

Supplementary Material

Genetic assessment of the rare freshwater shrimp *Caridina logemanni* endemic to Hong Kong and its hybridisation with a widespread congener

GLai Him Chow^A, *Ling Ming Tsang*^A, *Ka Hou Chu*^A and *Ka Yan Ma*^{A,B,*}

^ASimon F. S. Li Marine Science Laboratory, School of Life Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, PR China.

^BSchool of Ecology, Sun Yat-sen University, Shenzhen, PR China.

*Correspondence to: Email: majx26@mail.sysu.edu.cn

Table S1. Primers and PCR protocols used in the amplification of mitochondrial *COI* and *NaK* intron markers.

Locus	Primer	Primer sequence (5' to 3')	Source	PCR condition	PCR profile
<i>COI</i>	COI-CC-for	F: ATTTCGAGCAGAATTAGGGCA	Tsang <i>et al.</i> (2017)	1st round	Initial denaturation: 95°C;
	CysF	F: CTAGAATTGCAGTCTAGCGTC	Ma <i>et al.</i> (2021)	1×PCR reaction buffer,	3 min
	dgH2198	R: TAAACTTCAGGGTGACCAAARAAYCA	Folmer <i>et al.</i> (1994)	0.5 mM of MgCl ₂ ,	33 cycles of
				0.4 μM of each primer,	denaturation: 95°C; 30s
				400 μM of dNTPs,	annealing: 50°C; 40s
				1.5 U of Taq polymerase	extension: 72°C; 1 min
				(TaKaRa),	Final extension:
				0.4 μL of template DNA,	72°C; 3 min
				MilliQ H ₂ O to 25 μL	1st round
<i>NaK</i> intron	F (1st round)	F: GCCTTCTTCTCCACCAACGCCGTTGAAGG	Modified from Michez <i>et al.</i> (2009)	2nd round	Initial denaturation: 95°C;
	nr (1st round)	R: ATAGGGTGATCTCCAGTRACCAT	Tsang <i>et al.</i> (2008)	Same as 1st round, except	3 min
	int254F (2nd round)	F: AACCCCCATTGCCAAGGAAA	Present study	template DNA replaced by	33 cycles of
	int674R (2nd round)	R: CAGCAGACCCTCAGGCAC	Present study	0.25 μL of 5× diluted	denaturation: 95°C; 30s
	int655R (2nd round)	R: CACRTTTGCAACRATAATRCCA	Present study	PCR products in 25 μL	annealing: 56–0.3°C/cycle
			mixture	(20 cycles)+50°C (10	
				cycles); 40 s	extension: 72°C; 2 min
					Final extension:
					72°C; 5 min
					2nd round
					Same as <i>COI</i>

References

- Folmer, O., Black, M. B., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.
- Ma, K. Y., Chow, L. H., Tsang, L. M., De Grave, S., and Chu, K. H. (2021). Contrasting population structures of freshwater atyid shrimps in Hong Kong and their conservation implications. *Marine and Freshwater Research* **72**(11), 1667–1678 [doi:10.1071/MF21069](https://doi.org/10.1071/MF21069).
- Michez, D., Patiny, S., and Danforth, B. N. (2009). Phylogeny of the bee family Melittidae (Hymenoptera: Anthophila) based on combined molecular and morphological data. *Systematic Entomology* **34**, 574–597. [doi:10.1111/j.1365-3113.2009.00479.x](https://doi.org/10.1111/j.1365-3113.2009.00479.x)
- Tsang, L. M., Ma, K. Y., Ahyong, S. T., Chan, T.-Y., and Chu, K. H. (2008). Phylogeny of Decapoda using two nuclear protein-coding genes: origin and evolution of the Reptantia. *Molecular Phylogenetics and Evolution* **48**, 359–368. [doi:10.1016/j.ympev.2008.04.009](https://doi.org/10.1016/j.ympev.2008.04.009)
- Tsang, L. M., Tsoi, K. H., Chan, S. K. F., Chan, T. K. T., and Chu, K. H. (2017). Strong genetic differentiation among populations of the freshwater shrimp *Caridina cantonensis* in Hong Kong: implications for conservation of freshwater fauna in urban areas. *Marine and Freshwater Research* **68**, 187–194. [doi:10.1071/MF15377](https://doi.org/10.1071/MF15377)

Table S2. Primers and PCR protocol used in the amplification of microsatellites.

Multiplex	Locus	Fluorescent dye	Primer sequences (5' to 3')	PCR condition	PCR profile
A	O17	ROX	F: AGGTAGCTGAGCGATGACC R: CTGCCCTTCGCTGTTCAGT	2× Type-it multiplex PCR master mix, 0.2 μM of each ROX-labeled and FAM-labeled primer, 0.4 μM of each HEX-labeled primer, 200 μM of dNTPs, 0.9 μL of template DNA, MilliQ H ₂ O to 6.25 μL	Initial denaturation: 95°C; 5 min 30 cycles of denaturation: 95°C; 30 s annealing: 54°C; 90 s extension: 72°C; 30 s Final extension: 60°C; 30 min
	N15	FAM	F: ACGCATGATGGAAAGGCAA R: TCACAAAGTCACGACTAAGAT		
B	N9	ROX	F: TGTGTTGGCAGATTTCGTCT R: GGCATGCTTAAACACATCCT		
	C8	FAM	F: GGCACAGTAAACAATGCGCT R: TAACAGCCGGTTGAGAGGC		
	C20	HEX	F: AGAGGCGATGGTTGGCATA R: GGTGCTCGACCGGTA ACTA		
C	N11	ROX	F: TTCAGTCAGCCAAACGACC R: TGTTGCTAAGTGTGCCTATTCT		

Table S3. Sampling localities, abbreviations and genetic diversity indices for mitochondrial *COI* data of *C. logemanni* and *C. cantonensis*.

Sampling localities	Abb.	<i>n</i>	<i>N_a</i>	<i>h</i>	π	Tajima's <i>D</i>	Fu's <i>F_s</i>	Latitude (N)	Longitude (E)
<i>C. logemanni</i>									
Overall		39	7	0.7045 ± 0.0501	0.04630 ± 0.02331	3.63977	17.12533		
North-eastern New Territories (1)	NHS	19	5	0.5906 ± 0.1185	0.02461 ± 0.01318	0.11724	7.61287		
North-eastern New Territories (2)	NWC	20	4	0.2842 ± 0.1284	0.01687 ± 0.00928	-1.43289	7.37209		
<i>C. cantonensis</i>									
Overall		145	18	0.8674 ± 0.0119	0.00733 ± 0.00418	-1.03204	-2.33678		
Fung Wong Wat	NFW	28	4	0.3254 ± 0.1102	0.00157 ± 0.00135	-1.21638	-0.43122	22°29.123'	114°18.333'
Ha Miu Tin	NHMT	30	5	0.3080 ± 0.1075	0.00098 ± 0.00099	-1.42765	-2.81424*	22°30.000'	114°16.000'
Luk Keng	NLG	26	3	0.4954 ± 0.0766	0.00113 ± 0.00109	-0.01682	0.04503	22°31.207'	114°13.684'
Nam Chung Lo Uk	NLU	10	4	0.7333 ± 0.1005	0.01444 ± 0.00843	2.02912	3.95049	22°31.087'	114°12.442'
Wu Kau Tang	NWKT	32	4	0.6754 ± 0.0474	0.00210 ± 0.00163	0.67883	0.37458	22°30.331'	114°14.540'
Wilson Trail Section 9	NWN	19	4	0.5088 ± 0.1171	0.00478 ± 0.00309	0.90916	1.99245	22°29.350'	114°13.470'

Abb., abbreviation; *n*, sample size; *N_a*, number of haplotype; *h* (mean ± s.d.), haplotype diversity; π (mean ± s.d.), nucleotide diversity. *, *P* < 0.05.

Data of *C. cantonensis* includes that from Wong *et al.* (2019).

Table S4. Sampling localities, abbreviations and genetic diversity indices for nuclear *NaK* intron data of *C. logemanni* and *C. cantonensis*.

Sampling localities	Abb.	<i>n</i>	<i>Na</i>	<i>h</i>	π	Tajima's <i>D</i>	Fu's <i>F_s</i>	<i>H_O</i>	<i>H_E</i>
<i>C. logemanni</i>									
Overall		66	3	0.5408 ± 0.0368	0.01178 ± 0.00666	1.56564	10.18031		
North-eastern New Territories (1)	NHS	26	2	0.2708 ± 0.0990	0.00587 ± 0.00385	0.09840	5.60494	0.308	0.271
North-eastern New Territories (2)	NWC	40	3	0.2718 ± 0.0869	0.00541 ± 0.00357	-1.09582	3.83878	0.200	0.272
<i>C. cantonensis</i>									
Overall		324	13	0.7889 ± 0.0137	0.01700 ± 0.00907	0.82159	5.77908		
Fung Wong Wat	NFW	72	2	0.0548 ± 0.0366	0.00102 ± 0.00115	-1.54647*	1.26871	0.056	0.055
Ha Miu Tin	NHMT	52	1	0.0000 ± 0.0000	0.00000 ± 0.00000	/	/	/	/
Luk Keng	NLG	48	3	0.2899 ± 0.0762	0.00407 ± 0.00287	-0.07242	2.95566	0.250	0.290
Nam Chung Lo Uk	NLU	48	4	0.5115 ± 0.0700	0.01453 ± 0.00804	0.70720	8.67185	0.375	0.511
Wu Kau Tang	NWKT	56	5	0.6052 ± 0.0604	0.01122 ± 0.00641	0.82673	5.41387	0.607	0.605
Wilson Trail Section 9	NWN	48	3	0.5665 ± 0.0582	0.01131 ± 0.00647	2.55256	8.72826	0.458	0.467

Abb., abbreviation; *n*, sample size; *Na*, number of allele; *h* (mean ± s.d.), genotype diversity; π (mean ± s.d.), nucleotide diversity; *H_O*, observed heterozygosity; *H_E*, expected heterozygosity. Asterisks in *H_O* and *H_E* were obtained from the Hardy–Weinberg equilibrium test. *, *P* < 0.05; **, *P* < 0.01.

Table S5. Sampling localities, abbreviations and genetic diversity indices at six microsatellite loci for *C. logemanni*.

Sampling localities	Abb.		Locus						
			Mean	C8	C20	N9	N11	N15	O17
Total		<i>Na</i>		11	4	14	11	8	21
Northeastern New Territories (1)	NWC	<i>n</i>	14.333	11	15	7	18	17	18
		<i>Na</i>	5.833	8	1	9	5	3	9
		<i>Ae</i>	3.822	5.261	1.000	7.538	2.906	1.908	4.320
		<i>H_O</i>	0.606	0.545	0.000	1.000	0.778	0.647	0.667
		<i>H_E</i>	0.596	0.810	0.000	0.867	0.656	0.476	0.769
		<i>P_{HW}</i>		0.044	/	0.783	0.752	0.158	0.890
		<i>F</i>	-0.048	0.327	/	-0.153	-0.186	-0.360	0.133
Northeastern New Territories (2)	NHS	<i>n</i>	14.833	11	15	17	16	16	14
		<i>Na</i>	9.167	8	4	10	9	6	18
		<i>Ae</i>	5.261	4.246	1.724	4.817	4.452	2.327	14.000
		<i>H_O</i>	0.605	0.909	0.267	0.588	0.625	0.313	0.929
		<i>H_E</i>	0.709	0.764	0.420	0.792	0.775	0.570	0.929
		<i>P_{HW}</i>		0.964	0.500	0.000*	0.000*	0.004*	0.647
		<i>F</i>	0.180	-0.189	0.365	0.258	0.194	0.452	0.000
Grand mean		<i>Na</i>	7.500						
		<i>Ae</i>	4.542						
		<i>H_O</i>	0.606						
		<i>H_E</i>	0.652						
		<i>F</i>	0.076						

Abb., abbreviation; *n*, sample size; *Na*, number of allele; *Ne*, number of effective alleles; *H_O*, observed heterozygosity; *H_E*, expected heterozygosity; *P_{HW}*, probability of deviation from HWE; *F*, fixation index. *, loci which deviate from HWE after sequential Bonferroni correction ($\alpha = 0.05$).

Table S6. Sampling localities, abbreviations and genetic diversity indices at six microsatellite loci for *C. cantonensis*.

Sampling localities	Abb.		Locus						
			Mean	C8	C20	N9	N11	N15	O17
Total		<i>Na</i>		18	19	19	19	20	38
Fung Wong Wat	NFW	<i>n</i>	29.500	29	29	30	30	30	29
		<i>Na</i>	10.833	12	5	9	10	5	24
		<i>Ae</i>	7.122	8.806	1.972	6.498	6.360	1.933	17.163
		<i>Ho</i>	0.779	0.828	0.690	0.933	0.900	0.600	0.724
		<i>He</i>	0.749	0.886	0.493	0.846	0.843	0.483	0.942
		<i>P_{HW}</i>		0.626	0.626	0.984	0.913	0.855	0.145
		<i>F</i>	-0.086	0.066	-0.399	-0.103	-0.068	-0.243	0.231
Ha Miu Tin	NHMT	<i>n</i>	27.333	28	28	28	27	25	28
		<i>Na</i>	7.667	10	5	10	6	5	10
		<i>Ae</i>	3.677	5.426	1.960	6.730	4.378	1.453	2.113
		<i>Ho</i>	0.642	0.893	0.679	0.786	0.704	0.360	0.429
		<i>He</i>	0.628	0.816	0.490	0.851	0.772	0.312	0.527
		<i>P_{HW}</i>		0.870	0.689	0.851	0.343	1.000	0.000*
		<i>F</i>	-0.047	-0.095	-0.385	0.077	0.088	-0.154	0.186
Luk Keng	NLG	<i>n</i>	19.000	21	16	21	21	14	21
		<i>Na</i>	8.667	10	11	10	10	6	5
		<i>Ae</i>	5.374	6.891	4.785	6.891	5.880	3.843	3.955
		<i>Ho</i>	0.793	0.857	0.875	0.952	0.857	0.929	0.286
		<i>He</i>	0.803	0.855	0.791	0.855	0.830	0.740	0.747
		<i>P_{HW}</i>		0.446	0.888	0.078	0.000*	0.669	0.000*
		<i>F</i>	0.018	-0.003	-0.106	-0.114	-0.033	-0.255	0.618
Nam Chung Lo Uk	NLU	<i>n</i>	21.500	23	14	23	23	22	24
		<i>Na</i>	8.000	10	6	12	12	4	4
		<i>Ae</i>	4.909	6.187	2.925	8.015	8.602	1.328	2.395
		<i>Ho</i>	0.661	0.696	0.714	0.870	1.000	0.273	0.417
		<i>He</i>	0.681	0.838	0.658	0.875	0.884	0.247	0.582
		<i>P_{HW}</i>		0.893	0.941	0.276	0.863	0.997	0.061

Sampling localities	Abb.	Locus							
		Mean	C8	C20	N9	N11	N15	O17	
Total	<i>Na</i>		18	19	19	19	20	38	
	<i>F</i>	0.023	0.170	-0.085	0.006	-0.132	-0.105	0.285	
Wu Kau Tang	NWKT	<i>n</i>	22.167	23	21	21	23	22	23
	<i>Na</i>	6.500	10	3	6	5	6	9	
	<i>Ae</i>	3.512	6.335	1.340	3.600	1.938	2.077	5.781	
	<i>Ho</i>	0.498	0.652	0.286	0.238	0.435	0.682	0.696	
	<i>He</i>	0.608	0.842	0.254	0.722	0.484	0.519	0.827	
	<i>P_{HW}</i>		0.265	0.900	0.000*	0.087	0.934	0.000*	
	<i>F</i>	0.119	0.226	-0.125	0.670	0.102	-0.315	0.159	
Wilson Trail Section9	NWN	<i>n</i>	20.833	23	24	15	24	22	17
	<i>Na</i>	8.667	14	4	6	11	6	11	
	<i>Ae</i>	4.821	9.043	1.889	4.091	4.780	2.402	6.721	
	<i>Ho</i>	0.606	0.826	0.458	0.267	0.750	0.864	0.471	
	<i>He</i>	0.724	0.889	0.470	0.756	0.791	0.584	0.851	
	<i>P_{HW}</i>		0.255	0.096	0.000*	0.791	0.627	0.000*	
	<i>F</i>	0.127	0.071	0.026	0.647	0.052	-0.480	0.447	
Grand mean	<i>Na</i>	8.389							
	<i>Ae</i>	4.902							
	<i>Ho</i>	0.663							
	<i>He</i>	0.699							
	<i>F</i>	0.026							

Abb., abbreviation; *n*, sample size; *Na*, number of allele; *Ne*, number of effective alleles; *Ho*, observed heterozygosity; *He*, expected heterozygosity; *P_{HW}*, probability of deviation from HWE; *F*, fixation index. C8, C20, N9, N11 and N15 data are from Ma *et al.* (in review). *, loci which deviate from HWE after sequential Bonferroni correction ($\alpha = 0.05$).

Table S7. Detection of null alleles for microsatellite markers in each *C. logemanni* and *C. cantonensis* populations by MICRO-CHECKER.

Population	Locus					
	C8	C20	N9	N11	N15	O17
<i>C. logemanni</i>						
NWC	+					
NHS			+		+	
<i>C. cantonensis</i>						
NFW						+
NHMT						
NLG						+
NLU						
NWKT			+			+
NWN	+		+			

+, possible presence of null alleles.

Table S8. AMOVA results of *C. logemanni* and *C. cantonensis* based on (a) *COI*, (b) *NaK* intron, and (c) microsatellite data.

Source of variation	Variance components	Percentage of variation
(a)		
<i>C. logemanni</i>		
Among populations	9.8854	70.81*
Within populations	4.07496	29.19
<i>C. cantonensis</i>		
Among populations	1.28502	67.15*
Within population	0.62861	32.85
(b)		
<i>C. logemanni</i>		
Among populations	2.06063	69.53*
Within populations	0.903	30.47
<i>C. cantonensis</i>		
Among populations	2.02105	65.51*
Within populations	1.0641	34.49
(c)		
<i>C. logemanni</i>		
Among populations	0.24074	13.06*
Within populations	1.60205	86.94
<i>C. cantonensis</i>		
Among populations	0.26318	13.52*
Within populations	1.68313	86.48

*, $P < 0.001$.

Table S9. Results of one-tail Wilcoxon heterozygosity excess test (*P*-values) and mode-shift indicator test of *C. logemanni* and *C. cantonensis* populations.

Population	Wilcoxon test	
	SMM	TPM
<i>C. logemanni</i>		
NWC	0.57813	0.42188
NHS	0.99219	0.99219
<i>C. cantonensis</i>		
NFW	0.57813	0.57813
NHMT	0.97656	0.94531
NLG	0.57813	0.34375
NLU	0.94531	0.57813
NWKT	0.99219	0.92188
NWN	0.98438	0.96094

SMM, stepwise mutation model; TPM, two-phase model. *, population-mutational model pair with significant heterozygosity excess.

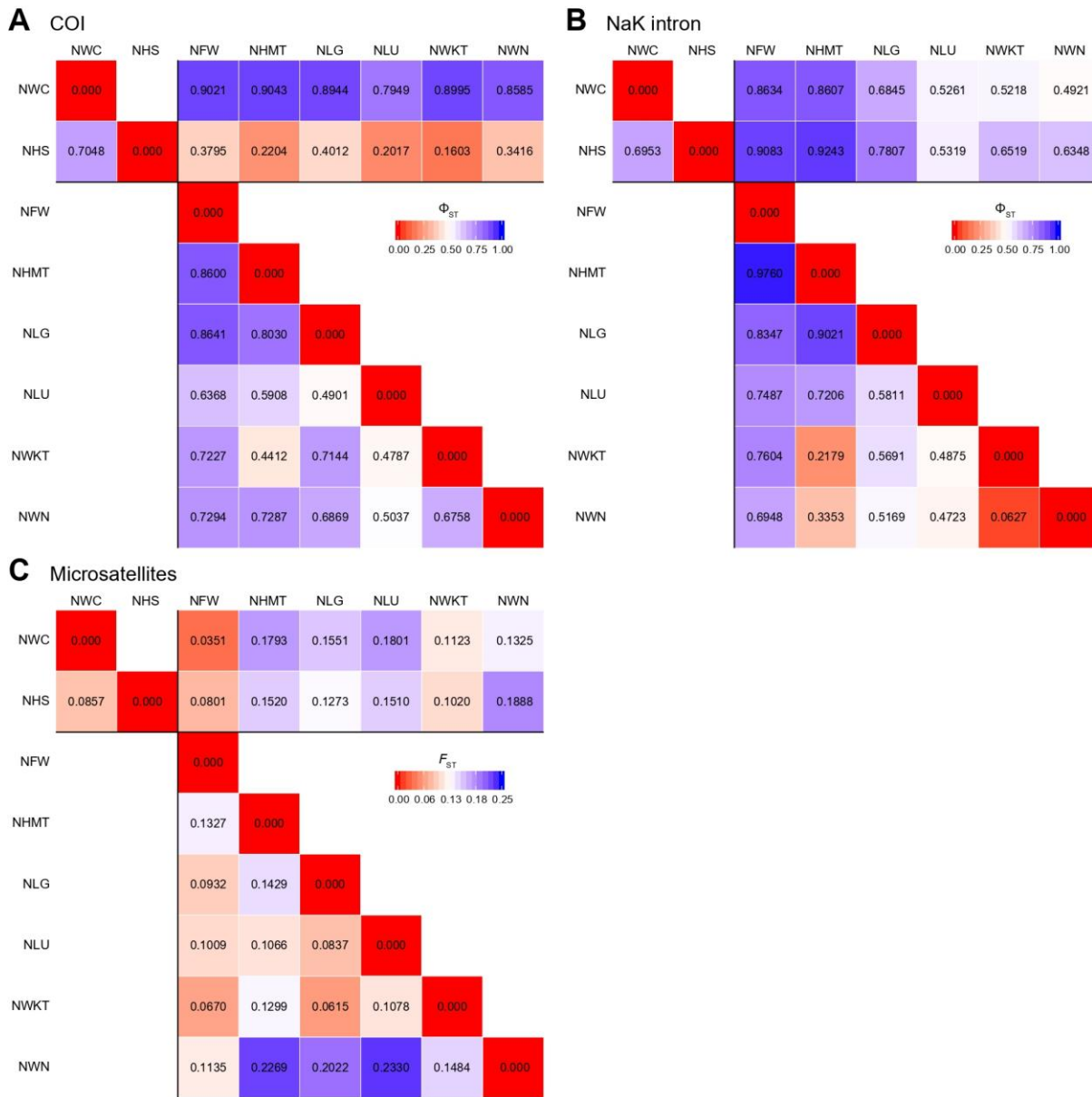


Figure S1. Heatmaps of (a) pairwise Φ_{ST} values based on *COI* and (b) *NaK* intron data, and (c) pairwise F_{ST} values based on microsatellite data, among *C. logemanni* (NWC, NHS) and *C. cantonensis* populations (NFW, NHMT, NLG, NLU, NWKT). Inter- and intra-specific comparisons are indicated above and below diagonal respectively. All values are significant after sequential Bonferroni corrections.

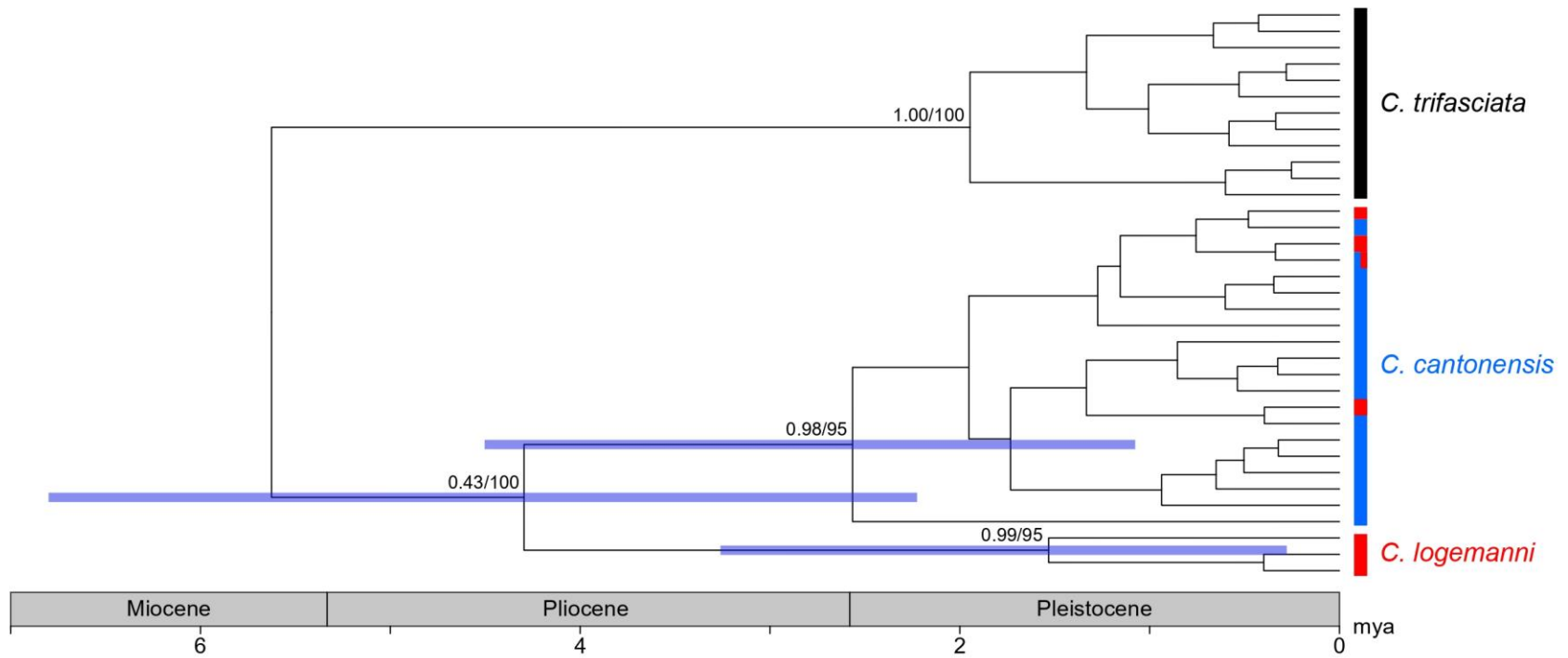


Figure S2. Bayesian phylogenetic tree of *C. logemanni* and *C. cantonensis* based on *COI* haplotypes. Branch support values are indicated as ‘posterior probability/bootstrap value’. Colours represent the species harbouring the haplotype: blue, *C. cantonensis*; red, *C. logemanni*; black, outgroup *C. trifasciata*. Blue bars at nodes represent the 95% HPD of the divergence time estimates.