Conservation Genetics
https://doi.org/10.1007/s10592-019-01189-x

RESEARCH ARTICLE

Population genomics of the peripheral freshwater fish *Polynemus melanochir* (Perciformes, Polynemidae) in a changing Mekong Delta

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

Received: 19 March 2018 / Accepted: 24 April 2019

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 examine the population structure of an important food fish within the MD, *Polynemus melanochir*, using a panel of 1735 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 among the sites sampled across the MD. In contrast to the lack of genetic structure, high levels of relatedness were found, including 26 putatively related pairs, 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 on *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. 2007a; 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.

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

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10592-019-01189-x) contains supplementary material, which is available to authorized users.

✉ K. E. Carpenter
kcarpent@odu.edu

[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, VA 23529-0266, USA

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

Published online: 29 April 2019

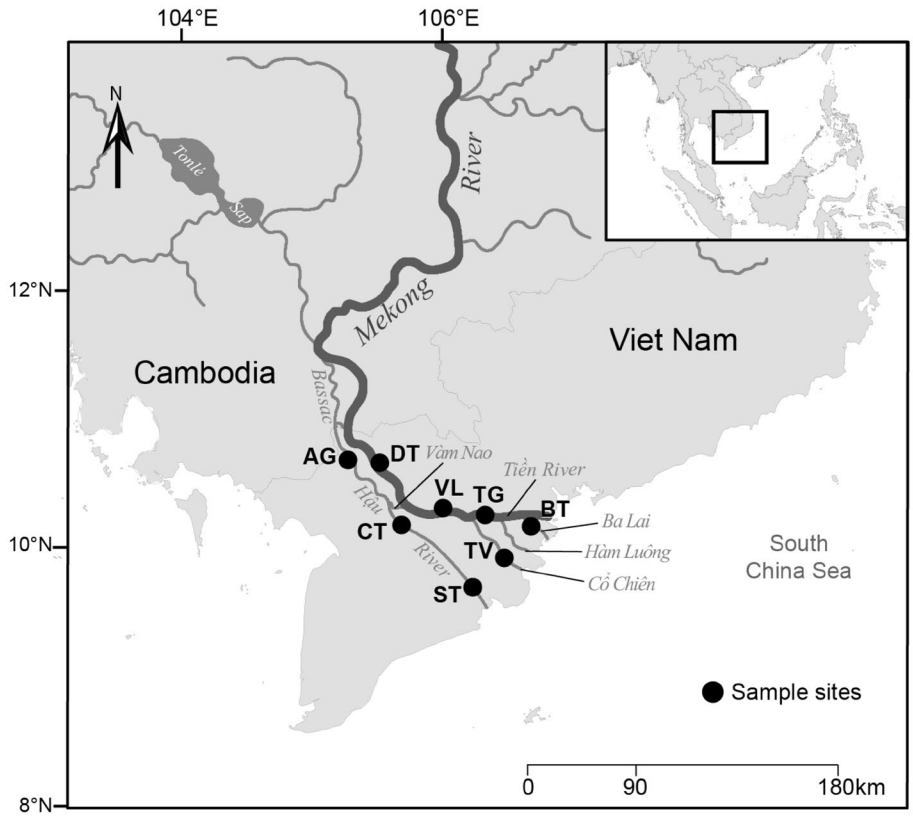

Fig. 1 Sampling map of *Polynemus melanochir* in Mekong Delta, Vietnam. Sampling sites: An Giang (AG), Dong Thap (DT), Can Tho (CT), Vinh Long (VL), Tien Giang (TG), Ben Tre (BT), Tra Vinh (TV), and Soc Trang (ST). *Inset* Sampling region (black box) within Southeast Asia

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/s, while on the Bassac it was between about 1000 to 4000 m^3/s 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 (Nguyen and Savenije 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,* was previously reported from the region as *P. borneensis* (Rainboth 1996), but this name is now considered a junior synonym of *P. melanochir* (Kottelat 2013). In the Asian mainland, this species is found primarily in the MD although some specimens were found upstream marginally in Cambodia. An extremely rare subspecies, *Polynemus melanochir dulcis*, restricted to the Tonle Sap, has been described with *P. melanochir melanochir* designated as the subspecies that is common in the MD (Motomura and Sabaj 2002; Motomura and Tsukawaki 2006). This common subspecies is also reported from Kalimantan, Borneo (Motomura 2004), although no molecular studies are known that help confirm its taxonomic status as a conspecific. We use *P. melanochir* to refer to the subspecies in the MD. *Polynemus melanochir* is considered a secondary freshwater fish because it 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 such as *P. melanochir* are derived from marine ancestors and 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, although separate statistics are not reported for this species. There is no specific habitat or biological information published for this species, however 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 to at least 25 cm standard length (Motomura 2004).

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 2012). 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. 2007b; Dasgupta et al. 2009; Hak et al. 2016; Nguyen 2017), including impacts that would restrict the distribution of *P. melanochir* in the MD (James et al. 2003; Koehn et al. 2011).

A variety of population genetic indices are useful in the application of conservation genetics for management of species. For example, knowledge on effective population size can give an indication if populations are viable in ecological and evolutionary timeframes, giving managers a framework for urgency of action (Frankham et al. 2014; Franklin et al. 2014). Information on the structure of populations including connectivity across a species range and source sink dynamics are important for understanding potential impacts of human-induced fragmentation or overexploitation in parts of a species range (Gido et al. 2016; Scribner et al. 2016), and population structure of estuarine and nearshore polynemids has been demonstrated using mitochondrial gene sequences (e.g., Chenoweth and Hughes 2003; Sun et al. 2013). In addition, an allozyme study showed population structure for the freshwater congeneric *Polynemus paradiseus* in India (Nahar et al. 2015), but no studies are known for *P. melanochir* in the Mekong Basin. While a handful of studies have shown varying levels of population structure within and between the main branches of the Mekong River 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 Sunnucks 2012), no known published studies to date have examined the genetic structure of fishes within the MD. The purpose of this study was to examine the population structure of *P. melanochir* across the 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 by native speakers 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

Table 1 *Polynemus melanochir* sample site information and genetic diversity in the Vietnamese Mekong River 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 (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

concentrations ≥ 3 ng/µL were selected for library preparation. Aliquots of 100 ng were bound to AMPureXP (Agencourt) beads using a 2:1 template to bead volume ratio, and any non-DNA products were removed with an ethanol wash.

The purified and concentrated 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 (10 ×), 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 min. 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 min, followed by 8 cycles (98 °C for 20 s, 60 °C for 15 s, 72 °C for 30 s), a final extension at 72 °C for 5 min 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 min, 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/cholllenbeck/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 in order to filter these out to conform to assumptions of neutrality and avoid misleading signals in population structure in downstream analyses (O'Leary et al. 2018). 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 or 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. 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. Any outliers identified by both Lositan and BayeScan were removed,

resulting in a putatively neutral SNP panel that was used for analysis.

Relatedness

Relatedness between sampled individuals was estimated with the R package 'related' (Pew et al. 2015) using two maximum likelihood estimators (Milligan 2003; Wang 2007) and a point estimator (Ritland 1996) that have all been shown to be robust to errors common in genotyping-by-sequencing methods such as allelic dropout (Attard et al. 2018). Parameters were set to allow for inbreeding and an error rate of 0.001. For all estimators, 95% confidence intervals were calculated using 500 bootstraps for each pairwise comparison. Five simulations were conducted with each estimator using 100 randomly selected loci from the neutral SNP panel to examine the performance of these estimators at distinguishing different levels of relatedness. For each replicate, 100 parent–offspring, 100 full-sib, 100 half-sib, and 100 unrelated pairs were simulated (see Supplementary Material, Fig. S3), and an (r) threshold based on the ability of the three estimators to distinguish relatedness in our simulations was used to identify related pairs. Any putatively related pairs above this cutoff were broken by removing one individual for analyses that will be biased by high levels of relatedness (such as PCAs and STRUCTURE analyses).

Genetic diversity 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 and Van Tienderen 2004).

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. NeEstimator was run with two datasets, one containing all individuals and one with putative related pairs broken as described above in order to avoid positive bias in estimates due to underestimating relatedness in the overall population (Waples and Anderson 2017). Effective population size was calculated for all sampled sites individually, for all the individuals combined into a single population, and for the Tien (DT, VL, TG, BT, TV) and Hau (AG, CT, ST) rivers.

Analyses of population structure

Pairwise comparisons of F_{ST} values between *P. melanochir* populations were computed using the neutral SNP panel 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 putative parent–offspring or 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 and Hochberg 1995). In addition to the results presented here, F_{ST} and principal component analyses were also run on putatively divergent outliers (reported in Supplementary Material).

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 (PCoA) were performed using the R package 'adegenet' (Jombart and Ahmed 2011). This analysis provides a graphic description of the allelic divergence among populations in multivariate space. Since high levels of relatedness strongly impact analyses using allele frequencies, the PCA and PCoA reported here were run on the neutral SNP panel following the removal of one individual from each parent–offspring or sibling pair. For comparison, a PCA performed with all individuals can be found in Supplemental Material (Fig. S4).

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). MIGRATE-n's implementation of coalescent theory measures migration $4 \times N_e$ generations in the past (Beerli 2009; Samarasin et al. 2016). 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 $T4 = 1.0 \times 10^6$. 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 sites (Tien Source) (9) migration occurring among all sites found in each river, however migration only occurs from the Hau River sites to Tien River sites (Hau Source). The most likely model was chosen using the Bezier approximation score produced by Migrate-n and migrants per generation for the chosen model were calculated according to Beerli (2009).

Results

Reference assembly and SNP filtering

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

Using the Freebayes tools for SNP calling, the initial dataset consisted of 459,374 raw SNPs. Following extensive filtering, the final dataset consisted of 1738 putative SNPs in 184 individuals (Table 1). Information on individuals removed at each step of filtering and data analysis is presented in Table S1.

Outlier loci detection

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

Relatedness

Simulation results indicated an increasing likelihood of the three estimators successfully identifying related individuals at the half-sibling level or greater beyond a threshold of $(r) > 0.125$ which was used as a cutoff for identifying putatively related pairs (see Supplementary Material, Figs. S2, S3). Analyses of relatedness using the full panel of 1735 neutral SNPs revealed 26 pairs of individuals related at the parent–offspring or sibling level (Table 2). Related pairs occurred most often within individual sampling sites with five or more related pairs sampled in the BT, CT, DT and TV populations, and a single pair of related individuals from AG. In addition, three related pairs were found between sites DT, CT, TG and TV, with the pair closest to our relatedness coefficient cutoff found between CT and TV (Table 2).

Genetic diversity and effective population size

Across eight sampled populations, *P. melanochir* showed average levels of observed 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 heterozygosity 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).

Estimates of N_e for individual sampling sites are reported in Supplementary Material and ranged from 18.3 (DT) to infinite (TG; Table S4). Combining individuals from all sites as a single population, N_e was estimated at 290.4, and at 147.4 and 170.0 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 among sites after breaking sibling pairs were not significant (Table 4).

Clustering analysis in the program STRUCTURE detected no population differentiation among 8 sampling

Table 2 Putatively related pairs of *Polynemus melanochir*

Source	Pair	Removed	Ritland (95% CI)	Milligan (95% CI)	Wang (95% CI)
Within sites	AG07–AG106	AG07	0.9201 (0.8673–0.9827)	0.9669 (0.9546–0.9777)	0.9672 (0.9553–0.9786)
	BT08–BT26	BT08, BT11, BT16, BT17, BT21	0.8741 (0.7974–0.9476)	0.9251 (0.9106–0.9394)	0.9254 (0.9109–0.9403)
	BT11–BT41		1.0347 (0.9533–1.124)	1 (1–1)	1 (1–1)
	BT16–BT27		0.922 (0.8578–0.9908)	0.9759 (0.9655–0.9846)	0.9762 (0.9659–0.9847)
	BT17–BT51		0.5042 (0.4471–0.5652)	0.5047 (0.4626–0.5445)	0.5054 (0.4633–0.5435)
	BT21–BT28		0.9107 (0.8479–0.9884)	0.9572 (0.9451–0.9681)	0.9575 (0.9454–0.9685)
	CT02–CT11	CT02, CT11, CT42	0.9598 (0.8783–1.0393)	0.8287 (0.8053–0.8544)	0.8298 (0.8056–0.8546)
	CT02–CT42		0.9819 (0.9198–1.0546)	0.9673 (0.9567–0.9783)	0.9674 (0.9578–0.9788)
	CT02–CT46		0.9885 (0.8874–1.0842)	0.6977 (0.6701–0.7255)	0.6985 (0.6713–0.7265)
	CT11–CT42		0.9603 (0.8868–1.0409)	0.8063 (0.783–0.8346)	0.8083 (0.7851–0.8355)
	CT11–CT46		0.9572 (0.8548–1.0622)	0.5947 (0.5587–0.6291)	0.5953 (0.5591–0.6279)
	CT42–CT46		0.9832 (0.8875–1.0854)	0.6784 (0.6502–0.7085)	0.6792 (0.6515–0.709)
	DT02–DT04	DT02, DT05, DT15, DT48	0.7226 (0.6636–0.7908)	0.6875 (0.6585–0.7163)	0.689 (0.6595–0.7172)
	DT03–DT05		0.7419 (0.6575–0.8278)	0.5522 (0.52–0.5869)	0.5533 (0.5197–0.5861)
	DT05–DT31		0.1999 (0.1365–0.2745)	0.1477 (0.0985–0.1974)	0.1477 (0.0985–0.1973)
	DT15–DT59		0.8839 (0.8035–0.9891)	0.7172 (0.6843–0.7465)	0.7183 (0.6864–0.7475)
	DT48–DT56		0.9065 (0.8119–0.9949)	0.6603 (0.6325–0.6915)	0.6613 (0.6334–0.693)
	TV01–TV42	TV01, TV04, TV10, TV12, TV31	0.9666 (0.8868–1.0508)	0.8275 (0.8057–0.8475)	0.8286 (0.8063–0.8492)
	TV04–TV12		0.9808 (0.9032–1.0603)	0.8718 (0.8527–0.8941)	0.8726 (0.8528–0.8957)
	TV04–TV44		0.956 (0.8851–1.038)	0.776 (0.7508–0.8025)	0.7769 (0.7508–0.8036)
	TV10–TV41		0.9944 (0.9184–1.0828)	0.9476 (0.9358–0.9601)	0.948 (0.9366–0.9605)
	TV12–TV44		0.9502 (0.8717–1.0257)	0.8827 (0.8657–0.9019)	0.883 (0.8668–0.9024)
	TV31–TV43		1.0407 (0.9468–1.1262)	0.9734 (0.962–0.982)	0.9736 (0.9625–0.9826)
Between sites	CT46–TV17	CT46	0.0477 (0.0355–0.1208)	0.1255 (0.0669–0.1685)	0.1279 (0.0671–0.1699)
	DT12–TG26	TG26	0.7776 (0.71–0.8571)	0.858 (0.8375–0.876)	0.8586 (0.8381–0.8757)
	DT50–CT09	CT09	0.9663 (0.8911–1.0499)	0.8348 (0.812–0.8574)	0.8356 (0.8127–0.8581)

Coefficients of relatedness (*r)* with 95% confidence intervals in parentheses are presented for both a point estimator (Ritland 1996) and two maximum likelihood estimators (Milligan 2003; Wang 2007)

Table 3 Estimates of *Polynemus melanochir* effective population size (N_e) calculated from 1735 neutral SNPs

Putative siblings included		Putative siblings removed	
Data set (*N*)	N_e (95% CIs)	Data set (*N*)	N_e (95% CIs)
Hau River (71)	147 (145–150)	Hau River (65)	793 (722–879)
Tien River (113)	170 (168–172)	Tien River (98)	1621 (1441–1850)
All sites (184)	290 (286–295)	All sites (163)	2056 (1878–2270)

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. Results for both the putative siblings included (n = 184) and putative siblings pairs broken (n = 163) and datasets are shown. 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

sites, with an optimal *K* value of 3 as determined by the Evanno method (Fig. 2). All but two individuals were assigned with high probably to a single lineage. An individual from CT was assigned with a q-score of 84.46% to a second lineage and an individual from AG was assigned with a q-score of 56.27% to a third. 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 16.8% 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.

Table 4 Pairwise F_{ST} values from 1735 neutral SNPs of *Polynemus melanochir* in the Mekong Delta, calculated with putative parent–offspring/sibling pairs included (lower diagonal; n = 184) and with these pairs broken (upper diagonal; n = 163)

Pop	AG	DT	CT	VL	TG	BT	TV	ST
AG	–	0.003	0.000	0.000	0.000	-0.001	0.003	0.002
DT	**0.008**	–	0.002	0.002	0.001	0.003	0.004	0.002
CT	**0.009**	**0.011**	–	0.002	0.000	0.002	0.001	0.002
VL	0.002	**0.008**	**0.007**	–	0.001	0.002	0.001	0.000
TG	0	**0.005**	**0.006**	0.001	–	0.000	0.002	0.002
BT	0.002	**0.010**	**0.010**	**0.005**	**0.003**	–	0.002	0.003
TV	**0.008**	**0.013**	**0.013**	**0.007**	**0.007**	**0.010**	–	0.003
ST	**0.003**	**0.008**	**0.008**	0	**0.002**	**0.005**	**0.008**	–

Significant F_{ST} values are in bold

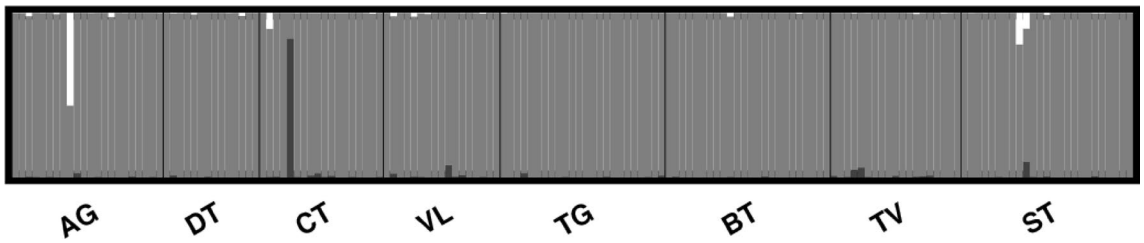

Fig. 2 Bar plot of *Polynemus melanochir* showing individual assignments to inferred clusters using the neutral SNP panel in the program STRUCTURE. Each genotype is represented by a single vertical bar partitioned into segments representing the estimated membership fractions in *K* clusters (optimal *K* = 3)

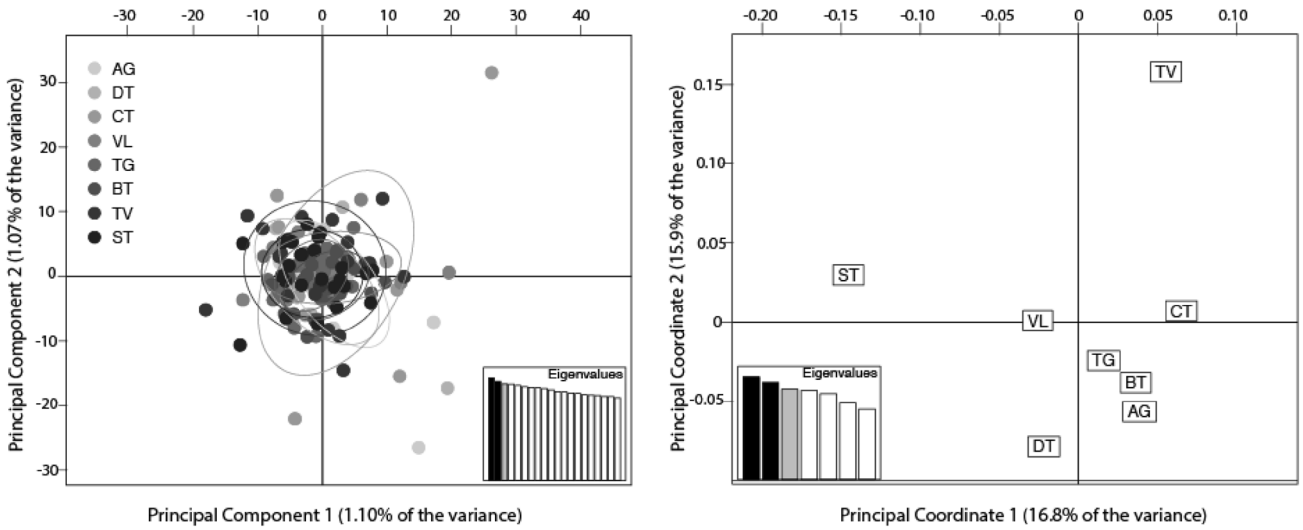

Fig. 3 Principal component (PCA, left) and principal coordinate (PCoA, right) analyses for *Polynemus melanochir* using neutral loci (related individuals removed). Insets show eigenvalues for the first 20 axes of the PCA and all axes for the PCoA

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 of the Hau and Tien Rivers, as well as up- and downstream. Results showed the 'Rivers Separate' model was the most supported (Bezier approximation score of − 644,468; Table S5), 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

Table 5 Estimates of mutation-scaled migration (M) and mutation-scaled effective population size ($\boldsymbol{\theta}$) 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 $(\theta \times M)/4$

River	Source	Sink	μ scaled migration (M)	μ scaled N_e (θ)	M/gen.
Hau	CT	AG	19,740	0.00057	3
	ST	AG	19,367	0.00057	3
	AG	CT	19,260	0.00057	3
	ST	CT	5060	0.00057	1
	AG	ST	19,727	0.00057	3
	CT	ST	19,767	0.00057	3
Tien	DT	BT	15,647	0.00103	5
	TG	BT	15,353	0.00103	4
	TV	BT	15,313	0.00103	4
	VL	BT	14,807	0.00103	4
	BT	DT	14,727	0.00097	4
	TG	DT	15,407	0.00097	4
	TV	DT	15,300	0.00097	4
	VL	DT	13,753	0.00097	4
	BT	TG	15,047	0.00077	3
	DT	TG	15,753	0.00077	4
	TV	TG	14,820	0.00077	3
	VL	TG	13,860	0.00077	3
	BT	TV	15,633	0.00003	3
	DT	TV	15,847	0.00003	3
	TG	TV	16,407	0.00003	3
	VL	TV	13,500	0.00003	3
	BT	VL	15,673	0.00117	5
	DT	VL	16,380	0.00117	5
	TG	VL	16,087	0.00117	5
	TV	VL	15,380	0.00117	5

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 panmictic population in the Vietnamese Mekong Delta with a low overall effective population size. Many of the pairwise F_{ST} comparisons between 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 migrate freely between rivers or that reproduction may occur upstream of the Vam Nao tributary where active flow from the Tien to the Hau occurs. The few species of Polynemidae for which information is available, including one species of *Polynemus*, are broadcast spawners. Assuming *P. melanochir* also uses this reproductive strategy, then passive dispersal would mostly be downstream, and panmixia could have occurred through egg dispersal within each of the respective main branches and between them through the Vam Nao tributary. Another possibility is that adult migration drives high levels of gene flow within and between rivers. The swimming capability of these demersal fishes is unknown but upstream swimming to some extent during high flow periods should be possible and even more feasible during the low flow periods of the dry season. The species may also inhabit and travel through the many canals that connect these two river systems (Vo 2012).

The two sampled fish with higher percentages of other putative lineages indicated by the STRUCTURE analysis both had genotyping rates of approximately 90% and observed and expected heterozygosity levels on par with the average, so their assignment is not likely an artefact of outlying values of missingness or heterozygosity. Because this is a complicated taxonomic group, it is possible that our sampling captured a rare migrant of the Tonle Sap subspecies or an F1 hybrid of two extant subspecies. In addition to *P. melanochir*'s potential for adult migration, this region experiences seasonal flood pulses which are characterized by large-scale, reversible flows between the Tonle Sap and the mainstream Mekong (Junk et al. 1989; Arias et al. 2013) which could also promote the exchange of individuals or eggs from upstream sources to the MD. The Migrate-n analysis indicates predominant migration within but not between the Tien and Hau rivers, however these historic patterns were not detectable in other analyses and may not reflect contemporary patterns of gene flow influenced by anthropogenic changes to MD hydrology such as the building of canals between tributaries. The equalization of downstream flow below the Vam Nao suggests that fish have the same ability to navigate upstream in both the Tien and Hau, and this may explain the lack of genetic differentiation within both rivers. *Polynemus melanochir* can therefore be considered a single fisheries and conservation management unit across the MD.

The small F_{ST} values and significant differences among populations that disappear if siblings are removed from calculations may reflect fairly high rates of localized reproduction and a low overall effective population size. The presence of parent–offspring or sibling pairs in half of the sampled sites may indicate that local recruitment originates from a limited pool of successful reproductive adults. The effective population size calculated with all sites combined is less than the lower limit of 500 considered sufficiently resilient to long-term population changes (Jamieson and Allendorf

2012). The effective population sizes per site were even smaller (less than 70 for most sites containing sibling pairs), but since population structure analyses provide strong evidence for a single panmictic population, these calculations are for analytically forced 'populations,' and an estimate of population size from all sites combined is more reliable. The presence of high levels of relatedness within sample sites, which will strongly affect N_e estimates, may be a consequence of non-random sampling of the overall population by fishers that resulted in a higher likelihood of capturing siblings or otherwise related individuals. Alternatively, related fish may not be randomly distributed throughout the MD so the related pairs sampled within five of our eight sampled sites could be an accurate reflection of a true biological phenomenon in this species. Without knowledge of the life history and reproductive strategy of *P. melanochir*, it is impossible to distinguish which of these is more likely to drive the levels of relatedness measured within sampled sites.

The management implications of our findings have both promising and alarming components. Panmixia in the population indicates that fragmentation is limited and that localized threats may not harm the overall population. However, the low estimated effective population sizes are worrisome in that they indicate this species may not be resilient to long-term changes in the MD environment (Jamieson and Allendorf 2012). The minimum viable population of 500 to maintain evolutionary potential and hence ability to adapt to changing environments is above the effective population size detected in *P. melanochir* for the entire sampled region. Therefore, management of this species should be careful to ensure that overfishing does not further erode the effective population size of the existing population. More should also be done to understand the effects of pollution and other threats to this species to avoid pushing it further below viable population sizes and if necessary, mitigate existing threats. In addition, this species is very actively exploited and although an important component of fisheries production, the fisheries biology of this species is largely unknown and should be studied to understand viability of populations.

This is one of the first studies on population genetics of MD fishes and its findings introduce many potential future questions to be addressed. For example, this study did not have results that directly address questions of the influence increasing salinity regimes may have on populations of freshwater fishes in the MD. The average elevation above sea level for the MD is 0.3 to 0.7 m and sea level rise between 1985 and 2010 was measured at 3 mm per year with predictions that the entire region will be inundated with salt water in less than a century (Hak et al. 2016). This will undoubtedly threaten freshwater fishes such a *P. melanochir* that predominate in the region. Low heterozygosity levels in two of the three sites sampled closest to the sea (Table 1) give a faint but inconclusive hint that selection may be occurring at some of these sites, and the high salinity levels at these sites (Vo 2012) may be a selective pressure factor. A comparison of upstream and downstream sites using a greater coverage of the genome may show divergent selection if genes involved in osmoregulation can be sampled. Transcriptomic studies examining differences in gene expression relative to salinity tolerance in upstream and downstream populations of fishes would also be informative in light of predicted increases in saltwater intrusion into the MD. In addition, although considered mostly an estuarine species restricted to the MD, the freshwater dependence of *P. melanochir* is still not well understood, especially in context of the rare subspecies restricted to the Tonle Sap. Genetic studies should be expanded to include the Tonle Sap subspecies and the putative conspecific found in Borneo to better understand the limits of this species (we attempted to collect in Kalimantan, Borneo for this study but were not successful).

The increasing threats and predicted changes in the region require a more comprehensive view of the natural history of resident fishes. Detailed information on feeding habits, growth and reproductive biology would provide a more complete understanding of the population ecology of this species. 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. These results also provide important baseline measures of diversity that can be used for future genetic analyses aimed at monitoring responses of *P. melanochir* in the face of climate change and increased resource use in the MD.

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.

References

Adamson EA, Hurwood DA, Baker AM, Mather PB (2009) Population subdivision in Siamese mud carp *Henicorhynchus siamensis* in the Mekong River basin: implications for management. J Fish Biol 75(6):1371–1392

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

Alliance Delta (2011) Mekong Delta water resources assessment studies Vietnam-Netherlands Mekong Delta Masterplan project. Deltares, Delft

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

Arias ME, Cochrane T, Norton D, Killeen TJ, Khon P (2013) The flood pulse as the underlying driver of vegetation in the largest wetland and fishery of the Mekong basin. Ambio 42(7):864–876

Attard CR, Beheregaray LB, Möller LM (2018) Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: avoiding the trap of precise bias. Mol Ecol Resour 18(3):381–390

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

Beerli PE (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. In: Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi (eds) Population genetics for animal conservation. The Cambridge University Press, Cambridge, pp. 42–79

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

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

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

Campbell, IC (2012). Biodiversity of the Mekong Delta. In: The Mekong Delta system. Springer, Dordrecht, pp. 293–313

Chenoweth SF, Hughes JM (2003) Oceanic interchange and nonequilibrium population structure in the estuarine dependent Indo Pacific tasselfish *Polynemus sheridani*. Mol Ecol 12(9):2387–2397

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

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

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

Dasgupta S, Laplante B, Meisner C, Wheeler D, Yan J (2009) The impact of sea level rise on developing countries: a comparative analysis. Clim Chang 93:379–388

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

Dutta D, Alam J, Umeda K, Hayashi M, Hironaka S (2007) A two-dimensional hydrodynamic model for flood inundation simulation: a case study in the lower Mekong river basin. Hydrol Process 21:1223–1237

Earl DA, von Holdt BM (2012) Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2):359–361

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

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

Frankham R, Bradshaw CJ, Brook BW (2014) Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biol Conserv 170:56–63

Franklin IR, Allendorf FW, Jamieson IG (2014) The 50, 500 rule is still valid: reply to Frankham et al. Biol Conserv 176:284–285

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

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. arXiv Preprint http://arxiv.org/abs/1207.3907

Gido KB, Whitney JE, Perkin JS, Turner TF (2016) Fragmentation, connectivity and fish species persistence in freshwater ecosystems. Conserv Fresh Fish 292–323

Goldberg CS, Waits LP (2010) Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. Mol Ecol Resour 10:304–313

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

Hak D, Nadaoka K, Bernado LP, Le Phu V, Quan NH, Toan TQ, Trung NH, Van Ni D, Van PD (2016) Spatio-temporal variations of sea level around the Mekong Delta: their causes and consequences on the coastal environment. Hydrol Res Lett 10:60–66

Hurwood DA, Adamson EA, Mather PB (2008) Evidence for strong genetic structure in a regionally important, highly vagile cyprinid (*Henicorhynchus lobatus*) in the Mekong River Basin. Ecol Freshw Fish 17(2):273–283

James KR, Cant B, Ryan T (2003) Responses of freshwater biota to rising salinity levels and implications for saline water management: a review. Aust J Bot 51:703–713

Jamieson IG, Allendorf FW (2012) How does the 50/500 rule apply to MVPs? Trends Ecol Evol 27(10):578–584

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071

Junk W, Bayley PB, Sparks RE (1989) The flood pulse concept in river floodplain systems. In: Dodge DP (ed) Proceedings of the international large river symposium (LARS). Can B Fish Aquat Sci 106:110–127

Koehn JD, Hobday AJ, Pratchett MS, Gillanders BM (2011) Climate change and Australian marine and freshwater environments, fishes and fisheries: synthesis and options for adaptation. Mar Freshw Res 62:1148–1164

Kottelat M (2013) The fishes of the inland waters of Southeast Asia: a catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles B Zool, Supplement 27:1–663

Le AT, Chu TH, Miller F, Bach TS (2007a) Flood and salinity management in the Mekong Delta, Vietnam. In: Be, TT, Sinh BT, Miller F (eds) Challenges to sustainable development in the Mekong Delta: regional and national policy issues and research needs. Bangkok, Thailand

Le TV, Nguyen HN, Wolanski E, Tran TC, Haruyama S (2007) The combined impact on the flooding in Vietnam's Mekong River delta of local man-made structures, sea level rise, and dams upstream in the river catchment. Estuar Coast Shelf S 71:110–116

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv Preprint http://arxiv.org/abs/1303.3997

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14):1754–1760

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22(13):1658–1659

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4:792–794

Milligan BG (2003) Maximum-likelihood estimation of relatedness. Genetics 163:1153–1167

Motomura H (2004) FAO species catalogue. Threadfins of the world (family Polynemidae). An annotated and illustrated catalogue of polynemid species known to date. FAO, Rome

Motomura H, Sabaj MH (2002) A new subspecies, *Polynemus melanochir dulcis*, from Tonle Sap Lake, Cambodia, and redescription of *P. m. melanochir* Valenciennes in Cuvier and Valenciennes, 1831 with designation of a neotype. Ichthyol Res 49:181–190

Motomura H, Tsukawaki S (2006) New species of the threadfin genus *Polynemus* (Teleostei: Polynemidae) from the Mekong River basin, Vietnam, with comments on the Mekong species of *Polynemus*. Raffles B Zool 54:459–464

Nahar A, Siddik MA, Alam MA, Chaklader MR (2015) Population genetic structure of paradise threadfin *Polynemus paradiseus* (Linnaeus, 1758) revealed by allozyme marker. Int J Zool Res 1(2):48–56

Nguyen N (2017) Historic drought and salinity intrusion in the Mekong Delta in 2016: lessons learned and response solutions. Vietnam Sci Technol 1:93–96

Nguyen AD, Savenije HH (2006) Salt intrusion in multi-channel estuaries: a case study in the Mekong Delta. Vietnam. Hydrol Earth Syst Sc Discuss 10:743–754

Nguyen TT, Sunnucks P (2012) Strong population genetic structure and its management implications in the mud carp *Cirrhinus molitorella*, an indigenous freshwater species subject to an aquaculture and culture-based fishery. J Fish Biol 80(3):651–668

Nguyen LA, Verreth JA, Leemans HB, Bosma RH, De Silva S (2016) A decision tree analysis to support potential climate change adaptations of striped catfish (*Pangasianodon hypophthalmus* Sauvage) farming in the Mekong Delta, Vietnam. Tropicultura 34(Special):105–15

Noh S, Choi M, Kim E, Dan NP, Thanh BX, Van Ha NT, Sthiannopkao S, Han S (2013) Influence of salinity intrusion on the speciation and partitioning of mercury in the Mekong River Delta. Geochim Cosmochim Acta 106:379–390

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018) These aren't the loci you'e looking for: principles of effective SNP filtering for molecular ecologists. Mol Ecol 27(16):3193–3206

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

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959

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

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genet Res 67(2):175–185

Samarasin P, Shuter BJ, Wright SI, Rodd FH (2016) The problem of estimating recent genetic connectivity in a changing world. Conserv Biol 31(1):126–135

Scribner KT, Lowe WH, Landguth E, Luikart G, Infante DM, Whelan GE, Muhlfeld CC (2016) Applications of genetic data to improve management and conservation of river fishes and their habitats. Fisheries 41(4):174–188

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

So N, Van Houdt JK, Volckaert FA (2006) Genetic diversity and population history of the migratory catfishes Pangasianodon hypophthalmus and Pangasius bocourti in the Cambodian Mekong River. Fish Sci 72(3):469–476

Sun X, Xu D, Lou B, Zhang T, Xin J, Guo Y, Ma S (2013) Genetic diversity and population structure of *Eleutheronema rhadinum* in the East and South China Seas revealed in mitochondrial COI sequences. Chin J Oceanol Limn 31(6):1276–1283

Ta TK, Nguyen VL, Tateishi M, Kobayashi I, Tanabe S, Saito Y (2002) Holocene delta evolution and sediment discharge of the Mekong River, southern Vietnam. Quaternary Sci Rev 21:1807–1819

Takagi AP, Ishikawa S, Nao T, Song SL, Hort S, Thammavong K, Saphakdy B, Phomsouvanhm A, Nishida M, Kurokura H (2010) Genetic differentiation and distribution routes of the bronze feather back *Notopterus notopterus* (Osteoglossiformes: Notopteridae) in Indochina. Biol J Linn Soc 101(3):575–582

Takagi AP, Ishikawa S, Nao T, Limsong S, Hort S, Thammavong K, Saphakdy B, Phomsouvanhm A, Nishida M, Kurokura H (2011) Population structure of the climbing perch, *Anabas testudineus*, in the lower Mekong River basin. Fish Manag Ecol 18(2):145–153

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

Vo, KT (2012) Hydrology and hydraulic infrastructure systems in the Mekong Delta, Vietnam. In: The Mekong Delta system. Springer, Dordrecht, pp. 49–81

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

Waples RS, Anderson EC (2017) Purging putative siblings from population genetic data sets: a cautionary view. Mol Ecol 26(5):1211–1224