Conservation Genetics

Population Genomics of the Peripheral Freshwater Fish Polynemus melanochir (Perciformes, Polynemidae) in a Changing Mekong Delta

--Manuscript Draft--

Manuscript Number:	COGE-D-18-00063
Full Title:	Population Genomics of the Peripheral Freshwater Fish Polynemus melanochir (Perciformes, Polynemidae) in a Changing Mekong Delta
Article Type:	Original Research
Keywords:	Vietnam; population genetics; RADSeq; river fisheries
Corresponding Author:	Kent Carpenter, PhD Old Dominion University College of Sciences Norfolk, UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Old Dominion University College of Sciences
Corresponding Author's Secondary Institution:	
First Author:	Binh Thuy Dang, PhD
First Author Secondary Information:	
Order of Authors:	Binh Thuy Dang, PhD
	Quyen Ha Dang Vu, M.S.
	Ellen E. Biesack, M.S.
	Tinh Vu Doan, M.S.
	Tran Linh Thuoc, PhD
	Amanda S. Ackiss, PhD
	Brian L. Stockwell, PhD
	Kent E. Carpenter, PhD
Order of Authors Secondary Information:	
Funding Information:	National Science Foundation (US) (0730256) — Dr. Kent E. Carpenter
	National Academy of Sciences (PEER 2-7) — Dr. Binh Thuy Dang
Abstract:	The Mekong River is a vital fisheries resource supporting millions of people in mainland Southeast Asia. However, numerous threats have the potential to negatively impact fish populations in this region including overfishing, pollution, climate change and increased urban, agriculture and upstream hydropower development. Although a few studies have examined the population genetic structure of fishes within the upper Mekong River, no known studies have explored that of fishes within the Mekong Delta (MD). Here, we examined the population structure of an important food fish within the MD Polynemus melanochir, using a panel of 1,735 single nucleotide polymorphisms (SNPs) generated by restriction site-associated DNA (RAD) sequencing across eight locations on the Tien (Mekong) and Hau (Bassac) Rivers in Vietnam. Pairwise FST values, principal component analysis and Structure analysis all indicate high levels of gene flow are occurring between the sites sampled across the MD. In contrast to the lack of genetic structure, high levels of relatedness were found, including 14 pairs of full siblings, as well as an effective population size (Ne) of less than 500 across the MD. While panmixia indicates that fragmentation of this population is not presently an important threat, a low Ne estimate suggests this species may not be resilient to long-term environmental changes in the MD. The reliance of P. melanochir as a food

	resource may be contingent on management and mitigation of low effective population sizes.

OLD DOMINION UNIVERSITY
Department of Biological Sciences, Norfolk, Virginia 23529-0266
Phone: 757-683-3595 FAX: 757-683-5283

17 March 2018

A. Rus Hoelzel, Editor
Conservation Genetics

Dear Dr. Hoelzel,

I am writing to resubmit the original research manuscript "Population Genomics of the Peripheral Freshwater Fish *Polynemus melanochir* (Perciformes, Polynemidae) in a Changing Mekong Delta" by Dang *et al.* for publication in *Conservation Genetics* as encouraged by the editorial board in a 6 Feb. 2018 editor's decision. The manuscript has undergone major revisions in content and style to address comments returned by an outside reviewer. These revisions included a complete reworking of the introduction and discussion, the addition of more life history information for the species in question, and homogenization of the writing style and wording. We also incorporated new results, obtained by allowing for inbreeding in our relatedness analyses based on the previous observation of several related pairs within the dataset.

We would like to reiterate the appropriateness of this research for *Conservation Genetics*. Our results find that although *P. melanochir* appears to form a panmictic population throughout the Vietnamese Mekong River Delta, it also maintains a small effective population size of less than 500 across the entire sampled region. The riverine habitats of the Delta are at risk from a variety of environmental threats that could impact fish populations in coming decades. This study is also the first to target this region for population genetic analysis using high-throughput methods, and it introduces both the need for and possibilities of population genetic research in the Delta.

We certify that this manuscript was produced through original and previously unpublished research and that we have no conflicts of interest to divulge. We would like to kindly suggest that Dr. G. Bernardi be the editor for this submission as he is familiar with our original submission.

Sincerely,

Kent Carpenter (Corresponding author on behalf of all co-authors)
Professor & Eminent Scholar

Population Genomics of the Peripheral Freshwater Fish *Polynemus melanochir* (Perciformes, Polynemidae) in a Changing Mekong Delta

Dang, B.T[1], Vu, Q.H.D.[1,2], Biesack, E.E.[3], Doan, T.V.[4], Truong, T.O.[1], Tran, T.L.[2], Ackiss, A.S.[3], Stockwell, B.L.[3], Carpenter K.E.[3,*]

[1] Department of Biology, Institute of Biotechnology and Environment, Nha Trang University, 02 Nguyen Dinh Chieu, Nha Trang, Vietnam

[2] Faculty of Biology – Biotechnology, University of Science, Vietnam National University Ho Chi Minh City, 227 Nguyen Van Cu Str., Dist. 5, Ho Chi Minh City, Vietnam

[3] Department of Biological Sciences, Old Dominion University, Norfolk, Virginia 23529-0266 USA

[4] Department of Software Engineering, IT Faculty, Nha Trang University, 02 Nguyen Dinh Chieu, Nha Trang, Vietnam

*Corresponding author: Kent E. Carpenter
Email: kcarpent@odu.edu
Office Phone: (757) 683-4197
ORCID iD: 0000-0003-3618-1811

Acknowledgements

We thank the many people from the communities of the Mekong River Delta who helped us collect tissues from fish markets and team members of Biodiversity and Conservation, Institute for Biotechnology and Environment, Nha Trang University for project support. We are very grateful to Christopher Bird and Sharon Furiness Magnuson who provided help and advice in sequencing and data analysis. We thank Min Dong, Wirawan Purwanto, Terry Stilwell and John Pratt of the Old Dominion University High Performance Computing Facility for computational assistance. This project was funded by the United States Agency for International Development supported Partnerships for Enhanced Research Project 2-7 under USAID Cooperative Agreement AID-OAA-A-11-00012 and project OISE-0730256 sponsored by the United States National Science Foundation.

Abstract

The Mekong River is a vital fisheries resource supporting millions of people in mainland Southeast Asia. However, numerous threats have the potential to negatively impact fish populations in this region including overfishing, pollution, climate change and increased urban, agriculture and upstream hydropower development. Although a few studies have examined the population genetic structure of fishes within the upper Mekong River, no known studies have explored that of fishes within the Mekong Delta (MD). Here, we examined the population structure of an important food fish within the MD *Polynemus melanochir*, using a panel of 1,735 single nucleotide polymorphisms (SNPs) generated by restriction site-associated DNA (RAD) sequencing across eight locations on the Tien (Mekong) and Hau (Bassac) Rivers in Vietnam. Pairwise F_{ST} values, principal component analysis and Structure analysis all indicate high levels of gene flow are occurring between the sites sampled across the MD. In contrast to the lack of genetic structure, high levels of relatedness were found, including 14 pairs of full siblings, as well as an effective population size (*Ne*) of less than 500 across the MD. While panmixia indicates that fragmentation of this population is not presently an important threat, a low *Ne* estimate suggests this species may not be resilient to long-term environmental changes in the MD. The reliance of *P. melanochir* as a food resource may be contingent on management and mitigation of low effective population sizes.

Keywords: Vietnam; population genetics; RADSeq; river fisheries

Introduction

The Mekong Delta (MD) of Vietnam is an ecosystem experiencing numerous threats whose impacts on fish populations are poorly understood. Urbanization, agricultural runoff, pollution and climate change are important factors affecting biodiversity and food security in this rice- and fish-basket region of Southeast Asia (Le et al. 2007; Allen et al. 2012; Smajgl et al. 2015; Nguyen et al. 2016). Understanding the population structure of fishes within this region may be key to mitigating the detrimental effects that various threats may have on these aquatic resources. A handful of studies have shown varying levels of population structure within and between the main branches of the Mekong in Cambodia and Lao PDR (So et al. 2006; Hurwood et al. 2008; Adamson et al. 2009; Takagi et al.2010; Takagi et al. 2011; Nguyen and Sunnocks 2012). However, no known published studies to date have examined the genetic structure of fishes within the MD.

The MD is a complex ecosystem with many factors potentially influencing the structure of fish populations. A main feature of the hydrology in the delta is the two anastomosing main branches of the Mekong River (Vo 2012). The eastern branch is typically referred to as the Mekong or the Tien within Vietnam. The Tien branches into four widespread primary outflows to the South China (East) Sea (Fig. 1). The western branch is referred to as the Bassac (mostly in Cambodia) and Hau (mostly in Vietnam) and has several close-set outflowing branches. The eastern and western regions of the delta have different underlying geological structures (Ta et al. 2002) and their flow rates differ widely in the northern delta. During the same four-month period in 2002, the discharge rate from the Mekong was between 11,000 to nearly 30,000 m^3/sec, while on the Bassac it was between about 1,000 to 4,000 m^3/sec at approximate latitudinal equivalent gauging stations in Cambodia (Dutta et al. 2007). A report by Delta Alliance (2011) also states that the ratio of total water flow between the Tien and Hau rivers is 80/20 respectively. However, water flows from the Tien to the Hau via the Vam Nao tributary north of Can Tho, Vietnam (Fig. 1), effectively equalizing the rate of flow between the two rivers downstream (Vo 2012). The heavy freshwater discharge of the Mekong River into the MD, especially during the wet season, renders it a strongly riverine estuary (Nguyuen and Savenjie 2006), but during the dry season salt water intrudes as much as 70 to 160 km inland (Noh et al. 2013; Gugliotti et al. 2017).

The Blackhand Paradise Fish, *Polynemus melanochir*, is a secondary freshwater fish that tolerates very limited saline concentrations. It is also classified as a peripheral freshwater fish because most of the species in the family are marine fishes (Motomura 2004) whose euryhaline representatives adapted to estuarine and freshwater environments. This species is common in the MD and an esteemed food fish caught by seines, trawls and set nets (Rainboth 1996 as *P. borneensis*, now considered a synonym of *P. melanochir*; Motomura 2004) although separate statistics are not reported for this species. There is no specific habitat or biological information published for this species although members of the family are known to be demersal and congeners are all freshwater species typically found in muddy waters (Motomura 2004). The very elongate pectoral-fin filaments of members of this genus are sense organs that help detect benthic invertebrates in limited visibility fresh and estuarine waters. They are caught in large numbers suggesting schooling behavior and their maximum size is reported at around 25 cm total length. Its primary Asian mainland distribution is the MD although some specimens were found upstream marginally in Cambodia and a rare, separate subspecies was found in Tonle Sap (Motomura and Sabaj 2002; Motomura and Tsukawaki 2006).

There are many threats to *P. melanochir* and overall fish biodiversity in the MD, including a rapidly growing human population, intensification of agriculture with increasing pesticide and fertilizer use, changes in flow rates due to numerous planned upstream hydropower dams and increasingly intense fishing pressure (Campbell, 2009). Climate change is expected to increase salinity intrusion dramatically over the next century with concomitant economic and ecological consequences (Cruz et al. 2007; Le et al. 2007; Dasgupta et al. 2009; Hak et al. 2016; Nguyen 2016), including impacts that would restrict the distribution of *P. melanochir* in the MD (James 2003; Koehn et al. 2011). The purpose of this study was to examine the population structure of *P. melanochir* across these two main tributaries of the MD to better understand fish population connectivity within this region and the potential impacts of human activities on the persistence of this important food fish.

Materials and Methods

Sampling sites and tissue collection

Polynemus melanochir were collected from markets located at local riversides or from fishing boats docked at markets from the Bassac (Hau) and Mekong (Tien) Rivers in the Vietnamese Mekong Delta (MD). Small, local markets supplied by short-ranging fishing boats were targeted, and all vendors were interviewed to confirm that fish were caught locally. A total of 245 individuals were sampled from upstream (An Giang, Dong Thap), midstream (Can Tho, Vinh Long, Tien Giang) and downstream localities (Tra Vinh, Ben Tre, Soc Trang; Table 1, Fig. 1). All fin clips and tissue samples were taken from fresh fish and preserved in 95% ethanol immediately after sampling.

DNA extraction and digestion

Genomic DNA was extracted from preserved tissue samples using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's instructions and treated with RNase (100 mg/mL) to remove residual RNA. Instead of a single, high volume elution, extracted DNA was eluted four separate times (100 µL each) to target higher DNA quality in the later elutions. All elutions were visualized using gel electrophoresis (1% agarose) to identify the best elution, with sharp, high molecular weight bands and no smear. The concentrations of the best elution for each sample were measured using a Qubit® 2.0 Fluorometer (Invitrogen). Samples with DNA concentrations ≥ 3 ng/µL were selected for library preparation, and 100 ng of each was then purified using AMPureXP (Agencourt) beads using a 2:1 template to bead volume ratio with the beads left in.

Purified DNA from each individual was simultaneously digested with the isoschizomeric restriction enzymes *MboI* and *Sau3AI* (New England Biolabs). Digestions were performed in 25 µL reactions consisting of 2.5 µL Cut Smart Buffer (10X), 0.5 µL *MboI* and 0.5 µL *Sau3AI* (5 unit/µL), and 21.5 µL of DNA template. Digestions were incubated at 37 °C for 3 h to overnight and then at 65 °C for 20 mins. The AMPureXP beads in the digested samples were reactivated with PEG solution (10 g PEG, 7.3 g NaCl, and water to 49 mL), and libraries were cleaned and then eluted from the beads in 20.1 µL Illumina Resuspension Buffer.

Library preparation

Cleaned libraries were inserted directly into the Illumina TruSeq Nano DNA library Prep kit following the Sample Preparation v2 Guide starting with the "Perform End Repair" step for one-third volume reactions (Supplement S1 - Toonen et al. 2013). Libraries were end repaired, 350 bp size-selected by Illumina SP beads, 3' ends were adenylated, and Illumina adapters were ligated to the digested genomic DNA sample. PCR reactions were performed in 15 µL reactions consisting of 1.5 µL Illumina PCR Primer Cocktail, 6 µL Illumina Enhanced PCR Mix, 1.875 µL ddH_2O and 5.625 µL DNA library. Reactions were carried out in thermocyclers (Icycler, Biorad) under the following temperature program: initial denaturation at 95 °C for 3 mins, followed by 8 cycles (98 °C for 20s, 60 °C for 15s, 72 °C for 30s), a final extension at 72 °C for 5 mins and a 4 °C hold. The 400 - 500 bp PCR fragments (of which 120 bp are the ligated adapters) were inspected using a 1.5% agarose gel with ethidium bromide run at 90 V for 30 mins, and bands were visualized under a UV transilluminator. PCR products were purified using SP Beads (1:1) and quantified using qPCR. DNA libraries were sequenced as paired-end 100 bp runs on a HiSeq 2500/4000 system (Illumina).

SNP discovery and filtering

Read processing was implemented using dDocent v2.0 (Puritz et al. 2014). Raw fastq files were trimmed using Trimmomatic v0.33 (Bolger et al. 2014) to simultaneously remove Illumina adapter sequences and any base that has a quality score of less than 10 (Toonen et al. 2013). Surviving forward and reverse reads were clustered and input into *de novo* reference assembly in Rainbow v2.0.2 (Chong et al. 2012) and CD-HIT v4.6.1 (Fu et al. 2012, Li and Godzik 2006) based on overall sequence similarity (90% by default). Quality-trimmed reads were mapped to the reference using BWA v0.7.12 (Li and Durbin 2009*)* with the MEM algorithm (Li 2013). Conversion from SAM to BAM files was performed using SAMTOOLS v0.1.19 (Li and Durbin 2009), and the output was further restricted to reads with mapping quality above 10.

SNP calling was performed using Freebayes v0.9.21 (Garrison and Marth 2012) with default parameters set by dDocent. Raw SNP files were concatenated into a single variant call format (VCF) file using vcftools v0.1.13 (Danecek et al. 2011). The raw SNP calls were then filtered with vcftools and vcffilter. Primary filtering steps

included: minor allele frequency (MAF > 0.05), minimum mean depth (≥ 5 mean DP ≤ 10), INDEL loci (this step decomposed insertion and deletion genotypes), Hardy-Weinberg Equilibrium (HWE with $p < 0.001$), mean quality score ($Q > 30$), max-missing (to apply a genotype call rate of 95% across all individuals), and number of variants (restricted to biallelic SNPs). Secondary filtering steps included keeping loci based on allelic balance (AB > 0.3), mean mapping quality (0.9 < MQM/MQMR < 1.05), and proportion of alternate alleles (0.05 < PAIRED/PAIREDR < 1.75). Putative SNPs were submitted to rad_haplotyper (https://github.com/chollenbeck/rad_haplotyper) to remove possible paralogs. Finally, SNP data was subjected an overall heterozygosity filter to remove loci and individuals exhibiting high heterozygosity (> 0.6) to get the validated SNP panel.

Outlier loci detection

Two methods were used to detect loci putatively under selection. The Lositan Selection Workbench (Antao et al. 2008) identified outlier loci displaying unusually high and low values of F_{ST} by comparing observed F_{ST} values with values expected under neutrality (Beaumont and Nichols 1996). Loci with F_{ST} values higher and lower than 95% of the neutral distribution were considered to be under divergent or balancing selection, respectively. An initial run was performed with 50,000 simulations and all loci, using the mean neutral F_{ST} as a preliminary value. In addition, the Bayesian approach of BayeScan v2.1 (Foll and Gaggiotti 2008) was employed to estimate the posterior probability of a given locus being under the effect of selection (Foll et al. 2008). BayeScan was run using default settings, and loci putatively under divergent selection were defined as those with a false discovery rate (FDR) < 5% and alpha-values significantly > 0 (i.e. with Q-values smaller than 0.05), while loci putatively under balancing selection had alpha-values significantly smaller than 0. Loci identified by both Lositan and BayeScan were collated into a single panel of outlier loci, which were then removed from the main SNP panel, resulting in a putatively neutral SNP dataset for analysis.

Genetic diversity, relatedness, and effective population size

Numbers of alleles (N_A), effective numbers of alleles (N_E), expected (H_e) and observed (H_o) heterozygosity, and inbreeding coefficients (G_{IS}) were calculated for each sampled population and over all populations across the MD using GenoDive v.2.0b27 (Miermans & Van Tienderen 2004).

High levels of relatedness can impact analyses of population structure and estimates of population size, so relationships between individuals were estimated with the R package 'related' (Pew et al. 2015) using the dyadic (Milligan 2003) and triadic (Wang 2007) maximum likelihood estimators and allowing for inbreeding. For both estimators 95% confidence intervals were calculated with 500 bootstrap events for each pairwise comparison. Potential clones were identified as exhibiting a related value (r) > 0.90, and clone pairs were broken by removing one individual from remaining analyses.

Estimates of effective population size (N_e) were generated with NeEstimator v2b (Do et al. 2014) using the linkage disequilibrium method, with a minor allele frequency cutoff of 0.05. Effective population size was calculated for all sampled sites individually, for the Tien (DT, VL, BT, TV) and Hau (AG, CT, ST) rivers, and for all the individuals combined into a single population.

Analyses of population structure

Pairwise comparisons of F_{ST} values between *P. melanochir* populations were computed in GenoDive with 10,000 iterations to test for significant differentiation among sampled sites. These comparisons were repeated following the removal of one individual from each full sibling pair to examine the influence of relatedness on differentiation. All p-values underwent false discover rate (FDR) correction to avoid false positives resulting from multiple comparisons (Benjamini & Hochberg 1995).

We tested for population connectivity and structure in the program Structure v2.3.4 (Pritchard et al. 2000) following the removal of one individual from each pair of full siblings. Closely related individuals may have a significant effect on clustering analyses, so these individuals were eliminated in these analyses (Goldberg and Waits 2010). Structure uses a model-based Bayesian clustering method to infer the number

of lineages, *K*, in a dataset. Structure was run to test *K* values of 1 through 8 with 10,000 iterations of burn-in followed by 5,000 Markov Chain Monte Carlo (MCMC) steps, using the correlated allele frequencies admixture model. The optimal value of *K* was evaluated using the Evanno method (Evanno et al, 2005) by Structure Harvester v0.6.94 (Earl and vonHoldt 2012).

A principal component analysis (PCA) and principal coordinate analysis (PCoAs) were performed using the R package 'adegenet' (Jombart and Ahmed 2011) following the removal of one individual from each full sibling pair. This analysis provides a graphic description of the genetic divergence among populations in multivariate space.

Historic migration rates

Historic gene flow between populations was estimated using the Bayesian inference implemented in MIGRATE-n v3.6.11 (Beerli and Felsenstein 2001). Sample sizes were reduced for each population to obtain 178 loci genotyped in 100% of individuals used for the analysis (Table S1). The run was performed using 500,000 recorded genealogies sampled every 100 steps, preceded by a burn-in of 20,000. Four hot chains were used with temperatures: T1 = 1.0, T2 = 1.5, T3 = 3.0 and T4 = 1.0×10^6. After optimization, the maximum mutation-scaled effective populations size (θ) prior was set at 0.1 while the maximum mutation-scaled migration (*M*) prior was set at 20,000. Nine hypotheses of migration among populations were tested: (1) symmetric migration rates between all sites (Panmixia Model), (2) non-symmetric migration rates between all sites (Full Model) (3) migration between all sites within each of the rivers (Hau and Tien), but no migration between rivers (Rivers Separate), (4) migration occurring only between neighboring, downstream sites and between rivers (Downstream Open), (5) migration occurring only between neighboring, downstream sites but no migration between rivers (Downstream Closed), (6) migration occurring only between neighboring, upstream sites and between rivers (Upstream Open), (7) migration occurring only between neighboring, upstream sites but no migration between rivers (Upstream Closed), (8) migration occurring among all sites found in each river, however migration only occurs from the Tien River sites to Hau River

289 sites (Tien Source), (9) migration occurring among all sites found in each river,
290 however migration only occurs from the Hau River sites to Tien River sites (Hau
291 Source). The most likely model was chosen using the Bezier approximation score
292 produced by Migrate-n and migrants per generation for the chosen model were
293 calculated according to Beerli (2009).

294

295 **Results**

296 *Reference assembly and SNP filtering*

297 RAD sequencing efforts for *Polynemus melanochir* generated a total of 358,189,296
298 paired-end, 101 bp reads, which, when filtered and aligned to create a catalogue,
299 resulted in a total of 82,116 RAD tags used to generate a >4X coverage *de novo*
300 reference of 20,385,313 bp.

301 Using the Freebayes tools for SNP calling, the initial dataset consisted of 459,374 raw
302 SNPs. Following extensive filtering, the final dataset consisted of 1738 putative SNPs
303 in 184 individuals. Information on individuals removed at each step of filtering and
304 data analysis is presented in Table S1**.**

305

306 *Outlier loci detection*

307 BayeScan identified three SNPs as outliers (FDR ≤ 0.05) from the panel of 1738
308 putative SNPs used to detect selection footprints. Lositan identified five SNPs as
309 candidates for positive selection (F_{ST} simulated < F_{ST} sample), three of which were
310 the same loci identified by BayeScan. In addition, Lositan identified 38 SNPs as
311 candidates for balancing selection, however none of the loci identified by Lositan as
312 candidates for either balancing or divergent selection survived FDR correction. The
313 three loci detected as outliers putatively under positive selection by BayeScan (Fig.
314 S1**)** were removed from the SNP panel and the 1735 remaining loci were assumed to
315 be neutral.

316

317 *Genetic diversity, relatedness, and effective population size*

318 Across eight sampled populations, *P. melanochir* showed average levels of observed
319 and expected heterozygosity of 0.303 ± 0.003 and 0.344 ±0.003, respectively.

Observed heterozygosity within sites ranged from 0.274 (ST) to 0.344 (TG), and expected from 0.341 (TV) to 0.345 (BT). The average number of alleles was 2, as constrained by filtering, and effective number of alleles 1.554 ± 0.007 (Table 1). Inbreeding coefficients ranged from 0.054 (DT) to 0.201 (ST), with an overall GIS for all individuals at 0.117 (Table 1).

Analyses of genetic relationships between individuals revealed 14 pairs of putative full siblings and two pairs of putative cousins (Table 2) following clone removal (Table S1). Full siblings occurred most abundantly within the DT and TV populations, with the other sibling pairs occurring within CT and BT and between sites separated by up to more than 200 km (Fig. 1). Cousin pairs were also found within DT and between CT and TV (Table 2).

Estimates of N_e ranged from 18.3 (DT) to infinite (TG; Table S2). N_e was estimated at 461.8 for all individuals across all sites as a single population, and at 285.5 and 263.2 for the Hau and Tien rivers, respectively (Table 3).

Analyses of population structure

Pairwise F_{ST} comparisons of the geographically defined populations were very small (≤ 0.012) but often significant (19 out of a total of 28 comparisons). All significant comparisons retained significance after FDR correction. Among Hau River sites, the downstream site ST was the most divergent from other sites, with all comparisons with other sites significant (Table 4; Fig. 1). Of Tien River sites, the upstream site DT was the most divergent from other sites, with all comparisons with other sites significant, followed by the downstream site TV, with six significant of seven total comparisons (Table 4; Fig. 1). All pairwise comparisons between sites after breaking sibling pairs were not significant (Table S3).

Clustering analysis in the program Structure detected no population differentiation among 8 sampling sites, with a *K* value of 3 determined by the Evanno method (Fig. 2). The majority of individuals were assigned with high probably to a single cluster. The four individuals with higher percentages of the additional two clusters contained

some of the lowest values of heterozygosity and higher amounts of missing data. The principal component analysis (PCA) revealed a similar lack of population differentiation as revealed by Structure (Fig. 3)**.** The first axis explains 1.10% of the variation and the second axis explains an additional 1.07%.

Results from a principal coordinate analysis (PCoA) showed the populations from upstream and midstream (AG, DT, CT, and TG) closely grouped with the estuarine site of the Tien River (BT). The downstream populations TV and ST are the most distinct from other sampled sites in the MD. The first axis accounts for 16.9% of the variation and the second axis accounts for an additional 15.9%. However, the dimensions of the PCoA are incredibly small (d = 0.05) so drawing any conclusions about relative placement of sampling localities is not possible.

Historic migration rates

To examine the historic migratory patterns of *P. melanochir* across the MD, we tested nine different models in Migrate-n, which allowed fish to move in and out Hau and Tien Rivers, as well as up- and downstream. Results showed the 'Rivers Separate model was the most supported (Bezier approximation score of -644468; Table S4), in which bidirectional migration was maintained among all sites within each river but not between rivers. Mean numbers of migrants per generation (*M* / gen.) along each pathway for the Rivers Separate model are presented in Table 5. Migration within the Hau River ranged from one to three migrants per generation and migration within the Tien River ranged from two to five migrants per generation (Table 5). Migration levels are seemingly chaotic, with no direction in particular (downstream or upstream) supporting more migration.

Discussion

Results indicate that *Polynemus melanochir* maintains a single nearly panmictic population in the Vietnamese Mekong Delta (MD) with a low overall effective population size. Many of the pairwise F_{ST} comparisons among sampled sites were significant but these values were very low, and Structure, PCA and PCoA analyses showed no regional clustering of sites. This indicates that *P. melanochir* is able to

384 migrate freely between rivers or that reproduction may occur upstream of the Vam
385 Nao tributary where active flow from the Tien to the Hau occurs. The few species of
386 Polynemidae for which information is available, including one species of *Polynemus*,
387 are broadcast spawners. Assuming *P. melanochir* also uses this reproductive strategy,
388 then passive dispersal would mostly be downstream, and panmixia could have
389 occurred through migration within each of the respective main branches and egg
390 dispersal through the Vam Nao tributary, or adult migration between rivers. The
391 swimming capability of these demersal fishes is unknown but upstream swimming to
392 some extent during high flow periods should be possible and even more feasible
393 during the low flow periods of the dry season. The species may also inhabit and travel
394 through the many canals that connect these two river systems (Vo 2012). The
395 Migrate-n analysis indicates predominant migration within but not between the Tien
396 and Hau rivers, however these historic patterns were not detectable in other analyses
397 and may be becoming less important to population structure with contemporary canal
398 migration and other man-induced changes in flow patterns. The equalization of
399 downstream flow below the Vam Nao suggests that fish have the same ability to
400 navigate upstream in both the Tien and Hau, and this may explain the similar lack of
401 genetic differentiation within both rivers.
402
403 The small F_{ST} values and significant differences among populations that disappear if
404 siblings are removed from calculations may reflect fairly high rates of localized
405 reproduction and a low overall effective population size. The presence of full siblings
406 and cousins in half of the sites sampled indicates that local recruitment originates
407 from a limited pool of successful reproductive adults. The effective population size
408 calculated with all sites combined is less than the lower limit of 500 considered
409 sufficiently resilient to long-term population changes (Jamieson and Allendorf 2012).
410 The effective population sizes per site were even smaller (less than 40 for sites
411 containing sibling pairs), but since population structure analyses provide strong
412 evidence for a single panmictic population, these calculations are for analytically
413 forced ‘populations,’ and an estimate of population size from all sites combined is
414 more reliable. The presence of high levels of relatedness within sample sites, which
415 will strongly affect N_e estimates, may also be a result of unintentional sampling of
416 schools by fishers, which may have a higher likelihood of containing siblings. If this
417 species is a broadcast spawner, it is likely that fish sampled across all sites in the MD

418 are randomly mixed. In this case, the presence of related pairs within sites would be
419 further evidence of an overall small population size.
420
421 The management implications of our findings have both promising and alarming
422 components. Panmixia in the population indicates that fragmentation is limited and
423 that localized threats may not harm the overall population. However, the low
424 estimated effective population sizes are worrisome in that they indicate this species
425 may not be resilient to long-term changes in the MD environment (Jamieson and
426 Allendorf 2012). The minimum viable population of 500 to maintain evolutionary
427 potential and hence ability to adapt to changing environments is above the effective
428 population size detected in *P. melanochir* for the entire sampled region. Therefore
429 management of this species should be careful to ensure that overfishing does not
430 further erode the effective population size of the existing population. More should
431 also be done to understand the effects of pollution and other threats to this species to
432 avoid pushing it further below viable population sizes and if necessary, mitigate
433 existing threats.
434
435 This study is one of the first studies on population genetics of MD fishes and its
436 findings introduce many potential future questions to be addressed. For example, this
437 study did not have results that directly address questions of the influence increasing
438 salinity regimes may have on populations of freshwater fishes in the MD. The average
439 elevation above sea level for the MD is 0.3 to 0.7 m and sea level rise between 1985
440 and 2010 was measured at 3 mm per year with predictions that the entire region will
441 be inundated with salt water in less than a century (Hak et al. 2016). This will
442 undoubtedly threaten freshwater fishes such a *P. melanochir* that predominate in the
443 region. Low heterozygosity levels in two of the three sites sampled closest to the sea
444 (Table 1) give a faint but inconclusive hint that selection may be occurring at some of
445 these sites, and the high salinity levels at these sites (Vo 2012) may be a selective
446 pressure factor. A comparison of upstream and downstream sites using a greater
447 coverage of the genome may show divergent selection if genes involved in
448 osmoregulation can be sampled. Transcriptomic studies examining differences in gene
449 expression relative to salinity tolerance in upstream and downstream populations of
450 fishes would also be informative in light of predicted increases in saltwater intrusion
451 into the MD. Most importantly, the increasing threats and predicted changes in the

region require a more comprehensive view of the natural history of resident fishes. Managing and mitigating threats to these ecologically and economically important components of delta biodiversity will require a greater understanding of the population structure of a wide range of representative habitat specialists and generalists. This study of *P. melanochir* demonstrates that important ecological information of MD fishes can be gained from conservation genetic studies using advanced genomics.

459 **References**
460
461 Adamson EA, Hurwood DA, Baker AM, Mather PB (2009) Population subdivision in
462 Siamese mud carp *Henicorhynchus siamensis* in the Mekong River basin:
463 implications for management. J Fish Biol 75(6):1371-92

464 Allen DJ, Smith KG, Darwall WR (2012) The status and distribution of freshwater
465 biodiversity in Indo-Burma. IUCN, Gland, Switzerland

466 Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a
467 workbench to detect molecular adaptation based on a Fst-outlier
468 method. BMC Bioinformatics 9:323

469 Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of
470 population structure. P Roy Soc Lond B Bio 263:1619–1626

471 Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix
472 and effective population size in *n* subpopulations by using a coalescent
473 approach. P Natl Acad Sci USA 98:4563–4568

474 Benjamini Y, Hochberg Y (1995). Controlling the false discovery rate: a practical and
475 powerful approach to multiple testing. J R Stat Soc Series B 57(1): 289–300.

476 Bolger AM, Lohse M and Usadel B (2014) Trimmomatic: A flexible trimmer for
477 Illumina sequence data. Bioinformatics 30(15): 2114–2120

478 Campbell, IC (2012). Biodiversity of the Mekong Delta. In The Mekong Delta
479 System (pp. 293-313). Springer Netherlands.

480 Chong Z, Ruan J and Wu CI (2012) Rainbow: an integrated tool for efficient
481 clustering and assembling RAD-seq reads. Bioinformatics 28(21): 2732–2737

482 Cruz RV, Harasawa H, Lal M, Wu S, Anokhin Y, Punsalmaa B, Honda Y, Jafari M,
483 Li C, Huu Ninh N. Asia in Parry M L, Canziani OF, Palutikof JP, van der
484 Linden PJ and Hanson CE eds Climate change 2007: impacts, adaptation and
485 vulnerability. Contribution of Working Group II to the fourth assessment
486 report of the Intergovernmental Panel on Climate Change, pp. 469–506.
487 Cambridge University Press, Cambridge, UK

488 Csilléry K, Johnson T, Beraldi D, Clutton-Brock T, Coltman D, Hansson B, Spong G,
489 Pemberton M (2006) Performance of marker-based relatedness estimators in
490 natural populations of outbred vertebrates. Genetics. 173, 2091-2101.

491 Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE,
492 Lunter G, Marth GT, Sherry ST, McVean G and Durbin R (2011) The variant
493 call format and VCFtools. Bioinformatics 27(15): 2156–2158

494 Dasgupta S, Laplante B, Meisner C, Wheeler D, Yan J (2009) The impact of sea level
495 rise on developing countries: a comparative analysis. Climatic Change 93:379-
496 88

497 Delta Alliance (2011) Mekong Delta water resources assessment studies. Vietnam-
498 Netherlands Mekong Delta Materplan project. Deltares, Delft, The

499 Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ and Ovenden JR (2014)
500 NeEstimator v2: re-implementation of software for the estimation of
501 contemporary effective population size (N_e) from genetic data. Mol Eco
502 Resour 14:209-214

503 Dutta D, Alam J, Umeda K, Hayashi M, Hironaka S (2007) A two-dimensional
504 hydrodynamic model for flood inundation simulation: a case study in the
505 lower Mekong river basin. Hydrol Process 21:1223-37

506 Earl DA and vonHoldt BM (2012) STRUCTURE HARVESTER: a website and
507 program for visualizing STRUCTURE output and implementing the Evanno
508 method. Conserv Genet Resour 4(2):359–361

509 Evanno G, Regnaut S and Goudet J (2005) Detecting the number of clusters of
510 individuals using the software structure: a simulation study. Mol Ecol
511 14(8):2611–2620

512 Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0). An integrated
513 software package for population genetics data analysis. Evol Bioinform 1:47-
514 50

515 Foll M and Gaggiotti O (2008) A genome-scan method to identify selected loci
516 appropriate for both dominant and codominant markers: a Bayesian
517 perspective. Genetics 180(2):977–993

518 Fu F, Tarnita CE, Christakis NA, Wang L, Rand DG and Nowak MA (2012)
519 Evolution of in-group favoritism. Sci Rep-UK 2:1–6

520 Garrison E and Marth G (2012) Haplotype-based variant detection from short-read
521 sequencing. arXiv Preprint arXiv:1207.3907

522 Goldberg CS, Waits LP (2010) Quantification and reduction of bias from sampling
523 larvae to infer population and landscape genetic structure. Mol Ecol
524 Resour 10:304–13

525 Guo B, DeFaveri J, Sotelo G, Nair A, Merilä J (2015) Population genomic evidence
526 for adaptive differentiation in Baltic Sea three-spined sticklebacks. BMC Biol
527 13:19

528 Gugliotta M, Saito Y, Nguyen VL< Ta TK, Nakashima R, Tamura T, Uehara K,
529 Katsuki K, Yamaoto S (2017) Process regime, salinity, morphological, and
530 sedimentary trends along the fluvial to marine transition zone o fthe mixed-
531 energy Mekong River dealta, Vietnam. Continental Shelf Research 147:7-26

532 Hak D, Nadaoka K, Bernado LP, Le Phu V, Quan NH, Toan TQ, Trung NH, Van Ni
533 D, Van PD (2016) Spatio-temporal variations of sea level around the Mekong
534 Delta: their causes and consequences on the coastal environment. Hydrol Res
535 Lett 10:60-6

536 Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-
537 generation RAD sequencing identifies thousands of SNPs for assessing
538 hybridization between rainbow and westslope cutthroat trout. Mol
539 Ecol Resc. 11(s1):117-22

540 Hurwood DA, Adamson EA, Mather PB (2008) Evidence for strong genetic structure
541 in a regionally important, highly vagile cyprinid (*Henicorhynchus lobatus*) in
542 the Mekong River Basin. Ecol Freshw Fish 17(2):273-83

543 James KR, Cant B, Ryan T (2003) Responses of freshwater biota to rising salinity
544 levels and implications for saline water management: a review. Aust J Bot
545 51:703-13

546 Jamieson, IG, & Allendorf, FW (2012). How does the 50/500 rule apply to
547 MVPs?. *Trends in Ecology & Evolution*, *27*(10), 578-584.

548 Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-
549 wide SNP data. Bioinformatics 27: 3070-3071
550
551 Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction
552 on measures of allelic richness. Mol Ecol Notes 5: 187–189
553 Kalinowski ST, Wagner AP, and Taper ML (2006) ML-RELATE: a computer
554 program for maximum likelihood estimation of relatedness and relationship.
555 Mol Ecol Notes 6: 576–579
556 Koehn JD, Hobday AJ, Pratchett MS, Gillanders BM (2011) Climate change and
557 Australian marine and freshwater environments, fishes and fisheries: synthesis
558 and options for adaptation. Mar Freshwater Res 62:1148-64
559 Le AT, Chu TH, Miller F, Bach TS (2007) Flood and salinity management in the
560 Mekong Delta, Vietnam. In Be, TT, Sinh BT, Miller F (eds) Challenges to
561 sustainable development in the Mekong Delta: Regional and national policy
562 issues and research needs. Bangkok, Thailand
563 Le TV, Nguyen HN, Wolanski E, Tran TC, Haruyama S (2007) The combined impact
564 on the flooding in Vietnam's Mekong River delta of local man-made
565 structures, sea level rise, and dams upstream in the river catchment. Estuar
566 Coast Shelf S 71:110-6
567 Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with
568 BWA-MEM. arXiv Preprint arXiv:1303.3997
569 Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-
570 Wheeler transform. Bioinformatics 25(14):1754–1760
571 Li W and Godzik A (2006) Cd-hit: A fast program for clustering and comparing large
572 sets of protein or nucleotide sequences. Bioinformatics 22(13):1658–1659
573 Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular
574 markers. Genetics 152(4):1753–1766
575 Meirmans PG and Van Tienderen PH (2004) GENOTYPE and GENODIVE: two
576 programs for the analysis of genetic diversity of asexual organisms. Mol Ecol
577 Notes 4:792-794
578 Milligan BG (2003) Maximum-likelihood estimation of relatedness. Genetics 163:
579 1153-1167
580 Motomura H (2004) FAO species catalogue. Polynemid fishes of the world (family
581 Polynemidae). An annotated and illustrated catalogue of polynemid species
582 known to date. FAO, Rome
583 Motomura H, Sabaj MH (2002) A new subspecies, *Polynemus melanochir dulcis*,
584 from Tonle Sap Lake, Cambodia, and redescription of *P. m. melanochir*
585 Valenciennes in Cuvier and Valenciennes, 1831 with designation of a neotype.
586 Ichthyol Res 49:181-90
587 Motomura H, Tsukawaki S (2006) New species of the threadfin genus Polynemus
588 (Teleostei: Polynemidae) from the Mekong River basin, Vietnam, with
589 comments on the Mekong species of *Polynemus*. Raffles B Zool 54:459-64
590 Nguyen AD, Savenije HH (2006) Salt intrusion in multi-channel estuaries: a case
591 study in the Mekong Delta, Vietnam. Hydrol Earth Syst Sc Discuss 10:743-54

592 Nguyen TT, Sunnucks P (2012) Strong population genetic structure and its
593 management implications in the mud carp *Cirrhinus molitorella*, an
594 indigenous freshwater species subject to an aquaculture and culture-based
595 fishery. J Fish Biol 80(3):651-68

596 Nguyen LA, Verreth JA, Leemans HB, Bosma RH, De Silva S (2016) A Decision
597 Tree Analysis to Support Potential Climate Change Adaptations of Striped
598 Catfish (*Pangasianodon hypophthalmus* Sauvage) Farming in the Mekong
599 Delta, Vietnam. Tropicultura 34(Special):105-15

600 Nguyen N (2017) Historic drought and salinity intrusion in the Mekong Delta in
601 2016: Lessons learned and response solutions. Vietnam Science and
602 Technology 1:93-6

603 Noh S, Choi M, Kim E, Dan NP, Thanh BX, Van Ha NT, Sthiannopkao S, Han S
604 (2013) Influence of salinity intrusion on the speciation and partitioning of
605 mercury in the Mekong River Delta. Geochimica et Cosmochimica Acta
606 106:379-90

607 Pavey SA, Gaudin J, Normandeau E, Dionne M, Castonguay M, Audet C, Bernatchez
608 L (2015) RAD sequencing highlights polygenic discrimination of habitat
609 ecotypes in the panmictic American eel. Current Biology 25(12):1666-71.

610 Peakall R and Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel.
611 Population genetic software for teaching and research-an update.
612 Bioinformatics 28(19):2537–2539

613 Pew J, Muir PH, Wang J and Frasier TR (2015) Related: an R package for analysing
614 pairwise relatedness from codominant molecular markers. Mol Ecol Resour
615 15:557–561

616 Pritchard JK, Stephens M and Donnelly P (2000) Inference of Population Structure
617 Using Multilocus Genotype Data. Genetics 155(2):945-959

618 Pujolar JM, Jacobsen MW, Als TD, Frydenberg J, Munch K, Jónsson B, Jian JB,
619 Cheng L, Maes GE, Bernatchez L, Hansen MM (2014) Genome-wide
620 single-generation signatures of local selection in the panmictic European eel.
621 Mol Ecol 23:2514-28

622 Puritz JB, Hollenbeck CM, Gold JR (2014) *dDocent*: a RADseq, variant-calling
623 pipeline designed for population genomics of non-model organisms. PeerJ 2:
624 e431

625 Smajgl A, Toan TQ, Nhan DK, Ward J, Trung NH, Tri LQ, Tri VP, Vu PT (2015)
626 Responding to rising sea levels in the Mekong Delta. Nat Clim Change 5:167

627 So N, Van Houdt JK, Volckaert FA (2006) Genetic diversity and population history
628 of the migratory catfishes Pangasianodon hypophthalmus and Pangasius
629 bocourti in the Cambodian Mekong River. Fisheries Sci 72(3):469-76

630 Ta TK, Nguyen VL, Tateishi M, Kobayashi I, Tanabe S, Saito Y (2002) Holocene
631 delta evolution and sediment discharge of the Mekong River, southern
632 Vietnam. Quaternary Sci Rev 21:1807-19

633 Takagi AP, Ishikawa S, Nao T, Limsong S, Hort S, Thammavong K, Saphakdy B,
634 Phomsouvanhm A, Nishida M, Kurokura H (2011) Population structure of the
635 climbing perch, *Anabas testudineus*, in the lower Mekong River basin.
636 Fisheries Manag Ecol 18(2):145-53

637 Takagi AP, Ishikawa S, Nao T, Song SL, Hort S, Thammavong K, Saphakdy B,
638 Phomsouvanhm A, Nishida M, Kurokura H (2010) Genetic differentiation and
639 distribution routes of the bronze featherback *Notopterus notopterus*
640 (Osteoglossiformes: Notopteridae) in Indochina. Biol J Linn Soc 101(3):575-
641 82

642 Vo, KT (2012). Hydrology and hydraulic infrastructure systems in the Mekong Delta,
643 Vietnam. In *The Mekong Delta System* (pp. 49-81). Springer Netherlands.

644 Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR and
645 Bird CE (2013) ezRAD: a simplified method for genomic genotyping in non-
646 model organisms. PeerJ 19(1):e203

647 Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise
648 relatedness. Genet Res 89(03):135–153

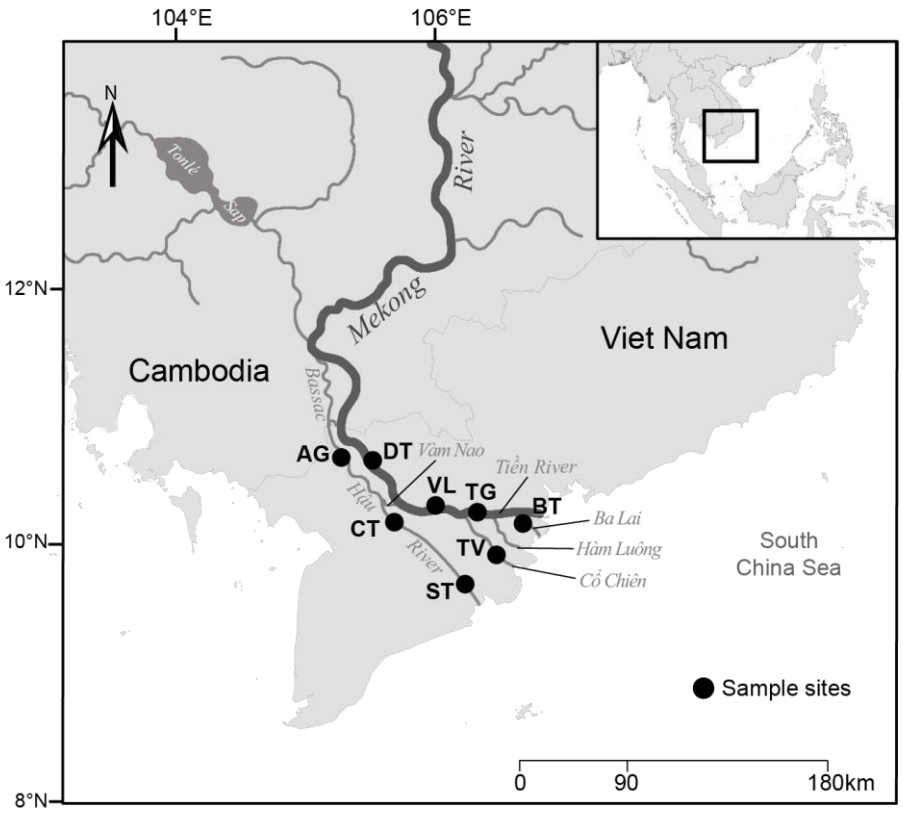

Fig. 1: Sampling map of *Polynemus melanochir* in Mekong Delta, Vietnam. Sampling sites: An Giang (AG), Dong Thap (DT), Can Tho (CT), Vinh Long (VL), Tien Giang (TG), Ben Tre (BT), Tra Vinh (TV), and Soc Trang (ST). INSET: Sampling region (black box) within Southeast Asia.

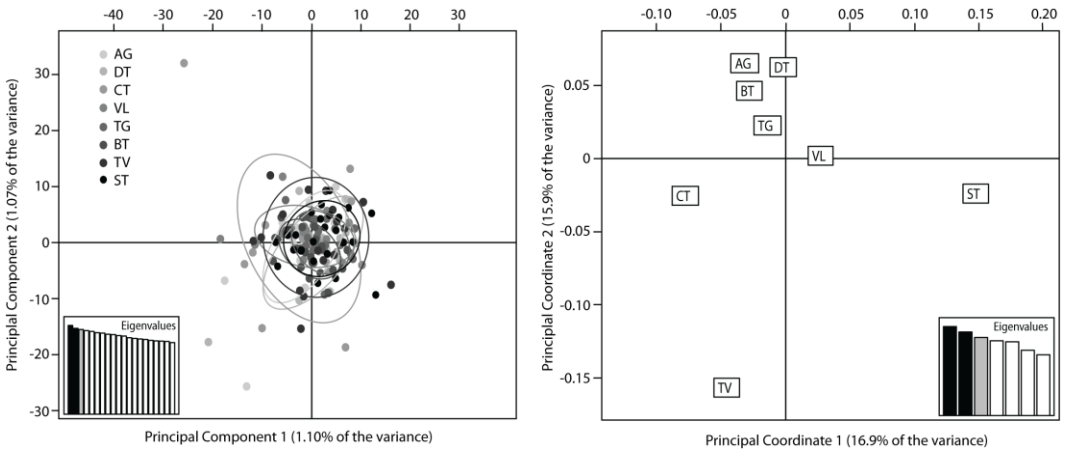

Fig. 2: Principal component (PCA, left) and principal coordinate (PCoA, right) analyses for *Polynemus melanochir* using neutral loci (related individuals removed). Insets show eigenvalues for the first 20 axes of the PCA and all axes for the PCoA.

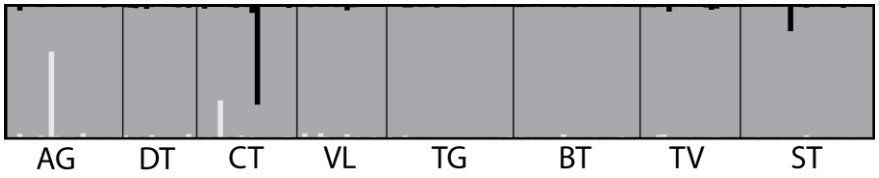

Fig. 3: Bar plot of *Polynemus melanochir* showing individual assignments to inferred clusters using the neutral SNP panel in the program STRUCTURE. Each genotype is represented by a single vertical bar partitioned into segments representing the estimated membership fractions in *K* clusters (optimal *K* = 3)

Table 1. *Polynemus melanochir* sample site information and genetic diversity in the Vietnamese Mekong Delta, including number of samples collected (***N***), number of alleles (N_A), effective number of alleles (N_E), observed and expected heterozygosity (***Ho/He***), inbreeding coefficient (G_{IS}), and the percent of polymorphic loci.

River location	Sampling sites and population code	Geographic coordinates		*N*	N_A	N_E	*Ho*	*He*	G_{IS}	% polymorphic SNPs
Upstream	An Giang (AG)	10°28'N	105°12'E	26	1.987	1.569	0.299	0.344	0.133	97.44
	Dong Thap (DT)	10°35'N	105°36'E	26	1.979	1.567	0.325	0.343	0.054	97.2
Midstream	Can Tho (CT)	10°02'N	105°45'E	40	1.989	1.564	0.282	0.343	0.178	98.28
	Vinh Long (VL)	10°06'N	106°01'E	25	1.979	1.562	0.296	0.343	0.137	97.74
	Tien Giang (TG)	10°26'N	106°43'E	26	1.997	1.574	0.344	0.345	0.003	98.23
Downstream	Ben Tre (BT)	10°08'N	106°29'E	33	1.994	1.572	0.326	0.345	0.056	98.54
	Tra Vinh (TV)	09°47'N	106°20'E	33	1.986	1.561	0.280	0.341	0.180	98.53
	Soc Trang (ST)	9°32'N	105°56'E	36	1.992	1.565	0.274	0.343	0.201	98.92
Overall	-	-	-	245	2.000	1.554	0.303	0.344	0.117	98.11

Table 3: Pairs of putative siblings of *Polynemus melanochir* from relatedness analysis. Coefficients of relatedness (r) with 95 % confidence intervals in parentheses are presented for both the **dyadml** likelihood estimator (Wang, 2007) and the **trioml** likelihood estimator (Milligan, 2003). The most likely relationship for each pair is also shown.

Specimen Pairs	**Groupings**	**Dyadml**	**Trioml**	**Relationship**
TV_12/TV_44	TV-TV	**0.883** (0.864 - 0.901)	**0.883** (0.864 - 0.901)	Full sibling
TV_04/TV_12	TV-TV	**0.873** (0.853 - 0.895)	**0.873** (0.8.52 - 0.893)	Full sibling
DT_12/ TG_26	DT-TG	**0.859** (0.838 - 0.877)	**0.859** (0.838 - 0.878)	Full sibling
DT_50/CT_09	DT-CT	**0.836** (0.815 - 0.858)	**0.836** (0.811 - 0.857)	Full sibling
TV_01/TV_42	TV-TV	**0.829** (0.805 - 0.852)	**0.829** (0.807 - 0.851)	Full sibling
CT_11/CT_42	CT-CT	**0.808** (0.782 - 0.833)	**0.808** (0.784 - 0.834)	Full sibling
TV_04/TV_44	TV-TV	**0.777** (0.750 - 0.801)	**0.777** (0.749 - 0.805)	Full sibling
DT_15/DT_59	DT-DT	**0.718** (0.689 - 0.752)	**0.718** (0.688 - 0.746)	Full sibling
DT_02/ DT_04	DT-DT	**0.689** (0.660 - 0.721)	**0.689** (0.656 - 0.724)	Full sibling
CT_42/CT_46	CT-CT	**0.679** (0.651 - 0.710)	**0.679** (0.649 - 0.708)	Full sibling
DT_48/DT_56	DT-DT	**0.661** (0.638 - 0.692)	**0.661** (0.633 - 0.690)	Full sibling
CT_11/CT_46	CT-CT	**0.595** (0.557 - 0.632)	**0.595** (0.558 - 0.634)	Full sibling
DT_03/DT_05	DT-DT	**0.553** (0.519 - 0.588)	**0.553** (0.521 - 0.587)	Full sibling
BT_17/ BT_51	BT-BT	**0.505** (0.467 - 0.539)	**0.505** (0.468 - 0.549)	Full sibling
DT_05/ DT_31	DT-DT	**0.148** (0.100 - 0.195)	**0.148** (0.101 - 0.193)	Cousin
CT_46/TV_17	CT-TV	**0.128** (0.08 7 -0.169)	**0.128** (0.080 - 0.171)	Cousin

Table 4: Estimates of *Polynemus melanochir* effective population size (N_e) calculated from 1,735 neutral SNPs. For the "all sites" datasets, all individuals from each sample site were included. The "Hau River" dataset included only samples collected from An Giang, Can Tho and Soc Trang and the "Tien River" dataset included only samples from Dong Thap, Vinh Long, Tien Giang, Ben Tre, and Tra Vinh. Sample sizes (***N***) are presented in parentheses with the name of each analysis and 95% confidence intervals are presented in parentheses with estimates of N_e.

Data Set (*N*)	***N*$_e$ (95% CIs)**
Hau River (69)	**286** (276 – 296)
Tien River (107)	**263** (258 – 269)
All Sites (n = 176)	**462** (452 – 472)

Table 2. Pairwise values of F_{ST} (below the diagonal) and their respective p-values (above the diagonal) of *Polynemus melanochir*. Bold values indicate significant differences between populations ($p < 0.034$ as corrected by FDR).

Pop	AG	DT	CT	VL	TG	BT	TV	ST
AG	-	**0.0004**	**0.0292**	0.3404	0.6540	0.9064	**0.0094**	**0.0291**
DT	**0.0075**	-	**0.0135**	**0.0020**	**0.0117**	**0.0001**	**0.0005**	**0.0001**
CT	**0.0044**	**0.0083**	-	0.0909	0.0688	**0.0044**	**0.0083**	**0.0022**
VL	0.0004	**0.0080**	0.0040	-	0.2373	**0.0311**	0.0530	0.3979
TG	-0.0003	**0.0050**	0.0030	0.0007	-	0.3220	**0.0038**	**0.0148**
BT	-0.0013	**0.0079**	**0.0050**	**0.0022**	0.0004	-	**0.0017**	**0.0018**
TV	**0.0059**	**0.0120**	**0.0081**	0.0057	**0.0057**	**0.0064**	-	**0.0030**
ST	**0.0018**	**0.0081**	**0.0054**	0.0003	**0.0018**	**0.0028**	**0.0064**	-

Table 6: Estimates of mutation-scaled migration (***M***) and mutation-scaled effective population size (***θ***) for each sampled population of *Polynemus melanochir*, between each site for the Separate Rivers Model of migration, and number of migrants per generation (***M* / gen.**) for each migration pathway (source-sink), calculated by the formula (*θ*M*)/4.

River	Source	Sink	μ scaled migration (*M*)	μ scaled N_e (*θ*)	*M* / gen.
Hau	CT	AG	19740	0.00057	3
	ST	AG	19367	0.00057	3
	AG	CT	19260	0.00057	3
	ST	CT	5060	0.00057	1
	AG	ST	19727	0.00057	3
	CT	ST	19767	0.00057	3
Tien	DT	BT	15647	0.00103	5
	TG	BT	15353	0.00103	4
	TV	BT	15313	0.00103	4
	VL	BT	14807	0.00103	4
	BT	DT	14727	0.00097	4
	TG	DT	15407	0.00097	4
	TV	DT	15300	0.00097	4
	VL	DT	13753	0.00097	4
	BT	TG	15047	0.00077	3
	DT	TG	15753	0.00077	4
	TV	TG	14820	0.00077	3
	VL	TG	13860	0.00077	3
	BT	TV	15633	0.00003	3
	DT	TV	15847	0.00003	3
	TG	TV	16407	0.00003	3
	VL	TV	13500	0.00003	3
	BT	VL	15673	0.00117	5
	DT	VL	16380	0.00117	5
	TG	VL	16087	0.00117	5
	TV	VL	15380	0.00117	5

Supplementary Material

Click here to view linked References

Click here to access/download

Supplementary Material

SupplementalInformation_DangBT.docx