681.1) / *Tribolodon nakamurai* (NC 008651.1) > *Eutaeniophorus festivus* (NC 012043.1) / *Ataxolepis apus* (NC 012045.1) > *Takifugu oblongus* (NC 011634.1) / *Takifugu ocellatus* (NC 011635.1) > *Takifugu obscures* (NC 011626.1) / *Takifugu pardalis* (NC 011627.1) / *Takifugu chinensis* (NC 011633.1) / *Takifugu poecilonotus* (NC 011621.1) / *Takifugu exascurus* (NC 011622.1) / *Takifugu stictonotus* (NC 011629.1) / Takifugu snyderi (NC 011630.1) > *Hippoglossus hippoglossus* (NC 009709.1) / *Hippoglossus s*tenolepis (NC 009710.1) > *Trachurus trachurus* (NC 006818.1) / *Trachurus japonicus* (NC 002813.1) > *Cetostoma regani* (NC 004389.1) / *Eutaeniophorus sp*. 033 Miya (NC 004390.1) > *Scobinichthys granulates* (NC 011926.1) / *Nelusetta ayraudi* (NC 011921.1) > *Acanthaluteres brownii* (NC 011947.1) / *Meuschenia hippocrepis* (NC 011956.1) > *Xenobalistes tumidipectoris* (NC 011321.1) / *Xanthichthys auromarginatus* (NC 011945.1)
Amphibia (Am)	> *Ambystoma andersoni* (NC 006888.1) / *Ambystoma tigrinum tigrinum* (NC 006887.1) / *Ambystoma mexicanum* (NC 005797.1)
Aves (Av)	> *Gallus gallus* (NC 001323.1) / *Gallus gallus spadiceus* (NC 007235.1) / *Gallus gallus gallus* (NC 007236.1) / *Gallus gallus bankiva* (NC 007237.1)
Eutheria (Eu)	> *Bison bison* (NC 012346.1) / *Bos indicus* (NC 005971.1) > *Canis lupus lupus* (NC 009686.1) / *Canis lupus familiaris* (NC 002008.4) > *Loxodonta africana* (NC 000934.1) / *Mammuthus primigenius* (NC 007596.2) > *Mus musculus castaneus* (NC 012387.1) / *Mus musculus* (NC 005089.1) / *Mus musculus domesticus* (NC 006914.1) > *Mus musculus molossinus* (NC 006915.1) / *Mus musculus musculus* (NC 010339.1) > *Lama guanicoe* (NC 011822.1) / *Lama glama* (NC 012102.1) > *Delphinus capensis* (NC 012061.1) / *Stenella attenuata* (NC 012051.1) / *Tursiops aduncus* (NC 012058.1) / *Tursiops truncatus* (NC 012059.1) / *Stenella coeruleoalba* (NC 012053.1) / *Sousa chinensis* (NC 012057.1) > *Phoca largha* (NC 008430.1) / *Phoca hispida* (NC 008433.1) / *Phoca vitulina* (NC 001325.1) > *Ursus thibetanus mupinensis* (NC 008753.1) / *Ursus americanus* (NC 003426.1) > *Ursus thibetanus thibetanus* (NC 011118.1) / *Ursus thibetanus* (NC 009971.1)

Prokaryotes	
Crenarchaeota	> *Geogemma indica* str.296 (DQ492260.1) / *Hyperthermus butylicus* str. DSM 5456 (NC 008818.1) > *Metallosphaera sedula* str. DSM 5348 (NC 009440.1) / *Metallosphaera prunae* str. Ron12/II = DSM 10039 (X90482.1) / *Ignisphaera aggregans* str. AQ1.S1 (DQ060321.1) > *Pyrobaculum calidifontis* str. JCM 11548 (NC 009073.1) / *Thermoproteus tenax* str. YS44 (AY538162.2) / *Pyrobaculum organotrophum* str. JCM 9190 (AB063647.1) / *Pyrobaculum islandicum* str. DSM 4184 (NC 008701.1) / *Pyrobaculum neutrophilum* str. NZ1 (X81886.1) > *Sulfolobus acidocaldarius* str. DSM 639 (NC 007181.1) / *Sulfolobus thuringiensis* str. Ron12/I (X90485.1) / *Sulfolobus solfataricus* (X03235.1) > *Sulfolobus tokodaii* str. 7 (NC 003106.2) / *Sulfolobus islandicus* str. M.14.25 (CP001400.1) > *Desulfurococcus kamchatkensis* str. 1221n (EU167539.2) / *Desulfurococcus mobilis* (M36474.1) / *Staphylothermus marinus* (X99560.1)
Tenericutes	> Candidatus *Mycoplasma turicensis* str. IT226 (EU839977.1) / *Mycoplasma haemofelis* str. IT14 (EU839978.1) > Uncultured Anaeroplasma sp. clone EMP E27 (EU794313.1) / Uncultured Anaeroplasma sp clone EMPL44 (EU794312.1) > Candidatus *Phytoplasma australiense* isolate NZ09156 (FJ943262.1) / *Papaya phytoplasma* TW (AJ919994.2) / Salt cedar witches'-broom phytoplasma str. SCWB1 (FJ432664.1) / *Iranian plum phytoplasma* PJ15 (FJ409624.1) / Candidatus *Phytoplasma solani* isolate 25-5-48 (EU814640.1) > Russian olive witches'-broom phytoplasma str. Uromia (EU886968.1) / *Aster yellows phytoplasma* (AY549311.2) / Mulberry dwarf phytoplasma str. NY (EU999736.1) / Paulownia witches'-broom phytoplasma str. PaWB-Beijing (FJ263621.1) / Candidatus *Phytoplasma asteris* (FJ429364.1) > Jujube witches'-broom phytoplasma str. NY1 (EU999739.1) / Candidatus *Phytoplasma ulmi* str. EYCZ1 (EU184021.1) > *Mycoplasma wenyonii* Fengdu (FJ375309.1) / *Mycoplasma suis* str. SC (FJ263943.1) > *Mycoplasma iguanae* str. 2327 (EU859973.1) / *Mycoplasma sp.* A1802T str. A1802 (FM165075.1) / *Mycoplasma arthritidis* str. 158L3-1 (CP001047.1) / *Mycoplasma canadense* str. 466 (EU925158.1) / *Mycoplasma sp*. 237IAT str. 237IA (FM165077.1) / *Mycoplasma zalophi* str. CSL 7520 (FJ226581.1) > *Mycoplasma citelli* str. RG1D-L (EU925165.1) / *Mycoplasma columborale* str. 18303 (EU859977.1) > *Ureaplasma urealyticum* serovar 10 str. ATCC 33699 (NC 011374.1) / *Ureaplasma urealyticum* serovar 12 str. ATCC 33696 (AAZT01000001.1) / Uncultured *Ureaplasma sp*. clone VE12C09 (GQ179714.1) > *Mycoplasma fermentans* str. PG-18 (FJ226561.1) / *Mycoplasma phocirhinis* str. CSL 7475-4 (EU827598.1) / *Mycoplasma canimucosale* str. HRC689 (EU797451.1)

Viruses	
Lentivirus	> *Human immunodeficiency virus 2* (NC 001722.1) / *Simian-Human immunodeficiency virus* (NC 001870.1) > *Simian immunodeficiency virus SIV-mnd 2* (NC 004455.1) / *Human immunodeficiency virus 1* (NC 001802.1) > *Visna/Maedi virus* (NC 001452.1) / *Ovine lentivirus* (NC 001511.1) / *Caprine arthritis encephalitis virus* (NC 001463.1)
Papillomaviridae	> *Bovine papillomavirus 5* (NC 004195.1) / *Bovine papillomavirus 8* (NC 009752.1) > *Mastomys coucha papillomavirus 2* (NC 008519.1) / *Old World harvest mouse papillomavirus* (NC 008582.1) > *Deer papillomavirus* (NC 001523.1) / *Common chimpanzee papillomavirus 1* (NC 001838.1) > *Human papillomavirus type 26* (NC 001583.1) / *European elk papillomavirus* (NC 001524.1) > *Human papillomavirus type 5* (NC 001531.1) / *Human papillomavirus type 99* (NC 012745.1) / *Human papillomavirus RTRX7* (NC 004761.1) > *Human papillomavirus type 9* (NC 001596.1) / *Human papillomavirus type 113* (NC 012748.1) > *Caretta caretta papillomavirus 1* (NC 011530.1) / *Chelonia mydas papillomavirus 1* (NC 011531.1) > *Human papillomavirus type 24* (NC 001683.1) / *Human papillomavirus type 98* (NC 012744.2) > *Equine papillomavirus 1* (NC 004194.1) / *Equus caballus papillomavirus 1* (NC 003748.1)
Rhabdoviridae	> *Lettuce yellow mottle virus* (NC 011532.1) / *Lettuce necrotic yellows virus* (NC 007642.1) > *Snakehead rhabdovirus* (NC 000903.1) / *Viral hemorrhagic septicemia virus* (NC 000855.1) > *Infectious hematopoietic necrosis virus* (NC 001652.1) / *Hirame rhabdovirus* (NC 005093.1) > *Vesicular stomatitis Indiana virus* (NC 001560.1) / *Tupaia virus* (NC 007020.1) > *Australian bat lyssavirus* (NC 003243.1) / *Mokola virus* (NC 006429.1) / *European bat lyssavirus 2* (NC 009528.1) / *European bat lyssavirus 1* (NC 009527.1) / *Rabies virus* (NC 001542.1) > *Maize mosaic virus* (NC 005975.1) / *Taro vein chlorosis virus* (NC 006942.1)

Supplementary Table S9. Comparison of molecular methods for species identification.

Method		Time/run	Cost/independent read or run (US$)
Synthetic chain-terminator chemistry (Sanger method)		1-3 hours[a]	0.80[b]
Next-generation sequencing platforms	Pyrosequencing	7-8 hours[a,c]	8,439[c]
	Clonal single molecule array	2-5 days[a,c,d]	8,950[c] 6,300[d]
	Sequencing-by-ligation	3-5 days[a,c]	17,447[c] 7,700[d]
Real-time PCR		1 day[e]	450[e]
SPInDel by automated fluorescent DNA fragment length analysis		1 hour	0.76[f]

[a] Hodgson, J. The Evolution of Sequencing Technology. Science posters. http://www.sciencemag.org/products/

[b] Ryan, D. *et al.* (2007) *Trends Biotechnol.* **25,** 385-389.

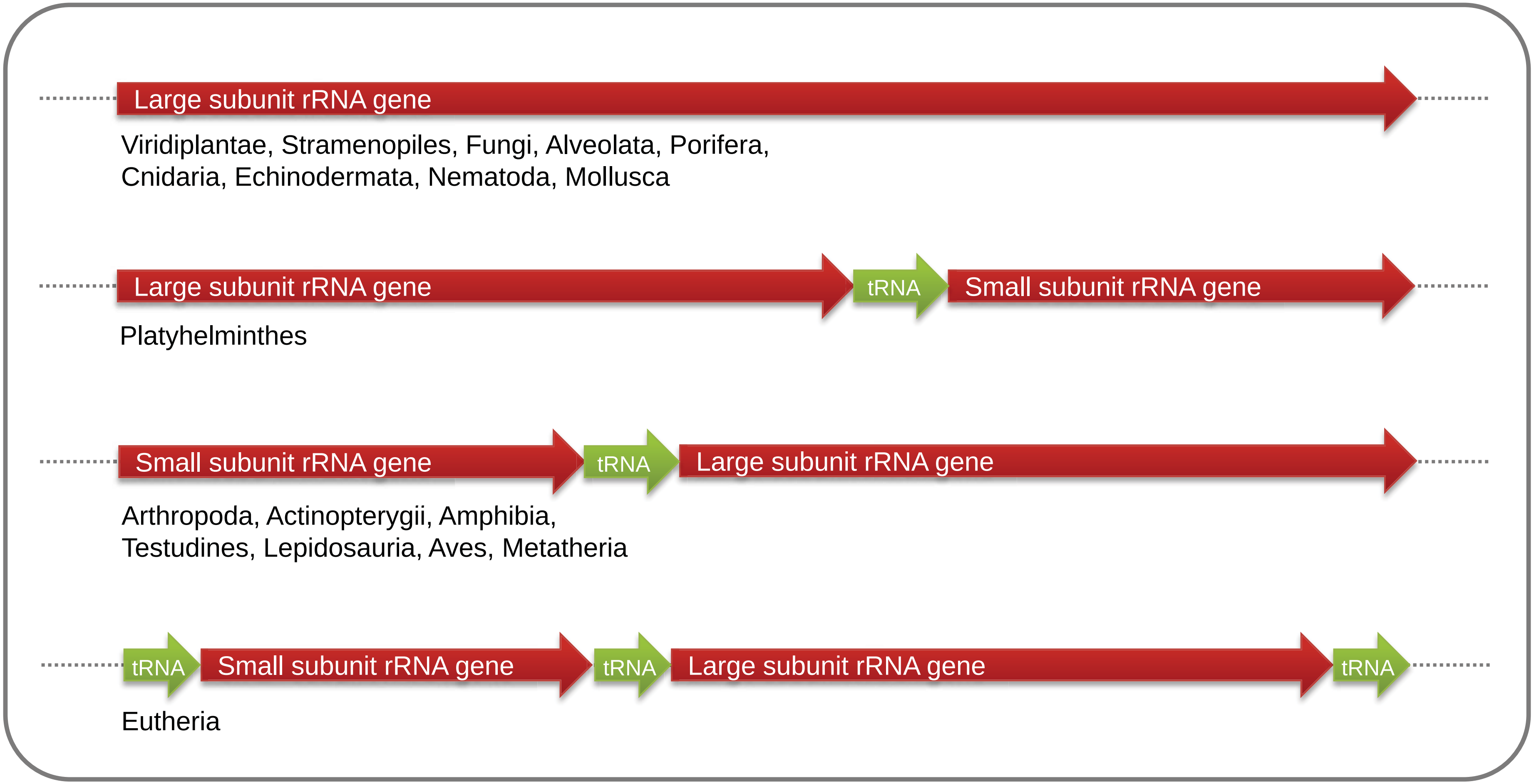
[c] Mardis, E.R. (2008) *Trends Genet.* **24,** 133-141.

[d] Hert, D.G., Fredlake, C.P. and Barron, A.E. (2008) *Electrophoresis*. **29**, 4618–4626.

[e] Ahmed, F.E (2002) *Trends Biotechnol*. **20**, 215-223.

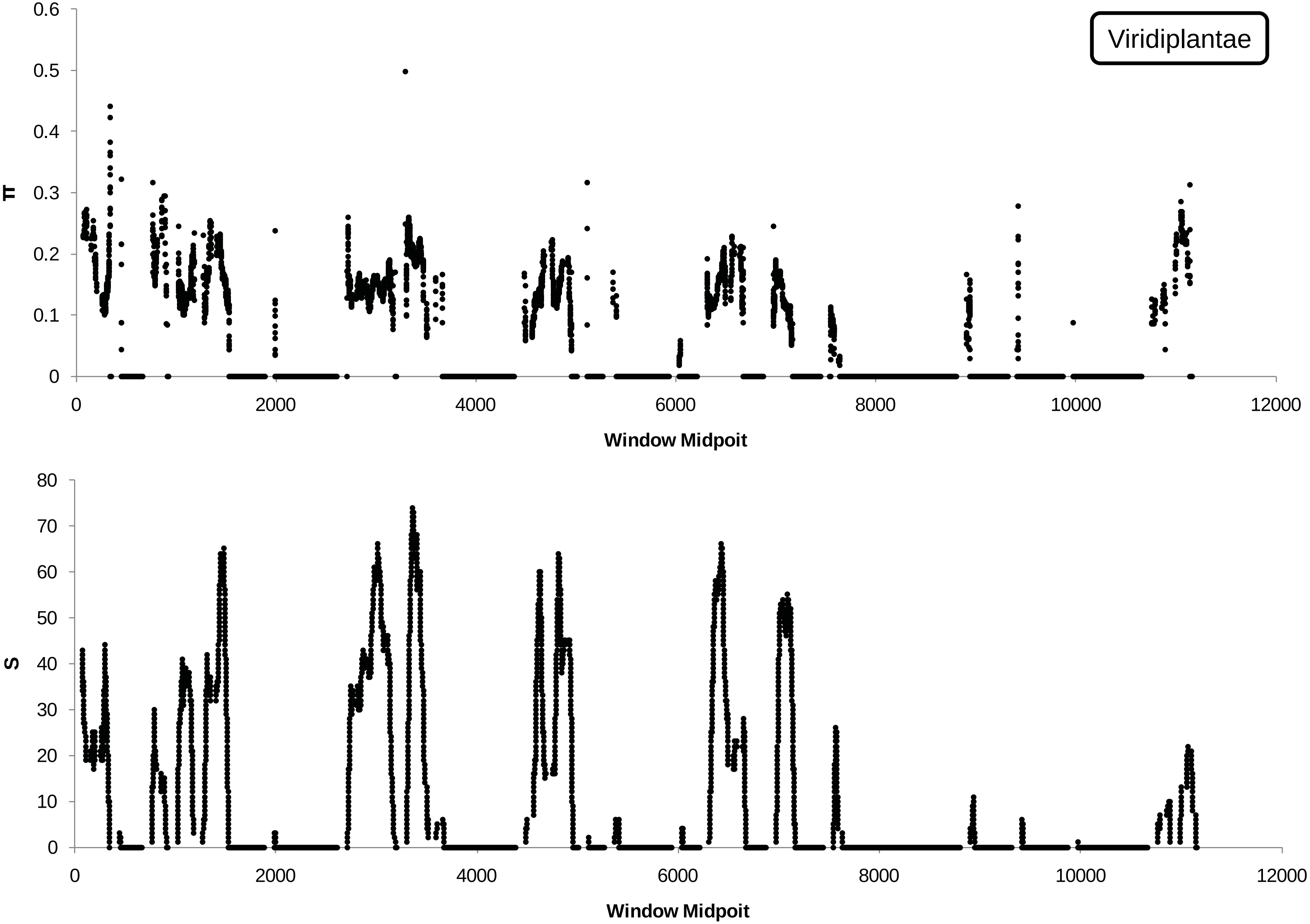
[f] Our estimate includes instrument maintenance, labor, reagents and disposables costs.

Supplementary Figure S1. Ribosomal RNA gene regions used in the development of the SPInDel method. The scheme represents the most common genomic architecture observed in 18 eukaryotic groups (introns and rare genomic rearrangements are not represented). The large subunit rRNA gene was annotated on reference sequences as 26S rRNA, rrn26 rRNA or 26Srrn (in Viridiplantae), 21S rRNA or 23S rRNA (in Fungi), rnl b1 rRNA or rnl b2 rRNA (in some Alveolata species with two rRNA gene copies) and rnl rRNA, rrnL rRNA, LSU rRNA, l-rRNA and 16S rRNA in most taxonomic groups. The small subunit rRNA gene was also annotated as s-rRNA, rrnS rRNA, rnr1 rRNA and 12S rRNA.

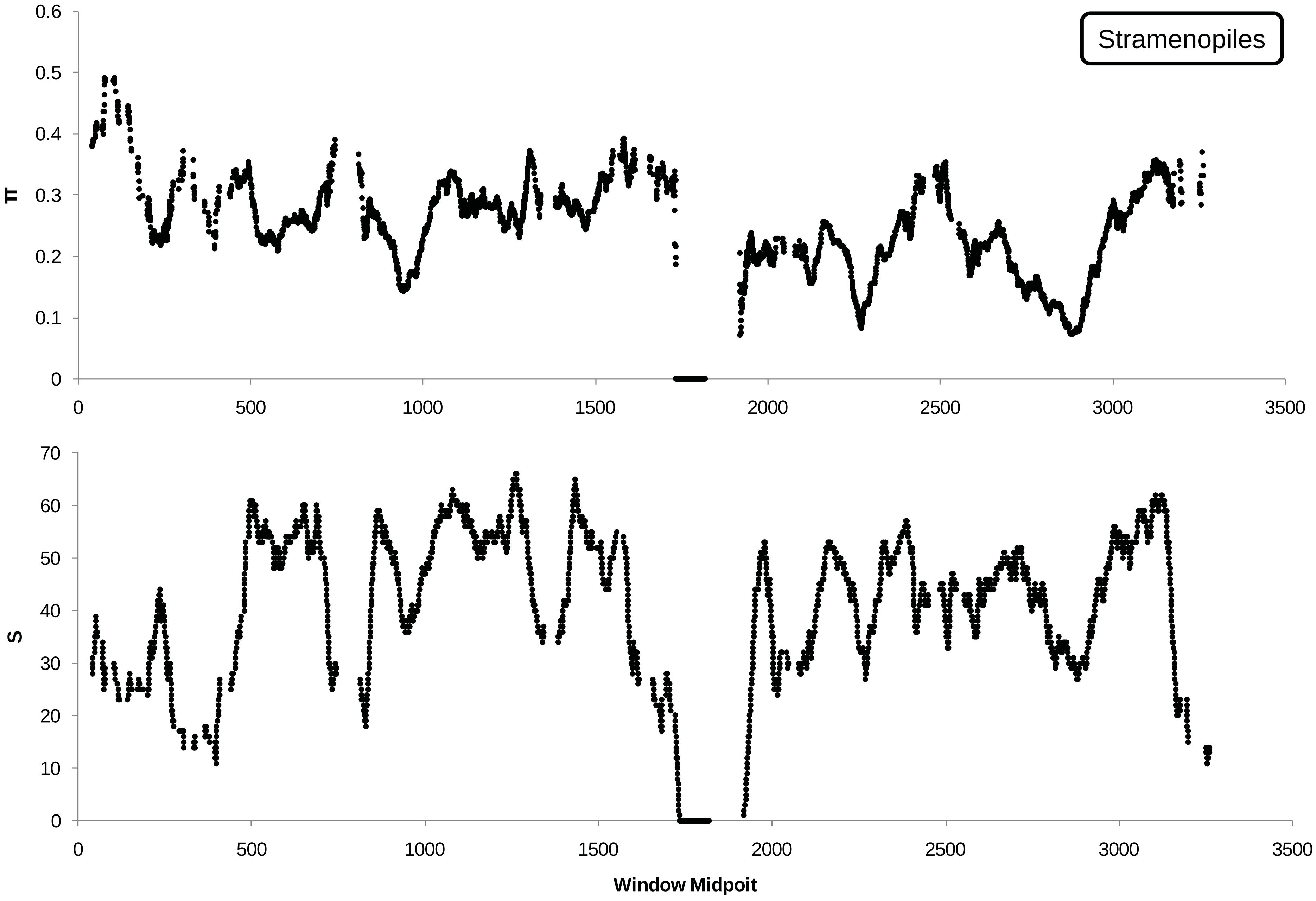


Supplementary Figure S2. Sliding window analysis of nucleotide diversity (π) and the number of segregating sites (S) across ribosomal RNA gene alignments in 18 eukaryotic groups. A window of 100 nucleotides (sites with alignment gaps were counted in the window length) was slid along the alignment in one nucleotide steps. Values for each parameter were assigned to the nucleotide at the midpoint of each window.

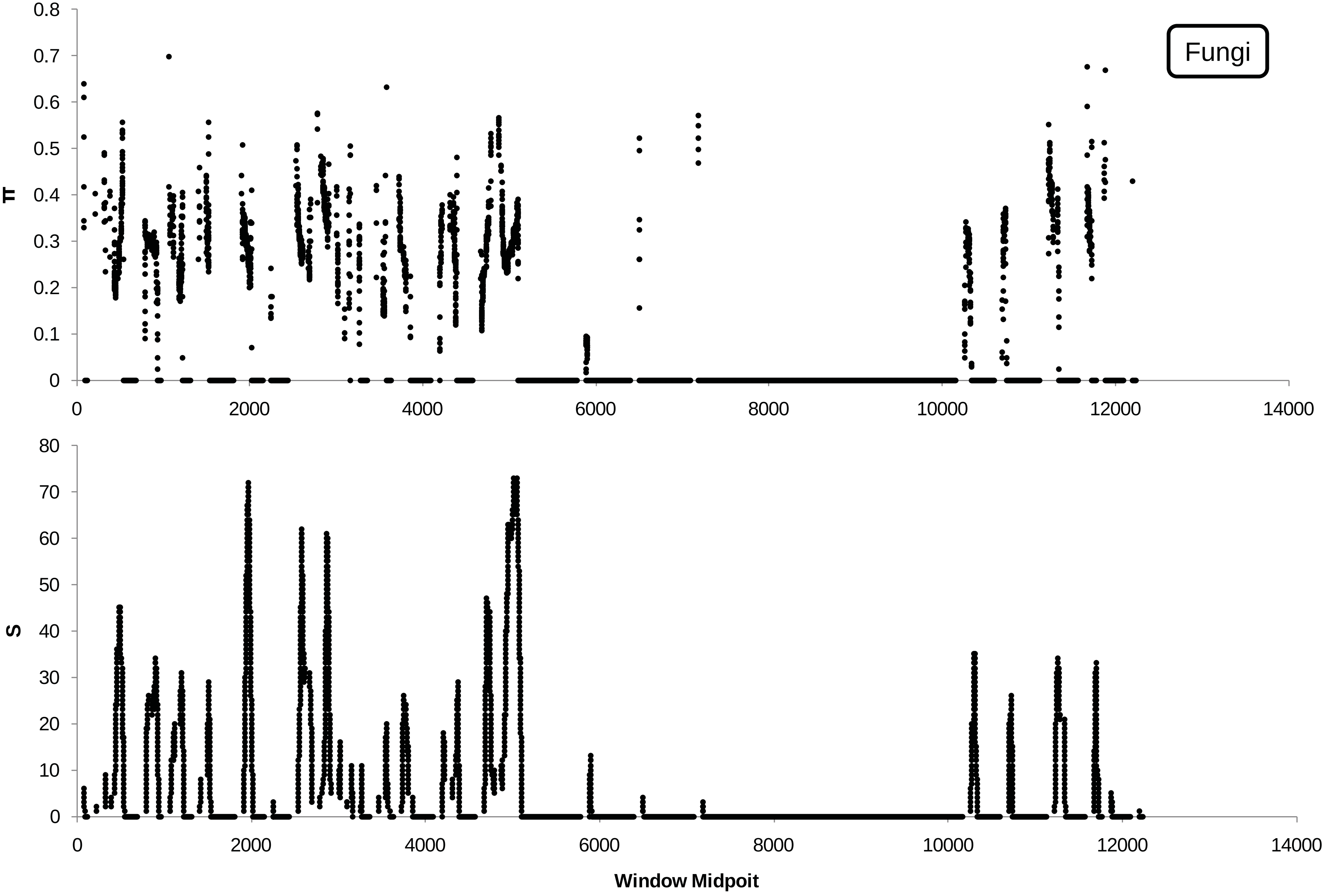
Supplementary Figure S2



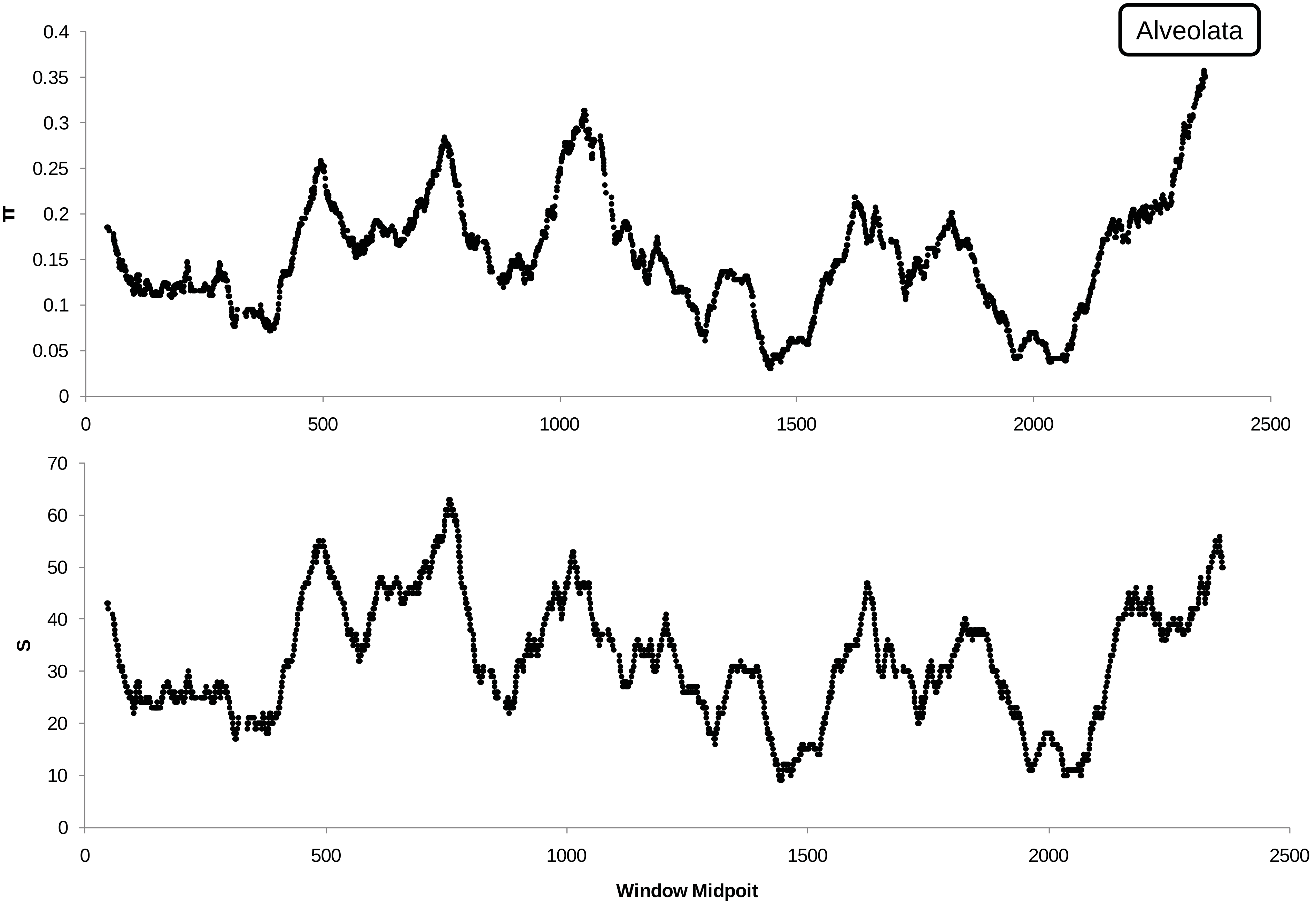
Supplementary Figure S2 (cont.)



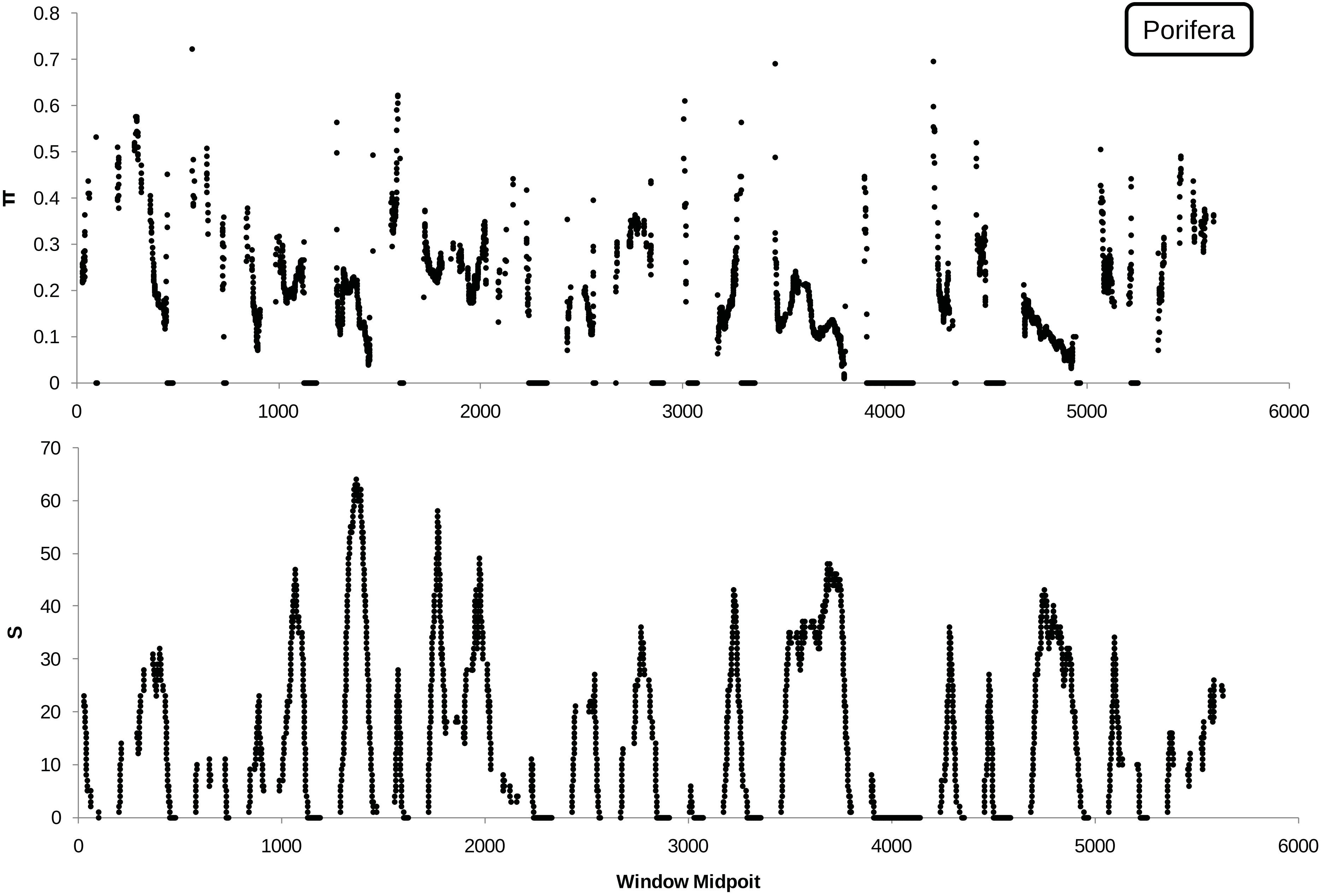
Supplementary Figure S2 (cont.)



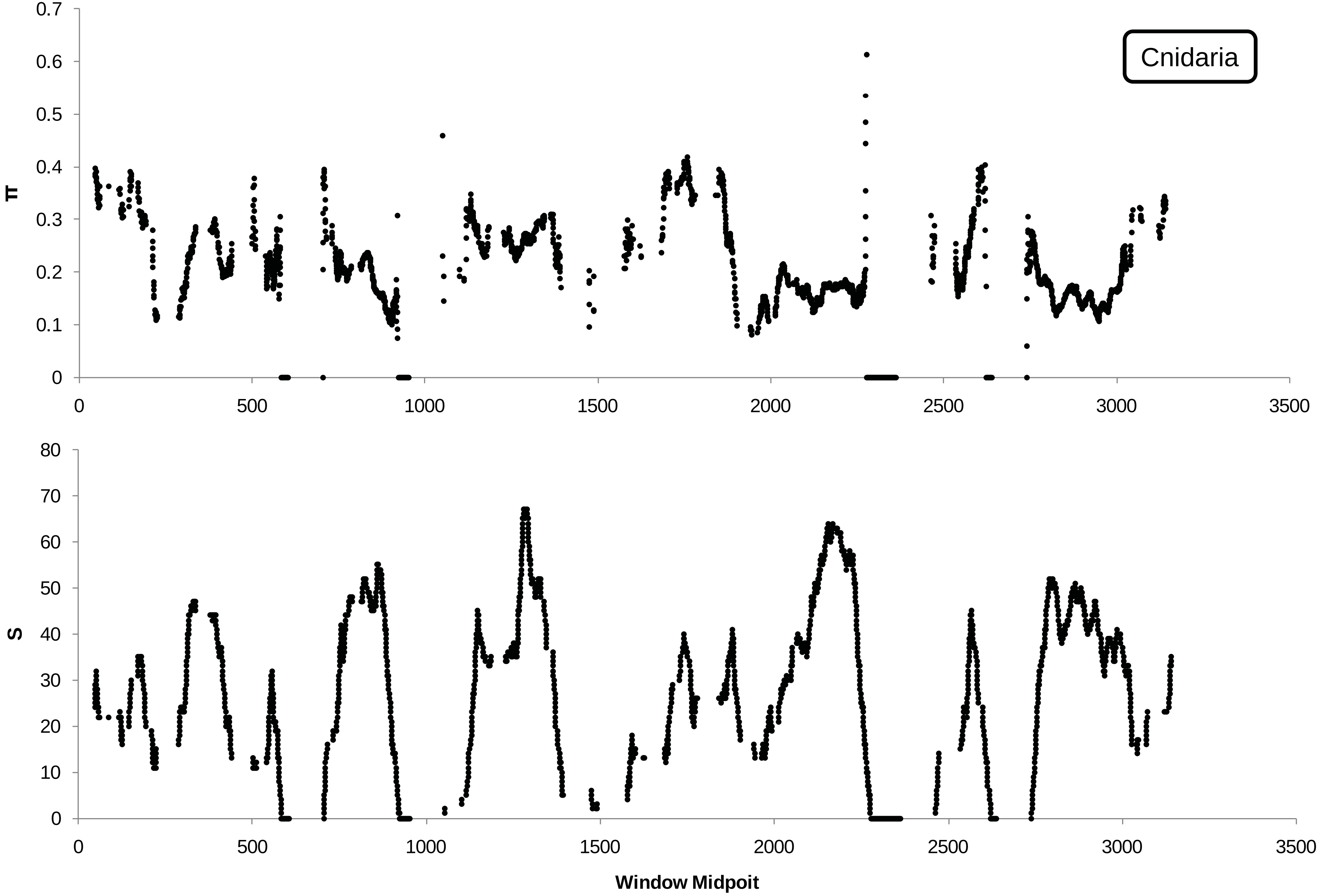
Supplementary Figure S2 (cont.)



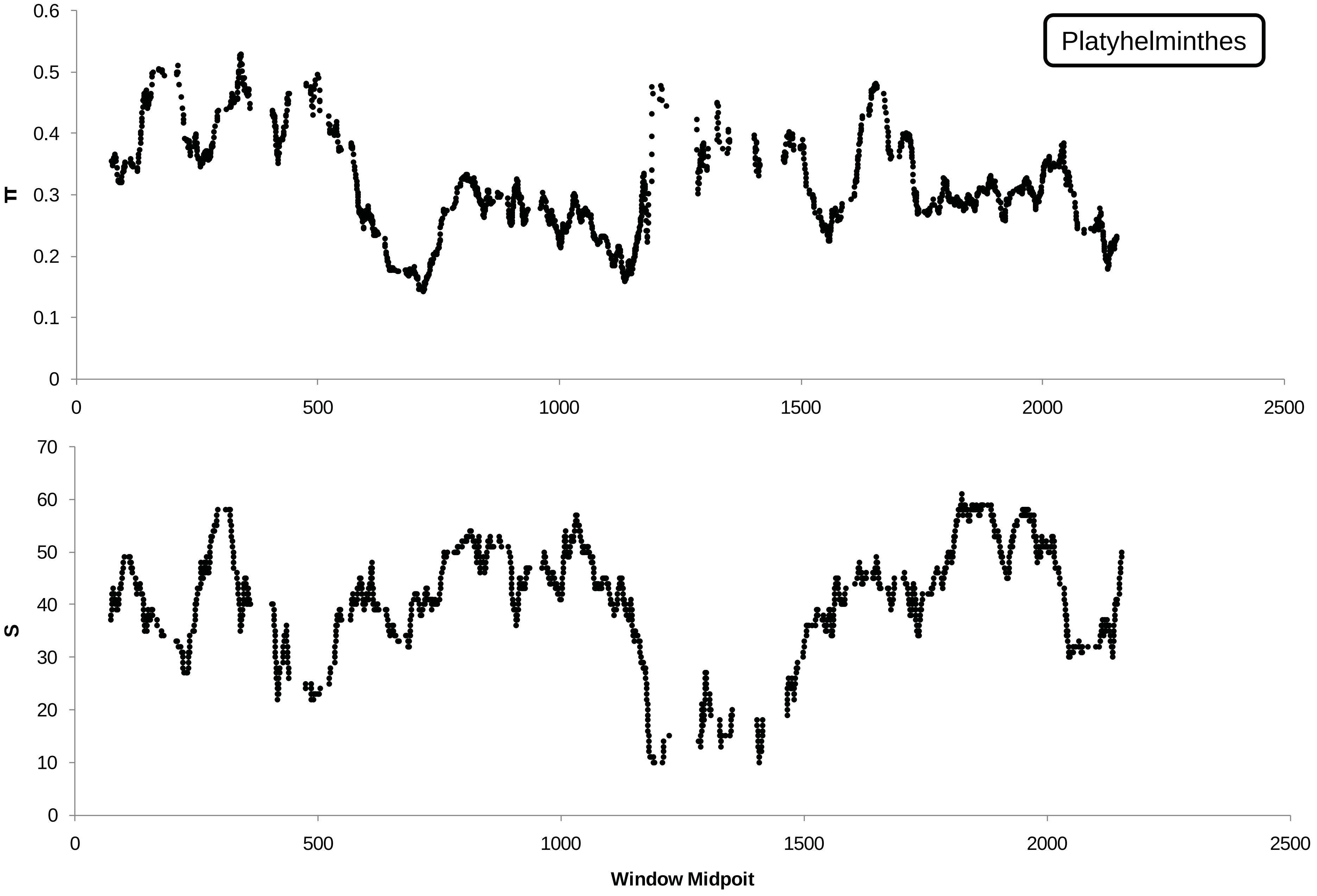
Supplementary Figure S2 (cont.)



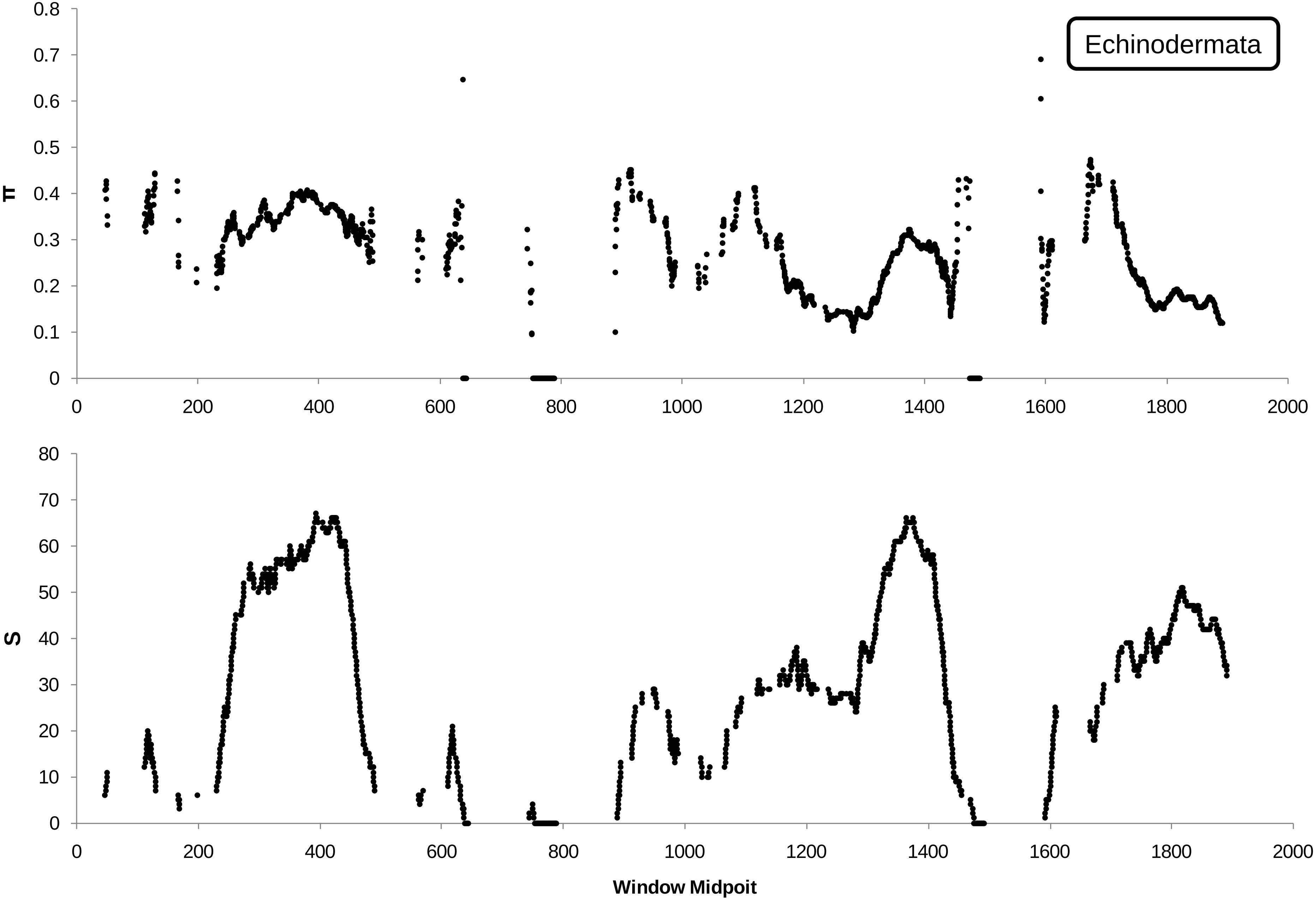
Supplementary Figure S2 (cont.)



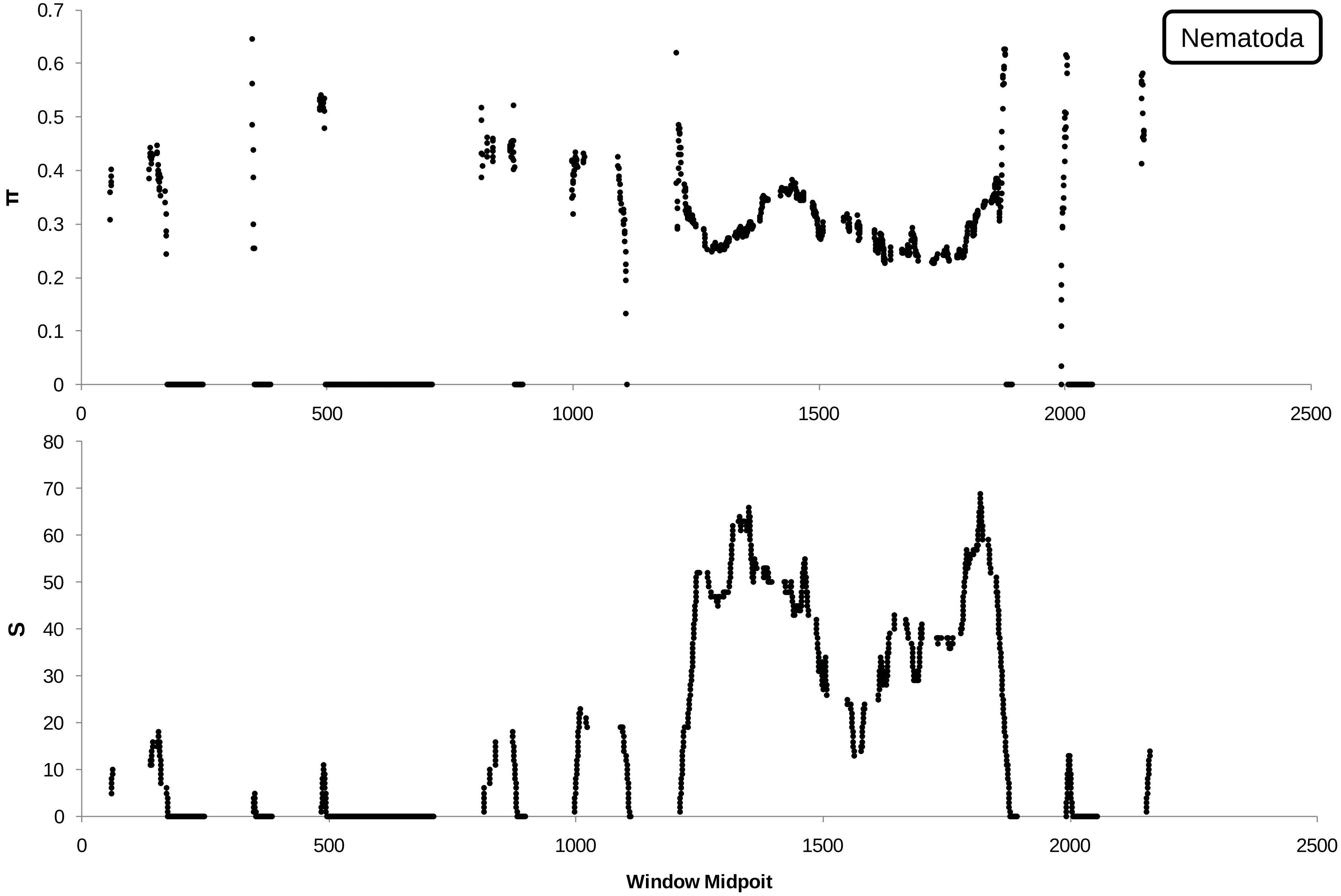
Supplementary Figure S2 (cont.)



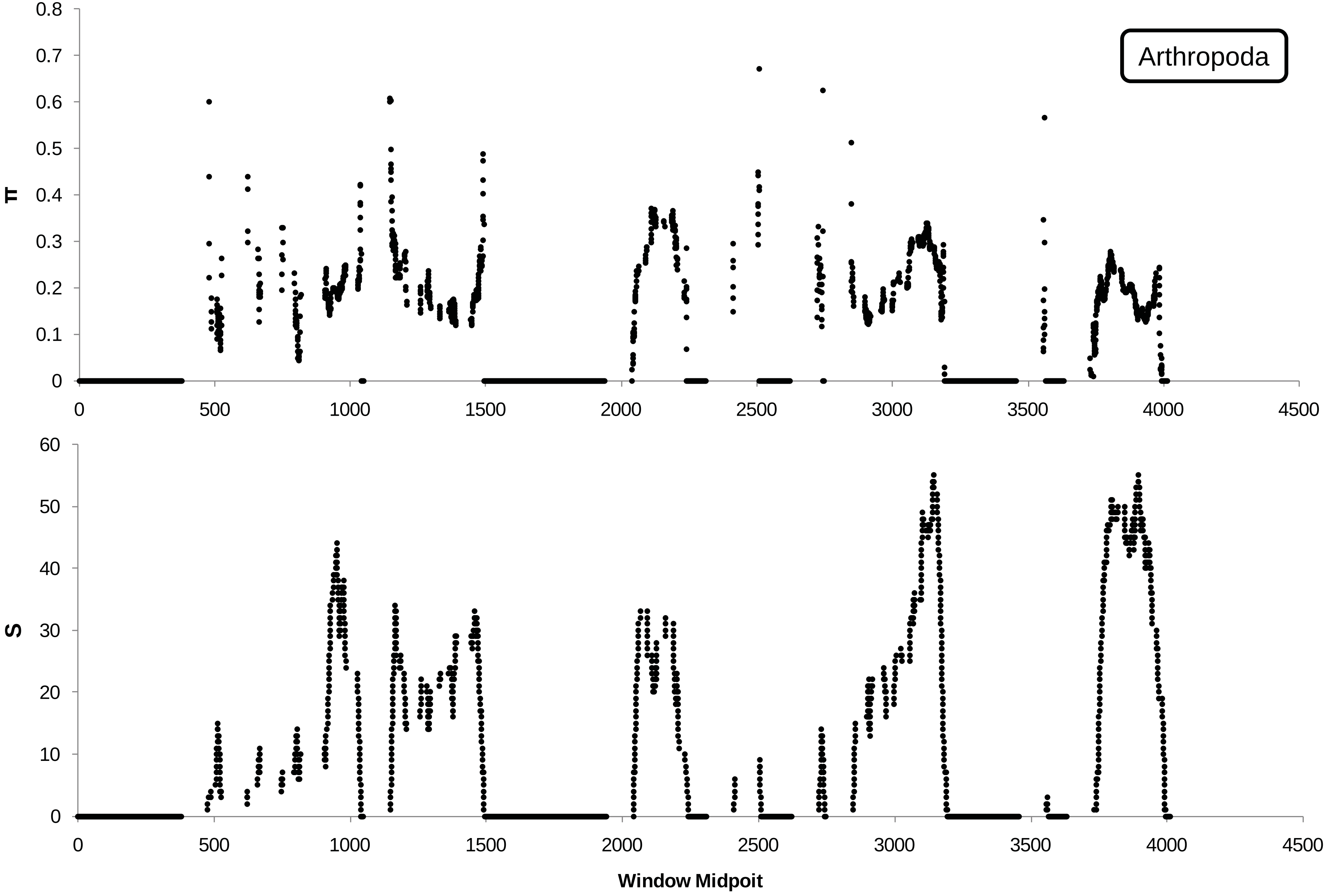
Supplementary Figure S2 (cont.)



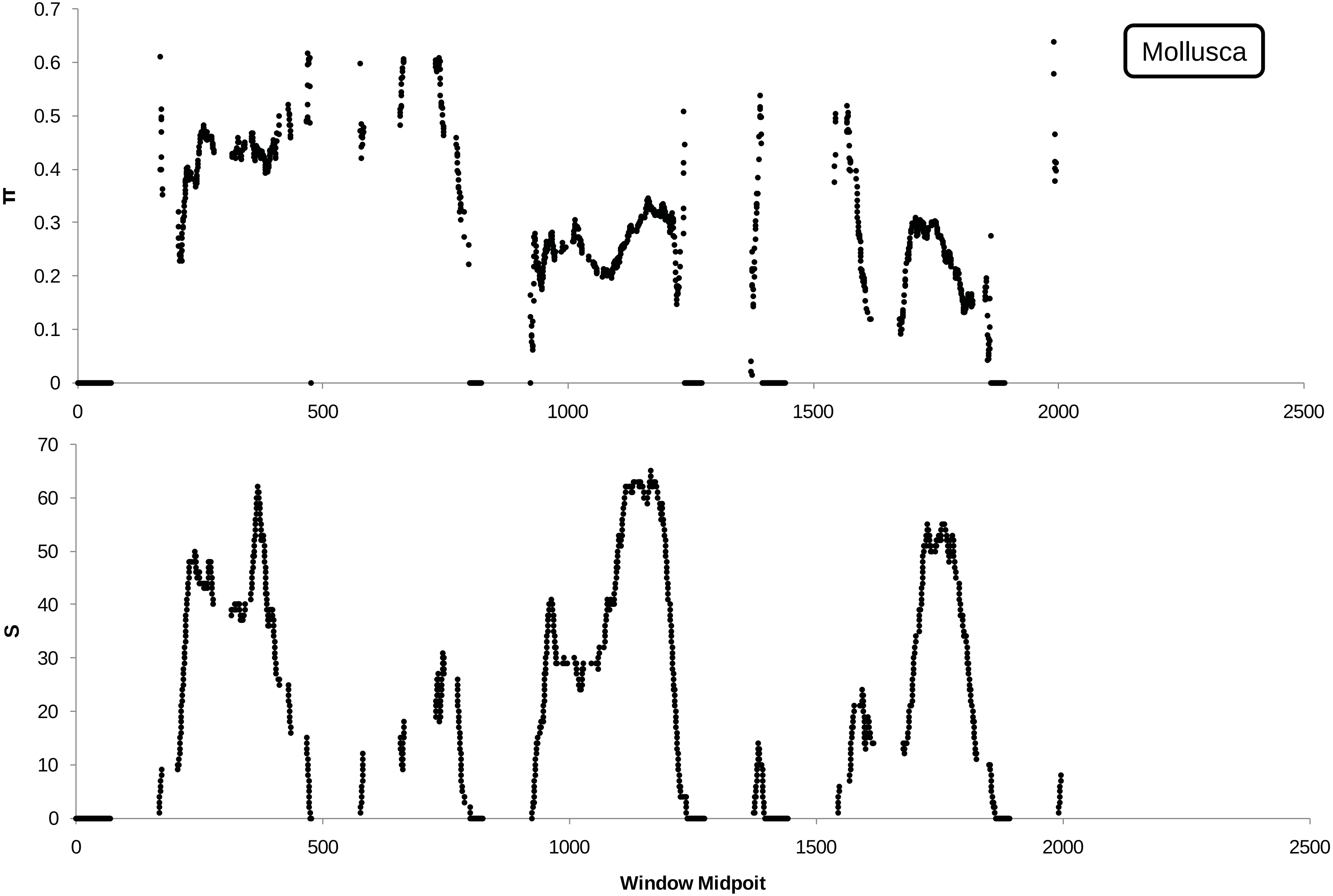
Supplementary Figure S2 (cont.)



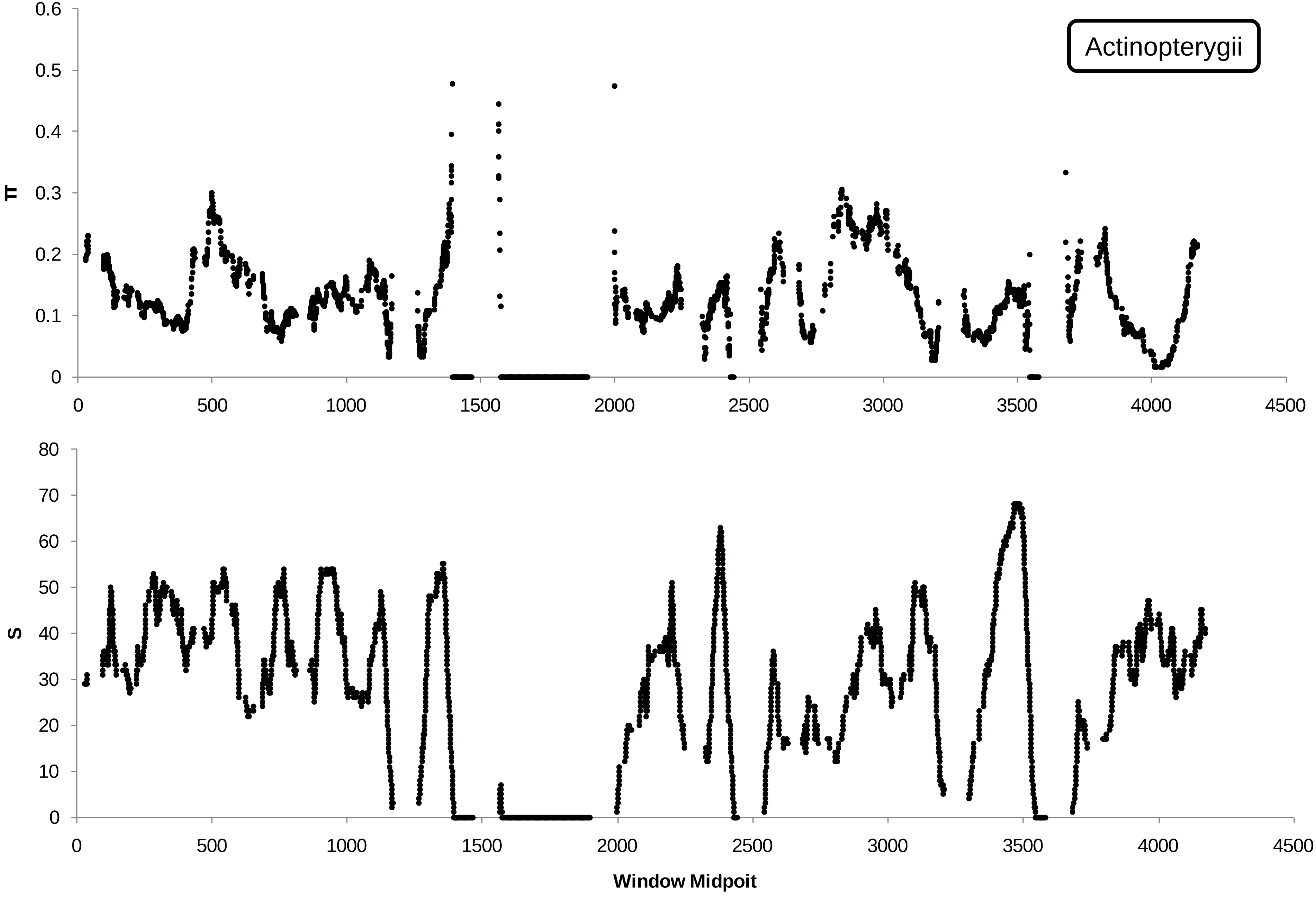
Supplementary Figure S2 (cont.)



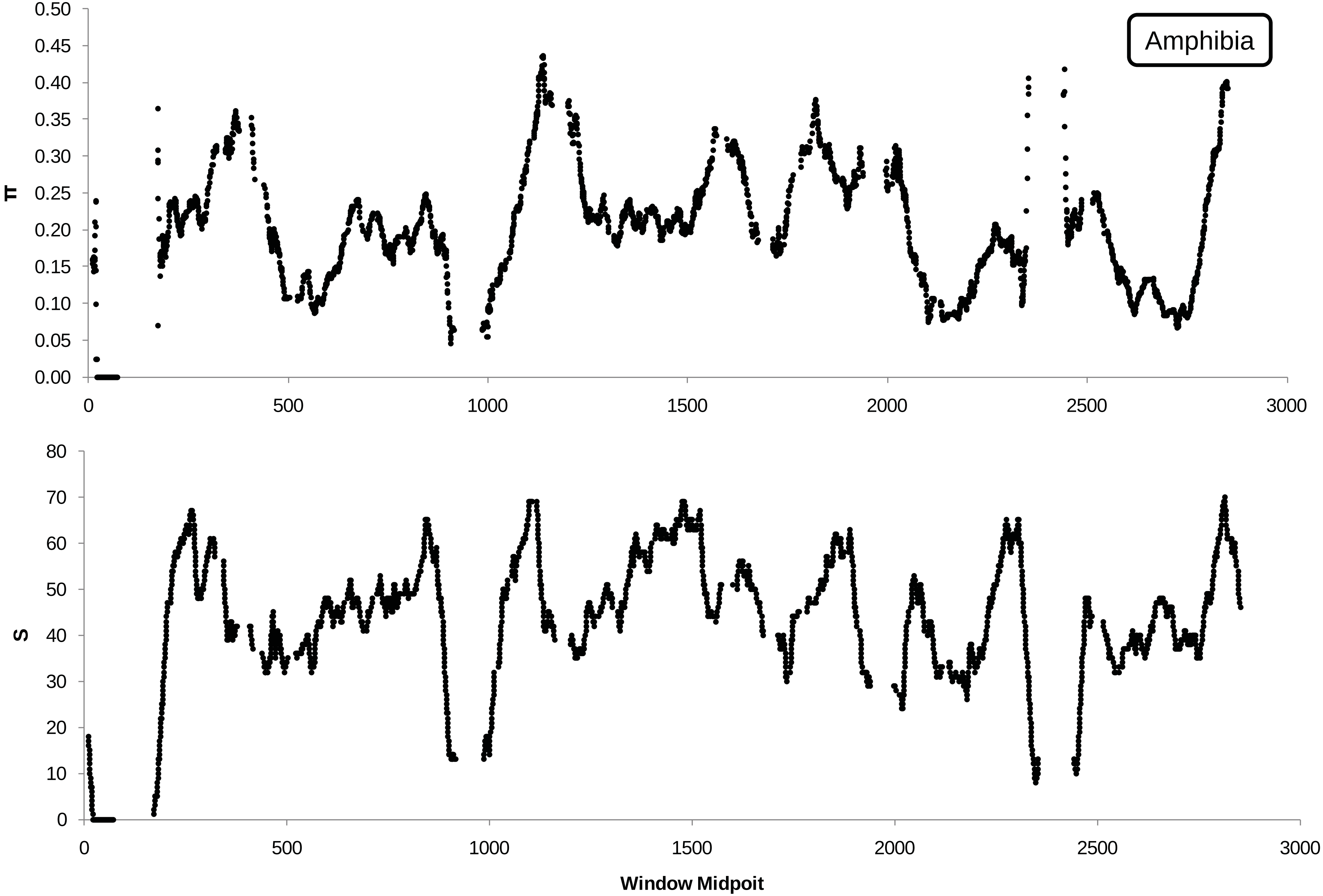
Supplementary Figure S2 (cont.)



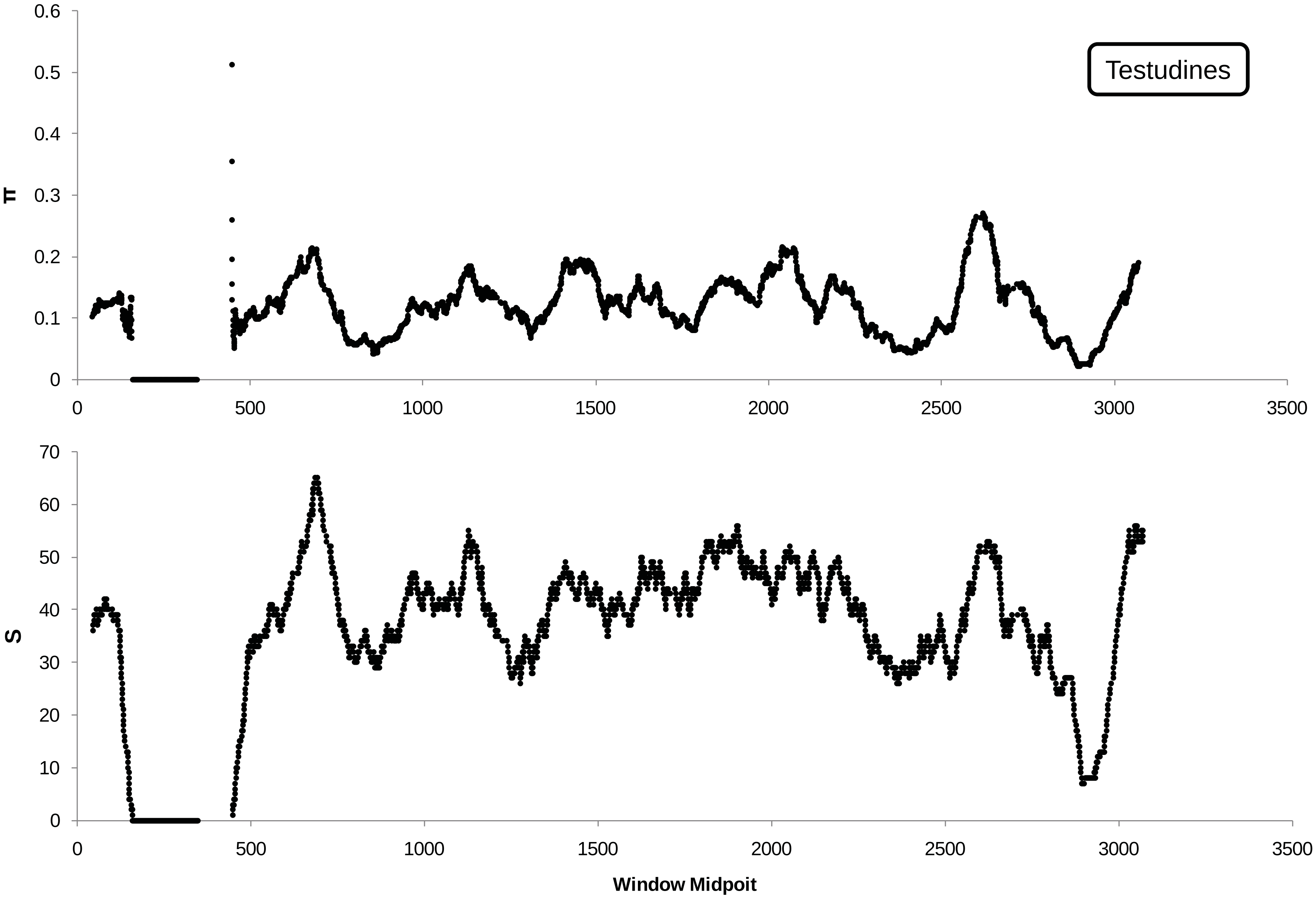
Supplementary Figure S2 (cont.)



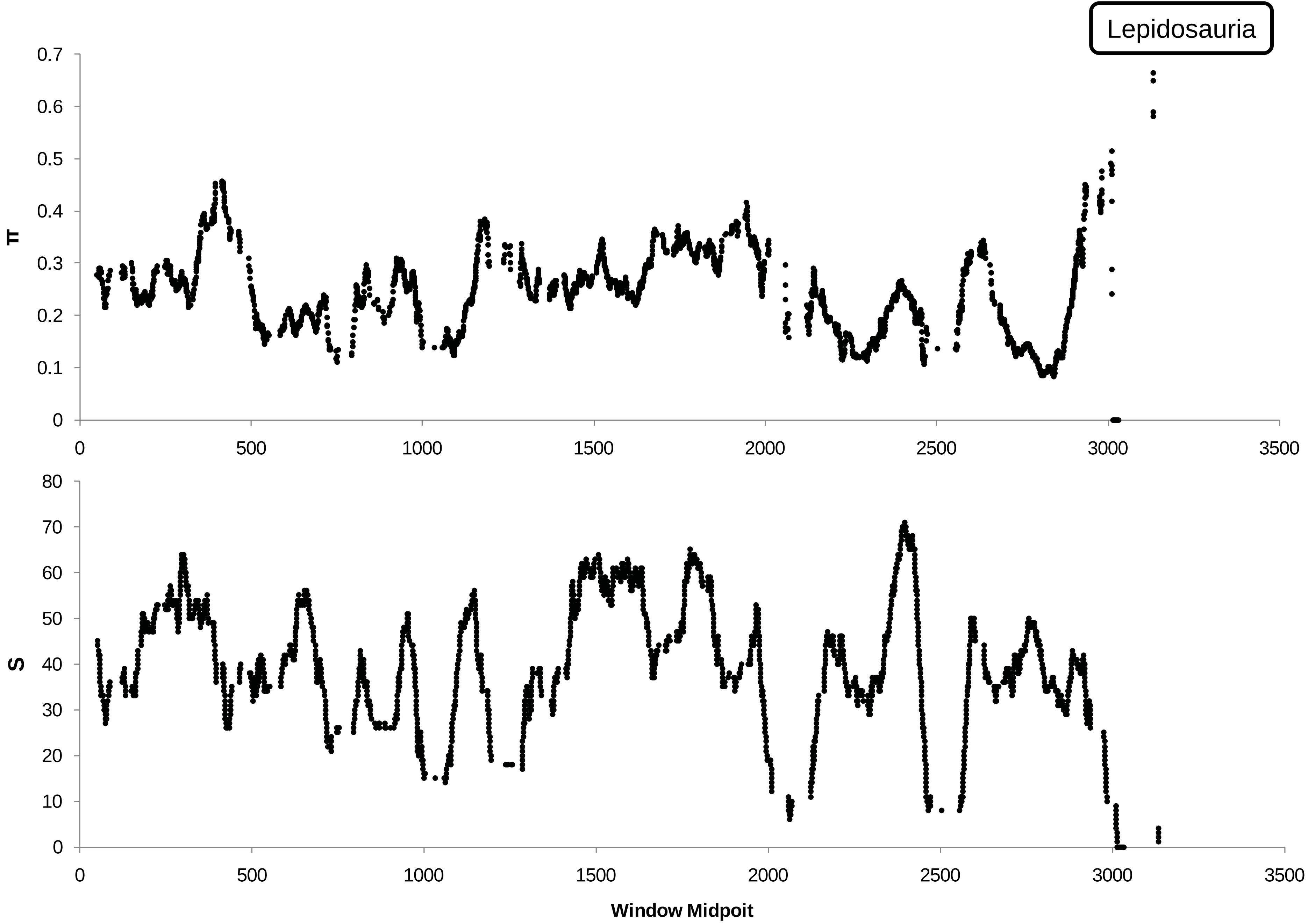
Supplementary Figure S2 (cont.)



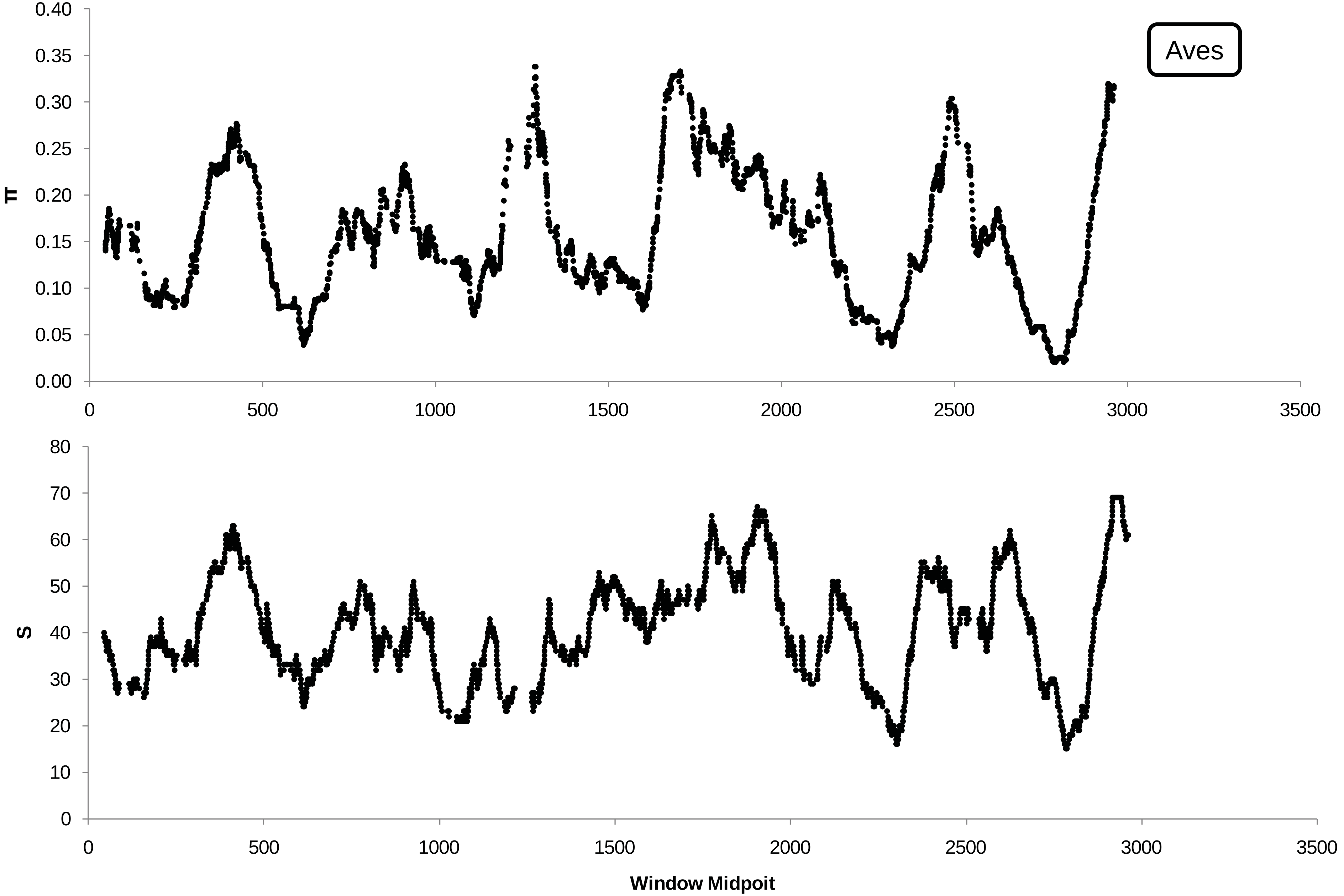
Supplementary Figure S2 (cont.)



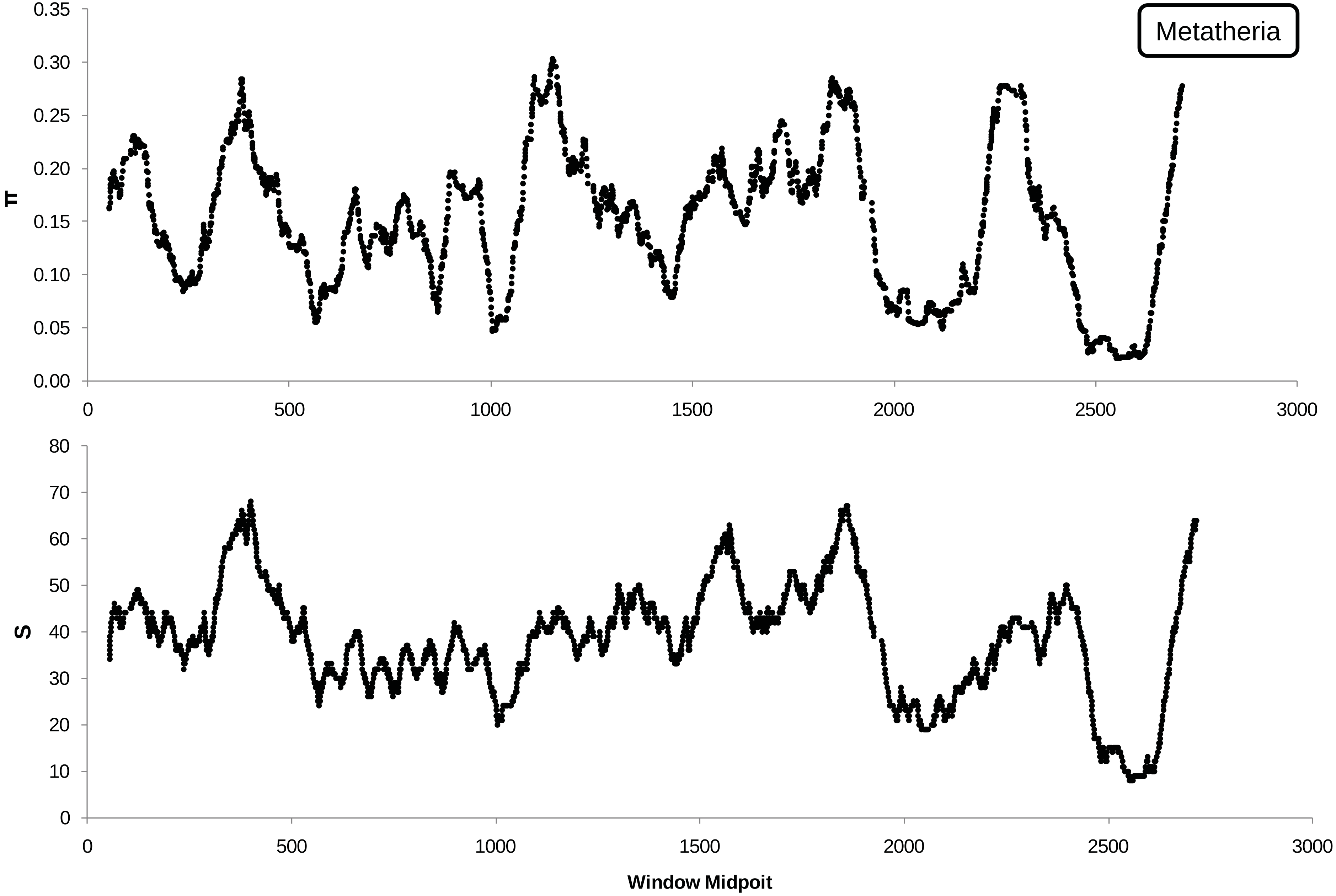
Supplementary Figure S2 (cont.)



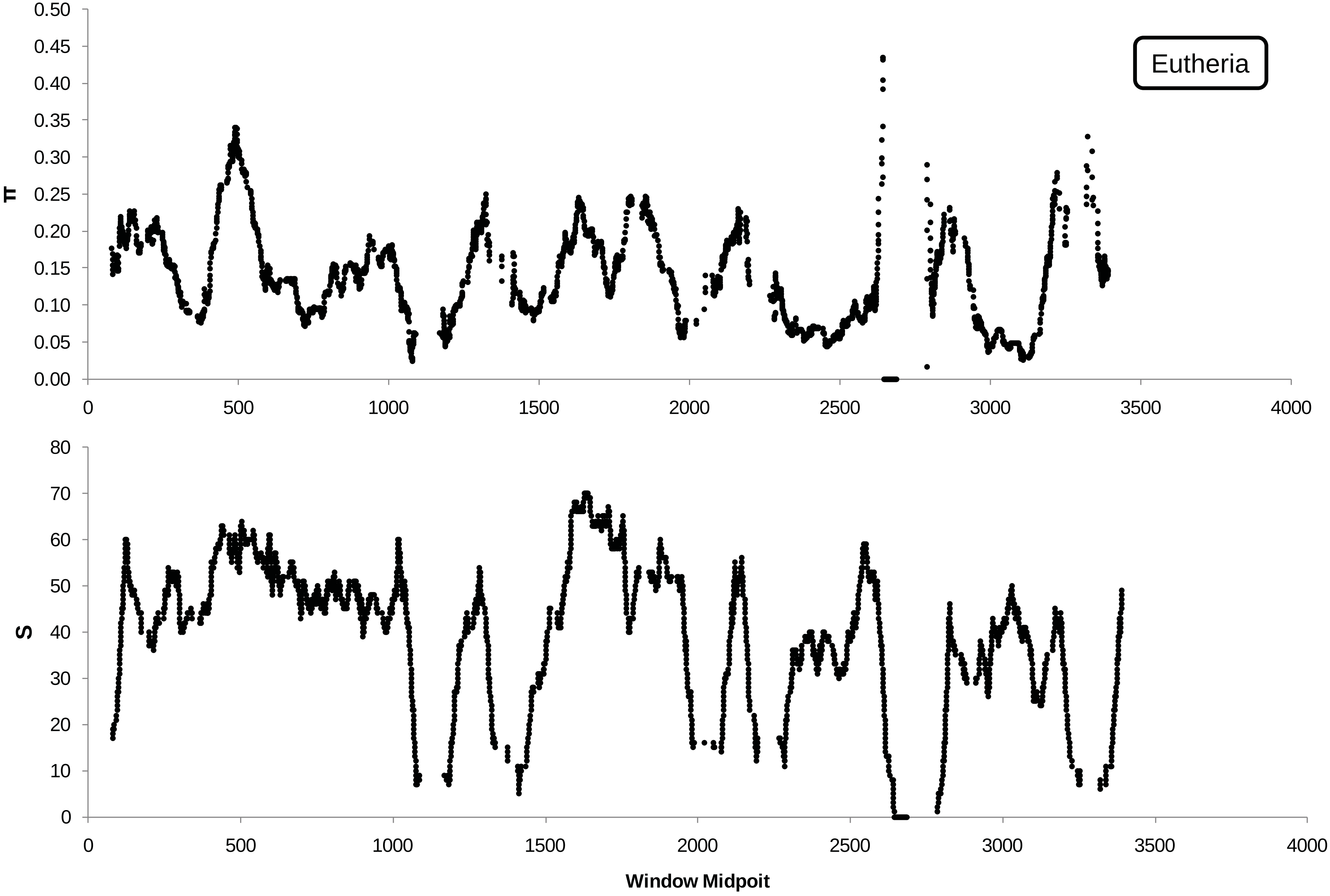
Supplementary Figure S2 (cont.)



Supplementary Figure S2 (cont.)



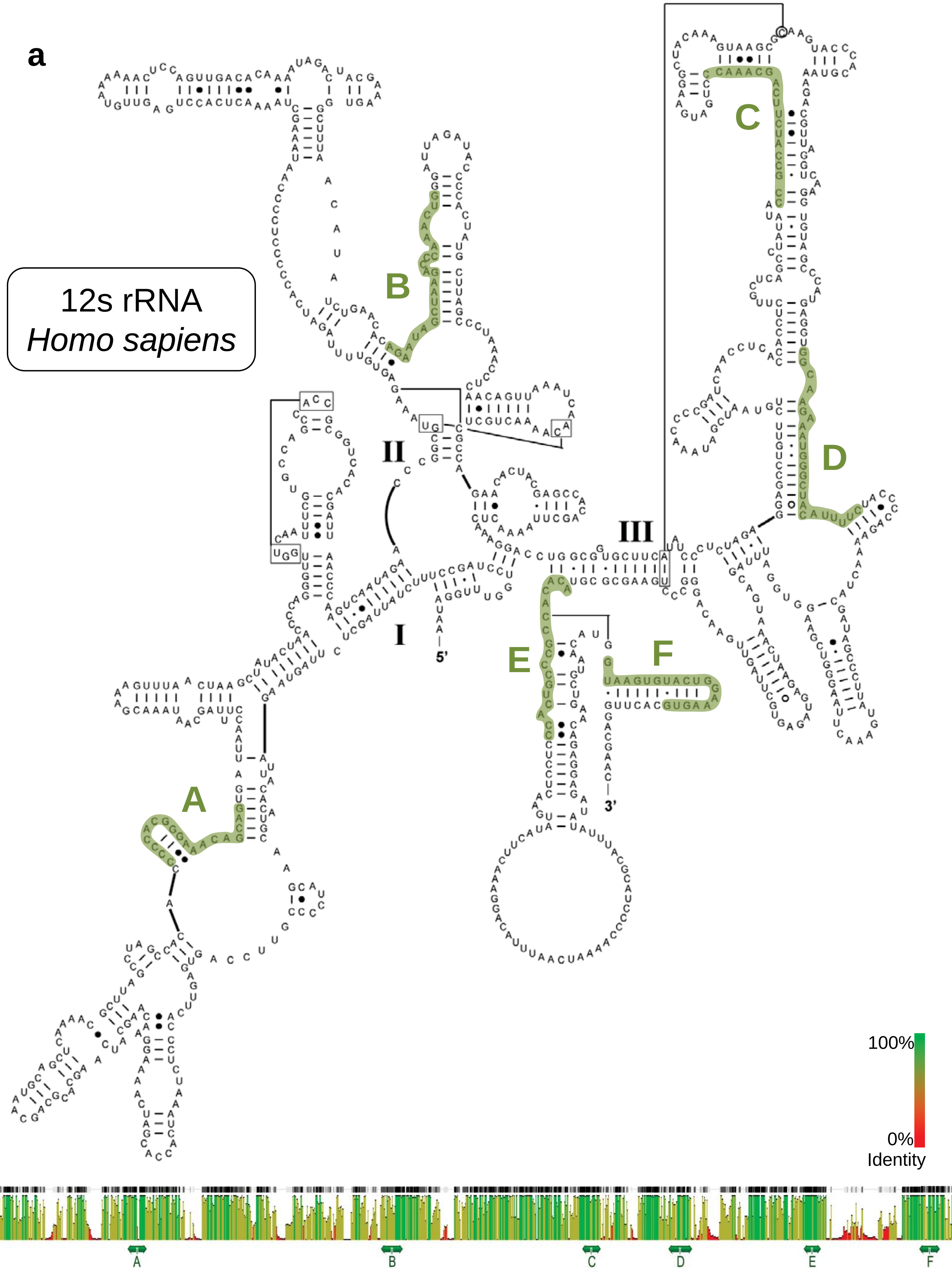
Supplementary Figure S2 (cont.)



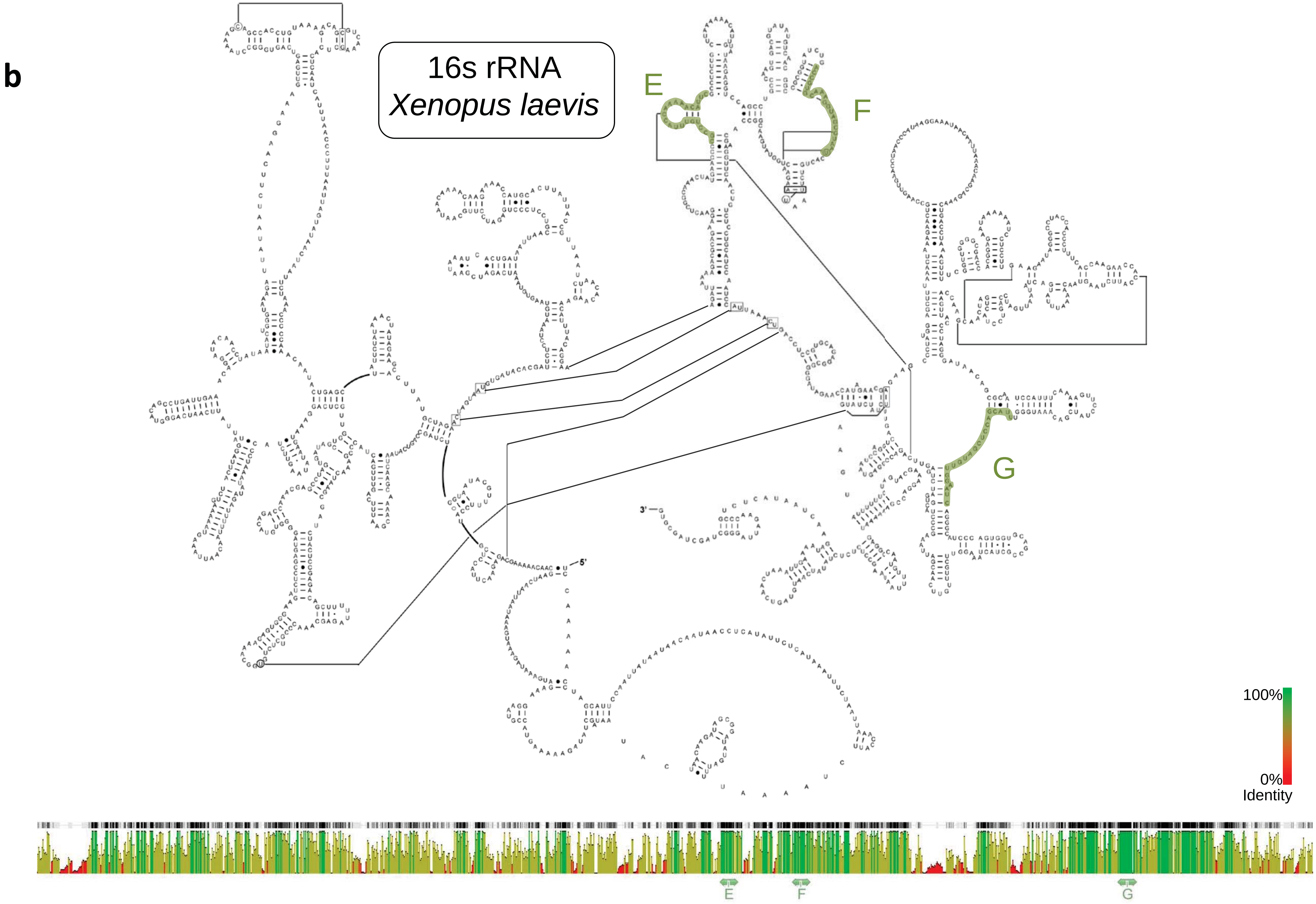
Supplementary Figure S3. SPInDel conserved regions on ribosomal RNA (rRNA) secondary structures. (**a**) *Homo sapiens* mitochondrial 12S rRNA secondary structure showing the location of eutherian conserved regions A to F. (**b**) *Xenopus laevis* 16S rRNA secondary structure showing the location of SPInDel conserved amphibian regions E, F and G. Identity plots represent the distribution of conserved (green bars) and variable (red bars) sites across the sequence alignment of rRNA genes in Eutheria (n = 237) and Amphibia (n = 84) groups. Identity plots were obtained in Geneious software, and the rRNA secondary structure diagrams were retrieved from the Comparative RNA Web (CRW) Site (http://www.rna.ccbb.utexas.edu).

Supplementary Figure S3

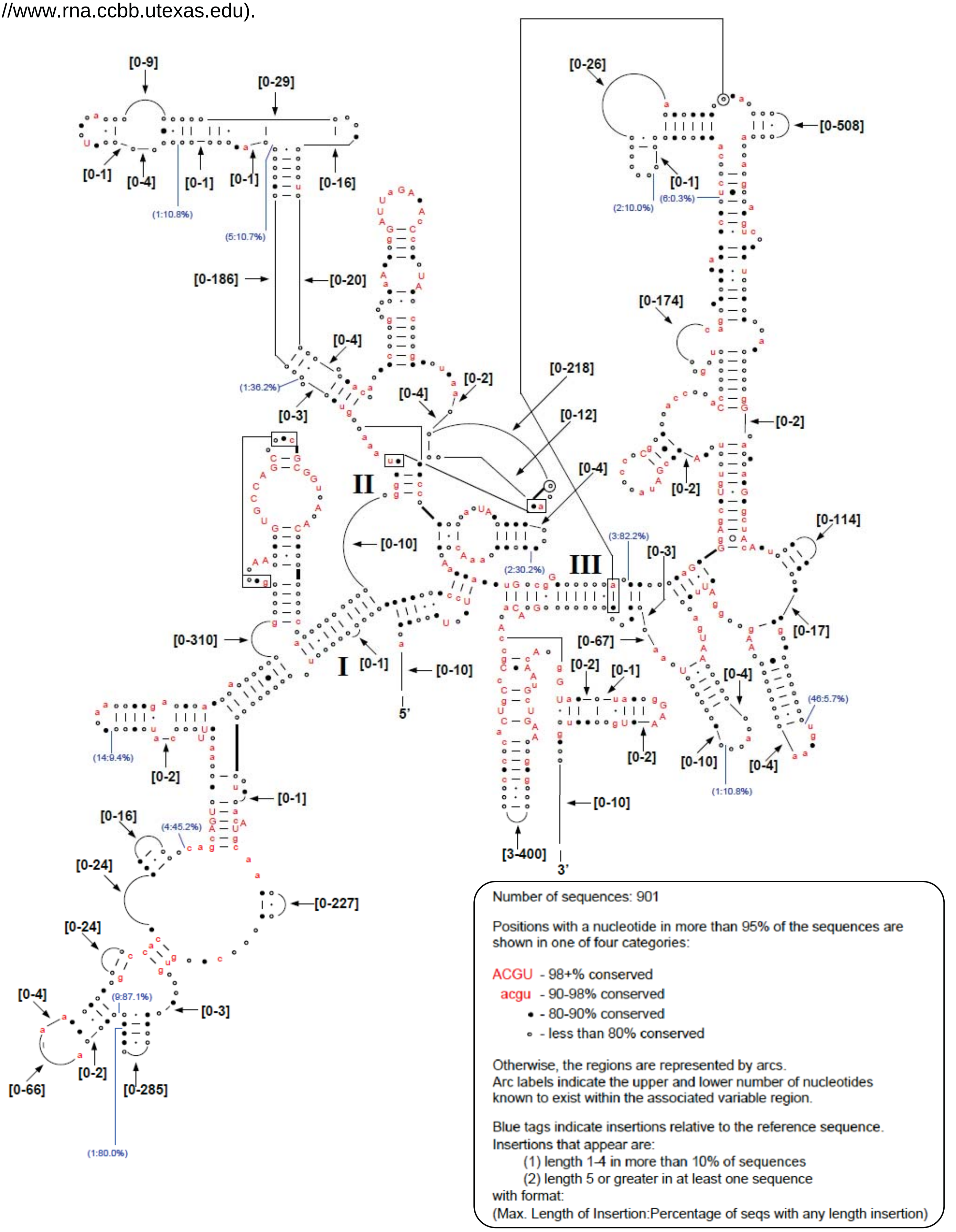
a



Supplementary Figure S3 (cont.)



Supplementary Figure S4. Phylogenetic conservation obtained from 901 eukaryotic mitochondrial small subunit rRNA aligned sequences superimposed onto the *Zea mays* (X00794) rRNA secondary structure model. Arcs and blue tags indicate regions with variable lengths. The diagram is available on the Comparative RNA Web (CRW) Site (http://www.rna.ccbb.utexas.edu).



Number of sequences: 901

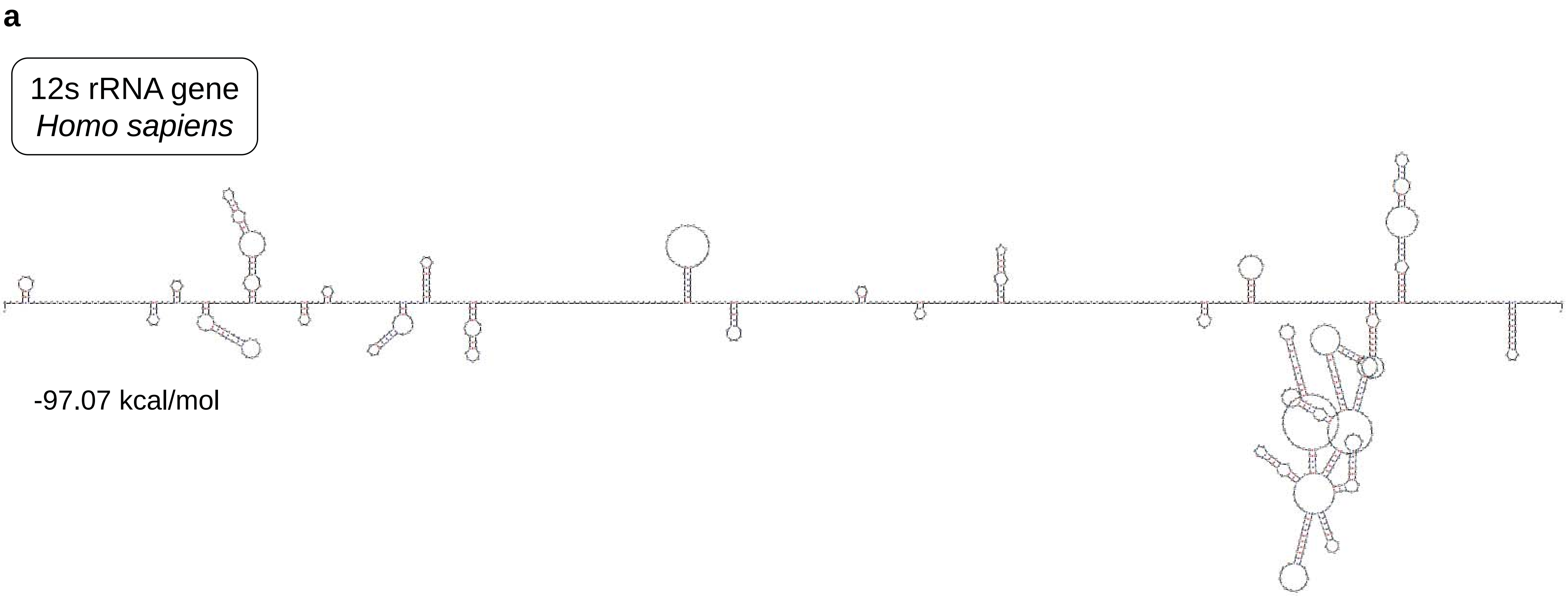
Positions with a nucleotide in more than 95% of the sequences are shown in one of four categories:

ACGU - 98+% conserved
acgu - 90-98% conserved
● - 80-90% conserved
○ - less than 80% conserved

Otherwise, the regions are represented by arcs.
Arc labels indicate the upper and lower number of nucleotides known to exist within the associated variable region.

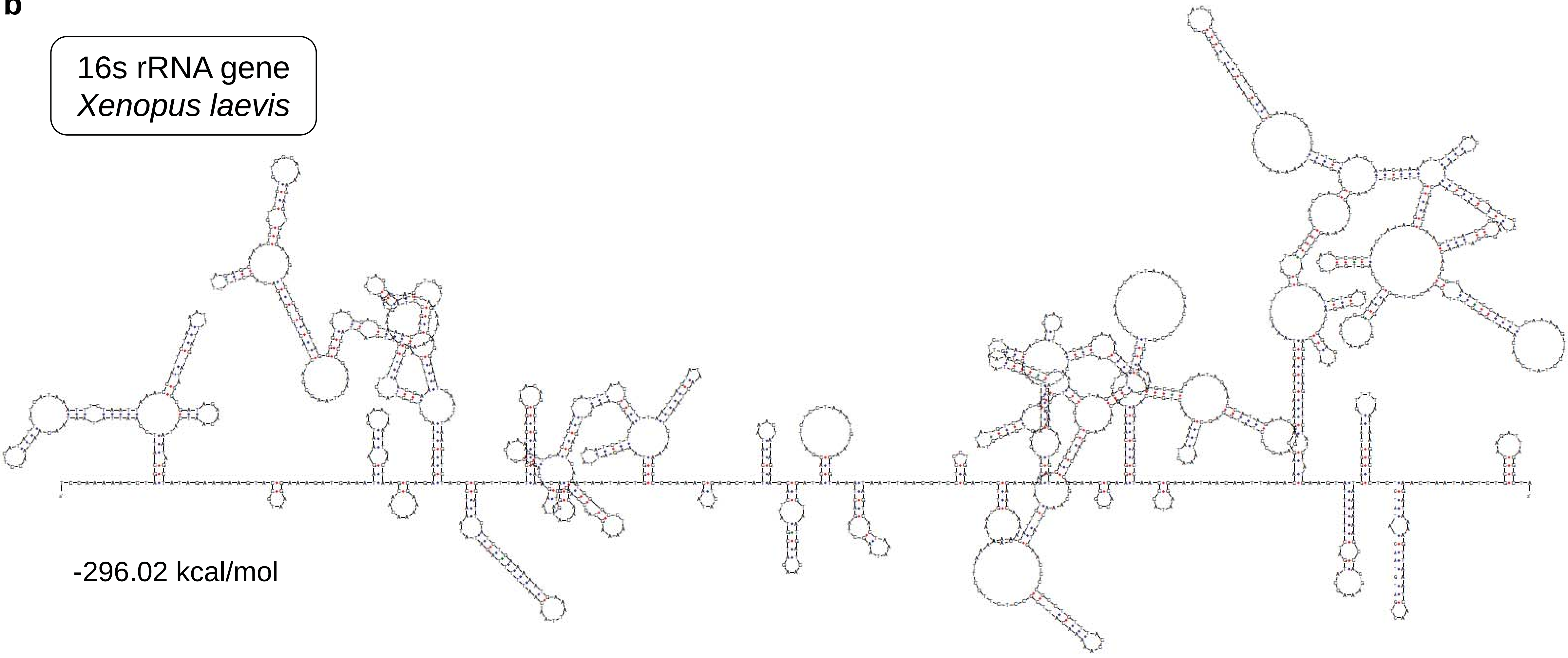
Blue tags indicate insertions relative to the reference sequence.
Insertions that appear are:
(1) length 1-4 in more than 10% of sequences
(2) length 5 or greater in at least one sequence
with format:
(Max. Length of Insertion:Percentage of seqs with any length insertion)

Supplementary Figure S5. Putative DNA secondary structures of mitochondrial ribosomal RNA genes. (**a**) DNA secondary structure for the *Homo sapiens* 12S rRNA gene (accession number J01415) at 37°C. (**b**) DNA secondary structure for the *Xenopus laevis* 16S rRNA gene (accession number M10217) at 20°C. Folding energies are indicated. Structures were predicted using the *mfold* web server.

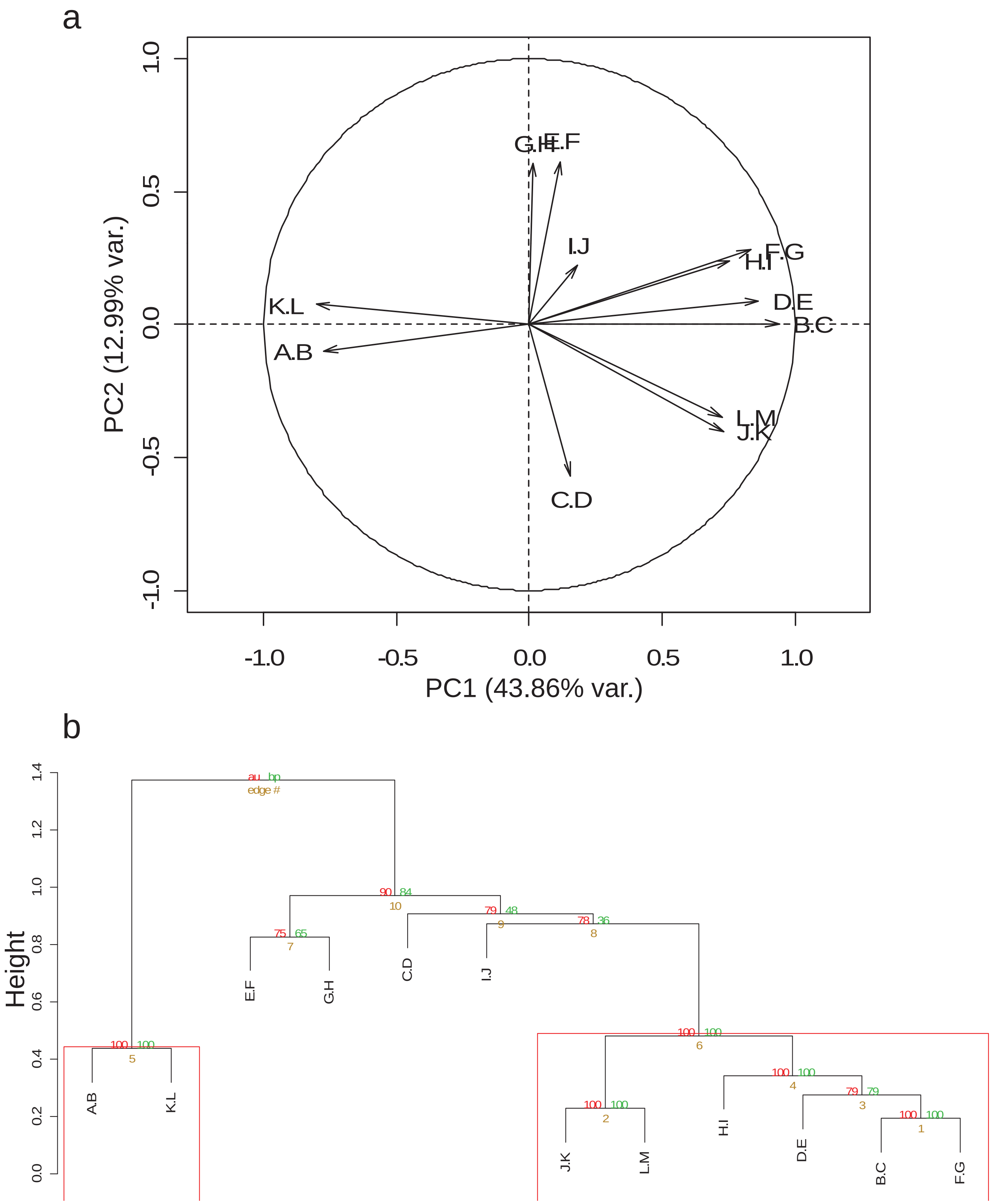


Supplementary Figure S5 (cont.)

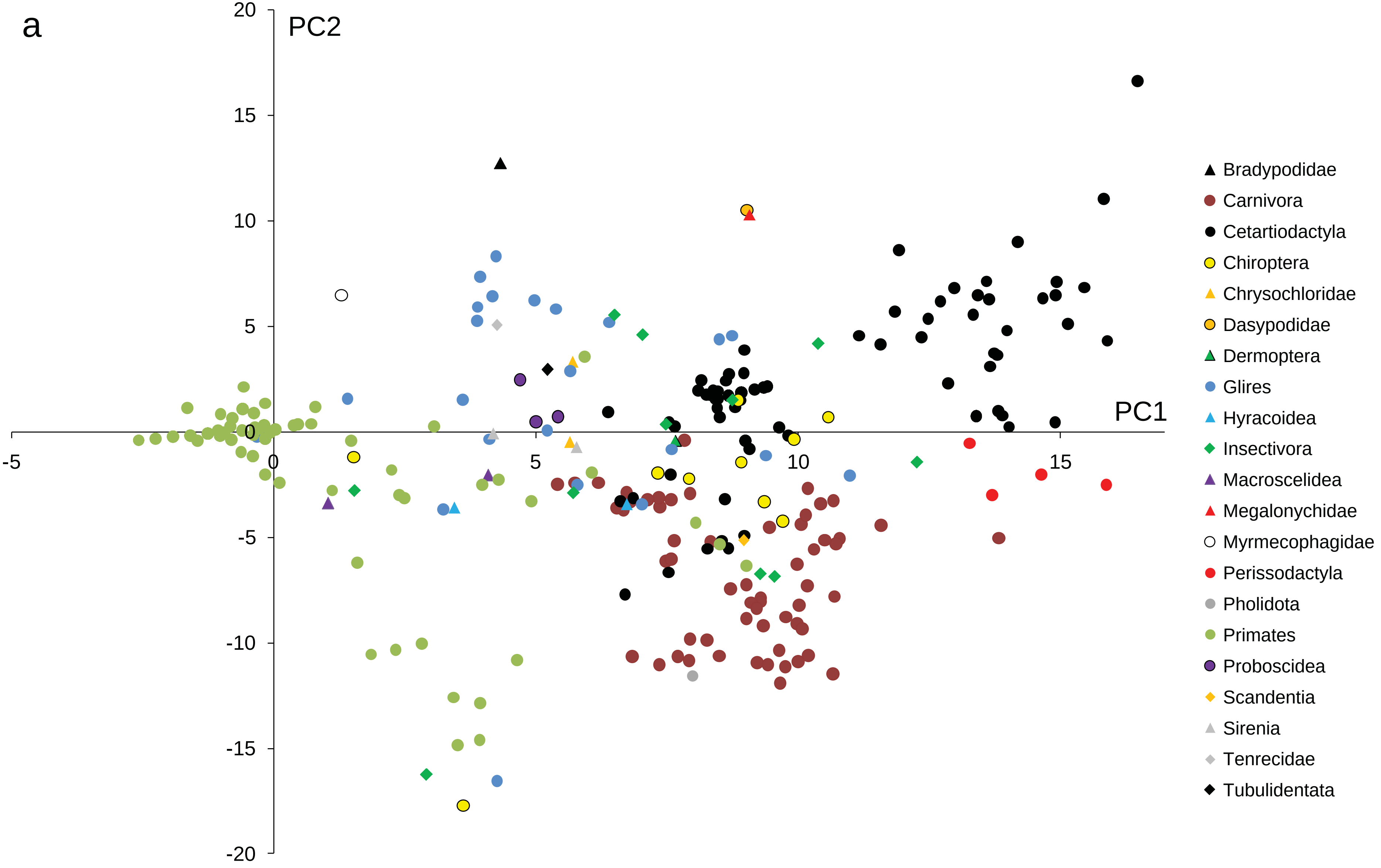
b



Supplementary Figure S6. Correlation between SPInDel hypervariable regions in Eutheria. (**a**) Principal component analysis. (**b**) Hierarchical clustering analysis. The hierarchical clustering was performed using average linkage correlation as a metric. Bootstrap probabilities (BP) were estimated with 1000 bootstraps.

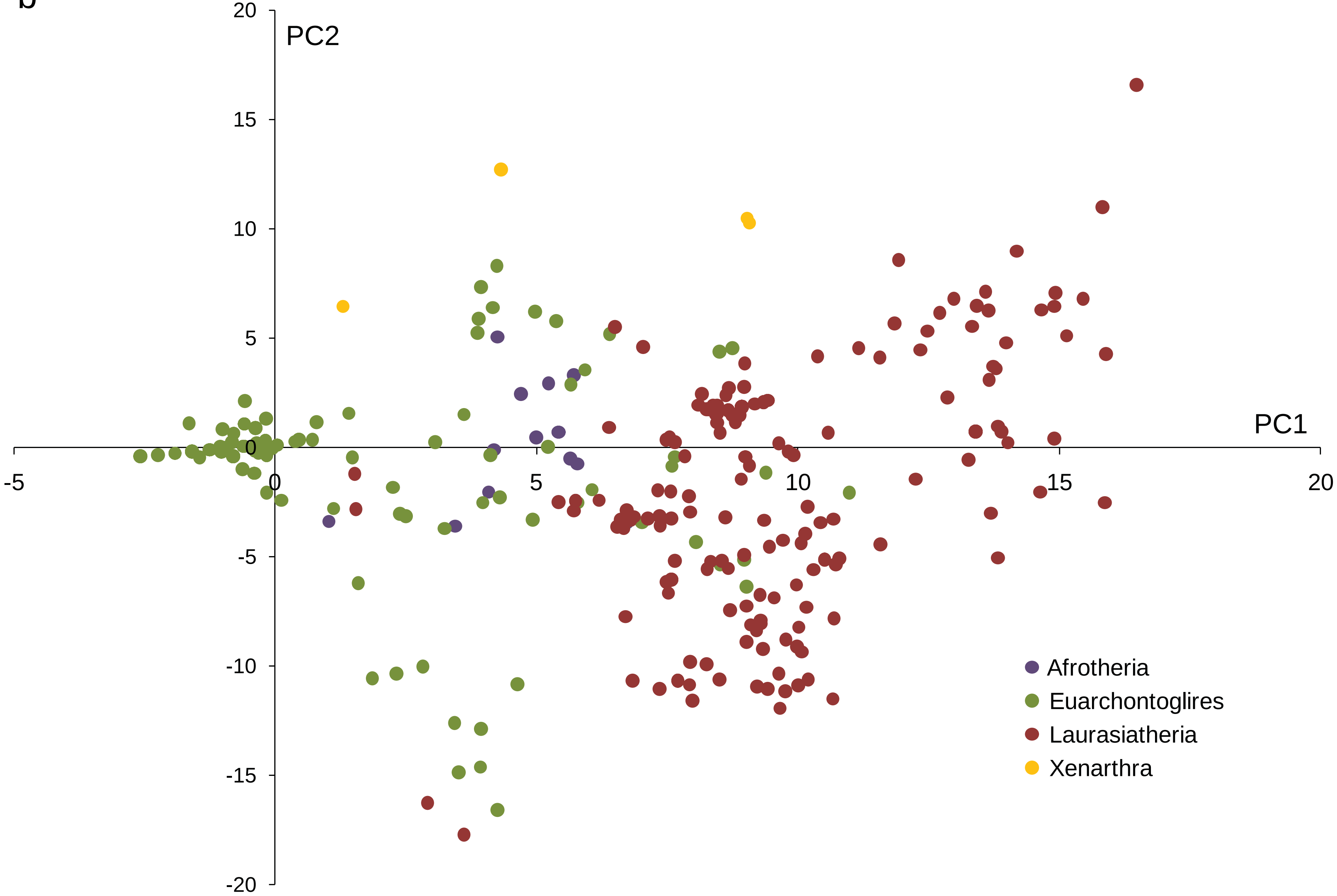


Supplementary Figure S7. Discrimination of 237 eutherian species using standard SPInDel profiles. The matrix of pairwise distances measured by the number of different loci between profiles was used for principal component (a, b) and multidimensional scaling (c) analyses. Species were labeled according to their taxonomic categories at two different levels. The input matrix was generated in the SPInDel workbench.



Supplementary Figure S7 (cont.)

b



Supplementary Figure S7 (cont.)