Supporting Information

Hedtke et al. 10.1073/pnas.1106742108

SI Materials and Methods

Primer Development. To identify nuclear markers appropriate for phylogenetic analysis, we tested a candidate set of "universal" primers for single-copy loci (1). Primers were additionally developed using an α-amylase gene (*amy*) sequence from *Corbicula* (GenBank accession no. AF468016). Primer pairs were tested on DNA from North America using PCR annealing temperatures ranging from 44 °C to 56 °C; PCR products were then separated on an ethidium-bromide stained agarose gel. Primer pairs that produced consistent products were optimized by varying temperature and concentration of magnesium chloride and/or BSA. Successful amplifications were cleaned using a Viogene Gel-M gel extraction kit (Viogene Biotek) and sequenced using an Applied Biosystems 3100 automated sequencer. On the basis of initial sequences of products amplified using a primer pair for an intron of the α-subunit of adenosine triphosphate synthase (*atps-α*) (1), additional primers were designed and tested. From 5′ to 3′, the primer sequences for the markers used in this study are as follows: AmyE3c-f: ACA TGG TGC AAC GAT GC; AmyE4a-r: TGA TAA CCA CAT CTA CCA AG; ATPSaSH1f: GTG CCC ATY GGW AGA GGA CAG AGA G; and ATPSaSH4r: TGA TGG TGT CAA TAG CAA TGG CAG T.

Reaction conditions were as follows: 2.5 µL Thermopol 10× buffer with $MgCl_2$ (NEB), 2.5 µL 25 mM dNTPs, 0.75 µL each 10 mM primer, 1 µL BSA (*amy* only), 0.2 µL NEB Taq polymerase, and 3–5 µL DNA extract, brought to a total volume of 25 µL with double-distilled water. PCR conditions were as follows: 94 °C 1:30 min, 35 cycles 94 °C 1 min, 46 °C (*amy*)/52 °C (*atps-α*) 1 min, 72 °C 2 min, and 72 °C 5 min. To sequence individual alleles, we cloned PCR products using the Invitrogen TOPO TA Cloning kit with pCR 2.1-TOPO vector (Invitrogen Life Technologies) and sequenced the amplified product from individual clones.

Testing for Recombinant Sequences. During PCR, recombination can occur between alleles of an individual and confound phylogenetic analysis (2). We used the program RDPv3.26 (3) to test all sequences for possible PCR-mediated recombination within individual PCR products using the following implemented methods: RDP, GENECONV, Bootscan/Recscan, MaxChi, Chimaera, SiScan, and 3seq (4–10). In no case was recombination inferred between sequences found in an individual, suggesting that PCR-mediated recombination was absent or minimal in these data sets.

Phylogenetic Analyses. The model of evolution for alignments was determined under the Akaike Information Criterion (AIC) using ModelTest v3.7 (11) for maximum likelihood analyses, and MrModelTest v2.3 (12) for Bayesian analyses, using the correction for sample size (i.e., the AICc), with the number of bases in the alignment being the sample size.

We estimated the maximum likelihood estimate (MLE) for all trees using GARLI v0.96 (13), performing 20 search replicates for each alignment. We ran 1,000 nonparametric bootstrap replicates using GARLI under the same search settings as those used to determine the MLE. For Bayesian phylogenetic analysis, we performed four replicate runs with four chains each using MrBayes v.3.1.2 (14). We set the exponential rate parameter for the distribution of the prior probability on branch length to 0.01 rather than the default of 0.1, because the default branch length prior in MrBayes can lead to branch lengths 10 to >100 times larger than the MLE estimate if distances between sequences are expected to be small (15). Each run continued for 5,000,000 generations and was sampled every 1,000 trees and parameters. We assessed convergence using MrConverge (as described in ref. 16) and discarded all tree and parameters sample before convergence as burn-in. We used FigTree v1.1 (17) to visualize trees and create initial graphics files.

Hypothesis Testing. The posterior probability of alternative hypotheses was determined by filtering the postconvergence Bayesian posterior sample using backbone constraints in PAUP* v.4b (18). Because topologies are sampled in proportion to their posterior probability once stationarity has been reached, the proportion of trees within the post burn-in sample that matches a given constraint tree represents the probability of that topology. For each of the three genes, we tested the monophyly of alleles within each androgenetic individual, the monophyly of all alleles from all androgenetic taxa, support for individual bipartitions, and pairwise relationships between sexual taxa.

1. Jarman SN, Ward RD, Elliott NG (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Mar Biotechnol (NY)* 4:347–355.
2. Bradley RD, Hillis DM (1997) Recombinant DNA sequences generated by PCR amplification. *Mol Biol Evol* 14:592–593.
3. Martin DP, Williamson C, Posada D (2005) RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* 21:260–262.
4. Smith JM (1992) Analyzing the mosaic structure of genes. *J Mol Evol* 34:126–129.
5. Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology* 265:218–225.
6. Gibbs MJ, Armstrong JS, Gibbs AJ (2000) Sister-scanning: A Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16:573–582.
7. Martin D, Rybicki E (2000) RDP: Detection of recombination amongst aligned sequences. *Bioinformatics* 16:562–563.
8. Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc Natl Acad Sci USA* 98:13757–13762.
9. Martin DP, Posada D, Crandall KA, Williamson C (2005) A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* 21:98–102.
10. Boni MF, Posada D, Feldman MW (2007) An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 176:1035–1047.
11. Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
12. Nylander JAA (2004) *MrModelTest v2.* (Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden).
13. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence data sets under the maximum likelihood criterion. Ph.D. dissertation (University of Texas at Austin, TX).
14. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
15. Brown JM, Hedtke SM, Lemmon AR, Lemmon EM (2010) When trees grow too long: Investigating the causes of highly inaccurate bayesian branch-length estimates. *Syst Biol* 59:145–161.
16. Brown JM, Lemmon AR (2007) The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. *Syst Biol* 56:643–655.
17. Rambaut A (2006) *FigTree v1.2.2* (Institute of Evolutionary Biology, Univ of Edinburgh, UK).
18. Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4* (Sinauer Associates, Sunderland, MA).

Table S1. Posterior probabilities of monophyletic groups in three different gene trees of *Corbicula*

Bipartition or clade tested	cox-I	amy	atps-α
Monophyly of androgenetic individuals			
C. sp. A	n/a	0.14	0.53
C. sp. B	n/a	0.01	0.20
C. sp. C	n/a	0	0.0
C. fluminea Korea	n/a	0	0.14
C. fluminea Thailand	n/a	0	0.13
C. fluminea Taiwan	n/a	0.01	0.05
C. fluminea Philippines	n/a	n/a	0.53
Monophyly of all androgenetic individuals	0.03	0	0
Posterior probability of sexual taxa being sister			
(*C. loehensis, C. matannensis*)	0.97	0.99	0
(*C. loehensis, C. moltkiana*)	0.02	0	0.33
(*C. madagascariensis, C. moltkiana*)	0.71	0	n/a
(*C. matannensis, C. sandai*)	0.01	0	0.99
(*C. moltkiana, C. sandai*)	0.21	0	0
Posterior probability of sexual taxa being closest sexual relatives, removing androgenetic taxa			
(*C. loehensis, C. matannensis*)	0.97	0.99	0
(*C. loehensis, C. moltkiana*)	0.02	0	0.99
(*C. madagascariensis, C. moltkiana*)	0.76	0	n/a
(*C. matannensis, C. sandai*)	0.01	0	0.99
(*C. moltkiana, C. sandai*)	0.21	0.99	0

Table S2. Species identification, geographic location, and collection number for each individual *Corbicula* specimen, and number of clones sequenced and alleles inferred for the third intron of the α-amylase gene (*amy*) and a putative intron of the α subunit of adenosine triphosphate synthase (*atps-α*)

Species	Country of origin	Other names recently used or species	Collection number	No. of clones *amy*	No. of alleles *amy*	No. of clones *atps-α*	No. of alleles *atps-α*	Evidence for sexual/asexual
C. fluminea	Korea		UMMZ 266690	7	2	6	3	Lack of genetic variation (1)
C. fluminea	Taiwan		ZMB 170096a	7	2	5	3	Unreduced sperm (2)
C. fluminea	Thailand		UMMZ 266691	6	2	7	3	Species identification
C. fluminea	Philippines		ZMB 103026	—	—	8	2	Species identification
C. loehensis	Indonesia		ZMB 190582	6	2	5	1	Reproductive morphology (3)
			ZMB 190768	5	2	5	1	
C. madagascariensis	Madagascar	*C. africana* (4)	UMMZ 255293	4	2	—	—	Unknown
C. matannensis	Indonesia		ZMB 191042	4	1	4	1	Reproductive morphology (3)
C. moltkiana	Indonesia		ZMB 103024	3	2	3	1	Reproductive morphology (3)
C. sandai	Japan		UMMZ 266689	5	2	6	1	Reproductive morphology (5)
C. sp. C	Netherlands	Possibly form S (4), *C. fluminalis* (6)	fff2	6	2	7	2	Markers identical to *C. sp.* C (this study)
C. sp. A	Netherlands	Possibly form R (4), *C. fluminea* (6)	ggg1	7	2	8	1	Markers identical to *C. sp.* A (this study)
C. sp. A	United States	Form A (7), white (8), *C. fluminea* (9), *C. cf leana* (10)	xx11	—	—	8	3	Reproductive morphology (7), lack of genetic variation (7, 8, 10)
			qq1	8	2	—	—	
C. sp. B	United States	Purple (8), form B (7), *C. cf fluminea* (10), *C. fluminea* (9)	rr1	3	2	—	—	Reproductive morphology (7), lack of genetic variation (7, 8, 10)
			yy12	7	2	6	3	
C. sp. C	Argentina		U3	—	—	3	2	Sperm morphology (7)
			U4	5	2	—	—	
C. tobiae	Indonesia		ZMB 103027	4	1	6	1	Unknown

1. Park JK, Kim W (2003) Two *Corbicula* (Corbiculidae: Bivalvia) mitochondrial lineages are widely distributed in Asian freshwater environment. *Mol Phylogenet Evol* 29:529–539.
2. Komaru A, Konishi K (1999) Non-reductional spermatozoa in three shell color types of the freshwater clam *Corbicula fluminea* in Taiwan. *Zoolog Sci* 16:105–108.
3. Glaubrecht M, von Rintelen T, Korniushin AV (2003) Toward a systematic revision of brooding freshwater *Corbicula* in southeast Asia (Bivalvia, Veneroida): On shell morphology, anatomy, and molecular phylogenetics of endemic taxa from islands in Indonesia. *Malacologia* 45:1–40.
4. Pfenninger M, Reinhardt F, Streit B (2002) Evidence for cryptic hybridization between different evolutionary lineages of the invasive clam genus *Corbicula* (Veneroida, Bivalvia). *J Evol Biol* 15:818–829.
5. Hachiri S, Higashi S (1970) Spermiogenesis in the freshwater mussel *Corbicula sandai*. *J Educ Dept Shiga Univ Nat Sci* 20:35–39.
6. Renard E, Bachmann V, Cariou ML, Moreteau JC (2000) Morphological and molecular differentiation of invasive freshwater species of the genus *Corbicula* (Bivalvia, corbiculidea) suggest the presence of three taxa in French rivers. *Mol Ecol* 9:2009–2016.
7. Lee T, Siripattrawan S, Ituarte CF, Ófoighil D (2005) Invasion of the clonal clams: *Corbicula* lineages in the New World. *Am Malacol Bull* 20:113–122.
8. Hillis DM, Patton JC (1982) Morphological and electrophoretic evidence for two species of *Corbicula* (Bivalvia: Corbiculidae) in North America. *Am Midl Nat* 108:74–80.
9. Britton JC, Morton B (1979) *Corbicula* in North America: The evidence reviewed and evaluated. *Proceedings of the First International Corbicula Symposium*, ed Britton JC (Texas Christian Univ Research Foundation, Forth Worth, TX), pp 247–287.
10. Hedtke SM, Stanger-Hall K, Baker RJ, Hillis DM (2008) All-male asexuality: Origin and maintenance of androgenesis in the Asian clam Corbicula. *Evolution* 62:1119–1136.