

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](https://doi.org/10.1111/mec.13929)

Published: 01/01/2017

Document Version
Peer reviewed version

[Link to publication](#)

Citation for published version (APA):

Feutry, P., Berry, O., Kyne, P. M., Pillans, R. D., Hillary, R. M., Grewe, P. M., Marthick, J. R., Johnson, G., Gunasekera, R. M., Bax, N. J., & Bravington, M. (2017). Inferring contemporary and historical genetic connectivity from juveniles. *Molecular Ecology*, 26(2), 444-456. <https://doi.org/10.1111/mec.13929>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Biological Sciences – Genetics**

2
3 **Inferring contemporary and long-term genetic connectivity from**
4 **juveniles.**

5
6
7 Pierre Feutry^{a,b*}, Oliver Berry^c, Peter M. Kyne^b, Richard D. Pillans^d, Rich Hillary^a,
8 Peter M. Grewe^a, James R. Marthick^e, Grant Johnson^f, Rasanthi M. Gunasekera^a,
9 Nicholas J. Bax^{a,g}, Mark Bravington^a

10
11 **a CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart 7000, Tasmania,**
12 **Australia**

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

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

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

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

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

23 **g Institute for Marine and Antarctic Science, University of Tasmania, Private Bag**
24 **129, Hobart 7001, Tasmania, Australia**

25
26 *** Corresponding author:**

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

29
30 **Keywords: marine dispersal, philopatry, kinship, reproductive movements,**
31 **threatened species, *Glyphis glyphis***

32 **Abstract**

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

56

57 **Significance Statement**

58

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

71

72

73 **Introduction**

74

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

89

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

104

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

121

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

136

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

153

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

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

174

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

180

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

192

193

194 **Material and methods**

195

196 *Sampling and DNA extraction*

197

198 *Glyphis glyphis* samples were collected between January 2012 and December
199 2014 in the Alligator Rivers system (South Alligator n=82; East Alligator n=6;
200 West Alligator n=1) and the Adelaide River (n=142) of the Northern Territory
201 (NT), and the Wenlock River system (n=125) of Queensland (QLD), northern
202 Australia (Fig. 1). Sharks were caught by rod and line or gillnet. Each shark was
203 measured, sexed and a small fin clip was taken from the inner pectoral fin before
204 it was released at the site of capture. Sampled sharks were from the size range
205 49–195 cm total length (TL) representing neonates through to subadults. Size at
206 birth is ~50–65 cm TL (31) and with a median size of sampled sharks of 82.75
207 cm TL, most represented neonates or 1+ juveniles. Sharks were sampled under
208 Northern Territory Fisheries Special Permit S17/3252, Kakadu National Park
209 Research Permit RK805, Queensland Fisheries General Research Permit 163582,
210 and Charles Darwin University Animal Ethics Committee A11041. DNA was
211 extracted using either the DNeasy Blood and Tissue kits (Qiagen) or the
212 NucleaMag Tissue kits (Macherey-Nagel).

213

214 *SNP genotyping*

215

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

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

225

226 *SNP filtering*

227

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

238

239 Departure from Hardy-Weinberg equilibrium (HWE) was then tested for each
240 locus within each sampling location using the "HWE.test.genind" function in the
241 Adegenet R package (36) and the false discovery rate method was applied to
242 control for multiple comparison testing (37).

243 We used the R package OutFlank (38) to identify outlier loci putatively under the
244 influence of directional selection. The approach implemented in Outflank is
245 based on an improved method for deriving the null distribution of population
246 differentiation for neutral loci. It results in fewer false positives than other
247 outlier tests, which are more influenced by the effects of demographic history
248 (39). We ran Outflank with 5% left and right trim for the null distribution of F_{ST} ,
249 minimum heterozygosity for loci of 0.1, and a 5% false discovery rate (q value).

250

251 *Mitogenome sequencing*

252

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

262

263 *Kinship analyses and fish filtering*

264

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

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

273

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

280

281 *Population structure analysis*

282

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

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

291

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

307

308 **Results**

309

310 *SNP filtering*

311

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

317 downstream kinship and population analysis consisted of 1330 SNPs.
318 Descriptive statistics including allelic richness (AR), observed heterozygosity
319 (H_o), expected heterozygosity (H_e) and inbreeding coefficient (F_{is}) are given in
320 Supplementary Material S3.

321

322 *Kinship analyses*

323

324 A total of 72 FS pairs (94 unique individuals) were identified, of which 12, 11 and
325 49 originated from the Adelaide, Alligator and Wenlock Rivers, respectively. No
326 cross-river FS pairs were identified. A total of 145 HS pairs (179 unique
327 individuals) were identified, 44 within the Adelaide River, 14 within the Alligator
328 Rivers, 69 within the Wenlock River, and 18 split across the Adelaide and
329 Alligator Rivers (Table 1).

330

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

341

342 *Population differentiation*

343

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

357

358 *Clustering analyses*

359

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

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

372

373 **Discussion**

374

375 For the first time in any elasmobranch species, whole mitogenome sequencing
376 and genotyping-by-sequencing genome scans have been used in combination to
377 analyse population connectivity. Our results reveal that a significant fraction of
378 the *G. glyphis* individuals analysed from all three rivers were close kin (26% FS;
379 50% HS). This observation permits direct estimation of recent adult (breeding)
380 and juvenile movements in this threatened species. In addition, the identification
381 of kin means that long-term connectivity estimated from population subdivision
382 can be made from juveniles without the family sampling bias that may be
383 common in population genetic datasets (14).

384

385 *Direct estimate of contemporary connectivity*

386

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

395

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

409

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

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

423

424 Given that the juveniles don't or very rarely move between rivers, the
425 distribution of HS provides insight into the movements of adults between
426 reproductive events. Out of the 121 cross-cohort HS pairs, 103 (85%) were
427 captured within the same river, indicating that in most cases at least one parent
428 returned to reproduce in the same river across breeding seasons. The remaining
429 18 (15%) cross-cohort HS pairs were shared between the Adelaide and the
430 Alligator Rivers. In these cases, at least one parent had moved between these
431 rivers (or their associated mating aggregation areas if gamete exchange occurs
432 outside the river) to reproduce. Van Diemen Gulf is a relatively small system, and
433 it is possible that adults from different rivers flowing in the gulf mix in this area.
434 In contrast to the cross-cohort HS pairs, same-cohort HS pairs were never
435 captured between rivers. Assuming females only breed once a year, this suggests
436 that males do not reproduce with females going to pup in different rivers within
437 the same year. Based on the variability in reproductive periodicity of Australian
438 carcharhinids of similar or smaller size, minimum reproductive periodicity
439 would be annual (48, 49), or potentially biennial given large size at maturity
440 (50). Hence, it is likely that the Adelaide and Alligator Rivers populations have
441 different mating aggregation areas. Once fish can be aged accurately,
442 reproductive periodicity could be determined by examining the time gap
443 between HS pairs.

444

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

450

451 *Population structure when sampling families*

452

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

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

465

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

477

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

491

492

493 *Direct versus indirect connectivity estimates and management implications*

494

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

506

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

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

514

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

523

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

535

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

546

547 Acknowledgments:

548

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

561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609

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, Burrridge 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.

- 659 37. Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: a
660 practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. Ser. A.*
661 (*Stat. Soc.*) 57(1):289-300.
- 662 38. Whitlock MC & Lotterhos KE (2015) Reliable detection of loci responsible
663 for local adaptation: inference of a null model through trimming the
664 distribution of *F_{st}*. *Am. Nat.* 186:S24-36.
- 665 39. Lotterhos KE & Whitlock MC (2015) The relative power of genome scans
666 to detect local adaptation depends on sampling design and statistical
667 method. *Mol. Ecol.* 24(5):1031-1046.
- 668 40. Chen X, Liu M, Grewe PM, Kyne PM, & Feutry P (2014) Complete
669 mitochondrial genome of the Critically Endangered speartooth shark
670 *Glyphis glyphis* (Carcharhiniformes: Carcharhinidae). *Mitochondr. DNA*
671 25(6):431-432.
- 672 41. Jones OR & Wang J (2010) COLONY: a program for parentage and sibship
673 inference from multilocus genotype data. *Mol. Ecol. Resour.* 10(3):551-
674 555.
- 675 42. Pembleton LW, Cogan NO, & Forster JW (2013) StAMPP: an R package for
676 calculation of genetic differentiation and structure of mixed - ploidy level
677 populations. *Mol. Ecol. Resour.* 13(5):946-952.
- 678 43. Evanno G, Regnaut S, & Goudet J (2005) Detecting the number of clusters
679 of individuals using the software STRUCTURE: a simulation study. *Mol.*
680 *Ecol.* 14(8):2611-2620.
- 681 44. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, & Mayrose I (2015)
682 CLUMPAK: a program for identifying clustering modes and packaging
683 population structure inferences across K. *Mol. Ecol. Resour.*
- 684 45. Kanno Y, Vokoun JC, & Letcher BH (2011) Sibship reconstruction for
685 inferring mating systems, dispersal and effective population size in
686 headwater brook trout (*Salvelinus fontinalis*) populations. *Conserv. Genet.*
687 12(3):619-628.
- 688 46. Tillett BJ, Meekan MG, Field IC, Hua Q, & Bradshaw CJ (2011) Similar life
689 history traits in bull (*Carcharhinus leucas*) and pig-eye (*C. amboinensis*)
690 sharks. *Mar. Freshwat. Res.* 62(7):850-860.
- 691 47. Australian Government (2015) Our north, our future: white paper on
692 developing Northern Australia. Available at:
693 [http://industry.gov.au/ONA/WhitePaper/Documents/northern_australia](http://industry.gov.au/ONA/WhitePaper/Documents/northern_australia_white_paper.pdf)
694 [white_paper.pdf](http://industry.gov.au/ONA/WhitePaper/Documents/northern_australia_white_paper.pdf).
- 695 48. Chin A, Simpfendorfer C, Tobin A, & Heupel M (2013) Validated age,
696 growth and reproductive biology of *Carcharhinus melanopterus*, a widely
697 distributed and exploited reef shark. *Mar. Freshwat. Res.* 64(10):965-975.
- 698 49. Harry AV, Tobin AJ, & Simpfendorfer CA (2013) Age, growth and
699 reproductive biology of the spot-tail shark, *Carcharhinus sorrah*, and the
700 Australian blacktip shark, *C. tilstoni*, from the Great Barrier Reef World
701 Heritage Area, north-eastern Australia. *Mar. Freshwat. Res.* 64(4):277-
702 293.
- 703 50. Mcauley RB, Simpfendorfer C, Hyndes G, & Lenanton R (2007)
704 Distribution and reproductive biology of the sandbar shark, *Carcharhinus*
705 *plumbeus* (Nardo), in Western Australian waters. *Mar. Freshwat. Res.*
706 58(1):116-126.

- 707 51. Dudgeon CL, *et al.* (2012) A review of the application of molecular
708 genetics for fisheries management and conservation of sharks and rays. *J.*
709 *Fish Biol.* 80(5):1789-1843.
- 710 52. Daly-Engel TS, *et al.* (2012) Global phylogeography with mixed-marker
711 analysis reveals male-mediated dispersal in the endangered scalloped
712 hammerhead shark (*Sphyrna lewini*). *Plos One* 7(1):e29986.
- 713 53. Pardini AT, *et al.* (2001) Sex-biased dispersal of great white sharks.
714 *Nature* 412(6843):139-140.
715
716

717 Table 1. Intra and inter river number of full-sibling pairs (above) and cross-
 718 cohort half-sibling pairs + same-cohort half-sibling pairs (below).
 719

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

720
 721 Table 2. Mitogenome-based pairwise Φ_{ST} for all individuals (above) and the
 722 dataset without full-sibling and half-sibling pairs (below).
 723

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

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

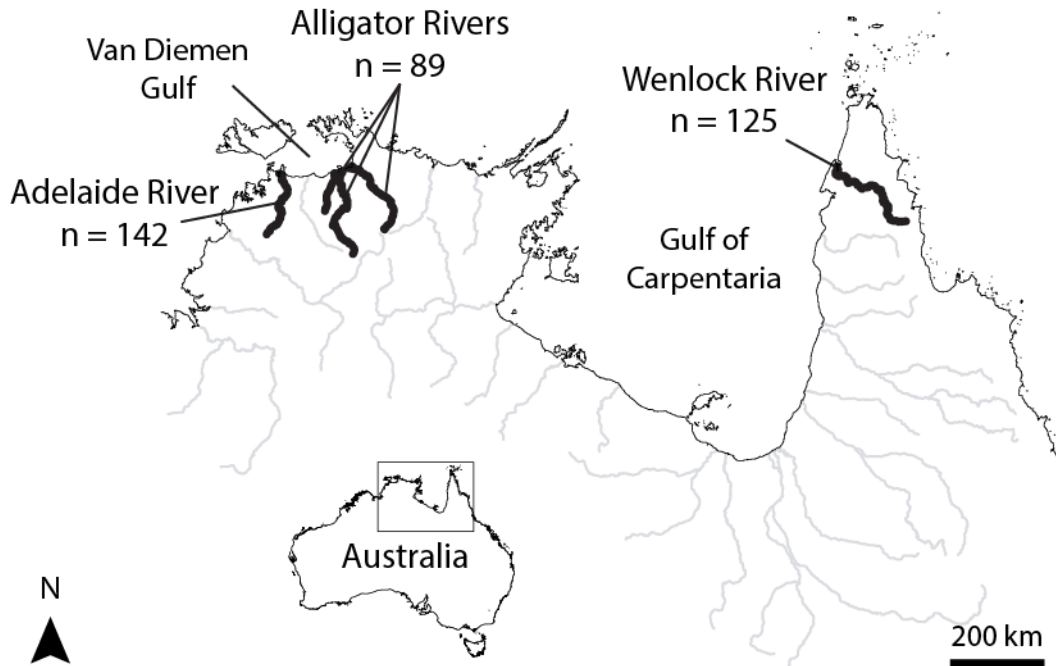
725
 726 Table 3. Nuclear SNP-based pairwise F_{ST} for all individuals (above) and the
 727 dataset without full-sibling and half-sibling pairs (below).
 728

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

729 ^{NS} P-value > 0.05; ** P-value < 0.0001

730
 731

732



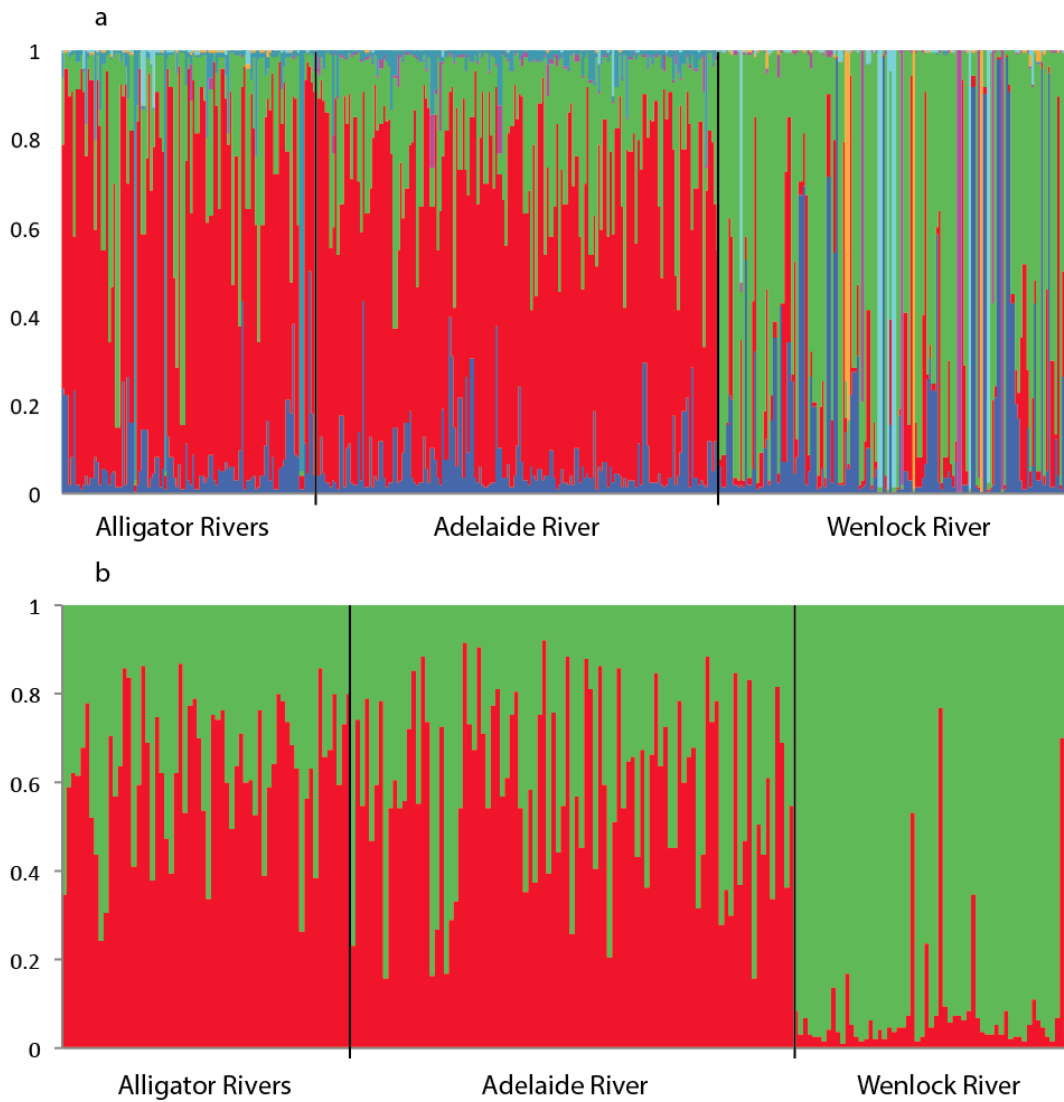
733

734

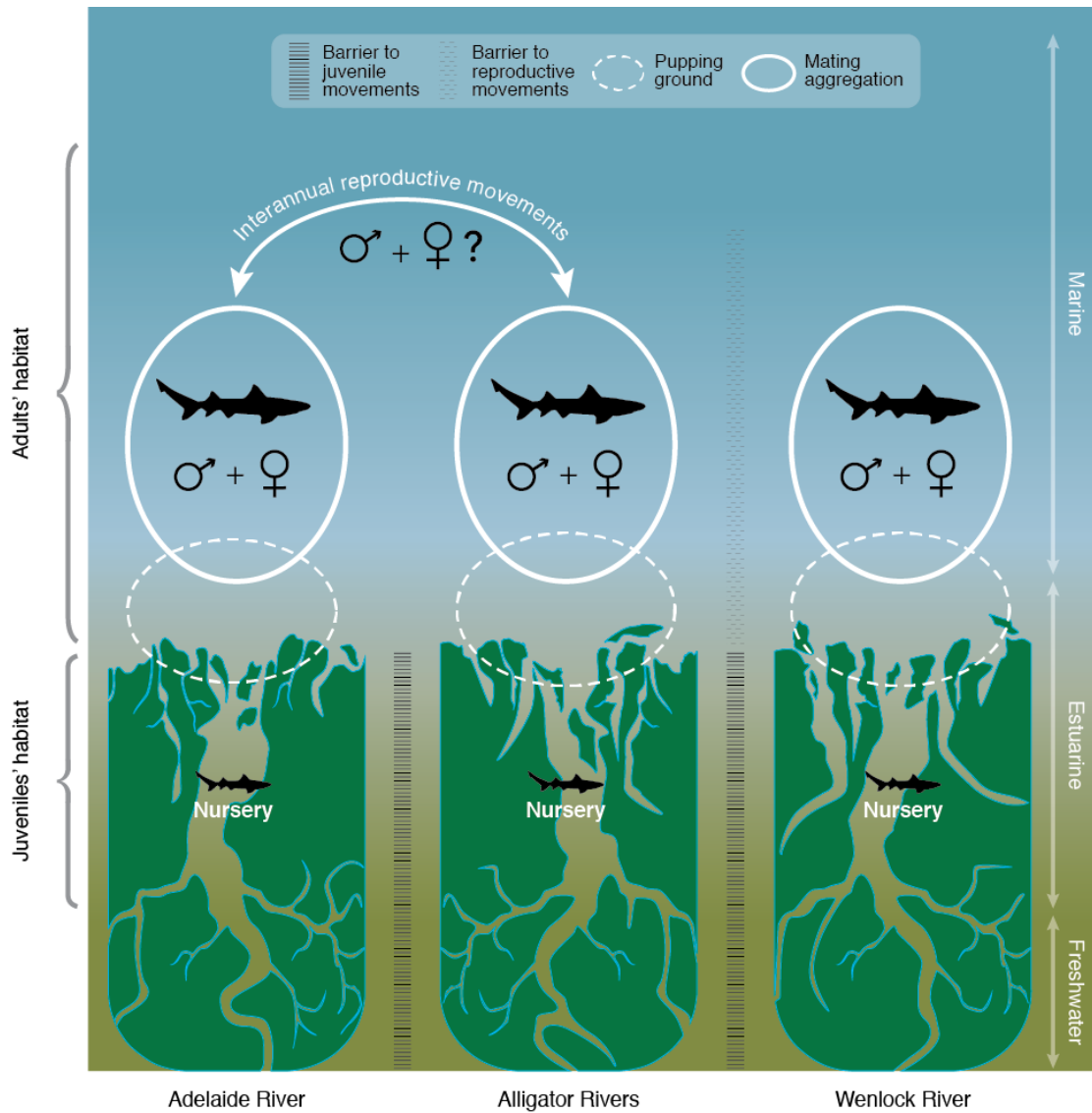
735

736

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



737 Alligator Rivers Adelaide River Wenlock River
 738 Figure 2. *Glyphis glyphis* STRUCTURE admixture analysis. Each cluster (K) is
 739 designated by a different colour. Each vertical bar represents one individual,
 740 partitioned according to admixture proportion from each cluster. a) Analysis of
 741 dataset with all samples, most likely K=7. b) Analysis of dataset without full-
 742 sibling and half-sibling pairs, most likely K=2.
 743



744
745
746
747

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