**Supplemental Information for:**

Genome-wide SNPs redefines species boundaries and conservation units in the freshwater mussel genus *Cyprogenia* of North America

**Kyung Seok Kim and Kevin J. Roe**\*

Iowa State University, Department of Natural Resource Ecology and Management, Ames, IA, 50011, USA

Running Title: Fine-scale population structure in *Cyprogenia*

\* Corresponding author:

Kevin J. Roe

Iowa State University, Department of Natural Resource Ecology and Management, Ames, IA, 50011, USA

E-mail: kjroe@iastate.edu

**INDEX**

1. **ddRAD loci and quality/sample filtering**
2. **STACKS parameter selection and SNP discovery**
3. **Exploration of Evolutionary process of *Cyprogenia***
4. **References**
5. **Supplementary Tables**
6. **Supplementary Figures**

**I. ddRAD loci and quality/sample filtering**

Process\_radtags in the STACKS pipeline (Catchen et al. 2011; 2013) was used to retain the equal size of 85 bp among sequence reads after barcode removal (-t 85), to clean data by removing any sequence reads with an uncalled base (-c), and with low quality score (-q) and to rescue barcodes and restriction enzyme cut sites (-r). NGS sequencing for 238 mussel specimens generated a total of 1,822,173,521 sequence reads for 5 lanes (364,434,704 reads/lane), with an average of 6,207,131 reads per tagged individual after demultiplexing. Of these, 45.2% (average of 2,806,551 per individual) were retained after discarding any reads with low quality scores, uncalled base, ambiguous barcodes and ambiguous rad\_tags. There were no major differences in number of ddRAD raw sequence reads obtained between different indexes within each lane and between different lanes (Fig. S2). We performed several additional quality filtering approaches. Initially, we filtered out 9 of 238 individuals by removing individuals that produced below 25% of the average sequence reads per individual (2,806,551) and with low quality sequence reads using options implemented in STACKS Process\_radtags. We constructed concatenated sequences of SNPs recovered from loci that were present in >70% of remaining samples (i.e. 160 of 229) with the population parameter (--phylip\_var) and removed an additional 57 individuals that had more than 40% of alleles coded as “N” among those concatenated sequences. Lastly, among replicate samples, the replicate with a lower proportion of “N” was selected for further analyses. In summary, a total of 66 out of the 238 samples neither passed the criterion of 25% of mean number of sequence reads retained per sample, nor the criterion of 40% of ambiguous nucleotide in concatenated SNP sequences (>525 for 1,312 SNPs at -p 160 for 229 samples). After removing replicate samples, a total of 156 samples (148 *Cyprogenia* + 8 *Dromus*) were retained.

**II. STACKS parameter selection and SNP discovery**

A wrapper program (denovo\_map.pl) was used to explore the effect of combination of different core parameters (-m, -M, -n) for de novo assembly within STACKS on number of RAD and SNP loci recovered and SNP error rates. A total of 8 replicates (4 replicates from *C. aberti* and4 replicates from *C. stegaria*), that were sequenced from two independent libraries were used to run the wrapper program multiple times with a range of parameter combinations (-m: 2-10, -M: 2-10, -n: 0-7, --max\_locus\_stacks: 2-6). For each run, only one parameter was varied, with the remaining set to m = 3, M = 2, n = 0 as a default. The maximum number of stacks at a single de novo locus (--max\_locus\_stacks) was set to 3 (default) to control for confounding loci that may arise from short, sequencing error-based stacks or from repetitive sequences 1. We further explored the SNP calling model, by comparing the default SNP model (where error rate varies freely) and the bounded model, testing different values (0.05, 0.1, 0.2, and 0.3) for the upper bound (sequencing error upper bound). Lastly, we applied corrections to genotype and haplotype calls using STACKS RX program (rxstacks). Outputs were analyzed to estimate the number of RAD loci and SNPs recovered, and SNP error rates using R scripts written by Mastretta\_Yanes *et al*.1, and to explore the clustering pattern of replicate pairs, the percentage missing data and the proportion of variance using principal component analysis (PCA) (adegenet package in R) using R v. 3.3.3 2.

Once the parameter set with the best performance was selected, we manually ran the STACKS programs, i.e. ustacks, cstacks, sstacks, rxstacks. Specifically, the program cstacks was used to build catalog loci, a set of consensus loci, from 23 samples (20 *Cyprogenia* + 3 *Dromus* samples) that contain the highest number of loci and represent each sampling location and species. Subsequently, STACKS RX program (rxstacks) was used to make corrections to genotype and haplotype calls in individual samples based on population wide data using the following parameters: default SNP model, confounded loci filtering (--conf\_lim 0.25), haplotype pruning (--prune\_haplo) and minimum log likelihood required to keep a catalog locus (--lnl\_lim -10.0). Loci with alleles with minor frequency of less than 0.05 (--min\_maf 0.05) were filtered out and a data subset consisting of only the first SNP per ddRAD-Seq locus (--write\_sinlge\_snp) were retained to minimize effects of linkage disequilibrium.

Across all explored parameter profiles, the number of loci retained ranged from 46,544 to 491,466 and the number of SNPs ranged from 20,386 to 113,024 (Fig. S2A). In general, three main parameter components that control the minimal coverage (-m), and number of mismatches allowed between loci when processing a single individual (-M) and building the catalogue of loci (-n) greatly contributed to the variance of the number of loci retained and SNPs. As the mean coverage per locus increase (the min. coverage -m from 2 to 10), both the number of loci retained and SNPs substantially decreased (Fig. S2B), whereas increasing the number of mismatches allowed between loci when building the catalogue of loci (-n from 0 to 7), produced decreased number of loci retained but increased SNPs (Fig. S2B). Increasing the number of mismatches allowed between loci when processing a single individual (-M) did not affect the number of loci retained but increased the number of SNPs. On the other hand, a parameter that controls the maximum number of stacks at a single de novo locus (-max\_locus\_stacks), minimally contributed to the amount of data among different parameter settings.

Various parameters generated different levels of the SNP error rate, ranging from 5.1% to 29.9% depending on replicate samples (Fig. S3). SNP errors between replicates can be caused by allelic dropout due to low coverage, or by the allowance of error-based variation due to PCR and/or sequencing errors during de novo assembly. Indeed, SNP error rates substantially increased as minimum number of reads required to create a stack (-m) and number of mismatches allowed between loci when processing a single individual (-M) increase (Fig. S3A, B). The lowest mean SNP error rates (8.6%) was detected in a parameter set of -m 3, -M 2, and -n 3 (Fig. S3C). Again, a parameter, -max\_locus\_stacks (2-6), did not affect the SNP error rates (Fig. S3D). We further explored SNP calling models with the best performance among replicates using a selected parameter set (-m 3, -M 2, -n 3). SNP calling models did not affect the number of loci retained. However, increasing the value of the upper bounded model (from 0.05 to 0.3) revealed slight decrease in number of SNP loci, with the lowest number of SNP loci in the default SNP model. On the other hand, mean SNP error rate was the lowest in the default SNP model (8.6%), but highest in the parameter --bound\_high 0.05 (11.3%). Application of correction program (rxstacks) and the selection of the first SNP per locus (--write\_single\_snp) further improved SNP error rates from 5.7% (without correction) to 5.2% (with correction) (Fig. S3E), but traded off the recovery of some SNP loci (from 31,945 to 25,806). PCA clustering with correction resulted in increased power in Proportion of Variance in first two components (from 39.5% without correction to 44.2% with correction) as well as in proportion of missing data (from 60.62 % without correction to 54.34 % with correction). In general, PCA analysis showed that most replicates were clustered together regardless of parameter settings, but corrected data resulted in minimizing differences between replicates and increasing discrimination between species (data not shown).

Therefore, results from various exploratory analyses in STACKS revealed that assembly parameters of -m = 3, -M = 2, -n = 3, --max\_locus\_stacks = 3, and default SNP model with application of correction program (rxstacks) and selection of the first SNP per locus performed better than other parameter settings and therefore, this parameter setting was chosen for de novo RADseq assembly and SNP discovery, and for downstream genetic analyses.

**III. Exploration of Evolutionary process and demographic history in *Cyprogenia***

Plausible evolutionary scenarios and demographic history for each Highland region were explored using the Approximate Bayesian Computation (ABC) approach implemented in DIYABC 2.1 3.

When a dataset includes large number of specimens from many sampling sites with genetic structure, reorientation of samples according to genetic units defined by objective genetic clustering methods can substantially reduce the number and complexity of evolutionary scenarios to be compared using ABC simulations 4,5.

Here, we designed and compared plausible evolutionary scenarios for the two *C. aberti* regions (*C. aberti*\_Ouachita and *C. aberti*\_Ozark) using clades from the phylogenetic trees as well as genetic clusters defined by STRUCTURE. Bayesian model-based clustering analyses showed clear structure for *C. aberti*\_Ouachita and *C. aberti*\_Ozark, but not for specimens from *C. stegaria*\_Eastern (see results below). Therefore, for *C. aberti*\_Ouachita and *C. aberti*\_Ozark, six evolutionary scenarios were proposed using a three-population model, i.e. three scenarios with divergence (merge) model and three scenarios with admixture (split) model, implemented in DIYABC 2.1 3. For example, divergence (merge) scenarios were designed with a recent divergence of a population from its genetically closer population independently from the remaining one. Admixture (split) scenarios were designed to reflect that a recent population was generated by the result of admixture between the other two populations. Three populations, i.e. Rivers, in each highland were switched with the possible combinations to account for the scenario. The most supported evolutionary scenario for each of the two highlands was selected from comparisons of scenarios through the ABC computation of the posterior probabilities and the selected scenario was assessed by its goodness of fit with principal component analysis and posterior checking of summary statistics using the Model checking option implemented in DIYABC (Fig. S6).

**IV. References**

1. Mastretta-Yanes, A. *et al.* Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol. Ecol. Resour.* **15**, 28–41 (2015).

2. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (2012).

3. Cornuet, J.-M. *et al.* DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187–1189 (2014).

4. Ascunce, M. S. *et al.* Global Invasion History of the Fire Ant Solenopsis invicta. *Science (80-. ).* **331**, 1066–1068 (2011).

5. Lombaert, E. *et al.* Complementarity of statistical treatments to reconstruct worldwide routes of invasion: The case of the Asian ladybird Harmonia axyridis. *Mol. Ecol.* **23**, 5979–5997 (2014).

**V. Supplementary Tables**

**Table S1** Sample information of freshwater mussels (*Cyprogenia sp.* and *Dromus sp.)* used in this study

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Species | Sample ID | Tree OTU | River, State | Sampling method | Main Clade (subclade) | Genetic cluster -subcluster | Conglutinate ID (Color) | NCBI\_SRA | Reference/Collector |
| *Cyprogenia stegaria* | CST\_01 | Eastern1 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060972 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_02 | Eastern2 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060973 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_03 | Eastern3 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060974 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_04 | Eastern4 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060975 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_05 | Eastern5 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060976 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_06 | Eastern6 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060977 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_07 | Eastern7 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060978 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_08 | Eastern8 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060979 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CST\_09 | Eastern9 | Salt River, KY | Mantle | A | I-1 |  | SAMN09060980 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | UAUC1499 | Eastern11 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060981 | Serb, 2006 |
| *Cyprogenia stegaria* | UAUC1500 | Eastern12 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060982 | Serb, 2006 |
| *Cyprogenia stegaria* | CCB\_01 | Eastern13 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060983 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CCB\_02 | Eastern14 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060984 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CCF\_01 | Eastern15 | Clinch River, TN | Mantle | A (2) | I-2 |  | SAMN09060985 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CCF\_03 | Eastern17 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060986 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CCF\_04 | Eastern18 | Clinch River, TN | Mantle | A (2) | I-2 |  | SAMN09060987 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CCF\_05 | Eastern19 | Clinch River, TN | Mantle | A | I-1 |  | SAMN09060988 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_01 | Eastern21 | Green River, KY | Mantle | A | I-1 |  | SAMN09060989 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_02 | Eastern22 | Green River, KY | Mantle | A | I-1 |  | SAMN09060990 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_03 | Eastern23 | Green River, KY | Mantle | A | I-1 |  | SAMN09060991 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_04 | Eastern24 | Green River, KY | Mantle | A | I-1 |  | SAMN09060992 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_05 | Eastern25 | Green River, KY | Mantle | A | I-1 |  | SAMN09060993 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_06 | Eastern26 | Green River, KY | Mantle | A | I-1 |  | SAMN09060994 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_07 | Eastern27 | Green River, KY | Mantle | A | I-1 |  | SAMN09060995 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_08 | Eastern28 | Green River, KY | Mantle | A | I-1 |  | SAMN09060996 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_09 | Eastern29 | Green River, KY | Mantle | A | I-1 |  | SAMN09060997 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CGE\_10 | Eastern30 | Green River, KY | Mantle | A | I-1 |  | SAMN09060998 | Grobler et al., 2011 |
| *Cyprogenia stegaria* | CLR\_01 | Eastern31 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09060999 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_02 | Eastern32 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061000 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_03 | Eastern33 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061001 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_04 | Eastern34 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061002 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_05 | Eastern35 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061003 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_06 | Eastern36 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061004 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_07 | Eastern37 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061005 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_09 | Eastern39 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061006 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_10 | Eastern40 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061007 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_11 | Eastern41 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061008 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_12 | Eastern42 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061009 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_13 | Eastern43 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061010 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_14 | Eastern44 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061011 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_15 | Eastern45 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061012 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_17 | Eastern46 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061013 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_18 | Eastern47 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061014 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_19 | Eastern48 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061015 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_20 | Eastern49 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061016 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_21 | Eastern50 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061017 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_22 | Eastern51 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061018 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_23 | Eastern52 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061019 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_24 | Eastern53 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061020 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLR\_25 | Eastern54 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061021 | Chong et al., 2016 |
| *Cyprogenia stegaria* | CLRd\_16 | Eastern55 | Licking River, KY | cytology brushes | A | I-1 |  | SAMN09061022 | Chong et al., 2016 |
| *Cyprogenia aberti* | COR\_02 | Ouachi2 | Ouachita River, AR | Mantle | C-1 | III-1 | R\_OuaOC4 (Red) | SAMN09061023 | Chong et al., 2016 |
| *Cyprogenia aberti* | COR\_03 | Ouachi3 | Ouachita River, AR | Mantle | C-1 | III-1 | R\_OuaOC5 (Red) | SAMN09061024 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSL\_01 | Ouachi20 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC9 (Red) | SAMN09061025 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSL\_02 | Ouachi21 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC12 (Brown) | SAMN09061026 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_01 | Ouachi22 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC13 (Brown) | SAMN09061027 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_02 | Ouachi23 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC14 (Brown) | SAMN09061028 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_03 | Ouachi24 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC15 (Brown) | SAMN09061029 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_04 | Ouachi25 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC16 (Brown) | SAMN09061030 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_05 | Ouachi26 | Saline River, AR | Mantle | C-2 | III-2 | B\_SalOC17 (Brown) | SAMN09061031 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_06 | Ouachi27 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC10 (Red) | SAMN09061032 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_07 | Ouachi28 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC11 (Red) | SAMN09061033 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_08 | Ouachi29 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC12 (Red) | SAMN09061034 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_09 | Ouachi30 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC13 (Red) | SAMN09061035 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSR\_10 | Ouachi31 | Saline River, AR | Mantle | C-2 | III-2 | R\_SalOC14 (Red) | SAMN09061036 | Chong et al., 2016 |
| *Cyprogenia aberti* | UAUC1837 | Ouachi33 | Caddo River, AR | Mantle | C-1 | III-1 |  | SAMN09061037 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1838 | Ouachi34 | Caddo River, AR | Mantle | C-1 | III-1 |  | SAMN09061038 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1839 | Ouachi35 | Caddo River, AR | Mantle | C-1 | III-1 |  | SAMN09061039 | John Harris |
| *Cyprogenia aberti* | UAUC2374 | Ouachi36 | Ouachita River, AR | Mantle | C-1 | III-1 |  | SAMN09061040 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2375 | Ouachi37 | Ouachita River, AR | Mantle | C-1 | III-1 |  | SAMN09061041 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2377 | Ouachi38 | Ouachita River, AR | Mantle | C-1 | III-1 |  | SAMN09061042 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2378 | Ouachi39 | Ouachita River, AR | Mantle | C-1 | III-1 |  | SAMN09061043 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2381 | Ouachi40 | Caddo River, AR | Mantle | C-1 | III-1 |  | SAMN09061044 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2383 | Ouachi41 | Caddo River, AR | Mantle | C-1 | III-1 |  | SAMN09061045 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2736 | Ouachi42 | Saline River, AR | Mantle | C-2 | III-2 |  | SAMN09061046 | Serb, 2006 |
| *Cyprogenia aberti* | CBR\_07 | Ozark5 | Black River, MO | cytology brushes | B-1 | 1II-1 | R\_BlaOZ1 (Red) | SAMN09061047 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_09 | Ozark7 | Black River, MO | cytology brushes | B-1 | II-1 | R\_BlaOZ2 (Red) | SAMN09061048 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_10 | Ozark8 | Black River, MO | Mantle | B-1 | II-1 | R\_BlaOZ3 (Red) | SAMN09061049 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_13 | Ozark11 | Black River, MO | cytology brushes | B-1 | II-1 | B\_BlaOZ1 (Brown) | SAMN09061050 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_14 | Ozark12 | Black River, MO | cytology brushes | B-1 | II-1 | B\_BlaOZ2 (Brown) | SAMN09061051 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_15 | Ozark13 | Black River, MO | Mantle | B-1 | II-1 | B\_BlaOZ3 (Brown) | SAMN09061052 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_18 | Ozark16 | Black River, MO | Mantle | B-1 | II-1 |  | SAMN09061053 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_19 | Ozark17 | Black River, MO | Mantle | B-1 | II-1 |  | SAMN09061054 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBR\_27 | Ozark22 | Black River, MO | cytology brushes | B-1 | II-1 |  | SAMN09061055 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_02 | Ozark24 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061056 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_03 | Ozark25 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061057 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_04 | Ozark26 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061058 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_05 | Ozark27 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061059 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_06 | Ozark28 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061060 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_07 | Ozark29 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061061 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_08 | Ozark30 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061062 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_09 | Ozark31 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061063 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_10 | Ozark32 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061064 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_11 | Ozark33 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061065 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_12 | Ozark34 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061066 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_13 | Ozark35 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061067 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_14 | Ozark36 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061068 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_16 | Ozark38 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061069 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_18 | Ozark40 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061070 | Chong et al., 2016 |
| *Cyprogenia aberti* | CBW\_19 | Ozark41 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061071 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_01 | Ozark42 | St. Francis River, MO | Mantle | B-2 | II-2 | R\_StFOZ6 (Red) | SAMN09061072 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_02 | Ozark43 | St. Francis River, MO | Mantle | B-2 | II-2 | R\_StFOZ7 (Red) | SAMN09061073 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_03 | Ozark44 | St. Francis River, MO | Mantle | B-2 | II-2 | R\_StFOZ8 (Red) | SAMN09061074 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_04 | Ozark45 | St. Francis River, MO | Mantle | B-2 | II-2 | B\_StFOZ4 (Brown) | SAMN09061075 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_05 | Ozark46 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ5 (Brown) | SAMN09061076 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_06 | Ozark47 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ6 (Brown) | SAMN09061077 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_07 | Ozark48 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ7 (Brown) | SAMN09061078 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_08 | Ozark49 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ8 (Brown) | SAMN09061079 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_09 | Ozark50 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ9 (Brown) | SAMN09061080 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_10 | Ozark51 | St. Francis River, MO | Mantle | B-2 | II-2 | B\_StFOZ10 (Brown) | SAMN09061081 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_13 | Ozark54 | St. Francis River, MO | cytology brushes | B-2 | II-2 | B\_StFOZ11 (Brown) | SAMN09061082 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_18 | Ozark59 | St. Francis River, MO | cytology brushes | B-2 | II-2 |  | SAMN09061083 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_19 | Ozark60 | St. Francis River, MO | cytology brushes | B-2 | II-2 |  | SAMN09061084 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_21 | Ozark62 | St. Francis River, MO | Mantle | B-2 | II-2 |  | SAMN09061085 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSF\_24 | Ozark65 | St. Francis River, MO | cytology brushes | B-2 | II-2 |  | SAMN09061086 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_01 | Ozark70 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061087 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_02 | Ozark71 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061088 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_05 | Ozark73 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061089 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_06 | Ozark74 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061090 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_07 | Ozark75 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061091 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_08 | Ozark76 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061092 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_09 | Ozark77 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061093 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_10 | Ozark78 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061094 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_12 | Ozark79 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061095 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_13 | Ozark80 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061096 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_14 | Ozark81 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061097 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_15 | Ozark82 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061098 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_16 | Ozark83 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061099 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_17 | Ozark84 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061100 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSPd\_04 | Ozark86 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061101 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSP\_11 | Ozark89 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061102 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_01 | Ozark96 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061103 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_02 | Ozark97 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061104 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_03 | Ozark98 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061105 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_04 | Ozark99 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061106 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_05 | Ozark100 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061107 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_06 | Ozark101 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061108 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_07 | Ozark102 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061109 | Chong et al., 2016 |
| *Cyprogenia aberti* | CSS\_08 | Ozark103 | Spring River, AR | Mantle | B-1 | II-1 |  | SAMN09061110 | Chong et al., 2016 |
| *Cyprogenia aberti* | UAUC1446 | Ozark104 | St. Francis River. MO | Mantle | B-2 | II-2 |  | SAMN09061111 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1535 | Ozark105 | Spring River, KS | Mantle | B-3 | II-3 |  | SAMN09061112 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1536 | Ozark106 | Current, AR | Mantle | B-1 | II-1 |  | SAMN09061113 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1647 | Ozark107 | Fall River, KS | Mantle | B-3 | II-3 |  | SAMN09061114 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1650 | Ozark108 | Buffalo, AR | Mantle | B-1 | II-1 |  | SAMN09061115 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC1652 | Ozark109 | Strawberry, AR | Mantle | B-1 | II-1 |  | SAMN09061116 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2724 | Ozark110 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061117 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2725 | Ozark111 | Black River, AR | Mantle | B-1 | II-1 |  | SAMN09061118 | Serb, 2006 |
| *Cyprogenia aberti* | UAUC2726 | Ozark112 | Strawberry, AR | Mantle | B-1 | II-1 |  | SAMN09061119 | Serb, 2006 |
| *Dromus dromas*  | DRO\_1R | Dromus1 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061120 | This study |
| *Dromus dromas*  | DRO\_2R | Dromus2 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061121 | This study |
| *Dromus dromas*  | DRO\_3W | Dromus3 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061122 | This study |
| *Dromus dromas*  | DRO\_4U | Dromus4 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061123 | This study |
| *Dromus dromas*  | DRO\_6R | Dromus6 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061124 | This study |
| *Dromus dromas*  | DRO\_7R | Dromus7 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061125 | This study |
| *Dromus dromas*  | T358 | Dromus11 | Clinch River, TN | cytology brushes | D | IV |  | SAMN09061126 | This study |
| *Dromus dromas*  | T360 | Dromus13 | Clinch River, TN | cytology brushes | D | IV |   | SAMN09061127 | This study |

**Table S2** Pairwise *F*ST’s among 10 rivers throughout distribution range of genus *Cyprogenia* based on 7,243 SNP loci. *F*ST values from raw data below diagonal and *F*ST values from corrected data above diagonal

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Highland |  | Eastern | Ozark | Ouachita |
|  | River | Clinch | Green | Licking | Salt | Black | Spring | St.Francis | Caddo | Ouachita | Saline |
| Eastern | Clinch | ***-*** | 0.026 | ***0.026*** | 0.019 | ***0.197*** | ***0.259*** | ***0.244*** | ***0.271*** | ***0.248*** | ***0.215*** |
|  | Green | 0.051 | ***-*** | 0.003 | 0.006 | ***0.241*** | ***0.304*** | ***0.287*** | ***0.307*** | ***0.286*** | ***0.248*** |
|  | Licking | ***0.050*** | 0.007 | ***-*** | 0.005 | ***0.238*** | ***0.298*** | ***0.283*** | ***0.301*** | ***0.283*** | ***0.250*** |
|  | Salt | 0.059 | 0.012 | 0.012 | ***-*** | ***0.212*** | ***0.275*** | ***0.257*** | ***0.278*** | ***0.259*** | ***0.226*** |
| Ozark | Black | ***0.506*** | ***0.508*** | ***0.513*** | ***0.508*** | ***-*** | ***0.018*** | ***0.030*** | ***0.134*** | ***0.119*** | ***0.098*** |
|  | Spring | ***0.535*** | ***0.537*** | ***0.539*** | ***0.537*** | ***0.041***  | ***-*** | ***0.024*** | ***0.174*** | ***0.160*** | ***0.140*** |
|  | St.Francis | ***0.563*** | ***0.558*** | ***0.554*** | ***0.561*** | ***0.106*** | ***0.073*** | ***-*** | ***0.173*** | ***0.155*** | ***0.136*** |
| Ouachita | Caddo | ***0.612*** | ***0.595*** | ***0.580*** | ***0.605*** | ***0.308*** | ***0.354*** | ***0.402*** | ***-*** | 0.007 | 0.036 |
|  | Ouachita | ***0.609*** | ***0.594*** | ***0.582*** | ***0.604*** | ***0.311*** | ***0.357*** | ***0.400*** | 0.046 | ***-*** | 0.030 |
|  | Saline | ***0.632*** | ***0.619*** | ***0.603*** | ***0.630*** | ***0.342*** | ***0.397*** | ***0.443*** | ***0.206*** | ***0.177*** | ***-*** |

A total of 146 individuals except two geographically isolated individuals (Ozark\_105 and Ozark\_107) were included in this analysis. Significance was tested by genic differentiation for each population pair from exact G test using default Markov chain parameters (Dememorisation: 10000, Batches: 100, Iterations per batch: 5000). Bold Italicized indicates highly significant *F*ST value. Corrected data refers to data that missing data were replaced with randomly drawn alleles based on the overall allele frequencies. Significance level was adjusted for multiple testing using Bonferroni correction.

**Table S3** Posterior distribution of the current and ancestral effective population sizes (*Ne*) estimates (Median and 95% credible interval) for each *Cyprogenia aberti* sample site from best fit scenario for evolutionary process using ABC simulation implemented in DIYABC

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Highlands | Parameter (*Ne*) | Prior distribution | Median | Quantile5% | Quantile95% |
|  | BL | Uniform[10 - 2.0 x104] | 2.09 x 102 | 5.29 x 102 | 8.15 x 103 |
|  | SP | Uniform[10 - 2.0 x104] | 6.38 x 103 | 1.93 x 103 | 1.54 x 104 |
| Ozark | SF | Uniform[10 - 2.0 x104] | 4.23 x 102 | 1.49 x 102 | 1.53 x 103 |
|  | OZKa | Uniform[10 - 2.0 x104] | 1.99 x 104 | 1.95 x 104 | 2.00 x 104 |
|  | OU | Uniform[10 - 2.0 x104] | 1.36 x 103 | 2.62 x 102 | 7.99 x 103 |
|  | CA | Uniform[10 - 2.0 x104] | 4.31 x 102 | 1.26 x 102 | 2.48 x 103 |
| Ouachita | SA | Uniform[10 - 2.0 x104] | 8.09 x 102 | 3.11 x 102 | 2.56 x 103 |
|  | OUAa | Uniform[10 - 2.0 x104] | 1.93 x 104 | 1.69 x 104 | 1.99 x 104 |

Probable evolutionary scenarios were designed and selected based on genetic clusters identified with STRUCTURE (Details in Supplementary Information). The prior distribution and interval are given for each parameter. BL: Black, SP: Spring, SF: St. Francis, OZKa: Ancestral Ozark Highland, OU: Ouachita, CA: Caddo, SA: Saline, OUAa: Ancestral Ouachita Highland.

Best fit scenario for Ouachita Highland; Scenario 4: Admixture (split) model, where Ouachita population originated from an admixture between Caddo, and Saline rivers, Best fit scenario for Ozark Highland; Scenario 1: Divergence (merge) model, Black population is derived from Spring, and Spring population originated from St. Francis.

**VI. Supplementary Figures**

**Figure S1.** Flowchart of experimental design from library construction to data analyses in this study. Flowchart shows different datasets (e.g. raw reads, phylogenetic and population genetic RAD loci datasets), data processing steps (e.g. de-multiplexing, loci assembly, loci filtering, and selection of optimal STACKS parameter), genetic analyses (e.g. phylogenetics, population genetic, outlier analyses for conglutinate egg color, and evolutionary model using ABC simulations).

 **Figure S2.** Number of ddRAD raw sequence reads and single-nucleotide polymorphisms based on STACKS core parameter settings.

A: Number of ddRAD raw sequence reads obtained between different indexes within each lane and between different lanes

B: Total number of ddRAD loci (Red dotted line) and single-nucleotide polymorphisms (Grey dotted line) obtained using various STACKS core parameter settings. For each run, only one parameter varied, with the remaining set to m = 3, M = 2, n = 0 and max\_locus\_stacks (--mls) = 3



A



 B

**Figure S3.** Effect of different settings for STACKS core parameters and program components on the single nucleotide polymorphism (SNP) error rate.

A: -m = 2-10, B: -M = 2-10, C: -n = 0-7, D: --max\_locus\_stacks = 2-6 E: SNP calling models (--bound\_high 0.05 vs default SNP model), correction program (rxstacks). and the selection of the first SNP per locus (--write\_single\_snp)

For each run, only one parameter varied (shown on the x axis), with the remaining set to m = 3, M = 2, n = 0 and max\_locus\_stacks (mx.lcs) = 3



 A B

 C D



 E

**Figure S4.** Phylogenetic relationships of freshwater mussel species based on maximum likelihood (ML) analysis based on concatenated sequences of SNPs obtained from four different filtering criteria.

ML trees from different filtering criteria, -p 78, -p 94, -p 109, nd -p 125, where -p refers to the minimum number of samples that a locus must be present in for inclusion of the locus in the phylogenetic analysis. Trees were constructed based on 7,440, 4,395, 2,014, and 511, informative sites respectively. All datasets selected the same TVM+R3 as a best-fit model according to BIC implemented in IQtree. Correspondence of each color to sampling site is represented in Fig. 3.



 -p 78 -p 94 -p 109 -p 125

**Figure S5.** Phylogenetic trees for 31 *C. aberti* specimens with identified egg colors of conglutinate lure.

Maximum likelihood (ML) tree (A) and Bayesian tree (B) were constructed based on 1,173 informative sites to examine phylogenetic relationships for 33 specimens, including 31 *C. aberti* specimens identified by conglutinate egg colors, and two *C. stegeria* specimens used as an outgroup. The TVM+I+G4 model of sequence evolution was selected following a Bayesian Information Criterion (BIC) implemented in IQtree 1.5.6. Red branch refers to specimens with red color of conglutinate lure and brown branch to specimens with brown color of conglutinate lure)



A



B

**Figure S6.** Results of scenario comparison among 6 evolutionary scenarios and the principal component model check analyses using DIYABC.

The best fit evolutionary scenario for each of three Highlands was selected from comparisons of six evolutionary scenarios through the ABC computation of the posterior probabilities of each scenario and by their goodness of fit to data sets with principal component analysis and posterior checking of summary statistics using the Model checking option implemented in DIYABC. The selected scenario fit our data as shown in principal component analyses.

A: Ouachita Highland (Ouachita, Caddo, and Saline Rivers,): A best fit evolutionary scenario: admixture (split) scenario 4 [Ancestral ® Saline and Caddo (divergence) ® Ouachita (admixture)]

B: Ozark Highland (Black, Spring, and St. Francis Rivers): A best fit evolutionary scenario: divergence (merge) scenario 1 [Ancestral ® St. Francis and Spring (divergence) ® Black from Spring (divergence)]

Arrow shows plausible evolutionary trajectory from ancestral population to current populations.



A

****

B

**Figure S7.** Comparison of various estimates (Mean±SD) of ddRAD loci from 110 destructive (mantle biopsy) and 46 non-destructive (cytology brush) samples obtained after STACKS process\_radtags running.

Number of retained loci (Retained) was obtained after subtraction of loci with no radtag (NoRadtags) and low quality (LowQuality) from total number of ddRAD loci generated.



**Figure S8.** Mapping results of destructive and nondestructive genetic samples using FastQ\_Screen

ddRAD fastq sequences of two randomly selected were mapped to four genome sequences and PhiX sequences.

A: CSF\_21 (Sample using destructive method: 3,023,365 loci)

B: CSF\_19 (Sample using non-destructive cytology brush: 3,881,449 loci)

Genome sequences of a bivalve, *Mytilus galloprovincialis*, are available from GenBank (accession: LNJA000000000). Genome sequences of human (*Homo sapiens*), and yeast (*Saccharomyces\_cerevisiae*) are available from ftp://ftp.ensembl.org/pub/current\_fasta. Genome sequences of a bacterium (*Escherichia coli*) are available from GenBank (accession: U00096.3). PhiX sequences are available from Refseq accession NC\_001422.1



A

****

B