MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 5877–5885 doi: 10.1111/mec.13441

Selection and sex-biased dispersal in a coastal shark: the influence of philopatry on adaptive variation

D. S. PORTNOY,* J. B. PURITZ,* C. M. HOLLENBECK,* J. GELSLEICHTER,† D. CHAPMAN‡ and J. R. GOLD*

*Department of Life Sciences, Marine Genomics Laboratory, Harte Research Institute, Texas A&M University-Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412, USA, †University of North Florida, 1 UNF Drive, Jacksonville, FL 32224, USA, ‡Stony Brook University, Stony Brook, NY 11776, USA

Abstract

Sex-biased dispersal is expected to homogenize nuclear genetic variation relative to variation in genetic material inherited through the philopatric sex. When site fidelity occurs across a heterogeneous environment, local selective regimes may alter this pattern. We assessed spatial patterns of variation in nuclear-encoded, single nucleotide polymorphisms (SNPs) and sequences of the mitochondrial control region in bonnethead sharks (*Sphyrna tiburo*), a species thought to exhibit female philopatry, collected from summer habitats used for gestation. Geographic patterns of mtDNA haplotypes and putatively neutral SNPs confirmed female philopatry and male-mediated gene flow along the northeastern coast of the Gulf of Mexico. A total of 30 outlier SNP loci were identified; alleles at over half of these loci exhibited signatures of latitude-associated selection. Our results indicate that in species with sex-biased dispersal, philopatry can facilitate sorting of locally adaptive variation, with the dispersing sex facilitating movement of potentially adaptive variation among locations and environments.

Keywords: elasmobranchs, genome scan, localized adaptation, male-mediated gene flow

Received 13 July 2015; revision received 27 October 2015; accepted 27 October 2015

Introduction

Sex-biased dispersal arises when individuals of one sex exhibit site fidelity (philopatry), while individuals of the opposite sex are prone to disperse (Pusey 1987). This occurs in a wide variety of vertebrate taxa (e.g. birds, Clarke *et al.* 1997; mammals, Lawson Handley & Perrin 2007) and is thought to result from fitness differences between the sexes associated with local competition for resources (including mates), inbreeding avoidance, and/or parental investment (Gandon 1999; Perrin & Mazalov 2000). There also is a relationship between mating system and which sex is dispersive; monogamous species feature territorial males and dispersive females, while polygamous species feature female philopatry and male dispersal (Greenwood 1980).

Correspondence: David S. Portnoy, Fax: 361-825-2025; E-mail: david.portnoy@tamucc.edu

Dispersal and resulting gene flow acts as a homogenizing force across the genome, opposed by the processes of genetic drift and disruptive selection. The level of gene flow necessary to counteract genetic drift can be relatively small, as large populations experience little drift and only a few migrants are required in small populations (Wright 1931; Slatkin 1985). Disruptive selection, on the other hand, is capable of generating divergence in specific genomic regions, even when gene flow is high, if the strength of selection is high relative to the number of immigrants and/or the patterns of immigration are nonrandom in relation to local environmental conditions (Endler 1973; Slatkin 1987; Garant *et al.* 2007). Sex-biased dispersal, via gene flow through the dispersive sex, has a homogenizing effect on biparentally inherited nuclear variation; uniparentally inherited markers not under disruptive selection (e.g. heterologous sex chromosomes, mtDNA) sort through the philopatric sex and may depart from homogeneity at a greater rate through time (Avise 1994). If habitats of a philopatric species vary in environmental conditions,

the homogenizing effects of sex-mediated gene flow may be counteracted in specific genomic regions if localized selection leads to increased reproductive success for the philopatric sex (Lenormand 2002). Finally, philopatric behaviour by one of the sexes can reduce the strength of migration, facilitating local adaptation (Slatkin 1987).

Studies in several species of live-bearing sharks have revealed spatial genetic patterns (homogeneity in nuclear-encoded microsatellites and heterogeneity in maternally inherited mtDNA) consistent with female philopatry and male-mediated gene flow (Portnoy & Heist 2012; Chapman *et al.* 2015). Females in these species exhibit considerable parental investment, giving birth after long gestation periods to small litters of fully developed offspring, suggesting that return to a favourable habitat could enhance embryonic growth during gestation (Economakis & Lobel 1998; Driggers *et al.* 2014) as well as provide predictable access to food and shelter from predators (Heupel *et al.* 2007). It also is known that habitats used by the same species for gestation and/or parturition may differ substantially, even at small spatial scales (DiBattista *et al.* 2007; Feldheim *et al.* 2014). Based on the above, coastal philopatric sharks represent a good model system to assess possible effects that localized adaptation may have on genomewide patterns of variation in the context of sex-asymmetric gene flow.

We assessed spatial patterns of variation in nuclear-encoded single nucleotide polymorphisms (SNPs) and sequences of the mitochondrial control region in bonnethead sharks (*Sphyrna tiburo*), a species thought to exhibit female-biased philopatry (Driggers *et al.* 2014). Bonnetheads are common seasonal residents in coastal and estuarine waters of the western Atlantic Ocean (Atlantic), including the Gulf of Mexico (Gulf), and are known to use nearshore habitat for gestation and parturition (Compagno 1984; Driggers *et al.* 2014). Bonnetheads in the Atlantic and Gulf migrate seasonally, and a variety of life stages are commonly found in bays, estuaries, and nearshore waters from May to November (Cortes *et al.* 1996; Ulrich *et al.* 2007). The species has a short gestation period of 4–5 months (Parsons 1993), with parturition occurring in the late summer to early fall and mating occurring shortly thereafter (Manire & Rasmussen 1997; Ulrich *et al.* 2007). Sperm storage is necessary, as ovulation does not occur until spring (Manire *et al.* 1995). Bonnetheads mature between 1–7 years (Lombardi-Carlson *et al.* 2003; Frazier *et al.* 2014) and females give birth to 2–14 (avg. ~9) fully developed pups (Frazier *et al.* 2013). Unlike other coastal sharks, the observed migratory behaviour does not appear to be associated with the use of nursery areas (Heupel *et al.* 2007), but instead may be related to increasing food availability for gestating females and gaining access to potential mates for males (Driggers *et al.* 2014). Significant differences in life history among bonnetheads across small geographic regions have been documented in several studies; differences found between samples from the eastern Gulf include size at age, growth rate, and size and age at maturity (Parsons 1993; Carlson & Parsons 1997; Lombardi-Carlson *et al.* 2003). In addition, studies in both the Atlantic and eastern Gulf have shown site fidelity by adult bonnetheads (particularly females) to particular estuaries or bays during the summer months, on intra- and interannual timescales (Heupel *et al.* 2006; Driggers *et al.* 2014).

We sampled adult and subadult animals from three localities along the west coast of Florida (eastern Gulf of Mexico) and one locality off the coast of North Carolina (western Atlantic Ocean). Sample localities in the Gulf were selected because of identified latitudinal differences in life history parameters among bonnetheads in the region (Lombardi-Carlson *et al.* 2003); the sample from the Atlantic was included to have a sample outside the Gulf and because of identified differences in life history between bonnetheads in the Gulf and Atlantic (Frazier *et al.* 2014). We used a ddRAD approach (Peterson *et al.* 2012) to genotype individuals at thousands of nuclear-encoded SNPs, permitting a search for spatial differences in genomic regions putatively under selection; inclusion of putatively neutral SNPs and mtDNA sequences allowed us to assess further whether dispersal in bonnetheads is sex-biased.

Materials and methods

Tissues (fin clips) from 134 bonnetheads sampled between 1998 and 2000 from four nearshore localities (Fig. 1) were used in the study. Samples were obtained during the summer months (May to September) when mature individuals are in areas used for gestation, parturition, and mating. Individuals sampled were mostly a mix of mature females and males.

Double-digest RAD (ddRAD) libraries were prepared following Peterson *et al.* (2012); details of the protocol may be found in the Appendix S1 (Supporting information). Libraries were sequenced on two lanes of an Illumina HiSeq 2000 DNA sequencer. The first library was sequenced as a paired-end run for reference contig assembly in order to facilitate downstream bioinformatics inference. The second library was sequenced as a single-end run, as a cost-effective manner to genotype SNPs. The *dDocent* pipeline (Puritz *et al.* 2014) was used for reference contig assembly, read mapping, and SNP genotyping. Default parameters were used for each step, with the exception of contig assembly, where a customized script was used to mitigate the high levels

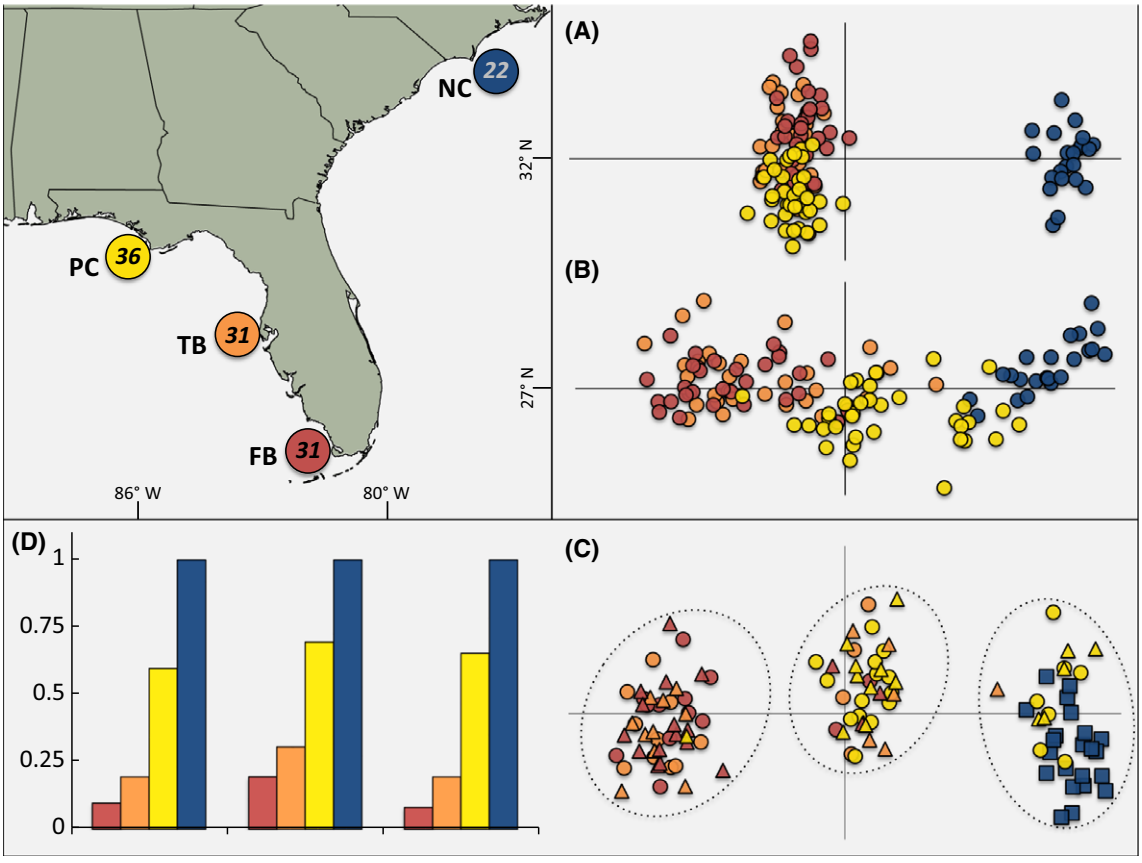

Fig. 1 Samples of bonnethead sharks obtained off North Carolina (NC, blue), Florida Bay (FB, red; 18 males, 13 females), Tampa Bay (TB, orange; 17 males, 14 females) and Panama City (PC, yellow; 15 males, 21 females). Results of discriminant analysis of principle components for (A) putatively neutral N-SNP loci, (B) outlier O-SNP loci putatively under selection, with group membership defined by sample locality, and (C) outlier O-SNP loci putatively under selection, with group membership based on *k*-means clustering. Females are coded as circles, males as triangles, and individuals of unknown sex as squares. Representative allele frequencies (D) of three O-SNP loci (left to right, E66074, E109425, E106435) that contributed ~24% to the distribution of individuals along the *X*-axis. Colours represent sample locations for all figures. SNP, single nucleotide polymorphism.

of repeats and duplications expected in large genomes. The initial set of data consisted of 648 035 variant SNP loci across 147 920 fragments.

The entire mitochondrial control region (1134 bp) was amplified using primers Pro-L and 282H (Keeney *et al.* 2003); details of the protocol may be found in the Appendix S1 (Supporting information). Electropherotograms were examined by eye, aided by GENEIOUS v.7.1 (Biomatters Ltd.); all sequences were trimmed to 1064 bp due to occasional nonspecific amplification on the 3′ end that made accurate base calling difficult.

Single nucleotide polymorphisms were extensively filtered before further analysis. The initial raw data set was filtered to remove all genotypes with <5 reads per individual and loci called in <75% of all individuals. Consequently, only the top 90% of individuals in genotype call rate were retained. The resulting data set contained 121 individuals. SNPs were then filtered to meet the following criteria: presence in 97.5% of individuals across the data set, minor allele frequency >5% across the data set, and conformance to expectations of Hardy–Weinberg equilibrium (HWE). Additional parameters considered during filtering included allele balance within heterozygous individuals, SNP quality to depth ratio, percentage of contribution from forward and reverse reads, maximum mean read depth across individuals, and removal of possible paralogs (Details on SNP filtering are described in Appendix S1, Supporting information). The final, filtered data set consisted of 5914 SNPs spread across 3967 fragments.

Genetic diversity (nuclear genome) within each locality was assessed as the mean nucleotide diversity (π) across all SNPs, using VCFTOOLS (Danecek *et al.* 2011). Homogeneity of π across localities was assessed using analysis of variance (ANOVA) and Tukey–Kramer HSD independent contrasts as implemented in JMP® v.11 (SAS Institute Inc.). Genetic diversity (mtDNA) was assessed as mean nucleon (*h*) and nucleotide diversity (π) within each

locality, using ARLEQUIN v.3.5.1.2 (Excoffier & Lischer 2010).

Relatedness of individuals within each locality was assessed in VCFTOOLS, using the statistic developed by Yang *et al.* (2010). Two individuals in the sample from Florida Bay (FB) possessed high relatedness to each other (0.61) relative to the average relatedness (−0.045) across all individuals, suggesting these two individuals shared parents. The individual with more missing data was removed from subsequent SNP-based analyses to avoid possible issues with consanguinity. SNPs were then organized into haplotypes (loci), using a custom Perl script that produces output in GENEPOP format. During haplotyping, a total of 23 loci were excluded from further analysis; 12 were identified as possible paralogs and 11 could not be haplotyped in more than 90% of individuals assayed. GENEPOP files were converted to BAYESCAN format, using PGDSPIDER v.2.0.7 (Lischer & Excoffier 2012), and BAYESCAN (Foll & Gaggiotti 2008) was used to identify individual outlier loci by assessing fit to different models of selection. The program was run with all default values, with the exception of 30 pilot runs and a thinning interval of 50; significance of outlier loci was determined using a *q*-value that directly corresponded to a false discovery rate (FDR) of 0.05. Loci were then divided into two sets: one that contained putatively neutral SNPs (N-SNP loci) and one that contained outlier SNPs (O-SNP loci) putatively under selection.

Geographic homogeneity among localities in N-SNP and O-SNP loci was tested using single-level analysis of molecular variance (AMOVA), as implemented in GENODIVE v.2.0 (Meirmans & van Tienderen 2004). Pairwise F_{ST} values (both nuclear data sets) were estimated using GENODIVE; significance of pairwise F_{ST} values was assessed by permuting individuals between samples 10 000 times. Homogeneity of mtDNA haplotypes among localities was tested using single-level AMOVA, as implemented in ARLEQUIN. Distances were calculated using a Kimura 2-parameter model (Kimura 1980), as selected by JMODELTEST v. 2.1.4 (Guindon & Gascuel 2003; Darriba *et al.* 2012). Pairwise Φ_{ST} values were estimated using ARLEQUIN, with significance determined by permuting individuals between samples 10 000 times. Correction for multiple testing was implemented using the FDR procedure (Benjamini & Hochberg 1995).

Discriminant Analysis of Principle Components (DAPC; Jombart *et al.* 2010) was carried out on both N-SNP and O-SNP loci, using the ADEGENET package (Jombart & Ahmed 2011) in R v.3.0.2 (R Development Core Team 2013), with group membership defined by locality. DAPC also was carried out on O-SNP loci, with group membership inferred using *k*-means clustering (MacQueen 1967); contribution of O-SNP loci to genetic clustering was then inferred from loading variables used in each discriminant function. For all O-SNP loci, the reference contig, assembled from paired-end reads, was screened against the NCBI nucleotide-read database, using the BLASTN algorithm (Altschul *et al.* 1990). The top three hits with *E*-values < 0.01 were recorded.

Results

Summary statistics for SNPs and mtDNA are given in Table S1 (Supporting information); GenBank accession numbers and geographic distribution of mtDNA haplotypes are given in Table S2 (Supporting information). Estimated mean nucleotide diversity (π) across all SNP loci per sample (±SE) varied from 0.296 (±0.002) in the sample from North Carolina (NC) to 0.319 (±0.002) in the sample from FB. Mean estimates of π differed significantly across samples ($F_3 = 21.483$, $P < 0.001$), with mean π in NC being significantly lower than in the other samples (Tukey–Kramer HSD—$P < 0.001$). The same pattern was observed in haplotype diversity of mtDNA sequences; estimated diversity was lower in NC ($h = 0.719 \pm 0.077$), while h values did not differ among the other three samples.

A total of 30 haplotypes, containing 49 O-SNPs, were identified as candidate loci under selection ($q < 0.05$); the remaining SNPs (5865 scattered across 3910 haplotypes) were consistent with a neutral model. A total of 72 alleles were identified among the 30 O-SNP loci; 21 loci were bi-allelic, while nine were multi-allelic (Table S3, Supporting information). Significant heterogeneity among all four localities in all three marker types was detected by AMOVA (Table S4, Supporting information); the proportion of the total genetic variance explained by geography (locality) was 0.79% (N-SNP loci), 7.77% (mtDNA haplotypes), and 27.07% (O-SNP loci). Pairwise estimates of F_{ST} and Φ_{ST} (Table 1) revealed differences among the three marker types. For N-SNP loci, allele frequencies in NC differed significantly from those in FB, TB (Tampa Bay), and PC (Panama City); allele frequencies in the latter three were homogeneous. For mtDNA, the haplotype distribution in NC differed significantly from those in FB, TB, and PC; estimates of Φ_{ST} between FB and PC differed significantly from one another, while those between FB and TB and TB and PC were homogeneous. Allele frequencies of O-SNP loci in both NC and PC differed significantly from one another and from those in FB and TB, while allele frequencies in FB and TB were homogeneous. Significant heterogeneity among the three localities in the Gulf also was detected by AMOVA for mtDNA haplotypes ($F_{ST} = 0.027$, $P = 0.033$) and O-SNP loci ($F_{ST} = 0.157$, $P = 0.000$), but not for N-SNP loci ($F_{ST} = 0.0003$, $P = 0.151$).

Table 1 Below diagonal: pairwise F_{ST} values for putatively neutral SNP loci (N-SNP) and for outlier SNP loci putatively under selection (O-SNP), and pairwise Φ_{ST} values for mtDNA haplotypes (mtDNA), between samples of bonnetheads obtained off North Carolina (NC), Florida Bay (FB), Tampa Bay (TB) and Panama City (PC). Above diagonal: probability (*P*) values; those significant after correction for multiple comparisons are given in bold

	N-SNP					O-SNP					mtDNA			
	NC	FB	TB	PC		NC	FB	TB	PC		NC	FB	TB	PC
NC	—	**<0.001**	**<0.001**	**<0.001**	NC	—	**<0.001**	**<0.001**	**<0.001**	NC	—	**<0.001**	**0.001**	**0.014**
FB	**0.019**	—	0.317	0.038	FB	**0.543**	—	0.382	**<0.001**	FB	**0.234**	—	0.158	**0.011**
TB	**0.021**	0.000	—	0.344	TB	**0.462**	0.000	—	**<0.001**	TB	**0.161**	0.014	—	0.406
PC	**0.021**	0.001	0.000	—	PC	**0.180**	**0.244**	**0.177**	—	PC	**0.064**	**0.055**	0.000	—

Analysis of N-SNP loci, using DAPC and with prior group membership defined by locality, revealed two distinct clusters along the primary (*X*) axis (Fig. 1A); one was comprised of individuals from NC, while the other contained individuals from the three localities in the Gulf. Analysis of O-SNP loci, with prior group membership defined by locality, revealed a different pattern along the primary axis (Fig. 1B). Twelve individuals from PC clustered with individuals in the sample from NC, while the remaining individuals formed a second cluster; both clusters were more diffuse than in the analysis of N-SNP loci. When prior group membership of O-SNP loci was inferred using *k*-means clustering, three distinct clusters were revealed in DAPC analysis (Fig. 1C). One cluster contained primarily individuals from NC and PC and one individual from TB; one cluster contained individuals from the Gulf, primarily from PC; and one cluster contained mostly individuals from FB and TB and one individual from PC. The primary (*X*) axis described 99.6% of the variance. Allele frequencies at three representative O-SNP loci (Fig. 1D) clearly reveal a clinal, north–south (latitudinal) pattern in allele frequencies. The correlation between allele (haplotype) frequencies at each O-SNP locus and latitude was then evaluated using standard least squares regression as implemented in JMP v.11. Alleles at 17 O-SNP loci were correlated ($P \leq 0.05$) with latitude and explained 56.9% of the variation along the primary axis, while 18 O-SNP loci had r^2 values ≥ 0.90 and explained 75.6% of the variation along the *X*.

Eight of the 30 O-SNP loci had no sequence counterpart in GenBank; the remaining 22 were highly similar (*E*-value < 0.01) to several DNA sequences (Table S5, Supporting information). Frequent 'hits' included sequence similarities to clones or contigs in other species, and to annotated genomic regions of known immune response proteins (e.g. cytokines MIP-3 and interleukin-1β and a T cell receptor), putative regulatory elements (e.g. zinc-finger proteins, Hox genes), and SINE-type sequences.

Discussion

The significant difference in N-SNP loci between bonnetheads from the Atlantic and Gulf indicates genetically distinct populations with little to no gene flow between the two regions. This geographic pattern has been observed in other marine taxa (Avise 1992; Gold & Richardson 1998; Gold *et al.* 2009) including coastal sharks (Portnoy *et al.* 2014) and supports results from a recent mtDNA assessment of population structure in the bonnethead (Escatel-Luna *et al.* 2015). This pattern is hypothesized to stem from biogeographic processes associated with the Florida Current and/or narrowing of the continental shelf in southeastern Florida (Portnoy *et al.* 2014). The absence of significant divergence in N-SNP loci among the three localities in the Gulf is consistent with gene flow occurring between the Florida Keys (FB) and north-central Florida (PC).

Asymmetry in geographic patterns of variation between N-SNP loci (homogeneous) and mtDNA haplotypes (heterogeneous) among bonnetheads from the Gulf is consistent with female philopatry and male-biased dispersal (Melnick & Hoelzer 1992). Because mtDNA is haploid and uniparentally inherited, a greater magnitude of divergence at mtDNA compared to nuclear loci is to be expected (Birky 2001). Similar patterns are documented in several shark species (Portnoy & Heist 2012; Chapman *et al.* 2015) and interannual tag-and-recapture studies of bonnetheads (Driggers *et al.* 2014) demonstrate strong site fidelity of females to specific estuaries. The pattern of mtDNA haplotype variation among bonnetheads in the Gulf indicates an isolation-by-distance effect rather than complete isolation as mtDNA haplotypes in the intermediate sample locality (TB) did not differ significantly from those in sample localities (PC and FB) at the geographic extremes. This also suggests that female bonnetheads may stray from preferred localities but most likely to neighbouring ones.

The largest proportion of the genetic variance explained by locality (geography) was due to O-SNP loci. In theory,

outlier loci can reflect genomic regions associated with local adaptive differences (Nielsen *et al.* 2009; Allendorf *et al.* 2010) or genomic regions that have diverged more than expected over time via a nonadaptive process such as genetic drift (Hedrick 2011). However, genetic drift is a genomewide effect (Luikart *et al.* 2003) and the significant correlations between allele frequencies at O-SNP loci and latitude and the complete absence of any clinal pattern in N-SNP loci indicate that the observed geographic pattern of O-SNP loci stems from localized divergent selection. The greater similarity in allele frequencies at outlier O-SNP loci between PC and NC also supports divergent selection associated with latitude as the two localities are situated at more northerly latitudes yet are at the geographic extreme of possible (homogenizing) gene flow among the localities studied.

Signatures of latitude-driven selection are common given that natural phenomena (e.g. climate, diurnal cycle) impact distributions of biological organisms, and that selection is imposed by the local biotic environment and interactions between a focal population and other organisms (Kawecki & Ebert 2004). Examples of well-known latitude-specific effects on marine fish include demographic traits such as growth rate (Conover & Present 1990) and host–parasite/pathogen systems (Poulin & Morand 2000). A few of the O-SNP loci found in this study did have sequence similarities to regions of genes putatively involved in regulation and development, and there are significant latitudinal differences in growth rate and size at age among bonnetheads in the region of the Gulf sampled (Lombardi-Carlson *et al.* 2003). A larger proportion of the O-SNP loci had sequence similarities to regions of genes involved in immune response. This result might reflect latitudinal variation in parasite infectivity (Poulin & Morand 2000) and increased infectivity of parasites to sympatric hosts rather than allopatric hosts of the same species (Morand *et al.* 1996). Some caution in interpreting these data, however, is advisable, in part because the O-SNP loci sequences were small in size, and in part because the majority of SNPs recovered using a ddRAD approach are not within protein-coding genes (Baxter *et al.* 2011). Further, while we found a general correlation of allele frequencies with latitude for O-SNPs, this does not demonstrate causation as other factors may be equally or more important. As an example, the spatial sampling encompasses both the warm-temperate and tropical provinces along the Florida coast, and differences in allele frequencies could reflect differences in ecology and climate.

Occurrence of philopatry in association with a nonrandom pattern of geographic variation in small genomic regions was reported recently (Stiebens *et al.* 2013) in a study of variation in MHC alleles among philopatric loggerhead turtles in the Cape Verde Archipelago. Both mtDNA haplotypes and MHC alleles were structured genetically among nesting islands, but only nuclear-encoded microsatellites followed a geographic pattern, in this case one of isolation by distance indicative of restricted male dispersal. In our study, only females appeared structured geographically along the western coast of Florida. These and tagging data (Driggers *et al.* 2014) where >95% of interannual bonnethead returns to the same estuary were female indicate that bonnethead males are less philopatric than females, and that maintenance of localized adaptive alleles in bonnetheads may occur through female matrilines. Thus, selection and sex-specific philopatry can interact to sort adaptive nuclear alleles across geographic space.

Association of spatially discrete matrilines and localized genomic regions under selection suggest that female genotype and philopatry to gestational areas may increase offspring fitness as a maternal effect (Mousseau & Fox 1998; Badyaev & Uller 2009). This is consistent with a review of parental effects in species with sex-asymmetric dispersal and a model that showed that selective pressure to develop locally adaptive parental effects is high when dispersal is sex-biased (Revardel *et al.* 2010). Unfortunately, studies of parental effects in sharks are limited (Hussey *et al.* 2010) despite a female reproductive biology (long gestation, live birth) in several species that suggests occurrence of important maternal effects.

Sex-specific philopatry reduces overall dispersal and consequently may redistribute genetic diversity among rather than within subpopulations or demes. In bonnethead sharks, homogeneity of N-SNP loci across geographic localities within the Gulf demonstrates that genetic diversity was partitioned equally within and among demes, indicating that extensive male dispersal was enough to overcome drift processes. In contrast, strong differentiation at a small subset of nuclear genes among samples collected at gestational areas indicates that localized selection was sufficiently strong to outweigh the homogenizing force of dispersal and gene flow. Thus, while female philopatry in bonnethead sharks may promote maintenance of adaptive alleles in specific localities, gene flow mediated by males or straying females could move potentially adaptive variation among environments (Slatkin 1987; Garant *et al.* 2007). Given local environmental heterogeneity on larger temporal scales, the maintenance and movement of potentially adaptive variation across the landscape likely facilitates species persistence (Bowen & Roman 2005).

Acknowledgements

We thank C.A. Manire for providing several samples from Florida and P. Bentzen and two anonymous reviewers for

helpful comments and suggestions. This work was supported by funds provided by the College of Science and Engineering at Texas A&M University-Corpus Christi and by the Environmental Protection Agency (Award #R826128-01-0 to C.A. Manire). Although the research described in this article was supported in part by the United States Environmental Protection Agency, it has not been subjected to the agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. Field sampling in Florida was conducted under research permits issued by the Florida Fish and Wildlife Conservation Commission to Mote Marine Laboratory. This article is publication number 10 of the Marine Genomics Laboratory at Texas A&M University-Corpus Christi and number 104 in the series Genetic Studies in Marine Fishes.

References

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **111**, 697–709.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Avise JC (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos*, **63**, 62–76.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, Inc., New York, New York.

Badyaev AV, Uller T (2009) Parental effects in ecology and evolution: mechanisms, processes and implications. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **364**, 1169–1177.

Baxter SW, Davey JW, Johnston JS *et al.* (2011) Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE*, **6**, e19315.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B (Statistical Methodology)*, **57**, 289–300.

Birky CWJR (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics*, **35**, 125–148.

Bowen BW, Roman J (2005) Gaia's handmaidens: the Orlog model for conservation biology. *Conservation Biology*, **19**, 1037–1043.

Carlson JK, Parsons GR (1997) Age and growth of the bonnethead shark, *Sphyrna tiburo*, from northwest Florida, with comments on clinal variation. *Environmental Biology of Fishes*, **50**, 331–341.

Chapman DC, Feldheim KA, Papastamatiou YP, Hueter RE (2015) There and back again: a review of residency and return migrations in sharks, with implications for population structure and management. *Annual Review of Marine Science*, **7**, 547–570.

Clarke AL, Sæther B-E, Røskaft E (1997) Sex biases in avian dispersal: a reappraisal. *Oikos*, **79**, 429–438.

Compagno LJV (1984) FAO Species Catalogue. Vol. 4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 2—Carcharhiniformes. *FAO Fisheries Synopsis*, **125**, 251–655.

Conover DO, Present TMC (1990) Countergradient variation in growth rate: compensation for length of the growing season among Atlantic silversides from different latitudes. *Oecologia*, **83**, 316–324.

Cortes E, Manire CA, Hueter RE (1996) Diet, feeding habits and diel feeding chronology of the bonnethead shark, *Sphyrna tiburo*, in southwest Florida. *Bulletin of Marine Science*, **58**, 353–367.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

DiBattista JD, Feldheim KA, Gruber SH, Hendry AP (2007) When bigger is not better: selection against large size, high condition and fast growth in juvenile lemon sharks. *Journal of Evolutionary Biology*, **20**, 201–212.

Driggers WB III, Frazier BS, Adams DH *et al.* (2014) Site fidelity of migratory bonnethead sharks *Sphyrna tiburo* (L. 1758) to specific estuaries in South Carolina, USA. *Journal of Experimental Marine Biology and Ecology*, **459**, 61–69.

Economakis AE, Lobel PS (1998) Aggregation behavior of the grey reef shark, *Carcharhinus amblyrhynchos*, at Johnston Atoll, Central Pacific Ocean. *Environmental Biology of Fishes*, **51**, 129–139.

Endler JA (1973) Gene flow and population differentiation. *Science*, **19**, 243–250.

Escatel-Luna E, Adams DH, Uribe-Alcocer M, Islas-Villanueva V, Díaz-Jaimes P (2015) Population genetic structure of the bonnethead shark, *Sphyrna tiburo*, from the western North Atlantic Ocean based on mtDNA sequences. *Journal of Heredity*, **106**, 355–365.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Feldheim KA, Gruber SH, DiBattista JD *et al.* (2014) Two decades of genetic profiling yields first evidence of natal philopatry and long-term fidelity to parturition sites in sharks. *Molecular Ecology*, **23**, 110–117.

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

Frazier B, Gelsleichter J, Gonzalez De Acevedo M (2013) *Preliminary data on the reproductive biology of the bonnethead (Sphyrna tiburo) from the southeast U.S. Atlantic coast*. SEDAR34-WP-22. SEDAR, Charleston, South Carolina, pp. 11.

Frazier BS, Driggers WB III, Adams DH, Jones CM, Loefer JK (2014) Validated age, growth and maturity of the bonnethead *Sphyrna tiburo* in the western North Atlantic Ocean. *Journal of Fish Biology*, **85**, 688–712.

Gandon S (1999) Kin competition the cost of inbreeding and the evolution of dispersal. *Journal of Theoretical Biology*, **200**, 345–364.

Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, **21**, 434–443.

Gold JR, Richardson LR (1998) Population structure in greater amberjack, *Seriola dumerili*, from the Gulf of Mexico and the western Atlantic Ocean. *Fisheries Bulletin*, **96**, 767–778.

Gold JR, Saillant E, Ebelt ND, Lem S (2009) Conservation genetics of gray snapper (*Lutjanus griseus*) in U.S. waters of the northern Gulf of Mexico and western Atlantic Ocean. *Copeia*, **2009**, 277–286.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**, 1140–1162.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Hedrick PW (2011) *Genetics of Populations*. Jones and Bartlett, Sudbury, Massachusetts.

Heupel MR, Simpfendorfer CA, Collins AB, Tyminski JP (2006) Residency and movement patterns of bonnethead sharks, *Sphyrna tiburo*, in a large Florida estuary. *Environmental Biology of Fishes*, **76**, 47–67.

Heupel MR, Carlson JK, Simpfendorfer CA (2007) Shark nursery areas: concepts, definition, characterization and assumptions. *Marine Ecology Progress Series*, **337**, 287–297.

Hussey NE, Wintner SP, Dudley SFJ, Cliff G, Cocks DT, MacNeil MA (2010) Maternal investment and size-specific reproductive output in carcharhinid sharks. *Journal of Animal Ecology*, **79**, 184–193.

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

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Keeney D, Heupel MR, Heuter RE, Heist EJ (2003) Genetic heterogeneity among blacktip shark, *Carcharhinus limbatus*, continental nurseries along the U.S. Atlantic and Gulf of Mexico. *Marine Biology*, **143**, 1039–1046.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Lawson Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology*, **2007**, 1559–1578.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183–189.

Lischer HEL, Excoffier L (2012) PGDSpider: ab automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Lombardi-Carlson LA, Cortés E, Parsons GR, Manire CA (2003) Latitudinal variation in life-history traits of bonnethead sharks, *Sphyrna tiburo* (Carcharhiniformes: Sphyrnidae) from the eastern Gulf of Mexico. *Marine and Freshwater Research*, **54**, 875–883.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

MacQueen JB (1967) Some methods for classification and analysis of multivariate observations. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 281–297.

Manire CA, Rasmussen LE (1997) Serum concentrations of steroid hormones in the mature male bonnethead shark, *Sphyrna tiburo*. *General and Comparative Endocrinology*, **107**, 414–420.

Manire CA, Rasmussen EL, Hess DL, Hueter RE (1995) Serum steroid hormones and the cycle of the female bonnethead shark, *Sphyrna tiburo*. *General and Comparative Endocrinology*, **97**, 366–376.

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

Melnick DJ, Hoelzer GA (1992) Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *International Journal of Primatology*, **13**, 379–393.

Morand S, Manning SD, Woolhouse MEJ (1996) Parasite-host coevolution and geographic patterns of parasite infectivity and host susceptibility. *Proceedings of the Royal Society of London B: Biological Sciences*, **263**, 119–128.

Mousseau TA, Fox CW (1998) The adaptive significance of maternal effects. *Trends in Ecology and Evolution*, **13**, 403–406.

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology*, **18**, 3128–3150.

Parsons GR (1993) Geographic variation in reproduction between two populations of the bonnethead shark, *Sphyrna tiburo*. *Environmental Biology of Fishes*, **38**, 25–35.

Perrin N, Mazalov V (2000) Local competition, inbreeding and the evolution of sex-biased dispersal. *The American Naturalist*, **155**, 116–127.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Portnoy DS, Heist EJ (2012) Molecular markers: progress and prospects for understanding reproductive ecology in elasmobranchs. *Journal of Fish Biology*, **80**, 1120–1140.

Portnoy DS, Hollenbeck CM, Belcher CN *et al.* (2014) Contemporary population structure and post-glacial genetic demography in a migratory marine species, the blacknose shark, *Carcharhinus acronotus*. *Molecular Ecology*, **23**, 5480–5495.

Poulin R, Morand S (2000) The diversity of parasites. *Quarterly Review of Biology*, **75**, 277–293.

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

Pusey AE (1987) Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology and Evolution*, **2**, 295–299.

R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org.

Revardel E, Franc A, Petit RJ (2010) Sex-biased dispersal promotes adaptive parental effects. *BMC Evolutionary Biology*, **10**, 217.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393–430.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Stiebens VA, Merino SE, Roder C, Chain FJJ, Lee PLM, Eizaguirre C (2013) Living on the edge: how philopatry maintains adaptive potential. *Proceedings of the Royal Society of London B: Biological Sciences*, **280**, 20130305.

Ulrich G, Jones CM, Driggers WB III, Drymon M, Oakley D, Riley C (2007) Habitat utilization, relative abundance and seasonality of sharks in the estuarine and near shore waters

of South Carolina. *American Fisheries Society Symposium*, **50**, 125–139.
Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.
Yang J, Benyamin B, McEvoy BP *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, **42**, 565–569.

D.S.P., J.B.P. and J.R.G. had responsibility for data collection and analysis and primary responsibility for writing of the manuscript. All other authors have reviewed and contributed to the current version of the manuscript. C.M.H. participated in data collection and analysis, and J.G. and D.C. obtained the samples.

Data accessibility

GenBank accession numbers for mtDNA sequences may be found in Table S2 (Supporting information). Demultiplexed, raw sequencing reads: Short Read Archive (Bioproject accession #PRJNA286089). The final SNP data set, in VCF format, the neutral and outlier haplotype data sets, in GENEPOP format, and a script to reproduce bioinformatic filtering: Dryad doi:10.5061/dryad.7k4c1.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary of diversity statistics for 5914 SNPs and sequences (1064 base pairs) of the mitochondrial control region for samples of bonnetheads from North Carolina (NC) and three localities along the Gulf Coast of Florida: Florida Bay (FB), Tampa Bay (TB), and Panama City (PC).

Table S2 Distribution of haplotypes and GenBank Accession numbers for mitochondrial control region sequences from samples of bonnetheads off North Carolina (NC) and three locations along the Gulf Coast of Florida: Florida Bay (FB), Tampa Bay (TB) and Panama City (PC).

Table S3 Results of standard least squares regression of allele frequencies at outlier loci by latitude: loci are organized as bi-allelic and multi-allelic.

Table S4 Results of analysis of molecular variance (AMOVA) for all three data sets.

Table S5 Results of BLAST search for sequence similarity of SNP containing loci.

Appendix S1 Methods.

Electronic Supplementary Material

Supplemental Methods:

ddRAD Library Methods

DNA was extracted using Mag-Bind Tissue DNA kits (Omega Bio-Tek) and digested with *Eco*RI and *Msp*I. A barcoded adapter was ligated to *Eco*RI restriction sites and a common adapter was ligated to *Msp*I restriction sites, using equimolar quantities of each digested sample. Samples were then pooled into four 'index' libraries consisting of ~34 individuals each and size selected using a Pippin Prep DNA size selection system (Sage Science Inc.). Fragments were selected using a mean size of 375 bp, with a 'tight' selection window (± 37 bp). Illumina flow-cell adapter sequences and index-specific identifiers were added to each index library, using 12 cycles of PCR.

mtDNA Sequencing

Thirty-microlitre PCR reactions contained 1X reaction buffer (pH 8.5), 2 mM MgCl2, 0.25 mM of each dNTP, 250 pmol of each primer, 0.05 U/μL *Taq* polymerase and 1.0 μL of template. Reaction conditions consisted of initial denaturation at 95°C for 2 min followed by 35 cycles at 94°C for 30 s, 54° for 60 s, and 72°C for 65 s, and a final extension at 72°C for 10 min. Amplified products were sent to Beckman Coulter to be cleaned and sequenced bi-directionally.

Bioinformatic Analysis

The combination of paired-end (PE) and single-end (SE) libraries called for customization to the default *dDocent* pipeline (Puritz *et al*. 2014). PE read files were placed in a working directory and a modified version of *dDocent* (version 1.0) was run with a cutoff value of 2 and a clustering % of 0.95. There were three modifications for reference contig assembly to help deal with a large repetitive genome.

1. During the clustering command of *Rainbow* (Chong *et al*. 2012), the mismatch parameter was changed to 2 and the –L command was implemented.
2. During the div command of *Rainbow* (Chong *et al*. 2012), the minimum frequency for a new variant was changed from 0.2 to 0.05.
3. During the merge command of *Rainbow* (Chong *et al*. 2012), the minimum number of reads to assemble was lowered from 5 to 3 and the maximum number of divided clusters to merge and the maximum number of reads to assemble was raised from 300 to 500.

After reference assembly was completed, SE reads were returned to the working directory. The read mapping portion of *dDocent* (version 1.0) was modified again to help deal with a large repetitive genome. The default clipping penalties were changed to (20,5), the mismatch parameter was lowered from 4 to 3, and the gap opening penalty was lowered from 6 to 5. Additionally, reads with more than 20% clipping were removed with AWK after mapping. These settings would enhance the ability for highly polymorphic reads to map to the reference, but also remove reads that have only have a small matching portion. Afterwards, default values of *dDocent* (version 1.0) were used to call to call variants. Raw variant calls were subjected to several filtering steps to reduce false positives. A script to reproduce the filtering steps is available at Dryad (doi:10.5061/dryad.7k4c1). Raw variants were filtered sequentially via VCFtools (Danecek *et al.* 2011) or custom bash scripts, using the following steps:

1. Loci were removed that had a minor allele count of less than 2, a PHRED quality score of less than 20, and a call rate of less than 50%.
2. All genotypes with less than 5 reads were changed to missing.
3. Loci were removed that now had less than a 75% call rate, and a minor allele frequency of less than 0.01.
4. 13 individuals (out of 134) were removed for having more than 40% missing data, using the script filter_missing_ind.sh (https://github.com/jpuritz/dDocent/blob/master/scripts/filter_missing_ind.sh).
5. The highly related individual (as described in main text; LK_007) was removed. Loci called in less than 97.5% of individuals and with a minor allele frequency of less than 5% were removed.

After this point, variant calls were filtered using a custom script (FB_filters_Bhead.sh; Dryad: doi:10.5061/dryad.7k4c1) that utilizes vcflib (https://github.com/ekg/vcflib) and VCFtools (Danecek *et a*l. 2011) to filter loci based on FreeBayes INFO criteria and depth:

1. Loci were removed if the average allele balance at heterozygous genotypes was less than 28% (i.e., if a genotype had 100X coverage, there would have to be an average of 28 or more reads from the alternate allele across all heterozygous genotypes). Additionally, if the quality sum of the reference or alternate allele was 0, the locus was removed. This removes sites that have a large portion of spurious heterozygous genotype calls.

2. Loci were then removed if the quality score was less than half of the total depth. With FreeBayes, excessive depth can give inflated quality scores.
3. Loci were removed if the ratio between the mean mapping quality of the alternate and reference allele was less than 0.9 or more than 1.05.
4. Loci were then removed if the majority of reads did not come from only one read orientation. Our insert size was much larger than our PE read lengths, so true RAD loci should not have forward and reverse reads that overlap.
5. Loci also were removed based on the status of properly paired reads. True variants should have reads coming from all properly paired reads, or only from reads that are not properly paired (some RAD loci do not assemble well for the PE read, leaving only forward reads). However, false variants tend to have properly paired reference reads and not properly paired alternate reads. Loci were retained if more than 0.05% of reference reads were properly paired and less than 0.05% of alternate reads were properly paired and vice versa.
6. Of the remaining loci, the average depth (and standard deviation) across all individuals was calculated. Loci that have a depth greater than the average depth plus on SD are removed if the quality score is less than 2 times the depth. This filter is based off results reported in (arXiv:1404.0929v1) by Li (2014).
7. Only loci that were in the bottom 90% of mean depth (less than 79.94) were kept to remove any possible paralogs or repetitive genomic regions.

After these two filter steps, loci were filtered based on locality-based tests of Hardy Weinberg equilibrium. Loci were removed that had a *P* value of less than 0.001 in at least 25% of the populations, using the script filter_hwe_by_pop.pl (https://github.com/jpuritz/dDocent/blob/master/scripts/filter_hwe_by_pop.pl). Variant calls were then decomposed into SNP and INDEL calls, using vcflib; INDELS were then removed with VCFtools to produce a VCF file of SNP only calls. SNPs were then filtered again for paralogs by removing loci that had more than 3 SNPs within 5 bp and loci that had had more than 4 SNPs within 17 bp. This final set of SNP calls was used for all subsequent analyses.

Supplemental References

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

2. Danecek P, Auton A, Abecasis G, *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156-2158.
3. Li H (2014) Towards better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics*, **30**, 2843-2851.
4. Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, **2**, e431.

Table S1. Summary of diversity statistics for 5,914 SNPs and sequences (1,064 base pairs) of the mitochondrial control region for samples of bonnetheads from North Carolina (NC) and three localities along the Gulf Coast of Florida: Florida Bay (FB), Tampa Bay (TB), and Panama City (PC). *n* - sample size; π - nucleotide diversity; *H* – number of haplotypes; *h* – nucleon diversity.

SNPs	**NC**	**FB**	**TB**	**PC**
n	22	32	31	36
π	0.2996	0.3189	0.3163	0.3168
mtDNA				
n	23	24	27	25
H	7	17	17	16
h	0.719 ± 0.077	0.938 ± 0.039	0.940 ± 0.031	0.947 ± 0.029
π	0.0010 ±0.0008	0.0025 ±0.0015	0.0022 ±0.0014	0.0026±0.0016

Table S2. Distribution of haplotypes and GENBANK Accession numbers for mitochondrial control region sequences from samples of bonnetheads off North Carolina (NC) and three locations along the Gulf Coast of Florida: Florida Bay (FB), Tampa Bay (TB) and Panama City (PC).

	NC	FB	TB	PC	Genbank Accession #
H1	6	1	2	2	KT031755
H2	11	1	6	3	KT031756
H3	1	-	-	-	KT031757
H4	1	-	-	-	KT031758
H5	1	-	-	-	KT031759
H6	2	6	2	5	KT031760
H7	1	-	-	-	KT031761
H8	-	-	-	1	KT031762
H9	-	-	-	2	KT031763
H10	-	-	-	1	KT031764
H11	-	-	-	1	KT031765
H12	-	-	-	2	KT031766
H13	-	-	-	1	KT031767
H14	-	-	1	1	KT031768
H15	-	-	-	1	KT031769
H16	-	-	-	1	KT031770
H17	-	2	3	1	KT031771
H18	-	1	1	1	KT031772
H19	-	-	-	1	KT031773
H20	-	-	-	1	KT031774
H21	-	-	1	-	KT031775
H22	-	-	2	-	KT031776
H23	-	-	1	-	KT031777
H24	-	1	-	-	KT031778
H25	-	-	1	-	KT031779
H26	-	1	-	-	KT031780
H27	-	1	-	-	KT031781
H28	-	1	-	-	KT031782
H29	-	1	-	-	KT031783
H30	-	-	1	-	KT031784
H31	-	-	1	-	KT031785
H32	-	1	-	-	KT031786
H33	-	1	-	-	KT031787
H34	-	1	-	-	KT031788

H35	-	-	1	-	KT031789
H36	-	-	1	-	KT031790
H37	-	1	-	-	KT031791
H38	-	1	-	-	KT031792
H39	-	2	-	-	KT031793
H40	-	-	1	-	KT031794
H41	-	-	1	-	KT031795
H42	-	1	-	-	KT031796
H43	-	-	1	-	KT031797
H44	-	1	-	-	KT031798
Total	23	25	27	25	

Table S3. Results of standard least squares regression of allele frequencies at outlier loci by latitude: loci are organized as bi-allelic and multi-allelic. Values of $r^2 > 0.90$ and $P < 0.05$ are italicized and bolded. %X and %Y are percentage contribution to abscissa and ordinate from DAPC analysis, using k-means clustering.

Haplotype	***r*²**	***P*-value**	% X	%Y
Bi-allelic				
E28275:1/2	***0.93***	***0.037***	7.54	0.08
E29236:1/2	***0.92***	***0.042***	0.42	0.62
E43805:1/2	***0.90***	***0.050***	6.22	0.08
E53310:1/2	0.85	0.075	2.56	1.10
E61036:1/2	0.49	0.299	0.02	0.16
E64126:1/2	***0.91***	***0.048***	0.56	0.02
E65274:1/2	0.76	0.130	0.54	1.80
E66074:1/2	***0.92***	***0.043***	8.94	0.04
E68107:1/2	0.76	0.127	2.64	3.56
E82240:1/2	***0.90***	0.053	7.50	0.70
E87901:1/2	0.83	0.089	0.46	4.36
E92875:1/2	***0.93***	***0.035***	2.68	0.00
E94551:1/2	***0.92***	***0.041***	8.28	0.02
E95266:1/2	***0.90***	***0.049***	0.90	0.14
E102269:1/2	0.84	0.084	0.02	10.58
E106435:1/2	***0.92***	***0.041***	8.28	0.02
E107376:1/2	0.61	0.222	0.18	1.28
E113131:1/2	0.86	0.073	2.10	1.24
E117105:1/2	0.86	0.071	3.52	0.38
E167424:1/2	0.87	0.069	2.04	1.76
E195370:1/2	0.83	0.089	3.54	1.12
Multi-allelic				
E52101:1/…	0.67	0.179	0.22	0.26
:2/…	0.29	0.457	0.07	13.98
:3/…	0.06	0.763	0.55	10.46
E69589:1/…	0.76	0.126	0.36	2.44

:2/…	0.87	0.070	0.32	3.18
:3/…	***0.92***	***0.042***	0.00	0.05
E73988:1/…	***0.92***	***0.038***	2.81	0.09
:2/…	***0.92***	***0.043***	3.48	0.07
:3/…	0.84	0.085	0.04	0.31
E75833:1/…	0.17	0.593	0.05	4.02
:2/…	**0.90**	0.054	1.05	1.14
:3/…	**0.94**	***0.029***	0.62	0.88
E94553:1/…	0.87	0.069	2.31	0.21
:2/…	0.79	0.112	0.58	0.07
:3/…	***0.92***	***0.042***	0.57	0.53
E110379:1/…	0.35	0.405	0.24	18.46
:2/…	0.89	0.057	0.86	7.43
:3/…	***0.93***	***0.034***	2.08	2.92
E71001:1/…	***0.92***	***0.039***	1.09	0.07
:2/…	0.73	0.146	0.05	1.18
:3/…	0.79	0.109	0.63	1.07
:4/…	0.67	0.182	0.00	0.04
E109425:1/…	***0.95***	***0.028***	2.35	1.15
:2/…	***0.92***	***0.038***	4.37	0.11
:3/…	0.45	0.331	0.34	0.48
:4/…	0.02	0.862	0.00	0.00
E131866:1/…	***0.90***	0.052	3.91	0.01
:2/…	0.89	0.055	0.04	0.06
:3/…	***0.93***	***0.037***	1.92	0.24
:4/…	0.72	0.150	0.15	0.03

Table S4. Results of analysis of molecular variance (AMOVA) for all three data sets. Data include % of variation (%), degrees of freedom (df), and sum of squares (SS).

AMOVA using Outlier Loci (O-SNPs):

Source of Variation	**Nested in**	**%**	**df**	**SS**	F_{ST}	***P*-value**
Among Individuals	Population	72.93	236	1245.48	--	--
Among Population	--	27.07	3	365.10	0.271	<0.00001

AMOVA using Neutral Loci (N-SNPs):

Source of Variation	**Nested in**	**%**	**df**	**SS**	F_{ST}	***P*-value**
Among Individuals	Population	99.21	236	163871.48	--	--
Among Population	--	0.79	3	3068.96	0.008	<0.00001

AMOVA using mtDNA haplotypes:

Source of Variation	**Nested in**	**%**	**df**	**SS**	Φ_{ST}	***P*-value**
Among Individuals	Population	92.23	95	74.52		--
Among Populations	--	7.77	3	7.26	0.078	<0.00001

Table S5. Results of BLAST search for sequence similarity of SNP containing loci: E-value, expected number of hits at random; PI. Percent identity.

Locus	Description	E-value	PI	Accession #	Hit start	Hit end
Di-Allelic						
E28275_L100	zebrafish DNA clone CH211-168G2 in linkage group 9	1.06E-03	83.7	BX005152	12226	12274
	zebrafish DNA clone DKEY-159B16 in linkage group 9	1.06E-03	83.7	CR407600	70212	70164
	Mustelus manazo DNA, HE1 SINE clone: Mm 3	3.72E-03	70.8	AB027718	181	86
E29236_L100	*Triakis scyllium* TNFRSF6B gene for tumor necrosis factor	9.96E-17	83.5	AB282596	1771	1681
	Carcharhinus leucas microsatellite C105	3.48E-16	83.0	KJ916108	55	147
	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	1.21E-15	82.4	AB174766	2935	3025
E43805_L100	*Ginglymostoma cirratum* clone GC_BA-557	1.69E-07	84.7	AC164927	65082	65024
	Ginglymostoma cirratum clone GC_BA-557B6	1.69E-07	84.7	AC164927	83026	82978
	Carcharhinus plumbeus RAG1 (partial) and RAG2 (complete)	1.69E-07	84.7	AY172838	8908	8966
E53310_L100	*Carcharhinus isodon* clone *Cis*172 microsatellite	1.06E-03	89.7	JQ365996	56	94
	Predicted: *Equus caballus* Zinc finger protein 226 (ZNF226)	1.3E-02	87.5	XM_005596327	2151	2112
	Predicted: *Equus caballus* Zinc finger protein 226 (ZNF226)	1.3E-02	87.5	XM_005596328	1916	1877
E61036_L100	None					
E64126_L100	None					
E65274_L142	*Mustelus manazo* DNA, HE1 SINE, clone:Mm 2	6.29E-13	85.9	AB027717	38	108
	Triakis scyllium IL-1 gene for interleukin-1β	7.66E-12	86.6	AB074142	3145	3211
	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	2.67E-11	84.5	AB174766	3231	3161
E66074_L100	None					
E68107_L100	Predicted*: Latimeria chalumnae* Zinc finger protein 850-like (LOC1 02360397)	2.67E-11	80.7	XM_005995770	1396	1314
	Predicted: *Colius striatus* Zinc finger protein 501-like (LOC10455 9532)	1.14E-09	80.0	XM_010204559	482	403
	Predicted: *Xenopus* (*silurana) tropicalis* Zinc finger protein 84-like (LOC101730679)	3.97E-09	76.6	XM_004919267	571	478
E87901_L109	*Carcharhinus sorrah* microsatellite *Cs*08	5.88E-26	87.4	AY545211	752	643
	Carcharhinus plumbeus T cell receptor gamma (TCRG)	2.05E-25	86.4	FJ854492	17993	18102
E92875_L100	*Stegastes partitus* Zinc finger and scan domain-protein 2-like (LOC103371773)	3.72E-03	86.0	XM_008301224	1162	1120
	Stegastes partitus Zinc finger and scan domain-protein 2-like (LOC103371773)	3.72E-03	86.0	XM_008301232	782	740
E94551_L100	*Mustelus manazo* DNA, HE1 SINE clone:Mm2	7.66E-12	92.6	AB027717	65	12

	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	7.66E-12	92.6	AB174766	3204	3257
	Mustelus manazo DNA, HE1 SINE clone:Mm 3	3.26E-10	90.7	AB027718	65	12
E95266_L148	*Callorhinchus milii* clone P02H01.kidney.K065	1.3E-02	91.4	JX052830	4	38
	Callorhinchus milii clone, NADH dehydrogenase (Ubiquinone) 1 beta subcomplex	1.3E-02	91.4	JX209076	25	59
	Callorhinchus milii NADH Dehydrogenase (Ubiquinone) 1 beta subcomplex	1.3E-02	91.4	NM_001292677	1	35
E102269_L100	None					
E106435_L100	*Mustelus manazo* DNA, HE1 SINE, clone:Mm 2	3.71E-22	86.9	AB027717	239	144
	Mustelus manazo DNA, HE1 SINE, clone:Mm 4	3.71E-22	88.9	AB027719	231	143
	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	3.71E-22	87.9	AB174766	3030	3125
E107376_L100	*Carcharhinus plumbeus* T cell receptor gamma (TCRG)	1.06E-03	72.3	FJ854492	25169	25069
E113131_L101	*Ginglymostoma cirratum* Zinc finger protein 112, β2m , βrd2 genes	3.72E-03	84.8	AB571627	68553	68598
E117105_L100	None					
E195370_L100	*Brugia pahangi* genome assembly B	3.72E-03	93.9	LK964241	31402	31434
Multi-allelic						
E52101_L100	*Scyliorhinus canicula* Cluster_HOXD sequence	5.00E-15	83.0	FQ032660	58525	58612
	Mustelus manazo DNA, HE1 SINE, clone:Mm 2	2.00E-12	81.0	AB027717	52	136
	Carcharhinus plumbeus RAG1 (partial) and RAG2 (complete)	1.0E-10	80.0	AY172838	3341	3425
E69589_L100	*Carcharhinus plumbeus* T cell receptor gamma (TCRG)	4.0E-09	83.0	FJ854492	18165	18105
	Carcharhinus sorrah microsatellite CS08 sequence	6.00E-07	90.0	AY545211	601	648
	Carcharhinus plumbeus Ig lambda light chain gene, COMPLETE CDS	6.00E-07	90.0	U34992	6243	6290
E75833_L160	*Sphyrna lewin* GRLN Gene for preproghrelin, COMPLETE CDS	3.26E-10	79.5	AB254130	3838	3925
	Mustelus manazo DNA, HE1 SINE, clone:Mm 2	3.97E-09	78.7	AB027717	57	145
	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	4.83E-08	77.5	AB174766	3212	3124
E82240_L100	*Homo sapiens* BAC clone RP11-334C6 from chromosome 7	1.3E-02	84.4	AC073418	55859	55815
	Pan troglodytes BAC clone CH251-734F1 from chromosome 7	1.3E-02	84.4	AC190230	179380	179336
	Pan troglodytes BAC clone CH251-541E24 from chromosome 7	1.3E-02	84.4	AC192728	31156	31112
E94553_L100	*Botryotinia fuckeliana* T4 supercontig_34_1 genomic supercontig	1.3E-02	84.0	FQ790278	92859	92814
	Botryotinia fuckeliana B05.10 hypothetical protein (BC1G_13291)	1.3E-02	84.0	XM_001548305	1089	1044
	Torpedo marmorata mRNA fragment for acetylcholinesterase c-term	4.53E-02	96.6	X13173	1745	1773
E71001_L100	None					
E73988_L143	*Scyliorhinus canicula* cluster_HOXBS	2.00E-14	79.0	FQ032659	65153	65252
	Triakis scyllium IL-1 gene for interleukin-1β	2.00E-13	79.0	AB074142	899	1002

	Scyliorhinus canicula cluster_HOXD sequence	1.00E-10	78.0	FQ032660	15011	15107
E109425_L100	None					
E110379_L124	None					
E131866_L100	*Triakis scyllium* IL-1 gene for interleukin-1β	2.34E-18	81.7	AB074142	1087	984
	Triakis scyllium MIP3 gene for macrophage inflammatory protein-3α	9.96E-17	81.6	AB174766	3443	3540
	Ginglymostoma cirratum clone GC_Ba-678C3	4.24E-15	80.6	AC165195	14263	14360
E167414_L100	*Callorhinchus milii* gamma-aminobutyric acid type B receptor	3.05E-04	83.0	XM_007884901	1038	1090