Next-generation genomic shotgun sequencing indicates greater genetic variability in the mitochondria of...

Article *in* Conservation Genetics Resources · March 2014

DOI: 10.1007/s12686-014-0296-3

CITATIONS
3

READS
27

4 authors, including:

Michael S Eackles
United States Geological Survey
16 PUBLICATIONS **355** CITATIONS
SEE PROFILE

Jay Richard Stauffer
Pennsylvania State University
230 PUBLICATIONS **3,762** CITATIONS
SEE PROFILE

Conservation Genet Resour (2015) 7:9–11
DOI 10.1007/s12686-014-0296-3

TECHNICAL NOTE

Next-generation genomic shotgun sequencing indicates greater genetic variability in the mitochondria of *Hypophthalmichthys molitrix* relative to *H. nobilis* from the Mississippi River, USA and provides tools for research and detection

John J. Miller · Michael S. Eackles · Jay R. Stauffer · Tim L. King

Received: 23 May 2014 / Accepted: 8 August 2014 / Published online: 19 September 2014

Abstract We characterized variation within the mitochondrial genomes of the invasive silver carp (*Hypophthalmichthys molitrix*) and bighead carp (*H. nobilis*) from the Mississippi River drainage by mapping our Next-Generation sequences to their publicly available genomes. Variant detection resulted in 338 single-nucleotide polymorphisms for *H. molitrix* and 39 for *H. nobilis*. The much greater genetic variation in *H. molitrix* mitochondria relative to *H. nobilis* may be indicative of a greater North American female effective population size of the former. When variation was quantified by gene, many tRNA loci appear to have little or no variability based on our results whereas protein-coding regions were more frequently polymorphic. These results provide biologists with additional regions of DNA to be used as markers to study the invasion dynamics of these species.

Keywords Asian carp, *Hypophthalmichthys molitrix* · *Hypophthalmichthys nobilis* · Mitochondrial genome · Single-nucleotide polymorphism (SNP) · Invasive

The invasion of the Asian carp *Hypophthalmichthys moltrix* (silver carp) and *Hypophthalmichthys nobilis* (bighead carp) into the Laurentian Great Lakes is of concern to conservationists due to potential impacts on aquatic ecosystems. *H. molitrix* and *H. nobilis* have been shown to alter plankton communities (Burk et al. 1986; Berday et al. 2005) and compete with native filter feeding fish (Irons et al. 2007) with which they have overlapping diets (Sampson et al. 2009). Thus, an invasion of the Great Lakes by *Hypophthalmichthys* spp. could have a substantial impact on these ecosystems. The development of control or eradication strategies for these carp will require the characterization of the associated migration, colonization, and extinction processes among nascent populations yet little is known about the genetics of these fish in North America. Heritable markers can represent a valuable tool for studying these phenomena (i.e., Makhrov and Bolotov 2006; King et al. 2011) and aid in the detection of these species from environmental samples via eDNA (Jerde et al. 2011). To further address the need for additional molecular tools, we mapped our Next-Generation Sequencing reads to the publicly available mitochondrial genomes of Li et al. (2009b) and characterized the variability within these species. The previously undetected variation identified in this research provides additional markers for distinguishing individuals, populations, and species as well as for eDNA detection. All of which are important for informing control efforts.

DNA from multiple individuals was pooled separately for each species (Supplemental material 1 for information on numbers and locations) and sequenced using Ion Torrent Proton and Personal Genome Machine as well as Roche 454 GS Jr. platforms. Low quality ends were removed from the reads using TrimmingReads.pl from NGS QC Toolkit (Patel and Jain 2012) requiring a Phred score of at least 25 and a sequence length of at least 40 nucleotides. Redundant reads were removed using Picard (http://picard.sourceforge.net) MarkDuplicates with the REMOVE DUPLICATES option.

Electronic supplementary material The online version of this article (doi:10.1007/s12686-014-0296-3) contains supplementary material, which is available to authorized users.

J. J. Miller (✉) · J. R. Stauffer
School of Forest Resources, Pennsylvania State University, University Park, PA 16802, USA
e-mail: millerjohnjames.miller@gmail.com

J. J. Miller · M. S. Eackles · T. L. King
Aquatic Ecology Branch, U. S. Geological Survey, Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430, USA

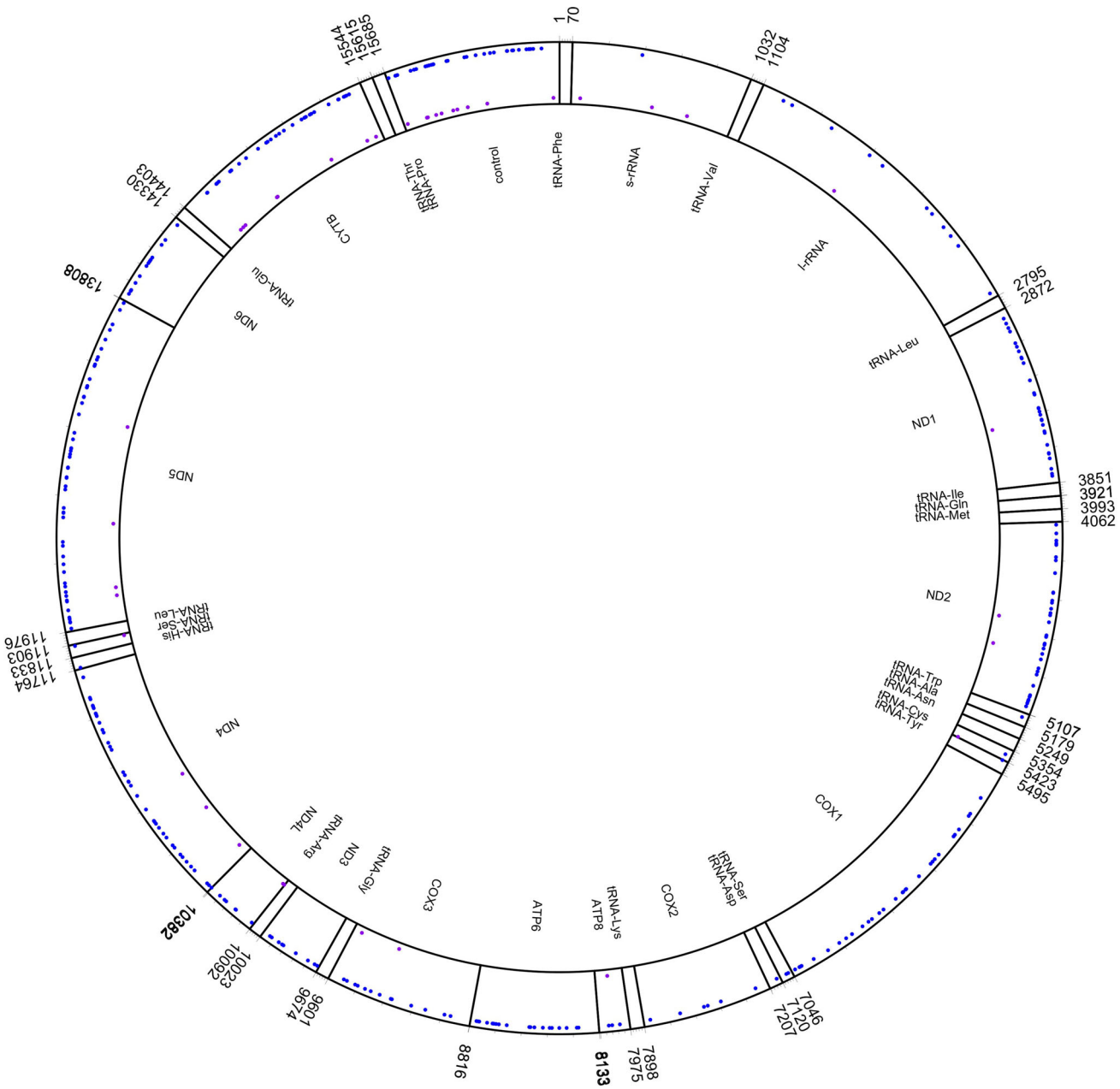

Fig. 1 Ideogram showing gene borders for *H. molitrix*, gene names, and approximate locations of SNPs for both species. The outer ring of *dots* represents the locations of variant sites in *H. molitrix* while the inner ring of *dots* represents the locations of variant sites in *H. nobilis*

We used BWA-MEM (Li and Durbin 2010) to map reads to the conspecific published mitochondrial genomes (Li et al. 2009b) of *H. nobilis* (NC_010194) and *H. molitrix* (NC_010156). Mapped reads will be submitted to the Sequence Read Archive at NCBI (National Center for Biotechnology Information). SAMtools/bcftools (Li et al. 2009a) was used to call variants with a base quality cutoff of at least 25 and a mapping quality cutoff of at least 30 with a mapping quality coefficient downgrade of 50. Picard and SAMtools utilities were used to merge and interconvert file formats as appropriate. Only single-nucleotide polymorphisms (SNPs) with a vcf quality (QUAL column) greater than 100 and coverage greater than 50 were counted/reported in this paper. Perl was used to parse vcf files and statistical information regarding the mapping and variant calling was gathered and calculated using Perl and R version 3.1.0 (R Core Team 2014) and Picard. The R package "circlize" (Gu et al. 2014) was used to draw the ideogram showing the approximate locations of SNPs. The Fisher's exact test was calculated with R using average polymorphism frequency multiplied by locus length as expected values.

Variation within the two carp mitochondrial genomes was elucidated by mapping NGS reads to the genome sequences of the same species reported by Li et al. (2009b). Of 72.8 million *H. molitrix* reads 127,893 mapped to its mitochondrial genome (NC_010156) resulting in 100 % coverage of this genome to a mean coverage depth of 852X (median = 873, SD = 257, min = 16, max = 1,533). Variant detection resulted in a total of 338 SNPs within the *H. molitrix* mitochondrial genome accounting for 2.0 % of residues. Of 82.7 million *H. nobilis* reads, 96,869 mapped to its mitochondrial genome (NC_010194) covering 100 % of this genome with a mean depth of 645X (median = 666, SD = 191, min = 30, max = 1,144). Variant detection resulted in a total of 39 polymorphic sites or 0.2 % of the residues. The greater mitochondrial genetic variation in *H. molitrix* relative to *H. nobilis* illustrated, along with the approximate locations of the SNPs, in Fig. 1 (see Supplemental material 2 for exact locations) is indicative of a greater female effective population size of North American *H. molitrix* relative to *H. nobilis*, possibly due to a larger number of independent introductions of *H. molitrix* and is consistent with the results of King et al. (2011).

Closer inspection of Fig. 1 indicates the density of SNPs within tRNA loci was lower than for the remainder of each genome. To test this hypothesis, Fisher's exact tests were conducted. The *H. molitrix* data indicated a significantly reduced level of polymorphism within the tRNA regions ($p = 0.0013$). However, the same pattern was not supported in the *H. nobilis* data ($p = 0.67$) possibly due to the lower overall frequency of polymorphisms in the *H. nobilis* mitochondrial genome. The reduced variation in tRNA sequences could be due to greater selective pressure on these loci. As is indicated in Fig. 1, most regions of the *H. molitrix* mitochondrial genome, with the exception of the tRNAs, and s-rRNA are replete with polymorphisms. The regions of the *H. nobilis* mitochondrial genome with the highest density of polymorphisms are cytochrome B and the controls region but other regions with lower SNP density can be found in Fig. 1 and Supplemental material 2. These polymorphic regions provide additional markers to study the population genetics and phylogeography of the invasion process and for the eDNA detection of these invasive species.

Acknowledgments We would like to thank Duane Chapman, U.S. Geological Survey, for providing tissue samples and Joe Boland and Dave Roberson, National Cancer Institute, for providing access to the Ion Torrent Proton sequencing platform. In addition, we would like to acknowledge Jim Grazio of the Pennsylvania Department of Environmental Protection and Sharon Gross of the U.S. Geological Survey Ecosystems Mission Area Invasive Species Program for funding this project. We wish to thank Dolly Coykendall for providing useful comments on an earlier version of the manuscript. Use of trade, product, or firm names does not imply endorsement by the U.S. Government.

References

Berday N, Zaoui D, Lamrini A, Abi M (2005) Impact of feeding activity of silver carp on plankton removal from a high-rate pond effluent. Water Qual Res J Canada 40:191–201

Burk JS, Bayne DR, Rea H (1986) Impact of silver and bighead carps on plankton communities of channel catfish ponds. Aquaculture 55:59–68

Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) Circlize implements and enhances circular visualization in R. Bioinformatics. doi:10.1093/bioinformatics/btu393

Irons KS, Sass GG, McClelland MA, Stafford JD (2007) Reduced condition factor of two native fish species coincident with invasion of non-native Asian carps in the Illinois River, U.S.A. Is this evidence for competition and reduced fitness? J Fish Biol 71(Supplement D):258–273

Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) "Sight-unseen" detection of rare aquatic species using environmental DNA. Conserv Let 4:150–157

King TL, Eackles MS, Chapman DC (2011) Tools for assessing kinship, population structure, phylogeography, and interspecific hybridization in Asian carps invasive to the Mississippi River, USA: isolation and characterization of novel tetranucleotide microsatellite DNA loci in silver carp *Hypophthalmichthys molitrix*. Conserv Genet Res 3:397–401

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26:589–595

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009a) The sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079

Li SF, Xu JW, Yang QL, Wang CH, Chen Q, Chapman DC, Lu G (2009b) A comparison of complete mitochondrial genomes of silver carp *Hypophthalmichthys molitrix* and bighead carp *Hypophthalmichthys nobilis*: implications for their taxonomic relationship and phylogeny. J Fish Biol 74:1787–1803

Makhrov AA, Bolotov IN (2006) Dispersal routes and species identification of freshwater animals in Northern Europe: a review of molecular evidence. Russian J Genet 42:1101–1115

Patel RK, Jain M (2012) NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE 7(2):e30619. doi:10.1371/journal.pone.0030619

R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Sampson SJ, Chick JH, Pegg MA (2009) Diet overlap among two Asian carp and three native fishes in backwater lakes on the Illinois and Mississippi Rivers. Biol Invasions 11:483–496