Supplementary Material

Gene Arrangement and Adaptive Evolution in the Mitochondrial Genomes of Terrestrial Sesarmid Crabs *Geosesarma faustum* and *Geosesarma penangensis*

Nyok-Sean Lau[1], Ka-Kei Sam[1], Amirrudin B. Ahmad[2,3], Khadijah-Ahmad Siti[2,3], Abdul Wahab Ahmad Zafir[4] and Alexander Chong Shu-Chien[1,5*]

[1]Centre for Chemical Biology, Universiti Sains Malaysia, Bayan Lepas, 11900 Penang, Malaysia.

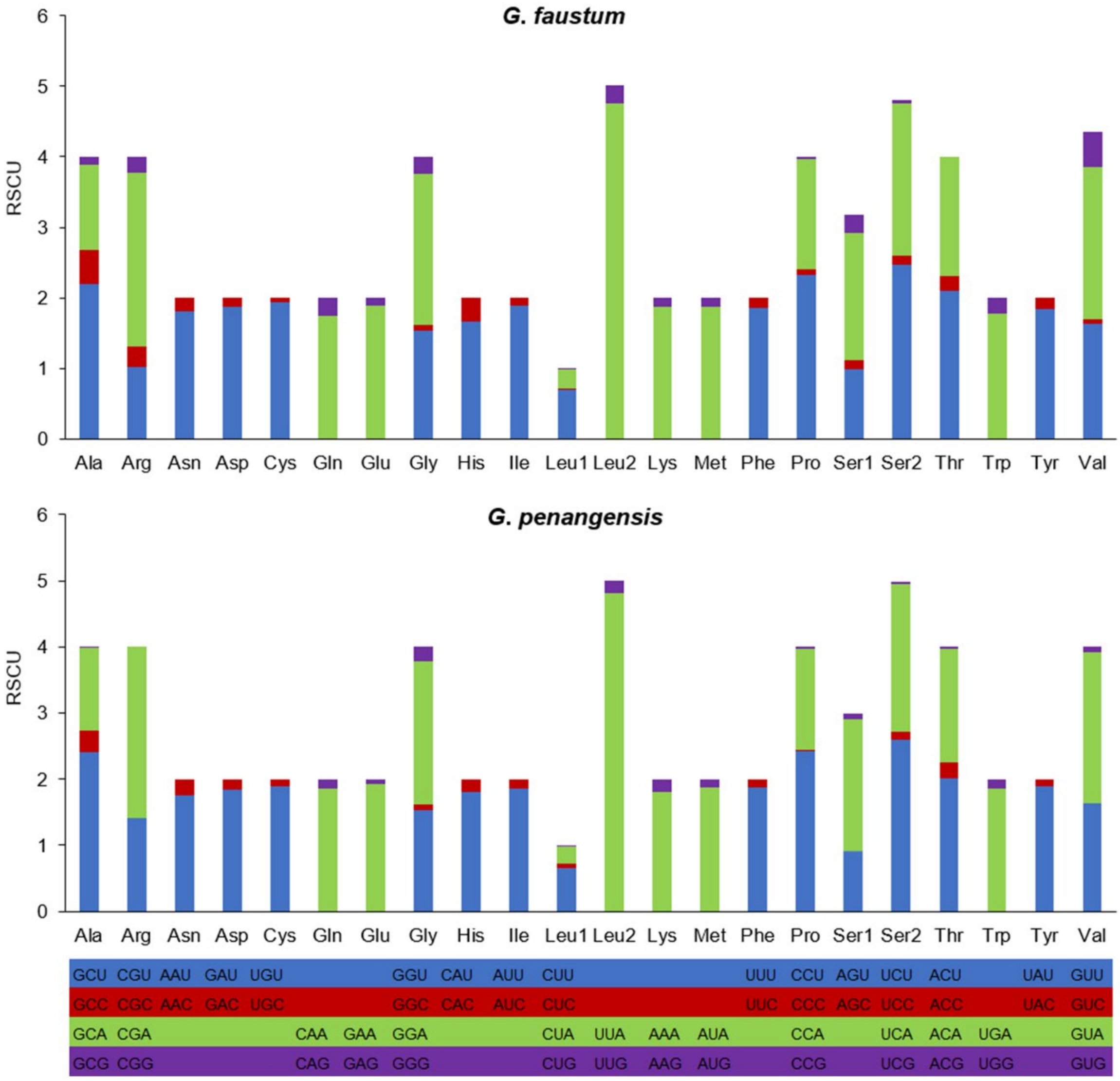
[2]Institute of Tropical Biodiversity and Sustainable Development, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia.

[3]Faculty of Science and Marine Environment, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia.

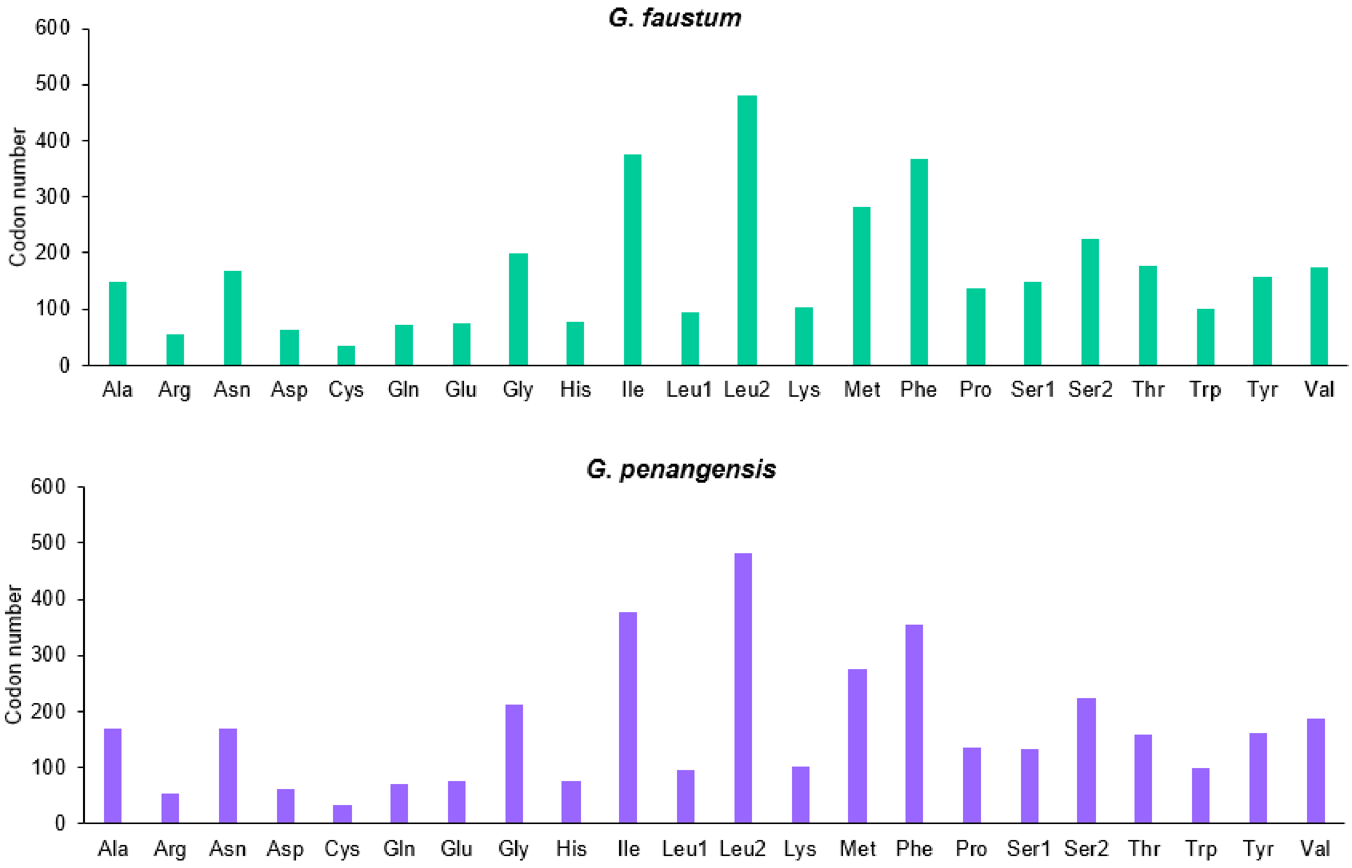
[4]The Habitat Foundation, Bukit Bendera, Penang, 11300, Malaysia.

[5]School of Biological Sciences, Universiti Sains Malaysia, Minden, 11800 Penang, Malaysia.

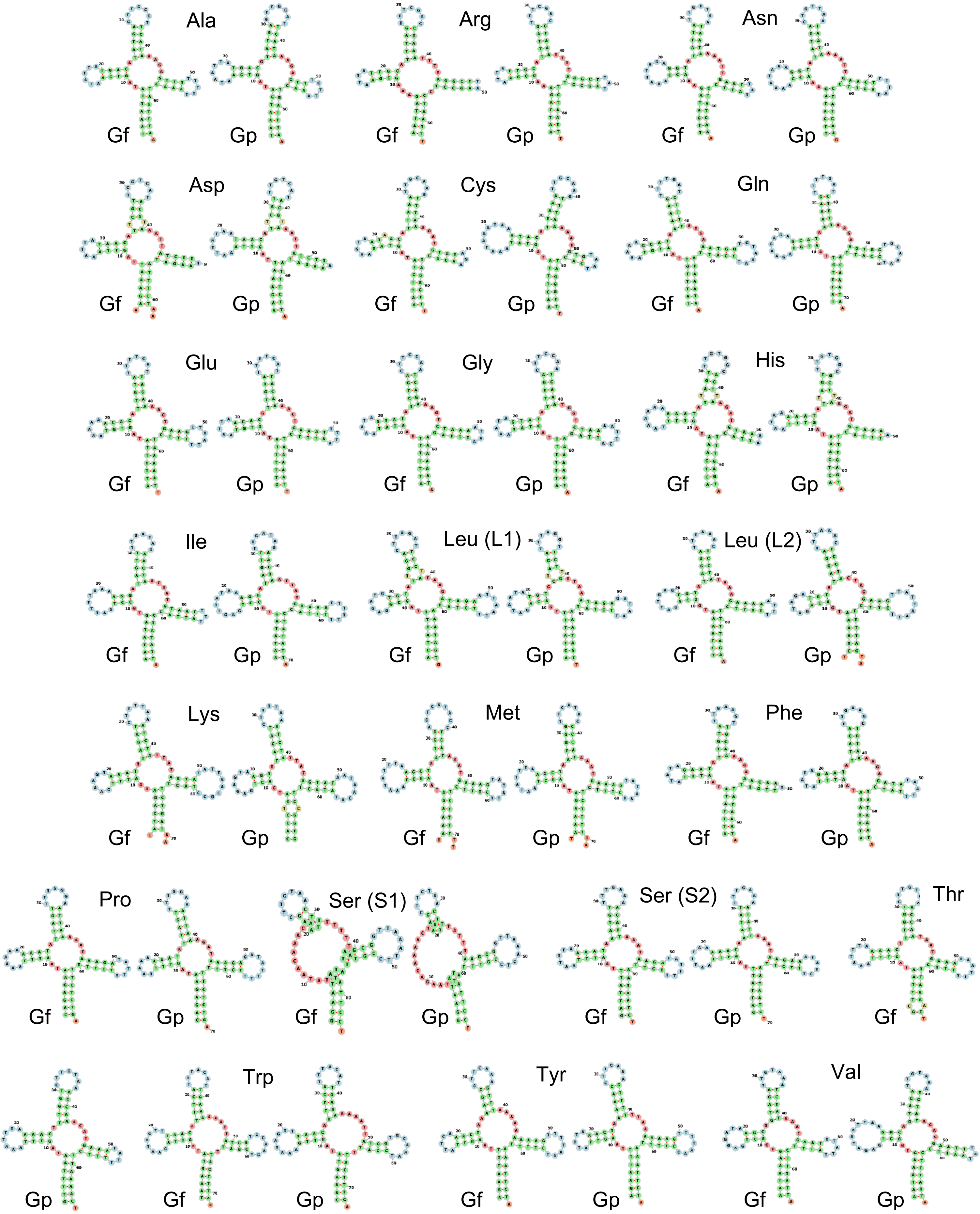
[*]Correspondence:
Alexander Chong Shu-Chien
alex@usm.my



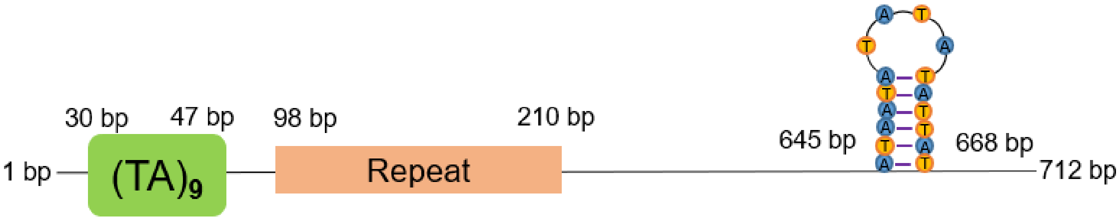
Supplementary Figure S1 Relative synonymous codon usage in the *G. faustum* and *G. penangensis* mitogenomes.



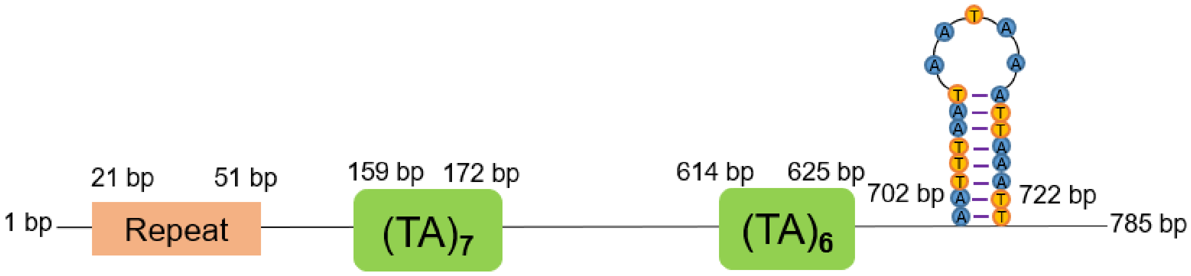
Supplementary Figure S2 Amino acid composition in the *G. faustum* and *G. penangensis* mitogenomes.



Supplementary Figure S3 Secondary structures of transfer RNA genes of *G. faustu*m (Gf) and *G. penangensis* (Gp) mitogenomes.

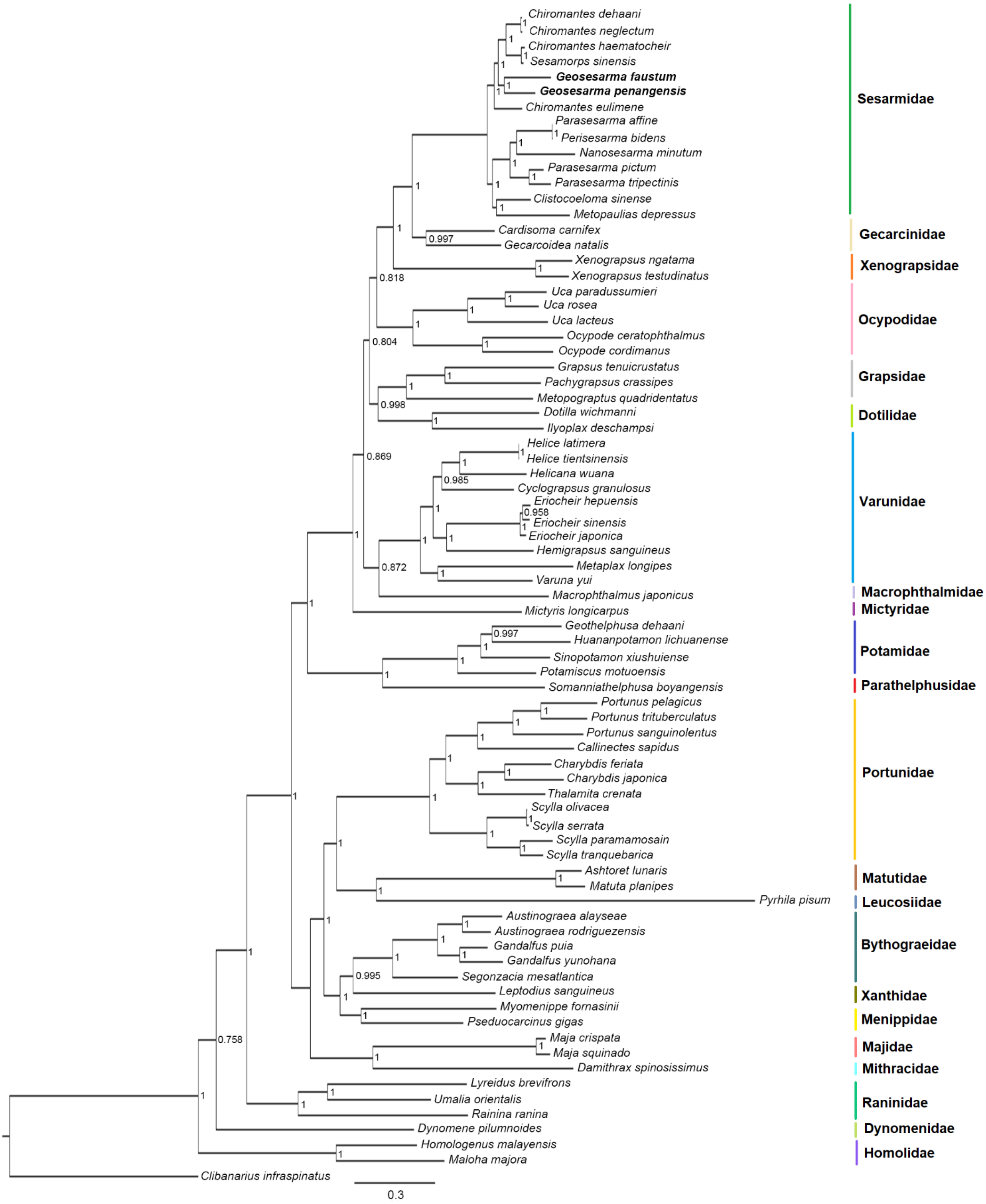


G. faustum



G. penangensis

Supplementary Figure S4 Control regions of *G. faustum* and *G. penangensis*.



Supplementary Figure S5 Phylogenetic tree inferred from nucleotide sequences of 13 protein-coding genes of the mitogenomes using Bayesian inference method. Numbers on nodes indicate posterior probability values, and sequences generated in this study are indicated in bold font.

Supplementary Table S1 List of Brachyura species analyzed in this study and their GenBank accession numbers.

Family	Species	Accession number
Bythograeidae	*Austinograea alayseae*	KC851803
Bythograeidae	*Austinograea rodriguezensis*	JQ035658
Bythograeidae	*Gandalfus puia*	KR002727
Bythograeidae	*Gandalfus yunohana*	EU647222
Bythograeidae	*Segonzacia mesatlantica*	KY541839
Dotillidae	*Dotilla wichmanni*	MH183129
Dotillidae	*Ilyoplax deschampsi*	JF909979
Dynomenidae	*Dynomene pilumnoides*	KT182070
Gecarcinidae	*Cardisoma carnifex*	MF461623
Gecarcinidae	*Gecarcoidea natalis*	MH816962
Grapsidae	*Grapsus tenuicrustatus*	KT878721
Grapsidae	*Metopograpsus quadridentatus*	MH183127
Grapsidae	*Pachygrapsus crassipes*	KC878511
Homolidae	*Homologenus malayensis*	KJ612407
Homolidae	*Moloha majora*	KT182069
Leucosiidae	*Pyrhila pisum*	KU343210
Macrophthalmidae	*Macrophthalmus japonicus*	KU343211
Majidae	*Maja crispata*	KY650651
Majidae	*Maja squinado*	KY650652
Matutidae	*Ashtoret lunaris*	LK391941
Matutidae	*Matuta planipes*	MG756601
Menippidae	*Myomenippe fornasinii*	LK391943
Menippidae	*Pseudocarcinus gigas*	AY562127
Mictyridae	*Mictyris longicarpus*	LN611670
Mithracidae	*Damithrax spinosissimus*	KM405516
Ocypodidae	*Ocypode ceratophthalmus*	LN611669
Ocypodidae	*Ocypode cordimanus*	KT896743
Ocypodidae	*Uca lacteus*	MH796169
Ocypodidae	*Uca paradussumieri*	MN072633
Ocypodidae	*Uca rosea*	MN072632
Parathelphusidae	*Somanniathelphusa boyangensis*	KU042042
Portunidae	*Callinectes sapidus*	AY363392
Portunidae	*Charybdis feriata*	KF386147
Portunidae	*Charybdis japonica*	FJ460517
Portunidae	*Portunus pelagicus*	KR153996
Portunidae	*Portunus sanguinolentus*	KT438509
Portunidae	*Portunus trituberculatus*	AB093006
Portunidae	*Scylla olivacea*	FJ827760
Portunidae	*Scylla paramamosain*	JX457150
Portunidae	*Scylla serrata*	HM590866
Portunidae	*Scylla tranquebarica*	FJ827759

Portunidae	*Thalamita crenata*	LK391945
Potamidae	*Geothelphusa dehaani*	AB187570
Potamidae	*Huananpotamon lichuanense*	KX639824
Potamidae	*Potamiscus motuoensis*	KY285013
Potamidae	*Sinopotamon xiushuiense*	KU042041
Raninidae	*Lyreidus brevifrons*	KM983394
Raninidae	*Ranina ranina*	AB752308
Raninidae	*Umalia orientalis*	KM365084
Sesarmidae	*Chiromantes dehaani*	MH593563
Sesarmidae	*Chiromantes eulimene*	MN877385
Sesarmidae	*Chiromantes haematocheir*	MH457175
Sesarmidae	*Chiromantes neglectum*	KX156954
Sesarmidae	*Clistocoeloma sinense*	KU589292
Sesarmidae	*Geosesarma faustum*	MZ725940
Sesarmidae	*Geosesarma penangensis*	MZ725941
Sesarmidae	*Metopaulius depressus*	KX118277
Sesarmidae	*Nanosesarma minutum*	MH899177
Sesarmidae	*Parasesarma affine*	MH310444
Sesarmidae	*Parasesarma pictum*	MG580780
Sesarmidae	*Parasesarma tripectinis*	KU343209
Sesarmidae	*Perisesarma bidens*	KY808394
Sesarmidae	*Sesarmops sinensis*	KR336554
Varunidae	*Cyclograpsus granulosus*	LN624373
Varunidae	*Eriocheir hepuensis*	FJ455506
Varunidae	*Eriocheir japonica*	FJ455505
Varunidae	*Eriocheir sinensis*	KP126617
Varunidae	*Helicana wuana*	KX344898
Varunidae	*Helice latimera*	KU589291
Varunidae	*Helice tientsinensis*	KR336555
Varunidae	*Hemigrapsus sanguineus*	KX456205
Varunidae	*Metaplax longipes*	MH899176
Varunidae	*Varuna yui*	MG756602
Xanthidae	*Leptodius sanguineus*	KT896744
Xenograpsidae	*Xenograpsus ngatama*	KY985236
Xenograpsidae	*Xenograpsus testudinatus*	EU727203
Diogenidae	*Clibanarius infraspinatus*	NC025776

Supplementary Table S2 Substitution saturation tests.

Gene regions	*Iss*	*Iss.cSym*[a]	p^b	*Iss.cAsym*[c]	p^d
1st codons	0.275	0.808	0.0000	0.554	0.0000
2nd codons	0.138	0.808	0.0000	0.554	0.0000
3rd codons	0.709	0.808	0.0000	0.554	0.0000
All codons	0.353	0.818	0.0000	0.572	0.0000

NumOUT = 32

[a] Index of substitution saturation assuming a symmetrical tree.

[b] Significant difference between *Iss* and *Iss.cSym* (two-tailed test).

[c] Index of substitution saturation assuming an asymmetrical tree.

[d] Significant difference between *Iss* and *Iss.cAsym* (two-tailed test).

Supplementary Table S3 Composition and skewness of mitogenomes in *Sesarmidae* species.

Species	A%	T%	G%	C%	A + T%	AT-skew	GC-skew	Length (bp)
Mitogenome								
Chiromantes dehaani	37.46	38.23	9.49	14.83	75.69	−0.010	−0.220	15,917
Chiromantes eulimene	37.13	38.40	9.71	14.76	75.53	−0.017	−0.208	15,894
Chiromantes haematocheir	37.33	38.28	9.44	14.95	75.61	−0.013	−0.226	15,899
Chiromantes neglectum	37.42	38.21	9.51	14.86	75.63	−0.010	−0.220	15,920
Clistocoeloma sinense	37.11	38.62	9.36	14.92	75.72	−0.020	−0.229	15,706
Geosesarma faustum	38.75	39.75	8.27	13.23	78.49	−0.013	−0.230	15,880
Geosesarma penangensis	38.3	40.14	8.51	13.06	78.44	−0.023	−0.211	15,955
Metopaulias depressus	37.94	39.38	8.71	13.97	77.32	−0.019	−0.232	15,765
Nanosesarma minutum	37.97	39.71	8.91	13.40	77.69	−0.022	−0.201	15,637
Parasesarma affine	36.60	38.23	10.07	15.10	74.83	−0.022	−0.200	15,638
Parasesarma pictum	36.60	39.00	9.82	14.57	75.60	−0.032	−0.194	15,611
Parasesarma tripectinis	36.23	37.98	10.11	15.67	74.22	−0.024	−0.216	15,612
Perisesarma bidens	36.61	38.29	10.06	15.13	74.81	−0.021	−0.202	15,641
Sesarmops sinensis	37.37	38.31	9.39	14.93	75.68	−0.012	−0.228	15,905
Protein-coding genes								
Chiromantes dehaani	30.99	42.89	13.25	12.87	73.88	−0.161	0.015	11,175
Chiromantes eulimene	30.48	43.22	13.72	12.59	73.69	−0.173	0.043	11,172
Chiromantes haematocheir	30.91	42.99	13.23	12.87	73.90	−0.163	0.014	11,175
Chiromantes neglectum	30.94	42.87	13.27	12.92	73.81	−0.162	0.013	11,169
Clistocoeloma sinense	31.04	43.16	13.25	12.56	74.20	−0.163	0.027	11,196
Geosesarma faustum	32.68	44.22	11.71	11.39	76.90	−0.150	0.014	11,178
Geosesarma penangensis	32.58	44.25	11.90	11.27	76.84	−0.152	0.027	11,181
Metopaulias depressus	31.96	43.72	12.54	11.77	75.68	−0.155	0.032	11,136
Nanosesarma minutum	32.04	44.03	12.30	11.62	76.07	−0.158	0.029	11,142
Parasesarma affine	30.31	42.65	13.84	13.21	72.95	−0.169	0.023	11,166
Parasesarma pictum	30.70	43.43	13.37	12.50	74.13	−0.172	0.034	11,166
Parasesarma tripectinis	30.11	42.25	13.93	13.71	72.36	−0.168	0.008	11,163
Perisesarma bidens	30.33	42.60	13.82	13.25	72.94	−0.168	0.021	11,166
Sesarmops sinensis	30.96	43.01	13.17	12.86	73.97	−0.163	0.012	11,175

Species	A%	T%	G%	C%	A + T%	AT-skew	GC-skew	Length (bp)
tRNAs								
Chiromantes dehaani	37.70	38.04	13.59	10.67	75.75	−0.004	0.120	1472
Chiromantes eulimene	37.41	38.02	13.92	10.65	75.43	−0.008	0.133	1465
Chiromantes haematocheir	36.90	37.85	14.43	10.82	74.74	−0.013	0.143	1469
Chiromantes neglectum	37.64	38.04	13.65	10.67	75.68	−0.005	0.123	1472
Clistocoeloma sinense	38.56	37.69	13.66	10.09	76.24	0.011	0.150	1486
Geosesarma faustum	39.47	39.47	11.86	9.20	78.94	0.000	0.126	1467
Geosesarma penangensis	39.05	38.91	12.26	9.78	77.96	0.002	0.112	1493
Metopaulias depressus	39.71	38.69	12.24	9.36	78.40	0.013	0.133	1463
Nanosesarma minutum	39.24	39.10	12.49	9.16	78.34	0.002	0.154	1473
Parasesarma affine	38.22	37.81	13.71	10.26	76.03	0.005	0.144	1481
Parasesarma pictum	38.29	38.16	13.46	10.08	76.45	0.002	0.144	1478
Parasesarma tripectinis	38.14	37.40	13.75	10.71	75.54	0.010	0.124	1484
Perisesarma bidens	38.34	38.07	12.94	10.65	76.42	0.004	0.097	1484
Sesarmops sinensis	37.22	37.36	13.91	11.52	74.57	−0.002	0.094	1467
rRNAs								
Chiromantes dehaani	39.71	40.59	13.13	6.56	80.31	−0.011	0.333	2,163
Chiromantes eulimene	40.46	40.00	12.67	6.87	80.46	0.006	0.297	2,155
Chiromantes haematocheir	40.30	40.11	12.59	7.01	80.41	0.002	0.285	2,169
Chiromantes neglectum	39.78	40.52	13.09	6.61	80.30	−0.009	0.329	2,162
Clistocoeloma sinense	39.85	40.36	12.96	6.83	80.21	−0.006	0.310	2,168
Geosesarma faustum	40.83	41.82	11.54	5.81	82.65	−0.012	0.330	2,219
Geosesarma penangensis	41.62	40.90	11.32	6.16	82.52	0.009	0.295	2,208
Metopaulias depressus	41.98	40.17	11.64	6.21	82.14	0.022	0.304	2,044
Nanosesarma minutum	42.69	39.68	11.51	6.13	82.37	0.037	0.305	2,155
Parasesarma affine	40.74	39.20	12.58	7.48	79.93	0.019	0.254	2,138
Parasesarma pictum	40.38	39.32	13.24	7.06	79.70	0.013	0.305	2,167
Parasesarma tripectinis	40.96	39.11	12.93	7.00	80.07	0.023	0.298	2,273
Perisesarma bidens	41.64	40.58	10.84	6.94	82.22	0.013	0.219	1,513
Sesarmops sinensis	39.65	39.76	13.34	7.25	79.41	−0.001	0.296	1,821
Control region								

Species	A%	T%	G%	C%	A + T%	AT-skew	GC-skew	Length (bp)
Chiromantes dehaani	46.35	36.04	8.94	8.67	82.39	0.125	0.016	727
Chiromantes eulimene	48.30	33.85	7.79	10.06	82.15	0.176	−0.127	706
Chiromantes haematocheir	46.44	37.18	7.25	9.13	83.62	0.111	−0.115	745
Chiromantes neglectum	46.36	36.08	8.64	8.92	82.44	0.125	−0.016	729
Clistocoeloma sinense	43.42	39.47	6.58	10.53	82.89	0.048	−0.231	684
Geosesarma faustum	43.46	39.10	7.59	9.85	82.56	0.053	−0.129	711
Geosesarma penangensis	44.59	39.11	6.88	9.43	83.69	0.065	−0.156	785
Metopaulias depressus	41.06	40.58	7.32	11.04	81.63	0.006	−0.203	833
Nanosesarma minutum	47.06	37.30	7.32	8.32	84.36	0.116	−0.064	697
Parasesarma affine	43.36	38.94	7.37	10.32	82.30	0.054	−0.167	678
Parasesarma pictum	43.33	38.10	7.14	11.43	81.43	0.064	−0.231	630
Parasesarma tripectinis	43.94	37.12	7.77	11.17	81.06	0.084	−0.180	528
Perisesarma bidens	43.51	38.79	7.23	10.47	82.30	0.057	−0.183	678
Sesarmops sinensis	46.07	37.15	7.46	9.32	83.22	0.107	−0.111	751

Supplementary Table S4 Summary of *G. faustum* and *G. penangensis* mitogenomes.

G. faustum							*G. penangensis*						
Gene	Direction	Location	Length (bp)	Start codon	Stop codon	Intergenic nucleotide	Gene	Direction	Location	Length (bp)	Start codon	Stop codon	Intergenic nucleotide
cox1	F	1-1534	1534	ATG	T	0	*cox1*	F	1-1539	1539	ATG	TAA	0
trnL2	F	1536-1601	66	−	−	1	*trnL2*	F	1535-1602	68	−	−	−5
cox2	F	1608-2295	688	ATG	T	6	*cox2*	F	1609-2296	688	ATG	T	6
trnK	F	2296-2366	71	−	−	0	*trnK*	F	2297-2365	69	−	−	0
trnD	F	2366-2428	63	−	−	−1	*trnD*	F	2366-2431	66	−	−	0
atp8	F	2429-2587	159	ATG	TAA	0	*atp8*	F	2432-2590	159	ATG	TAA	0
atp6	F	2587-3255	669	ATA	TAA	−1	*atp6*	F	2587-3258	672	ATA	TAA	−4
cox3	F	3255-4046	792	ATG	TAA	−1	*cox3*	F	3258-4049	792	ATG	TAA	−1
trnG	F	4046-4111	66	−	−	−1	*trnG*	F	4049-4114	66	−	−	−1
nad3	F	4109-4462	354	ATA	TAA	-3	*nad3*	F	4121-4465	345	ATA	TAA	6
trnA	F	4465-4530	66	−	−	2	*trnA*	F	4468-4532	65	−	−	2
trnR	F	4538-4600	63	−	−	7	*trnR*	F	4538-4601	64	−	−	5
trnN	F	4601-4665	65	−	−	0	*trnN*	F	4604-4672	69	−	−	2
trnS1	F	4670-4733	64	−	−	4	*trnS1*	F	4675-4740	66	−	−	2
trnE	F	4735-4800	66	−	−	1	*trnE*	F	4741-4806	66	−	−	0
trnH	R	4805-4868	64	−	−	4	*trnH*	R	4811-4873	63	−	−	4
trnF	R	4871-4933	63	−	−	2	*trnF*	R	4874-4938	65	−	−	0
nad5	R	4946-6676	1731	ATA	TAA	12	*nad5*	R	4945-6687	1743	ATA	TAA	6
nad4	R	6686-8050	1365	ATG	TAA	9	*nad4*	R	6689-8053	1365	ATG	TAA	1
nad4L	R	8044-8346	303	ATG	TAA	−7	*nad4L*	R	8047-8349	303	ATG	TAA	−7
trnT	F	8356-8423	68	−	−	9	*trnT*	F	8358-8423	66	−	−	8
trnP	R	8424-8491	68	−	−	0	*trnP*	R	8424-8493	70	−	−	0
nad6	F	8500-8997	498	ATA	TAA	8	*nad6*	F	8502-8999	498	ATA	TAA	8
cob	F	9003-10,131	1129	ATA	T	5	*cob*	F	9003-10,134	1132	ATA	T	3

trnS2	F	10,133-10,199	67	−	−	1	*trnS2*	F	10,135-10,204	70	−	−	0
nad1	R	10,222-11,169	948	GTG	TAA	22	*nad1*	R	10,224-11,162	939	ATA	TAA	19
trnL1	R	11,203-11,270	68	−	−	33	*trnL1*	R	11,203-11,270	68	−	−	40
rrnL	R	11,216-12,609	1394	−	−	−55	*rrnL*	R	11,246-12,609	1364	−	−	−25
trnV	R	12,612-12,683	72	−	−	2	*trnV*	R	12,611-12,683	73	−	−	1
rrnS	R	12,684-13,508	825	−	−	0	*rrnS*	R	12,684-13,527	844	−	−	0
CR	−	13,509-14,220	712	−	−	0	CR	−	13,528-14,312	785	−	−	0
trnQ	R	14,221-14,289	69	−	−	0	*trnI*	F	14,313-14,383	71	−	−	0
trnI	F	14477-14,544	68	−	−	187	*trnQ*	R	14,484-14,555	72	−	−	100
trnM	F	14,598-14,669	72	−	−	53	*trnM*	F	14,667-14,737	71	−	−	111
nad2	F	14669-15,679	1011	ATG	TAG	−1	*nad2*	F	14,738-15,745	1008	ATG	TAG	0
trnW	F	15,679-15,747	69	−	−	−1	*trnW*	F	15,744-15,812	69	−	−	−2
trnC	R	15,752-15,815	64	−	−	4	*trnC*	R	15,818-15,884	67	−	−	5
trnY	R	15,816-15,880	65	−	−	0	*trnY*	R	15,887-15,955	69	−	−	2

Supplementary Table S5 Composition and skewness of *G. faustum* (Gf) and *G. penangensis* (Gp) mitogenomes.

	A%		T%		G%		C%		A + T%		AT-skew		GC-skew	
	Gf	Gp	Gf	Gp	Gf	Gp	Gf	Gp	Gf	Gp	Gf	Gp	Gf	Gp
Mitogenome	38.75	38.30	39.75	40.14	8.27	8.51	13.23	13.06	78.49	78.44	−0.013	−0.023	−0.230	−0.211
PCGs	32.67	32.58	44.24	44.26	11.71	11.89	11.38	11.27	76.91	76.84	−0.150	−0.152	0.014	0.011
cox1	31.68	31.25	40.09	39.77	13.82	14.42	14.41	14.55	71.77	71.02	−0.117	−0.120	−0.021	−0.004
cox2	35.32	34.45	40.99	39.68	11.19	11.05	12.50	14.83	76.31	74.13	−0.074	−0.071	−0.055	−0.146
atp8	35.85	36.48	45.91	52.83	3.14	2.52	15.09	8.18	81.76	89.31	−0.123	−0.183	−0.655	−0.529
atp6	34.08	34.08	41.41	41.07	9.12	9.23	15.40	15.63	75.49	75.15	−0.097	−0.093	−0.256	−0.257
cox3	32.83	29.80	40.03	42.80	11.99	13.38	15.15	14.02	72.85	72.60	−0.099	−0.179	−0.116	−0.023
nad3	33.33	33.62	44.07	46.09	9.04	8.41	13.56	11.88	77.40	79.71	−0.139	−0.156	−0.200	−0.171
nad5	34.43	35.97	44.43	43.83	13.81	13.31	7.34	6.88	78.86	79.80	−0.127	−0.098	0.306	0.318
nad4	34.07	33.33	45.20	45.93	13.77	14.21	6.96	6.52	79.27	79.27	−0.140	−0.159	0.329	0.371
nad4L	32.34	33.33	48.84	48.18	15.18	13.20	3.63	5.28	81.19	81.52	−0.203	−0.182	0.614	0.429
nad6	29.72	30.92	53.41	50.00	5.02	5.62	11.85	13.45	83.13	80.92	−0.285	−0.236	−0.405	−0.411
cob	32.15	31.71	43.05	42.76	9.65	10.60	15.15	14.93	75.20	74.47	−0.145	−0.148	−0.221	−0.170
nad1	28.80	29.82	47.78	46.96	15.61	15.44	7.81	7.77	76.58	76.78	−0.248	−0.223	0.333	0.330
nad2	31.45	30.75	48.17	49.21	7.12	7.14	13.25	12.90	79.62	79.96	−0.210	−0.231	−0.301	−0.287
tRNAs	39.40	39.05	39.40	38.91	12.00	12.26	9.20	9.78	78.80	77.96	0.000	0.002	0.132	0.112
rRNAs	40.78	41.62	41.91	40.90	11.49	11.32	5.81	6.16	82.69	82.52	−0.014	0.009	0.328	0.295
CR	43.46	44.49	39.10	39.11	7.59	6.88	9.85	9.43	82.56	83.69	0.053	0.065	−0.129	−0.156