Inferring contemporary and historical genetic connectivity from juveniles

Feutry, Pierre; Berry, Oliver; Kyne, Peter M.; Pillans, Richard D.; Hillary, Richard M.; Grewe, Peter M.; Marthick, James R.; Johnson, Grant; Gunasekera, Rasanthi M.; Bax, Nicholas J.; Bravington, Mark

Published in:
Molecular Ecology

DOI:
10.1111/mec.13929

Published: 01/01/2017

Document Version
Peer reviewed version

Biological Sciences – Genetics

Inferring contemporary and long-term genetic connectivity from juveniles.

Pierre Feutry[a,b*], Oliver Berry[c], Peter M. Kyne[b], Richard D. Pillans[d], Rich Hillary[a], Peter M. Grewe[a], James R. Marthick[e], Grant Johnson[f], Rasanthi M. Gunasekera[a], Nicholas J. Bax[a,g], Mark Bravington[a]

a CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart 7000, Tasmania, Australia

b Research Institute for the Environment and Livelihoods, Charles Darwin University, Ellengowan Drive, Darwin 0909, Northern Territory, Australia

c CSIRO Oceans and Atmosphere, PMB 5, Floreat 6014, Western Australia, Australia

d CSIRO Oceans and Atmosphere, 41 Boggo Road, Dutton Park 4102, Queensland, Australia

e Menzies Institute for Medical Research, University of Tasmania, 17 Liverpool Street, Hobart 7000, Tasmania, Australia

f Department of Primary Industry and Fisheries, Aquatic Resource Research Unit, GPO Box 3000, Darwin 0801, Northern Territory, Australia

g Institute for Marine and Antarctic Science, University of Tasmania, Private Bag 129, Hobart 7001, Tasmania, Australia

*** Corresponding author:**

Pierre Feutry, CSIRO, Castray Esplanade, Hobart, Tasmania, 7000, Australia. Phone: +61362325095, Email: Pierre.Feutry@csiro.au

Keywords: marine dispersal, philopatry, kinship, reproductive movements, threatened species, *Glyphis glyphis*

Abstract

Understanding population connectivity using molecular markers has broad application in natural resource management. The most popular approach has been indirect estimates of connectivity derived from allele frequencies. More recently, the spatial distribution of parent-offspring and full-sibling (FS) pairs has been used to provide direct estimates of larval or juvenile movements. In combination, these approaches potentially provide contemporary and long-term connectivity estimates. Here we combine indirect estimates from whole mitogenome sequences and nuclear SNPs with direct estimates of adult and juvenile movements from FS and half-sibling (HS) data for the Critically Endangered Speartooth Shark, *Glyphis glyphis*. Over 350 juveniles were captured from the three river systems in tropical northern Australia where this species is found. None of the 72 FS and 24 same-cohort HS pairs of juvenile sharks were captured in different rivers, suggesting strong river fidelity in juveniles. In contrast, 18 of the 121 cross-cohort HS pairs identified were captured between the two closest river systems (c. 150 km apart) demonstrating recent male breeding movements between these rivers, but not more widely. Mitogenomic analyses revealed river specific long-term female reproductive philopatry. Allele frequency differences in the nuclear SNP data were observed between the river systems. However, between the two closest river systems, this only reflected the river fidelity in juveniles since it was not evident when FS and HS pairs were excluded from the dataset. Accounting for juvenile river fidelity, female philopatry and the presence of two distinct gene pools is important for the management of this threatened species.

Significance Statement

Measuring population connectivity is a critical task in conservation biology. It allows understanding how impacts on one population will affect another. For example, if one population is reduced to undesirable levels, over which time frame surrounding populations will help restoring it. Thanks to considerable efforts and advances in methodology, genetic markers can now provide reliable long-term estimates of population connectivity. However, scientists are still limited in their ability to estimate contemporary connectivity, arguably the most practical time frame for management. Here, we tackle this issue for the Endangered Speartooth Shark by leveraging the connectivity information contained in the spatial distribution of juvenile kin pairs. Importantly, this new approach can be achieved by sampling at a single time period without the need to sample often inaccessible adults.

Introduction

The management of natural resources relies on an understanding of distributions and population boundaries (1). Smaller, less productive populations can become threatened while larger and more productive populations can be exploited sustainably (2). Also, populations that have been genetically isolated for many generations are more susceptible to irreversible decline than more connected populations that can be buffered by their connectivity. Defining the level of connectivity between populations has challenged fishery and conservation scientists for many decades (3). Methodological advances, particularly in genetics (4), have got closer to the question fishery and conservation scientists have been asking: are two geographically separated populations of the same species connected such that a decline in one will affect the other, or conversely, if one population is reduced to undesirable levels, will the other population help restore it within a practical management time frame of a small number of generations?

Genetic markers are commonly used to identify and measure the strength of population boundaries through the application of two main analytical frameworks. In the first, the extent of population differentiation between spatially or temporally separated samples is evaluated by quantifying differences in allele frequencies with metrics such as F_{ST} and its analogues (5-8). This indirect approach is a relatively powerful way to detect restricted gene flow but is limited to cases where prior knowledge of putative population boundaries is available and can be tested. Secondly, as larger genetic datasets and more powerful computers have become available, unsupervised clustering algorithms have increasingly been used to provide indirect delineation of population boundaries (9-12). Broadly speaking, these methods assign individuals to groups that best meet Hardy-Weinberg and gametic-phase disequilibrium expectations (12). Because they are not reliant on *a priori* defined population boundaries they have the potential to detect cryptic population structure.

One limitation of these two indirect approaches is the difficulty to interpret these indirect estimates in an ecological context as they reflect processes integrated over evolutionary timeframes (13). A further limitation of these approaches is the upward bias in population subdivision caused by family members within samples used to infer population structure. Sampling a large number of progeny from a small pool of reproducing adults can produce an 'Allendorf-Phelps effect', that is, highly significant measures of population differentiation without reproductive isolation (14, 15). It can also erroneously produce a signal of population subdivision when clustering algorithms, such as the one implemented in the software package STRUCTURE, are used (16), termed here 'Anderson-Dunham effect'. To minimise this effect, the population sample should adequately represent breeders in the putative populations of interest (14). In practice, however, this can be difficult to achieve, since closely related individuals aggregate in many species (17-19). This is most problematic if juveniles are sampled because they have had fewer opportunities to disperse, or because they obtain benefits from gregariousness (20).

More recently, genetic markers have been used to provide direct estimates of population connectivity by mapping the spatial distribution of parent-offspring pairs (21-23). In contrast to indirect methods, direct estimates offer a clearly defined brief timeframe over which to measure spatial processes because the distance between parents and their offspring must accrue between the offspring's birth and capture (21). This approach is particularly useful for characterising dispersal kernels, identifying the drivers of dispersal in juveniles and to investigate contemporary recruitment dynamic (24). However, direct methods typically cannot determine whether dispersing offspring contribute to subsequent generations, or how consistent the observed movements are over the long term. Since both contemporary and long-term spatial processes are relevant to species management, the simultaneous application of both direct and indirect methods should be a highly desirable approach (25). Particularly if inferences can be made from the same dataset

Recent improvements in sequencing methods now permit the genotyping of hundreds of individuals at thousands of loci (26) and whole mitogenomes instead of single mitochondrial markers (27, 28). This can benefit both indirect and direct approaches to assessing population connectivity (4). More markers will for example increase the ability to detect low levels of population differentiation (15). The main factor limiting the use of direct estimates of genetic connectivity studies is sampling. Good estimates derived from parent-offspring distribution require the sampling of a significant proportion of the adults and juveniles of each population, which is only possible for small populations with well-defined distributional ranges. With more markers, direct methods can also reveal kinship beyond parent-offspring, potentially removing the need to sample adults (29). The spatial distribution of cross-cohort half-sibling pairs for example provides insight into their parents' breeding movements. Hence, access to adults is not required and sampling can be done in areas such as nurseries, where juveniles aggregate and boundaries may be understood.

The Speartooth Shark, *Glyphis glyphis* (Carcharhinidae), belongs to a poorly-known and highly threatened group of river sharks, whose taxonomy, distributions, population structure and conservation status are only now beginning to be resolved (28, 30-32). *Glyphis glyphis* is of high conservation concern and is classified as Critically Endangered on the Australian *Environment Protection and Biodiversity Conservation Act* 1999. This assessment was mostly based on infrequent collections across a restricted distribution, suggesting low population abundance. Understanding population boundaries and abundance is central to effective management of the species. *Glyphis glyphis* is currently known from three river systems within tropical Australia flowing into Van Diemen Gulf and the Gulf of Carpentaria where they inhabit large tidal river systems, estuaries and coastal environments (31, 33). Until recently only juveniles and sub-adults had been found. The first adults of the species were recorded in 2014 in southern Papua New Guinea (32) and 2015 in Australia (R.D. Pillans, unpubl. data). It is suspected that adults occur in the marine and coastal zone of northern Australia, possibly entering estuaries and rivers to give birth, as neonates can be reliably found during parturition season from October to

December, in upper tidal reaches of rivers (Pillans *et al.* 2010; P.M. Kyne *et al.* unpubl. data). Because adults can't be reliably caught, understanding of the species' biology relies heavily on the study of juveniles (28).

A recent mitogenomic study suggested female reproductive philopatry (28), but the extent of male dispersal remains unknown. Such information is critical to direct management of this threatened species, given its occurrence in only a limited number of river systems. Strong population structure would suggest that management would need to focus at the level of the individual river.

Here, we combined whole mitogenome sequencing and genome scans to investigate the population structure of *G. glyphis*. We infer juvenile and adult contemporary connectivity from the spatial distribution of full- and half-siblings, and contrast it with indirect longer-term estimates of genetic connectivity to provide management-relevant information on the spatial scale of movement in this threatened species. Specifically, we determine whether juveniles move between river systems (putative populations); whether adults (separately for males and females) breed with adults from more than one river system; and the degree of bias in indirect methods caused by the failure to account for familial structure. This is achieved by sampling at a single time period without the need to sample often inaccessible adults.

Material and methods

Sampling and DNA extraction

Glyphis glyphis samples were collected between January 2012 and December 2014 in the Alligator Rivers system (South Alligator n=82; East Alligator n=6; West Alligator n=1) and the Adelaide River (n=142) of the Northern Territory (NT), and the Wenlock River system (n=125) of Queensland (QLD), northern Australia (Fig. 1). Sharks were caught by rod and line or gillnet. Each shark was measured, sexed and a small fin clip was taken from the inner pectoral fin before it was released at the site of capture. Sampled sharks were from the size range 49–195 cm total length (TL) representing neonates through to subadults. Size at birth is ~50–65 cm TL (31) and with a median size of sampled sharks of 82.75 cm TL, most represented neonates or 1+ juveniles. Sharks were sampled under Northern Territory Fisheries Special Permit S17/3252, Kakadu National Park Research Permit RK805, Queensland Fisheries General Research Permit 163582, and Charles Darwin University Animal Ethics Committee A11041. DNA was extracted using either the DNeasy Blood and Tissue kits (Qiagen) or the NucleaMag Tissue kits (Macherey-Nagel).

SNP genotyping

The SNP genotyping was done using DArTseq™, a new implementation of sequencing of complexity reduced representations (34). The protocol used in this study mostly followed that described by Grewe*, et al.* (35). The only difference being that two complexity reduction methods were used instead of

one, PstI-SphI and PstI-NspI, in order to generate more markers. The SNP calling was done with DArT PLD's proprietary software DArTsoft14. DArTsoft14 uses scoring consistency derived from technical sample replicates (i.e. samples processed twice from DNA library preparation to SNP calling) to optimise its algorithm parameters (35).

SNP filtering

The data set used for population analysis consisted of 75bp fragments containing one or more SNP. When multiple polymorphisms were found on the same 75bp fragment (RAD contig), a single SNP was randomly chosen to represent that locus avoiding linkage disequilibrium between close loci. Prior to population analysis, loci were further screened by excluding loci not scored for all individuals, with minor allele frequency (MAF) lower than 0.02, reproducibility lower than 0.99 (approximately 10% of the individuals were genotyped twice and the reproducibility represented the proportion of the replicate pairs for which the genotyping is consistent) and with average sequencing depth lower than 10x.

Departure from Hardy-Weinberg equilibrium (HWE) was then tested for each locus within each sampling location using the "HWE.test.genind" function in the Adegenet R package (36) and the false discovery rate method was applied to control for multiple comparison testing (37).
We used the R package OutFlank (38) to identify outlier loci putatively under the influence of directional selection. The approach implemented in Outflank is based on an improved method for deriving the null distribution of population differentiation for neutral loci. It results in fewer false positives than other outlier tests, which are more influenced by the effects of demographic history (39). We ran Outflank with 5% left and right trim for the null distribution of F_{ST}, minimum heterozygosity for loci of 0.1, and a 5% false discovery rate (q value).

Mitogenome sequencing

The mitogenomes of 93 *G. glyphis* included in this study were sequenced as part of previous work (28). Another 56 were amplified and sequenced following the same protocol (Genbank Accession, XXX-XXX). In short, the mitogenomes were amplified in 2 overlapping fragments. The PCR products were then purified with Agencourt AMPure XP magnetic beads (Beckman Coulter) and prepared with Nextera XT DNA Sample Preparation kits (Illumina) for sequencing on a Miseq (Illumina). Reads were trimmed, filtered and mapped onto the reference sequence (40) using default parameters for the low sensitivity and no fine tuning options in GENEIOUS PRO (v. 8.1.7).

Kinship analyses and fish filtering

COLONY (v. 2.0.5.8) (41) was used to identify full-sibling (FS) and half-sibling (HS) relationships from the nuclear DNA data. Five independent runs using 'update allele frequency' were carried out with default analysis parameters, except for the mating system, which was set as female and male polygamy with

potential inbreeding. Pairs of FS or HS with uncertainty probability estimate above 0.95 were considered true sibships. Cross-cohort HS were determined by comparing capture dates and fish length to growth rate estimates derived from recaptures.

To address potential bias from family sampling (14, 16), identical population analyses were carried out on both the all individuals (ALL) and without FS or HS (No_Sib) sample sets. To create the No_Sib dataset, one individual from each sibling pair was randomly discarded from the ALL dataset. When some individuals belong to more than one pair of FS or HS, those discarded were chosen so as to maximize the number of individuals preserved.

Population structure analysis

ARLEQUIN (v. 3.5.1.3) was used to calculate pairwise fixation indexes (Φ_{ST}) between each pair of rivers and test for reproductive female philopatry. Tamura-Nei was used as the model of nucleotide evolution in the AMOVA and to calculate Φ_{ST} values. Contemporary female reproductive philopatry was tested using an approximate likelihood ratio test based on cross-cohort HS haplotypes. Details for this test are provided in Supplementary Material S1.

Pairwise F_{ST} (6) and associated p-values were derived from the SNP data using the R package StAMPP and 10,000 bootstraps (42).

To further evaluate whether the nuclear genetic variation was partitioned geographically, a model-based clustering approach was used as implemented in STRUCTURE (v. 2.3.4) (12). Runs were done on the CSIRO Accelerator Cluster "Bragg", which consists of 128 Dual Xeon 8-core E5-2650 compute nodes. STRUCTURE seeks to group individuals in such a way that the groups maximize conformity to Hardy-Weinberg and linkage equilibrium. We ran STRUCTURE across values for K (number of clusters) between 1 and 8, and evaluated the fit of the data to different values of K. The fits of alternative models were evaluated with the Delta K method (43) implemented in CLUMPAK (44) and based on 20 independent runs for each value of K. All runs incorporated a 200,000 iterations burn in followed by 500,000 clustering iterations. We ensured the adequacy of the run length by checking the runtime likelihood and alpha for stability. For all runs we assumed that allele frequencies were correlated between sampling sites and allowed for admixture. All runs were completed with and without inclusion of prior location information (LOCPRIOR).

Results

SNP filtering

The DArTsoft14 pipeline delivered 2198 and 1948 SNPs for the PstI-SphI and PstI-NspI complexity reduction methods, respectively (Supplementary Material S2). These SNPs were then combined into a single SNP dataset for quality filtering and analysis. A total of 1330 SNPs passed all quality control filtering steps. No outlier SNP was detected using Outflank so that the SNP dataset for

downstream kinship and population analysis consisted of 1330 SNPs. Descriptive statistics including allelic richness (AR), observed heterozygosity (Ho), expected heterozygosity (He) and inbreeding coefficient (Fis) are given in Supplementary Material S3.

Kinship analyses

A total of 72 FS pairs (94 unique individuals) were identified, of which 12, 11 and 49 originated from the Adelaide, Alligator and Wenlock Rivers, respectively. No cross-river FS pairs were identified. A total of 145 HS pairs (179 unique individuals) were identified, 44 within the Adelaide River, 14 within the Alligator Rivers, 69 within the Wenlock River, and 18 split across the Adelaide and Alligator Rivers (Table 1).

Growth rate derived from recapture data ranged from 18.2 to 36.5 cm.year^{-1} for fish smaller than 85 cm TL (N=4) and from 6.3 to 7.4 cm.year^{-1} for fish larger than 85 cm TL (N=2). Fish from 18 HS pairs with length differences less than 7 cm were captured fewer than 150 days apart and classified as same-cohort. Fish from another six pairs of HS with length differences ranging 14–19 cm and captured between 200 and 400 days apart were also classified as same-cohort. None of these 24 same-cohort HS pair had fish captured in different rivers. Given the amount of time between captures, the length difference and the growth rate observed, fish from all other HS pairs were unlikely to be born at the same time and were thus considered cross-cohort (Table 1).

Population differentiation

Measures of population differentiation based on whole mitogenomes and nuclear SNPs are given in Tables 2 and 3, respectively. All pairwise mitogenome-based measures of population differentiation were statistically significant, independent of whether the FS and HS were included in the analyses or not (Table 2). Private haplotypes were found in each river, but at least one haplotype per river was found at another sampling site (Supplementary Material S4). Population differentiation between each river pair was also supported by nuclear SNPs, except for the Adelaide and Alligator Rivers after the FS and HS were discarded (Table 3). SNP-based pairwise F_{ST} were higher for the ALL dataset than the No_Sib dataset. The F_{ST} between Adelaide and Alligator Rivers was an order of magnitude lower and became non significant, whereas F_{ST} between Adelaide/Alligator and Wenlock Rivers roughly decreased by a factor of two but remained significantly different from zero (Table 3).

Clustering analyses

Only the ALL dataset showed clear evidence of genetic differentiation among the rivers and this was manifest as a division between Adelaide/Alligator Rivers and Wenlock River. The delta K analysis indicated K=7 as the best fit (delta K = 2.61), but 5 small clusters consisted of full and half-siblings (Fig. 2a). These results were consistent whether location priors were included or not (Supplementary material S5). The only signal of population structure remaining in the No_Sib

dataset was the distribution of q-values at K=2, which distinguished Wenlock samples from Adelaide and Alligator samples when location information was included as prior (Fig. 2b). This signal disappeared when location priors were not included in the analyses. L(K) was stable and did not support K=2 as the best fit whether the location information was included as prior or not (Supplementary material S5).

Discussion

For the first time in any elasmobranch species, whole mitogenome sequencing and genotyping-by-sequencing genome scans have been used in combination to analyse population connectivity. Our results reveal that a significant fraction of the *G. glyphis* individuals analysed from all three rivers were close kin (26% FS; 50% HS). This observation permits direct estimation of recent adult (breeding) and juvenile movements in this threatened species. In addition, the identification of kin means that long-term connectivity estimated from population subdivision can be made from juveniles without the family sampling bias that may be common in population genetic datasets (14).

Direct estimate of contemporary *connectivity*

The spatial distribution of FS pairs has previously been used to infer the movements of juvenile fishes (45). In addition, parent-offspring pairs had been used to infer the extent of larval dispersal (21-23). This is the first time adult movement is inferred from the spatial distribution of HS pairs. This is a considerable improvement for connectivity studies, for threatened species in particular, where adults are rare and not easily sampled. In the case of *G. glyphis*, only two adults have been caught in Australia as part of a scientific study (R.D. Pillans, unpubl. data).

We identified over 200 *G. glyphis* full- and half-sibling pairs with a high degree of certainty, made possible by the large number of SNP loci analysed. Full-sibling pairs were only captured within the same river suggesting that juveniles remain in the natal river for some time. Because age and growth data are not available for *G. glyphis*, the age-at-length of juvenile Bull Shark *Carcharhinus leucas* reported by Tillett, Meekan, Field, Hua and Bradshaw (46) is the best proxy available. These two species are sympatric in northern Australian rivers, have a similar life history including the use of river systems as nursery areas, and have similar size at birth and maximum sizes. The largest *G. glyphis* full-sibling identified in the current study was thus estimated to be 6 years old; suggesting that the use of river nurseries last several years for juveniles. Age data for *G. glyphis* would be required to estimate more accurately the extent of their presence in natal rivers.

Extended residency within the limited spatial habitat of these natal rivers may increase susceptibility to anthropogenic impacts. However, neither of the NT river system in this study has commercial line or net fisheries, and therefore pressure is greatly reduced in comparison to some adjacent coastal areas. In Queensland, commercial net and crab fisheries that are known to capture

juvenile *G. glyphis* overlap with the species distribution in the Wenlock River system as well as in coastal environments. The extent of capture of juveniles in rivers by recreational fishers is unknown, but illegal captures of this protected species have been recorded in the NT (Kyne and Feutry XXXX) and Queensland (R.D. Pillans, unpubl. data). Furthermore, the scale of Indigenous harvest is unknown. Future plans for further agricultural development of northern Australia and associated increased water demand (47) will likely have implications for the riverine habitats of these threatened species.

Given that the juveniles don't or very rarely move between rivers, the distribution of HS provides insight into the movements of adults between reproductive events. Out of the 121 cross-cohort HS pairs, 103 (85%) were captured within the same river, indicating that in most cases at least one parent returned to reproduce in the same river across breeding seasons. The remaining 18 (15%) cross-cohort HS pairs were shared between the Adelaide and the Alligator Rivers. In these cases, at least one parent had moved between these rivers (or their associated mating aggregation areas if gamete exchange occurs outside the river) to reproduce. Van Diemen Gulf is a relatively small system, and it is possible that adults from different rivers flowing in the gulf mix in this area. In contrast to the cross-cohort HS pairs, same-cohort HS pairs were never captured between rivers. Assuming females only breed once a year, this suggests that males do not reproduce with females going to pup in different rivers within the same year. Based on the variability in reproductive periodicity of Australian carcharhinids of similar or smaller size, minimum reproductive periodicity would be annual (48, 49), or potentially biennial given large size at maturity (50). Hence, it is likely that the Adelaide and Alligator Rivers populations have different mating aggregation areas. Once fish can be aged accurately, reproductive periodicity could be determined by examining the time gap between HS pairs.

It is significant that no cross-cohort HS pairs were shared between the Alligator/ Adelaide Rivers emptying into Van Diemen Gulf and the more distant Wenlock River emptying into the eastern Gulf of Carpentaria. Adult breeding movements on scales of ˜150 km therefore seem commonplace in *G. glyphis*, but non-existent or very rare over distance an order of magnitude higher.

Population structure when sampling families

Previously, whole mitogenome sequencing of *G. glyphis* had revealed female philopatry (28) which is common in sharks (51), but had not provided insight into the movements of males, nor been able to discount the effects of sampling kin. Nuclear markers provide the ability to take the understanding of population structure of *G. glyphis* a step further because they reflect both male and female mediated gene flow. In addition they permit identification of kin, whose presence has the potential to drive an upward bias in apparent population subdivision (14, 16), including in a previous study on *G. glyphis* by Feutry*, et al.* (28). In this case, the removal of FS and HS pairs did not greatly affect pairwise fixation indexes Φ_{ST} (average <5% absolute difference in Φ_{ST}), demonstrating the observed

population differentiation was due to female reproductive philopatry and not bias from family sampling.

In contrast to the mitogenome data, the presence of close relatives in the nuclear SNP dataset substantially increased the signal of population sub-division revealed by F_{ST} and STRUCTURE analyses. The presence of FS and HS in the sample created an upwards bias in the estimation of F_{ST} between the Adelaide and Alligator rivers and an overestimation of the number of populations identified by STRUCTURE, as predicted by Allendorf and Phelps (14) and Anderson and Dunham (16) respectively. The significant population differentiation initially identified in the full dataset between the Adelaide and Alligator Rivers was due entirely to bias from FS and HS. Similarly, the STRUCTURE analysis overestimated the sample partitioning with groups of FS and HS forming independent cluster (16).

The contrast between nuclear and mtDNA markers indicates sex-biased dispersal. Sex-biased dispersal has previously been reported in other sharks (52, 53) and has important implications for management. Daly-Engel *et al.* (2012) stated that the use of female or biparentally-inherited loci only can mislead conclusions with regards to management units. While mitochondrial markers showed structuring between the Adelaide and Alligator Rivers, the use of nuclear SNP loci indicated that these rivers are part of the same gene pool. Importantly though, as females exhibit river specific reproductive philopatry, this gene flow could not compensate for the loss of females from a specific river, so the female population of each river stills needs to be managed as though it is an isolated population. The Van Diemen Gulf population should be managed as a separate unit to the isolated Wenlock River population. The relationship of these populations to the species in Papua New Guinea (PNG) should be examined.

Direct versus indirect connectivity estimates and management implications

The contrasted F_{ST} and STRUCTURE results between the ALL and No_Sib datasets highlight the importance of inferring sibship when juveniles are sampled for population structure studies. Both direct and indirect estimates of population connectivity support the Adelaide and Alligator Rivers as part of the same nuclear gene pool, whereas the Wenlock River likely has a strong degree of demographic independence, at least for the generation of adults who produced the juveniles included in this study. F_{ST} values between Wenlock and Adelaide/Alligator Rivers was low but significant, whereas little evidence for the Wenlock River to host a different population was found using STRUCTURE. This highlights the limited ability of the STRUCTURE clustering approach to detect demographically meaningful breaks.

Direct estimates of connectivity have two main advantages over indirect methods. The first one is a known timeframe for the movements. FS and same-cohort HS provide information for the current generation of juveniles. The exact period of time covered depends on the age of the juveniles. Given appropriate sampling, potential between river movements could be inferred for each year

class. Cross-cohort HS provide information about their parents' movements between breeding events.

The second advantage is the information about migration rate between populations that lies in the distribution of HS pairs. In the non-spatial context, Bravington, Skaug and Anderson (29) have outlined how these data can be used to estimate sex-specific abundance and survival rates in a modified mark-recapture framework called close-kin mark-recapture (CKMR). An extension of this framework into the spatial domain would be able to utilise the migratory and abundance related information in these data to separate the two, and obtain quantitative estimates of between river migration rates.

As presented here, one advantage of the indirect estimate over direct estimates of connectivity is the ability to provide information about sex biased gene flow. However, given mitochondrial DNA is maternally inherited, if all HS had their mitogenomes sequenced and sufficient haplotypic diversity was uncovered, it would be possible to determine if there was a bias in the sex of the parent shared by the HS pair distributed in different rivers. Genetic markers located on sex specific chromosomes could also help in this task. For example, in heterogametic organisms, markers on the sex chromosomes can be used to identify whether male HS pairs have different fathers (29). This approach can provide direct evidence of which parent has been moving and sex specific movement rates, whereas indirect estimates only provide evidence of sex biased gene flow.

Combined information from direct and indirect connectivity estimates enables us for the first time to detect intergenerational between river movement and breeding patterns from a single contemporaneous sample. For *G. glyphis* in northern Australia: (i) juveniles do not move between river systems during riverine residencies (possibly >6 years); (ii) females predominantly pup in the same river; but (iii) reproducing males may move between breeding aggregations for river systems closer than 150km apart, although data on where breeding aggregations occur are lacking. This has implications for the conservation of this Critically Endangered species, in both the management and potential mitigation of increasing demands on their environment.

Acknowledgments:

This study was supported by the Marine Biodiversity Hub, a collaborative partnership supported through funding from the Australian Government's National Environmental Research Program (NERP). PF was partly supported by the North Australia Marine Research Alliance (NAMRA). PMK was partly supported by the NERP Northern Australia Hub. We thank the Traditional Owners and staff of Kakadu National Park, especially Anne O'Dea and Khan Spokes, and Northern Territory Department of Primary Industry and Fisheries, especially Thor Saunders and Mark Grubert for field assistance. The contribution of a large number of field volunteers is also acknowledged. Kate Buckley, Karen Devitt, Rhian Evans, Sharon Every, Charlotte Klempin and Roy Tipiloura, in particular provided significant field assistance. Thanks to CSIRO Scientific Computing for the use of the Bragg Cluster

References

1. Waples RS & Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* 15(6):1419-1439.
2. Musick J, Burgess G, Cailliet G, Camhi M, & Fordham S (2000) Management of sharks and their relatives (Elasmobranchii). *Fisheries* 25(3):9-13.
3. Kalinowski ST (2004) Genetic polymorphism and mixed-stock fisheries analysis. *Can. J. Fish. Aquat. Sci.* 61(7):1075-1082.
4. Gagnaire PA*, et al.* (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol. Appl.* 8(8):769-786.
5. Raymond M & Rousset F (1995) An exact test for population differentiation. *Evolution* 49(6):1280-1283.
6. Weir BS & Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38(6):1358-1370.
7. Wright S (1951) The genetical structure of populations. *Annals of eugenics* 15(1):323-354.
8. Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*:395-420.
9. Alexander DH, Novembre J, & Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19(9):1655-1664.
10. Corander J, Waldmann P, Marttinen P, & Sillanpää MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* 20(15):2363-2369.
11. Dawson KJ & Belkhir K (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genet. Res.* 78(01):59-77.
12. Pritchard JK, Stephens M, & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-959.
13. Kool JT, Moilanen A, & Treml EA (2013) Population connectivity: recent advances and new perspectives. *Landscape Ecol.* 28(2):165-185.
14. Allendorf FW & Phelps SR (1981) Use of allelic frequencies to describe population structure. *Can. J. Fish. Aquat. Sci.* 38(12):1507-1514.
15. Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* 89(5):438-450.
16. Anderson EC & Dunham KK (2008) The influence of family groups on inferences made with the program Structure. *Mol. Ecol. Resour.* 8(6):1219-1229.
17. Hansen MM, Nielsen EE, & Mensberg KLD (1997) The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. *Mol. Ecol.* 6(5):469-474.
18. Oremus M, Poole MM, Albertson GR, & Baker CS (2012) Pelagic or insular? Genetic differentiation of rough-toothed dolphins in the Society Islands, French Polynesia. *J. Exp. Mar. Biol. Ecol.* 432–433:37-46.
19. Richard KR, Dillon MC, Whitehead H, & Wright JM (1996) Patterns of kinship in groups of free-living sperm whales (*Physeter macrocephalus*)

610 revealed by multiple molecular genetic analyses. *Proc. Natl. Acad. Sci.*
611 93(16):8792-8795.
612 20. Wilson EO (1975) *Sociobiology: the new synthesis* (Harward University
613 Press) p 697.
614 21. Jones GP, Planes S, & Thorrold SR (2005) Coral reef fish larvae settle close
615 to home. *Curr. Biol.* 15(14):1314-1318.
616 22. Planes S, Jones GP, & Thorrold SR (2009) Larval dispersal connects fish
617 populations in a network of marine protected areas. *Proc. Natl. Acad. Sci.*
618 106(14):5693-5697.
619 23. Christie MR*, et al.* (2010) Larval connectivity in an effective network of
620 marine protected areas. *Plos One* 5(12):e15715.
621 24. Cowen RK & Sponaugle S (2009) Larval dispersal and marine population
622 connectivity. *Marine Science* 1.
623 25. Berry O, England P, Marriott RJ, Burridge CP, & Newman SJ (2012)
624 Understanding age‐specific dispersal in fishes through hydrodynamic
625 modelling, genetic simulations and microsatellite DNA analysis. *Mol. Ecol.*
626 21(9):2145-2159.
627 26. Davey JW*, et al.* (2011) Genome-wide genetic marker discovery and
628 genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12(7):499-
629 510.
630 27. Feutry P*, et al.* (2015) Whole mitogenome sequencing refines population
631 structure of the Critically Endangered sawfish *Pristis pristis*. *Mar. Ecol.*
632 *Prog. Ser.* 533:237.
633 28. Feutry P*, et al.* (2014) Mitogenomics of the Speartooth Shark challenges
634 ten years of control region sequencing. *BMC Evol. Biol.* 14:232.
635 29. Bravington MV, Skaug HJ, & Anderson EC (In press) Close-kin mark-
636 recapture methods. *Stat. Sci.*
637 30. Li C*, et al.* (2015) DNA capture reveals transoceanic gene flow in
638 endangered river sharks. *Proc. Natl. Acad. Sci.* 112(43):13302-13307.
639 31. Pillans RD, Stevens JD, Kyne PM, & Salini J (2010) Observations on the
640 distribution, biology, short-term movements and habitat requirements of
641 river sharks *Glyphis spp.* in northern Australia. *Endanger. Species Res.*
642 10:321-332.
643 32. White WT*, et al.* (2015) Rediscovery of the Threatened River Sharks,
644 *Glyphis garricki* and *G. glyphis*, in Papua New Guinea. *Plos One*
645 10(10):e0140075.
646 33. Kyne PM (2014) Threatened fishes and marine turtles of Kakadu National
647 Park (with notes on marine mammals). *Kakadu National Park Landscape*
648 *Symposia Series. Symposium 7: Conservation of Threatened Species, 26–27*
649 *March 2013, Bowali Visitor Centre, Kakadu National Park*, eds Winderlich S
650 & Woinarski J (Internal Report 623, Supervising Scientist, Darwin), pp 58-
651 74.
652 34. Altshuler D*, et al.* (2000) An SNP map of the human genome generated by
653 reduced representation shotgun sequencing. *Nature* 407(6803):513-516.
654 35. Grewe PM*, et al.* (2015) Evidence of discrete yellowfin tuna (*Thunnus*
655 *albacares*) populations demands rethink of management for this globally
656 important resource. *Sci. Rep.* 5:16916.
657 36. Jombart T (2008) adegenet: a R package for the multivariate analysis of
658 genetic markers. *Bioinformatics* 24(11):1403-1405.

37. Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. A. (Stat. Soc.)* 57(1):289-300.
38. Whitlock MC & Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of Fst. *Am. Nat.* 186:S24-36.
39. Lotterhos KE & Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24(5):1031-1046.
40. Chen X, Liu M, Grewe PM, Kyne PM, & Feutry P (2014) Complete mitochondrial genome of the Critically Endangered speartooth shark *Glyphis glyphis* (Carcharhiniformes: Carcharhinidae). *Mitochondr. DNA* 25(6):431-432.
41. Jones OR & Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol. Ecol. Resour.* 10(3):551-555.
42. Pembleton LW, Cogan NO, & Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed‐ploidy level populations. *Mol. Ecol. Resour.* 13(5):946-952.
43. 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.
44. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, & Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.*
45. Kanno Y, Vokoun JC, & Letcher BH (2011) Sibship reconstruction for inferring mating systems, dispersal and effective population size in headwater brook trout (*Salvelinus fontinalis*) populations. *Conserv. Genet.* 12(3):619-628.
46. Tillett BJ, Meekan MG, Field IC, Hua Q, & Bradshaw CJ (2011) Similar life history traits in bull (*Carcharhinus leucas*) and pig-eye (*C. amboinensis*) sharks. *Mar. Freshwat. Res.* 62(7):850-860.
47. Australian Government (2015) Our north, our future: white paper on developing Northern Australia. *Available at:* http://industry.gov.au/ONA/WhitePaper/Documents/northern_australia_white_paper.pdf.
48. Chin A, Simpfendorfer C, Tobin A, & Heupel M (2013) Validated age, growth and reproductive biology of *Carcharhinus melanopterus*, a widely distributed and exploited reef shark. *Mar. Freshwat. Res.* 64(10):965-975.
49. Harry AV, Tobin AJ, & Simpfendorfer CA (2013) Age, growth and reproductive biology of the spot-tail shark, *Carcharhinus sorrah*, and the Australian blacktip shark, *C. tilstoni*, from the Great Barrier Reef World Heritage Area, north-eastern Australia. *Mar. Freshwat. Res.* 64(4):277-293.
50. Mcauley RB, Simpfendorfer C, Hyndes G, & Lenanton R (2007) Distribution and reproductive biology of the sandbar shark, *Carcharhinus plumbeus* (Nardo), in Western Australian waters. *Mar. Freshwat. Res.* 58(1):116-126.

51. Dudgeon CL, *et al.* (2012) A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *J. Fish Biol.* 80(5):1789-1843.
52. Daly-Engel TS, *et al.* (2012) Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). *Plos One* 7(1):e29986.
53. Pardini AT, *et al.* (2001) Sex-biased dispersal of great white sharks. *Nature* 412(6843):139-140.

Table 1. Intra and inter river number of full-sibling pairs (above) and cross-cohort half-sibling pairs + same-cohort half-sibling pairs (below).

Rivers	Adelaide	Alligators	Wenlock
Adelaide	*42+2*\ 12	0	0
Alligators	*18+0*	*9+5*\ 11	0
Wenlock	*0+0*	0+0	52+17 \ 49

Table 2. Mitogenome-based pairwise Φ_{ST} for all individuals (above) and the dataset without full-sibling and half-sibling pairs (below).

Rivers	Adelaide	Alligators	Wenlock
Adelaide		0.24705**	0.70517**
Alligators	0.24352**		0.23768*
Wenlock	0.67673**	0.25928*	

* P-value < 0.01; ** P-value < 0.0001

Table 3. Nuclear SNP-based pairwise F_{ST} for all individuals (above) and the dataset without full-sibling and half-sibling pairs (below).

Rivers	Adelaide	Alligators	Wenlock
Adelaide		0.00095**	0.00458**
Alligators	0.00008^{NS}		0.00493**
Wenlock	0.00279**	0.00285**	

$_{NS}$ P-value > 0.05; ** P-value < 0.0001

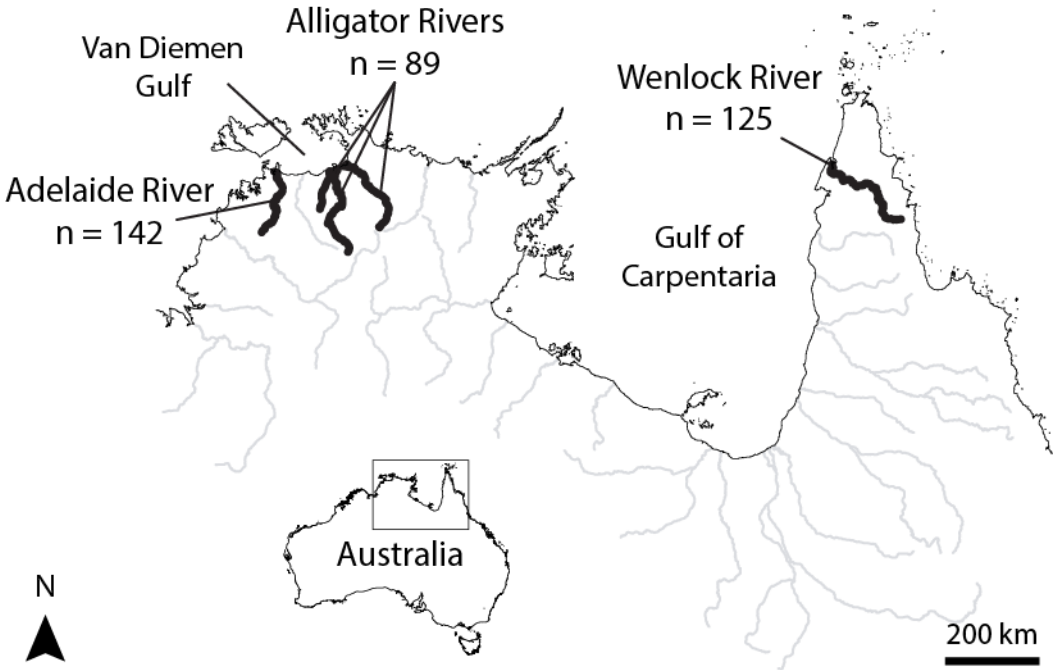


Figure 1. *Glyphis glyphis* sampling locations and sample size in northern Australia.

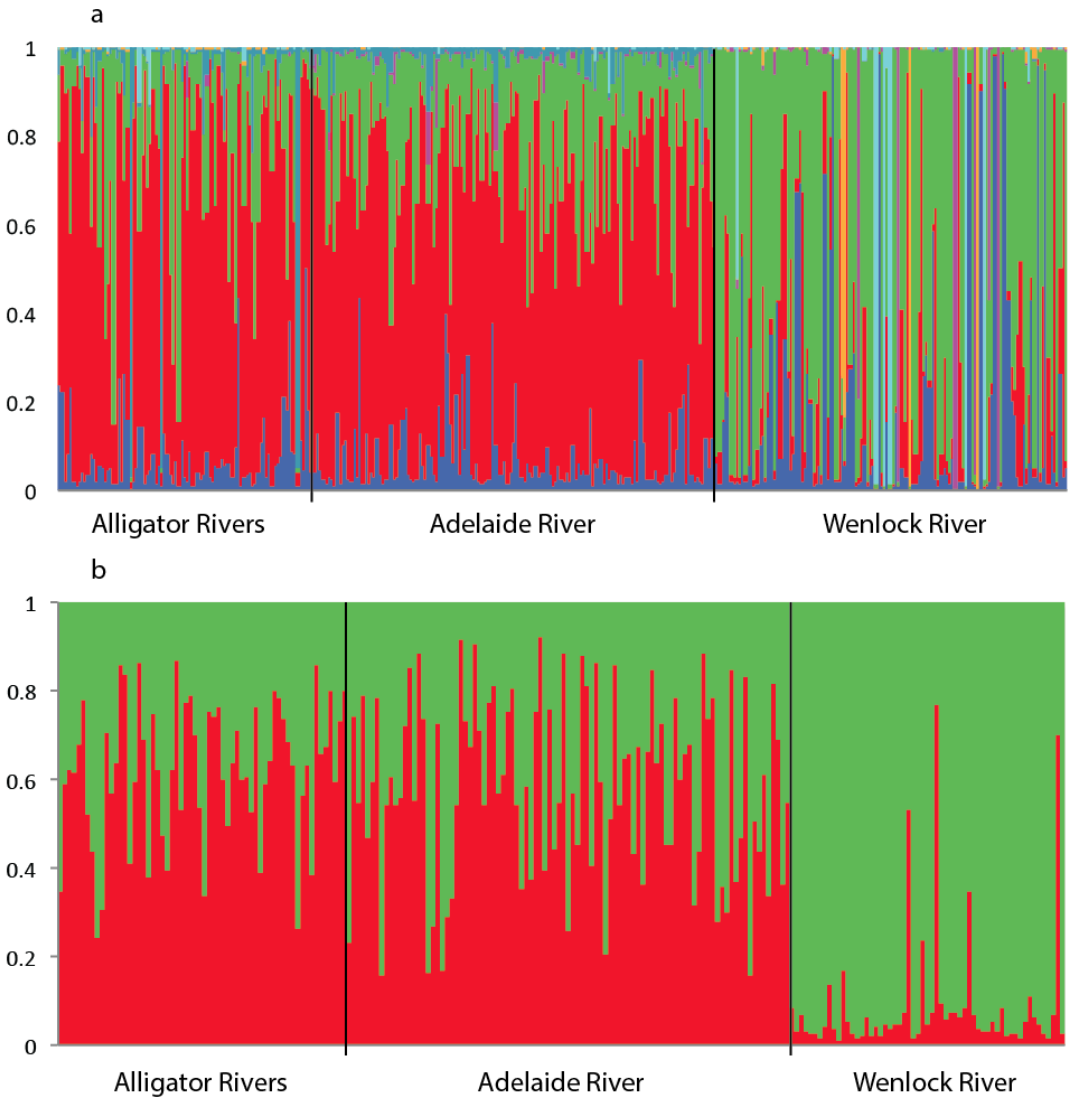
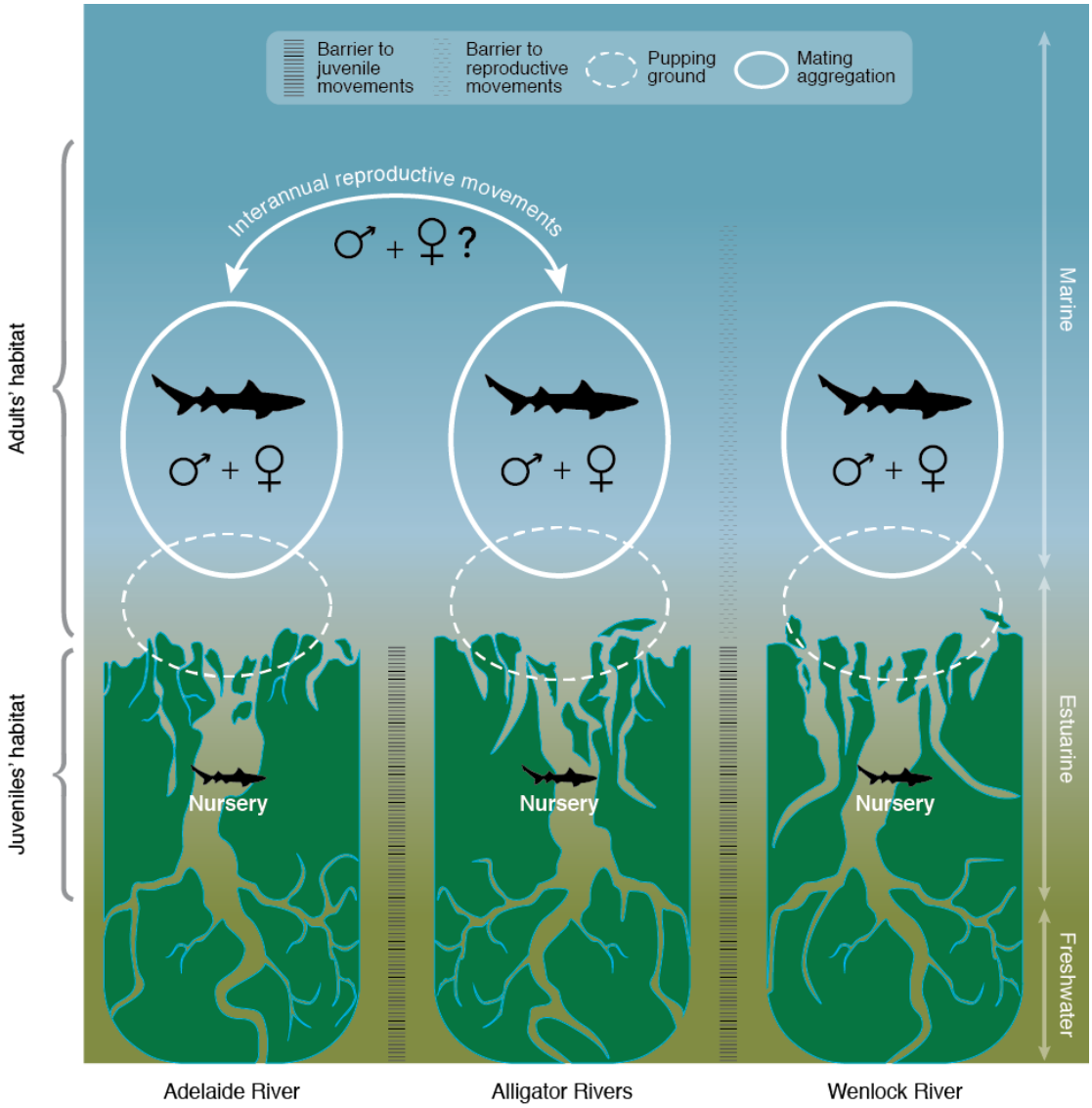


Figure 2. *Glyphis glyphis* STRUCTURE admixture analysis. Each cluster (K) is designated by a different colour. Each vertical bar represents one individual, partitioned according to admixture proportion from each cluster. a) Analysis of dataset with all samples, most likely K=7. b) Analysis of dataset without full-sibling and half-sibling pairs, most likely K=2.



744
745 Figure 3. Schematic representation of *Glyphis glyphis* movements as inferred
746 from spatial distribution of full- and half-sibling pairs and population structure
747 analyses of whole mitogenomes and nuclear genome scans.