data

MDPI

Data Descriptor

Novel Molecular Resources to Facilitate Future Genetics Research on Freshwater Mussels (Bivalvia: Unionidae)

Nathan A. Johnson [1,*] and Chase H. Smith [2,3,*]

1 Wetland and Aquatic Research Center, U.S. Geological Survey, 7920 NW 71st Street, Gainesville, FL 32653, USA
2 Biology Department, Baylor University, Waco, TX 76798, USA
3 Department of Integrative Biology, University of Texas at Austin, 2415 Speedway, Austin, TX 78712, USA
* Correspondence: najohnson@usgs.gov (N.A.J.); chase.smith@austin.utexas.edu (C.H.S.)

Received: 8 July 2020; Accepted: 28 July 2020; Published: 30 July 2020

Abstract: Molecular data have been an integral tool in the resolution of the evolutionary relationships and systematics of freshwater mussels, despite the limited number of nuclear markers available for Sanger sequencing. To facilitate future studies, we evaluated the phylogenetic informativeness of loci from the recently published anchored hybrid enrichment (AHE) probe set Unioverse and developed novel Sanger primer sets to amplify two protein-coding nuclear loci with high net phylogenetic informativeness scores: *fem-1 homolog C* (FEM1) and *UbiA prenyltransferase domain-containing protein 1* (UbiA). We report the methods used for marker development, along with the primer sequences and optimized PCR and thermal cycling conditions. To demonstrate the utility of these markers, we provide haplotype networks, DNA alignments, and summary statistics regarding the sequence variation for the two protein-coding nuclear loci (FEM1 and UbiA). Additionally, we compare the DNA sequence variation of FEM1 and UbiA to three loci commonly used in freshwater mussel genetic studies: the mitochondrial genes *cytochrome c oxidase subunit 1* (CO1) and *NADH dehydrogenase subunit 1* (ND1), and the nuclear *internal transcribed spacer 1* (ITS1). All five loci distinguish among the three focal species (*Potamilus fragilis*, *Potamilus inflatus*, and *Potamilus purpuratus*), and the sequence variation was highest for ND1, followed by CO1, ITS1, UbiA, and FEM1, respectively. The newly developed Sanger PCR primers and methodologies for extracting additional loci from AHE probe sets have great potential to facilitate molecular investigations targeting supraspecific relationships in freshwater mussels, but may be of limited utility at shallow taxonomic scales.

Dataset: https://doi.org/10.5066/P9Q3CFL5.

Keywords: anchored hybrid enrichment; conservation genetics; endangered species; mitochondrial DNA; nuclear DNA; primer design; Potamilus; Sanger sequencing

1. Summary

Recent studies utilizing molecular tools have been integral in resolving the evolutionary history of freshwater mussels (Bivalvia: Unionida), despite a heavy reliance on the Sanger sequencing of mitochondrial DNA (mtDNA). Nearly all phylogenetic studies on freshwater mussels have largely relied on the commonly used marker *internal transcribed spacer 1* (ITS1) to incorporate inference from the nuclear genome [1–6], which was first used in freshwater mussels nearly 20 years ago [7]. Even though researchers continue to utilize this locus, the number of studies reporting issues related to excessive heterozygosity, primarily due to length polymorphisms, continues to increase [1,2,6,8–14]. Other nuclear loci, such as *histone H3* and *28S*, have been utilized in freshwater mussel phylogenetic studies; however, these markers are well known to show limited diversity at shallow taxonomic scales and have primarily been used to resolve deep level phylogeny [15–20].

In recent years, the decreasing costs of next-generation sequencing platforms have significantly increased the ability to generate molecular supermatrices in non-model taxa [21,22], including freshwater mussels [23,24]. In particular, the recently developed anchored hybrid enrichment (AHE) probe set Unioverse [23] has drastically improved the ability to resolve phylogeny in freshwater mussels. The Unioverse probe set consists of 811 protein-coding loci derived from genomic and transcriptomic resources across Bivalvia that can be captured across all freshwater mussels to resolve phylogenetic relationships. Despite the decreasing costs of next-generation sequencing, the utilization of AHE probe sets can be cost-prohibitive for small-scale projects or molecular investigations that incorporate hundreds of individuals to investigate intra- or interspecific relationships. However, AHE probe sets offer opportunities for the development of primers for the amplification and Sanger sequencing of select protein-coding loci that can be used for small-scale projects.

Here, we evaluated the phylogenetic informativeness of loci in the Universe probe set and report the development of novel primer pairs for the amplification of two protein-coding nuclear genes *fem-1 homolog C* (FEM1) and *UbiA prenyltransferase domain-containing protein 1* (UbiA). To demonstrate the utility of these markers and facilitate their use in future studies, we provide the PCR primer sequences, optimized PCR conditions and thermal cycling parameters, haplotype networks, DNA alignments, and summary statistics regarding sequence variation for the two protein-coding nuclear loci (FEM1 and UbiA) and three loci that are commonly used in studies in freshwater mussels: the mitochondrial genes *cytochrome c oxidase subunit 1* (CO1) and *NADH dehydrogenase subunit 1* (ND1), and the nuclear ITS1 locus. All five loci distinguish among the three focal species (*Potamilus fragilis*, *Potamilus inflatus*, and *Potamilus purpuratus*) and should be amplifiable across the subfamily Ambleminae. The observed sequence variation was highest for ND1, followed by CO1, ITS1, UbiA, and FEM1, respectively (Figure 1). We also provide the detailed methodology used in the marker selection to expedite the identification of additional candidate loci and primer development from available AHE data. The newly developed Sanger PCR primers and methodologies for extracting additional loci have great potential to facilitate molecular investigations targeting supraspecific relationships in freshwater mussels, but may be of limited utility at shallow taxonomic scales.

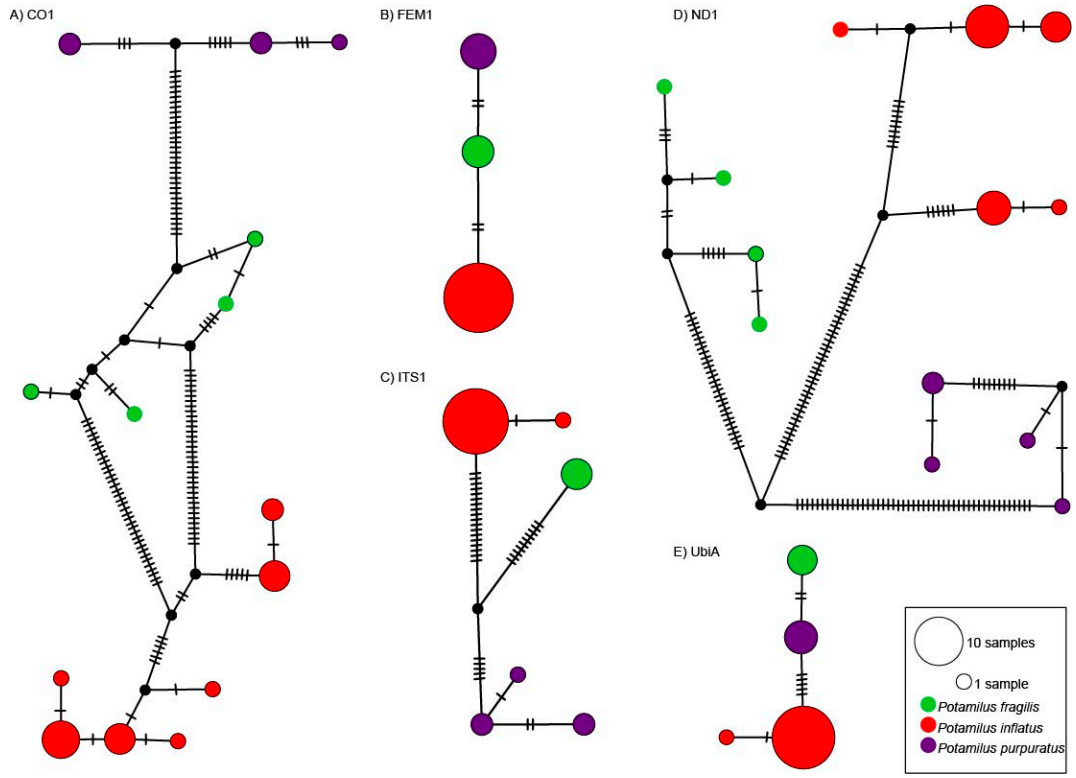

Figure 1. TCS haplotype networks for (**A**) cytochrome c oxidase subunit 1 (CO1), (**B**) fem-1 homolog C (FEM1), (**C**) internal transcribed spacer 1 (ITS1), (**D**) NADH dehydrogenase subunit 1 (ND1), and (**E**) UbiA prenyltransferase domain-containing protein 1 (UbiA). Each colored circle represents a unique haplotype, the colors correspond to individual species, the black circles represent unsampled haplotypes, and the hash marks indicate nucleotide differences between haplotypes.

2. Data Description

2.1. Specimen Details

All the metadata related to the specimens used in this study, including the collection location, GPS coordinates, and museum catalog numbers, are provided (https://doi.org/10.5066/P9Q3CFL5) [25].

2.2. Molecular Data

We present the DNA sequence data from five markers: the mitochondrial genes CO1 and ND1, the nuclear non-coding marker ITS1, and the protein-coding nuclear genes FEM1 and UbiA. Our five-locus DNA alignment consisted of 3368 bp of mitochondrial and nuclear sequence data (CO1 = 657 bp; ND1 = 900 bp; FEM1 = 501 bp; UBiA = 765 bp; ITS1 = 545 bp). The number of loci sequenced for each individual varies from two to five loci, with all loci available for 28 individuals (Table 1). The specific sample sizes for each locus are as follows: CO1 (n = 102); ND1 (n = 103); FEM1 (n = 29); UBiA (n = 29); and ITS1 (n = 31). A subset of individuals was chosen for the additional nDNA loci due to the high prevalence of multiple copies at ITS1 and low genetic diversity at FEM1 and UbiA. All the DNA alignment files are available in Phylip format (.phy), with the first line indicating the number of taxa and number of nucleotides and subsequent lines containing a taxon identifier, catalog number, and GenBank Accession number in the first column (each separated by underscore), and the DNA sequence in the second column. The file names are as follows: CO1.phy; ND1.phy; FEM1.phy; UbiA.phy; ITS1.phy; and 5_locus.phy (https://doi.org/10.5066/P9Q3CFL5) [25].

Table 1. Collection information and GenBank or SRA accession numbers for all the specimens and loci analyzed in this study. Museum abbreviations are as follows: UA—Alabama Museum of Natural History; UF—Florida Museum.

Taxon	ID	Drainage	Source	CO1	ND1	ITS1	FEM1	UbiA
Potamilus fragilis	LfraAla001	Mobile	UF438237	MT662019	MT669665	MT661766	MT669798	MT669771
Potamilus fragilis	LfraAmi040	Pontchartrain	UF439330	MT662020	MT669666	MT661773	MT669778	MT669751
Potamilus fragilis	LfraAmi041	Pontchartrain	UF439352	MT662021	MT669667			
Potamilus fragilis	LfraAmi042	Pontchartrain	UF439352	MT662022	MT669668			
Potamilus fragilis	LfraPrl043	Pearl	UF439332	MT662023	MT669669			
Potamilus fragilis	LfraPrl044	Pearl	UF439332	MT662024	MT669670	MT661780	MT669785	MT669758
Potamilus fragilis	LfraPrl045	Pearl	UF439365	MT662025	MT669671			
Potamilus fragilis	LfraPrl046	Pearl	UF439343	MT662026	MT669672			
Potamilus fragilis	LfraPrl047	Pearl	UF439343	MT662027	MT669673			
Potamilus fragilis	LfraPrl048	Pearl	UF439343	MT662028	MT669674			
Potamilus fragilis	LfraAmi057	Pontchartrain	UF439529	MT662029	MT669675			
Potamilus fragilis	LfraAmi058	Pontchartrain	UF439529	MT662030	MT669676			
Potamilus fragilis	LfraAmi059	Pontchartrain	UF439529	MT662031	MT669677			
Potamilus fragilis	LfraMob063	Mobile	UF439528	MT662033	MT669679			
Potamilus fragilis	LfraMob064	Mobile	UF439528	MT662032	MT669678			
Potamilus fragilis	LfraMob065	Mobile	Uncataloged	MT662034	MT669680	MT661792	MT669797	MT669770
Potamilus inflatus	PinfMob001	Mobile	UF439131	MT662002	MT669647	MT661768	MT669773	MT669746
Potamilus inflatus	PinfMob002	Mobile	UF439131	MK044952	MK045103	MK036203	MT669774	MT669747
Potamilus inflatus	PinfMob003	Mobile	UF439131	MT662003	MT669648	MT661769	MT669775	MT669748
Potamilus inflatus	PinfMob004	Mobile	UF439131	MK044953	MK045104	MK036204	SRR10579071	SRR10579071
Potamilus inflatus	PinfMob005	Mobile	UF439131	MT662004	MT669649	MT661770	MT669776	MT669749
Potamilus inflatus	PinfMob006	Mobile	UF439131	MT662005	MT669650	MT661771	MT669777	MT669750
Potamilus inflatus	PinfAmi010	Pontchartrain	UF439530	MT662006	MT669651	MT661774	MT669779	MT669752
Potamilus inflatus	PinfAmi011	Pontchartrain	UF439530	MT662007	MT669652	MT661775	MT669780	MT669753
Potamilus inflatus	PinfAmi012	Pontchartrain	UF439531	MT662008	MT669653	MT661776	MT669781	MT669754
Potamilus inflatus	PinfAmi013	Pontchartrain	UF439532	MT662009	MT669654	MT661777	MT669782	MT669755
Potamilus inflatus	PinfAmi014	Pontchartrain	UF439532	MT662010	MT669655	MT661778	MT669783	MT669756
Potamilus inflatus	PinfAmi015	Pontchartrain	UF439533	MT662011	MT669656	MT661779	MT669784	MT669757
Potamilus inflatus	PinfMob019	Mobile	UF439514	MT662012	MT669657	MT661783	MT669788	MT669761
Potamilus inflatus	PinfMob020	Mobile	UF439514	MT662013	MT669658	MT661784	MT669789	MT669762
Potamilus inflatus	PinfMob021	Mobile	UF439514	MT662014	MT669659	MT661785	MT669790	MT669763
Potamilus inflatus	PinfMob022	Mobile	UF439514	MT662015	MT669660	MT661786	MT669791	MT669764
Potamilus inflatus	PinfMob023	Mobile	UF439514	MT662016	MT669661	MT661787	MT669792	MT669765

Table 1. *Cont.*

Taxon	ID	Drainage	Source	CO1	ND1	ITS1	FEM1	UbiA
Potamilus inflatus	PinfMob017	Mobile	UF439513	MT662017	MT669662	MT661788	MT669793	MT669766
Potamilus inflatus	PinfMob018	Mobile	UF439513	MT662018	MT669663	MT661789	MT669794	MT669767
Potamilus inflatus	PinfMob016	Mobile	UA2696		MT669664	MT661781	MT669786	MT669759
Potamilus purpuratus	PpurPas001	Pascagoula	UF438434	MT662035	MT669681			
Potamilus purpuratus	PpurPrl022	Pearl	UF439145	MT662036	MT669682			
Potamilus purpuratus	PpurPrl023	Pearl	UF439145	MK044960	MK045111	MK036211	MT669799	MT669772
Potamilus purpuratus	PpurPrl024	Pearl	UF439145	MK044961	MK045112	MK036212		
Potamilus purpuratus	PpurPrl025	Pearl	UF439145	MT662037	MT669683			
Potamilus purpuratus	PpurPrl026	Pearl	UF439145	MT662038	MT669684	MT661767		
Potamilus purpuratus	PpurAmi038	Pontchartrain	UF439452	MT662039	MT669685			
Potamilus purpuratus	PpurAmi039	Pontchartrain	UF439452	MT662040	MT669686			
Potamilus purpuratus	PpurAmi040	Pontchartrain	UF439452	MT662041	MT669687			
Potamilus purpuratus	PpurAmi041	Pontchartrain	UF439452	MT662042	MT669688			
Potamilus purpuratus	PpurAmi042	Pontchartrain	UF439452	MT662043	MT669689			
Potamilus purpuratus	PpurAmi043	Pontchartrain	UF439453	MT662044	MT669690			
Potamilus purpuratus	PpurAmi044	Pontchartrain	UF439453	MT662045	MT669691			
Potamilus purpuratus	PpurAmi045	Pontchartrain	UF439453	MT662046	MT669692	MT661772	SRR10579081	SRR10579081
Potamilus purpuratus	PpurAmi046	Pontchartrain	UF439453	MT662047	MT669693			
Potamilus purpuratus	PpurAmi047	Pontchartrain	UF439453	MT662048	MT669694			
Potamilus purpuratus	PpurAmi048	Pontchartrain	UF439454	MT662049	MT669695			
Potamilus purpuratus	PpurAmi049	Pontchartrain	UF439454	MT662050	MT669696			
Potamilus purpuratus	PpurAmi050	Pontchartrain	UF439454	MT662051	MT669697			
Potamilus purpuratus	PpurAmi051	Pontchartrain	UF439454	MT662052	MT669698			
Potamilus purpuratus	PpurPrl052	Pearl	UF439456	MT662053	MT669699			
Potamilus purpuratus	PpurPrl053	Pearl	UF439456	MT662054	MT669700			
Potamilus purpuratus	PpurPrl054	Pearl	UF439457	MT662055	MT669701			
Potamilus purpuratus	PpurPrl055	Pearl	UF439457	MT662056	MT669702			
Potamilus purpuratus	PpurPrl056	Pearl	UF439457	MT662057	MT669703			
Potamilus purpuratus	PpurPrl057	Pearl	UF439457	MT662058	MT669704			
Potamilus purpuratus	PpurPrl058	Pearl	UF439457	MT662059	MT669705			
Potamilus purpuratus	PpurPrl059	Pearl	UF439456	MT662060	MT669706			
Potamilus purpuratus	PpurPrl060	Pearl	UF439456	MT662061	MT669707			
Potamilus purpuratus	PpurPrl061	Pearl	UF439456	MT662062	MT669708			
Potamilus purpuratus	PpurPrl062	Pearl	UF439456	MT662063	MT669709			

Table 1. *Cont.*

Taxon	ID	Drainage	Source	CO1	ND1	ITS1	FEM1	UbiA
Potamilus purpuratus	PpurPrl063	Pearl	UF439456	MT662064	MT669710			
Potamilus purpuratus	PpurPrl064	Pearl	UF439458	MT662065	MT669711			
Potamilus purpuratus	PpurPrl065	Pearl	UF439459	MT662066	MT669712			
Potamilus purpuratus	PpurPrl066	Pearl	UF439459	MT662067	MT669713			
Potamilus purpuratus	PpurPrl067	Pearl	UF439459	MT662068	MT669714			
Potamilus purpuratus	PpurPrl068	Pearl	UF439459	MT662069	MT669715			
Potamilus purpuratus	PpurPrl069	Pearl	UF439459	MT662070	MT669716			
Potamilus purpuratus	PpurMob081	Mobile	UA62	MT662071	MT669717			
Potamilus purpuratus	PpurMob082	Mobile	UA2469	MT662072	MT669718			
Potamilus purpuratus	PpurMob083	Mobile	UA2510	MT662073	MT669719			
Potamilus purpuratus	PpurMob084	Mobile	UA2562	MT662074	MT669720			
Potamilus purpuratus	PpurMob085	Mobile	UA2740	MT662075	MT669721	MT661782	MT669787	MT669760
Potamilus purpuratus	PpurMob086	Mobile	UA3100	MT662076	MT669722			
Potamilus purpuratus	PpurMob087	Mobile	UA3123	MT662077	MT669723			
Potamilus purpuratus	PpurMob088	Mobile	UA3205	MT662078	MT669724			
Potamilus purpuratus	PpurMob089	Mobile	UA3417	MT662079	MT669725			
Potamilus purpuratus	PpurMob090	Mobile	UA3482	MT662080	MT669726			
Potamilus purpuratus	PpurPas097	Pascagoula	UF439510	MT662081	MT669727			
Potamilus purpuratus	PpurPas098	Pascagoula	UF439510	MT662082	MT669728			
Potamilus purpuratus	PpurPas099	Pascagoula	UF439510	MT662083	MT669729			
Potamilus purpuratus	PpurPas100	Pascagoula	UF439510	MT662084	MT669730			
Potamilus purpuratus	PpurPas101	Pascagoula	UF439510	MT662085	MT669731	MT661790	MT669795	MT669768
Potamilus purpuratus	PpurPas102	Pascagoula	UF439510	MT662086	MT669732			
Potamilus purpuratus	PpurPas103	Pascagoula	UF439510	MT662087	MT669733			
Potamilus purpuratus	PpurMob107	Mobile	UF439527	MT662088	MT669734	MT661791	MT669796	MT669769
Potamilus purpuratus	PpurMob108	Mobile	UF439527	MT662089	MT669735			
Potamilus purpuratus	PpurMob109	Mobile	UF439527	MT662090	MT669736			
Potamilus purpuratus	PpurMob110	Mobile	UF439527	MT662091	MT669737			
Potamilus purpuratus	PpurMob111	Mobile	UF439527	MT662092	MT669738			
Potamilus purpuratus	PpurMob112	Mobile	UF439527	MT662093	MT669739			
Potamilus purpuratus	PpurMob113	Mobile	UF439527	MT662094	MT669740			
Potamilus purpuratus	PpurMob114	Mobile	UF439527	MT662095	MT669741			
Potamilus purpuratus	PpurMob115	Mobile	UF439527	MT662096	MT669742			
Potamilus purpuratus	PpurMob116	Mobile	UF439527	MT662097	MT669743			
Potamilus purpuratus	PpurMob117	Mobile	UF439527	MT662098	MT669744			
Potamilus purpuratus	PpurMob118	Mobile	UF439527	MT662099	MT669745			

3. Methods

3.1. Taxon Sampling and DNA Extraction

We present molecular data on 103 specimens representing *Potamilus fragilis* (n = 22), *Potamilus inflatus* (n = 14), and *Potamilus purpuratus* (n = 67) used in Smith and Johnson [26] (Table 1). All specimens were collected from four Gulf of Mexico river drainages in the southeastern United States: Mobile, Pascagoula, Pearl, and Pontchartrain. Genomic DNA was extracted from mantle tissue clips from vouchered individuals using the Qiagen PureGene DNA extraction kit with the standard extraction protocol (Qiagen, Hilden, Germany).

3.2. Novel Primer Design and Gene Annotation

We compiled data from a recent study [24] utilizing the AHE probe set Unioverse to develop novel primer sets for amplifying protein-coding nuclear loci for use in the freshwater mussel genus *Potamilus*. To screen for loci in the dataset that were informative at shallow phylogenetic scales, we measured the net phylogenetic informativeness (PI) using an arbitrary time scale [27]. This methodology has been used in previous studies to calculate the power of individual loci in AHE datasets [28,29]. First, we reconstructed a phylogeny from a concatenated alignment of probe loci using IQ-TREE v 1.6.11 [30,31], and the consensus tree was arbitrarily dated with a molecular clock (i.e., tips = time 0, root = time 1) using the program PATHd8 [32]. A concatenated alignment partitioned by the probe and the ultrametric tree from PATHd8 were uploaded into the web application PhyDesign [33] (http://phydesign.townsend.yale.edu/) to estimate the PI using the HyPhy substitution rates algorithm with the GTR model of nucleotide evolution and empirical base frequencies [34]. We used the R script PhyDesign.r [29] to identify specific nucleotide positions in the alignment with unusually high substitution rates that could be contributing phylogenetic noise. Nucleotide positions with rate values higher than five were removed from the alignment manually and the filtered matrices were re-uploaded to PhyDesign as above for a final analysis.

Three nucleotides were removed from the dataset due to unusually high substitution rates (rate value > 5 = "phantom spikes"). In the filtered dataset, the probe regions with a 100% capture efficiency across Ambleminae had an average net PI of 4.62 and ranged from 0.32 to 23.09 (Table 2). Using the results from PhyDesign, we selected two candidate loci for primer development and PCR validation: locus 156 and locus 412. Locus 156 and locus 412 exhibited a 100% capture efficiency in our dataset, had suitable candidate primers that could be cross amplified across *Potamilus*, displayed high levels of average PI (9.22 and 11.61, respectively), and were able to discriminate our focal species. We were unable to develop compatible primers for the other candidate loci with high net PI scores (e.g., locus 70 and locus 413).

We used BLASTX [35] to annotate the gene and protein names for our candidate loci [36]. Briefly, the probe region sequences of both loci for *P. inflatus* were searched against the non-redundant protein database using BLASTX, which returned 172 and 118 BLAST hits for locus 156 and locus 412, respectively. Locus 156 was identified as *UbiA prenyltransferase domain-containing protein 1*, and the highest homology was to genes in the marine bivalves *Crassostrea virginica* (74.62%) and *C. gigas* (72.31%). Locus 412 was identified as a *fem-1 homolog*, and the highest homology was to genes in the unionid bivalves *Hyriopsis schlegelii* (99.44%) and *H. cumingii* (98.89%), and the marine bivalves *Mizuhopecten yessoensis* (87.22%), *Pecten maximus* (87.22%), *C. virginica* (86.11%), and *C. gigas* (86.11%). There were inconsistencies regarding whether the region was a *fem-1 homolog A* or *fem-1 homolog C*. All the blast hits except for *H. cumingii* and *H. schlegelii* indicated the sequence was representative of *fem-1 homolog C;* therefore, we annotated the locus as *fem-1 homolog C*.

Table 2. Average, minimum, and maximum net phylogenetic informativeness (PI), and the time at maximum PI for the 55 loci with a 100% capture efficiency across Ambleminae. Time at max PI represents an arbitrary time scale, with values closer to zero providing the maximum phylogenetic signal at shallower taxonomic scales. Loci are ordered from the highest to lowest average net PI.

Locus	Average Net PI	Min PI	Max PI	Time at Max
L413	23.09598	1.103651	30.20921	0.99
L70	15.90158	1.681899	18.11504	0.63
L412	11.60713	0.527607	15.10124	0.93
L162	9.295725	0.444904	11.765	0.82
L156	9.220128	0.430839	12.28036	0.99
L369	8.041972	0.287906	11.83626	0.99
L573	7.960265	0.326746	10.91673	0.99
L747	7.601916	0.306527	10.45607	0.99
L97	6.869637	0.30823	8.97996	0.94
L564	6.554754	0.390707	7.890758	0.73
L370	6.372669	0.229417	9.187468	0.99
L593	6.171768	0.317327	7.673508	0.78
L113	5.689908	0.30191	7.02746	0.77
L108	5.613172	0.290217	7.103136	0.91
L776	5.583852	0.185192	8.420968	0.99
L426	5.407149	0.224737	7.267066	0.99
L745	5.263546	0.171265	8.105492	0.99
L34	5.071216	0.375813	5.909839	0.53
L541	4.995603	0.222499	6.497481	0.9
L16	4.822584	0.198933	6.458898	0.96
L222	4.821583	0.175282	6.977383	0.99
L184	4.815454	0.16821	7.114062	0.99
L663	4.754253	0.279931	5.731063	0.62
L319	4.509462	0.347447	5.428627	0.99
L381	4.260574	0.382234	5.004245	0.42
L486	3.80863	0.146694	5.4417	0.99
L636	3.676559	0.126859	5.45936	0.99
L19	3.654261	0.18187	4.57087	0.8
L667	3.442098	0.152375	4.484831	0.9
L697	3.300729	0.302709	3.804831	0.43
L155	3.250609	0.116582	4.730085	0.99
L420	2.901364	0.133287	3.700777	0.82
L772	2.867933	0.098412	4.267794	0.99
L576	2.704247	0.22991	3.072973	0.48
L464	2.661726	0.092384	3.933712	0.99
L543	2.647827	0.106438	3.637347	0.99
L210	2.636395	0.086601	4.013408	0.99
L602	2.564771	0.089917	3.799166	0.99
L327	2.426569	0.31026	3.106579	0.29
L586	2.333785	0.085107	3.364878	0.99
L539	2.309995	0.11608	2.882391	0.79
L488	2.270353	0.08539	3.247081	0.99
L728	2.24181	0.071619	3.456425	0.99
L240	2.131098	0.07367	3.139256	0.99
L568	2.10264	0.079918	2.940695	0.99
L274	1.981736	0.065524	3.011985	0.99
L358	1.872561	0.059073	2.900606	0.99
L600	1.735105	0.058428	2.616461	0.99
L212	1.659549	0.063007	2.330338	0.99
L686	1.643869	0.056209	2.461919	0.99
L237	1.508957	0.049352	2.299572	0.99
L729	1.368267	0.043306	2.118198	0.99
L183	1.142517	0.036308	1.75823	0.99
L544	0.873877	0.02454	1.445329	0.99
L510	0.317174	0.009299	0.511776	0.99

3.3. PCR and Sequencing

PCRs were conducted using a 25 µL mixture of the following: molecular grade water (9.5 µL), MyTaq^{TM} Red Mix (12.5 µL; Bioline, London, UK), primers (1.0 µL each), and DNA template (100 ng). The primers for all loci and thermal cycling conditions for CO1, ND1, and ITS1 are reported in Table 3. The thermal cycling conditions for FEM1 and UbiA were as follows: an initial denaturation at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 51/60 °C (FEM1/UbiA) for 30 s, and 72 °C for 90 s. The products were sent to Molecular Cloning Laboratories (McLAB, South San Francisco, CA, USA) for bi-directional sequencing on an ABI 3730. Geneious v 10.2.3 [37] was used to assemble the contigs and edit chromatograms, and the sequences were aligned in Mesquite v 3.61 [38] using MAFFT v 7.311 [39]. The loci were aligned independently using the L-INS-i method in MAFFT and translated into amino acids to ensure the absence of stop codons and gaps.

Table 3. PCR primer sequences and cycling conditions used in this study.

Locus	Primers	Source	Conditions
CO1	F: 5′-GTTCCACAAATCATAAGGATATTGG-3′ R: 5′-TACACCTCAGGGTGACCAAAAAACCA-3′	Campbell et al. (2005) [40]	Johnson et al. 2018 [22]
ND1	F: 5′-TGGCAGAAAAGTGCATCAGATTAAAGC-3′ R: 5′-CCTGCTTGGAAGGCAAGTGTACT-3′	Serb et al. (2003) [41]	Serb et al. 2003 [41]
ITS1	F: 5′-AAAAAGCTTCCGTAGGTGAACCTGCG-3′ R: 5′-AGCTTGCTGCGTTCTTCATCG-3′	King et al. (1999) [7]	King et al. 1999 [7]
FEM1	F: 5′- GTRATGGAGTATCGCAGTGT-3′ R: 5′-ACRCTYTTCCTGTTAACATC-3′	This study	This study
UbiA	F: 5′- TTTACTCCTGTTGCACTTGGGA-3′ R: 5′-AGCATCTGTCATGAAGACTCCAAC-3′	This study	This study

3.4. Sequence Variation and Haplotype Analysis

We created haplotype networks (Figure 1) and calculated the nucleotide diversity, number of haplotypes, number of segregating sites, and number of parsimony-informative sites (Table 4) to compare the amounts of sequence variation across all five loci used in this study. The TCS haplotype networks and sequence variation statistics were calculated using PopART 1.7 [42]. All five loci distinguish among the three focal species (Figure 1). The sequence variation was highest for ND1, followed by CO1, ITS1, UbiA, and FEM1, respectively (Table 4). Despite selecting loci from the AHE probe set with a high net PI, the level of sequence variation remains low when compared to mtDNA and ITS1, suggesting the limited utility of the probes at intraspecific levels.

Table 4. Summary of diversity indices based on all five loci utilized in this study. Abbreviations and symbols are as follows: *cytochrome c oxidase subunit 1* (CO1); *fem-1 homolog C* (FEM1); *internal transcribed spacer 1* (ITS1); *NADH dehydrogenase subunit 1* (ND1); *UbiA prenyltransferase domain-containing protein 1* (UbiA); sample size (*n*); nucleotide diversity (π); number of haplotypes (nh); number of segregating sites (S); and number of parsimony-informative sites (P).

Locus	*n*	π	nh	S	P
CO1	28	0.047864	14	82	78
FEM1	28	0.003431	3	4	4
ITS1	28	0.039169	6	333	31
ND1	28	0.062111	13	146	138
UbiA	28	0.004002	4	8	7

4. User Notes

All the data and metadata described in this study are at https://doi.org/10.5066/P9Q3CFL5 [25], and all the novel GenBank accessions for this study were as follows: CO1: MT662002–MT662099;

FEM1: MT669773–MT669799; ITS1: MT661766–MT661792; ND1: MT669647–MT669745; and UbiA: MT669746–MT669772 (Table 1).

Author Contributions: Conceptualization, N.A.J. and C.H.S.; methodology, N.A.J. and C.H.S.; validation, N.A.J. and C.H.S.; formal analysis, N.A.J. and C.H.S.; investigation, N.A.J. and C.H.S.; resources, N.A.J. and C.H.S.; data curation, N.A.J. and C.H.S.; writing—original draft preparation, N.A.J.; writing—review and editing, C.H.S.; visualization, N.A.J.; supervision, N.A.J.; project administration, N.A.J.; funding acquisition, N.A.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the U.S. Fish and Wildlife Service and U.S. Geological Survey.

Acknowledgments: The authors thank John Pfeiffer for providing preliminary data and advice during the marker development stage of this project and Matt Cannister for assistance with preparing the metadata file for sharing on ScienceBase. Special thanks to Jeff Powell for help obtaining funding, which was provided by the U.S. Fish and Wildlife Service and U.S. Geological Survey. The specimens utilized in this study were either from museum collections or collected under the U.S. Fish and Wildlife Service permit TE 697819-4. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Smith, C.H.; Johnson, N.A.; Pfeiffer, J.M.; Gangloff, M.M. Molecular and morphological data reveal non-monophyly and speciation in imperiled freshwater mussels (*Anodontoides* and *Strophitus*). *Mol. Phylogenet. Evol.* **2018**, *119*, 50–62. [CrossRef] [PubMed]
2. Johnson, N.A.; Smith, C.H.; Pfeiffer, J.M.; Randklev, C.R.; Williams, J.D.; Austin, J.D. Integrative taxonomy resolves taxonomic uncertainty for freshwater mussels being considered for protection under the U.S. Endangered Species Act. *Sci. Rep.* **2018**, *8*, 1–16. [CrossRef] [PubMed]
3. Keogh, S.M.; Simons, A.M. Molecules and morphology reveal 'new' widespread North American freshwater mussel species (Bivalvia: Unionidae). *Mol. Phylogenet. Evol.* **2019**, *138*, 182–192. [CrossRef]
4. Inoue, K.; McQueen, A.L.; Harris, J.L.; Berg, D.J. Molecular phylogenetics and morphological variation reveal recent speciation in freshwater mussels of the genera *Arcidens* and *Arkansia* (Bivalvia: Unionidae). *Biol. J. Linn. Soc.* **2014**, *112*, 535–545. [CrossRef]
5. Inoue, K.; Hayes, D.M.; Harris, J.L.; Christian, A.D. Phylogenetic and morphometric analyses reveal ecophenotypic plasticity in freshwater mussels *Obovaria jacksoniana* and *Villosa arkansasensis* (Bivalvia: Unionidae). *Ecol. Evol.* **2013**, *3*, 2670–2683. [CrossRef] [PubMed]
6. Campbell, D.C.; Lydeard, C. The genera of Pleurobemini (Bivalvia: Unionidae: Ambleminae). *Am. Malacol. Bull.* **2012**, *30*, 19–38. [CrossRef]
7. King, T.L.; Eackles, M.S.; Gjetvaj, B.; Hoeh, W.R. Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia: Unionidae): Conservation implications of range discontinuity. *Mol. Ecol.* **1999**, *8*, S65–S78. [CrossRef]
8. Jones, J.W.; Neves, R.J.; Ahlstedt, S.A.; Hallerman, E.M. A holistic approach to taxonomic evaluation of two closely related endangered freshwater mussel species, the oyster mussel *Epioblasma capsaeformis* and tan riffleshell *Epioblasma florentina walkeri* (Bivalvia: Unionidae). *J. Molluscan Stud.* **2006**, *72*, 267–283. [CrossRef]
9. Campbell, D.C.; Johnson, P.D.; Williams, J.D.; Rindsberg, A.K.; Serb, J.M.; Small, K.K.; Lydeard, C. Identification of 'extinct' freshwater mussel species using DNA barcoding. *Mol. Ecol. Resour.* **2008**, *8*, 711–724. [CrossRef]
10. Pfeiffer, J.M.; Johnson, N.A.; Randklev, C.R.; Howells, R.G.; Williams, J.D. Generic reclassification and species boundaries in the rediscovered freshwater mussel '*Quadrula*' *mitchelli* (Simpson in Dall, 1896). *Conserv. Genet.* **2016**, *17*, 279–292. [CrossRef]
11. Smith, C.H.; Johnson, N.A.; Pfeiffer, J.M.; Gangloff, M.M. Molecular and morphological data to facilitate future research on freshwater mussels (Bivalvia: Unionidae: Anodontinae). *Data Brief.* **2018**, *17*, 95–104. [CrossRef] [PubMed]
12. Smith, C.H.; Johnson, N.A.; Havlik, K.; Doyle, R.D.; Randklev, C.R. Resolving species boundaries in the critically imperiled freshwater mussel species, *Fusconaia mitchelli* (Bivalvia: Unionidae). *J. Zool. Syst. Evol. Res.* **2020**, in press. [CrossRef]

13. Smith, C.H.; Johnson, N.A.; Inoue, K.; Doyle, R.D.; Randklev, C.R. Integrative taxonomy reveals a new species of freshwater mussel, *Potamilus streckersoni* sp. nov. (Bivalvia: Unionidae): Implications for conservation and management. *Syst. Biodivers.* **2019**, *17*, 331–348. [CrossRef]
14. Grobler, P.J.; Jones, J.W.; Johnson, N.A.; Beaty, B.; Struthers, J.; Neves, R.J.; Hallerman, E.M. Patterns of genetic differentiation and conservation of the Slabside Pearlymussel, *Lexingtonia dolabelloides* (Lea, 1840) in the Tennessee River drainage. *J. Molluscan Stud.* **2006**, *72*, 65–75. [CrossRef]
15. Lopes-Lima, M.; Froufe, E.; Do, V.T.; Ghamizi, M.; Mock, K.E.; Kebapçı, Ü.; Klishko, O.; Kovitvadhi, S.; Kovitvadhi, U.; Paulo, O.S.; et al. Phylogeny of the most species-rich freshwater bivalve family (Bivalvia: Unionida: Unionidae): Defining modern subfamilies and tribes. *Mol. Phylogenet. Evol.* **2017**, *106*, 174–191. [CrossRef]
16. Pfeiffer, J.M.; Sharpe, A.E.; Johnson, N.A.; Emery, K.F.; Page, L.M. Molecular phylogeny of the Nearctic and Mesoamerican freshwater mussel genus *Megalonaias*. *Hydrobiologia* **2018**, *811*, 139–151. [CrossRef]
17. Jeratthitikul, E.; Phuangphong, S.; Sutcharit, C.; Prasankok, P.; Kongim, B.; Panha, S. Integrative taxonomy reveals phenotypic plasticity in the freshwater mussel *Contradens contradens* (Bivalvia: Unionidae) in Thailand, with a description of a new species. *Syst. Biodivers.* **2019**, 1–14. [CrossRef]
18. Whelan, N.V.; Geneva, A.J.; Graf, D.L. Molecular phylogenetic analysis of tropical freshwater mussels (Mollusca: Bivalvia: Unionoida) resolves the position of *Coelatura* and supports a monophyletic Unionidae. *Mol. Phylogenet. Evol.* **2011**, *61*, 504–514. [CrossRef]
19. Huang, X.-C.; Su, J.-H.; Ouyang, J.-X.; Ouyang, S.; Zhou, C.-H.; Wu, X.-P. Towards a global phylogeny of freshwater mussels (Bivalvia: Unionida): Species delimitation of Chinese taxa, mitochondrial phylogenomics, and diversification patterns. *Mol. Phylogenet. Evol.* **2019**, *130*, 45–59. [CrossRef]
20. Huff, S.W.; Campbell, D.; Gustafson, D.L.; Lydeard, C.; Altaba, C.R.; Giribet, G. Investigations into the phylogenetic relationships of freshwater pearl mussels (Bivalvia: Margaritiferidae) based on molecular data: Implications for their taxonomy and biogeography. *J. Molluscan Stud.* **2004**, *70*, 379–388. [CrossRef]
21. Ekblom, R.; Galindo, J. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* **2011**, *107*, 1–15. [CrossRef] [PubMed]
22. Ellegren, H. Genome sequencing and population genomics in non-model organisms. *Trends Ecol. Evol.* **2014**, *29*, 51–63. [CrossRef] [PubMed]
23. Pfeiffer, J.M.; Breinholt, J.W.; Page, L.M. Unioverse: Phylogenomic resources for reconstructing the evolution of freshwater mussels (Unionoida). *Mol. Phylogenet. Evol.* **2019**, *137*, 114–126. [CrossRef]
24. Smith, C.H.; Pfeiffer, J.M.; Johnson, N.A. Comparative phylogenomics reveal complex evolution of life history strategies in a clade of bivalves with parasitic larvae (Bivalvia: Unionoida: Ambleminae). *Cladistics* **2020**. [CrossRef]
25. Johnson, N.A.; Smith, C.H. Novel genetic resources to facilitate future molecular studies in freshwater mussels (Bivalvia: Unionidae). U.S. Geological Survey Data Release 2020. Available online: https://doi.org/10.5066/P9Q3CFL5 (accessed on 29 July 2020).
26. Smith, C.H.; Johnson, N.A. A comparative phylogeographic approach to facilitate recovery of an imperiled freshwater mussel (Bivalvia: Unionida: *Potamilus inflatus*). *Diversity* **2020**, *12*, 281. [CrossRef]
27. Townsend, J.P. Profiling phylogenetic informativeness. *Syst. Biol.* **2007**, *56*, 222–231. [CrossRef]
28. Bogarín, D.; Pérez-Escobar, O.A.; Groenenberg, D.; Holland, S.D.; Karremans, A.P.; Lemmon, E.M.; Lemmon, A.R.; Pupulin, F.; Smets, E.; Gravendeel, B. Anchored hybrid enrichment generated nuclear, plastid and mitochondrial markers resolve the *Lepanthes horrida* (Orchidaceae: Pleurothallidinae) species complex. *Mol. Phylogenet. Evol.* **2018**, *129*, 27–47. [CrossRef]
29. Fragoso-Martínez, I.; Salazar, G.A.; Martínez-Gordillo, M.; Magallón, S.; Sánchez-Reyes, L.; Moriarty Lemmon, E.; Lemmon, A.R.; Sazatornil, F.; Granados Mendoza, C. A pilot study applying the plant anchored hybrid enrichment method to New World sages (*Salvia* subgenus *Calosphace*; Lamiaceae). *Mol. Phylogenet. Evol.* **2017**, *117*, 124–134. [CrossRef]
30. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]
31. Chernomor, O.; von Haeseler, A.; Minh, B.Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst. Biol.* **2016**, *65*, 997–1008. [CrossRef]
32. Britton, T.; Anderson, C.L.; Jacquet, D.; Lundqvist, S.; Bremer, K. Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **2007**, *56*, 741–752. [CrossRef] [PubMed]

33. López-Giráldez, F.; Townsend, J.P. PhyDesign: An online application for profiling phylogenetic informativeness. *BMC Evol. Biol.* **2011**, *11*, 1–4. [CrossRef] [PubMed]
34. Pond, S.L.K.; Frost, S.D.W.; Muse, S.V. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **2005**, *21*, 676–679. [CrossRef] [PubMed]
35. States, D.J.; Gish, W. Combined use of sequence similarity and codon bias for coding region identification. *J. Comput. Biol.* **1994**, *1*, 39–50. [CrossRef] [PubMed]
36. Krauthammer, M.; Rzhetsky, A.; Morozov, P.; Friedman, C. Using BLAST for identifying gene and protein names in journal articles. *Gene* **2000**, *259*, 245–252. [CrossRef]
37. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
38. Maddison, W.P.; Maddison, D.R. Mesquite: A modular system for evolutionary analysis. Version 3.61. 2019. Available online: https://www.mesquiteproject.org/ (accessed on 29 July 2020).
39. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
40. Campbell, D.C.; Serb, J.M.; Buhay, J.E.; Roe, K.J.; Minton, R.L.; Lydeard, C. Phylogeny of North American amblemines (Bivalvia, Unionoida): Prodigious polyphyly proves pervasive across genera. *Invertebr. Biol.* **2005**, *124*, 131–164. [CrossRef]
41. Serb, J.M.; Buhay, J.E.; Lydeard, C. Molecular systesmatics of the North American freshwater bivalve genus *Quadrula* (Unionidae: Ambleminae) based on mitochondrial ND1 sequences. *Mol. Phylogenet. Evol.* **2003**, *28*, 1–11. [CrossRef]
42. Leigh, J.W.; Bryant, D. POPART: Full-Feature software for haplotype network construction. *Methods Ecol. Evol.* **2015**, *6*, 1110–1116. [CrossRef]