Genomics versus mtDNA for resolving stock structure in the silky shark (*Carcharhinus falciformis*)

Derek W. Kraft [Corresp., 1], **Emily Conklin** [1], **Evan Barba** [1], **Melanie Hutchinson** [1, 2], **Robert J. Toonen** [1], **Zac H. Forsman** [1], **Brian W. Bowen** [1]

[1] Hawai'i Institute of Marine Biology, University of Hawai'i, Kaneohe, Hawaii, USA

[2] Joint Institute of Marine and Atmospheric Research, Pacific Islands Fisheries Science Center, NOAA, University of Hawai'i, Honolulu, Hawaii, USA

Corresponding Author: Derek W. Kraft
Email address: Kraftd@hawaii.edu

Conservation genetic approaches for elasmobranchs have focused on regions of the mitochondrial genome or a handful of nuclear microsatellites. High-throughput sequencing offers a powerful alternative for examining population structure using many loci distributed across the nuclear and mitochondrial genomes. These single nucleotide polymorphisms are expected to provide finer scale and more accurate population level data; however, there have been few genomic studies applied to elasmobranch species. The desire to apply next-generation sequencing approaches is often tempered by the costs, which can be offset by pooling specimens prior to sequencing (pool-seq). In this study, we assess the utility of pool-seq by applying this method to the same individual silky sharks, *Carcharhinus falciformis*, previously surveyed with the mtDNA control region in the Atlantic and Indian Oceans. Pool-seq methods were able to recover the entire mitochondrial genome as well as thousands of nuclear markers. This volume of sequence data enabled the detection of population structure between regions of the Atlantic Ocean populations, undetected in the previous study (inter-Atlantic mitochondrial SNPs F_{ST} values comparison ranging from 0.029 to 0.135 and nuclear SNPs from 0.015 to 0.025). Our results reinforce the conclusion that sampling the mitochondrial control region alone may fail to detect fine-scale population structure, and additional sampling across the genome may increase resolution for some species. Additionally, this study shows that the costs of analyzing 4,988 loci using pool-seq methods are equivalent to the standard Sanger-sequenced markers and become less expensive when large numbers of individuals (>300) are analyzed.

Genomics versus mtDNA for Resolving Stock Structure in the Silky Shark (*Carcharhinus falciformis*)

Derek W. Kraft[1*], Emily Conklin[1], Evan Barba[1], Melanie Hutchinson[1,2], Robert J. Toonen[1], Zac H. Forsman[1], Brian W. Bowen[1]

[1]Hawai‘i Institute of Marine Biology, University of Hawai‘i, 46-007 Lilipuna Road, Kāne‘ohe, HI 96744

[2]Joint Institute for Marine and Atmospheric Research, University of Hawai‘i, Pacific Islands Fisheries Science Center, NOAA

Corresponding Author:

Derek W. Kraft[1]

46-007 Lilipuna Rd, Kaneohe, HI 96744, USA

Email address: Kraftd@hawaii.edu

Abstract

Conservation genetic approaches for elasmobranchs have focused on regions of the mitochondrial genome or a handful of nuclear microsatellites. High-throughput sequencing offers a powerful alternative for examining population structure using many loci distributed across the nuclear and mitochondrial genomes. These single nucleotide polymorphisms are expected to provide finer scale and more accurate population level data; however, there have been few genomic studies applied to elasmobranch species. The desire to apply next-generation sequencing approaches is often tempered by the costs, which can be offset by pooling specimens prior to sequencing (pool-seq). In this study, we assess the utility of pool-seq by applying this method to the same individual silky sharks, *Carcharhinus falciformis*, previously surveyed with the mtDNA control region in the Atlantic and Indian Oceans. Pool-seq methods were able to recover the entire mitochondrial genome as well as thousands of nuclear markers. This volume of sequence data enabled the detection of population structure between regions of the Atlantic Ocean populations, undetected in the previous study (inter-Atlantic mitochondrial SNPs F_{ST} values comparison ranging from 0.029 to 0.135 and nuclear SNPs from 0.015 to 0.025). Our results reinforce the conclusion that sampling the mitochondrial control region alone may fail to detect fine-scale population structure, and additional sampling across the genome may increase resolution for some species. Additionally, this study shows that the costs of analyzing 4,988 loci using pool-seq methods are equivalent to the standard Sanger-sequenced markers and become less expensive when large numbers of individuals (>300) are analyzed.

Introduction

Many elasmobranchs around the globe have experienced devastating population declines due to overfishing in both target and non-target fisheries (Musick et al., 2000; Clarke et al., 2006; F. Ferretti et al., 2010; Heupel et al., 2014; Dulvy et al., 2014; Oliver et al., 2015; Dulvy & Trebilco, 2018). These species are especially vulnerable to overfishing due to life history traits such as late maturity, slow growth, low fecundity, and high juvenile mortality, which collectively result in low intrinsic rate of population increase (Baum et al., 2003; Dulvy et al., 2008). Elasmobranch populations take decades to recover from overfishing, and only if fishing pressure is relieved for an extended period (Stevens et al., 2000). Furthermore, many threatened and endangered elasmobranchs have little to no population genetic data that would assist in the resolution of management units (reviewed in Domingues et al., 2018a).

Genetically distinct populations are isolated management units known as stocks; however, stocks can be defined on a smaller scale than genetic populations through other criteria, such as an exclusive economic zone boundry (Carvalho & Hauser, 1994; Ovenden et al., 2015). Reduced gene flow indicates that if a population is overfished it will not be replenished by immigrants from surrounding populations. This is why managing on a genetic stock-by-stock basis is essential for successful maintenance of exploited species and is sorely needed for over-harvested elasmobranchs (Dizon et al., 1993; Heist, 2004; Tallmon et al., 2010).

For the past two decades the standard for examining population structure in elasmobranchs has been a section of the mitochondrial genome, usually the control region (mtCR) (e.g. Duncan et al. 2006; Hoelzel et al. 2006; Keeney & Heist 2006; Castro et al. 2007; Whitney et al. 2012; Clarke et al. 2015; reviewed in Domingues et al. 2018a). Though recent studies are moving

71 towards multi-marker approaches (Momigliano et al., 2017; Pazmiño et al., 2018: Green et al.,
72 2019), there is still a large body of literature focusing on mtCR. The mitochondrial genome has a
73 higher rate of mutation than most of the nuclear genome (Brown et al., 1979; Charlesworth &
74 Wright, 2001; Neiman & Taylor, 2009) and this rate of mutation is a key advantage in
75 vertebrates with slowly-evolving genomes (Avise et al., 1992; Martin et al., 1992).
76 Elasmobranch mtDNA studies to date have been successful in elucidating population partitions
77 and evolutionary divergences, but the maternal inheritance of mtDNA can limit conclusions
78 about gene flow in cases of sex-biased (usually male) dispersal. Both mtDNA and nuclear
79 markers often have concordant results in sedentary species (e.g., Lavery et al., 1996; Avise,
80 2004; Zink & Barrowclough, 2008; DiBattista et al., 2015) but, when examined alone, may miss
81 key components of population structure, particularly in migratory fauna (Pardini et al., 2001;
82 Bowen et al., 2005; Toews & Brelsford, 2012). When highly mobile elasmobranchs are
83 examined with both mtDNA and nuclear markers (usually microsatellites), a different picture
84 often emerges in which females are more resident and males are dispersive (Pardini et al., 2001;
85 Schultz et al., 2008; Portnoy, et al., 2010; Karl, Castro, Lopez, Charvet, & Burgess, 2011; Daly-
86 Engel et al., 2012; Portnoy et al., 2015: Bernard et al., 2017; Domingues et al., 2018b).
87 Identifying outlier SNPs in the nuclear genome can highlight genes possibly under selection, or
88 show functional responses to environmental changes that have important management
89 consequences (Barrio et al., 2016; Fischer et al., 2013; Guo et al., 2016; Jones et al., 2012).
90 Therefore, the combination of mitochondrial and nuclear markers can yield fundamental
91 ecological and evolutionary insights.
92

High-throughput sequencing is a powerful tool for revealing fine-scale population structure that may be missed by single locus studies (Andrews et al., 2016; Hohenlohe et al., 2018). However, this method can be costly, especially when examining many individuals as is typical of population genetic or phylogeography studies, and the perceived cost may prevent some from considering a high-throughput sequencing approach. For population genetics approaches based on differences in allele frequencies among populations, equimolar pooling of samples before sequencing is an affordable and accurate strategy for large-scale genetic analysis (Schlötterer et al., 2014). Several studies have successfully resolved population structure using a pooled site-associated DNA approach known as pool-seq, including some in commercially valuable marine species (e.g. Gautier et al., 2013; Mimee et al., 2015). Pool-seq provides estimates of allele frequencies for thousands of loci distributed across the genome simultaneously, which in some cases gives greater statistical power that can actually exceed the accuracy of allele frequency estimates based on individual sequencing (Futschik & Schlötterer, 2010, but also see Anderson et al. 2014). Therefore, a comparison of results between the standard mtCR analysis and high-throughput pool-seq is informative in evaluating the relative power and cost of the two approaches for examining population structure.

The silky shark (*Carcharhinus falciformis* (Müller & Henle, 1839)) is the second most commonly harvested shark on Earth (Oliver et al., 2015; Rice & Harley, 2013). They are one of the top contributors to the shark fin trade and the most common elasmobranch bycatch species in tuna purse-seine fisheries around the world (Cardenosa et al., 2018; Clarke et al., 2006; Oliver et al., 2015). This pelagic shark, formerly abundant in all tropical oceans, has declined by an estimated 85% in the last 20 years, and is now listed as vulnerable and declining by the

International Union for the Conservation of Nature (Rice & Harley, 2013; IUCN, 2017). Currently silky shark population assessments are conducted at the scale of regional fishery management organization, and conservation management measures are implemented at this scale in the absence of genetic or movement data to define population boundaries. Clarke et al. (2015) surveyed silky sharks across these regional management regions and found the western Atlantic was strongly differentiated from the Indian Ocean, but the North Atlantic, Gulf of Mexico, and Brazil could not be differentiated and appeared to comprise a single population. In contrast, using the same mtCR marker, Domingues et al. (2017) examined five regions across the Western Atlantic and found the North Western Atlantic was distinct from the South Western Atlantic. The difference between the two studies results from additional sampling in the South West Atlantic from further south than Clarke et al. (2015).

In an era where wildlife management needs far exceed the financial resources to address them, many seek to find the most accessible, robust, and economical means to define management units. In this study, we provide a direct comparison of population genetic analysis methods between Sanger sequencing of the mtCR region and high-throughput sequencing of regional pools of individuals. The same individuals from Clarke et al. (2015) were re-sequenced using pool-seq approaches. Regions re-sequenced included Gulf of Mexico, North West Atlantic, and Brazil, as well as one geographically distant location in the Red Sea (*Fig 1*). We focused this analysis on SNPs from the mitochondrial DNA as well as nuclear DNA. We did not analyze any microsatellite loci because they were not a part of Clarke et al. (2015). We then evaluate the economics of conducting pool-seq relative to conventional Sanger sequencing of these same

138 individuals. Ecological and management implications will be addressed in a subsequent
139 companion paper.
140

141 ## Materials & Methods

142 ***Sampling and sequencing***

143 A total of 143 silky shark fin clips or muscle sections were sampled from commercial or artisanal
144 fisheries across four geographic regions and are the same samples examined in Clarke et al.
145 (2015). Specifically, we sampled the Gulf of Mexico (GM, n =39), the North Atlantic (NA, n =
146 33), Brazil (BR, n = 34), and the Red Sea (RS, n = 37). These sample sizes are slightly lower
147 than Clarke et al. (2015). This reduction was due to DNA degradation over time and the need for
148 high-quality genomic DNA for pool-seq. This is contrary to the DNA quality needed for
149 amplifying a single marker from the mitochondrial control region. Additionally only a subset of
150 the Red Sea samples were randomly selected to keep sample sizes relatively similar.
151

152 DNA was extracted using Qiagen DNeasy Blood & Tissue kit (Qiagen, Mississauga, ON,
153 Canada), following manufacturer protocols. Extracted DNA quality was assessed visually by gel
154 electrophoresis and imaged using Gel Doc E-Z System (BIO RAD, Hercules, California, USA).
155 Only DNA aliquots with strong genomic DNA bands were further processed, while degraded or
156 overly digested DNA was discarded. Aliquots of high-quality DNA were quantified using an
157 AccuClear Ultra high sensitivity dsDNA quantitation kit (Biotium, Fremont CA, USA) and a
158 SpectroMax M2 (Molecular Devices, Sunnyvale, CA, USA). Libraries were pooled with an
159 equal amount of DNA (ng/µl) contributed per individual to minimize individual contribution
160 bias, totaling 2000 ng of DNA per library. Number of individuals per pool are displayed in *Fig 1*.

No PCR was performed to ensure individual DNA contribution was kept equal within and across libraries (Anderson et al., 2014). The rest of the library preparation followed the ezRAD library preparation protocol (Toonen et al. 2013; Knapp et al., 2016). This included DNA digested with DpnII restriction enzyme and adapters ligated using a Kapa hyper Prep Kit (Kapa Biosystems, Wilmington, MA, USA). Pooled libraries were sequenced using Illumina MiSeq (v3 2x300bp PE) at the Hawai‘i Institute of Marine Biology EPSCoR Core sequencing facility.

Genetic analyses

MultiQC was used to assess sequence quality scores, sequence length distributions, duplication levels, and overrepresented sequences (Ewels et al., 2016). To analyze the mitochondrial genome, a previously published mitochondrial genome from *Carcharhinus falciformis* was used as a reference (GeneBank accession number KF801102). Raw paired-end reads were trimmed with TRIMMOMATIC, mapped to the mitochondrial genome reference BWA (mem algorithm), and variants called using the dDocent bioinformatics pipeline, modified for pool-seq (Puritz et al. 2014, see below for details). Called SNPs were then analyzed with AssessPool (github.com/ToBoDev/assessPool, see below for details).

The bioinformatics pipeline included dDocent followed by AssessPool. Given that no reference genome was available, a reference was constructed using the dDocent *de novo* assembly and optimized utilizing the reference optimization steps provided on the dDocent assembly tutorial (http://ddocent.com/assembly/). Before assembly reads were trimmed using default settings and then an overlap (OL) assembly was performed, followed by clustering with CD-HIT with a –c parameter of 90% similarity. For mapping using BWA (mem algorithm) all match, mismatch, and gap open penalty score parameters were also default settings. Different parameters were

tested during optimization but did not improve mapping. Within-pool (K1) and between-pool (K2) minimum locus depth values selected for the *de novo* assembly did impact the results. dDocent provides graphical outputs to help select these values; however, testing a few different values of each is recommended to fully explore the potential of the data by balancing number of contigs by coverage depth (see ddocent.com/UserGuide for details). Selected values for K1 and K2 were 3 and 3 respectively. Once assembled, sequences were mapped, SNPs were called within the dDocent pipeline using FreeBayes, modified for SNP calling in pools (Garrison and Marth 2012, https://github.com/ekg/freebayes). Any contigs that aligned to the mitochondrial genome were removed from this nuclear dataset. The contigs that aligned specifically to the mitochondrial control region were saved for SNP validation to directly compare the results from this pool-seq approach to those previously reported by Clarke et al. (2015).

SNP calling with FreeBayes was optimized for pooled samples using the 'pooled-continuous' option, and minor allele frequency was decreased to 0.05 to capture alleles with frequency greater than 5% in the population (See Supplementary Material for code). The dDocent pipeline outputs SNPs in two variant call format files (.vcf), one being all raw SNPs (TotalRawSNPs.vcf) and another with filtered SNPs (Final.recode.vcf) however dDocent does not optimize filtering for pool-seq data. Therefore, the raw SNPs were processed with the pool-seq specific program AssessPool which uses VCFtools and vcflib to filter SNPs (Danecek et al., 2011). SNPs were processed with the following filters: minimum pool number of 2, minimum quality score of 20, minimum depth threshold of 30, maximum amount of missing data of 3, maximum allele length of 10, quality score to depth ratio of 0.25 as well as mean depth per site vs. quality score, and finally a maximum mean depth threshold of 1000 (*Table S1*). AssessPool then sends filtered

SNPs to either PoPoolation2 (Kofler et al., 2011) or poolfstat (Hivert et al., 2018). PoPoolation2 calculates mean pairwise F_{ST} values and significance in the form of p-values obtained using Fisher's exact test and combined using Fisher's method (as described in Ryman et al. 2006). Poolfstat (Hivert et al. 2018) takes a different approach, calculating F_{ST} values based on an analysis-of-variance framework (*sensu* Wier & Cockerham 1984) to eliminate biases associated with varying pool sizes. AssessPool then organizes, summarizes, and creates visualizations of the data using RStudio (RStudio Team 2020).

As a quality control test, sequences from Clarke et al. (2015) were downloaded from GenBank (accession numbers KM267565–KM267626), and SNPs from these data were compared directly to SNPs called within the control region of the mitochondrial pool-seq data generated here. Concordance of this validation set of SNPs was determined by Mantel test in R (Legendre & Legendre, 1998) comparing the matrices of pairwise F_{ST} values among populations.

Cost Analysis

The cost of pool-seq approach compared to Sanger sequencing of individual loci was calculated based on library preparation and sequencing cost at our facility. We did not include labor but calculated the total cost to generate sequence data from each sample included here from such expenses as the extraction, laboratory consumables, PCR amplification, library preparation, reaction clean-ups, quantification, quality control testing, and sequencing costs. These costs were translated into functions in RStudio (RStudio Team, 2020) where Sanger sequencing is a fixed rate per individual and pool-seq costs are fixed per flow cell on our MiSeq, but individual cost varies based on number of individuals and number of pooled regions per sequencing run. These functions were then plotted together for comparison.

230

231 ## Results

232 A total of 30.8 million reads were generated for the four geographic regions, which averaged 7.7
233 ± 3.0 million reads per pooled library. Results from the MutliQC assessment showed fairly
234 homogenous output between libraries in regard to sequence quality scores, GC and per base
235 sequence content, sequence length distributions, duplication levels, overrepresented sequences,
236 and adapter content. Once assembled, aligned, and mapped, 5,792 SNPs were resolved across the
237 mitochondrial and nuclear genomes combined. There were 4,103 biallelic SNPs, 168 were
238 multialleleic SNPs and 48 were insertions and deletions (INDELs). INDELs and multiallelic
239 SNPs remain a challenge for quantification software, so we restricted our analysis to biallelic
240 loci (Fracassetti et al. 2015). AssessPool creates visualizations of F_{ST} values and allows for visual
241 outlier inspection. No visual outliers were present and given these SNPs are distributed
242 haphazardly across the genome, they are assumed to be putatively neutral.

243

244 ***Mitochondrial Genome***

245 Analysis of the complete mitochondrial genome (17,774 bp) revealed 804 variable sites: 681
246 biallelic and 17 multiallelic SNPs. Because coverage in this dataset was fairly low on average,
247 most of these SNPs did not meet the filter threshold. After further filtering for the highest quality
248 markers, 30 SNPs were selected to calculate allele frequencies. Pairwise F_{ST} values were all
249 significant (*Fig 2*, *Table S2*). The Red Sea had much higher F_{ST} values (ranging from 0.367 to
250 0.745) than any inter-Atlantic comparison (ranging from 0.029 to 0.135). However, all
251 comparisons within the Atlantic still showed significant F_{ST} values, the highest being between

the North Atlantic and Brazil, and the lowest between Brazil and Gulf of Mexico (*Fig 2*, *Table S2*).

Nuclear loci

Our nuclear data showed 4,988 variants of which 3,422 were biallelic SNPs and 151 were multialleleic SNPs. A total of 346 SNPs remained after the same filtering process for the highest quality SNPs was applied as for the mitochondrial genome. Nuclear markers showed lower F_{ST} values between locations than the mitochondrial data, yet all P-values were still significant (*Fig 2*, *Table S2*). The Red Sea showed consistently higher F_{ST} values in comparison to inter-Atlantic comparisons except for the North Atlantic to Gulf of Mexico comparison, which showed the second highest mean F_{ST} value (*Fig 2*, *Table S2*). The highest value ($F_{ST} = 0.035$) was observed between Gulf of Mexico and the Red Sea, whereas the lowest ($F_{ST} = 0.014$) was between the North Atlantic and Brazil, which had the highest F_{ST} value within the Atlantic for the mitochondrial data.

SNP validation

SNPs called in the mitochondrial control region using the pool-seq protocol were compared with those reported in Clarke et al. (2015). Of the 34 SNPs in their study 14 of them had a minor allele count (MAC) of less than or equal to 3 and several were singletons. These SNPs are removed from the pool-seq data due to MAC SNP filter of >3 to remove sequencing errors that might be scored as rare alleles during high-throughput sequencing. Therefore, singletons or any rare allele represented fewer than 3 times in a population will inherently be removed from pool-seq data sets. Fortunately those rare alleles do not tend to overly impact Fst values and should

not bias interpretations of population structure (Bird et al. 2011; Toonen et al. 2011). Three SNPs were found in the Clarke study with a MAC of >3 that were not present in the pool-seq data; however, the remining 17 SNPs were all present in our data, plus one that was not found in the Clarke study (*Fig S1*). Despite the loss of these rare alleles from the SNP validation set, pairwise F_{ST} values estimated by both methods remained highly correlated (Mantel test, $r^2 = 0.96$, $p < 0.05$), and comparisons between the Red Sea and all three Atlantic populations showed the same relative magnitude between both methods.

Cost Analysis

The findings for cost analysis indicate that pool-seq reaches a threshold at approximately 300 individuals, after which this approach offers cheaper results than individual Sanger sequences. Furthermore, the cost is only twice as expensive at just over 100 individuals (*Fig 3a*). The pool-seq approach provides a far higher ratio of information for the cost, yielding greater population resolution. This cost assessment does not include analytical time, labor, or effort associated with pool-seq analyses such as access to computer resources and expertise with bioinformatic pipelines. However these costs are likely to decrease in the near future as bioinformatic pipelines are improved and become more widely available, for example as applications deployed via cloud based platforms such as Galaxy (https://usegalaxy.org/) or CyVerse (https://cyverse.org/). It is also important to note that the choice of pool-seq methodology has many caveats, which are discussed in greater detail in the 'considerations on pool-seq' section of the discussion below.

Discussion

Elasmobranchs are being harvested at unsustainable levels in several commercial fishing industries around the world. A fundamental step in successful management of any species is

297 resolving population boundaries so they can be managed on a genetic stock by stock basis. As
298 genetic sequencing technologies advance, there is greater opportunity to detect even small-scale
299 genetic differences between populations. When these differences amount to statically significant
300 allele frequencies at the population level, this indicates limited exchange among distinct stocks.
301
302 Here, we validate the utility of pool-seq using the same individuals as a previous study (Clarke et
303 al. 2015) and show that pool-seq recovers additional population structure relative to Sanger
304 sequencing of the mtDNA control region. Pool-seq was able to detect isolated populations
305 between the Gulf of Mexico, Western Atlantic, and along the Brazilian coast, where Clarke et al.
306 (2015) found no population structure. As expected, the Red Sea population was highly isolated
307 from Atlantic conspecifics using both approaches.
308
309 One advantage of this pool-seq approach is that we recover SNPs throughout the entire
310 mitochondrial genome along with thousands of additional nuclear loci that together provide
311 greater statistical power to detect finer scale population structure (Ryman & Palm 2006; Larsson
312 et al. 2009; Kurland et al. 2019). The pool-seq approach yielded significant genetic structure
313 among inter-Atlantic regions in both mtDNA and nuclear loci, whereas Sanger sequencing of the
314 mtCR lacked power to resolve significant differences among the same populations. The
315 congruence between the mitochondrial genome and nuclear loci reinforces the conclusion of
316 population structure among all regions sampled in this study.
317
318 In this case, pool-seq lived up to the promise of increased power to detect fine-scale structure,
319 but does it live up to the promise (Ferretti et al., 2013; Schlötterer et al., 2014) of being cost-

effective? Individual extraction costs remain fixed across both approaches and Sanger sequencing generally has a flat rate per individual, including PCR primers and reagents, and sequencing per individual per locus. In contrast, pool-seq has a flat sequencing cost determined by the number of reads generated from the high-throughput sequencing platform, plus a small additional cost per pool for the exact quantification of DNA for equimolar pooling and the library preparation for high-throughput sequencing. Comparing costs at our institution between a single Sanger sequencing marker and pool-seq on the Illumina MiSeq platform indicates pool-seq becomes less expensive when sample size of the study rises above 300 individuals. Although the cost per pool is essentially fixed, when higher numbers of individuals are included per pool, the price per individual analyzed is further reduced (*Fig 3b*). Our comparison here is limited to 12 pools due to the maximum number of reads per lane produced on the MiSeq platform. Therefore, analyzing more than 12 pools would require additional sequencing runs and result in a step increase in the cost per individual/pool, although this would differ among other Illumina machines (such as the HiSeq, NextSeq or NovaSeq) or other high-throughput sequencing platforms (such as the PacBio Sequel II). Larger numbers of pools could be run on some of these machines, but with differing individual read lengths and sequencing depths, which also bring other trade-offs. Likewise, samples can also be run with individual barcodes, therefore gaining the individual information lost by pooling specimens, but with increased initial setup and sequencing costs. There are so many options by which to apply these methods that we cannot possibly consider them all here, and the availability, cost, and trade-offs associated with each should be ideally considered by individuals when designing high-throughput sequencing projects. In our case, we considered only the options currently available to us through our campus sequencing core, and all these pool-seq price comparisons are to a single Sanger-

sequenced marker. Thus, when considering the information acquired from pool-seq compared to the cost from traditional single mitochondrial marker the price per individual advantage is massively amplified.

Considerations with pool-seq

As with any sequencing technique, there are still several factors to consider before deciding if pool-seq is appropriate for a particular study. Multiple reviews have been published on high-throughput and pool-seq approaches demonstrating pros, cons, and considerations with these methods, which are beyond the scope of this study. Interested readers should consult Perez-Enciso & Ferretti (2010), Futschik & Schlötterer (2010), Kofler et al. (2012), Ferretti et al. (2013), Schlötterer et al., 2014, Andrews & Luikart (2014), Andrews et al. (2016), and Kurland et al. (2019).

Pooling assumes individuals are from the interbreeding individuals within a single population of the same species. Therefore, care needs to be taken to avoid cryptic species, combining multiple populations (Wahlund effect), or other unintentional bias when selecting individuals to pool (Garnier-Géré & Chikhi 2013). For wide ranging pelagic species such as the blue shark or oceanic whitetip it seems reasonable to pool individuals from a larger area than it would be for small benthic species such as horn sharks, wobbegongs, or most rays. Population structure may be obscured if the geographic range per pool is too large or if there is complex population structure (*sensu* Bowen et al. 2005), because individuals from multiple sub-populations will be mixed into a single pool from which allele frequencies are calculated. Certainly pool-seq is not appropriate in all cases. It is a cost-saving approach for analyses based on allele frequencies

366 only, because individual information is lost by pooling, including haplotypes/genotypes and
367 linkage disequilibrium information. Also, pooling makes it difficult to distinguish between low
368 frequency alleles in the population and sequencing error. Therefore, careful filtering must be
369 applied to ensure only valid SNPs are analyzed instead of analyzing sequencing noise (Anand et
370 al., 2016; Schlötterer et al., 2014). Finally, the estimation of F_{ST} from pooled data remains a
371 subject of some debate, and new approaches and bias corrections are being actively developed
372 (Kofler et al. 2011; Hivert et al. 2018). To account for this uncertainty, we include analyses
373 based on both the original PoPoolation2 (Kofler et al., 2011) package and the newer poolfstat
374 (Hivert et al., 2018) that explicitly considers potential biases associated with varying pool sizes.
375 The two approaches yield slightly different F_{ST} values (see *Table S2*), however a comparison of
376 the two F_{ST} matrices shows strong correlation (Mantel r=0.991 for mitochondrial and r=0.978 for
377 nuclear data, $p < 0.05$). Therefore, only those F_{ST} values calculated by PoPoolation2 are reported
378 in the main text for ease of presentation.
379
380 Though pool-seq has been shown to be an affordable and reliable tool for population genomics
381 (Futschik & Schlötterer, 2010; Gautier et al., 2013; Rellstab et al. 2013; Konczal et al. 2014;
382 Schlötterer et al. 2014; Kurland et al. 2019), projects with larger budgets could allocate funds for
383 any of a variety of other genomic sequencing techniques such as individual RADseq libraries
384 (Hohenlohe et al. 2010), GBS (Narum et al. 2013), SNP arrays (Qi et al. 2017), bait capture
385 (Feutry et al. 2020), or low coverage genomewide sequencing (Therkildsen & Palumbi 2017).
386 These approaches allow for individual genotyping to examine questions that require individual-
387 level information and could provide a deeper assessment of populations. However it is also
388 important to consider not all labs can afford to generate genomic level data, especially in

389 developing countries, and having a cost-effective alternative to single marker studies will
390 continue to be invaluable to many.
391

392 ## Conclusions

393 The finding of population structure on the scale of North Atlantic/Gulf of Mexico/Brazil is
394 nearly unprecedented for a pelagic shark. Population structure in globally distributed sharks is
395 typically detected on a scale of ocean basins (Atlantic versus Indo-Pacific, Castro et al. 2007;
396 Graves & McDowell, 2015) and a few pelagic fishes have no population structure on a global
397 scale (e.g. Basking shark, *Cetorhinus maximus*, Hoelzel et al. 2006; Blue shark *Prionace glauca,*
398 Veríssimo et al. 2017; Wahoo, *Acanthocybium solandri* , Theisen et al. 2008). The resolution of
399 isolated populations on the scale of North Atlantic Ocean is more typical of coastal species than
400 pelagic species. The silky shark seems to be a pelagic species with a somewhat coastal
401 population structure. This has strong implications for international management because smaller
402 stocks imply smaller populations which are more readily depleted. At a minimum, these data
403 require rethinking a single population management approach for the Atlantic, and this pattern
404 needs to be investigated for this species across the Indo-Pacific as well.
405

406 Overall this study demonstrates pool-seq is a powerful and cost-effective tool for analyzing large
407 portions of the genome which the methods traditionally used for elasmobranchs could not
408 supply. Sharks and rays are an imperiled group of species that could benefit from advanced
409 genomic studies to outline appropriate management units. Finally, although the technology is
410 becoming cheaper and easier to apply, it is a common pitfall to assume everyone in the field can
411 afford, or must use, these approaches to produce defensible science. Bowen et al. (2014)
412 advocate judicious rather than wholesale application of genomic approaches as the most robust

course of study, particularly when considering the global inequities in available research budgets. Sanger sequencing is still more cost effective for small numbers of individuals, but as the number of individuals included in a study rise, the cost per individual reaches the point where high throughput sequencing studies can be cheaper than sequencing a single mitochondrial marker from each individual. We provide an example of just such a case here, and highlight the potential advantage of cost savings together with increased power for resolution of fine scale population structure. Though there is still additional cost of using cluster computer servers and bioinformatics programs, these cost are dropping as technology advances. When study organism and sampling strategies are assessed and implemented into the study design, pool-seq has great promise for augmenting the scientific foundations for management of marine recourses.

Acknowledgements

This study was made possible by the generous donation of specimens by Christopher R. Clarke, Mahmood Shivji, Stephen A. Karl, J.D. Filmalter, and Julia Spaet. We thank members of the ToBo Lab for sharing expertise, advice and discussions that contributed to this manuscript. Special thanks to Darren Lerner, Kim Holland, Carl Meyer, S. Gulak, D. Bethe, D. McCauley, C. Wilson, Guy Harvey Ocean Foundation, and Save Our Seas Foundation. This paper is funded in part by a cooperative agreement from the National Oceanic and Atmospheric Administration, Project R/SS-19PD, which is sponsored by the University of Hawai‘i Sea Grant College Program under Institutional Grant No. NA14OAR4170071 (B.W.B) from NOAA Office of Sea Grant, Department of Commerce. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. This is contribution #1821 from the Hawaii Institute of Marine Biology, contribution #XXXX from the Hawaii Sea Grant

Program, and contribution #11128 from the School of Ocean and Earth Science and Technology at the University of Hawaii.

Author Contributions:

Derek W. Kraft – Designed research, performed research, analyzed data, wrote paper, provided funding for research.

Emily Conklin – Created bioinformatics pipeline for data analysis and assisted in data analysis

Evan Barba – Created bioinformatics pipeline for data analysis and assisted in data analysis

Melanie Hutchinson – Helped obtain samples and provided biological insight

Robert J. Toonen – Contributed funding, contributed to research design, and contributed to data analysis and edited paper

Zac H. Forsman – Contributed to research design and data analysis

Brian W. Bowen – Contributed funding, contributed to research design, and edited paper

References

Anand, S., Mangano, E., Barizzone, N., Bordoni, R., Sorosina, M., Clarelli, F., Corrado, L., Boneschi, F. M., D'Alfonso, S., & De Bellis, G. (2016). Next generation sequencing of pooled samples: Guideline for variants' filtering. *Scientific Reports*, *6*(August), 1–9. https://doi.org/10.1038/srep33735

Anderson, E. C., Skaug, H. J., & Barshis, D. J. (2014). Next-generation sequencing for molecular ecology: A caveat regarding pooled samples. In *Molecular Ecology* (Vol. 23, Issue 3, pp. 502–512). https://doi.org/10.1111/mec.12609

Andrews, K.R., & Luikart, G. (2014). Recent novel approaches for population genomics data

analysis. *Molecular Ecology*, *23*(7), 1661–1667.

Andrews, Kimberly R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Avise, J. C., Bowen, B. W., Bermingham, E., Meylan, A. B., & Lamb, T. (1992). Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Molecular Biology & Evolution*, *9*, 457–473.

Barrio, A. M., Lamichhaney, S., Fan, G., Rafati, N., Pettersson, M., Zhang, H., Dainat, J., Ekman, D., Höppner, M., Jern, P., Martin, M., Nystedt, B., Liu, X., Chen, W., Liang, X., Shi, C., Fu, Y., Ma, K., Zhan, X., … Andersson, L. (2016). The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *ELife*, *5*(MAY2016), 1–32. https://doi.org/10.7554/eLife.12081

Baum, J. K., Myers, R. A., Kehler, D. G., Worm, B., Harley, S. J., & Doherty, P. A. (2003). Collapse and conservation of shark populations in the Northwest Atlantic. *Science*, *299*(5605), 389–392. https://doi.org/10.1126/science.1079777

Bernard, A. M., Horn, R. L., Chapman, D. D., Feldheim, K. A., Garla, R. C., Brooks, E. J., Gore, M. A., & Shivji, M. S. (2017). Genetic connectivity of a coral reef ecosystem predator: the population genetic structure and evolutionary history of the Caribbean reef shark (Carcharhinus perezi). *Journal of Biogeography*, *44*(11), 2488–2500. https://doi.org/10.1111/jbi.13062

Bird, C.E., Karl, S.A., Smouse, P.E. and Toonen, R.J. (2011). Detecting and measuring genetic differentiation. *Phylogeography and population genetics in Crustacea*, *19*(3):31-55.

Bowen, B. W., Bass, A. L., Soares, L., & Toonen, R. J. (2005). Conservation implications of

482 complex population structure : lessons from the loggerhead turtle (Caretta caretta).
483 *Molecular Ecology*, *14*, 2389–2402. https://doi.org/10.1111/j.1365-294X.2005.02598.x
484 Bowen, B.W., Shanker, K., Yasuda, N., Celia, M., Malay, M.C.M.D., von der Heyden, S.,
485 Paulay, G., Rocha, L.A., Selkoe, K.A., Barber, P.H. and Williams, S.T. (2014).
486 Phylogeography unplugged: comparative surveys in the genomic era. Bulletin of Marine
487 Science, 90(1):13-46.
488 Brown, W. M., George, M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial
489 DNA. *Proceedings of the National Academy of Sciences of the United States of America*,
490 *76*(4), 1967–1971. https://doi.org/10.1146/annurev.es.18.110187.001413
491 Cardeñosa, D., Fields, A. T., Babcock, E. A., Zhang, H., Feldheim, K., Shea, S. K. H., Fischer,
492 G. A., & Chapman, D. D. (2018). CITES-listed sharks remain among the top species in the
493 contemporary fin trade. *Conservation Letters*, *11*(4), 1–7.
494 https://doi.org/10.1111/conl.12457
495 Carvalho, G. R., & Hauser, L. (1994). Molecular genetics and the stock concept in fisheries.
496 *Reviews in Fish Biology and Fisheries*, *4*(3), 326–350. https://doi.org/10.1007/BF00042908
497 Castro, A. L. F., Stewart, B. S., Wilson, S. G., Hueter, R. E., Meekan, M. G., Motta, P. J.,
498 Bowen, B. W., & Karl, S. A. (2007). Population genetic structure of Earth's largest fish, the
499 whale shark (Rhincodon typus). *Molecular Ecology*, *16*(24), 5183–5192.
500 https://doi.org/10.1111/j.1365-294X.2007.03597.x
501 Charlesworth, D., & Wright, S. I. (2001). Breeding systems and genome evolution. *Current*
502 *Opinion in Genetics & Development*, *11*(6), 685–690. https://doi.org/10.1016/S0959-
503 437X(00)00254-9
504 Clarke, C. R., Karl, S. A., Horn, R. L., Bernard, A. M., Lea, J. S., Hazin, F. H., Prodöhl, P. A., &

505 Shivji, M. S. (2015). Global mitochondrial DNA phylogeography and population structure
506 of the silky shark, Carcharhinus falciformis. *Marine Biology*, *162*(5), 945–955.
507 https://doi.org/10.1007/s00227-015-2636-6
508 Clarke, S. C., McAllister, M. K., Milner-Gulland, E. J., Kirkwood, G. P., Michielsens, C. G. J.,
509 Agnew, D. J., Pikitch, E. K., Nakano, H., & Shivji, M. S. (2006). Global estimates of shark
510 catches using trade records from commercial markets. *Ecology Letters*, *9*(10), 1115–1126.
511 https://doi.org/10.1111/j.1461-0248.2006.00968.x
512 Daly-Engel, T. S., Seraphin, K. D., Holland, K. N., Coffey, J. P., Nance, H. A., Toonen, R. J., &
513 Bowen, B. W. (2012). Global phylogeography with mixed-marker analysis reveals male-
514 mediated dispersal in the endangered scalloped hammerhead shark (sphyrna lewini). *PLoS*
515 *ONE*, *7*(1), e29986. https://doi.org/10.1371/journal.pone.0029986
516 Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., Handsaker, R.
517 E., Lunter, G., Marth, G. T., Sherry, S. T., Mcvean, G., & Durbin, R. (2011). *The variant*
518 *call format and VCFtools*. *27*(15), 2156–2158.
519 https://doi.org/10.1093/bioinformatics/btr330
520 DiBattista, J. D., Waldrop, E., Rocha, L. A., Craig, M. T., Beruman, M. L., & Bowen, B. W.
521 (2015). Blinded by the bright: A lack of congruence between color morphs,
522 phylogeography, and taxonomy in a cosmopolitan Indo-Pacific butterflyfish, Chaetodon
523 auriga. *Journal of Biogeography*, *42*, 1919–1929.
524 Dizon, A. E., Lockyer, C., Perrin, W. F., Demaster, D. P., & Sisson, J. (1993). Rethinking the
525 stock concept: a phylogeographic approach. *Biological Conservation*, *64*(2), 176–177.
526 https://doi.org/10.1016/0006-3207(93)90670-V
527 Domingues, R. R., Hilsdorf, A. W. S., Shivji, M. M., Hazin, F. V. H., & Gadig, O. B. F. (2017).

528 Effects of the Pleistocene on the mitochondrial population genetic structure and
529 demographic history of the silky shark (*Carcharhinus falciformis*) in the western Atlantic
530 Ocean. *Reviews in Fish Biology and Fisheries*, *28*, 213 - 227.
531 https://doi.org/10.1007/s11160-017-9504-z
532 Domingues, R. R., Hilsdorf, A. W. S., & Gadig, O. B. F. (2018a). The importance of considering
533 genetic diversity in shark and ray conservation policies. *Conservation Genetics*, *19*, 501–
534 525. https://doi.org/10.1007/s10592-017-1038-3
535 Domingues, R. R., Bruels, C. C., Gadig, O. B. F., Chapman, D. D., Hilsdorf, A. W. S., & Shivji,
536 M. S. (2018b). Genetic connectivity and phylogeography of the night shark (*Carcharhinus*
537 *signatus*) in the western Atlantic Ocean: Implications for conservation management.
538 *Aquatic Conservation: Marine and Freshwater Ecosystems*, *29*(1), 102–114.
539 https://doi.org/10.1002/aqc.2961
540 Dulvy, N. K., Baum, J. K., Clarke, S., Compagno, L. J. V., Cortés, E., Domingo, A., Fordham,
541 S., Fowler, S., Francis, M. P., Gibson, C., Martínez, J., Musick, J. A., Soldo, A., Stevens, J.
542 D., & Valenti, S. (2008). You can swim but you can't hide: The global status and
543 conservation of oceanic pelagic sharks and rays. *Aquatic Conservation: Marine and*
544 *Freshwater Ecosystems*, *18*(5), 459–482. https://doi.org/10.1002/aqc.975
545 Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R.,
546 Carlson, J. K., Davidson, L. N. K., Fordham, S. V., Francis, M. P., Pollock, C. M.,
547 Simpfendorfer, C. A., Burgess, G. H., Carpenter, K. E., Compagno, L. J. V., Ebert, D. A.,
548 Gibson, C., Heupel, M. R., Livingstone, S. R., … White, W. T. (2014). Extinction risk and
549 conservation of the world's sharks and rays. *ELife*, *2014*(3), 1–34.
550 https://doi.org/10.7554/eLife.00590.001

551 Dulvy, N. K., & Trebilco, R. (2018). Size-Based Insights into the Ecosystem Role of Sharks and
552 Rays. In *Shark Research: Emerging Technologies and Applications for the Field and*
553 *Laboratory*.
554 Duncan, K. M., Martin, A. P., Bowen, B. W., & De Couet, H. G. (2006). Global phylogeography
555 of the scalloped hammerhead shark (Sphyrna lewini). *Molecular Ecology*, *15*(8), 2239–
556 2251. https://doi.org/10.1111/j.1365-294X.2006.02933.x
557 Ewels, P., Lundin, S., & Max, K. (2016). Data and text mining MultiQC : summarize analysis
558 results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048.
559 https://doi.org/10.1093/bioinformatics/btw354
560 Ferretti, F., Worm, B., Britten, G. L., Heithaus, M. R., & Lotze, H. K. (2010). Patterns and
561 ecosystem consequences of shark declines in the ocean. In *Ecology Letters* (Vol. 13, Issue
562 8, pp. 1055–1071). https://doi.org/10.1111/j.1461-0248.2010.01489.x
563 Ferretti, L., Ramos-Onsins, S. E., & Pérez-Enciso, M. (2013). Population genomics from pool
564 sequencing. *Molecular Ecology*, *22*(22), 5561–5576. https://doi.org/10.1111/mec.12522
565 Feutry, P., Devloo-Delva, F., Tran Lu Y, A., Mona, S., Gunasekera, R.M., Johnson, G., Pillans,
566 R.D., Jaccoud, D., Kilian, A., Morgan, D.L. and Saunders, T. (2020). One panel to rule
567 them all: DArTcap genotyping for population structure, historical demography, and kinship
568 analyses, and its application to a threatened shark. Molecular Ecology Resources. Online
569 Early, https://doi.org/10.1111/1755-0998.13204
570 Fischer, M. C., Rellstab, C., Tedder, A., Zoller, S., Gugerli, F., Shimizu, K. K., Holderegger, R.,
571 & Widmer, A. (2013). Population genomic footprints of selection and associations with
572 climate in natural populations of Arabidopsis halleri from the Alps. *Molecular Ecology*,
573 *22*(22), 5594–5607. https://doi.org/10.1111/mec.12521

574 Futschik, A., & Schlötterer, C. (2010). The next generation of molecular markers from massively
575 parallel sequencing of pooled DNA samples. *Genetics*, *186*(1), 207–218.
576 https://doi.org/10.1534/genetics.110.114397
577 Garnier-Géré, P., & Chikhi, L. (2013) Population subdivision, Hardy-Weinberg equilibrium, and
578 the Wahlund effect. eLS Books, Wiley Online Library Pub. 3
579 https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470015902.a0005446.pub3
580 Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing.
581 1–9.
582 Gautier, M., Foucaud, J., Gharbi, K., Cézard, T., Galan, M., Loiseau, A., Thomson, M., Pudlo,
583 P., Kerdelhué, C., & Estoup, A. (2013). Estimation of population allele frequencies from
584 next-generation sequencing data: Pool-versus individual-based genotyping. *Molecular*
585 *Ecology*, *22*(14), 3766–3779. https://doi.org/10.1111/mec.12360
586 Green, M. E., Appleyard, S. A., White, W., Tracey, S., Devloo-Delva, F., & Ovenden, J. R.
587 (2019). Novel multimarker comparisons address the genetic population structure of silvertip
588 sharks (Carcharhinus albimarginatus). *Marine and Freshwater Research*, *70*(7), 1007–1019.
589 https://doi.org/10.1071/MF18296
590 Guo, B., Li, Z., & Merilä, J. (2016). Population genomic evidence for adaptive differentiation in
591 the Baltic Sea herring. *Molecular Ecology*, *25*(12), 2833–2852.
592 https://doi.org/10.1111/mec.13657
593 Heist, E. J. (2004). *Biology of Sahrks and Their Relatives* (J. C. Carrier, J. A. Musick, & M. .
594 Heithaus (eds.)). CRC Press, Boca Raton, FL.
595 Heithaus, M. R., Frid, A., Wirsing, A. J., & Worm, B. (2008). Predicting ecological
596 consequences of marine top predator declines. *Trends in Ecology and Evolution*, *23*(4),

202–210. https://doi.org/10.1016/j.tree.2008.01.003

Heupel, M. R., Knip, D. M., Simpfendorfer, C. A., & Dulvy, N. K. (2014). Sizing up the ecological role of sharks as predators. *Marine Ecology Progress Series*, *495*, 291–298. https://doi.org/10.3354/meps10597

Hivert, V., Leblois, R., Petit, E. J., Gautier, M., & Vitalis, R. (2018). Measuring genetic differentiation from Pool-seq data. *Genetics*, *210*(September), 315–330. https://doi.org/10.1534/genetics.118.300900

Hoelzel, A. R., Shivji, M. S., Magnussen, J. E., & Francis, M. P. (2006). Low worldwide genetic diversity in the basking shark (Cetorhinus maximus). *Biology Letters*, *2*, 639–642. https://doi.org/10.1098/rsbl.2006.0513

Hohenlohe, P. A., Hand, B. K., Andrews, K. R., & Luikart, G. (2018). Population genomics provides key insights in ecology and evolution. *Population Genomics: Concepts, Approaches and Applications*, doi:10.1007/13836_2018_20. https://doi.org/10.1007/13836_2018_20

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics 6(2): e1000862.

IUCN, I. U. for the C. of N. (2017). *IUCN Red List of Threatened Species*. Version 2013.1. http://www.iucnredlist.org/about/overview

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*,

484(7392), 55–61. https://doi.org/10.1038/nature10944

Karl, S. A., Castro, A. L. F., Lopez, J. A., Charvet, P., & Burgess, G. H. (2011). Phylogeography and conservation of the bull shark (Carcharhinus leucas) inferred from mitochondrial and microsatellite DNA. *Conservation Genetics*, *12*(2), 371–382. https://doi.org/10.1007/s10592-010-0145-1

Karl, S. a, Bowen, B. W., & Avise, J. C. (1992). Global population genetic structure and male-mediated gene flow in the green turtle (Chelonia mydas): RFLP analysis of anonymous nuclear loci. *Genetics*, *131*, 163–173. https://doi.org/10.1139/z05-185

Keeney, D. B., & Heist, E. J. (2006). Worldwide phylogeography of the blacktip shark (Carcharhinus limbatus) inferred from mitochondrial DNA reveals isolation of western Atlantic populations coupled with recent Pacific dispersal. *Molecular Ecology*, *15*, 3669–3679.

Knapp, I. S. S., Puritz, J. B., Bird, C. E., Whitney, J. L., Sudek, M., Forsman, Z. H., & Toonen, R. J. (2016). *ezRAD- an accessible next-generation RAD sequencing protocol suitable for non-model organisms_v3.2*. https://doi.org/dx.doi.org/10.17504/protocols.io.e9pbh5n

Kofler, R., Betancourt, A. J., & Schlötterer, C. (2012). Sequencing of pooled DNA samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in Drosophila melanogaster. *PLoS Genetics*, *8*(1). https://doi.org/10.1371/journal.pgen.1002487

Kofler, R., Pandey, R. V., & Schlötterer, C. (2011). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, *27*(24), 3435–3436. https://doi.org/10.1093/bioinformatics/btr589

Konczal, M., Koteja, P., Stuglik, M.T., Radwan, J. and Babik, W., 2014. Accuracy of allele frequency estimation using pooled RNA-Seq. Molecular Ecology Resources, 14(2):381-

392.

Kurland, S., Wheat, C.W., de la Paz Celorio Mancera, M., Kutschera, V.E., Hill, J., Andersson, A., Rubin, C.J., Andersson, L., Ryman, N. and Laikre, L., (2019). Exploring a Pool-seq-only approach for gaining population genomic insights in nonmodel species. *Ecology and evolution*, *9*(19), pp.11448-11463.

Larsson, L.C., Charlier, J., Laikre, L. and Ryman, N.(2009). Statistical power for detecting genetic divergence—organelle versus nuclear markers. *Conservation Genetics*, *10*(5): 1255.

Legendre, P., & Legendre, L. (1998). *Numerical Ecology. 2nd English Edition.*

Martin, A. P., Naylor, G., & Palumbi, S. R. (1992). Rates of mitochondrial DNA evolution in sharks are slow compared to mammals. *Nature*, *357*, 153–155. https://doi.org/10.1038/357153a0

Mimee, B., Duceppe, M. O., Véronneau, P. Y., Lafond-Lapalme, J., Jean, M., Belzile, F., & Bélair, G. (2015). A new method for studying population genetics of cyst nematodes based on Pool-Seq and genomewide allele frequency analysis. *Molecular Ecology Resources*, *15*(6), 1356–1365. https://doi.org/10.1111/1755-0998.12412

Momigliano, P., Harcourt, R., Robbins, W. D., Jaiteh, V., Mahardika, G. N., Sembiring, A., & Stow, A. (2017). Genetic structure and signatures of selection in grey reef sharks (Carcharhinus amblyrhynchos). *Heredity*, *119*(3), 142–153. https://doi.org/10.1038/hdy.2017.21

Musick, J. A., Burgess, G., Cailliet, G., Camhi, M., & Fordham, S. (2000). Management of Sharks and Their Relatives (Elasmobranchii). *Fisheries*, *25*(3), 9–13. https://doi.org/10.1577/1548-8446(2000)025<0009:MOSATR>2.0.CO;2

666 Myers, R. A., Baum, J. K., Shepherd, T. D., Powers, S. P., & Peterson, C. H. (2007). Cascading
667 effects of the loss of apex predatory sharks from a coastal ocean. *Science*, *315*(5820), 1846–
668 1850. https://doi.org/10.1126/science.1138657
669 Neiman, M., & Taylor, D. R. (2009). The causes of mutation accumulation in mitochondrial
670 genomes. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1660), 1201–1209.
671 https://doi.org/10.1098/rspb.2008.1758
672 Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R. and Hohenlohe, P.A. (2013).
673 Genotyping-by-sequencing in ecological and conservation genomics. *Molecular*
674 *ecology*, *22*(11): 2841-2847.
675 Oliver, S., Braccini, M., Newman, S. J., & Harvey, E. S. (2015). Global patterns in the bycatch
676 of sharks and rays. *Marine Policy*, *54*, 86–97. https://doi.org/10.1016/j.marpol.2014.12.017
677 Ovenden, J. R., Berry, O., Welch, D. J., Buckworth, R. C., & Dichmont, C. M. (2015). Ocean ' s
678 eleven : a critical evaluation of the role of population , evolutionary and molecular genetics
679 in the management of wild fi sheries. *Fish Fisheries*, *16*(125–159).
680 https://doi.org/10.1111/faf.12052
681 Pardini, A. T., Jones, C. S., Noble, L. R., Kreiser, B., & Malcolm, H. (2001). Sex-biased
682 dispersal of great white sharks. *Nature*, *412*, 139–140.
683 Pazmiño, D. A., Maes, G. E., Green, M. E., Simpfendorfer, C. A., Hoyos-Padilla, E. M., Duffy,
684 C. J. A., Meyer, C. G., Kerwath, S. E., Salinas-De-León, P., & Van Herwerden, L. (2018).
685 Strong trans-Pacific break and local conservation units in the Galapagos shark
686 (Carcharhinus galapagensis) revealed by genome-wide cytonuclear markers. *Heredity*,
687 *120*(5), 407–421. https://doi.org/10.1038/s41437-017-0025-2
688 Perez-Enciso, M., & Ferretti, L. (2010). Massive parallel sequencing in animal genetics :

689 wherefroms and wheretos. *Animal Genetics*, *41*, 561–569. https://doi.org/10.1111/j.1365-
690 2052.2010.02057.x
691 Portnoy, D. S., McDowell, J. R., Heist, E. J., Musick, J. A., & Graves, J. E. (2010). World
692 phylogeography and male-mediated gene flow in the sandbar shark, Carcharhinus
693 plumbeus. *Molecular Ecology*, *19*, 1994–2010.
694 Portnoy, D.S., Puritz, J.B., Hollenbeck, C.M., Gelsleichter, J., Chapman, D. and Gold, J.R.
695 (2015). Selection and sex-biased dispersal in a coastal shark: The influence of philopatry on
696 adaptive variation. Molecular Ecology, 24(23), pp.5877-5885.
697 Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). *dDocent* : a RADseq, variant-calling
698 pipeline designed for population genomics of non-model organisms. *PeerJ*, *2*, e431.
699 https://doi.org/10.7717/peerj.431
700 Qi, H., Song, K., Li, C., Wang, W., Li, B., Li, L. and Zhang, G. (2017). Construction and
701 evaluation of a high-density SNP array for the Pacific oyster (Crassostrea gigas). *PLoS*
702 *One*, *12*(3), p.e0174007.
703 Rice, J., & Harley, S. (2013). Updated stock assessment of the Silky sharks in the Western and
704 Central Pacfific Ocean. *WCPFC-SC9-2013/ SA-WP-03*, *August*.
705 RStudio Team (2020). RStudio: Integrated Development Environment for R. RStudio, PBC,
706 Boston, MA
707 URL http://www.rstudio.com/
708 Ryman, N. and Palm, S., (2006). POWSIM: a computer program for assessing statistical power
709 when testing for genetic differentiation. *Molecular Ecology Notes*, *6*(3), pp.600-602.
710 Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals-
711 mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*,

712 *15*(11), 749–763. https://doi.org/10.1038/nrg3803
713 Schultz, J. K., Feldheim, K. A., Gruber, S. H., Ashley, M. V. H., McGovern, T. M., Bowen, B.
714 W., Bruber, S. H., Ashley, M. V. H., McGovern, T. M., & Bowen, B. W. (2008). Global
715 phylogeography and seascape genetics of the lemon sharks (genus Negaprion). *Molecular*
716 *Ecology*, *17*(24), 5336–5348. https://doi.org/10.1111/j.1365-294X.2008.04000.x
717 Stevens, J. D., Bonfil, R., Dulvy, N. K., & Walker, P. A. (2000). The effects of fishing on sharks,
718 rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. *ICES*
719 *Journal of Marine Science*, *57*(3), 476–494. https://doi.org/10.1006/jmsc.2000.0724
720 Tallmon, D. A., Gregovich, D., Waples, R., & Baker, C. S. (2010). When are genetic methods
721 useful for estimating contemporary abundance and detecting population trends? *Molecular*
722 *Ecology Resources*, *10*, 684–692.
723 Theisen, T. C., Bowen, B. W., Lanier, W., & Baldwin, J. D. (2008). High connectivity on a
724 global scale in the pelagic wahoo, Acanthocybium solandri (tuna family Scombridae).
725 *Molecular Ecology*, *17*, 4233–4247.
726 Therkildsen, N.O. and Palumbi, S.R. (2017). Practical low-coverage genomewide sequencing of
727 hundreds of individually barcoded samples for population and evolutionary genomics in
728 nonmodel species. Molecular ecology resources, 17(2):194-208.
729 Toonen, R.J. and Grosberg, R.K. (2011). Causes of chaos: spatial and temporal genetic
730 heterogeneity in the intertidal anomuran crab *Petrolisthes cinctipes*. *Phylogeography and*
731 *population genetics in Crustacea*, 2011: 75-107.
732 Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R.,
733 & Bird, C. E. (2013). ezRAD: a simplified method for genomic genotyping in non-model
734 organisms. *PeerJ*, *1*, e203. https://doi.org/10.7717/peerj.203

735 Weir, B.S. and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population
736 structure. Evolution 38(6):1358-1370.
737 Whitney, N. M., Robbins, W. D., Schultz, J. K., Bowen, B. W., & Holland, K. N. (2012).
738 Phylogeography of the whitetip reef shark (Triaenodon obesus): a sedentary species with a
739 broad distribution. *Journal of Biogeography*, *39*, 1144–1156.
740 Zink, R.M., Barrowclough, G. (2008). Mitochondrial DNA under siege in avian
741 phylogeography. *Molecular Ecology*, 17, 2107–2121.
742

Figure 1

Sample locations of *Carcharhinus falciformis* followed by sample size

Abbreviations: GM = Gulf of Mexico, NA = North Atlantic, BR = Brazil, RS = Red Sea.

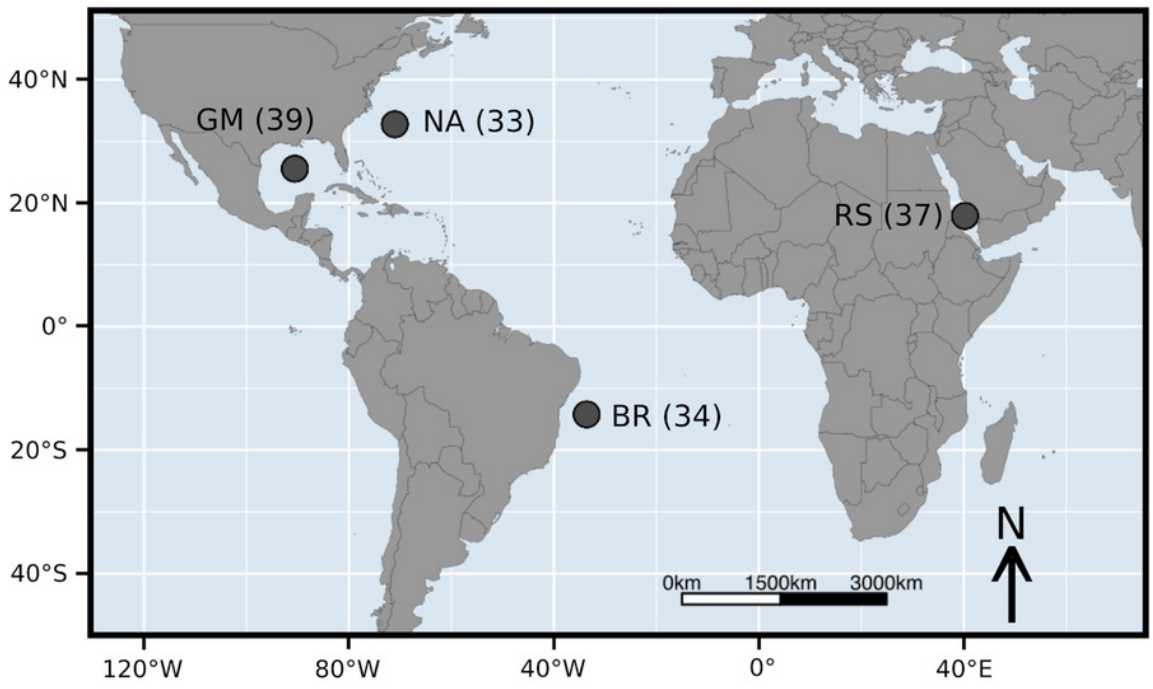


Figure 2

F_{st} heat map for mitochondrial and nuclear data compaired to Clarke et al. Phi_{st} and P-values

Figure 2. (A) Pairwise *FST* values generated by Pool-seq methods. Cool colors (top left) are *FST* values calculated from nuclear genome loci, warm colors (bottom right) are *FST* values from loci across the entire mitochondrial genome. All pairwise differences are significant ($p < 0.001$). **(B)** *PhiST* results from Clarke et. Al (2015) on the lower right triangle and P-values on the upper right triangle. Significant P-values and corresponding *PhiST* values in bold. Regional abbreviation are as follows; GM = Gulf of Mexico, BR = Brazil, NA = North Atlantic, RS = Red sea.

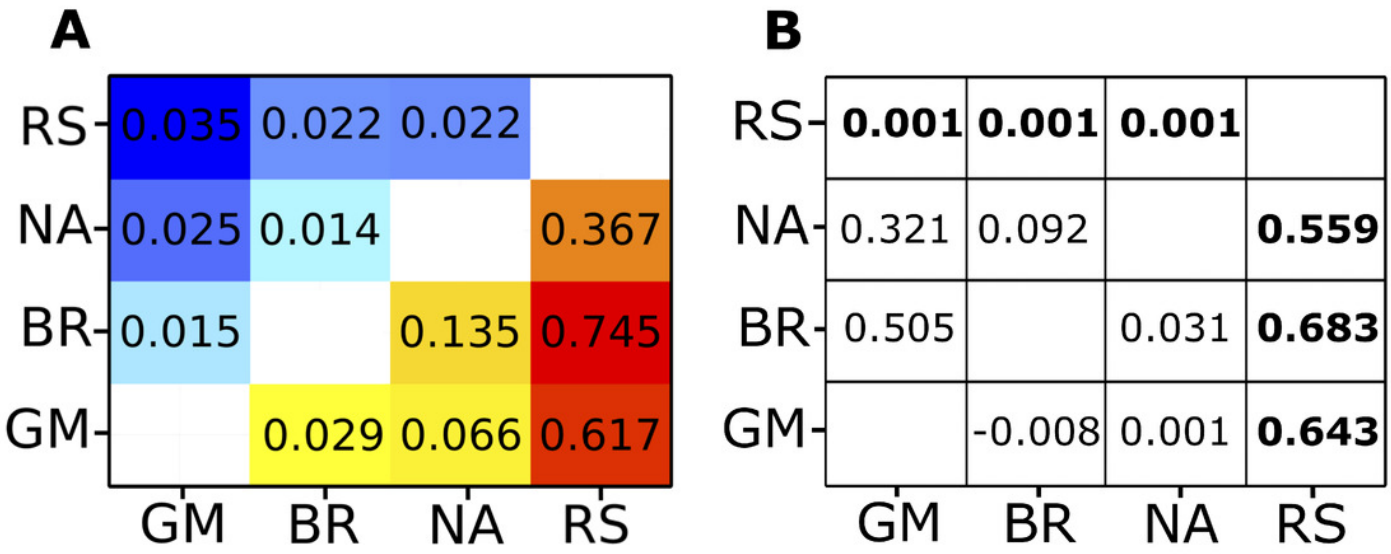
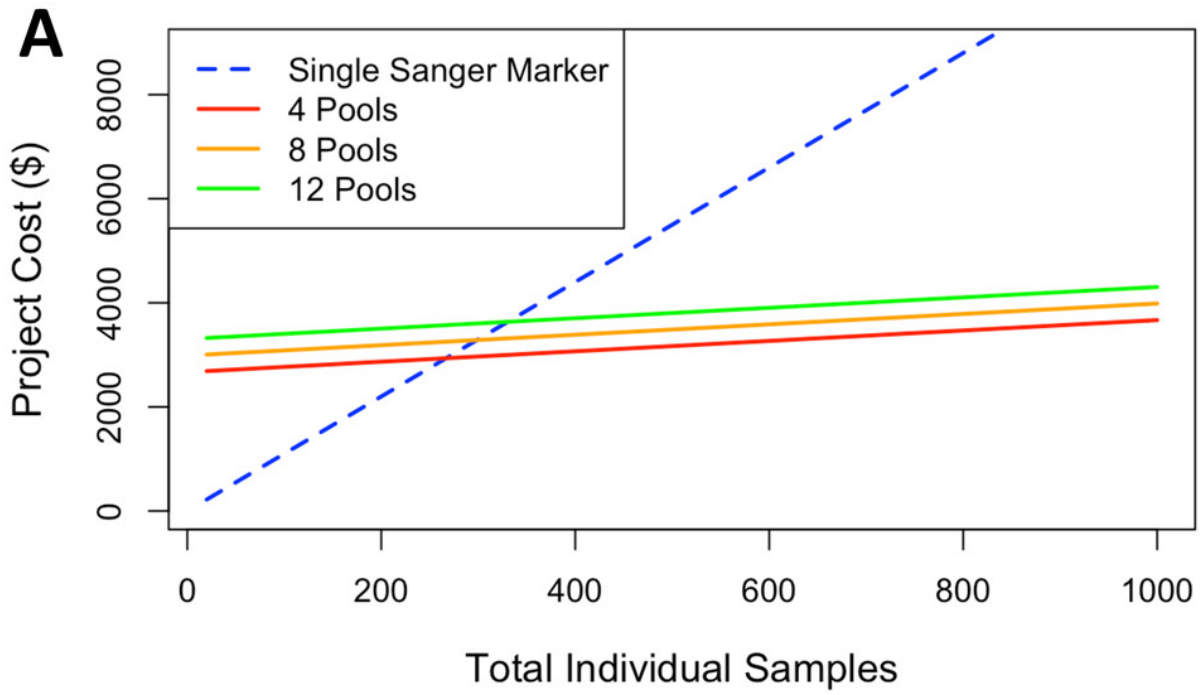
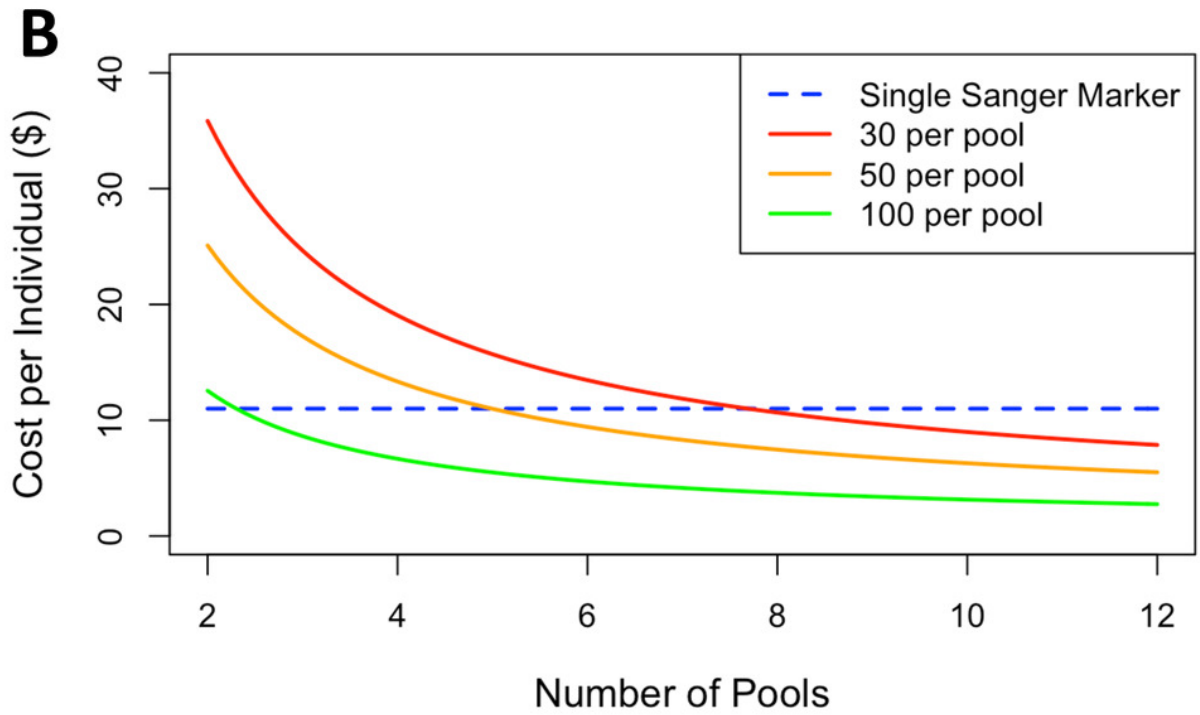

A
RS
NA
BR
GM
0.035
0.022
0.022
0.025
0.014
0.367
0.015
0.135
0.745
0.029
0.066
0.617
GM
BR
NA
RS
B
RS
NA
BR
GM
0.001
0.001
0.001
0.321
0.092
0.559
0.505
0.031
0.683
-0.008
0.001
0.643
GM
BR
NA
RS

Figure 3

Cost comparisons between sequencing projects using a single Sanger marker to projects using Pool-seq with varying numbers of pools.

(A) Sequencing costs comparing number of individuals to total cost between Sanger at our facility and three Pool-seq projects at our facility containing 4, 8, and 12 pools respectively, where pool sizes change with number of individuals. **(B)** Sequencing cost per individual with fixed pools across different number of Pools.


A
Single Sanger Marker
4 Pools
8 Pools
12 Pools
Project Cost ($)
0
2000
4000
6000
8000
0
200
400
600
800
1000
Total Individual Samples


B
Single Sanger Marker
30 per pool
50 per pool
100 per pool
Cost per Individual ($)
0
10
20
30
40
2
4
6
8
10
12
Number of Pools