2-2015

Population genomic inference of ecology, conservation, evolution, and demographic history of Atlantic seahorses and pipefishes (Syngnathidae)

Joel Thomas Boehm
Graduate Center, City University of New York

Population genomic inference of ecology, conservation, evolution, and demographic history of Atlantic seahorses and pipefishes (Syngnathidae)

By

J.T. Boehm

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2015

This manuscript has been read and accepted for the
Graduate Faculty in Biology in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Michael J. Hickerson_____

_____ _____
Date

Laurel Eckhardt_____

_____ _____
Date Executive Officer

Ana C. Carnaval_____

John Waldman_____

Rob DeSalle_____

Christopher P. Meyer_____
Supervisory Committee

The City University of New York

Abstract

Population genomic inference of ecology, conservation, evolution, and demographic history of Atlantic seahorses and pipefishes (Syngnathidae)

By

J.T. Boehm

Advisor: Dr. Michael J. Hickerson

In the Atlantic Ocean powerful directional ocean currents can play a significant role in the formation and persistence of marine species. Syngnathidae fishes have a sparse fossil record, high morphological plasticity, and many of these species are difficult to observe in the wild, therefore they frequently lack life history information and the status of regional lineages and species designations are often obscure. In this dissertation I explore the ecology, evolution, and conservation of primarily Atlantic seahorses (*Hippocampus*) and pipefish (*Syngnathus*) in four core chapters, using differing genetic datasets ranging from mitochondrial DNA to genome-wide RAD sequences. Most Synganthids have the potential to disperse passively by rafting on floating vegetation, and are direct developers, which is thought to limit their active mobility, yet many species have widespread distributions. The majority of genetic research on Syngnathids fishes has focused on Indo-Pacific species, however the Atlantic Ocean is home to dozens of species of pipefishes from nine genera and roughly 1/5th of the world's seahorses species. In Chapter 1, I use six loci to infer the species tree for all Atlantic seahorses and infer the demographic history and evolution of the "*Hippocampus erectus* complex." The results of this study support the establishment of an ancestral population of the *H. erectus* complex in the Americas, followed by the Amazon River outflow splitting it into Caribbean/North American *H. erectus* and South American *H. patagonicus* at a time of increased sedimentation and outflow. Following this split, colonization occurred across the Atlantic via the Gulf Stream currents with subsequent trans-Atlantic isolation. Based on the results of Chapter 1, the species *H. erectus* exhibited a panmictic genetic structure

from Latin America to temperate New York waters. However, inhabitants of the temperate region are considered by some ecologists to be tropical vagrants that only arrive during warm seasons from the southern provinces and perish as temperatures decline. Contrary to the findings of Chapter 1, in Chapter 2, I use thousands of RADseq loci and show strong support that temperate inhabitants are genetically diverged from southern populations and are composed of an isolated and persistent ancestral gene pool. The aim of Chapter 3 is to investigate how major current forces as well as climatic and geographic processes have shaped the evolutionary and demographic history of western Atlantic seahorses (*Hippocampus*) and pipefishes (*Syngnathus*). This Chapter takes a comparative approach across five codistributed species (two seahorses and three pipefishes). Genomic patterns of subpopulation divergence and post-divergence gene flow may be shared amongst fish species with similar life history traits, however ecological differences (i.e., macroclimatic tolerance and rafting propensity) may impact the rates of gene exchange and/or isolation times between subpopulations. The result of this study show how directional ocean currents and the life history trait of rafting propensity impacts population divergence and connectivity, and predicts gene flow directionality and magnitude in four out of five of the focal taxa. Lastly in Chapter 4, I use a molecular forensics approach to track the U.S. dried seahorse trade. Due to global exploitation, the genus *Hippocampus* are the only fish to have all species listed under the Convention of International trade of endangered species (CITES). Millions of individuals are traded each year for the use in traditional Chinese medicine as well as for souvenirs and crafts. Using "DNA barcoding," while mentoring high school and undergraduate students, we identified and compared specimens collected from two primary U.S. dried seahorse end-markets: 1) traditional Chinese medicine and, 2) Internet and coastal souvenir retailers. The results of this study found a significant contrast in both the species composition and size of individuals being sold between each market.

ACKNOWLEDGMENTS

I would like to thank the Sackler Institute for Comparative Genomics at the American Museum of Natural History and Rob DeSalle and George Amato in particular for providing laboratory space, support, and assistance. I also thank the following persons and institutions for providing invaluable assistance in obtaining specimens including D. Luzzatto; N. Dunham (Florida Fish and Wildlife Conservation Commission), T. Tuckey (Virginia Institute of Marine Science), T. Gardner (Atlantis Aquarium), The River Project (Pier 40, Manhattan), T. M. Grothues (Rutgers University), Clay Small (University of Oregon), Adam Jones and his lab members (Texas A&M University), Susan Pell (Brooklyn Botanical Gardens) and Kenyon Mobley (Max Planck Institute for Evolutionary Biology). Funding for this research was provided in part by CUNY Science Fellowship Program, The National Science Foundation Doctoral Dissertation Improvement Grant, Explorers Club Grant, Lerner-Gray Award in Marine Research, PSC-CUNY, Queens College Graduate Advisory Council, and Seymour Fogel Fund. I also give special thanks for the support, advice and encouragement of friends and fellow CUNY colleagues including Terry Demos, Stephen Harris, John Robinson, Katriina Ilves, Alexander Xue, Susan Tsang, Xin Chen, Tyler Joseph, Diego Alvarado-Serrano, Miguel Pinto, Michael Levandowsky, Nina Zain, Chris Anderson, Ana Cristina Fazza, as well as research collaborators Carole Baldwin (Smithsonian Institute), Sara Lourie (McGill University), Lucy Woodall (University of London), Peter Teske (University of Johannesburg), Nathan Putman (NOAA Southeast Fisheries Science Center), Matthieu Leray (Smithsonian Institute), and Sergio Floeter (Universidade Federal de Santa Catarina). In addition, a special thanks to my family for all their support, including Sylvia

Cohn, Claudia Boehm, Joel T. Boehm, Sr., Allison Boehm, Ashley Boehm, Andrew Boehm, my in-laws the Lee family, and especially my wife Narae Lee. Finally, I thank my advisor Mike Hickerson and the members of my Ph.D. committee, John Waldman, Ana Carnaval, Rob DeSalle and Chris Meyer, for their sound advice, insight, interest, and support throughout the course of this study.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Appendices

CHAPTER 1

Marine barriers and dispersal drive Atlantic seahorse diversification

ABSTRACT

To investigate how marine barriers shaped the demographic history of Atlantic seahorses (Syngnathidae: *Hippocampus*) we used range-wide sampling (n = 390) at mitochondrial and up to five nuclear DNA loci was carried out across the *Hippocampus erectus* species complex (*H. erectus* from the Caribbean/North America, *H. patagonicus* from South America and *H. hippocampus* from Europe and West Africa). Multi-species coalescent and approximate Bayesian computation (ABC) frameworks were used to estimate support of competing biogeographical hypotheses and demographic parameters, including lineage divergence times, effective population sizes and magnitudes of population size change. We identified four distinct lineages within the *H. erectus* complex. A posterior probability of 0.626 and corresponding Bayes factors ranging from 3.68 to 11.38 gave moderate to strong support for a basal divergence between South American populations of *H. patagonicus* and Caribbean/North American populations of *H. erectus* coincident with the inter-regional freshwater outflow of the Amazon River Barrier (ARB). Estimates of historical effective population sizes and divergence times indicate that European and West African populations of *H. hippocampus* expanded after colonization from a more demographically stable Caribbean/North American *H. erectus*. Our findings of trans-Atlantic colonization followed by isolation across a deep oceanic divide, and isolation across a freshwater barrier, may demonstrate a contrast in marine divide permeability for this group of rafters. Demographic inference supports the establishment of an ancestral population of the *H. erectus* complex in the Americas,

followed by the ARB splitting it into Caribbean/North and South American lineages at a time of increased sedimentation and outflow. Our estimates suggest that following this split, colonization occurred across the Atlantic via the Gulf Stream currents with subsequent trans-Atlantic isolation. These results illustrate that rafting can be a means of range expansion over large distances, but may be insufficient for sustaining genetic connectivity across major barriers, thereby resulting in lineage divergence.

Keywords: Atlantic Ocean, *Hippocampus*, marine barriers, ocean currents, rafting, coalescent, approximate Bayesian computation, species tree, demographic history

INTRODUCTION

For marine taxa with freely mobile adults or larvae, mounting genetic evidence demonstrates that mechanisms other than strict vicariant speciation may drive diversification, including soft vicariance, parapatric speciation, ecological speciation and founder colonization (Rocha et al., 2005b; Lessios & Robertson, 2006; Hickerson & Meyer, 2008; Rocha & Bowen, 2008; Luiz et al., 2012). These processes are undoubtedly affected by the permeability of marine divides, which are proposed to serve as 'biogeographical filters' rather than strict barriers (Vermeij, 1978; Floeter et al., 2007). Hence, the interplay between ocean currents, the emergence of oceanic barriers, organismal dispersal potential and organismal environmental tolerances all contribute to patterns of diversification in the oceans (Joyeux et al., 2001; Muss et al., 2001; Floeter et al., 2007; Rocha et al., 2008; Hellberg, 2009).

Seahorses (Syngnathidae: *Hippocampus*; nomenclature follows (Lourie et al., 1999; Piacentino & Luzzatto, 2004) are presumed to have a low dispersal potential because they are poor swimmers and lack a pelagic larval phase, yet many species have vast distributions (thousands of kilometres) with variable genetic connectivity, which suggests moderate to high long-distance dispersal capabilities (Lourie & Vincent, 2004; Lourie et al., 2005; Teske et al., 2005; Woodall et al., 2011). This apparent paradox probably arises from the ability to raft on floating mats of Sargassum and other macroalgae (Gunther, 1870; Casazza & Ross, 2008; Woodall et al., 2009), a mode of dispersal utilized by many fishes (Casazza & Ross, 2008; Luiz et al., 2012) and invertebrates (Ingolfsson, 1995; Theil & Gutow, 2005a; Fraser et al., 2011). Rafting has been shown to be a strong general predictor of colonization for coral reef fishes across the deep oceanic barrier of the equatorial mid-Atlantic (Luiz et al., 2012). Consistent with

this trend, recent studies have concluded that diversification in Hippocampus may be driven by rafting-mediated founder colonization followed by levels of gene flow insufficient for maintaining genetic connectivity with the original population (Lourie et al., 2005; Teske et al., 2005, 2007). Therefore, rafting-mediated dispersal can be differentiated into two general scales: (1) regional gene flow that links demes by recruitment between spatially distinct populations; and (2) large-scale colonization that results in range expansion across marine barriers, followed by divergence when subsequent dispersal events are too rare to prevent isolation (i.e. < 1 migrating individual per generation) (Slatkin, 1985; Paulay & Meyer, 2002).

Here, we use population-level sampling to test four competing historical biogeographical hypotheses that involve divergences across oceanic divides and a major freshwater outflow in three widespread sibling species of seahorses: *H. erectus* from Caribbean/North America, *H. hippocampus* from Europe and West Africa, and *H. patagonicus* from South America (hereafter the *H. erectus* complex). The distribution of the *H. erectus* complex encompasses most of the tropical and warm-temperate coasts of the Atlantic (Fig. 1.1a), allowing us to study the impact of two types of major marine barriers on the diversification of this species complex: (1) the deep oceanic divides of the mid-northern and equatorial Atlantic; and (2) the Amazon River Barrier (ARB), the largest freshwater outflow in the world (Hallam, 1997; Rocha et al., 2005b). Both barrier types are permeable, but to cross them poses considerable challenges, and may affect rafting dispersers in different ways. For oceanic crossing, currents can provide directionality to patterns of colonization from source regions (Gillespie et al., 2012), and rafting vectors are predicted to disperse primarily with the prevailing asymmetrical clockwise surface currents of the North Atlantic Gyre system (Fig. 1.1b). For freshwater

outflows, the extreme environmental shift created by the ARB acts as a break between some Caribbean and Brazilian faunas, while others show evidence of genetic connectivity (Joyeux et al., 2001; Rocha et al., 2002, 2005a).

In order to investigate how these two types of barriers shaped diversification in *Hippocampus*, we use coalescent methods with multi-locus population genetic data (mitochondrial DNA and up to five nuclear DNA loci, from 390 individuals throughout the *H. erectus* complex range) to obtain the relative probabilities of the following four competing hypotheses whose branching order is depicted in Fig. 1.2.

1. $\mathbf{H_{EAO}}$ – western Atlantic lineages were the product of equatorial current dispersal from a basal north-eastern Atlantic population followed by ARB isolation between Caribbean/ North American and South American lineages.

2. $\mathbf{H_{WAO}}$ – north-eastern Atlantic lineages were the product of Gulf Stream dispersal from the Caribbean/North American lineage subsequent to the ARB isolating Caribbean/North American and South American lineages.

3. $\mathbf{H_{CLK}}$ – north-eastern Atlantic lineages were the product of Gulf Stream dispersal from a basal Caribbean/North American lineage prior to a colonization of South American lineages from the north-eastern Atlantic.

4. $\mathbf{H_{META}}$ – a large connected metapopulation spanning the Americas and the north-eastern Atlantic became isolated to form all contemporary *H. erectus* complex lineages.

To estimate the statistical support for each of the four hypotheses, we use an approximate Bayesian computation (ABC) framework. Importantly, this approach yields relative timeframe estimates for the corresponding demographic events, and patterns of population stability and instability within each modeled hypothesis. To put the inferences

in a larger context, we also reconstruct the species tree for all Atlantic seahorse species. The majority of Atlantic *Hippocampus* species remain 'data deficient' under IUCN Red List evaluations (IUCN, 2012), and our study is therefore well timed to help inform conservation management, while providing insight into ecological and evolutionary processes that drive diversification of Atlantic fishes.

MATERIALS AND METHODS

Taxonomic background of study taxa and sampling

Many *Hippocampus* species have a high degree of intraspecific morphological plasticity linked to environmental gradients and habitats resulting in historical taxonomic confusion (Ginsburg, 1937; Lourie et al., 2004). Lourie et al. (2004) resolved many of these taxonomic issues and recognized seven species of Atlantic seahorses. *Hippocampus erectus* was consolidated into a single species with a confirmed distribution along the coastal western Atlantic from Venezuela to Canada (Lourie et al., 2004), with individuals as far south as Argentina tentatively included. Limited genetic data suggested that specimens from Argentina and Brazil appear distinct from north Atlantic specimens (Casey et al., 2004; Lourie et al., 2004; Luzzatto et al., 2012). In addition, our findings of evolutionary lineage independence (Moritz, 2002) and morphological evidence further support the distribution of a second species named *H. patagonicus* (Piacentino & Luzzatto, 2004) and we treat the lineage south of the ARB as *H. patagonicus* throughout this study (Fig. 1.1b).

Individuals were sampled throughout the range of the *H. erectus* complex (390 samples; listed in Appendix 1.1). In addition, specimens of *H. capensis* (n = 2), *H. reidi* (n = 4), *H. zosterae* (n = 3), *H. guttulatus* (n = 3), *H. algiricus* (n = 2) and non-Atlantic *H.*

ingens (eastern Pacific, n = 4) and *H. kelloggi* (Indo-Pacific, n = 4) were included in species tree (*BEAST) estimates (Heled & Drummond, 2010). Sequence divergence between samples of the trans-Isthmian geminate species *H. ingens* and *H. reidi* were used to estimate nuclear DNA (nDNA) and mitochondrial DNA (mtDNA) mutation rates. These species are assumed to have separated during the final rise of the Isthmus of Panama (3.2 Ma) (Coates et al., 1992; Teske et al., 2007), with genetic divergence occurring earlier (Hickerson et al., 2006), thereby yielding a range of upper and lower bound temporal rate estimates. Data for all species were combined from direct sequencing and NCBI database specimens (Appendix 1.1). When possible, non-invasive sampling was conducted by removing a fin clip from the lower dorsal fin. In total, sequence data from 412 individuals were included in this study. Newly generated sequences were deposited in GenBank (Appendix 1.1).

Coalescent-based reconstruction of species trees and demographic history

Phylogenetic relationships among Atlantic seahorses were reconstructed in *BEAST 1.53 using a multispecies coalescent approach that co-estimates multiple gene trees embedded in a shared species tree, while accounting for the stochastic gene tree/species tree incongruence expected with genetic drift (Edwards, 2009; Heled & Drummond, 2010). Our aligned data set was 3840 bp from three partial mtDNA gene regions [cytochrome b (cytb): 696 bp; cytochrome c oxidase subunit I: 652 bp; and control region: 371 bp] as well as five nuclear loci (aldolase: 186 bp; myh6: 711 bp; rhodopsin: 438 bp; Tmo4c4: 464 bp; and a partial S7 intron: 322 bp). For details of PCR and sequencing methods, see Appendix 1.1.

*BEAST accounts for the independent coalescent histories of each sampled locus, with the mtDNA genes being treated as a single non-recombining locus. The number of gene copies per locus per population averaged 4.8, and each nDNA locus was assumed to be unlinked. All loci were partitioned using the closest models selected by the Akaike information criterion (AIC) in jModelTest 2 (Darriba et al., 2012). *BEAST was run using a Yule prior with a strict molecular clock (estimated across loci) consisting of two independent chains each of 10^8 iterations and parameters logged every 10,000 samples. After discarding the first 2 9 10^7 iterations as burn-in, a 50% majority clade tree was obtained in TreeAnnotator 1.6.1 and visualized in FigTree 1.3.1 (Rambaut, 2012). Convergence was assessed with effective sample size (ESS) values (all values > 200) in Tracer 1.5 (Rambaut & Drummond, 2007). In addition, the final log file of the 335 most probable species trees after burn-in was visualized in DensiTree 1.45 (Bouckaert, 2010) to illustrate the statistical uncertainty associated with our species-tree estimate (Fig. 1.3). For independent gene trees, see Appendix 1.2.

In order to assess the relative probability of each of the four historical hypotheses for the *H. erectus* complex outlined in the Introduction (Fig. 1.2), we used an approximate Bayesian computation (ABC) approach deployed in the program DIYABC (Cornuet et al., 2008, 2010). This approach compresses the molecular data into informative summary statistics that are then compared to simulated data under our modeled hypotheses. Following simulation, Euclidean distances between the observed and simulated data set were computed using a local linear or polychotomous regression, and the closest 2% of the simulated data to the observed data were retained to estimate the posterior distributions (Beaumont et al., 2002; Cornuet et al., 2008). This allows for the ranking of modeled hypotheses based on approximate marginal likelihoods that were

statistically compared using Bayes factors (Jeffreys, 1961) to classify the best-fit model. Following this step, Monte Carlo validation was used to demonstrate accuracy and precision in the model choice procedure (Robert et al., 2011).

Within the ABC framework, four alternative hypotheses based on different species tree topologies (Fig. 1.2) were jointly simulated with population genetic parameters. These include: effective population sizes, divergence times and the timing and magnitude of changes in historical population size. Using sample sizes identical to those of the observed data, we generated 1 million simulations under each modeled hypothesis. Based on intraspecific analyses across all sampled populations (Appendix 1.3), Caribbean/North American *H. erectus* and South American *H. patagonicus* samples were considered independent evolutionary panmictic populations (Moritz, 2002), whereas eastern Atlantic *H. hippocampus* samples were regionally subdivided and constrained to be sister populations (West African and Europe). In all four historical scenarios, we explored three possible three-taxon population topologies plus a polytomy, each involving the split of an ancestral population into the three regional lineages with a subsequent split of the north-eastern Atlantic population into West African and European populations (Fig. 1.2).

As a preliminary check, we evaluated whether our models together with the chosen prior distributions were able to generate a subset of statistics close to our observed summary statistics. This was carried out by locating the observed value of each summary statistic within a principal components analysis (PCA) plane of 50,000 simulated data sets (Cornuet et al., 2010) (Appendix 1.2). All summary statistics used in our final analysis showed a good fit across models, and we kept the following global summary statistics for downstream analyses: (1) number of segregating sites and (2) mean pairwise

differences, which are well established as being jointly informative about past growth rates in the populations (Beaumont, 2010). Standard measures of allelic diversity included (3) the number of private segregating sites, and (4) the number of distinct haplotypes (mtDNA only). For pairwise interpopulation summary statistics, we included the mean of between-sample pairwise differences and the F_{ST} value between two samples (Appendix 1.3).

From the mtDNA locus, 390 individuals from the *H. erectus* complex were included (partial cytb – 391 bp) in the ABC analysis (Fig. 1.1a and Appendix 1.1). At the five autosomal loci, all sequences were phased probabilistically using the program PHASE in DNASP5 (Librado & Rozas, 2009). PHASE was run for 10,000 iterations, retaining results with a probability of > 95%. All base calls were unambiguously recovered, and final data sets included both phased gene copies from each sequenced individual (*H. erectus*, n = 8–46; *H. patagonicus*, n = 6–18; *H. hippocampus* West Africa, n = 6; and *H. hippocampus* Europe, n = 6–36). Genetic data under each of the four hypotheses were simulated under the exact number of gene copies per locus per population sample, which includes independent observed data sets for haploid (mtDNA) and diploid (nDNA) loci (Appendix 1.3). For the final analysis, we used uniform priors with a lower and an upper bound for population size of 10 to 2×10^6, and divergence times of 1×10^3 to 4×10^6 generations in the past for t1 and t2, and 1×10^3 to 7.25×10^6 generations in the past for t3. Generation time was assumed to be one year for all species, and all populations were allowed to vary to include a discrete size-change event. The mutation rate of each nuclear locus was drawn from a gamma distribution with 95% quantiles encompassing 1.6×10^9 to 1.6×10^{10} (in units of mutations per site per generation/lineage), while the mtDNA mutation rate was drawn from a uniform

distribution with 95% quantiles encompassing 5.5 x 10^9 to 7.5 x 10^9 (in units of mutations per site per generation/ lineage). Because there is great uncertainty in assumed mutation rates (Ho et al., 2011a, 2011b), we emphasize relative time estimates.

We computed the posterior probability of each hypothesis by performing a weighted polychotomous logistic regression on 2% of simulated data sets closest to the observed data (Cornuet et al., 2008, 2010). Subsequently, we used a local linear regression on 1% of the accepted closest simulated data sets conditional on the most likely hypothesis to estimate the posterior distributions of parameters. Due to the potential biases arising when using Bayes factors to conduct ABC model discrimination among non-nested models (Robert et al., 2011), we evaluated the power and accuracy of our ABC model selection using simulation validation techniques (Cornuet et al., 2008, 2010). Specifically, we simulated 100 test data sets (i.e. PODS; pseudo-observed data sets) under each of the four competing hypotheses and calculated the probability of type I and type II errors using the criteria of Cornuet et al. (2010), with conditions (i.e. tolerance, summary statistics and number of random prior iterations) and sample sizes identical to our empirically-based estimates and hypothesis choice.

RESULTS

Phylogenetic relationships of Atlantic seahorses

The *BEAST species tree recovered two well-supported clades (C1 and C2) (Fig. 1.3; posterior probability = 1.0), both of which are associated with trans-Atlantic biogeographical distributions, and bifurcations appear to follow major directional ocean currents (Fig. 1.1b) from basal to terminal species throughout Atlantic seahorse

diversification (Fig. 1.3). The focal lineages in this study, *H. erectus*, *H. hippocampus* and *H. patagonicus* (C2), form a monophyletic group (posterior probability 1.0) (Fig. 1.3), clearly placing *H. patagonicus* within the *H. erectus* complex, with a within-clade posterior probability of 0.823. Atlantic *Hippocampus* is most likely to have originated from Indo-Pacific lineages roughly 15 Ma, coincident with an ancient split at the closure of the Tethyan Seaway, followed by trans-Atlantic expansion (Teske et al., 2007). The oldest bifurcation between the north-eastern Atlantic *H. guttulatus* (Europe) and the north-western *H. zosterae* (Florida, USA) lacks support (posterior probability 0.303) and additional loci may be required to resolve this divergence.

Consistent with Teske et al. (2007), C1 contains a trans-Atlantic sister-species relationship between *H. algiricus* and *H. reidi* across the equatorial mid-Atlantic, with a basal placement of *H. capensis* (posterior probability 1.0) and is linked to an ancestral Indo-Pacific lineage (Fig. 1.3) (Teske et al., 2007). This topology is consistent with a 'colonization pathway' of dispersal around the southern tip of Africa that has been found for other species of Atlantic fishes with ancestral origins in the Indian Ocean (Bowen et al., 2006). C1 also includes the trans-isthmian geminate species *H. ingens* that is believed to have separated from *H. reidi* after the final rise of the Isthmus of Panama (Teske et al., 2007).

Demographic history inference and species tree concordance of the *Hippocampus erectus* complex

Both coalescent-based methods suggest the western Atlantic origin hypothesis (H_{WAO}; Fig. 1.2) for the *H. erectus* complex, in preference to the three other hypotheses. Posterior probability of support for H_{WAO} was 0.626 using DIYABC (model probabilities:

H_{EAO}, 0.170; H_{CLK}, 0.055; H_{META}, 0.149), with corresponding Bayes factors ranging from 3.68 to 11.38 giving moderate to strong support for H_{WAO} over the three hypotheses (Appendix 1.2). Additionally, model choice validation using PODS indicate that adequate power exists for selecting the true hypothesis among competing hypotheses (Cornuet et al., 2010) with an acceptable type I error rate (28%) and low type II error rate (8%).

Conditional on H_{WAO}, the ancestral population of the *H. erectus* complex was first isolated in the western Atlantic, supporting a divergence between South American populations (*H. patagonicus*) and Caribbean/North American populations (H. erectus) approximately 5.27 Ma [95% credibility intervals (CI): 2.74–7.25 Ma] (Table 1.1, Fig. 1.4). Subsequent to this isolation event, the European and West African populations (*H. hippocampus*) became isolated from Caribbean/ North American populations around 3.35 Ma (95% CI: 1.98–5.73 Ma). This second isolation event probably involved a substantial reduction in population size with the founding of the eastern Atlantic region, the effective population size (N_e) being 54,300 individuals (95% CI: 38,000–1,610,000) before the split into West African and European populations 484 ka (95% CI: 268 ka to 1.39 Ma). Following this split, subsequent population expansions occurred in Europe with an N_e of 1.7 million individuals (95% CI: 1,470,000–1,940,000) and a West African N_e of 829,000 individuals (95% CI: 404,000–1,680,000). In both eastern Atlantic populations, the posterior probabilities support the timings of major expansions to be recent and substantial. In contrast, the N_e of the Caribbean/North American population (*H. erectus*) of 1.92 million individuals (95% CI: 1,730,000–1,980,000) is inferred to be more stable throughout the various spitting events with some expansion. In addition, the N_e of the Caribbean/North American population (*H. erectus*) is roughly 10 times larger than the

South American population (*H. patagonicus*) of 191,000 individuals (95% CI: 86,000–691,000).

DISCUSSION

We detected two features that drove diversification in the Atlantic *H. erectus* complex: (1) a vicariant breakup of an ancestral population into relatively stable Caribbean/North American and South American lineages; and (2) a founder colonization of the north-eastern Atlantic from Caribbean/ North America via the Gulf Stream. The timing of the first event (5.27 Ma; 95% CI: 2.74–7.25 Ma) coincided with a sustained period of increasing Amazon River sediment deposition approximately 2.6–6.4 Ma (Figueiredo et al., 2009) and geological estimates of maximum freshwater outflow (Rocha, 2003) that isolated the regions of the Caribbean and Brazil (Fig. 1.4). The second set of events includes the trans-Atlantic dispersal and subsequent trans-Atlantic isolation (3.35 Ma; 95% CI: 1.98–5.73 Ma) consistent with an acceleration and reorganization of circulation patterns and Gulf Stream currents starting roughly 4.6 Ma (Coates et al., 1992; Haug & Tiedemann, 1998), coincident with the rising of the Isthmus of Panama (Coates et al., 1992; Muss et al., 2001; Teske et al., 2007; Lessios, 2008).

Allopatric divergence in the western Atlantic

All three species in the *H. erectus* complex appear to share similar widespread latitudinal distributions and tolerance across tropical to warm-temperate climate gradients, with limited genetic structuring within regions (Appendix 1.3). Our findings of relatively widespread gene flow within species (Caribbean to North America in *H. erectus*, and Brazil to Argentina in *H. patagonicus*) are in stark contrast to the ancient

divergence between South American and Caribbean/ North American populations (5.27 Ma; 95% CI: 2.74–7.25 Ma) coincident with the outflow of the Amazon River Barrier (ARB). The ARB accounts for over 2300 km of muddy coastline between the Amazon (Brazil) and Orinoco (Venezuela) rivers. The dramatic shift in substrate and oceanographical conditions coincides with the boundary between the Brazilian and the Caribbean biogeographical provinces (Collette & Rutzler, 1977; Rocha et al., 2005b). The divergence between *H. erectus* and *H. patagonicus* is a common distribution pattern shared with many western Atlantic taxa, including numerous Brazilian fishes isolated from their Caribbean sister species (Joyeux et al., 2001; Rocha, 2003; Briggs & Bowen, 2012). Although the soft sponge habitat below the ARB freshwater surface plume allows for connectivity in some fish species (Rocha et al., 2002), the divergence we find in seahorses is consistent with their absence in marine surveys, suggesting that syngnathids may not utilize this passage to maintain trans-barrier connectivity (Collette & Rutzler, 1977).

Geological estimates of the age of the ARB range from roughly 11.8 to 1.6 Ma (Rocha, 2003), with the earliest proposed date representing the initial presence of a large late Miocene submarine fan due to the Andean uplift (Hoorn et al., 1995). However, recent evidence supports the idea that the ARB reached its present form in the late Pliocene, with three clear stages of advancement: an initial stage of low sedimentation rates from 11.8 to 6.8 Ma; an increase in sedimentation from 6.8 to 2.4 Ma; and very high sedimentation rates from 2.4 Ma to the present (Figueiredo et al., 2009). Our divergence time estimates most closely align with the increase in ARB sediment deposition occurring 2.6 to 6.4 Ma, but it is likely that multiple determinants underlie the isolation of these lineages. These include the rise and fall of sea levels (Robertson et al., 2006), and fast-

flowing ocean currents that may further prevent rafting organisms from crossing the ARB. As westward-flowing equatorial currents meet eastern Brazil, the currents split into the north-westward-flowing Caribbean current and the southward-flowing Brazilian current (Fig. 1.1b). For seahorses, which are incapable of active long-distance migration, and are presumed to passively migrate by rafting, ocean current directionality may be particularly influential because Brazilian individuals would have to travel against the current before migrating westwards to bypass the ARB.

Founder colonization of the north-eastern Atlantic

Teske et al. (2007), provided strong phylogenetic evidence that rejected the hypothesis of trans-Atlantic vicariance from continental breakup 84 Ma for Atlantic species of *Hippocampus*, and instead proposed that trans-Atlantic sister species were the result of long-distance founder dispersal, following the final rise of the Isthmus of Panama. The closing of this land bridge isolated the western Atlantic from the eastern Pacific, contributing to a profound reorganization of ocean circulation and stability starting around 4.6 Ma (Haug & Tiedemann, 1998). Since closure, the North Atlantic Gyre's relatively stable currents have predominantly moved in an asymmetrical clockwise pattern (Fig. 1.1b), but with strong evidence of dynamic accelerations and decelerations, and latitudinal shifts that coincide with Pleistocene glacial and interglacial cycles (Kaneps, 1979; Keffer et al., 1988; Haug & Tiedemann, 1998). Accelerations could facilitate long-distance dispersal, whereas decelerations and latitude shifts could have isolated migrants at a time period consistent with our findings of trans-Atlantic divergence between *H. erectus* and *H. hippocampus* populations (3.35 Ma; 95% CI: 1.98–5.73 Ma).

Counter to the typical east–west pattern of post Pleistocene colonization found in temperate North Atlantic invertebrates (Wares et al., 2001; Ilves et al., 2010; Waltari & Hickerson, 2012), Briggs (1995) suggested that most trans-Atlantic species originated from west-to-east dispersal-driven diversification, although evidence supports the western Atlantic as an area of both origin and accumulation of biodiversity (Floeter et al., 2007; Rocha et al., 2008). For the *H. erectus* complex, whose distribution is associated with tropical to warm-temperate waters, the circularity of the North Atlantic Gyre system could have allowed dispersal to occur both westwards and eastwards across the Atlantic, presenting a possibility for multi-directional exchange that complicates the inference of ancestral source populations. However, we find the highest probability of support for western Atlantic origins (H_{WAO}; Figs 1.2 & 1.4). Under this scenario, historical population size estimates suggest an eastward founder colonization of the eastern Atlantic *H. hippocampus*, which expanded into present-day European and West African populations subsequent to isolation from the more demographically stable western Atlantic *H. erectus* populations. Our estimates of a much smaller founder population in the eastern Atlantic supports typical patterns of source-to-recipient dispersal dynamics, and although a large amount of inherent uncertainty exists when using ancestral area inference with phylogenetic results, our historical demographic parameter estimates add additional support for eastward expansion and founder colonization.

An alternative interpretation may also explain the inferred patterns of dispersal under the western Atlantic origin hypothesis. Rather than suggesting ARB vicariance followed by trans-Atlantic colonization of Europe/West Africa (Fig. 1.4), this history could have been characterized by widespread or intermittent ancestral connectivity (by way of circular beltlike dispersal; Fig. 1.1b), followed by isolation of eastern Atlantic

populations and a recovery from an ancestral population crash. Although this history is consistent with a population polytomy (H_{META}), it would also be consistent with H_{WAO} if isolation events were temporally staggered.

Regardless of which nuanced interpretation is correct, the fast-flowing Gulf Stream is likely to have been a major force in the diversification of the Atlantic *H. erectus* complex, and acts as a transporter of both early life history stages (i.e. larvae and young) and abundant aggregations of marine organisms associated with *Sargassum* rafts (Casazza & Ross, 2008) that can survive for generations as free-floating habitats (Theil & Gutow, 2005a, 2005b). The Gulf Stream has been proposed as the primary source of the ichthyofauna of Bermuda (Smith-Vaniz et al., 1999), the Caribbean mesopelagic fauna in the northern Sargasso Sea (Jahn, 1976) and some western Atlantic groups in the eastern Atlantic Azores archipelago (Avila et al., 2009). In addition, the Gulf Stream shapes the contemporary dispersal of numerous species, including sea turtles (Blumenthal et al., 2009) and American and European eels (Kleckner & McCleave, 1982), and may be responsible for the dispersal of the western Atlantic *H. erectus*, which was recently found in the Azores (Woodall et al., 2009).

Regarding European and West African populations, our findings indicate that these H. hippocampus populations diverged around 484 ka (95% CI: 268 ka to 1.39 Ma). Estimates of N_e of these descendant populations suggest that both contemporary populations are large and have undergone significant expansions since the original founder colonization (Fig. 1.4, Table 1.1). These findings are consistent with Woodall et al. (2011), who concluded that these expansions were probably affected by interglacial cycles in the late Pleistocene (190 to 21 ka). In addition, the Cape Verde frontal zone, where the southward-flowing Canary Current begins to shift westwards, has been

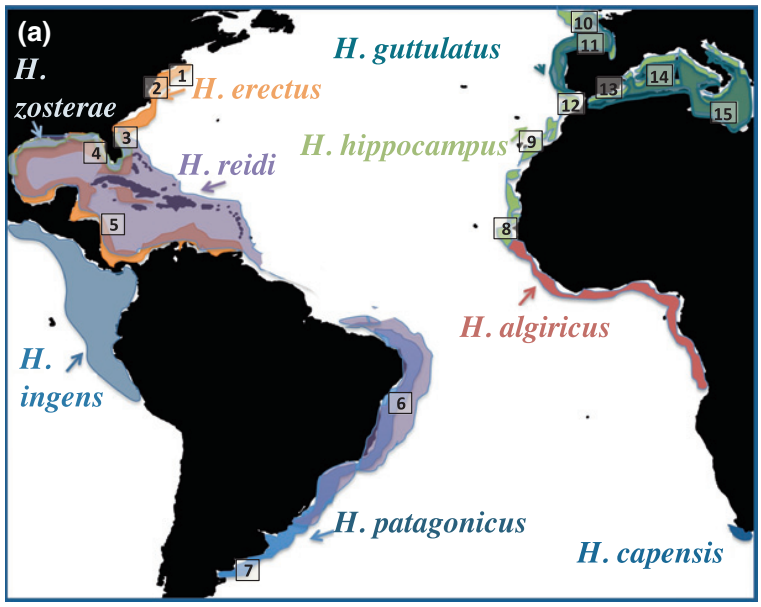
observed as a barrier to gene flow in several eastern Atlantic species, and may have contributed to the observed divergence (Vangriesheim et al., 2003; Woodall et al., 2011). Our estimates of trans-Atlantic and European/West African divergence times will be biased downwards if gene flow occurred after more ancient isolation events, a condition that is common in trans-Pacific ichthyofauna (Lessios & Robertson, 2006). Indeed, given the strong possibility that the Gulf Stream currents involved cyclical acceleration and deceleration, a more complex history involving demographic booms and busts and episodic periods of high gene flow should be considered an alternative hypothesis to be tested with population genomic-scale sampling. However, this sort of complex scenario is still consistent with isolation after trans-Atlantic dispersal.

Geographical patterns of Atlantic seahorse diversity

As the first reported study to include population-level sampling throughout the range of *H. erectus* and *H. patagonicus*, our reconstruction of species trees shows a clear monophyletic placement of *H. patagonicus* within the *H. erectus* complex and lineage independence from *H. erectus* (Figs 1.3 & 1.4). These results support the range restriction of *H. erectus* to an area north of the Orinoco outflow, and indicate that the species description of *H. patagonicus* (Argentina) (Piacentino & Luzzatto, 2004) may need to be extended to Brazilian populations.

Although both species have been shown to be under considerable fishing pressure (Dias et al., 2002; Vincent et al., 2011a), our estimates of long-term effective population sizes indicate that H. erectus was historically around 10 times more numerous than the South American species *H. patagonicus*. Currently, millions of seahorses are traded globally each year. Being Brazil's sixth most economically important ornamental export

(Vincent et al., 2011b), the demographic and evolutionary independence of *H. patagonicus* should qualify it as a unique genetic unit in need of further study and protection (Moritz, 2002). In addition, our results also support those of Woodall et al. (2011), who concluded that the West African population of H. hippocampus is genetically distinct from the European population. Like *H. patagonicus*, the West African population of *H. hippocampus* should also be considered a priority for conservation.



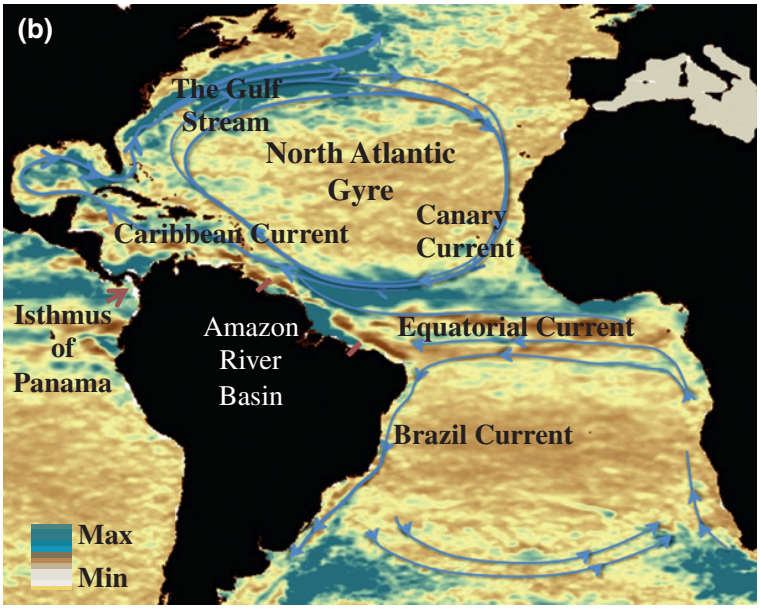


Figure 1.1 Atlantic seahorse distributions and major ocean currents. (a) Distributions modified from Teske et al. (2007) to include *Hippocampus patagonicus*. Small boxes 1–15 represent the collection locations of the *Hippocampus erectus* complex used in this study: (1) Hudson River Estuary and southern Long Island, New York (n = 31), Tuckerton and Little Bay, New Jersey (n = 13); (2) Chesapeake Bay (n = 22); (3) Atlantic Coast (Indian River and Jacksonville, Florida; n = 10); (4) Gulf Coast (Tampa Bay, south-east peninsula, offshore trawls; n = 28);(5) Caribbean Sea (n = 10); (6) Brazil (n = 7); (7) Argentina (Argentine Sea, San Antonio Bay and Mar de Plata; n = 14);(8) West Africa (Senegal; n = 26); (9) Canary and Porto Santo Islands (n = 44); (10) UK and Channel Islands (n = 39);(11) Bay of Biscay (n = 32); (12) South Iberia (n = 31);(13) southern coast, Spain (n = 18); (14) western Mediterranean (n = 44); and (15) eastern Mediterranean (n = 44). (b) Major ocean currents of North and South Atlantic gyre systems. The ocean environmental data layer is a five-year mean (1993–1997) of absolute velocity maximum (AVISO Remote Satellite data) transformed using Marine Geospatial Ecology Tools (MGET) 0.7 (Roberts et al., 2010) and visualized in ArcGIS 9.3 (ESRI, Redlands, CA). Blue colors represent current maxima, with arrows showing direction.

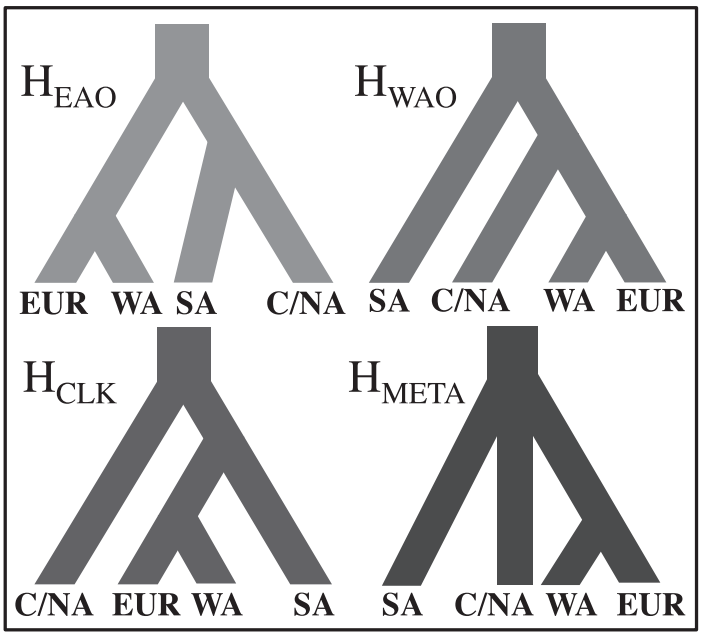


Figure 1.2 Four possible hypotheses explaining the branching history of the *Hippocampus erectus* complex: H_{EAO}, eastern Atlantic origin hypothesis; H_{WAO}, western Atlantic origin hypothesis; H_{CLK}, clockwise colonization hypothesis; H_{META}, metapopulation divergence hypothesis. C/NA, Caribbean/North America; SA, South America; WA, West Africa; EUR, Europe.

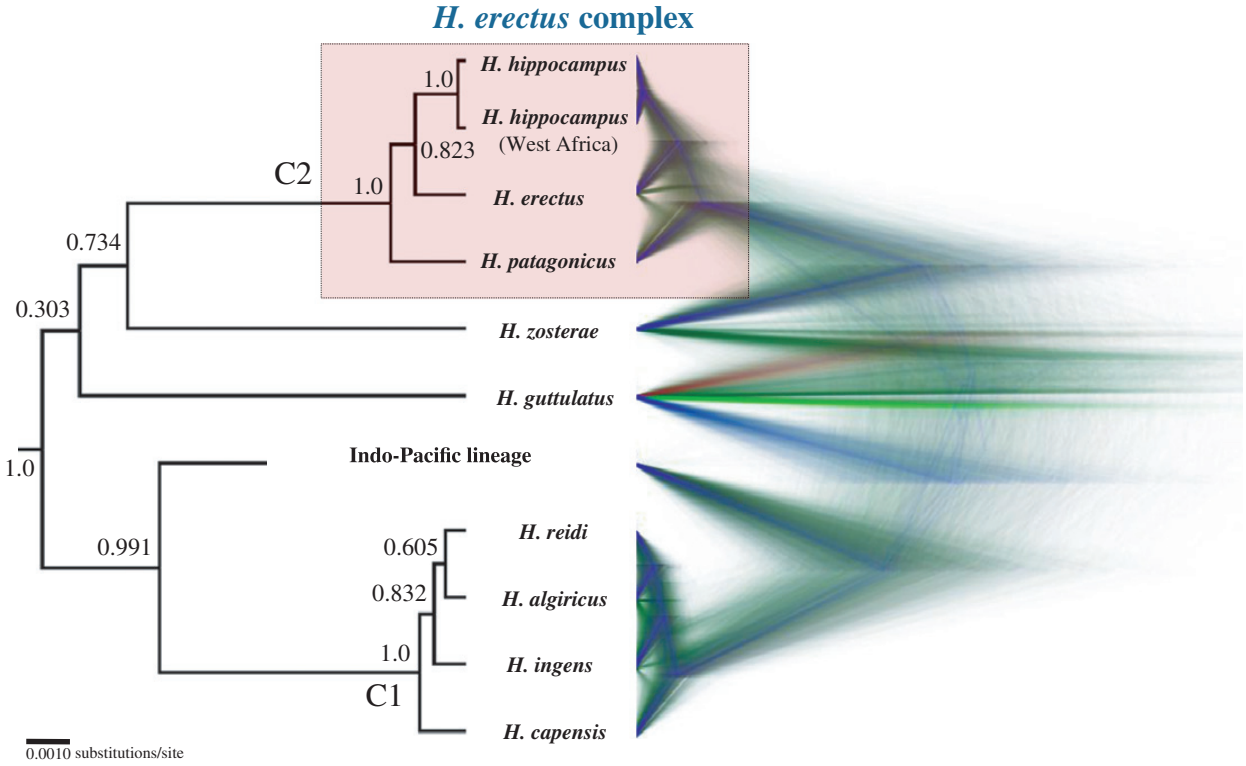


Figure 1.3 Species tree of the *Hippocampus erectus* complex and other Atlantic seahorses. Left: 50% majority-rule tree representing the probability of a given gene tree evolving within the branches of a species tree, generated in *BEAST. Right: 335 most probable topologies visualized in DensiTree v.1.45 to illustrate the statistical uncertainty of our species tree estimation. Greater topological agreement is visualized by a higher density of trees, whereas uncertainty in the height and distribution of nodes are represented by increased transparency.

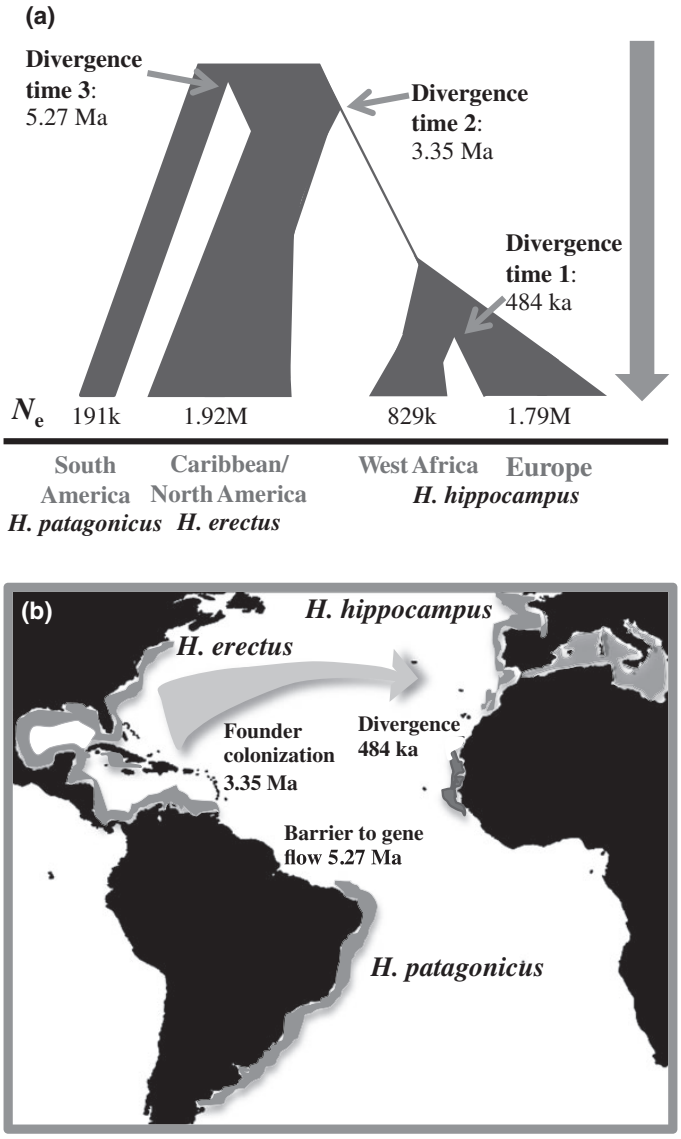


Figure 1.4 Most probable hypothesis (H_{WAO}) representing the demographic history of the *Hippocampus erectus* complex. (a) Tree illustration reconstructed from parameter values based on the mode of the posterior probabilities estimated from 1 million simulated data sets. Numerical values of effective population size (N_e) correspond to thousands (k) or millions (M) of individuals per population. (b) Visual representation of demographic inference.

Table 1.1 Parameter estimates based on posterior probabilities of the western Atlantic origin hypothesis (H_{WAO}) for the *Hippocampus erectus* complex.

Parameter	Mode	Quantiles 5%–95%
Modern N_e		
C/NA	1.92×10^6	1.73×10^6 to 1.98×10^6
SA	1.91×10^5	8.62×10^4 to 6.91×10^5
EU	1.79×10^6	1.47×10^6 to 1.94×10^6
WA	8.29×10^5	4.04×10^5 to 1.68×10^6
Ancestral N_e at time of size change		
EUWAt1sc	NC	
EUt1sc	3.39×10^5	1.01×10^5 to 1.79×10^6
EUAnt2sc	5.43×10^4	3.78×10^4 to 1.61×10^6
SAsc	NC	
C/NAsc	8.18×10^5	2.10×10^5 to 1.88×10^6
Ancestral populations		
A2sc	1.39×10^6	1.53×10^5 to 1.90×10^6
A1sc	2.28×10^5	5.19×10^4 to 1.64×10^6
Divergence times		
t1	4.84×10^5	2.76×10^5 to 1.43×10^6
t2	3.35×10^6	1.98×10^6 to 5.73×10^6
t3	5.27×10^6	2.74×10^6 to 7.25×10^6
Time of size change		
t2NA	1.25×10^6	1.24×10^5 to 4.37×10^6
t1EU	5.88×10^4	2.65×10^4 to 1.07×10^6
t1EUWA	6.80×10^4	3.55×10^4 to 1.36×10^6
t3SA	NC	

Table 1.1 legend: C/NA, *H. erectus*; SA, *H. patagonicus*; EU, *H. hippocampus* (Europe); WA, *H. hippocampus* (West Africa); EUAnt2, *H. hippocampus* effective population size (N_e) after t2 divergence; EUWAt1, *H. hippocampus* (West Africa) N_e after t1 divergence; EU t1, *H. hippocampus* (Europe) N_e after t1 divergence; SAsc, *H. patagonicus* N_e after t3 divergence; C/NAsc, *H. erectus* N_e after t2 divergence; A2sc, ancestral N_e of *H. erectus* and *H. hippocampus*; A1sc, ancestral N_e to all lineages. NC, posterior probability curves were non-informative (uniform).

CHAPTER 2

Population genomics reveals seahorses (*Hippocampus erectus*) of the western mid-Atlantic coast to be residents rather than vagrants

ABSTRACT

Understanding population structure and areas of demographic persistence and transients is critical for effective species management. However, direct observational evidence to address the geographic scale and delineation of ephemeral or persistent populations for many marine fishes is limited. The Lined seahorse (*Hippocampus erectus*) can be commonly found in three western Atlantic zoogeographic provinces, though inhabitants of the temperate northern Virginia Province are often considered tropical vagrants that only arrive during warm seasons from the southern provinces and perish as temperatures decline. Although genetics can locate regions of historical population persistence and isolation, previous evidence of Virginia Province persistence is only provisional due to limited genetic sampling (i.e., mitochondrial DNA and five nuclear loci). To test alternative hypotheses of historical persistence versus the ephemerality of a northern Virginia Province population we used a RADseq generated dataset consisting of 11,708 single nucleotide polymorphisms (SNP) sampled from individuals collected from the eastern Gulf of Mexico to Long Island, NY. Concordant results from genomic analyses all infer three genetically divergent subpopulations, and strongly support Virginia Province inhabitants as a genetically diverged and a historically persistent ancestral gene pool. These results suggest that individuals that emerge in coastal areas during the warm season can be considered "local" and supports offshore migration during the colder months. This research demonstrates how a large numbers of

genes sampled across a geographical range can capture the diversity of coalescent histories (across loci) while inferring population history. Moreover, these results clearly demonstrate the utility of population genomic data to infer peripheral subpopulation persistence in difficult-to-observe species.

Keywords: seahorse, marine fish, RADseq, western Atlantic, population persistence, population genomics

INTRODUCTION

In warmer seasons, the waters lining the concrete bulkheads, wooden piers, estuaries, and sandy beaches of the temperate Northeastern United State's mid-Atlantic coast become home to numerous tropical fish species (Curran, 1989; Briggs & Waldman, 2002). Over a century of research has cataloged the immigration of tropical vagrants or "strays" to these coastal mid-Atlantic waters. The majority of these individuals arrive due to passive planktonic dispersal in summer months, transported by ocean currents that circle north off the warm water mass of the Gulf Stream as it deflects northeast from U.S. towards Europe at roughly 35º latitude (Milstein & Thomas, 1976; Teixeira & Musick, 2001). This phenomenon positions Cape Hatteras as a delineation point between the zoogeographic Virginia and Carolina Provinces, each defined by distinct faunal endemism and unique macroclimatic conditions (Figure 2.1) (Briggs & Bowen, 2012; McCartney et al., 2013). Following this observation, studies of species found in both provinces suggest that Cape Hatteras acts as a "barrier" where intraspecific gene flow is reduced between provinces or alternatively acts as a northern latitudinal limit during the winter for species without cold thermal tolerance (McCartney et al., 2013; McBride, 2014). In this latter case, more sedentary tropical species that passively drift into the temperate Virginia Province during warmer months locally perish after cold winter temperatures advance.

Though many fishes exhibit wide thermal tolerance, ascertaining the true range of marine species can be challenging due to factors that include patchy distributions, cyclical population sizes, and seasonal movement patterns (Grosberg & Cunningham, 201AD; McBride, 2014). One species often associated with tropical vagrants in the Virginia Province is the Lined seahorse, *Hippocampus erectus* (Teixeira & Musick,

2001). Its status as a persistent subpopulation is uncertain primarily due to its nearshore absence during cold winter months and a scarcity of direct winter observations of individuals. *H. erectus* is commonly found in coastal zones in three zoogeographic provinces: Caribbean (tropical), Carolina (warm-temperate) and Virginia (temperate) (Figure 2.1). Some researchers suggest that long-distance rafting carries migrants northward to temporarily inhabit the Virginia Province as temperatures warm (Teixeira & Musick, 2001), a prediction supported by substantial observational evidence of long-distance rafting migration throughout its range (Fish, M.P. and Mowbray, 1970; Casazza & Ross, 2008). In contrast, other researchers suggest that localized active dispersal directed toward offshore migration for thermal refuge in continental shelf waters during late fall accounts for its winter absence (Able & Fahay, 1998). This hypothesis of seasonal localized migration is partially supported by the observation of inshore colonization of *H. erectus* as temperatures warm in April to June, characteristic of most temperately adapted fishes (Teixeira & Musick, 2001), and earlier than the July to September arrival typical for the majority of tropical strays (Briggs & Waldman, 2002; Howell & Auster, 2012).

Ecologists and evolutionary biologists often focus on questions at different temporal scales, but both fields are increasingly making use of genetic data to test hypotheses about population history, estimate the movement of individuals between local populations, and characterize the spatial distribution of genetic variation for effective species management (Grosberg & Cunningham, 201AD; Lowe & Allendorf, 2010; Hare et al., 2011). One example is the use of genetic data to examine source-sink dynamics (Pringle et al., ; Martinez-Solano & Gonzalez, 2008). True sink populations, even if annually persistent, require continual immigration from source populations and are

expected to exhibit genetic homogeneity with source populations or heterogeneity reflecting multiple sources of immigrants, while over time independent breeding subpopulations through random (genetic drift) or deterministic (natural selection) processes will exhibit distinct genetic divergence (Lowe & Allendorf, 2010). A number of studies have examined the biogeography and genetic divergence of *Hippocampus* species (Mobley et al., 2011). Most of this research has focused on Indo-Pacific species with genetic variation ranging from among localized South African estuaries (Teske et al., 2003), to widespread species complexes associated with rafting driven colonization (Teske et al., 2005), and differing levels of intraspecific divergence attributed to both ecological traits and biogeographic divides (Lourie & Vincent, 2004).

Here we test whether the presence of *H. erectus* individuals north of Cape Hatteras are the result of an ephemeral deme that is seasonally replenished from demographically persistent southern populations (H1; Hypothesis 1), or in contrast, there are persistent and isolated populations on either side of Cape Hatteras (H2; Hypothesis 2). A previous study of the *H. erectus* complex utilized mitochondrial DNA and five more slowly evolving nuclear loci across many individuals (N=129), yet rejected H2 in favor of H1, with little divergence and evidence of isolation across Cape Hatteras (Boehm et al., 2013). Now, with the decreasing cost of high-throughput sequencing, data can be sampled from across the autosomal genome to account for variations in mutation, coalescent history, and recombination, thereby facilitating a view of the complexity of a species evolutionary history with the potential to infer more recent divergence and/or populations differentiating in the presence of gene flow (Sousa & Hey, 2013).

To date, genome wide single nucleotide polymorphisms (SNP) datasets generated by restriction-site-associated DNA sequencing (i.e., RADseq) have been utilized to study

several fish species. Examples utilizing RADseq datasets include the support of cryptic differentiation between populations of the Baltic Sea herring (*Clupea herangus*) (Corander et al., 2013), the detection of hybrid individuals between trout species (Hohenlohe et al., 2011), genetic divergence of various sticklebacks populations (Hohenlohe et al., 2010; Catchen et al., 2013; Deagle et al., 2013), and robust phylogenetic resolution between African cichlid species (Wagner et al., 2013). To test the aforementioned competing hypotheses H1 and H2, we generated a genomic RADseq dataset consisting of 11,708 SNPs across individuals of *H. erectus* from the eastern Gulf of Mexico to Long Island, NY (Figure 2.1).

Although we base our inference from only 4-9 individuals per each of the three zoogeographic provinces (total individuals; N=23), data from large numbers of unlinked loci allow highly resolved inference even with few individuals (Felsenstein, 2006; Li & Jakobsson, 2012; Robinson et al., 2014). Moreover, given that outbred diploid genomes are comprised of recombining segments of DNA inherited from large pools of ancestors (Gronau et al., 2011), genome-level datasets should capture the diversity of coalescent histories (across loci) that reflects population history, such that information comes more from the number of loci sampled through the genome than from numbers of individuals per sampling locality (Lohse et al., 2011; Hearn et al., 2014).

MATERIALS AND METHODS

Sampling and bioinformatics

Samples of *H. erectus* ranged from the eastern Gulf of Mexico to New York State (N=23). Samples were collected from 2009-2013 from the following locations:

Apalachicola, FL, Tampa Bay, FL, Charlotte Harbor, FL, the Florida Keys, Jacksonville, FL, and Indian River Lagoon, FL, Chesapeake Bay, New Jersey, the Hudson River and Long Island, NY. The specimens collected in this study were carried out in accordance and approval of the Queens College Institutional Animal Care and Use Committee (IACUC) (Permit # 137), which approved all aspects of specimen use in this study. Domestic fishing of *Hippocampus* is neither under direct regulation within the United States nor under species protection and no specific permissions were required for these locations; however we collaborated with the following authorities for samples, and if standard collection permits were required they were issued for each collection location. The Florida specimens used in our study were collected under the authority of the Florida Fish and Wildlife (FFW) as part of the FFW: Southeast Area Monitoring and Assessment Program. Samples from the Chesapeake Bay were collected in collaboration with the Virginia Institute of Marine Science (VIMS), which is authorized to collect any fishes necessary for research under the Code of Virginia. Lastly, samples collected in New Jersey and New York were collected in collaboration with Rutgers University under the New Jersey Department of Environmental Protection and the New York State Department of Environmental Conservation (DEC) Special Licensing Unit, License No. 1638 with additional samples collected under DEC License No. 1405.

Sequenced samples were randomly chosen from a large number of individuals (N>100) over multiple collection years to ensure genomic similarity was not the result of non-independent relatedness. Total Genomic DNA was extracted using Puregene extraction (Qiagen) from tail muscle tissue and treated with RNAase A following standard protocols. Genomic DNA quality was checked on an agarose gel to ensure that the majority of DNA was >10,000bp and equalized to 30 ng/uL using Qubit Fluormetric

Quantitation (Invitrogen). Library construction and restriction site associated DNA (RAD) protocol followed (Baird et al., 2008; Lozier, 2014). Floragenex carried out library preparation and restriction associated digest (RAD) sequencing. Genomic DNA restriction digestion utilized the Sbfl enzyme and individual sequence adapters and barcode identifiers were ligated to genomic DNA prior to sequencing on the Illumina HiSeq platform. All sequences from cut sites resulted in single-end reads, which were demultiplexed and trimmed of adapters to 90bp fragment lengths.

Total reads per individual ranged from 1,264,862-4,736,299. The individual with the largest number of reads was processed to construct a de Novo pseudo reference genome, and reads for each individual were aligned using BOWTIE (Langmead et al., 2009). SAMTOOLS algorithms (Li et al., 2009) and custom Floragenex perl scripts were used to detect SNPs and call genotypes. SNP datasets were formatted in the variant call format (vcf) (Danecek et al., 2011). Initial genotyping required a minimum Phred quality score of 15, a minimum of 4x sequence coverage, with a minimum of 65% of individuals genotyped. Additional filtering was applied using R v.3 (R Development Core Team, 2004) to ensure a Phred score equal to a hard cutoff of q=20 (base call accuracy lower than 99%). To reduce the inclusion of false SNP discovery due to paralogous sequences or low quality genotype calls, vcftools was utilized to remove any sites with a minimum depth of 8x sequence coverage and maximum depth calculated in R based on the mean depth + 1.5 standards deviation (=295) across all sites. The final datasets resulted in a bi-allelic matrix of 11,708 genotypes (5777 90bp sequences) across individuals at all sites. For details on per individual raw reads, filtered, and analyzed reads see Table 2.1.

Population genomic analyses

A principal components analysis (PCA) was implemented to determine if sampled individuals reflect a history of differentiated populations by outputting individual coordinates along axes of genetic variation within a statistical framework (Patterson et al., 2006) that correspond to the first two principle components in Figure 2.2b. To further aide in assigning individuals to differentiated populations by inferring ancestry coefficients representing the proportions of each individual's genome that originated from a specified number of ancestral gene pools (K) we used the program sNMF (Frichot et al., 2014). The program sNMF estimates individual ancestry and population clustering by utilizing a sparse non-negative matrix factorization algorithm (sNMF) to compute least-squares estimates of ancestry coefficients. This software is capable of efficiently analyzing large bi-allelic datasets without loss of accuracy when compared with more commonly utilized programs STRUCTURE (Pritchard et al., 2000) and ADMIXTURE (Alexander et al., 2009) that use the same underlying model to infer ancestry coefficients. However, in contrast to the aforementioned programs, sNMF has significantly better computational efficiency and is robust to many of the demographic assumptions of Hardy-Weinberg and linkage equilibrium (Frichot et al., 2014; Harris et al., 2014). To verify the accuracy of this program Frichot et al. (2014) conducted an in-depth comparison with the software ADMIXTURE using simulated and empirical datasets and found concordant results across trials, while sNMF outperformed ADMIXTURE when population inbreeding (F_{IS}) was high. For our dataset ancestry coefficients (K) were estimated using sNMF to determine subpopulation membership by running 10 replicates of K 2-6 using a cross-entropy criterion (CEC). To evaluate the predictive capability and

error of the ancestry estimation algorithm, sNMF employs the CEC, which is comparable to the likelihood value implemented in the program ADMIXTURE. To select the best-supported ancestry coefficient, the lowest CEC value was represented by the K value (K=3). The ancestry coefficient plot (Figure 2.2c) was visualized using R v.3. For information on CEC values, as well as results obtained between sNMF and STRUCTURE on a subset of the total data (SNP=2000) see Appendix 2.1.

The program *Treemix* (Pickrell & Pritchard, 2012) was utilized to infer the phylogenetic relationships between sampled locations while accounting for ancestral admixture among populations. Specifically, *Treemix* incorporates a model to allow for population divergence in the presence of post-divergence admixture/migration (m) given that incorporation of this parameter can improve the likelihood fit of a bifurcating phylogeny. More specifically, the m parameter represents the proportion of admixture from one population to another (Gompert et al., 2014). The resulting phylogeny is based on a composite maximum likelihood of the local optimum tree, determined using a similar approach to (Felsenstein, 1981), with branch lengths proportional to the amount of genetic drift that has occurred per branch.

Population genetic statistics (Table 2.1; Figure 2.2d-2.2f) were generated using vcftools and calculated across all SNPs per individual. The calculation of F_{ST} utilized between subpopulations (Weir & Cockerhan, 1984) specifically accounts for differences in sample size and a small number of sampled individuals, and recent studies have shown that bi-allelic SNPs (>1000) using this approach will result in precise F_{ST} estimates (Willing et al., 2012).

To investigate the visual similarity between genetic and geographic distance from the PCA analysis (Fig. 2.2a and 2.2b), we conducted a test for isolation-by-distance

(IBD) to see if this pattern meets the expectation of genetic similarity decaying with geographic distance (Novembre et al., 2008) using the IBD program by Mantel's test (10,000 randomizations) of linearized FST (FST/(1-FST) and shoreline distance (km) (Jensen et al., 2005). Pairwise FST, calculated in vcftools, and distance of coastlines between sampling locations in kilometers was determined using Google Earth Tools. For this Mantel test, the centroid distance between sampling locations for the Gulf-Keys subpopulation was utilized and results indicated a non-significance correlation between geographic and genetic distance ($p = 0.4925$). See Appendix 2.1 for additional information and regression plots.

RESULTS AND DISCUSSION

Support for northern subpopulation divergence and isolation

Our results strongly support H2 over H1 with Virginia Province residents of *H. erectus* coming from a persistently breeding and isolated subpopulation, rejecting the categorization of it being composed of seasonal migrants. The sNMF-based estimates of ancestry coefficients support three distinct subpopulations with limited admixture (K=3) (Fig. 2.2c): 1) the Florida Keys-Gulf of Mexico (Gulf-Keys), 2) Indian River, Florida (South-Atlantic), and 3) Chesapeake-New York (North-Atlantic). The K=3 value reported in our study is considered robust as it exhibited the lowest CEC value across replicate runs of all K values (K=2-6). This substructure also visually emerges from the first two principle components of the PCA from the total amount of observed genomic variation (Fig. 2.2b). Here, the individuals from north of Cape Hatteras form a tight cluster, while individuals sampled from the Gulf-Keys and South-Atlantic form a cline

between the tightly clustered South-Atlantic individuals and an admixed set of Gulf-Keys individuals. Consistent with these results is the inferred population history that emerges from *Treemix*, which is concordant with long-term isolation of the North Atlantic sub-population with limited post-divergence admixture with southern subpopulations.

The elevated heterozygosity found in Gulf-Keys individuals (Fig. 2.3a) could be the result of admixture from un-sampled western Gulf/Caribbean individuals, which is also indicated from the sNMF analysis (Fig. 2.2c). However, this elevated heterozygosity could also be the result of a larger effective population size (Gazave et al., 2013). In contrast, the northern subpopulation shows a reduction in heterozygosity with an elevated level of singletons (Fig. 2.3b). This pattern indicates a possible demographic expansion after the last glacial maximum that is consistent with the likely unsuitable habitat in the Virginia Province during the late Pleistocene. This history of shifting habitat driven by climate change is suggested by palaeo-climatological research indicating that temperate environments north of Cape Hatteras were displaced southward (Cronin et al., 1981; McCartney et al., 2013), as well as the formation of the Chesapeake Bay 7.4-8.2 kya due to post-last glacial maximum sea level rise (Bratton et al., 2002).

Causes of divergence and isolation of the northern subpopulation

Given the strong evidence we report for Virginia Province inhabitants of *H. erectus* representing a persistently isolated independent subpopulation from other regional subpopulations, there are several conceivable non-mutually exclusive causes of this divergence. First, seagrass is a preferred breeding habitat of *H. erectus* and a long gap without coastal seagrasses exists along the Georgia and South Carolina coastlines

(roughly 600km) (Short et al., 2007). This barrier of unsuitable breeding habitat between northern Florida and the Virginia Province therefore likely results in the fish's rarity in this area, thereby increasing genetic isolation of the northern subpopulation (Wenner & Sedberry, 1989; Lourie et al., 2004). The confamilial pipefish *Syngnathus floridae* also shares a similar pattern of genetic divergence across this region of unsuitability, though the area of absence extends from the southern end of the Florida Peninsula to near Cape Hatteras, with the northern population extending from North Carolina to Chesapeake Bay (Mobley et al., 2010). Secondly, long-distance migration of *H. erectus* is observed to occur via *Sargassum* rafting driven by ocean currents (Casazza & Ross, 2008). Under this mode of migration, the northeastern deflection in ocean currents near Cape Hatteras toward the Mid-Atlantic may limit the arrival of southern migrants to the Virginia province. Lastly, individuals that do arrive from southern provinces may have a lower physiological tolerance to temperate conditions, reducing the chance of winter survival and also increasing the amount of genetic isolation. Selection correlated to the shift in macroclimate at Cape Hatteras has been observed in marine fishes (Hice et al., 2012), and future analysis of northern adaptation in *H. erectus* may help decouple the potential drivers of temperate subpopulation genetic isolation. Although our observed patterns of genetic isolation could have emerged via a continuous isolation-by-distance regime without clear breaks driving the isolation, a Mantel test resulted in a non-significant relationship between genomic and geographic distance (p=0.4925).

Support for local seasonal migration

Our results also support local offshore migration to account for the coastal absence of *H. erectus* from Virginia Province during winter months. While extreme temperature changes influence latitudinal movement of many species (Briggs & Waldman, 2002), substantial seasonal movement to and from provinces for *H. erectus* is unlikely due to its relatively weak swimming ability (Foster & Vincent, 2004). To avoid nearshore cold water temperatures, localized inshore-offshore migration has been reported for the confamilial pipefish (*Syngnathus fuscus*), which has similar life history traits to *H. erectus* (Lazzari & Able, 1990), and has also been suggested for some other species of *Hippocampus* (Foster & Vincent, 2004). As a qualitative comparison we examined abundance records from NOAA long-term offshore trawl surveys of *S. fuscus* (1972-2008; >90% 20 km off-coast; depth 10-20m) and found that they closely resemble that of *H. erectus,* further supporting intercontinental shelf overwintering (For additional details see Appendix 2.2). Regarding direct observation of this phenomena, a single record from divers in 1968 documented both species "hibernating" on the shelf substrates off Long Island, NY (Wicklund et al., 1968), where they resumed swimming several minutes after being brought to the surface. Many fish adapt to winter temperatures by decreasing energy demands and entering semi-torpidity (Ultsch, 1989), though no research has been conducted on the overwintering physiology of any Syngnathidae species. Nevertheless, localized overwintering in deeper waters may be an important component of *H. erectus*' life history and may also account for their winter absence in estuaries of the warm-temperate eastern Floridian Peninsula (i.e., Indian River Lagoon, FL; South-Atlantic).

Conclusions

Overall, our results demonstrate the utility of supplementing life history information with population genomic data when a small number of unlinked genetic loci may be insufficient to discern the range of persistence in difficult-to-observe fishes. Currently, the IUCN (World Conservation Union) Red List categorizes *H. erectus* as "vulnerable" based on it being commonly collected as by-catch and sold by trawl fishermen to supply the aquarium trade (Dias et al., 2002). Our results, throughout an extensive range of this species distribution, will help inform conservation, as well as captive breeding efforts, by strongly supporting northern Atlantic seahorses as a genetically distinct subpopulation. More broadly, because genomic data effectively samples a multitude of ancestors, even with a small number of sampled individuals, the approach taken in our study shows the promise of genomic data to infer population genetic structure in rare and/or difficult to obtain species.

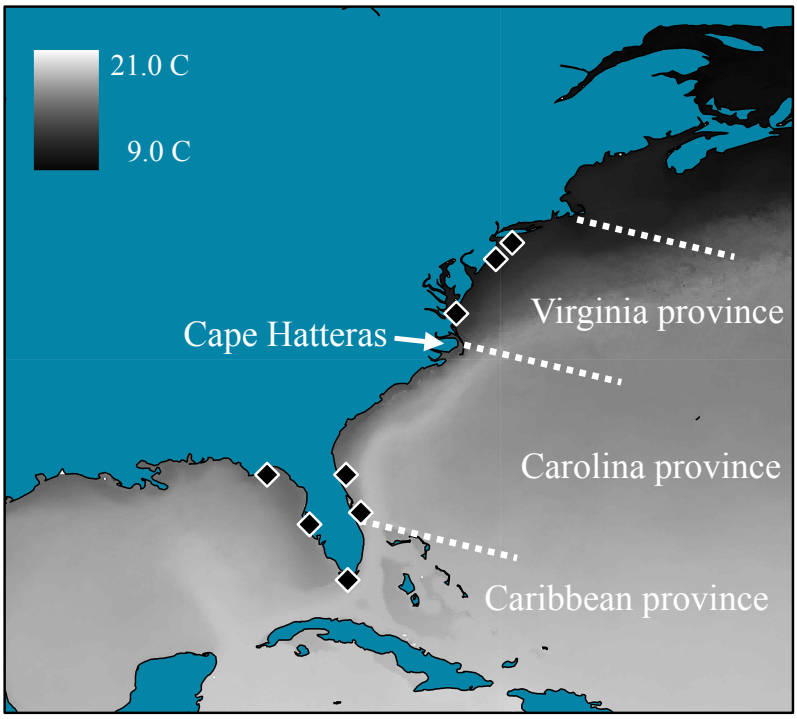


Figure 2.1. **Map of zoogeographic provinces, collection sites, and temperature variance.** Contrasting ocean minimum sea surface temperatures across zoogeographic provinces: generated in ARCGIS v.9.3 using the Bio-ORACLE long-term climatic dataset (Tyberghein et al., 2012). Collection sites from the northeastern Gulf of Mexico to New York State indicated by diamonds: Apalachicola, Tampa Bay-Charlotte Harbor, Florida Keys, Indian River Lagoon and Jacksonville, FL., Chesapeake Bay, New Jersey-New York.

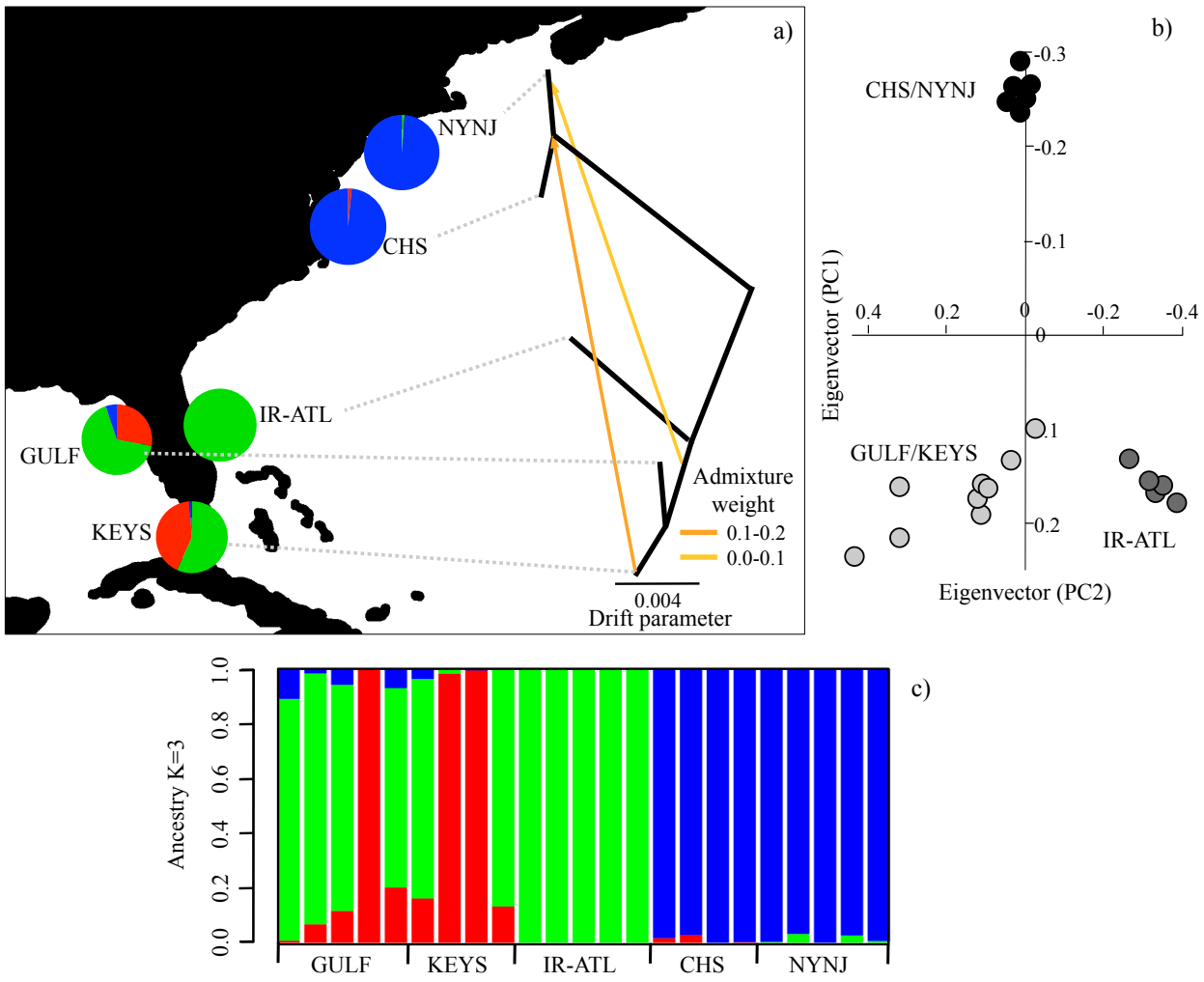


Figure 2.2. **Genomic variation across individuals and subpopulations.** (a) *Treemix* population tree with branch lengths scaled to the amount of genetic drift between regions and inferred proportion of historical ancestral genetic admixture (m=2) between southern and northern regions represented by arrows. Dotted lines do not represent branch length. (b) Principle component analysis. Black circles = Chesapeake Bay-New York, dark grey circles = Florida Atlantic coast, and light grey circles = Gulf of Mexico-Florida Keys. Pie diagrams (a) represent ancestry coefficient proportions derived from the sNMF ancestry plot (c). Each line of the sNMF plot represents one individual.

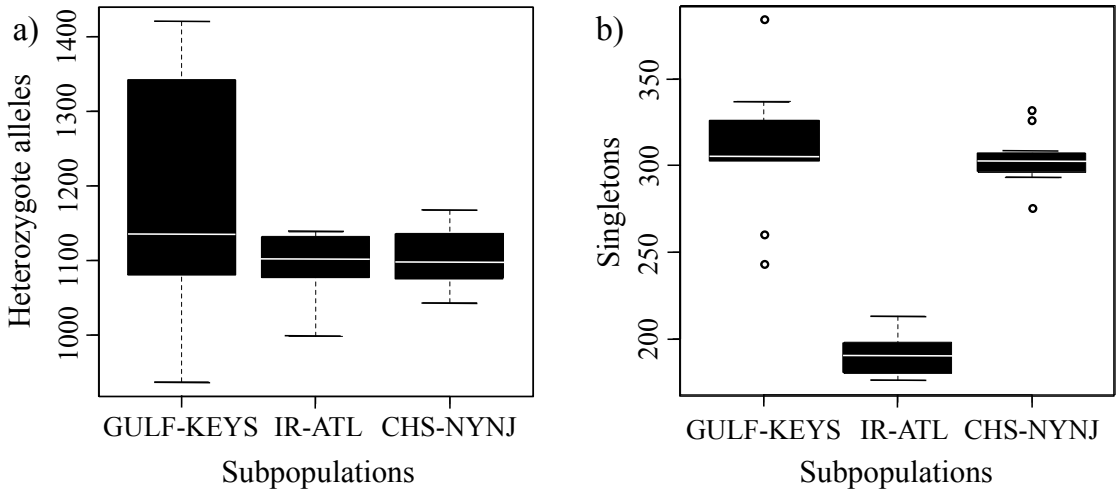


Figure 2.3. Distribution of heterozygote and singleton genotypes. Boxplots represent the range of observed heterozygote genotypes (a) and singleton genotypes per individual/per subpopulation (b).

Subpopulations	*n*	FST		SNPs	Mean He	Mean S
North-Atlantic	**9**	**North-Atlantic**	**South-Atlantic**	**11,708**	**1,104**	**304**
South-Atlantic	**5**	**0.1012**	**--------**	**11,708**	**1,090**	**191**
Gulf-Keys	**9**	**0.083**	**0.0454**	**11,708**	**1,183**	**309**

Table 2.1. **Population genomic summary statistics for each subpopulation.** FST, number of individuals (*n*), total number of SNPs, observed mean heterozygote alleles per individual/per subpopulation (He) and mean number of singletons (S) per individual/per subpopulation.

CHAPTER 3

Population genomics of codistributed pipefishes and seahorses using RAD sequencing: Ocean currents and life history predict gene flow asymmetry and population divergence

ABSTRACT

Powerful ocean currents and historical climatic fluctuations play a significant role in the diversification and population connectivity of coastal marine fishes. However, decoupling how these abiotic forces interact with life history traits in shaping demographic and evolutionary histories has remained challenging without comparative studies that sample at the population-genomic scale. One often over-looked trait is the ability of species to raft on floating vegetation, specifically *Sargassum spp*. that has been shown to be a significant determinate for successful long distance dispersal in seahorses (*Hippocampus*) and pipefish (*Syngnathus*), which are direct developers and are thought to lack long distance active migration. Rafting is passive, and the relative stability of major ocean currents over evolutionary timescales may cause historical gene flow to reflect ocean current directionality. To test hypotheses about how ecological and morphological traits affect gene flow magnitude and directionality after population colonization and/or divergence, we estimate these parameters using RADseq generated single nucleotide polymorphism (SNP) datasets from five codistributed species of western Atlantic Syngnathids. Using the site frequency spectrum (SFS) in a composite likelihood coalescent framework, these co-estimates indicate that gene flow directionality follows major ocean currents for both seahorse species and two of three pipefish species. Population divergence times and gene flow directionality and magnitude are also

consistent with inter-specific differences in ecological traits related to rafting propensity, environmental plasticity, and morphological traits being predictors of population history and regional patterns of connectivity.

Keywords: population genomics, RADseq, gene flow, ocean currents, fishes, seahorse and pipefish (Syngnathidae), *Sargassum,* rafting

INTRODUCTION

"*Inhabits all of the seas of the temperate and tropical regions. They are pelagic fishes which attach themselves to seaweed or other floating substances, and are liable to be carried by currents to great distances."* On the description of *Hipocampus* (seahorses), Gunther (1870).

Understanding the scale of dispersal and population connectivity requires accounting for multiple dispersal mechanisms (Kinlan & Gaines, 2003). In the ocean, paleoclimatic shifts and hydrodynamic forces play a significant role in the connectivity and genetic structure of coastal marine fauna (Briggs 1970; Avise 1992; Hewitt 1996; Avise & Walker 1998; Hare *et al*. 2005; Rocha *et al*. 2005). Yet, the impact of such abiotic forces is difficult to decouple from life history traits including zonation, habitat preference, swimming capacity, and passive migration propensity (Rocha et al., 2002, 2005b; McBride, 2014). Many marine fishes actively migrate, while most population genetic studies that have investigated passive dispersal have focused on early stages of pelagic planktonic larvae duration driven by ocean currents (Meyer et al., 2005; Weersing & Toonen, 2009; Crandall et al., 2012). However, results on the correlation between larval duration and genetic connectivity lacks consensus (but see, (Faurby & Barber, 2012), leading some studies to conclude that adult life history, ecological differences, and the transition between biogeographic regions can be as instrumental as oceanography in influencing population genetic structure (Rocha et al., 2007; Eytan & Hellberg, 2010; Riginos et al., 2011; Luiz et al., 2012).

Rafting with marine vegetation is an often-overlooked mode of dispersal for marine organisms, but for some species it can be linked to both ecological preference and oceanographic passive dispersal, as well as playing an important role in long-distance

founder colonization (de Queiroz, 2005; Fraser et al., 2011; Gillespie et al., 2012). Most research on rafting driven founder events focus on terrestrial island biogeography, although in the Atlantic Ocean several studies have concluded that the impact of current driven founder dispersal has lead to the colonization of oceanic islands by marine taxa (i.e., Bermuda and the Azores Archipelago) (Smith-Vaniz et al., 1999; Joyeux et al., 2001) as well as long-distance rafting dispersal via the Gulf Stream and Equatorial currents (Teske et al., 2007; Luiz et al., 2012; Boehm et al., 2013). In addition to long-distance colonization, many marine species have an ecological propensity to inhabitant and disperse with buoyant vegetative rafts of *Sargassum spp*. Unlike pelagic larvae, where connectivity or long-distance colonization may be driven by a small number of individuals that disperse beyond the mean of the most probable dispersal distance (Crandall et al., 2012), *Sargassum* rafts are positively buoyant, which may foster a higher magnitude of long-distance dispersal for inhabitants. Therefore, the distance of travel for coastal species that inhabit *Sargassum* can be driven by the directionality of ocean currents, while unpredictable events that move rafts beyond major oceanic currents to and from nearshore areas allow individuals to inhabit or vacate this habitat (Bortone et al., 1977).

While oceanic barriers to dispersal can play an important role in generating and maintaining patterns of marine diversity, very few “hard” allopatric marine barriers exist (Floeter et al., 2007; Rocha et al., 2007; Rocha & Bowen, 2008). The most notable exception in the western Atlantic is the final rise of the Isthmus of Panama that had a profound impact on the distribution and evolutionary trajectory of present day ichthyofauna (Briggs, 1995; Knowlton & Weigt, 1998; Lessios, 2008), which also had major impacts on oceanographic conditions and long-term directionality of major ocean

currents. For example, throughout much of the world, complex patterns of symmetrical currents are prominent (Treml et al., 2007; Ayre et al., 2009; Crandall et al., 2012). However, paleontological, biostratigraphic, and contemporary data support that tropical and warm-temperate North-Atlantic Belt currents have maintained a clockwise asymmetrical directionality since the rise of the Isthmus, (Kaneps, 1979; Keffer et al., 1988; Haug & Tiedemann, 1998). In the western Atlantic, these currents follow shoreline patterns in a predominately eastward direction, including the Gulf of Mexico's Loop current that transfers to the Florida current and into the Atlantic Gulf Stream (Briggs, 1974; Mullins et al., 1987; Lee et al., 1992; Haug & Tiedemann, 1998).

Due in part to the massive expanse of *Sargassum* making up the mid Atlantic's Sargasso Sea, the movement of *Sargassum* aggregations span over 5 million square kilometers of open waters in the Atlantic ocean (Laffoley et al., 2011). Because of the long-term directional stability of the North-Atlantic Gyre Belt that circles the Sargasso Sea, *Sargassum* dispersal is particularly pronounced. Based on Satellite imaging and empirical studies, contemporary dispersal of *Sargassum* aggregations are annually found in the western Gulf of Mexico (i.e., Texas coast) in the Spring (March-May) and subsequently follow an eastward current direction (Wells & Rooker, 2004; Gower et al., 2006; Gower & King, 2008, 2011b). While, many *Sargassum* aggregations remain within the Belt currents, winds and hydrodynamic eddies also cause rafts to move in and out of coastal zones (Bortone et al., 1977; Wells & Rooker, 2004; Tabone, 2011). This unpredictable movement can be probabilistically modeled using biophysical simulations to illustrate the tendency of rafts to reach a given shore or coastal area as well as movement within major currents (Figure 3.1). The more random weather driven movement allows fish larvae, juveniles, and adults the ability to occupy *Sargassum*

habitat. These inhabitants benefit from the availability of prey and shelter to increase the chance of arrival to new locations (Weis, 1968; Teske et al., 2005; Theil & Gutow, 2005b; Casazza & Ross, 2008; Gower & King, 2011). There are several studies on long-distance rafting driven dispersal (Hobday, 2000; Fraser et al., 2011; Luiz et al., 2012; Boehm et al., 2013) and *Sargassum* community composition (Weis, 1968; Fine, 1970; Wells & Rooker, 2004; Theil & Gutow, 2005a, 2005b), though to our knowledge the correlation between regional geographic population genetic structure, connectivity, and rafting dispersal within a model-based multi-species framework remains unexplored. While structure and regional isolation has been found with population genomic data in *Hippocampus erectus* (Boehm et al., *in press*), these results have not been used in the comparative framework we present here.

Depending on magnitude and direction, prevailing oceans currents are not always predictive of gene flow (Horne, 2014), and gene flow in some species will reflect anomalies (Crandall et al., 2012; Selkoe et al., 2014). Barriers may drive these anomalies to dispersal, which can be reinforced by paleoclimatic conditions that result in historical discontinuities in suitable habitat. While hard barriers such as the Isthmus are rare, permeable or "minor" barriers to dispersal associated with glacial cycles can impact habitat availability (Eytan & Hellberg, 2010), the magnitude of freshwater outflows (Floeter et al., 2007; Anderson et al., 2012, 2014; Portnoy & Gold, 2012), and lowering and rise of eustatic sea levels (Riginos et al., 2011; Ludt & Rocha, 2014). For example, during the last glacial maximum (LGM) it is estimated that lower sea levels reduced coastal shelf areas in the Caribbean basin by 89% from present day conditions (Bellwood & Wainwright, 2002; Eytan & Hellberg, 2010) and up to 92% in the Gulf of Mexico (Ludt & Rocha, 2014) thereby dramatically reducing the amount of available habitat as

sea levels approached the depth of continental shelf breaks. General observations from phylogeographic studies have determined that the lowering of sea levels may have strengthened the barrier of Florida Peninsula (Rocha et al 2007). This was first explored by Avise (1992) in a seminal study that demonstrated the real possibility of cryptic biodiversity in several marine species based on reciprocally monophyletic mitochondrial (mtDNA) lineages between the Gulf of Mexico and Atlantic Ocean populations. Other minor barriers in the Gulf of Mexico include cold freshwater pulses from the Mississippi River between 16 and 9 kya (Kennett & Shackleton, 1975; Aharon, 2003), and more recent Holocene changes in coastal formations and sediment deposits further impacting available habitat (Morton et al., 2000; Blum et al., 2001). For some species climatic events may strengthen permeable barriers, while species that can traverse "soft" barriers generally have extended dispersal ability or display reduced habitat specificity, and therefore lack population isolation throughout their range (Rocha et al., 2002; Marko, 2004; Lourie et al., 2005; Ayre et al., 2009).

Many seahorses (*Hippocampus*) and pipefish (*Syngnathus*) are presumed to be poor swimmers (Foster and Vincent 2005). Yet, in contrast to the findings of Avise (1992), the focal species in our study - *Syngnathus scovelli, S. floridae, S. louisianae, Hippocampus erectus* and *H. zosterae* - lack reciprocal monophyly between Atlantic and Gulf populations at the mtDNA locus (Appendix 3.2). In addition, several of these species display limited mtDNA genetic structure throughout their sampled distributions (i.e., within the Gulf of Mexico and along the western Atlantic coast). This pattern of limited population divergence is in contrast to many species with higher active swimming capacity that show more pronounced patterns mtDNA divergence (Anderson et al., 2012; Portnoy & Gold, 2012; McBride, 2014). These mtDNA-derived patterns might suggest

that our focal species successfully disperse across barriers and maintain range-wide population connectivity, and therefore have the capacity to transport over large distances either through limited planktonic durations or through rafting driven dispersal (Theil & Gutow, 2005a, 2005b). Alternatively, this could indicate that some areas experienced local extinction followed by recent re-recolonization that would erase patterns if accumulated mtDNA divergence (Cunningham & Collins, 1998). Syngnathids are direct developers with limited planktonic dispersal and commonly settle to benthic habitat shortly after birth, suggesting that rafting may be their primary means of transport. In support of this assumption, empirical studies of *Sargassum* communities have found that Syngnathids are among of the most abundant fish groups found in *Sargassum* habitat throughout the Gulf of Mexico and into the Atlantic (Fish and Mowbray, 1970; Bortone et al., 1977; Dawson & Vari, 1983; Wells & Rooker, 2004; Casazza & Ross, 2008).

Though, our five focal taxa have primarily overlapping distributions, they also possess ecological differences in macroclimatic tolerances and rafting propensity, such that patterns of genetic variation and demography may differ depending on species-specific ecological affinities (Dawson & Vari, 1983; Mobley et al., 2010). For example, previous comparative research on *Hippocampus* (seahorse) species concluded that ecological differences play an important role in determining population structure and response to historical climatic events (Lourie et al., 2005), and several other studies comparing marine invertebrate species also supported the role of ecology in shaping historical population structure (Ayre et al., 2009; Marko et al., 2010; Marko & Hart, 2011a; Crandall et al., 2012). Although there are several reasons to expect concordant ocean current driven movement among our focal species, species-specific life history traits as well as morphological differences may be non-mutually exclusive causes of

discordance in range-wide population structure and patterns of gene flow directionality and magnitude.

The comparison of multiple codistributed species provides a framework for investigating how dispersal, paleoclimatic changes, range expansions, divergence, and population connectivity shape present-day geographic patterns in genomic variation (Avise, 2008; Gillespie et al., 2012). Although large-scale comparative studies have been conducted using mtDNA loci and handful of autosomal loci, the sparse sampling of coalescent histories may hinder the results and conclusions of such studies. Comparative phylogeographic analyses have yet to be explored with genomic-scale population-level sampling, and here we take advantage of the greater statistical power provided by this amount of sampling that has only recently been affordable for non-model organisms (Davey et al., 2011, 2013; Glenn, 2011; McCormack et al., 2013).

To this end, we generated genomic data from five codistributed species throughout the Gulf of Mexico and along the western coast of the North Atlantic Ocean to test the following hypotheses: 1.) Persistent eastward ocean currents from the Gulf of Mexico to the Atlantic coast drive post-divergence asymmetrical gene flow directionality between subpopulations; 2.) Differences in life-history traits including body type and rafting propensity drive disparities in the magnitude and directionality of post-divergence gene flow between sub-populations; and 3.) The impact of climatic events on habitat availability drive times of divergence between subpopulations throughout the Gulf of Mexico, and between the Gulf and Atlantic coastal regions. While these three hypotheses are not mutually exclusive, their clear predictions allow our population-genomic sampling and our model-based approach to be a powerful means for understanding the fundamental forces underlying the geographic structuring of genomic variation in western

Atlantic Syngnathids. Although beyond the scope of this study, the rejection of the above hypotheses could suggest that selection may be acting to drive idiosyncratic patterns of gene flow and divergence, or heterogeneously acting on parts of the genome (i.e. "islands of divergence") (Pinho & Hey, 2010; Sousa & Hey, 2013; Cruickshank & Hahn, 2014).

With respect to the second hypothesis, we predict that seahorses (*Hippocampus*), will display post-divergence gene flow that is more congruent with ocean currents than pipefish (*Syngnathus*), because the latter might have higher capacity to actively migrate due to the presence of its caudal fin. Moreover, based on *Sargassum* fish community surveys, we predict that gene flow magnitude will be more pronounced in the species with greater rafting propensity (i.e., *S. louisianae*, *S. floridae*, and *H. erectus*) (Bortone et al., 1977; Dawson & Vari, 1983; Wells & Rooker, 2004; Hoffmayer & Franks, 2005; Casazza & Ross, 2008), rather than the euryhaline pipefish *S. scovelli* that inhabits low salinity and very shallow nearshore habitats (Dawson & Vari, 1983; Partridge et al., 2012) or the dwarf seahorse *H. zosterae* which has a very limited number of recorded rafting observations and shows site fidelity with seagrass beads and low mobility due to its small size that may limit its movement to and from rafting habitat (Strawn, 1958; Masonjones & Lewis, 1996). With respect to hypothesis 3, the historical availability of suitable habitat is likely to have differed between regions across the five focal taxa. For example, divergence times between subpopulations are predicted to be youngest across the Gulf to Atlantic coast because the presently suitable habitat between western (Gulf of Mexico) and the eastern Florida Peninsula (Atlantic coast) was only recently formed with the barrier islands that emerged in the Holocene (3-7kya) (Davis, 1992). However, species that are distribution across multiple zoogeographic provinces such as *H. erectus* and *S. floridae* are predicted to display deeper divergence times due to a greater

flexibility in ecological macroclimatic tolerances which may allow for resilience during climatic cycles (Briggs & Bowen, 2012, 2013).

To sample population genomic data for model-based inference of demographic histories parameterized by divergence times, population sizes (and size changes), and migration directionalities and magnitudes, we use restriction-site-associated DNA sequencing (RADseq) composed of several thousand genome-wide single nucleotide polymorphism (SNP) datasets per species (Table 1). In order to conduct model-based population genomic inference, one must assign sampled individual genotypes to populations that are directly incorporated into downstream coalescent-based models. To this end we use clustering methods to infer ancestral gene pools and visualize genomic similarity and divergence between individuals (i.e., PCA), as well as generated maximum likelihood phylogenies to infer the evolutionary relationship between subpopulations. Subsequently, we utilized the site frequency spectrum (SFS) across subpopulations to estimate historical demographic parameters (Gutenkunst et al., 2009; Locke et al., 2011; Excoffier et al., 2013). To this end, we specifically use the observed multidimensional SFS to estimate the aforementioned demographic parameters within a coalescent framework via composite likelihood (Excoffier et al., 2013). Although our analyses are reflective of an evolutionary temporal scale, we also simulate the biophysical movement of seasonal aggregates of *Sargassum* to quantify the potential probability and directionality of passive dispersal in this region to inform our predictions under hypotheses 1 and 2.

METHODS AND MATERIALS

Sygnathids, RADseq, and sampling strategy

Syngnathids have a sparse fossil record (Wilson & Orr, 2011), high levels of morphological plasticity (Lourie et al., 2004), and lack census data common to many economically important fisheries (Vincent, 1996; Baum et al., 2003; Evanson et al., 2011). Therefore, genetic data has been the primary means to help understand the ecology and evolution of these fishes over the last decade (Casey et al., 2004; Teske et al., 2004, 2007; Woodall et al., 2011; Boehm et al., 2013). These studies, as well as the majority of animal phylogeographic research, have focused on the use of mitochondrial DNA or a small number of nuclear loci as the primary data for inference that limits the power of statistical inference due to the large variability associated with gene co-ancestry times within populations (Avise, 2000; Hickerson et al., 2010). However, new methodological advancements to infer demographic processes have been matched with a rapid advancement of next-generation sequencing to generate large amounts of data for so-called non-model organisms. One increasingly common method is the use of restriction-site-associated DNA sequencing (RADseq) (Baird et al., 2008; Lozier, 2014) that allows homologous genome regions to be sequenced across multiple individuals and populations (Nosil et al., 2012; Carstens et al., 2013). The use of genomic data opens up a new frontier to study evolution of codistributed species with the potential to capture the complexity of species genome wide coalescent histories (Sousa & Hey, 2013; Robinson et al., 2014). However, while RADseq datasets have the potential for improved estimation of divergence and gene flow (Nosil et al., 2012; McCormack et al., 2013;

Sousa & Hey, 2013), this approach has yet to be deployed on multiple codistributed species.

It is now widely appreciated that traditional methods of gene flow estimation derived by calculating FST) are commonly faulty because of the assumption of symmetrical gene exchange and an equilibrium with respect to drift and migration (Slatkin, 1987; Whitlock & McCauley, 1999; Marko & Hart, 2011b). Furthermore, precise estimates of divergence in the presence of gene exchange can be a challenge due to evolutionary processes potentially causing genome wide variance in gene flow (Nosil et al., 2012; Carstens et al., 2013; Roux et al., 2013; Sousa & Hey, 2013) as well as the general challenge in disentangling patterns of gene flow from ancestral polymorphism (Lohse & Frantz, 2014). While coalescent theory demonstrates that sampling multiple unlinked loci can greatly strengthen parameters estimates (Nielsen & Wakeley, 2001; Adams & Hudson, 2004; Felsenstein, 2006; Wang & Hey, 2010; Huang et al., 2011) by providing replicate samples of the coalescent process (Hey & Nielsen, 2004; Hey, 2010; Sousa & Hey, 2013) deep genomic sampling also accommodates the sampling of fewer individuals (Felsenstein, 2006; Wang & Hey, 2010; Robinson et al., 2014) due to the fact that outbred diploid genomes are made up of recombining segments of DNA inherited from large pools of ancestors (Gronau et al., 2011).

A contrasting example of this substantial improvement in resolution of population history with fewer individuals and genomic-level sampling of populations can be taken from two studies of *H. erectus* (Boehm et al. 2013; *in press*). Specifically, genomic RADseq data was recently used to identify genetically distinct populations of the seahorse *H. erectus* (Boehm et al. 2014) not supported by a smaller number of loci and many more individuals (n=115, 6 loci vs. n=23, 5777 RADseq loci) (Boehm et al. 2013;

in press). Likewise, our present study utilizes 5-13 diploid individuals per subpopulation (Table 1), yet samples from 3,554-5,991 loci and (SNP from 6,272-13,672) across datasets. Although most of the focal species lack well-established mtDNA genetic lineages (Appendix 3.2), the deeper genomic sampling we deploy here has the potential to uncover complex and important details of species evolutionary and demographic histories not revealed by a single mtDNA locus.

For our sampling scheme, individuals for each species were collected throughout the Gulf of Mexico and along the Atlantic coast over multiple years (2009-2013). However, due to sampling difficulties and rarity in the area we were unable to obtain samples from the western Gulf of Mexico for *H. erectus*. For each species roughly 70-140 individuals were collected per species, verified morphologically, and sequenced for mitochondrial DNA (cytochrome oxidase subunit 1 or cytochrome b). From these samples, we randomly selected 17 to 24 individuals from throughout the sampled distributions for genomic sequencing. To reduce the chance that genomic variation or similarity was the result of relatedness we selected samples from across multiple years and different sampling locations.

RADseq library preparation, sequencing and bioinformatics processing

Extracted genomic DNA was removed from tail muscle tissue and treated with RNAase A following standard protocols (Puregene extraction, Qiagen) and confirmed for quality on an agarose gel to ensure that the majority of DNA was >10,000bp. All samples were equalized to 30 ng/uL using Qubit Fluormetric Quantitation (Invitrogen). Floragenex carried out library preparation and restriction-site-associated DNA

sequencing. Individually barcoded DNA was processed for genomic library construction and reduce genomic complexity followed (Baird et al. 2008, Lozier 2014) utilizing the restriction digestion enzyme SbfI (CCTGCA^GG). Following adapters and barcode identifier ligation, genomic DNA was sequenced on the Illumina HiSeq platform. All sequences from cut sites resulted in 90 bp single-end reads after being demultiplexed and trimmed of adapters using custom Floragenex scripts.

For each species the individual with the largest number of reads was processed to construct a de Novo pseudo reference genome. The individual with the largest number of reads was processed to construct a de Novo pseudo reference genome, and reads for each individual were aligned using BOWTIE (Langmead et al., 2009). SAMTOOLS algorithms (Li et al., 2009) were used to detect SNPs and call genotypes and outputted in the variant call format (vcf) (Danecek et al., 2011). Filtering was applied using R v.3 (R Development Core Team, 2004) to ensure a Phred score equal to a hard cutoff of q=20 (base call accuracy lower than 99%) and a minimum depth of 8x sequence coverage. To reduce the inclusion of false SNP discovery due to paralogous sequences, the vcftools software was used to remove any sites with a maximum depth calculated from the mean depth + 1.5 standard deviations calculated across all sites per species. The final datasets (Table 3.1) resulted in bi-allelic matrices across individuals with no missing data.

Many issues of confidence in genotype calling are caused due to low sequence depth (<20x coverage) (Nielson et al. 2012). However, our library preparation, sequencing coverage, and bioinformatics processing was similar to that of Lozier (2014) (mean coverage 145x (SD 65) and 235x (SD 95) for the 2 analyzed species) that compared genotype datasets called by SAMTOOLS, with an alternative method developed for low coverage next-generation sequencing (Nielsen et al., 2012) and found

little difference on impact of demographic inference or the potential for false genotype assignment when using high coverage datasets.

Genomic distance, ancestral gene pools, and phylogentic topologies

The estimation of intraspecific genetic divergence was carried out to inform the number of subpopulations used for model-based demographic inference and examine the concordance between principal component analyses, inference of ancestral gene pools, and population phylogenetic tree topologies. To visualize the major axes (PC1 and PC2) of genetic variation we utilized the software program smartpca (Patterson et al., 2006). For the inference of ancestry coefficients, representing the proportions of each individual's genome that originated from an independent ancestral gene pools (K), we used the program sNMF (Frichot et al., 2014). The program sNMF estimates individual ancestry by utilizing a sparse non-negative matrix factorization algorithm (sNMF) and is capable of efficiently analyzing large bi-allelic datasets without the loss of accuracy to the more commonly utilized programs such as STRUCTURE and ADMIXTURE. Genotypes were converted from vcf files to PCA and sNMF formats and no outliers were removed. PCA and sNMF both help assign individual genotypes to populations for downstream inference as well as being useful for visualizing continuous patterns of genetic differentiation (i.e., isolation by distance; (Engelhardt et al., 2010; Massatti & Knowles, 2014) Comparisons with the aforementioned programs (Frichot et al., 2014; Boehm et al., *in press*) found concordant results with significantly better computational efficiency. Moreover, unlike the model underlying the aforementioned programs, sNMF does not cluster ancestral gene pools based on the assumption of Hardy Weinberg or

linkage equilibrium. Ancestry coefficients (K) were estimated to determine subpopulation membership by running 10 replicates of K 1-6 using a cross-entropy criterion (CEC) for each species. To evaluate the predictive capability and error of the ancestry estimation algorithm sNMF employees the CEC that is comparable to the likelihood value implemented in the program ADMIXTURE. To select the best-supported ancestry coefficient the lowest CEC value represents the most probable K value (analogous to the highest Delta-K in the program STRUCTURE) (Evanno et al., 2005). Ancestry coefficient plots were visualized using R v.3.

The phylogentic relationships populations were inferred by implementing the program *Treemix* (Pickrell & Pritchard, 2012) that uses a maximum likelihood approach with the resulting phylogeny based on the local optimum tree, determined using a similar approach to Felsenstein (1981). *Treemix* incorporates a model to allow for population divergence in the presence of proportional post-divergence admixture/migration (m) given that incorporation of this parameter can improve the likelihood fit of a bifurcating phylogeny. However for our intraspecific topologies, the likelihood fitting reached a maximum likelihood (0.999-1.0), thereby limiting the ability for m to be incorporated into the tree graphs. The resulting phylogenies are represented with branch lengths scaled by the amount of genetic drift between populations (Appendix 3.1).

Sargassum **biophysical model simulations**

To illustrate the direction of major ocean currents while adding some predictive information on the potential of rafting dispersal, we simulated the dispersal of *Sargassum*. Annually seasonal aggregations of *Sargassum* begin to form in the Spring

(March-May) in the western Gulf of Mexico along the broad continental shelf of Texas (Wells & Rooker, 2004; Gower & King, 2011). We therefore used an ocean circulation model and particle-tracking software to estimate the likelihood of physical transport from this region to study sites. A particle release zone was defined between 26.5 - 30 N and 97.5 – 94 W. Virtual particles were tracked with Ichthyop (v.2) particle tracking software (Lett et al., 2008) within Global Hybrid Coordinate Ocean Model (HYCOM) output (Chassignet & Hurlburt, 2007). We used daily snapshots of Global HYCOM surface velocity at a spatial resolution of 0.08° and calculated particle trajectories each half hour using the Runge-Kutta fourth-order time-stepping method implemented by Ichthyop (Lett et al., 2008). These modeling tools have been widely used to simulate dispersal in Gulf of Mexico and western Atlantic (Putman et al., 2010, 2012a, 2012b; Putman & Naro-Maciel, 2013) and show a high-degree of concordance with *in situ* observations of ocean circulation, as measured by Lagrangian drifters (Fossette et al., 2012; Putman & He, 2013).

Between March 1 and May 30 (the period of peak *Sargassum* production (Gower & King 2011), 23000 virtual particles were released (250 per day) for 10 different years, 2004 – 2013. Each of these years, we recorded the number of particles that entered 0.5° x 0.5° zones centered on the 6 study sites (Fig. 3.1) within 9 months. The particle drift duration reflects observed seasonality in *Sargassum* abundance in the Gulf of Mexico and western Atlantic (Gower & King 2011) Two sets of simulations were run, one in which particles striking a coastline “beached” and another in which they "bounced" along the coast until currents changed to move them offshore. The beaching might be appropriate for modeling *Sargassum* drift (a passive process) whereas the bouncing might be more appropriate for modeling the animals within the *Sargassum* habitat, which likely have

some control over their movements that would allow them to escape such a fate (Putman et al., 2011, 2012b, 2014; Staaterman et al., 2012; Staaterman & Paris, 2013). For each simulation we determined the relative proportion of virtual particles entering the study sites (i.e., relative to the number of particles that entered any coastal site, what percentage were found at a given location).

Model-based coalescent analysis using the allele frequency spectrum

To infer the demographic history across our focal species, we utilized a coalescent simulation-based framework to infer parameters from the site frequency spectrum (SFS) implemented with the program fsc25 (Excoffier et al., 2013). This software uses a composite-likelihood approach to test and estimate parameters under historical population models of arbitrary complexity involving splits, migrations and size changes. To construct our models, subpopulation divisions were informed by observed genomic divergence (PCA), the inference of ancestral gene pools (sNMF) (Figure 3.2), and population tree topologies. To reduce linkage disequilibrium one SNP was selected per 90 bp tag and all SFS were folded. Model topologies were informed by the inference of maximum likelihood tree graphs using *Treemix* (Appendix 3.1) and range from 3 to 4 intraspecific populations per species. Parameters estimates included asymmetrical migration between populations, divergence times, effective populations size, and size change magnitudes.

Each parameter included an initial lower and upper bound to delimit search space although the composite likelihood approach does not place a constraint on the upper bound limit during the composite likelihood estimation of parameters following

coalescent genealogy simulations. Parameters within each model were free to vary; therefore to reduce the complexity of our analyses we did not compare different models with arbitrary subsets of parameters or different source and sink topologies. However, using simulated and empirical datasets, Excoffier et al., (2013) showed that precise parameter estimates were obtained in more complex models with up to ten populations. All final parameter estimates were selected based on the maximum likelihood estimate from 30-40 independent simulation runs per model. Each run consisted of 100,000 coalescent simulations and 40 cycles of the composite likelihood maximization algorithm for parameter estimations (Excoffier et al., 2013). To determine the parameter estimates with the highest likelihood we selected the model that had the smallest distance between the observed and expected multi SFS likelihood values across model runs for each species.

RESULTS

Reads and genotype (SNP) totals post bioinformatic processing

Across all sampled individuals we obtained a total of 439,997,022 single-end Illumina reads of 90 bp length. Total sequence reads per species ranged from 2,668,628-4,777,558 for *S. scovelli*, 2,723,877-5,378,986 for *S. floridae*, 3411094-5372202 for *S louisianae*, 1,264,862-4,736,299 for *H. erectus*, and 1,208,438-5,486,466 for *H. zosterae*. Following de novo assembly, genotype calling, and quality filtering across species the number of loci varied from 3,554 to 5,991 (median 4,651) and the number of polymorphic sites (SNPs) from 6,272 to 13,672 (median 11,225). For additional information see Table 3.1

Population genetic structure, ancestral gene pools and ML tree topologies

Unlike the lack of mtDNA lineage divergence across most of our focal species (Appendix 3.2), our genomic data suggests a history of subpopulation divergence and gene flow. The PCA and ancestry inference (sNMF) exhibited genetic divergence between individuals sampled from the western and eastern Gulf of Mexico, as well as on both sides of the Florida peninsula. Based the visual summaries of genomic distance, subpopulations ranged from tightly clustered individuals to genomic clines separated by smaller genetic distances between subpopulations from PCA results (PC1 and PC2; Fig. 3.2). Three of four species *S. floridae, S. scovelli* and *H. zosterae* all show a distinct cluster of genomic divergence in eastern Gulf of Mexico samples (Fig. 3.2). *H. erectus* and *S. scovelli* both exhibit subpopulation divergence on both sides of the Florida Peninsula, while the dwarf seahorse *H. zosterae* displays a cline from the northern Gulf to the Florida Keys, as well as a divergence with the Atlantic coast. Unlike the two seahorse species collected in the Florida Keys, the pipefish *S. floridae* exhibits a distinct genomic divergence between the tropical Florida Keys and all warm-temperate samples ranging from Texas to North Carolina. Another notable exception was *S. louisianae*, which shows a high degree of admixture between individuals from Texas and the western Gulf-Atlantic coast, independent from the northern Gulf of Mexico samples (Louisiana to Apalachicola, FL.).

The sNMF-based estimates of ancestry coefficients supported three distinct subpopulations (K=3) for 4 species (*H. zosterae, S. louisianae, H. erectus, S. scovelli*), with the exception of *S. floridae* (K=4). Depending on the species, the amount of admixture varied between subpopulations, ranging from distinctly isolated gene pools to some admixed populations. Most genomic admixture was associated with neighboring

populations. Nearly all cases of genomic subpopulation divergence were in contrast to the lack of mtDNA divergence with the exception of the dwarf seahorse *H. zosterae* (Appendix 3.2). These results are consistent with the *Treemix* phylogenies branch lengths scaled to the amount of genetic drift between subpopualtions. The basal population inferred from *Treemix* for our focal species were as follows: *S. louisianae*, northern Gulf of Mexico; *H. zosterae*, eastern Gulf of Mexico (i.e., Texas); *H. erectus*, western Gulf of Mexico; *S. scovelli,* Atlantic coast, FL; *S. floridae*, Florida Keys. The concordance across analyses were used to delineate ancestral gene pools (i.e, subpopulations) to inform the coalescent models for demographic inference and to define "source" populations as depicted in Figure 3.3 (not to scale) (for actual *Treemix* graphs see Appendix 3.1).

Demographic inference and gene flow direction

The coalescent-based parameter estimates of asymmetrical post-divergence gene flow (*Nm*) favored an overall trend of eastward Gulf of Mexico to Atlantic gene flow in 4 of our 5 species supporting the hypothesis that ocean currents influence gene flow directionality. Of the 16 pairwise *Nm* point estimates that showed gene exchange asymmetry, 81.25% are congruent with major ocean current directionality. Based on 36 intraspecific pairwise comparisons of post-divergence gene flow, 41.6% of *Nm* estimates are $\geq$1.0 and 19.4% between 0.1-0.99. The inference of asymmetrical migration, concordant with ocean current direction, was supported in both seahorse species and 2 of the 3 pipefish species. Table 3.2 and Figure 3.3 show that the magnitude of post-divergence gene exchange varied considerably between species, and was generally greater in species that are most commonly collected in *Sargassum* (*S. louisianae*, *S.*

floridae and *H. erectus*). *S. louisianae* in particular shows very high *Nm* values and shallow divergence based on the number of generations between subpopulation isolation.

Our coalescent inference of divergence times is based on an assumed generation time of 1-3 years (FishBase.org, Foster & Vincent, 2004; Mobley et al., 2010) except for *H. zosterae* whose generation time is assumed to range from 0.5-1 years (FishBase.org, Strawn, 1958). Four species (excluding *H. erectus* due to lack of eastern samples) showed subpopulation divergence between the eastern and western Gulf of Mexico that ranged from 5,376-16,128 to 13,461-40,383 kya, with *S. louisianae* showing the shallowest divergence and *H. zosterae* the greatest. Divergence times between the western Gulf and southern Atlantic populations ranged from 1,411-4,233 to 2,633-7899 kya for 3 species (*S. scovelli, H. zosterae and H. erectus*), while *S. louisianae* showed no divergence on either side of the Peninsula. For the two temperate populations, *H. erectus* and *S. floridae,* divergence times between these temperate populations and the eastern Gulf of Mexico were 14,804-44,412 and 2,633-7,899 kya, respectively.

For all species the effective population size (*Ne*) was greatest in the northern-eastern Gulf of Mexico; *H. erectus* (102,598), *H zosterae* (172,502), *S. floridae* (45,502), *S. louisianae* (31,628), and *S. scovelli* (64,685). In contrast, *H. erectus* (18,782), *S. louisianae* (7,103) and *H. zosterae* (39,150) populations along the Atlantic coast of the Florida Peninsula had the smallest inferred *Ne* values, and *S. floridae's* (5,041) northern Atlantic population. Table 2 shows all parameter estimates, including *Ne* values, *Ne* values prior to size change events, subpopulation divergence times, and *Nm* estimates.

Sargassum dispersal simulations

Dispersal simulations found evidence for oceanic connectivity between the Texas shelf and each of 6 study sites (i.e., areas of subpopulation divergence), regardless of whether particles "bounced" (i.e., were able to move back out of the coastal zone) or "beached" upon contacting a coastline. In general, the assumption that organisms continue to make progress with ocean currents when *Sargassum* enters nearshore waters ("bouncing") led to higher rates of long-distance dispersal (Figure 3.1b; Appendix 3.3). Although models predict large annual variation, some areas consistently received high inputs of particles and other areas consistently received low particle input. Mean values of connectivity between the Texas continental shelf and study sites was highest for the Indian River Lagoon (27.99% "bouncing", 10.29% "beaching"), Morehead City, NC (9.7% "bouncing", 5.58% "beaching"), Chandleur Sound, LA (7.30% "bouncing", 10.64% "beaching"), Tampa Bay, FL (4.95% "bouncing", 7.31% "beaching"), Pensacola, FL (1.72% "bouncing", 2.87% "beaching"), and Chesapeake Bay (0.90% "bouncing", 0.77% "beaching"). Figure 3.1b represents the probability of bouncing particles that would allow animals within the *Sargassum* vacate or colonize this habitat (Staaterman et al. 2012, Putman et al. 2012, Putman et al. 2014).

DISCUSSION

Here we demonstrate that ocean currents as well as life history traits impact the population connectivity of western Atlantic Syngnathid fishes. The majority of observed asymmetrical gene flow estimates support Hypothesis 1 suggesting that gene flow is

driven by passive dispersal and is congruent with eastward ocean current directionality. At the same time, our results also support Hypothesis 2 by suggesting that gene flow direction and magnitude may be further influenced by ecological traits related to rafting propensity and morphological differences among species supporting Hypothesis 2. The results were also somewhat consistent with Hypothesis 3 in that the heterogeneous divergence times we found could have been driven by historical changes in climate and consequent changes in species-specific habitat availability.

Gene flow asymmetry and oceanic current directionality

Population connectivity, divergence times and regional differences in genetic diversity in marine taxa may be impacted by nonrandom dispersal arising from ecological differences, hydrodynamic forces, environmental heterogeneity, and local extinction from climatic fluctuations (Rocha et al. 2002, Hare et al. 2005, Joyeux et al. 2001, Rocha et al. 2005, Eytan and Hellburg 2010). While studies of active migration focus primarily on economically important species (McBride, 2014), for many marine species that are difficult to observe, the magnitude of swimming capacity is often unknown. Under the assumption that Syngnathids primarily lack long-distance active migration, we tested the hypothesis that post-divergence gene flow between subpopulations will follow historically persistent eastward ocean current directionality. The inference of demographic parameters rejected Hypothesis 1 in only one of our five focal taxa indicating that passive dispersal may play an important role in connectivity and admixture between subpopulations of western Atlantic Syngnathids.

Asymmetrical migration was inferred for both seahorse species and two of the three pipefishes with *S. scovelli* being the lone exception. These results are consistent

with the results of biophysical model simulations that predict both the direction of S*argassum* dispersal and the capacity to reach nearshore zones throughout the sampled distribution (Figure 3.1). Model simulations of *Sargassum* movement primarily follow an eastward dispersal pattern and can disperse for extended periods of time (model time frame of nine months). These results are consistent with empirical evidence from satellite tracking as well as *Sargassum's* positive buoyancy and holopelagic ocean surface reproduction (Wells & Rooker, 2004; Gower & King, 2008, 2011). Though *Hippocampu*s was once thought to be a genus that strictly dispersed via rafting (Gunther, 1870), only some species spend a portion of their life rafting. As an explanation for this, individuals are thought to vacate or inhabit rafts as they move in and out of shorelines areas (Bortone et al., 1977), however random effects such as short planktonic movement must also be considered. For example, because of their small size at birth both pipefish and seahorses can enter the pelagic zone as plankton or enter pelagic neretic waters due to extreme weather conditions causing stable habitat to dislodge from substrate. However, once in the pelagic zone these individuals may have the potential to encounter and inhabit *Sargassum* rafts as has been shown with various marine invertebrate species (Cunningham & Collins, 1998).

Regardless of the mechanism, the probability of *Sargassum* rafts reaching shoreline areas in our biophysical simulations demonstrate connectivity between regions, though the probability to reach a given coastal zones varies both spatially and temporally (Figure 3.1b; Appendix Table 3.3). Rafting has several noted benefits to increase the survival of migrants, including shelter and the ability to support a rich community of species, which can increase the success of disperses to cross major barriers (Teske et al., 2005; Luiz et al., 2012). For example, a recent study looking at the diet of fishes within

Sargassum versus strictly pelagic individuals showed an increase in full versus empty stomachs (Vandendriessche & Messiaen, 2007) that likely reflects increased survival for rafters.

Ecological differences reflected in gene flow directionality and magnitude

Empirical evidence from several studies that cataloged fish communities of western Atlantic *Sargassum* found that Syngnathids are among the most abundant ichthyofauna rafters within the Gulf of Mexico (Bortone et al 1977, Wells and Rooker 2004), and also commonly present in rafts carried by the Atlantic Gulf Stream currents (Fish and Mowbrey 1972, Cassazza et al. 2008). Of the four species where asymmetrical gene exchange magnitude is favored, parameter estimates of gene flow (*Nm*) magnitude appear generally higher in species that are commonly found in rafting habitat (*S. louisianae*, *S. floridae* and *H. erectus*) (Table 3.3. and Fig. 3.4), suggesting that a species ecological propensity to raft, and not only stochastic rafting events, may result in long term migration trends in Syngnathid species.

Rafting has been correlated to the ability of species to traverse geographically expansive yet permeable barriers (Luiz et al., 2012). Long-distance migration of *H. erectus* is observed via *Sargassum* rafting, primarily within the Gulf Stream, and has been collected in *Sargassum* in the Mid-Atlantic's Bermuda island where *S. floridae* is also highly abundant (Smith-Vaniz et al., 1999). In contrast, *H. zostera* young are benthic at birth, and the presence of a prehensile tail used for grasping allows them to settle onto the surrounding vegetation shortly after they are released (Strawn 1958). This morphological trait is also present in *H. erectus* though this species is known to inhabit a

broader range of habitats, while *H. zosterae* is rarely found outside of seagrass beds due to their small size (2-3cm) and limited mobility (Masonjones et al. 2010).

S. louisianae in contrast shows very high levels of gene exchange between subpopulations. A recent study on fishes inhabiting *Sargassum* off of the coast of Texas (Wells and Rooker 2004) found that out of 25 fish species collected from inshore and offshore *Sargssum* rafts from southern to northern Texas, *S. louisianae* was one of the four most abundant species, making up 1,096 of the 10,518 fish collected in their study. This species also has the shallowest divergence between populations (Table 3.2), consistent with PCA, ancestry plots (Fig. 3.2), and low level of divergence inferred from *Treexmix* branch lengths (Appendix 3.1). Based on morphological differences, we predicted that pipefish (*Syngnathus*), due to the presence of a caudal fin to aid in swimming, would exhibit more genomic admixture incongruent with ocean currents then seahorses (*Hippocampus*) that utilize a prehensile tail to reduce movement and grasp stable habitat. For example, while asymmetrical gene flow is favored in *S. louisianae* this species also exhibits high levels of gene exchange in both directions from the inferred northern Gulf source population, which may be due in part to its large size as an adult (38 cm) in comparison to the focal species (*S. floridae* (25.0 cm), *S. scovelli* (18.3 cm), *H. erectus* (15.0 cm), *H. zosterae* (3 cm)) to aid in its active swimming ability.

Lastly, our results support gene flow in *S. scovelli* inconsistent with ocean currents, primarily between Atlantic and eastern Gulf populations, and no substantial ($Nm < 0.01$) gene exchange between eastern and western Gulf populations (Figure 3.4). Unlike the congeners in our study (*S. floridae* and *S. louisianae*), *S. scovelli* has not been recorded in offshore trawl or plankton collections or from depths greater than 6.1m (Dawson 1982). Moreover, as a euryhaline species it is the only species of *Syngnathus* in

our study range known to breed in freshwater and commonly inhabit low salinity estuaries (Dawson 1982). A recent study found evidence supporting migration from coastal populations into estuaries and bays, but that inhabitants of estuaries may be less prone to migrate seaward (Partridge et al., 2012).

Refugia and Colonization Directionality

Our results are also somewhat consistent with population genetic theory predicting that older and more persistent populations may have higher genetic diversity than more recently colonized populations. In line with theoretical expectations, genomic diversity was shown to be higher in the Gulf of Mexico than in Atlantic populations of *H. erectus* (Fig. 3.2; Boehm et al. *in press*). In addition, previous research on *S. floridae* found a similar pattern of increased genetic diversity in Gulf of Mexico population samples of microsatellite loci, supporting the conclusions that northern Atlantic populations recently expanded from the more stable Gulf region (Mobley et al. 2010). Our model-based inference supports these findings for warm-temperate/temeprate populations of *S. floridae,* and inferred a shallow divergence (2,633-7,899 kya) between the western Gulf of Mexico and the north Atlantic coast (i.e., the Carolinas) and a older divergence between eastern and western Gulf of Mexico of 5,376-16128 kya as well as a much larger effective size for both Gulf populations (Table 3.2). Northern colonization may be due to late-Pleistocene sea level changes and climatic conditions causing extinction and recolonization for a number of marine species in this temperate region (i.e., the Carolinas northward) (Bernatchez & Wilson, 1998; Grant & Bowen, 1998; Wares & Cunningham, 2001). Both northern population of *S. floridae* and *H. erectus* also

display much higher expansion magnitudes then southern Gulf populations (Table 3.2) further supporting the suggestion of eastward founder colonization from the Gulf with subsequent post-divergence gene flow.

Marine barriers, habitat availability, and genomic subpopulation divergence

Here we explore subpopulation divergence in relation to climatic changes and consequent changes in habitat availability, as well as the concordance with known minor barriers to dispersal in the western Atlantic. With some exception, our results support Hypothesis 3, suggesting that the impact of climatic events on habitat availability may drive divergence between subpopulations within the Gulf of Mexico and between the Gulf of Mexico and Atlantic coastal regions. Across most species, we find isolation between subpopulations between the western and eastern Gulf of Mexico, as well as between the western Gulf and the Atlantic (Figure 3.2). Inferences of divergence show that 45.5% of subpopulations split roughly 7-30 thousand generations ago (kga), whereas the rest show more shallow divergences from 1-7kga. The majority of shallow divergence times are between the Gulf and Atlantic coastal regions separated by the Florida Peninsula (Table 3.2).

Previous research supports several proposed dispersal barriers corresponding to major Atlantic Ocean currents (Briggs, 1974; Hallam, 1997; Rocha, 2003), yet most barriers are permeable and their impact will vary across taxa (Joyeux et al., 2001; Floeter et al., 2007; Rocha & Bowen, 2008). Despite the estimates of post-divergence gene flow in all five species, the divergence time estimates and the identification of distinct ancestral gene pools across the range of our focal species, both demonstrate that divergence is maintained in contemporary conditions of connectivity.

Though major ocean currents in the Atlantic have maintained some directionality that affected regional connectivity patterns across species, one explanation for population divergence is that the current estimates of regional isolation are the consequence of range expansion from refugia via colonization across unsuitable habitat into areas that were only suitable after the Last Glacial Maximum (LGM) due to coastal reorganization that occurred during the Holocene (Davis, 1992; Blum, 2003). In this latter phase, it is estimated that at 18 kya the Florida Peninsula extended outward to the continental shelf 125 meters lower then contemporary conditions (Ludt & Rocha, 2014). As sea levels rose and moved landward, the rate of rising sea level slowed roughly 7000 kya, which generated Holocene deposits resulting in the formation of most barrier islands of the Florida Peninsula (3000-7000 Kya) (Davis, 1992; Blum, 2003). At the same time, the warming conditions of the Holocene coincided with newly available habitat on the Atlantic coast north of the Florida peninsula, whereby colonization could have resulted with some isolation from the source refugia in the Gulf of Mexico, specifically for the temperate populations of *S. floridae* in the Carolinas (2,633-7,889 kya) and *H. erectus* North of Cape Hatteras (14,804-44,412 kya).

Most Syngnathids live primarily in sheltered seagrass during the breeding season (Dawson 1982, Forster and Vincent 2005, Mobley et al. 2010), and the shallow divergence times ranging from 1,861-5,583 to 3,733-11,198 kya between the Gulf and Atlantic populations are consistent with the Holocene coinciding with newly available habitat along the barrier island stretching 550 km of the north Atlantic coast as well the Carolina and mid-Atlantic coast for some Sygnathid taxa. Without this island barrier formation, colonization and range expansion would have been hindered at the Holocene because the eastern Peninsula would have been unprotected from the wave action and

lacked grass beds, similar to the contemporary conditions in the 700km area north of the Peninsula to the Carolinas, where a dearth of Syngnathids are located (Dawson & Vari, 1983). Moreover, with the exception of *S. scovelli* all Atlantic coastal populations have the smallest contemporary *Ne* and highest magnitude of population expansion, further supporting the inference of colonization (Table 3.2). For *S. scovelli* these processes may also account for potentially suitable habitat due to its ability to live in freshwater, and potentially inhabit outflows that form what would now be the barrier inlets between the ocean and bay side of the eastern Peninsula (Davis, 1992).

Isolation within the Gulf of Mexico

For most species, divergence times were concordant between eastern and western Gulf of Mexico subpopulations with the latter consistently having generally larger estimates of *Ne* as well as lower expansion magnitudes than Atlantic populations (Table 3.2). Despite lower sea levels at the last Glacial maximum, the historical persistence of fish populations within the Gulf of Mexico has been suggested to be the result of persistent refugial populations separated during the Pleistocene, followed by secondary contact at the Holocene (Portnoy and Gold 2013). The concordance of divergence times within the Gulf of Mexico across populations may be suggestive of two factors: 1) late Pleistocene cold pulses of freshwater discharge from the Mississippi River Delta (9-16 kya), and 2) potentially unsuitable habitat between northern Texas to west of the river Delta. Across most populations divergence times ranging from 5,376-16,128 (*S. floridae*) to 13,461-40,383 Kya (*H. zosterae*) (Table 3.2) that roughly correspond to the cold freshwater outflows during glacial melts (9-16 Kya) (Anderson et al., 2012).

Additionally, Dawson and Vari (1982) carried out the most comprehensive study of Atlantic Syngnathids throughout the Atlantic and based on catch records found that contemporary distributions indicate a near absence of specimens between northern Texas to west of the Mississippi River Delta. This may further reflect a lack of stable demes across this range, a pattern that could be reflected in our divergence time estimates if it has been historically persistent.

Divergence between zoogeographic provinces

The divergence times spanning different zoogeographic provinces (Briggs & Bowen, 2012) is particularly pronounced in two species, *H. erectus* and *S. floridae,* while there is little mtDNA divergence across the same province boundaries in both cases (Appendix 3.2). Both of these species span ocean provinces that range from stable tropical conditions to temperate zones that experience extreme seasonal temperature fluctuations. Even with some levels of gene flow between populations, as seen in *H. erectus*, selection may shape polymorphism and result in "islands of divergence" across the genome (Sousa & Hey, 2013), yet this would have to be tested explicitly (Roux et al., 2013). These estimates of pronounced autosomal divergence in these two species correspond with macroclimatic transition zones. For example, PCA and ancestral gene pools for *H. erectus* show temperate Virginia Province individuals form a tight genomic cluster, separate from the rest of the sampled distribution, and an inferred divergence time of 14,804-44,412 kya. A recent study on *H. erectus* (Boehm et al., *in press*) found that individuals represent a persistent population in the temperate Virginia Province north of Cape Hatteras (i.e., from Chesapeake Bay to New York State) and several hypotheses

were proposed for this divergence including the potential that southern inhabitants may have a lower physiological tolerance to temperate conditions thereby reducing the chance of overwintering survival.

Additionally both sister-species *S. floridae* and *S. louisianae* can be collected sympatrically throughout the much of the Gulf of Mexico (Dawson and Vari 1983). However, as a warm-temperately adapted species, *S. louisiane* is highly abundant along the Atlantic coast of the Florida Peninsula (i.e., Indian River Lagoon), while *S. floridae* is absent from the majority of the Atlantic coast of Florida, yet reemerges in the more temperate Carolinas (Dawson & Vari, 1983; Mobley et al., 2010). The absence of *S. floridae* from the majority of the Atlantic coast of Florida is enigmatic considering that it is commonly found sympatrically with *S. louisianae* in warm-temperate areas throughout the Gulf of Mexico, while in contrast individuals of *S. floridae* from the tropical Florida Keys form a distinct genomic cluster (depicted clearly in the PCA graph and ancestry plot; Fig. 3.2) and show an old divergence (29,041-87,123 kya) from all warm-temperate populations (ranging from Texas to North Carolina). One potential explanation for this divergence is that individuals in the tropical Caribbean Province may have a reduced thermal tolerance, and that selection may be acting on these individuals, restricting them to stable tropical conditions. This population may also account for its high abundance in tropical island of Bermuda, unlike the other two warm-temperately adapted congeners in this study (Dooley, 1972; Smith-Vaniz et al., 1999).

Future hypotheses and conclusions

A recent study by Cruickshank and Hahn (2014) suggest that genomic "islands" of relative divergence may represent loci involved in local adaptation without

introgression in the rest of the genomes of sister species. While the impact of selection remains beyond the scope of this study, the findings outlined in the previous section may allow for the generation of new selection driven hypotheses to be explored between subpopulations inhabiting different macroclimatic conditions where natural selection may be acting on a given population or heterogeneously across areas of the genome.

A lack of mtDNA genetic structure across most of our focal species is generally attributed to high dispersal ability, yet by utilizing genomic data, we observed many more discreet ancestral gene pools with post-divergence migration between populations. The overall patterns of our results suggest that ecological traits coupled with passive dispersal may play an important role in the distribution, isolation, and connectivity of western Atlantic Syngnathids. Inference of eastward post-divergence gene flow values supports the hypothesis that asymmetrical ocean currents may have driven the directional movement for the majority of these species, and also further supports that life history traits such as rafting propensity plays an important role in the dispersal of these primarily coastal inhabitants.

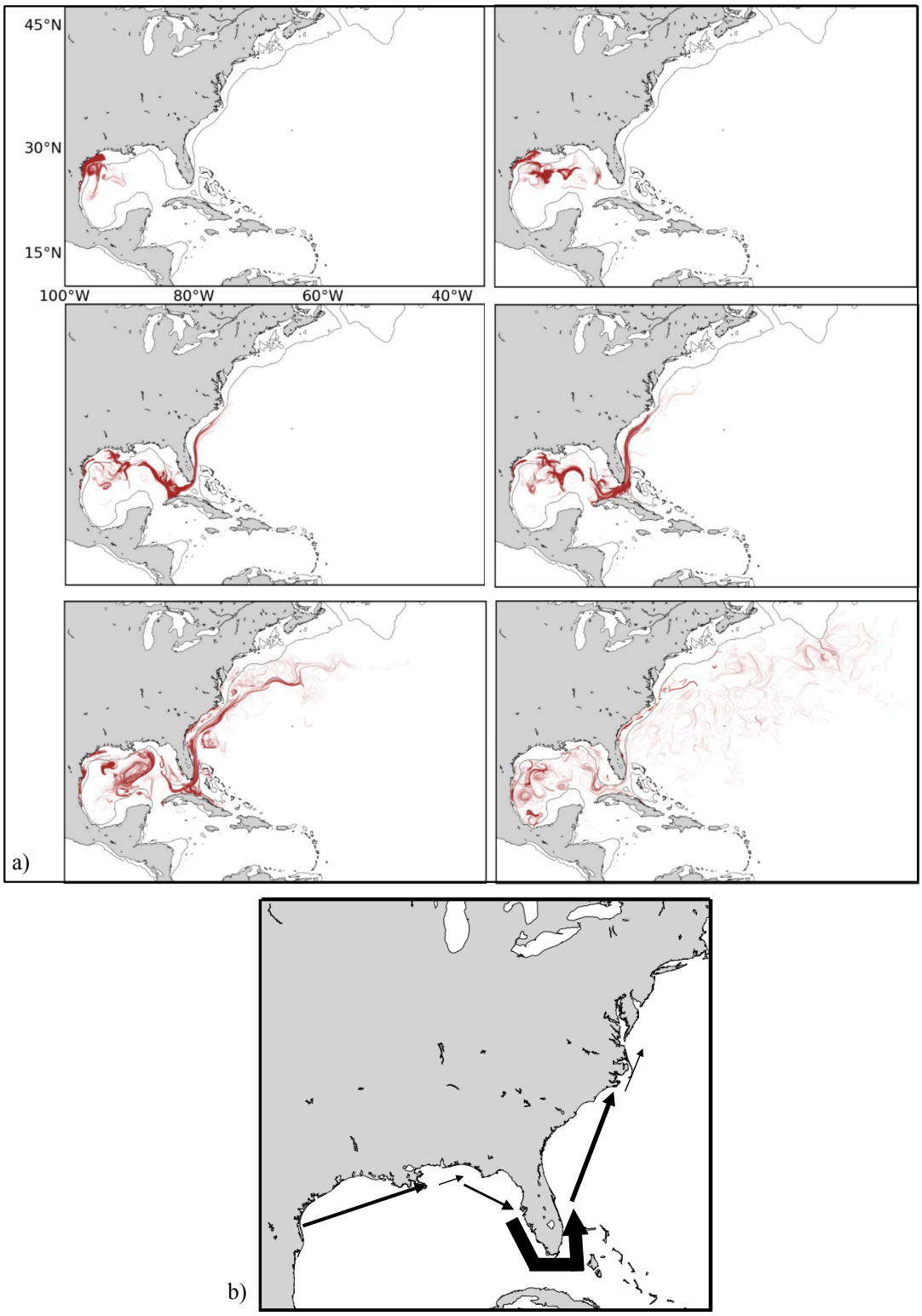


Figure 3.1 Biophysical model simulations of *Sargassum* dispersal from March-November. (a) Each image represents dispersal roughly every 1.5 months beginning in March and ending in November (2005). (b) Dispersal probabilities of *Sargassum*. Arrows sizes are proportional to the probability of *Sargassum* reaching each of the following locations: Chandeleur Sound, LA., Pensacola, FL., Tampa Bay, FL., Indian River Lagoon, FL, Moorehead City, NC., and Chesapeake Bay. Particles were released off of the coast of Texas from March-May. Probabilities based on mean values from 10 years of ocean circulation (2000-2009).

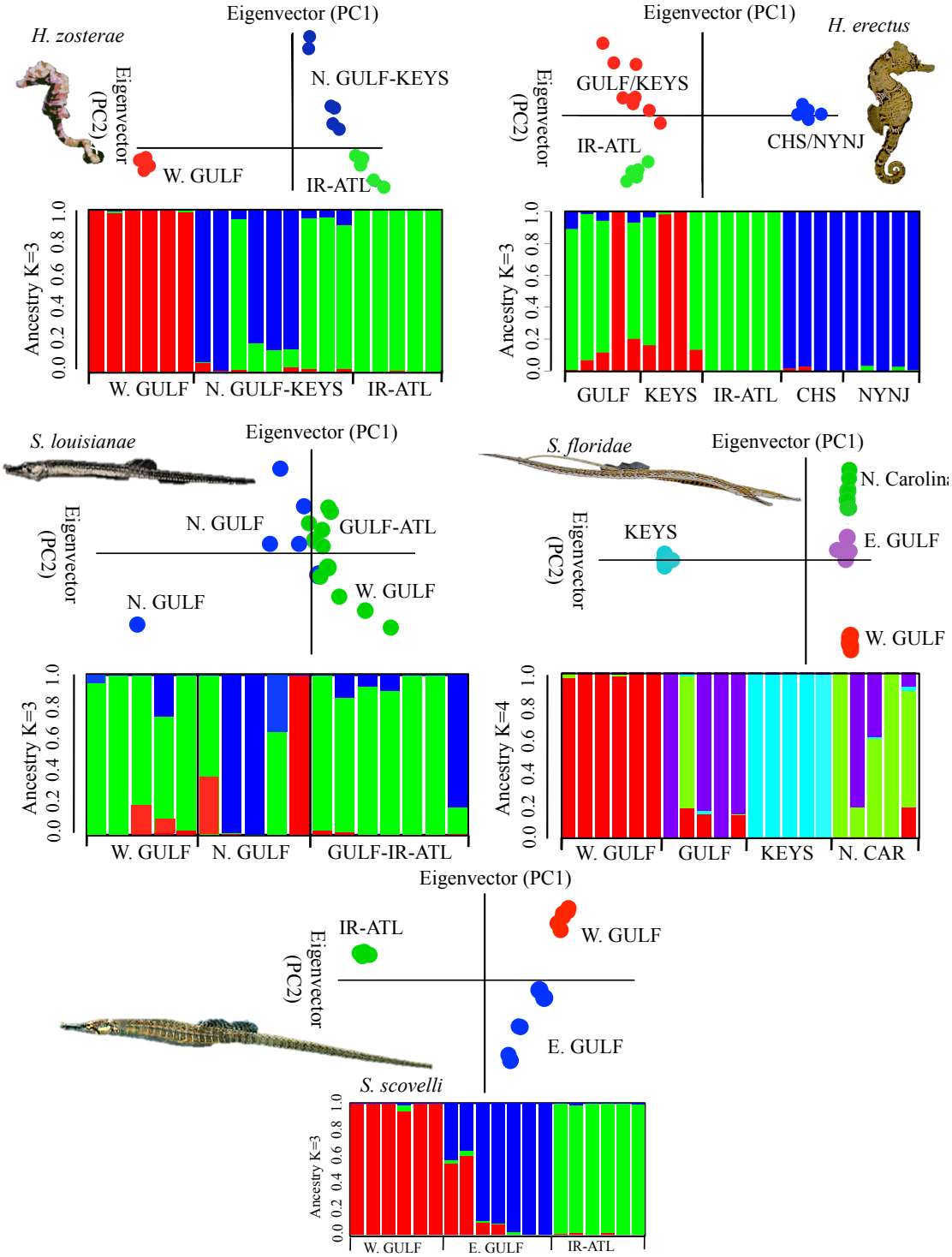


Figure 3.2 Genomic variations across individuals and subpopulations. Principle component analyses and ancestry plots derived from sNMF ancestry coefficients. Each line of the sNMF plot represents one individual. Population definitions: Chesapeake Bay = CHS; New York-New Jersey = NYNJ; Florida Atlantic coast = IR-ATL; Western Gulf of Mexico = W. GULF; Eastern Gulf of Mexico = E. GULF; Northern Gulf of Mexico = N. GULF; North Carolina = N. CAR; Florida Keys = KEYS.

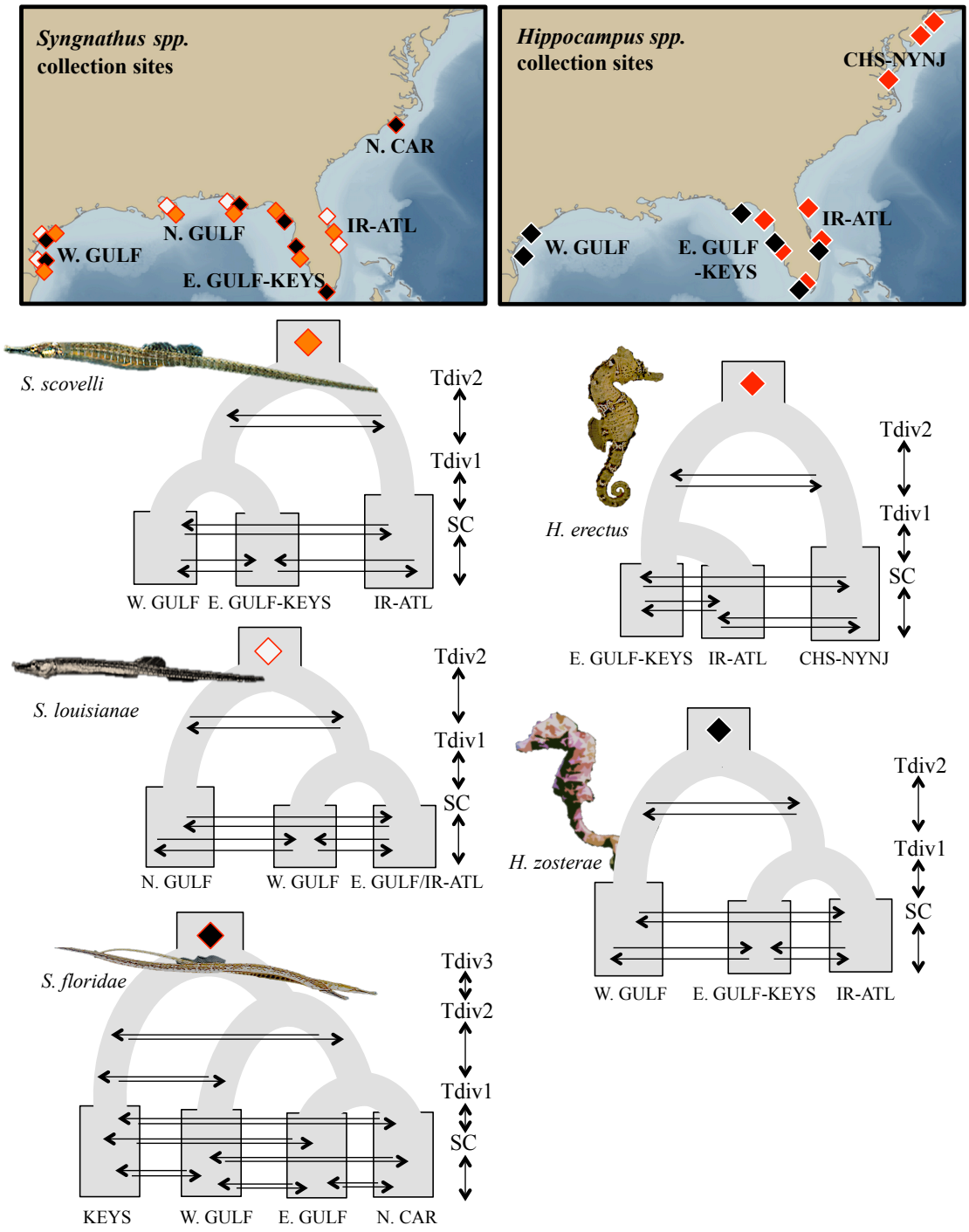


Figure 3.3 Graphic depictions of coalescent models per species and collection locations. Populations were defined based on the congruence between PCA and ancestry plot coefficients (Figure 3.2). Model topologies were informed based on *Treemix* population phylogenies (See Appendix 3.1 for *Treemix* results). The colored diamonds correspond to the collection locations per species. Tdiv1= divergence time one; Tdiv2= divergence time two; Tdiv3= divergence time three; SC = size change event with *Ne* values estimated for both modern and ancestral post-divergence population sizes. Arrows illustrate the direction of gene flow parameters (For parameter values see Tables 3.2 and 3.3 and Figure 3.4 for gene flow patterns). Population definitions: Chesapeake Bay = CHS; New York-New Jersey = NYNJ; Florida Atlantic coast = IR-ATL; Western Gulf of Mexico = W. GULF; Eastern Gulf of Mexico = E. GULF; Northern Gulf of Mexico = N. GULF; North Carolina = N. CAR; Florida Keys = KEYS.

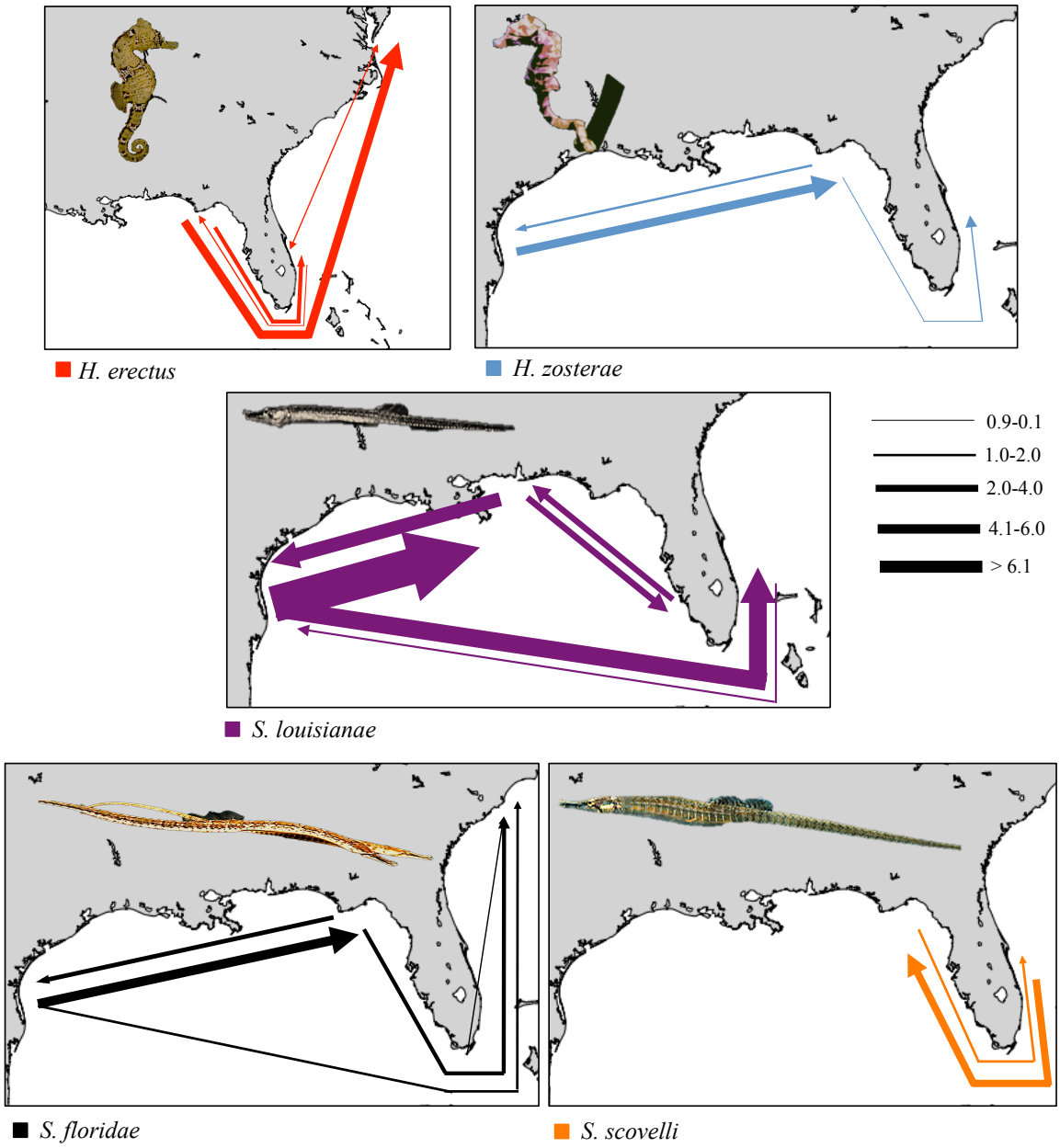


Figure 3.4: **Gene flow estimates from coalescent model simulations for each species.** Arrows scaled to Nm values (See Table 3.3 for parameter estimates). Nm values <0.1 not shown.

Table 3.1: Number of individuals sequenced per designated subpopulation (*n*), analyzed loci across all individuals at all analyzed (SNP) sites.

Subpopulations per species	*n*	Analyzed loci	Analyzed SNPs
Hippocampus erectus		5,777	11,708
Eastern Gulf (EG)	9		
South Atlantic (SA)	5		
North Atlantic (NA)	9		
H. zostera		5,991	13,672
Western Gulf (WG)	5		
Eastern Gulf (EG)	13		
South Atlantic (SA)	5		
Syngnathus scovelli		3,443	7,058
Eastern Gulf (EG)	6		
Western Gulf (WG)	7		
South Atlantic (SA)	6		
S. louisianae		3,554	6,272
Eastern Gulf (EG)	5		
Northern Gulf (NG)	5		
Western Gulf-South Atlantic (WGSA)	7		
S. floridae		4,615	11,225
Eastern Gulf (EG)	7		
Western Gulf (WG)	5		
North Atlantic (NA)	5		
Florida Keys (FK)	5		

Table 3.2: Composite likelihood parameter estimates inferred from the most probable model across 30-40 independent coalescent model runs per species.

Species	Populations	Modern *Ne*	*Ne* before size change	Divergence time (Tdiv1)	Divergence time (Tdiv2)	Divergence time (TDiv3)
H. erectus						
	Eastern Gulf (EG)	102,598	26,892	EG-SA divergence 1,861-5,583 kya	EG-NA divergence 14,804-44,412 kya	
	South Atlantic (SA)	18,782	519			
	North Atlantic (NA)	32,524	639			
H. zosterae						
	Western Gulf (WG)	57,873	44,313	EG-SA divergence 3,733-11,198 kya*	EG-WG divergence 13,461-40,383 kya	
	Eastern Gulf (EG)	172,502	18,092			
	South Atlantic (SA)	39,150	1,354			
S. floridae						
	Florida Keys (FK)	5,415	176	WG-NA divergence 2,633-7,899 kya	EG-WG divergence 5,376-16,128 kya	FK-EG divergence 29,041-87,123 kya
	Western Gulf (WG)	21,623	191			
	Eastern Gulf (EG)	45,502	22,296			
	North Atlantic (NA)	5,041	683			
S. louisianae						
	Northern Gulf (NG)	31,628	10,818	WG-EGSA divergence 1,411-4,233 kya	NG-WG divergence 7,423-22,269 kya	
	Western Gulf (WG)	18,421	8,466			
	Eastern Gulf-South Atlantic (EGSA)	2,376	84			
S. scovelli						
	South Atlantic (SA)	22,189	15,962	WG-EG divergence 2,424-7,272 kya	SA-EG divergence 8,454-25,362 kya	
	Eastern Gulf (EG)	64,685	26,023			
	Western Gulf (WG)	14,673	365			

*Divergence times are based on a generation time of 1-3 years per species with the exception of *H. zostera* (0.5-1.5 years).

Table 3.3: Composite likelihood gene flow (*Nm*) parameter values inferred from the most probable model across 30-40 independent coalescent model runs per species. The population prior to the arrow (>) indicates the source of gene flow to the receiving population (also see Figure 3.4).

Species				
H. erectus	Eastern Gulf (EG)	South Atlantic (SA)	North Atlantic (NA)	
	EG>SA	SA>EG	NA>EG	
Nm	1.135	0.435	0.1785	
	EG>NA	SA>NA	NA>SA	
Nm	4.425	0.515	0.51	
H. zosterae	Western Gulf (WG)	Eastern Gulf (EG)	South Atlantic (SA)	
	WG>EG	EG>WG	SA>WG	
Nm	4.225	1.215	0.015	
	WG>SA	EG>SA	SA>EG	
Nm	0.025	0.34	0.065	
S. floridae	Florida Keys (FK)	Western Gulf (WG)	Eastern Gulf (EG)	North Atlantic (NA)
	FK>WG	WG>EG	EG>WG	NA>WG
Nm	0.05	5.06	1.805	0.045
	FK>NA	WG>NA	EG>NA	NA>EG
Nm	0.01	0.1	3.475	2.1095
	FK>WA	WG>FK	EG>FK	NA>FK
Nm	0.115	0.06	0.08	0.04
S. louisianae	Western Gulf (WG)	Northern Gulf (NG)	Eastern Gulf-South Atlantic (EGSA)	
	WG>NG	NG>WG	EGSA>WG	
Nm	20.45	11.48	0.29	
	WG>EGSA	NG>EGSA	EGSA>NG	
Nm	15.02	4.86	3.45	
S. scovelli	South Atlantic (SA)	Eastern Gulf (EG)	Western Gulf (WG)	
	SA>EG	EG>SA	WG>SA	
Nm	3.67	1.22	0.02	
	SA>WG	EG>WG	WG>EG	
Nm	0.025	0.35	0.045	

CHAPTER 4

Comparing traditional Chinese medicine with Internet and coastal souvenir retailer markets of the United States seahorse trade: an example of DNA barcoding as a teaching tool for wildlife forensics

ABSTRACT

In a collaborative project between high school students, undergraduates, and research mentors, this study utilized DNA barcoding (identifying species from short homologous gene sequences) in a multi-year project (2009-2014) to evaluate the species composition of two prominent markets of dried seahorses (*Hippocampus*) in the United States: (1) traditional Chinese medicine (TCM), and (2) Internet-based auction sites and coastal souvenir retailers (ICR). Eight species were identified from TCM specimens (N=168), mostly of Asian Indo-Pacific origin with the exception of *H. algiricus* from western Africa and *H. ingens* from Latin America. In contrast, ICR samples (N=344) included five species, made up primarily of the Indo-Pacific *H. kuda* (51.5%) and the dwarf seahorse *H. zosterae* (40.7%) that is domestic to U.S. waters. The presence of *H. zosterae* differs from previous reports that have limited the scope of its exploitation to the live aquarium trade. eBay is a primary source of seahorse ICR and in 2012 an agreement was established to prohibit the sale of non-domestic seahorse species, therefore *H. zosterae* remains outside this agreement. However, all *Hippocampus* species are CITES listed (2004) due to overexploitation, and 167 signatory nations have agreed to trade reporting and the implementation of a 10-cm size limit to help maintain a more sustainable fishery. When we evaluated adherence to this size recommendation, the majority of TCM samples (85.7%) were >10-cm, in contrast with only 4.8% of ICR samples (excluding the dwarf species *H. zosterae*). The small size of ICR specimens

suggests non-size discriminatory bycatch as the most likely source. Moreover, to evaluate trade reporting we examined all available CITES import records and estimated that roughly 302,360 dried specimens were imported into the U.S. from 2004-2012. During this timeframe many U.S. imports remain listed as unidentified species (*Hippocampus spp*.) from confiscated shipments. Therefore, we also analyzed a recently confiscated shipment into the Port of New York, which contained two of the three most abundant species found in our TCM samples (*H. trimaculatus* and *H. spinosissimus*).

Keywords: dried seahorses, *Hippocampus,* traditional Chinese medicine, eBay, curio, DNA barcoding, molecular forensics, conservation genetics, CITES

INTRODUCTION

There is growing awareness that much of the world's fisheries are overexploited, but relatively little attention is paid to, or resources provided towards, the trade of non-food fisheries (Myers & Ottensmeyer, 2005). The growing popularity of traditional medicine has played a key role in the overexploitation and increasing global demand of non-food species including the genus *Hippocampus* (seahorses) (Alves & Rosa, 2007). Seahorses are commonly traded for traditional Chinese medicine (TCM), as aquarium pets, as well as for souvenirs and crafts (i.e., curio) (Foster & Vincent, 2004; Lourie et al., 2004; Vincent et al., 2011b). Historically, seahorses have been collected predominately throughout the Indo-Pacific; however, Southeast Asian countries and India have reported substantial declines in catch, causing the demand for seahorse fisheries to expand worldwide (Baum and Vincent 2005; Sanders et al. 2007; Vincent et al. 2011b). Tens of millions of seahorses are traded each year in as many as 80 countries (Vincent et al., 2013). Most seahorses enter the trade as incidental bycatch, leading to difficulties in tracking trade patterns and identifying species exploitation due to non-selective fishing with regard to both size of individuals as well as species collected (Hall et al. 2000; Baum et al. 2003; Baum and Vincent 2005; Meeuwig et al. 2006; Vincent et al. 2007).

Seahorses are found in temperate and tropical nearshore environments globally. Due to habitat loss and life history traits such as low active dispersal ability and low fecundity, many species are particularly vulnerable to overexploitation (Foster & Vincent, 2004; Vincent et al., 2011a). With the demand for seahorses outgrowing supply in many regions, concerns over fishery increases prompted the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) to list all

species of *Hippocampus* in Appendix II (Foster & Vincent, 2005) in an attempt to regulate exports that may threaten wild populations. This CITES listing was a significant achievement. Out of roughly 35,000 CITES-listed species only a dozen are marine fish groups, despite decades of decreasing or depleted fishery stocks worldwide (Reynolds et al., 2005; Worm et al., 2006, 2009; Vincent et al., 2013; McBride, 2014). Additionally, as a bycatch-dominated fishery the contemporary health of many seahorse populations are often unknown (Meeuwig et al., 2006). For example, of the 38 *Hippocampus* species listed on the International Union for the Conservation of Nature (IUCN) Red List of Threatened Species, 26 are considered "data deficient," 10 as "Vulnerable," one as "Endangered" and one of "Least concern." Of the 10 "vulnerable" species, all are exploited for trade, mostly as dried specimens, and more species are likely "vulnerable" but current population trends remain "unknown" for 29 species and 19 evaluations "need updating" (IUCN, 2014).

The evaluation of seahorse species most prominently being imported into the U.S. market, or collected domestically, can help to inform efforts in seahorse conservation and increases our baseline knowledge of species types, size, and overall trade dynamics and fishery pressures. In the United States, trade records indicate that considerable numbers of live and dried seahorses are imported annually from Asia (Vincent, 1996; Evanson et al., 2011; Vincent et al., 2011b), with an increasing market share in South America (Sanders et al., 2007), and Africa (Vincent et al., 2011b). TCM dominates the sale and purchase of dried specimens, though dried seahorses can also be purchased at souvenir and craft stores in coastal areas throughout the U.S. However, now with the increased use of the Internet for purchasing goods, platforms such as eBay and other Internet based resources make curio products readily available to purchase. This shift in product

availability makes dried seahorses easily obtainable across the U.S. and no longer limits sales to urban TCM markets or coastal areas. Because of the growing Internet trade, agreements were made to limit the sale of imported seahorse species on eBay in the U.S. and Europe (2012), and more recently in Australia (2013). The agreement with eBay was put in place to reduce exploitation of seahorses as curio, though this does not apply to the sale of domestic species.

In addition to eBay agreements and CITES listing, to help reduce the impact of non discriminatory bycatch, the Technical Animals Committee for CITES suggests that all signatory nations implement a recommendation for a 10-cm minimum size to allow the most exploited species to reach reproductive age before removal from the wild (Foster & Vincent, 2005). Because some fishers often make a substantial portion of their income catching seahorses (Pajaro et al., 1997; Vincent et al., 2007), this recommendation was put forward to allow fishers the ability to maintain an income while adopting more sustainable fisheries practices. Therefore, in addition to species identification the evaluation of specimen size can help address which developmental stages (i.e., juveniles or adults) particular species are being sold and in which end-market.

The use of DNA barcoding (identifying species from short homologous gene sequences) has become increasingly popularized over the last decade (Hebert et al., 2003a) and is now a commonly utilized tool to for the biological identification of animal products (Alacs et al., 2010). To date, DNA barcoding has been used across a range of marine-related questions including those involving the aquarium trade (Steinke et al., 2009), morphologically unidentifiable specimen fragments such as shark fins (Hoelzel, 2001; Holmes et al., 2009), skate products (Griffiths et al., 2013), tuna species labeling (Lowenstein et al. 2009), and caviar source-species identification (Doukakis et al., 2012).

As DNA barcoding becomes more accessible, students are able to conduct their own investigations, such as assessing the validity of fish being sold in New York City sushi restaurants (Schwartz, 2008). This allowed for the development of ongoing education programs such as the Urban Barcoding Project (urbanbarcodeproject.org), as well as the development of high school and undergraduate curriculum (Harris & Bellino, 2013) with molecular labs for high school students conducting marine research in both New York City and in Belize (biobelize.org). With the countless number of species being traded globally, it becomes difficult to sequence and identify a sufficient quantity of species with a limited number of biologists or paid research assistance. Therefore, we wanted to show the promise of using high school and undergraduate students to perform molecular work that can increase efficiency while also providing an authentic research experience.

Our study represents a collaborative effort between students and research mentors to identify and compare seahorse specimens from two primary markets: 1) traditional Chinese medicine (TCM) and, 2) Internet products and coastal souvenir retailers (ICR). In total, 532 samples were collected from 2009-2014. TCM samples were acquired from four Northeastern metropolitan "Chinatown" end-markets. Most ICR samples were purchased on eBay from sellers located in Florida and Texas, independent websites, and directly from souvenir shops along the Texas Gulf coast. We examine the contrasts between the species composition and size of specimens in each market. Lastly, CITES import records (2004-2012) were used to estimate the total number of individuals imported into the U.S., and compared with the species type identified in our samples. Currently a large number of imports remain listed as *Hippocampus spp*., many of which are confiscated. Therefore, we also identified a subset of individuals from a confiscated shipment into the Port of New York (2013). To carry out these goals we utilized two

mitochondrial DNA markers (cytochrome b or cytochrome oxidase subunit 1) as "barcode" genes for species identification purposes. By utilizing DNA barcoding to examine the seahorse trade our results further demonstrate that molecular forensics is an adaptable tool for training students to tackle molecular biology research and identifying species. To disseminate the use of DNA barcoding for in-depth wildlife trade investigations we have provided information on protocols as well as a workflow for establishing independent molecular based projects on Syngnathidae species and other wildlife products (Appendix 4.1).

METHODS AND MATERIALS

Sampling design

TCM samples were collected from Philadelphia, PA and the boroughs of Brooklyn, Queens, and Manhattan, NY. In total, samples were purchased from 13 stores (2-3 per area). When selecting TCM samples, specimens were chosen at random regardless of size. A minimum of 8 and maximum of 20 individuals were purchased at each location. If more than one size class was for sale, we selected an equal number of individuals at random from both groups. In some cases, if species were visibly different morphologically we selected across morphotypes so that all potential species were selected. Internet samples (10 stores) were purchased randomly. In addition, curio samples were purchased from four stores along the southern coast of Texas. Although we do not attempt to equate the number of species identified in our samples with the number being sold across the U.S. we did attempt to collect samples randomly and in relatively equal numbers. Moreover, when multiple size options were available from vendors we

made our best attempt to purchase evenly from all size groups so as not to bias the size comparison results. (For details see Appendix Table 4.2).

DNA extraction PCR and sequencing

This study was conducted primarily at the Sackler Institute for Comparative Genomics at the American Museum with three undergraduates and one high school student. In addition, three students conducted work at the newly established molecular biology lab at the High School for Environmental Studies (Harris & Bellino, 2013) in Manhattan, and two students from the Brooklyn Academy of Science and the Environment High School under the guidance of Dr. Susan Pell (Brooklyn Botanical Gardens).

A total of 517 specimens were purchased and identified from 2009-2014. In addition, a subset of 20 individuals provided by the U.S. Fish and Wildlife Service were sequenced for identification from a confiscated shipment into the Port of New York (2013).

For specimen identification, samples were sequenced for either a 696-base pair (bp) segment of mitochondrial (mtDNA) cytochrome b (cytb) or a 658bp section of cytochrome oxidase subunit 1 (CO1). The nucleotide sequences of cytb and CO1 contain species-specific information and have been shown to readily amplify from dried seahorse specimens (cytb; Sanders et al. 2007) and across metazoans (CO1; Hebert et al. 2003). Cytb is the dominant mtDNA gene used in studies of seahorse trade (Sanders et al. 2007), systematics (Casey et al., 2004; Teske et al., 2004; Teske & Beheregaray, 2009) and population genetic studies (Lourie & Vincent, 2004; Lourie et al., 2005; Lourie, 2006;

Woodall et al., 2011; Boehm et al., 2013) and is available on GenBank, while the Barcode of Life (BOLD) systems database now contains CO1 sequences for nearly all seahorse species.

Tissue was removed from tail muscle after removing the skin of the sample to avoid contamination of unknown fixatives sometimes used to preserve specimens. Although in many cases the morphology of specimens was severely degraded, students were instructed to sample from one side of the fish so the species could be verified morphologically if necessary due to the bilateral symmetry of the species. Total genomic DNA was extracted using standard extraction protocols (Qiagen DNeasy Blood and Tissue kit). PCR amplification for CO1 took place in 25µl reactions with 0.5µl of forward and reverse primers (10µM; (dgLCO1490: GGT CAA CAA ATC ATA AAG AYA TYG G and dgHCO2198: TAA ACT TCA GGG TGA CCA AAR AAY CA) (Meyer 2003), 0.2µl of Fisher taq polymerase, 2.5µl 10X Buffer, 1µl dNTP (10 mM) and 1µl of genomic DNA. PCR thermal cycling conditions were: 5 min at 95°C; 35 cycles of 30s at 95°C; 1 min at 48-50°C; 1 min at 72°C; and a final 10 min at 72°C. The amplification protocol for PCR reactions for cytochrome b was: 94°C for 5 min; 35-39 cycles of 94°C for 30 s, annealing temperature 50 to 52°C for 1 min, 72°C for 1 min; final extension 72°C for 10 min; final rest at 4°C. Seahorse specific primers were from Lourie et al. (2005): forward *shf2* TTG CAA CCG CAT TTT CTT CAG and reverse *shr2* CGG AAG GTG AGT CCT CGT TG. PCR products were visualized on 1.5% agarose gels stained with SYBR®Safe. Samples that amplified successfully were purified with Agencourt AMPure XP, sequenced with BigDye version 3.1 chemistry (Applied Biosystems, Inc. [ABI], Foster City, CA), and sequence data were collected using an ABI3730 genetic analyzer (ABI). All sequences were aligned using Geneious Pro v.6 (Drummond et al.,

2011).

Specimen identification

The sequences were identified using the BLASTn algorithm against the GenBank and BOLD databases and assigned species level status at >98%. This threshold was chosen based on pairwise distance comparisons of mitochondrial sequence divergence that confirmed the species identified in this study have a > 2% genetic distance (CO1; range 2.1-15%, cytb; 3.7-17.5%). The lower distance for CO1 (2.1%) corresponds to the closely related species *H. ingens* and *H. algircus*. We note that although some species in our samples could be identified based on clear morphological characters, other species take careful examination of morphometric counts and require taxonomic training, which was not always available to students. Moreover, many specimens due to various states of decay were not in optimal condition for morphological analysis. All TCM samples were sequenced and then identified. However, for Internet-curio specimens, many parcels or "lots" contained a single species, particularly consisting of either *H. kuda* or *H. zosterae*. For these parcels a subset of individuals were randomly chosen and sequenced (N=82). Remaining individuals were then verified following the morphological identification procedures outlined in The Guide to the Identification of Seahorses (Lourie et al., 2004).

Total height per individual

Total height was measured for each seahorse to examine if current CITES recommendations of a 10-cm minimum size limit were being met and in what proportion of each U.S. end-market. Total height followed recommendations of Lourie et al. (2004).

With the exception of *H. zosterae,* which typically ranges in adult size of 1.6-3.8 cm (Strawn, 1958), specimens were measured laterally from the distance from the tip of the prehensile tail (posterior) to the top of the coronet (anterior) using a flexible measuring device to account for variations in shape.

Calculating the number of seahorses per kilogram for U.S. import estimates

To estimate the total number of individuals being imported into the U.S. we utilized CITES records from 2004-2012 (the last year of updated records). All data is publically available from the CITES Trade database (trade.cites.org) using the following search commands: Year Range: 2004-2012, Importing countries: United States of America, All Sources, All Purposes, All Terms, Taxon Genus: *Hippocampus,* Comparative Tabulations: Gross Imports. When units are available most dried seahorses are reported in kilograms (kg), grams, or number of specimens. To derive an estimate of the number of seahorses per kg from our TCM samples 133 whole individuals were weighed. We calculated a mean weight of 2.78 g/seahorse (individuals ranged from 1.03-5.25 g/seahorse) equivalent to 359.7 individuals per kg. This mean was then combined with estimates of individuals/per kg from previous research studies reported in Evanson et al. (2011) based on 11 independent studies. These estimates per region are as follows: Australia (3.00 g/seahorse), Latin America (Atlantic) (2.42 g/seahorse), Latin America (Pacific) (3.51 g/seahorse), Malaysia (3.18 g/seahorse), Thailand (3.13 and 3.30 g/seahorse), Philippines (3.33 and 1.38 g/seahorse), India (1.50 g/seahorse), Indonesia (2.00 g/seahorse) and Vietnam (2.86 g/seahorse). When combined with our mean value of g/seahorse this resulted in a global value of 370.4 individuals/kg.

Based on this value we estimated the total number of individuals per species imported into the U.S. since the *Hippocampus* CITES listing became implemented. Estimates of the total number of individuals imported into the U.S. was calculated for each species in this study, all records labeled as *Hippocampus spp*. were combined, and all other species were also combined as a single estimate (i.e., *Hippocampus other*: composed of *H. fuscus, H. erectus, H. histrix, H. reidi* and *H. whitei*) (Fig. 4.3a). The one exception was *H. kuda* that was reported in individuals (i.e., "bodies") therefore, no conversion estimates were required. For all other species we consider our estimates to be conservative because we only utilized data that specified units (i.e., kilograms, grams or specimens) and excluded records where units were not reported, were recorded as "live," or as "derivatives" because this designation may be products that contain seahorse material (i.e., pills and capsules) that can be purchased for medicinal use, though the content portion of *Hippocampus* is unknown. We also present the total number of individuals imported per year by combining all imports (Fig. 4.3b). Figures 4.2 and 4.3 were generated in Microsoft Excel 2011 or R v.3.0.3 (R-project.org).

RESULTS

We found substantial differences in the species composition between the specimens from Internet-souvenir retailers (ICR) and traditional Chinese medicine (TCM) markets. We also show a significant difference (T-test; p-value=2.2×10^{-16}) between average sizes of specimens available for purchase between these markets. The difference in size between markets was further illustrated in the boxplot (Fig. 4.2b) by the lack of notch overlap between groups. Based on the height (cm) distribution of TCM and

ICR specimens, a lack of notch (i.e., the 95% confidence interval (CI) around the median) overlap rejects the null hypothesis that the two medians are equal (Chambers et al., 1983), further indicating a significant difference between the size of the specimens within each trade. Almost all TCM samples (85.7% of 168 individuals) were >10-cm. Of the 14.3% of TCM specimens that were under 10-cm, the mean size was 8.89 cm (range; 7-9.6 cm) (Fig. 4.2a and 4.2b), whereas the mean size of TCM >10-cm (85.7%) was 12.9 cm (range: 10-19.9 cm). In contrast, only 4.8% of 204 Internet-curio samples were >10-cm (10-12.1 cm), with a mean size less <10-cm of 6.3 cm (range: 3.1-9.7) cm (Fig. 4.2a and 2b). These estimates exclude *H. zosterae*, due to its small average size of maturity (1.6-3.8 cm) (Strawn, 1958).

Five species were identified in the Internet-curio market. The most abundant species from the Internet-curio market were mostly juvenile specimens (<10-cm) of *H. kuda* (51.5%) as well as a large number of the domestic dwarf seahorse *H. zosterae* (40.7%). Eight species were identified from our TCM samples with *H. trimaculatus, H. spinosissimus* and *H. ingens* making up 86.3% of the samples. In addition, we identified a subset of individuals from a confiscated shipment to *H. trimaculatus* (75%; N=15) and *H. spinosissimus* (25%; N=5). The total numbers of individuals per species/per market are summarized in Table 4.1. For the geographic distributions of each species see Figure 4.1.

To estimate total import abundance based on reported CITES records from 2004-2012, we combined our estimate with previous values from 11 studies (Evanson et al., 2011) for a global value of 370.4 individuals per kg. Based on this estimate we found that clearly reported imports of dried seahorses represented roughly 302,359 individuals imported from 2004-2012. Reported imports were presented for each species found in this study (Fig. 4.3a), and also include imports labeled as *Hippocampus spp*. All species

not found in our samples but reported as species-specific imports were grouped as a separate annual estimate *Hippocampus other*. Based on import records it is clear that imports range dramatically from year-to-year and between species. We also find that two (*H. spinosissimus* and *H. ingens*) of the three most commonly identified species in the U.S. TCM market make up only a small portion of the reported imports (Table 4.1 and Fig. 4.3a).

DISCUSSION

Contrasts between IRC and TCM markets and species exploitation trends

While only a glimpse into the overall U.S. seahorse trade, our results demonstrate a significant contrast between U.S. TCM and ICR end-markets. Most strikingly, we show a highly significant difference in terms of the size of individuals being sold between markets ($p=2.2x10^{-16}$), and a clear difference in the composition of the most abundant species. Nearly all ICR samples were under the 10-cm size recommended for sustainable harvest, and *Hippocampus kuda* and *H. zosterae* totaled 93.2% of all ICR samples. In contrast, most TCM samples were >10-cm with 86.3% of samples identified as *H. trimaculatus, H. spinosissimus* or *H. ingens* (Table 1).

Two previous studies examining the dried seahorse species composition in the U.S. focused on California (Sanders et al., 2007) (N=46), and TCM samples collected across the U.S. (N=200) by the International Fund for Animal Welfare (IFAW) (Vincent et al., 2011b). We found a surprisingly similar composition of our TCM samples and the IFAW Report, of which the same three species - *H. trimaculatus, H. spinosissimus* and *H. ingens* - made up the majority of samples (Table 4.1). In contrast, *H. ingens* made up

nearly two-thirds of the California study's samples (Sanders et al. 2007). All three of these species are considered vulnerable due to their abundance in the global trade as well as their collection as bycatch (Baum and Vincent 2005; Murugan et al. 2008). Furthermore, the Indo-Pacific species *H. trimaculatus* and *H. spinosissimus* populations have declined recently by as much as 30% (Foster & Vincent, 2004; Meeuwig et al., 2006; Morgan & Panes, 2007; Perry et al., 2010). Heavy fishing pressure has also caused a significant decline in the eastern Pacific *H. ingens* (Baum and Vincent 2005) with Latin America reporting 1.5 tons of dried specimen exports (roughly 500,000 individuals) in 2004 alone (Sanders et al. 2007).

Where our results most differed with previous research was the high abundance of *H. kuda* and the dwarf seahorse *H. zosterae* identified from ICR samples. The Indo-Pacific *H. kuda* is heavily exploited in the international TCM trade and many populations are considered threatened due to habitat degradation (Lourie et al., 2004; IUCN, 2014). Whereas, the dwarf seahorse *H. zosterae* was not found in TCM samples, and previous research has not listed this species as part of the dried seahorse trade (Lourie et al., 2004; Vincent et al., 2011b). However, based on our results dried *H. zosterae* specimens are easily purchased from sellers in the U.S. through eBay.

Exploitation of the domestic dwarf seahorse *Hippocampus zosterae*

As previously noted, eBay agreed that the sale of imported species are no longer allowed on the site. As a domestic *H. zosterae* remains outside of this agreement. Though, we did find Indo-Pacific species were still available for purchase (nearly all identified as *H. kuda)*, by reducing the sale of imported species, there is a chance that this

may create a higher demand for domestic seahorses. Though, we cannot deduce from our research what impact this policy has on the high numbers of *H. zosterae* collected in our samples, unquestionably this dwarf seahorse species is now easily purchased online.

H. zosterae is primarily found in association with *Zostera* and other seagrass species and due to its small size may be particularly susceptible to overfishing. Typically, most species of *Hippocampus* release roughly100–300 young (Lourie et al., 1999, 2004); however, in comparison *H. zosterae* has a relatively low fecundity and some individuals produce as few as 5 young per brood (Masonjones & Lewis, 1996).

Previous research has not placed *H. zosterae* as a species in the dried seahorse trade (Lourie et al., 2004); however, it is a popular species in the global aquarium market (Vincent, 1996; Wood, 2001). Florida is the only state that keeps records of individuals collected, and *H. zosterae* occupied the 2nd rank of the top 10 aquarium fishes exported from Florida (Wood, 2001). Reported catch landings from 1991-2000 equaled 469,757 individuals, with a mean yearly landing of 46,976 (annual range: 7,226-98,779) (Vincent et al. 2011b). Although the impact of the ICR fishery of *H. zosterae* populations is unknown, our results clearly show that the sale of this species is no longer limited to the aquarium industry. The most recent IUCN Red List update of this species was in 2003, and is currently listed as in "need of updating." Our results demonstrate a clear expansion in scope of this species exploitation, and may help to facilitate future management efforts and inform its conservation status.

Adherence of trades to CITES size limit recommendations

The 10-cm minimum height limit recommendation for seahorses by the CITES Technical Committee on Animals was based on the findings of Foster and Vincent (2005). The time of reproduction varies by species, therefore, the maximum height and size at first maturity was gathered for 32 seahorse species. They concluded that 15 of the 16 most traded species would be given some protection if this size standard was implemented by allowing individuals the potential to reproduce before being removed from the wild. The 167 signatory nations agreed to adhere to recommending this size limit to fishers to create a more sustainable seahorse trade.

Three of the species identified in our samples, *H. kelloggi, H. ingens* and *H. kuda,* may reach reproduction after 10-cm in height. While most TCM samples were >10-cm, the high abundance of juvenile individuals <10-cm in the Internet-curio market were mostly composed of *H. kuda* (93% of individuals excluding the dwarf species) (Fig 4.2a). While some fishers target seahorses directly, most seahorses in the international trade are still collected by non-selective fishing bycatch (Foster & Vincent, 2005; Vincent et al., 2011a). This form of fishing tends to be non-discriminatory in size selection and recommendations have been made to adapt sustainable methods to help combat the use of non-selective fishing gear for *Hippocampus* conservation (Bruckner et al., 2005). The international trade of seahorses is voluntarily regulated by the CITES signatory nations with implementation based on per-country management efforts. Therefore it is difficult to know what impact the CITES size minimum recommendation has on the adherence of TCM samples to meet 10-cm size limit. Though a difficult task, the most notable success has been carried out in the Philippines by Project Seahorse and their continued

partnerships with local communities to implement sustainable long-term fishery benefits (seahorse.fisheries.ubc.ca).

U.S. import abundance based on CITES records and confiscations

Vincent et al. (2011a) discussed the challenges faced in using CITES data to analyze trade in seahorses, but concluded that it "gives us an unparalleled tool to investigate the trade in seahorses and other listed species." Because of the public availability of all CITES records, we were able to use this data to estimate the number of individuals imported into the U.S. from 2004-2012. Based on a rough estimate, our results indicate that 302,359 dried individuals were imported into the U.S. from 2004-2012 (approximately 816 kg) using a mean estimate of 370.4 individuals per kg (See Materials and Methods for details).

Previous reports estimate the number of individuals per kg ranged dramatically per region (from 1.38 to 3.51 g/seahorse) (Marichamy et al., 1993; Vincent & Hall, 1996; Evanson et al., 2011). Therefore, if the upper and lower reported estimates were used instead of our global mean value the number of individuals imported into the U.S. would range from 232,645 to 591,818. Because of the variance of regional size estimates we emphasize that our results are only an approximation of number of individuals and are conservative due to the exclusion of any ambiguous import records. However, despite the uncertainty in this estimate, this information is still a useful benchmark for understanding relative abundance of dried seahorses being sold in the U.S. Previous estimates, prior to CITES, from the United States Department of the Interior (Fish and Wildlife Service) reported 755 kg imported from 1996-2000, imported primarily from the Philippines

(60%), China (30%) and Mexico (7%). In general, Asian countries remain the main exporters of seahorses (specifically, Thailand, mainland China and Hong Kong), though substantial new exports from Guinea, and Latin America have ben documented (Vincent et al., 2011b).

Based on our analysis of available CITES records it is clear that the total number of imported individuals into the U.S. varies dramatically between species (Fig. 4.3a) as well as from year to year (Fig. 4.3b). The Indo-Pacific species *H. trimaculatus, H. kuda* and *H. kelloggi* are the most reported species, while *H. barbouri, H. ingens,* and *H. spinosissimus* appear to make up only a small portion of reported U.S. imports. This result was somewhat surprising given that across studies *H. ingens* and *H. spinosissimus* are both consistently among the most abundant species found in U.S. TCM samples (Table 4.1 and Fig. 4.3a). Among species identified in our samples only the species *H. comes* remained unreported to CITES and specimens are likely shipped with other reported species, misidentified, or imported as unidentified (i.e., *Hippocampus spp*.). For example, of the 827 available import records in the CITES database 39.8% (329 records) were confiscated shipments, with the vast majority (85.1%, 280 records) recorded as *Hippocampus spp*. Of confiscated imports, 58.% of (192 records) were dried specimens, 31.9% (105 records) "derivatives," and 9.7% (32 records) "live" specimens.

The low abundance of reported *H. ingens* and *H. spinosissimus* may be due to undesignated species status (*Hippocampus spp*.) or unreported imports. For example, as an additional component of our research we sequenced a subset of individuals (N=20) from a confiscated shipment into the port of New York (*Hippocampus spp*.). Of these, 75% were identified as *H. trimaculatus* and 25% *H. spinosissimus*; again the two most abundant species found in our TCM samples (Table 4.1). Government agencies in the

U.S. are now utilizing molecular labs to identify potentially invasive or unidentified species, food products, and this work demonstrates the potential for this practice to work in tandem with CITES to validate the species composition of confiscated shipments.

Conclusions and recommendations

Our results demonstrate a significant contrast in both the size of individuals and species composition between U.S. TCM and ICR markets. Regarding specimen availability, TCM sellers do not appear to utilize online platforms for retail purposes and sales remain limited to urban metropolitan areas, whereas coastal souvenir stores are now linked in many cases to online commerce through eBay auctions as well as independent websites. Based on the prevalence and availability of the domestic dwarf seahorse *H. zosterae*, it is clear that this species is no longer only exploited for the live aquarium trade. Therefore, our findings may help inform its current Red List evaluation and status.

Size comparisons show that many individuals are still collected before being able to reach a size of first reproduction, indicating that a continued effort be made by signatory nations to emphasize the long-term benefit of a sustainable harvest to fishers by monitoring and reducing non-target bycatch (Bruckner et al., 2005). In addition, our results demonstrate that CITES is a valuable resource for helping to understand the volume and type of species being imported into the U.S.

Despite some criticism (Collins & Cruickshank, 2013), our study further demonstrates the utility of DNA barcoding as promising and adaptable tool for training students to tackle molecular biology research of wildlife trade products and gain experience with hands-on scientific investigation. For those interested in conducting this

form of research we have included a protocol, workflow, and information pointing to additional curricula and informational materials (Appendix 4.1). Because of the general utility of molecular wildlife identification by non-experts using DNA barcoding, moving forward studies such as ours may be a potential avenue to retroactively add valuable information to CITES import records, analyze random subsets of wildlife imports and exports, and further verify trade documentation that contributes to the evaluation of heavily exploited species.

Species	2010 TCM Philadelphia	2009-2013 TCM: New York	2011-2014 ECR ecommerce stores	2013 ECR Coastal stores: Texas	2013 Confiscated shipment	Sanders et al. (2007) TCM: California	Sanders et al. (2007) curio: California	2000 IFAW Report TCM	Total across reports
Hippocampus ingens (Girard 1859)	8	28	0	0	0	29	0	32	**97**
H. kelloggi (Jordan & Snyder 1902)	4	6	0	0	0	0	0	21	**31**
H. trimaculatus (Leach 1814)	3	47	0	11	15	3	1	58	**138**
H. spinosissimus (Weber 1913)	10	28	0	2	5	6	0	42	**93**
H. comes (Cantor 1850)	0	6	0	0	0	4	0	11	**21**
H. kuda (Bleeker 1852)	0	23	177	0	0	0	0	13	**213**
H. algiricus (Kaup 1856)	0	4	0	0	0	0	0	0	**4**
H. zosterae (Jordan & Gilbert 1882)	0	0	123	17	0	0	0	0	**140**
H. barbouri (Jordan & Richardson 1908)	0	1	0	14	0	0	9	5	**29**
H. spp.	0	0	0	0	0	2*	0	24**	**26**
Total	**N = 25**	**N = 143**	**N = 300**	**N = 44**	**N = 20**	**N = 44**	**N = 10**	**N = 206**	**N = 792**

Table 4.1 Species composition per market, region, and source. Species composition of samples collected for this study from TCM and Internet-curio stores and vendors. The composition of two previous studies Sanders et al. (2007) and International Fund for Animal Welfare (IFAW) (2000), presented in Vincent et al. (2011b). *Sander et al. (2007), reported one specimen that could not be distinguished between *H. algiricus* and *H. reidi*, and a single specimen of *H. hippocampus* (row *H. spp.*). ** The 24 individuals in the IFAW Report (row *H. spp.*) are composed of the following: *H. sp.* (N=11), *H. reidi* (N=5), *H. angustus* (N=5), *H. erectus* (N=3), *H. whitei* (N=1), *H. fuscus* (N=1). IFAW samples were collected from Chinatowns in Boston, MA., New York, NY., Washington, DC., Los Angeles, Oakland, and San Francisco, CA.

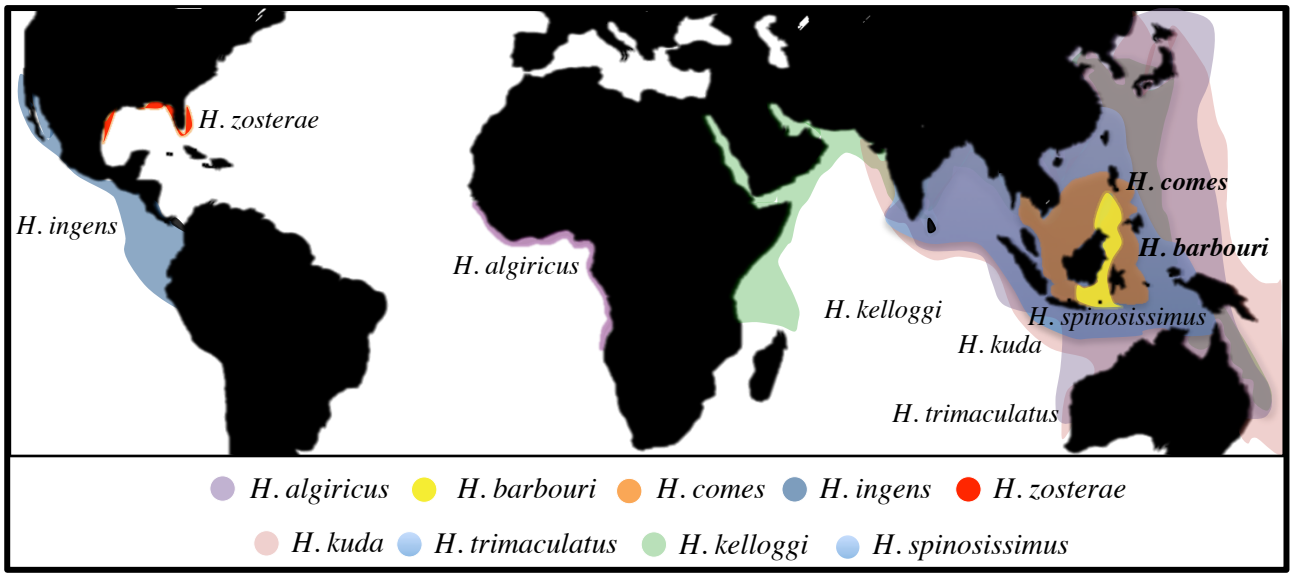


Figure 4.1 Map of seahorse species distributions identified in this study. Distributions of each species identified in this study based on Lourie et al. (2004) and verified from data points from Ocean Biographic Information Systems (OBIS: www.iobis.org).

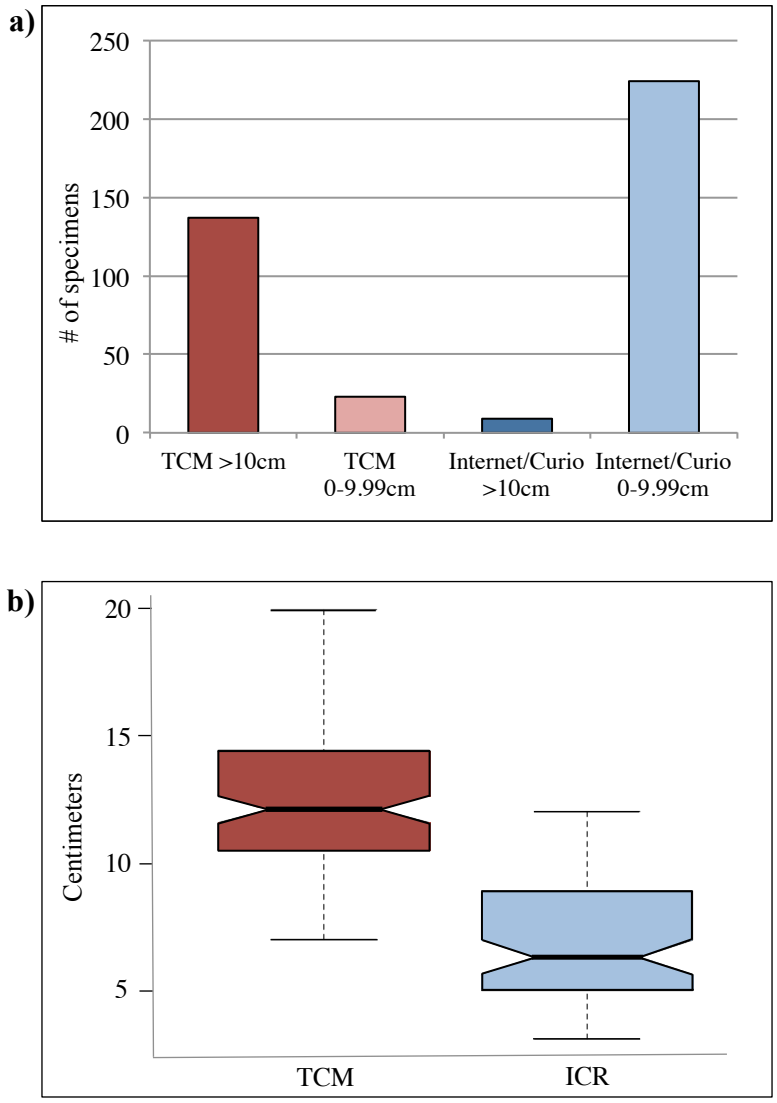


Figure 4.2 Size range and comparison of seahorse samples per end-market. (a) Specimens were divided into two categories >10-cm and < 10-cm for each market. Histogram values are presented in number of individuals sampled per category. (b) Samples of the dwarf seahorse *H. zosterae* were removed due to their small size at reproduction (2-3 cm). Boxplots illustrating the distribution of size variance across individuals per category between all TCM and ICR samples combined. The notches represent the 95% CI around the median for each group. The lack of notch overlap indicates a significant difference between groups.

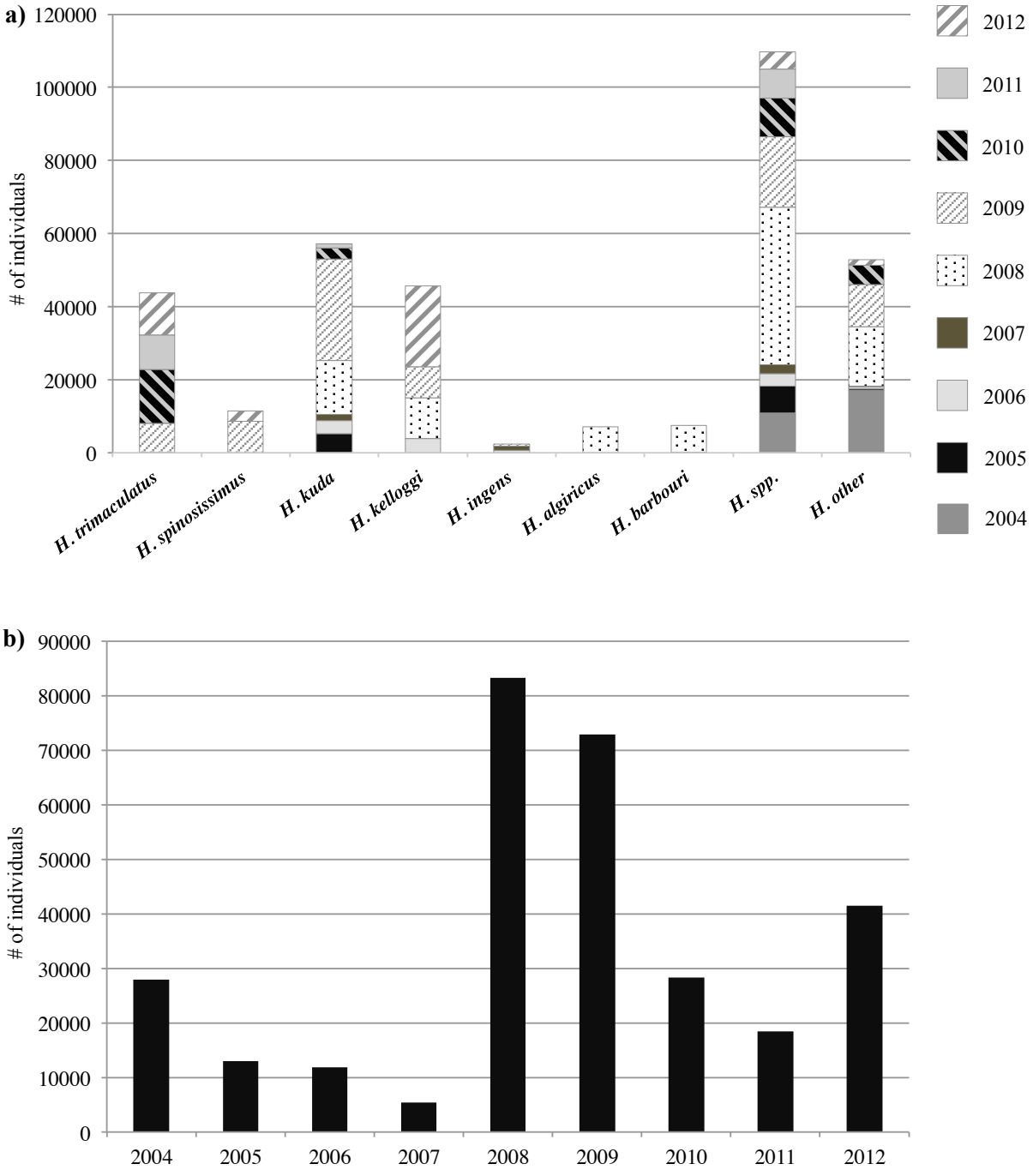


Figure 4.3 Estimates of seahorse import abundance per year derived from CITES import records. (a) Histogram of estimated individuals per species from 2004-2012. (b) Annual estimate of imported individuals.

APPENDIX 1.1

Appendix 1.1 (a) PCR and sequencing methods, (b) Samples, collection locations, collectors, and GenBank accession numbers.

Appendix 1.1a Mitochondrial markers sequenced include: cytochrome *b* (*cytb*) (Lourie et al., 2005; Woodall et al., 2011), cytochrome *c* oxidase subunit I (*COI*) (Folmer et al., 1994), and control region (D-loop) (Teske et al., 2003). All PCR reactions occurred in a total volume of 25 μL that included 2.5 μL of Fisher PCR 10× Buffer A, 0.05 mM of each dNTP, 0.5 U of *Taq* DNA Polymerase (Fisher), 0.5 μL of 10 μM each primer, and 1 μL of (10–25 ng) of DNA. The amplification protocol for PCR reactions of mtDNA was: 94 °C for 5 min; 35–39 cycles of 94 °C for 30 s, annealing temperature 50–52 °C for 1 min, 72 °C for 1 min; final extension 72 °C for 10 min; final rest at 4 °C. The amplification protocol for PCR reactions of nuclear loci was: 94 °C for 5 min; 35 cycles of 94 °C for 30 s, annealing temperature 50–62 °C for 45 s to 1 min (Teske & Beheregaray, 2009), 72 °C for 1 min; final extension 72 °C for 10 min; final rest at 4 °C. Nuclear DNA (nDNA) comprised five nuclear loci (aldolase, *myh6*, rhodopsin, *Tmo4c4* and a portion of the S7 ribosomal protein containing the first intron) (Teske & Beheregaray, 2009). The loci rhodopsin and *myh6* required nested PCR amplification (Teske & Beheregaray, 2009). We visualized PCR products on 1.5% agarose gels stained with SYBR Safe. Successful amplifications were purified with Agencourt AMPure XP, sequenced with BigDye version 3.1 chemistry (Applied Biosystems, Foster City, CA, USA), and sequenced using an ABI3730 Genetic Analyzer (ABI). All sequences were aligned with Geneious Alignment in GENEIOUS PRO 5.1.4 (Drummond et al., 2011).

Appendix 1.1b Samples used in this study, including collection locations (when applicable), collectors and/or museums, and GenBank accession numbers. Species IDs correspond to the following: HI, *Hippocampus ingens*; HZ, *H. zosterae*; HG, *H. guttulatus*; HA, *H. algiricus*; HC, *H. capensis*; HR, *H. reidi*; HK, *H. kelloggi*; HE, *H. erectus*; HP, *H. patagonicus*; HH, *H. hippocampus* (Europe); HHS, *H. hippocampus* (West Africa). The number of samples in the *H. erectus* complex include Caribbean/North America (*n* = 114), South America (*n* = 20), Europe (*n* = 229), and West Africa (*n* = 26) from a total of 15 populations.— = no data.

ID	Location	Collector	*cytb*	Control region	*COI*	Aldolase	*S7*	*Myh6*	Rhodopsin	*Tmo4c4*
HI1	Eastern Pacific (EP)	Teske, P./ GenBank	DQ288345	DQ288332	—	AY277365	AY277334	FJ905780	FJ905791	FJ905802
HI2	EP	Teske, P.	DQ288346	DQ288344	—	—	AY277333	KC811914	KC811902	—
HI3	EP	Teske, P.	AF192674	AY642329	KC851895	—	DQ288383	—	—	—
HI4	EP	GenBank	AF192673	KC851931	KC851894	—	—	—	—	—
HZ1	Florida	GenBank	AF192706	DQ288328	GQ502176	AY277371	DQ288380	FJ905781	—	FJ905803
HZ2	Florida	Baldwin, C.	KC851875	KC851927	KC851889	KC812043	KC811887	KC811912	KC811899	KC811849
HZ3	Florida	Dunham, N.	KC851876	KC851928	KC851890	KC812043	KC811888	KC811912	KC811900	KC811850
HG1	Europe	Woodall, L.	KC851873	KC851925	KC851886	KC812039	KC811885	KC811911	KC811897	KC811848
HG2	Europe	Woodall, L.	KC851874	KC851926	KC851887	KC812039	KC811886	KC811911	KC811898	KC811848
HG3	N/A	GenBank	EU547203	DQ288322	GQ502143	AY277361	AY277337	—	—	—
HA1	West Africa	Teske, P.	DQ288352	DQ288337	GQ502119	AY277366	AY277328	—	—	—
HA2	West Africa	Teske, P.	DQ288353	DQ288338	—	—	—	—	—	—
HC1	South Africa	Teske, P.	AF192652	AY149678	GQ502129	AY277357	AY277331	—	—	—
HC2	South Africa	Teske, P.	AF192650	AY149676	GQ502128	—	—	—	—	—
HR1	N/A	GenBank	DQ288349	DQ288335	GQ502165	AY277367	DQ288385	FJ905779	FJ905790	FJ905801
HR2	Florida	Baldwin, C.	KC851877	KC851929	KC851891	KC812044	KC811889	KC811913	KC811901	KC811851
HR3	N/A	GenBank	DQ288347	DQ288386	—	—	DQ288381	—	—	—
HR4	N/A	GenBank	DQ288343	DQ288384	—	—	—	—	—	—
HK1	N/A	GenBank	AF192675	AY629249	GQ502151	AY277351	AY277326	—	—	—
HK2	N/A	GenBank	AF192677	KC851930	GQ502149	AY277350	AY277325	—	—	—
HK3	N/A	GenBank	AF192676	—	GQ502150	—	—	—	—	—
HK4	N/A	GenBank	AF192678	—	KC851892	—	—	—	—	—
HE1	New York, USA (NY)	Boehm, J.	KC811915	—	—	—	—	—	—	—
HE2	NY	Boehm, J.	KC811916	—	—	—	—	—	—	—
HE3	NY	Boehm, J.	KC811917	—	—	—	—	—	—	—
HE4	NY	Boehm, J.	KC811918	—	—	—	—	—	—	—

HE5	NY	Boehm, J.	KC811919	—	—	—	—	—	—	—
HE6	NY	Boehm, J.	KC811920	—	—	—	—	—	—	—
HE7	NY	Boehm, J.	KC811921	—	—	—	—	—	—	—
HE8	NY	Boehm, J.	KC811922	—	—	—	—	—	—	—
HE9	NY	Boehm, J.	KC811923	—	—	—	—	—	—	—
HE10	NY	Boehm, J.	KC811924	—	—	—	—	—	—	—
HE11	NY	Boehm, J.	KC811925	—	—	—	—	—	—	—
HE12	NY	Boehm, J.	KC811926	—	—	—	—	—	—	—
HE13	NY	Boehm, J.	KC811927	—	—	—	—	—	—	—
HE14	NY	Boehm, J.	KC811928	—	—	—	—	—	—	—
HE15	NY	Boehm, J.	KC811929	—	—	—	—	—	—	—
HE16	NY	Boehm, J.	KC811930	KC851898	KC851878	KC812033	KC811853	—	KC811891	KC811838
HE17	NY	Boehm, J.	KC811931	—	—	KC812030	KC811860	—	—	—
HE18	NY	Gardner, T.	KC811932	—	—	—	—	—	—	—
HE19	NY	Gardner, T.	KC811933	—	—	—	—	—	—	—
HE20	NY	Gardner, T.	KC811934	KC851896	—	KC812030	KC811865	—	KC811893	—
HE21	NY	Gardner, T.	KC811935	—	—	KC812030	—	—	—	—
HE22	NY	Gardner, T.	KC811936	—	—	KC812033	—	—	—	—
HE23	NY	Gardner, T.	KC811937	—	—	—	—	—	—	—
HE24	NY	Gardner, T.	KC811938	—	—	—	—	—	—	—
HE25	NY	Gardner, T.	KC811939	—	—	—	—	—	—	—
HE26	NY	Gardner, T.	KC811940	—	—	—	—	—	—	—
HE27	NY	Gardner, T.	KC811941	—	—	—	—	—	—	—
HE28	NY	Gardner, T.	KC811942	—	—	—	—	—	—	—
HE29	NY	Gardner, T.	KC811943	—	—	—	—	—	—	—
HE30	NY	Gardner, T.	KC811944	—	—	—	—	—	—	—
HE31	NY	Gardner, T.	KC811945	—	—	—	—	—	—	—
HE32	New Jersey, USA (NJ)	Grotues, T.	KC811946	—	—	—	—	—	—	—
HE33	NJ	Grotues, T.	KC811947	—	—	KC812030	KC811866	—	—	—
HE34	NJ	Grotues, T.	KC811948	—	—	KC812030	KC811869	—	—	—
HE35	NJ	Grotues, T.	KC811949	—	—	—	—	—	—	—
HE36	NJ	Grotues, T.	KC811950	—	—	—	—	—	—	—
HE37	NJ	Grotues, T.	KC811951	—	—	—	—	—	—	—
HE38	NJ	Grotues, T.	KC811952	—	—	—	—	—	—	—
HE39	NJ	Grotues, T.	KC811953	—	—	—	—	—	—	—
HE40	NJ	Grotues, T.	KC811954	—	—	—	—	—	—	—

HE41	NJ	Grotues, T.	KC811955	—	—	—	—	—	—	—
HE42	NJ	Grotues, T.	KC811956	—	—	—	—	—	—	—
HE43	NJ	Grotues, T.	KC811957	—	—	—	—	—	—	—
HE44	NJ	Grotues, T.	KC811958	—	—	—	—	—	—	—
HE45	NJ	Grotues, T.	KC811959	—	—	—	—	—	—	—
HE46	Chesapeake Bay, USA (CHS)	Virginia Institute of Marine Science (VIMS)	KC811960	—	—	—	—	—	—	—
HE47	CHS	VIMS	KC811961	—	—	—	—	—	—	—
HE48	CHS	VIMS	KC811962	—	—	—	—	—	—	—
HE49	CHS	VIMS	KC811963	—	—	—	—	—	—	—
HE50	CHS	VIMS	KC811964	—	—	—	—	—	—	—
HE51	CHS	VIMS	KC811965	—	—	—	—	—	—	—
HE52	CHS	VIMS	KC811966	—	—	—	—	—	—	—
HE53	CHS	VIMS	KC811967	—	—	KC812030	KC811855	—	—	
HE54	CHS	VIMS	KC811968	—	—	—	—	—	—	—
HE55	CHS	VIMS	KC811969	—	—	—	—	—	—	—
HE56	CHS	VIMS	KC811970	—	—	—	—	—	—	—
HE57	CHS	VIMS	KC811971	—	—	—	—	—	—	
HE58	CHS	VIMS	KC811972	—	—	—	—	—	—	—
HE59	CHS	VIMS	KC811973	—	—	—	—	—	—	—
HE60	CHS	VIMS	KC811974	KC851900	KC851878	KC812030	KC811858	KC811903	KC811891	KC811839
HE61	CHS	VIMS	KC811975	—	—	—	—	—	—	—
HE62	CHS	VIMS	KC811976	—	—	—	—	—	—	—
HE63	CHS	VIMS	KC811977	—	—	KC812030	KC811856			
HE64	CHS	VIMS	KC811978	—	—	—	—	—	—	—
HE65	CHS	VIMS	KC811979	—	—	—	—	—	—	—
HE66	CHS	VIMS	KC811980			—	KC811864			
HE67	CHS	VIMS	KC811981	KC851898	KC851879	KC812033	KC811857	KC811904	KC811893	KC811840
HE68	Eastern Gulf of Mexico (GULF)	Dunham, N.	KC811982	—	—	—	—	—	—	—
HE69	GULF	Dunham, N.	KC811983	KC851897	KC851880	KC812030	—	KC811904	KC811891	KC811841
HE70	GULF	Dunham, N.	KC811984	—	—	—	—	—	—	—
HE71	GULF	Dunham, N.	KC811985	—	—	—		—	—	—
HE72	GULF	Dunham, N.	KC811986	—	—	—	—	—	—	—
HE73	GULF	Dunham, N.	KC811987	—	—	—	—	—	—	—

HE74	GULF	Dunham, N.	KC811988	—	—	—	—	—	—	—
HE75	GULF	Dunham, N.	KC811989	—	—	—	—	—	—	—
HE76	GULF	Dunham, N.	KC811990	—	—	—	—	—	—	—
HE77	GULF	Dunham, N.	KC811991	—	—	—	—	—	—	—
HE78	GULF	Dunham, N.	KC811992	—	—	KC812035	—	—	—	—
HE79	GULF	Dunham, N.	KC811993	KC851901	KC851879	KC812035	KC811870	KC811904	KC811892	—
HE80	GULF	Dunham, N.	KC811994	—	—	—	—	—	—	—
HE81	GULF	Dunham, N.	KC811995	—	—	—	—	—	—	—
HE82	GULF	Dunham, N.	KC811996	—	—	—	—	—	—	—
HE83	GULF	Dunham, N.	KC811997	—	—	—	—	—	—	—
HE84	GULF	Dunham, N.	KC811998	—	—	—	—	—	—	—
HE85	GULF	Dunham, N.	KC811999	—	—	—	—	—	—	—
HE86	GULF	Dunham, N.	KC812000	—	—	—	—	—	—	—
HE87	GULF	Dunham, N.	KC812001	—	—	—	—	—	—	—
HE88	GULF	Dunham, N.	KC812002	—	—	—	—	—	—	—
HE89	GULF	Dunham, N.	KC812003	—	—	KC812034	—	—	—	—
HE90	GULF	Dunham, N.	KC812004	—	—	—	—	—	—	—
HE91	GULF	Dunham, N.	KC812005	—	—	—	—	—	—	—
HE92	GULF	Dunham, N.	KC812006	—	—	—	—	—	—	—
HE93	GULF	Dunham, N.	KC812007	—	—	KC812033	—	—	—	—
HE94	GULF	Dunham, N.	KC812008	—	—	—	—	—	—	—
HE95	GULF	Dunham, N.	KC812009			KC812030	KC811854	—	—	—
HE96	GULF	Dunham, N.	KC812010			—	—	—	—	—
HE97	Florida, Atlantic (FLAT)	Dunham, N.	KC812011	KC851902	KC851880	KC812030	KC811863	—	—	—
HE98	FLAT	Dunham, N.	KC812012	—	—	KC812030	KC811868	—	—	—
HE99	FLAT	Dunham, N.	KC812013	—	—	—	—	—	—	—
HE100	FLAT	Dunham, N.	KC812014	—	—	—	—	—	—	—
HE101	FLAT	Dunham, N.	KC812015	—	—	—	—	—	—	—
HE102	FLAT	Dunham, N.	KC812016	—	—	—	—	—	—	—
HE103	FLAT	Dunham, N.	KC812017	—	—	KC812033	KC811867	—	—	—
HE104	FLAT	Dunham, N.	KC812018	—	—	KC812032	KC811862	—	—	—
HE105	FLAT	Dunham, N.	KC812019	KC851903	—	KC812030	KC811859	—	—	—
HE106	FLAT	Dunham, N.	KC812020	—	—	KC812031	KC811861	—	—	—

HE107	Caribbean Sea, Nicaragua (NIC)	Baum, J./ Redpath Museum, McGill University (RP)	KC812021	KC851899	KC851881	—	KC811852	—	—	KC811841
HE108	NIC	Baum, J./RP	KC812022	—	--	—	—	—	—	—
HE109	NIC	Baum, J./RP	KC812023	—	—	—	—	—	—	—
HE110	Caribbean Sea, Honduras (HON)	Baum, J./RP	KC812024	—	—	—	—	—	—	—
HE111	HON	Baum, J./RP	KC812025	—	—	—	—	—	—	—
HE112	Caribbean Sea, Mexico (MEX)	Baum, J./RP	KC812026	—	—	—	—	—	—	—
HE113	Caribbean Sea, Mexico (MEX)	Baum, J./RP	KC812027	—	—	—	—	—	—	—
HE114	MEX	Baum, J./RP	KC812028	—	—	—	—	—	—	—
HE115	MEX	Baum, J./RP	KC812029	—	—	—	—	—	—	—
HECB	Colombia	Teske, P.	DQ288341	—	—	—	—	—	—	—
HP1	Argentina	Lopez, A., RP-RM2858b	HM447028	—	KC851885	KC812042	—	—	—	—
HP2	Argentina	Lopez, A., RP - RM2859c	HM447026	KC851923	KC851885	KC812042	KC811871	—	—	—
HP3	Argentina	Lopez, A., RP - RM2859d	HM447016	KC851924	KC851885	KC812042	KC811874	KC811909	KC811894	KC811842
HP4	Argentina	Lopez, A., RP - RM2859b	HM447027	KC851920	KC851885	KC812042	KC811875	KC811910	KC811894	KC811842
HP5	Argentina	Lopez, A., RP - RM2858a	HM447025	KC851918	KC851885	KC812042	KC811873	—	—	—
HP6	Argentina	Lopez, A., RP - RM2859a	HM447015	—	—	—	—	—	—	—
HP7	Mar de Plata, Argentina (MPA)	Luzzatto,D.C	EU871944	—	—	—	—	—	—	—
HP8	MPA	Luzzatto,D.C	EU871944	—	—	—	—	—	—	—
HP9	MPA	Luzzatto,D.C	EU871944	—	—	—	—	—	—	—
HP10	MPA	Luzzatto,D.C	EU871944	—	—	—	—	—	—	—
HP11	San Antonio, Argentina (SAN)	Luzzatto,D.C	EU871945	—	—	—	—	—	—	—
HP12	SAN	Luzzatto,D.C	EU871945	—	—	—	—	—	—	—
HP13	SAN	Luzzatto,D.C	EU871945	—	—	—	—	—	—	—
HP14	SAN	Luzzatto,D.C	EU871945	—	—	—	—	—	—	—
HP15	Brazil	Rosa, I., RP – IR19	HM447024	KC851917	KC851885	KC812040	KC811876	KC811909	KC811894	KC811842
HP16	Brazil	Rosa, I., RP – IR10	HM447022	KC851919	—	KC812041	KC811872	—	—	—-
HP17	Brazil	Rosa, I., RP – IR1	HM447020	KC851921	—	KC812042	—	—	—	—

HP18	Brazil	Rosa, I., RP – IR3	HM447021	KC851922	—	KC812042	—	—	—	—
HP19	Brazil	Rosa, I., RP – IR43	HM447019	—	—	—	—	—	—	—
H. cf. erectus	Brazil	Casey, S.P./ GenBank	AF192660	—	—	—	—	—	—	—
HH/AH1	Ria Formosa, Portugal (RPO)	Woodall, L.	HQ437202	—	—	—	—	—	—	—
HH/AH2	RPO	Woodall, L.	HQ437207	—	—	—	—	—	—	—
HH/AH3	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH4	RPO	Woodall, L.	HQ437202	—	—	—	—	—	—	—
HH/AH5	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH6	RPO	Woodall, L.	HQ437199	—	—	—	—	—	—	—
HH/AH7	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH8	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH9	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH10	RPO	Woodall, L.	HQ437201	—	—	—	—	—	—	—
HH/AH11	RPO	Woodall, L.	HQ437203	—	—	—	—	—	—	—
HH/AH12	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH14	RPO	Woodall, L.	HQ437203	—	—	—	—	—	—	—
HH/AH15	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH16	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/AH22	RPO	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/AH23	RPO	Woodall, L.	HQ437215	—	—	—	—	—	—	—
HH/AH24	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/BH1	RPO	Woodall, L.	HQ437219	—	—	—	—	—	—	—
HH/BH2	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/BH3	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/BH4	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/PH2	RPO	Woodall, L.	KC851870	—	—	KC812036	KC811877	—	—	—
HH/PH3	RPO	Woodall, L.	KC851871	KC851908	KC851882	KC812036	KC811884	KC811906	—	—
HH/PH5	RPO	Woodall, L.	KC851872	—	—	KC812036	KC811877	—	—	—
HH/DH1	Portimao, Portugal (PPO)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/DH2	PPO	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/DH3	PPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/DH4	PPO	Woodall, L.	HQ437202	—	—	—	—	—	—	—
HH/DH5	PPO	Woodall, L.	HQ437197	—	—	—	—	—	—	—
HH/DH6	PPO	Woodall, L.	HQ437202	—	—	—	—	—	—	—

HH/GH1	South Coast, UK (AUK)	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH2	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH4	AUK	Woodall, L.	HQ437209	—	—	—	—	—	—	—
HH/GH6	AUK	Woodall, L.	HQ437208	—	—	—	—	—	—	—
HH/GH7	AUK	Woodall, L.	HQ437204	—	—	—	—	—	—	—
HH/GH8	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH9	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH11	AUK	Woodall, L.	HQ437200	—	—	—	—	—	—	—
HH/GH18	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH22	AUK	Woodall, L.	HQ437221	—	—	—	—	—	—	—
HH/GH23	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH26	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH27	AUK	Woodall, L.	HQ437232	—	—	—	—	—	—	—
HH/GH29	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH30	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH31	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH32	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH33	AUK	Woodall, L.	HQ437220	—	—	—	—	—	—	—
HH/GH34	AUK	Woodall, L.	HQ437218	—	—	—	—	—	—	—
HH/GH37	AUK	Woodall, L.	HQ437238	—	—	—	—	—	—	—
HH/GH38	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH39	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH40	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH41	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/GH42	AUK	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/GH43	AUK	Woodall, L.	HQ437239	KC851905	KC851882	KC812036	—	—	KC811895	KC811843
HH/GH44	AUK	Woodall, L.	HQ437206	KC851906	KC851883	KC812036	KC811878	—	—	KC811843
HH/GH45	AUK	Woodall, L.	HQ437206	—	—	KC812036	KC811884	—	—	—
HH/GH46	AUK	Woodall, L.	HQ437240	—	—	—	—	—	—	—
HH/GH47	AUK	Woodall, L.	HQ437241	—	—	—	—	—	—	—
HH/GH65	AUK	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH2	Israel (ISR)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH3	ISR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH4	ISR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH5	ISR	Woodall, L.	HQ437208	—	—	—	—	—	—	—
HH6	ISR	Woodall, L.	HQ437224	—	—	—	—	—	—	—

HH7	ISR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH8	Porto Santo (PMI)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH9	PMI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH1	Channel Islands (JCI)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH2	JCI	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/IH3	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH4	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH5	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH6	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH7	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/IH8	JCI	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH1	Malaga, Spain (MSP)	Woodall, L.	HQ437207	—	—	—	—	—	—	—
HH/JH2	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH8	MSP	Woodall, L.	HQ437210	—	—	—	—	—	—	—
HH/JH9	MSP	Woodall, L.	HQ437204	—	—	—	—	—	—	—
HH/JH10	MSP	Woodall, L.	HQ437211	—	—	—	—	—	—	—
HH/JH11	MSP	Woodall, L.	HQ437202	—	—	—	—	—	—	—
HH/JH12	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH13	MSP	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/JH14	MSP	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/JH15	MSP	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/JH16	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH17	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH18	MSP	Woodall, L.	HQ437198	KC851909	—	KC812036	KC811882	KC811906	KC811895	—
HH/JH19	MSP	Woodall, L.	HQ437198	KC851909	—	KC812037	KC811882	—	—	—
HH/JH20	MSP	Woodall, L.	HQ437242	—	—	KC812038	—	—	—	—
HH/JH21	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH22	MSP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/JH23	MSP	Woodall, L.	HQ437221	—	—	—	—	—	—	—
HH/LH1	France (SFR)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH3	SFR	Woodall, L.	HQ437226	—	—	—	—	—	—	—
HH/LH4	SFR	Woodall, L.	HQ437213	—	—	—	—	—	—	—
HH/LH5	SFR	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/LH6	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH7	SFR	Woodall, L.	HQ437224	—	—	—	—	—	—	—

HH/LH8	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH9	SFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/LH10	SFR	Woodall, L.	HQ437203	—	—	—	—	—	—	—
HH/LH11	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH12	SFR	Woodall, L.	HQ437210	—	—	—	—	—	—	—
HH/LH13	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH14	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH15	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH16	SFR	Woodall, L.	HQ437242	—	—	—	—	—	—	—
HH/LH18	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH19	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/LH20	SFR	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/LH21	SFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/MH1	La Rochelle, France (RFR)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/MH2	RFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/MH3	RFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/MH4	RFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/MH6	RFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/MH7	RFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/MH8	RFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NH1	Brest, France (BFR)	Woodall, L.	HQ437198	KC851907	KC851883	KC812036	KC811881	KC811906	KC811895	KC811846
HH/NH2	BFR	Woodall, L.	HQ437198	—	—	KC812036	KC811884	—	—	—
HH/NH3	BFR	Woodall, L.	HQ437198	—	KC851883	KC812036	KC811884	—	—	—
HH/NH4	BFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NH5	BFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NH6	BFR	Woodall, L.	HQ437236	—	—	—	—	—	—	—
HH/NH7	BFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NH8	BFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH1	Arcachon, France (AFR)	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH2	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH3	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH4	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH5	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH8	AFR	Woodall, L.	HQ437221	—	—	—	—	—	—	—

HH/OH9	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH10	AFR	Woodall, L.	HQ437221	—	—	—	—	—	—	—
HH/OH11	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH12	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH13	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH14	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/OH15	AFR	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/QH1	Napoli, Italy (NIT)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH2	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH3	NIT	Woodall, L.	HQ437242	—	—	—	—	—	—	—
HH/QH4	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH5	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH6	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH7	NIT	Woodall, L.	HQ437223	—	—	—	—	—	—	—
HH/QH8	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/QH10	NIT	Woodall, L.	HQ437225	—	—	—	—	—	—	—
HH/QH11	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/RH1	Le Croisic, France (CFR)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/RH2	CFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/RH3	CFR	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH1	Alicante, Spain (ASP)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH2	ASP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH3	ASP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH4	ASP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH5	ASP	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/TH6	Ria Formosa, Portugal (ASP)	Woodall, L.	HQ437213	—	—	—	—	—	—	—
HH/UH1	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/UH2	RPO	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/UH3	RPO	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/UH4	RPO	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/VH1	Gozo, Malta (GMA)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/VH2	GMA	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/VH3	GMA	Woodall, L.	HQ437204	—	—	—	—	—	—	—
HH/VH4	GMA	Woodall, L.	HQ437198	—	—	—	—	—	—	—

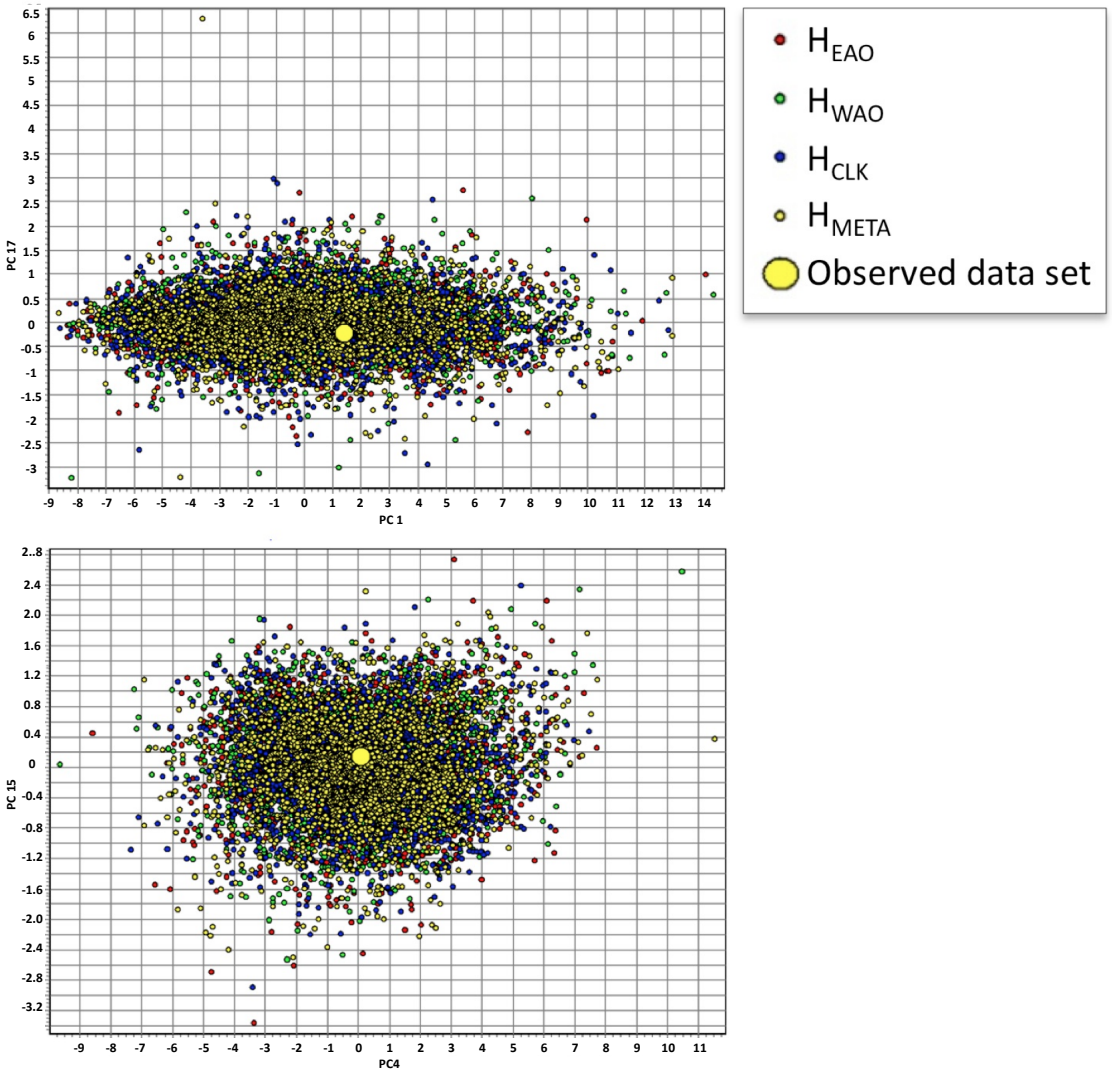
HH/VH5	GMA	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/WH1	Greece (GRE)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/WH2	GRE	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/WH3	GRE	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/WH4	GRE	Woodall, L.	HQ437206	—	—	—	—	—	—	—
HH/WH5	GRE	Woodall, L.	HQ437233	—	—	—	—	—	—	—
HH/WH6	GRE	Woodall, L.	HQ437233	—	—	—	—	—	—	—
HH/WH7	GRE	Woodall, L.	HQ437212	—	—	—	—	—	—	—
HH/WH8	GRE	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH1	Canary Islands (LCN)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH2	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH3	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH4	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH5	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH6	LCN	Woodall, L.	HQ437206	KC851910	—	KC812036	KC811884	—	—	—
HH/YH7	LCN	Woodall, L.	HQ437231	KC851912	—	KC812036	KC811884	—	KC811895	—
HH/YH8	LCN	Woodall, L.	HQ437198	—	—	KC812036	KC811877	—	—	—
HH/YH9	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH10	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH11	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH12	LCN	Woodall, L.	HQ437234	—	—	—	—	—	—	—
HH/YH13	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH14	LCN	Woodall, L.	HG437235	—	—	—	—	—	—	—
HH/YH15	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH16	LCN	Woodall, L.	HQ437224	—	—	—	—	—	—	—
HH/YH17	LCN	Woodall, L.	HQ437237	—	—	—	—	—	—	—
HH/YH18	LCN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/YH19	LCN	Woodall, L.	HQ437243	—	—	—	—	—	—	—
HHS/ZH1	Senegal, West Africa (SEN)	Woodall, L.	HQ437227	—	—	—	—	—	—	—
HHS/ZH2	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH3	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH4	SEN	Woodall, L.	HQ437228	—	—	—	—	—	—	—
HHS/ZH5	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH6	SEN	Woodall, L.	HQ437229	—	—	—	—	—	—	—

HHS/ZH7	SEN	Woodall, L.	HQ437244	—	—	—	—	—	—	—
HHS/ZH8	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH9	SEN	Woodall, L.	HQ437230	—	—	—	—	—	—	—
HHS/ZH10	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH11	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HHS/ZH12	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH13	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH14	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH15	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH16	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH17	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH18	SEN	Woodall, L.	HQ427216	—	—	—	—	—	—	—
HHS/ZH19	SEN	Woodall, L.	HQ437198	KC851914	KC851884	KC812036	KC811883	KC811907	KC811896	KC811844
HHS/ZH20	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	
HHS/ZH21	SEN	Woodall, L.	HQ437198	KC851916	KC851884	KC812036	KC811883	KC811907	KC811896	KC811844
HHS/ZH22	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH23	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH24	SEN	Woodall, L.	HQ437198	KC851915	—	KC812036	KC811879	KC811908	KC811896	KC811844
HHS/ZH25	SEN	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HHS/ZH28	SEN	Woodall, L.	HQ437217	—	—	—	—	—	—	—
HH/*H1	Riccione, Italy (RIT)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H2	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H3	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H4	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H5	RIT	Woodall, L.	HQ437222	—	—	—	—	—	—	—
HH/*H6	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H7	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H9	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H10	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H11	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H12	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H13	RIT	Woodall, L.	HQ437223	—	—	—	—	—	—	—
HH/*H14	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—

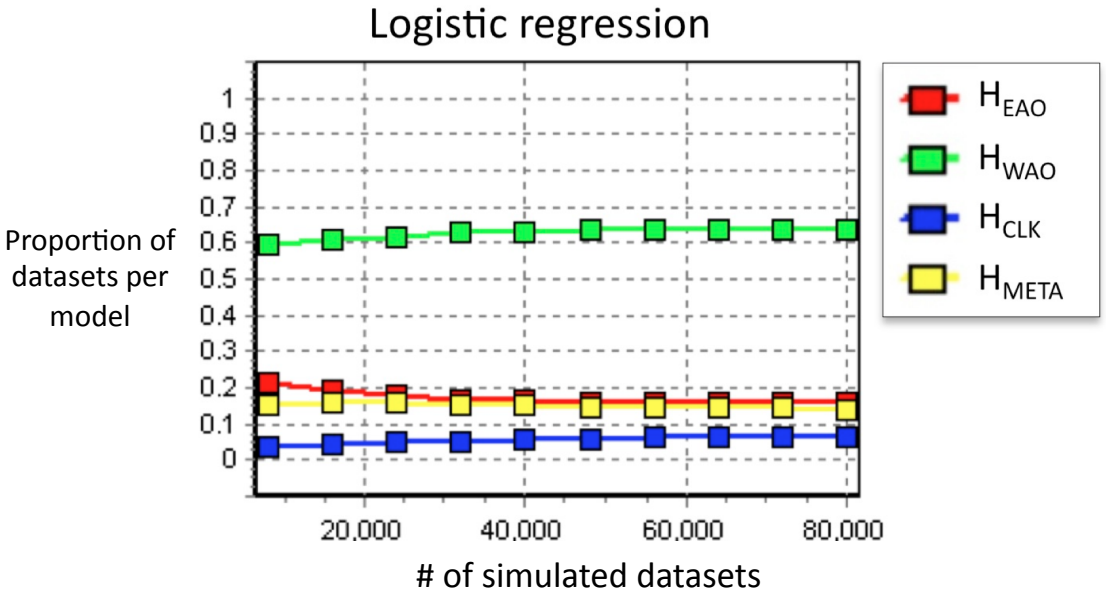
HH/*H15	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H16	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H17	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H18	RIT	Woodall, L.	HQ437214	—	—	—	—	—	—	—
HH/*H19	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H20	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H21	RIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/*H28	RIT	Woodall, L.	KC851867	—	—	KC812036	KC811884	—	—	—
HH/*H29	RIT	Woodall, L.	KC851868	—	—	KC812036	KC811880	—	—	KC811847
HH/*H30	RIT	Woodall, L.	KC851869	KC851912	KC851883	KC812036	KC811884	KC811905	—	KC811845
HH/NAH2	Napoli, Italy (NIT)	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NAH3	NIT	Woodall, L.	HQ437212	—	—	—	—	—	—	—
HH/NAH4	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NAH5	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NAH7	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NAH8	NIT	Woodall, L.	HQ437212	—	—	—	—	—	—	—
HH/NAH9	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—
HH/NAH10	NIT	Woodall, L.	HQ437198	—	—	—	—	—	—	—

Appendix 1.2 (a) Principal components analysis (b) model selection, and (c) gene trees.

Appendix 1.2a Principal components analysis (PCA) was carried out in DIYABC (Cornuet *et al*., 2008, 2010) to evaluate the proposed models and chosen prior distributions. PCAs were generated from simulated summary statistics (10,000) closest to our observed summary statistics. The below PCA planes show examples of the observed summary statistic (large yellow circle) in relation to the simulated data sets (all other data). In total, 72 observed summary statistic values were located within each PCA plane carried out in pairwise comparisons between each of the 72 values.

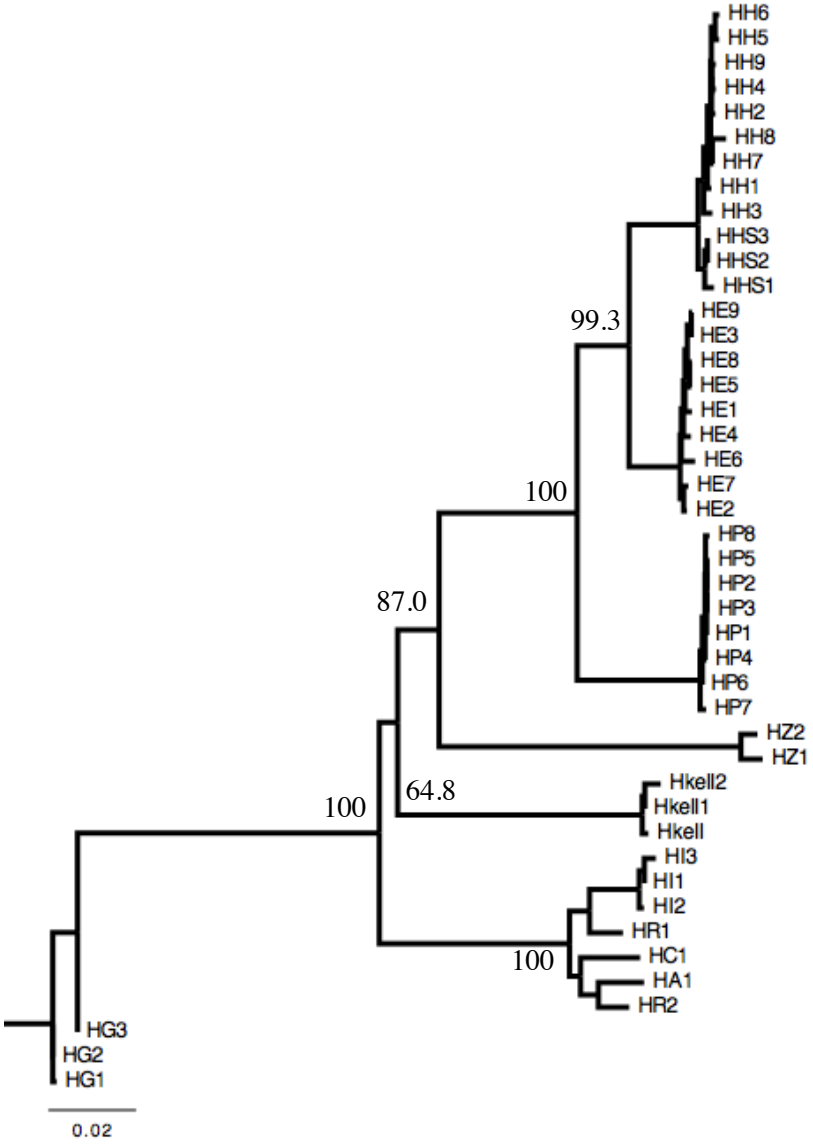


Appendix 1.2b Of the four million simulated datasets we selected (2%) 80,000 data sets closest to our observed data using a logistic regression procedure to select between the four modelled hypotheses. Modelled hypothesis H_{WAO} (green line) was consistently chosen over all other models (H_{EAO}, H_{CLK} and H_{META}). See Fig. 1.2 and the Introduction section for a detailed description of each hypothesis.

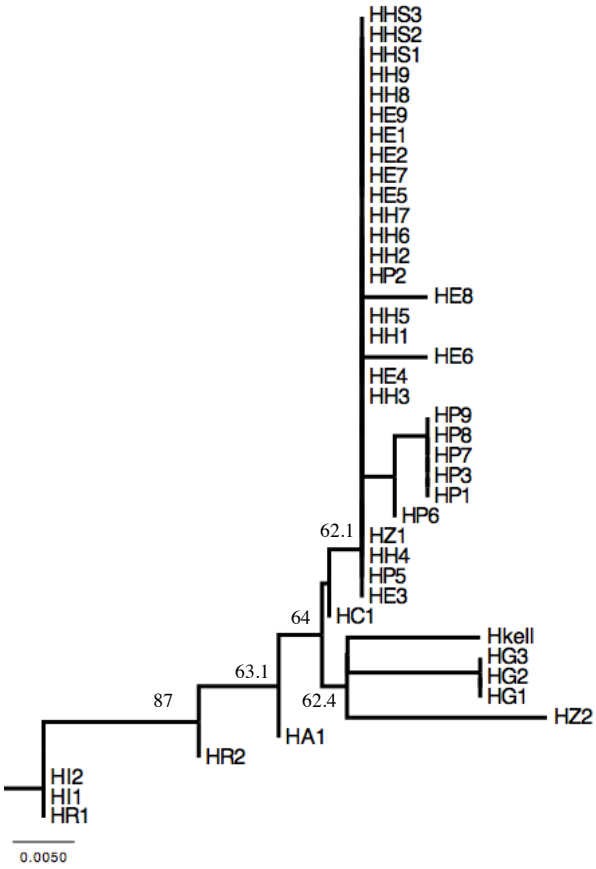


Appendix 1.2c Independent gene trees constructed from sequence data used in *BEAST analysis; estimated in PhyML v3 (Guindon et al., 2010) using 1000 bootstrap replicates. The best-fit model was determined using a likelihood ratio test, with likelihood scores evaluated using a standard Akaike information criterion (AIC) in JMODELTEST 0.1. Species IDs correspond to the following: HI, *Hippocampus ingens*; HZ, *H. zosterae*; HG, *H. guttulatus*; HA, *H. algiricus*; HC, *H. capensis*; HR, *H. reidi*; HK, *H. kelloggi*; HE, *H. erectus*; HP, *H. patagonicus*; HH, *H. hippocampus* (Europe); HHS, *H. hippocampus* (West Africa).

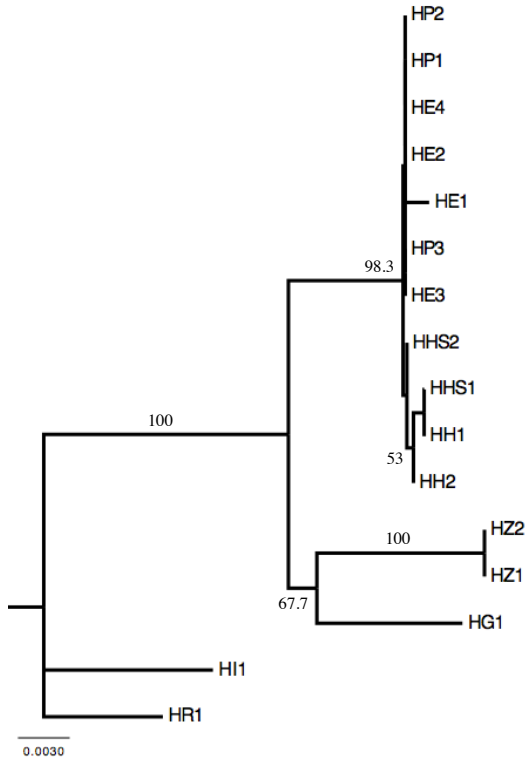
mtDNA concatenated



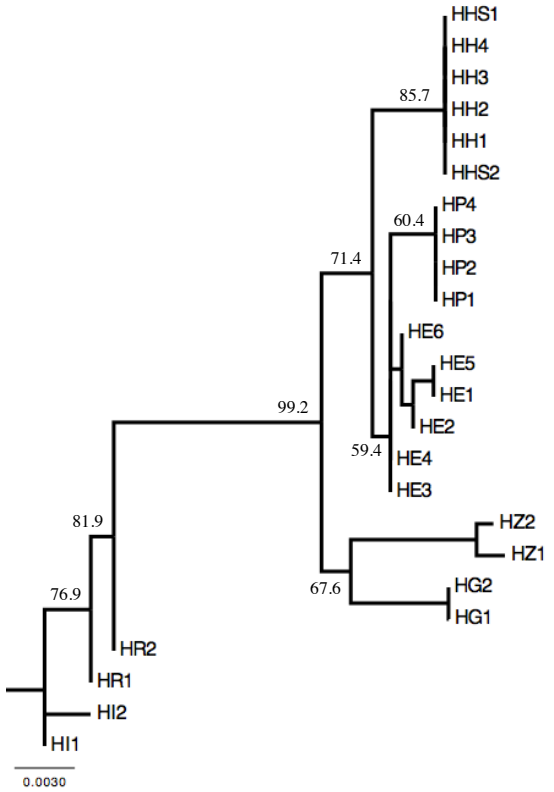
Aldolase



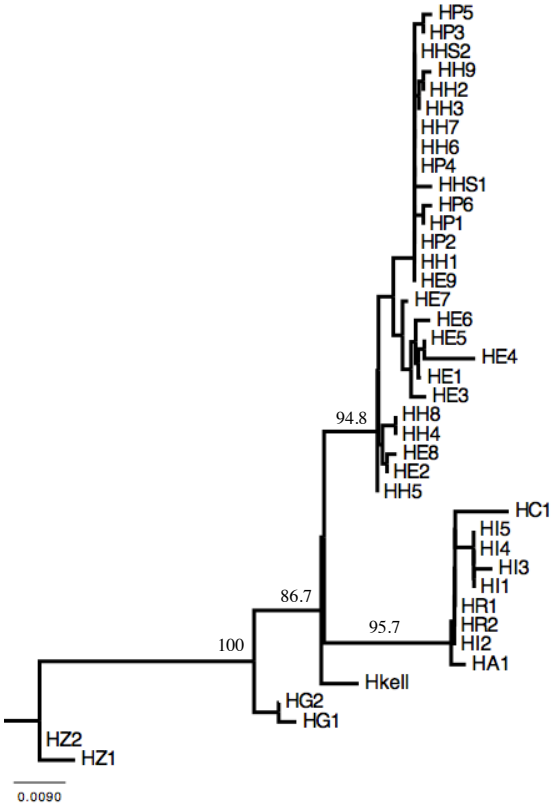
myh6



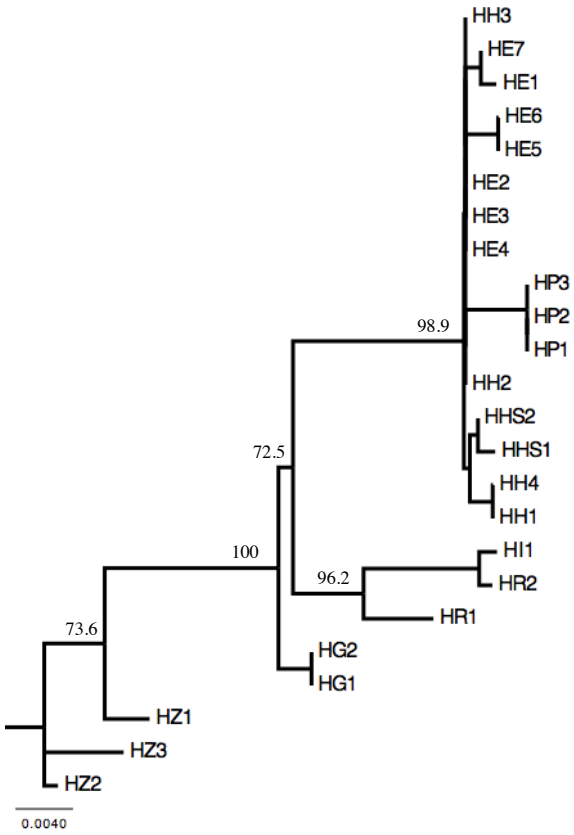
Rhodopsin



S7



Tmo4c4


HH3
HE7
HE1
HE6
HE5
HE2
HE3
HE4
HP3
HP2
HP1
HH2
HHS2
HHS1
HH4
HH1
HI1
HR2
HR1
HG2
HG1
HZ1
HZ3
HZ2
98.9
72.5
100
96.2
73.6
0.0040

Appendix 1.3 (a) Intraspecific population statistics and (b) observed summary statistics.

Appendix 1.3a Geographically sampled F_{ST}/Nm statistics were calculated to examine intraspecific population differences, and to help inform the sampling design of our coalescent analyses. Estimates of gene flow (migration rate, *Nm*) above the diagonal, and population pairwise F_{ST} below; (1) calculated from cytochrome *b* (*cytb*) 696 bp sequence data (North/Central America = 114) and (South America = 20) and (2) concatenated *cytb* + control region 935 bp (Europe–West Africa = 255). *Nm* values based on Slatkin (1995); linearized $t/M = F_{ST}/(1-F_{ST})$ calculated in ARLEQUIN 3.5.1.2 (Excoffier & Lischer, 2010). The previously suggested classification of *H. patagonicus* (Argentina) based on a morphological description (Piacentino & Luzzatto, 2004) is consistent with lineage independence from *H. erectus* (between *H. erectus* and *H. patagonicus* global mtDNA F_{ST} = 0.945), while showing panmixia across Brazil and Argentina samples (e.g. global mtDNA $Nm = infinite/F_{ST} = -0.053$). Likewise, near-panmixia can be strongly inferred across the six geographically sampled regions of *H. erectus* (global mtDNA $Nm = 3.72$ to *infinite*), with a moderate reduction in gene flow across Cape Hatteras (USA) at the mtDNA locus (Nm = 3.72–10.45). Given these F_{ST}-derived patterns of gene flow, we modelled *H. erectus* and *H. patagonicus* as two populations for the *BEAST and DIYABC (Cornuet *et al*., 2008, 2010) analyses. Using the same criteria, *H. hippocampus* samples were divided into two population samples (Europe and West Africa) for the same analyses, with pairwise mtDNA *Nm* values ranging from 0.81 to 1.09 between these two regions. This overall four-population modelling scheme was supported by a previous AMOVA on eight geographically sampled regions in the distribution of *H. hippocampus* (Woodall *et al*., 2011).

(1)

	New York–New Jersey *n* = 45	Chesapeake *n* = 22	Gulf Coast *n* = 28	Atlantic coast (Florida) *n* = 10	Central America *n* = 10	Brazil *n* = 6	Argentina *n* = 14
New York–New Jersey		82.35	7.32	10.06	10.49	0.04	0.05
Chesapeake	0.006		3.72	5.24	4.39	0.05	0.03
Gulf Coast	0.165	0.197		*infinite*	*infinite*	0.04	0.03
Atlantic coast	0.117	0.148	−0.032		*infinite*	0.02	0.03
Central America	0.080	0.101	0.043	0.005		0.05	0.03
Brazil	0.928	0.935	0.945	0.968	0.937		*infinite*
Argentina	0.916	0.917	0.933	0.951	0.902	−0.012	

(2)

	Channel *n* = 39	Biscay *n* = 32	S. Iberia *n* = 31	MSP *n* = 18	WMED *n* = 44	EMED *n* = 44	Islands *n* = 21	W. Africa *n* = 26
Channel		13.72	6.73	150.56	4.55	8.89	14.31	0.81
Biscay	0.032		3.22	7.45	2.04	2.96	2.97	0.89
S. Iberia	0.070	0.135		27.96	10.86	12.27	19.79	1.07
MSP	0.004	0.063	0.013		22.23	*infinite*	67.07	1.09
WMED	0.101	0.197	0.044	0.022		*infinite*	10.14	0.70
EMED	0.055	0.145	0.039	0.001	0.001		35.05	0.74
Islands	0.035	0.144	0.025	0.007	0.047	0.014		0.92
W. Africa	0.379	0.361	0.318	0.315	0.416	0.401	0.353	

Channel, south coast of UK; Biscay, Bay of Biscay; EMED, east Mediterranean Sea; Islands, Madeira and Canary Islands; MSP, south coast of Spain; S. Iberia, south-west Iberia; W. Africa, West Africa; WMED, West Mediterranean Sea.

Appendix 1.3b Observed summary statistics calculated in DIYABC.

Mitochondrial

	P1	P2	P3	P4
Number of haplotypes	30	4	32	7
Number of segregating sites	36	3	31	13
Mean pairwise differences	2.015	0.3895	0.944	2.3533
Private segregating sites	22	1	16	6

	P1 & P2	P1 & P3	P1 & P4	P2 & P3	P2 & P4	P3 & P4
Between-population pairwise differences	1.9692	1.1575	2.0298	0.9401	1.5918	0.9599
Pairwise population F_{ST}	0.8941	0.9164	0.8513	0.9607	0.9298	0.5723

Autosomal

	P1	P2	P3	P4
Number of segregating sites	6.4	1.6	3.2	1.6
Mean pairwise differences	1.3118	0.5134	0.8359	0.9
Private segregating sites	5.4	1.2	1.8	1.0

	P1 & P2	P1 & P3	P1 & P4	P2 & P3	P2 & P4	P3 & P4
Between-population pairwise differences	1.1859	1.1035	1.367	0.6689	0.7583	1.0524
Pairwise population F_{ST}	0.4262	0.3751	0.2979	0.6850	0.6527	−0.2487

P1, *H. erectus*; P2, *H. patagonicus*; P3, *H. hippocampus* Europe; P4, *H. hippocampus* W. Africa.

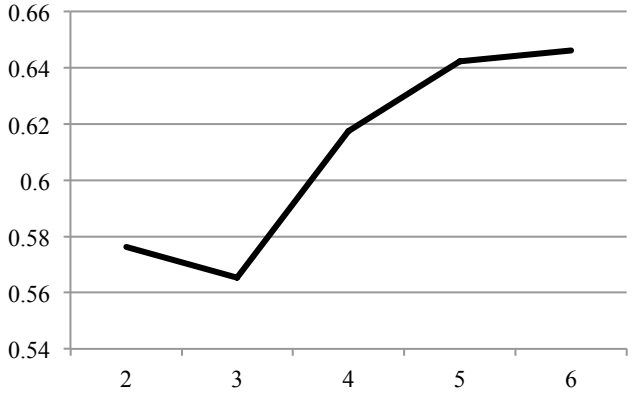
APPENDIX 2

Appendix 2.1

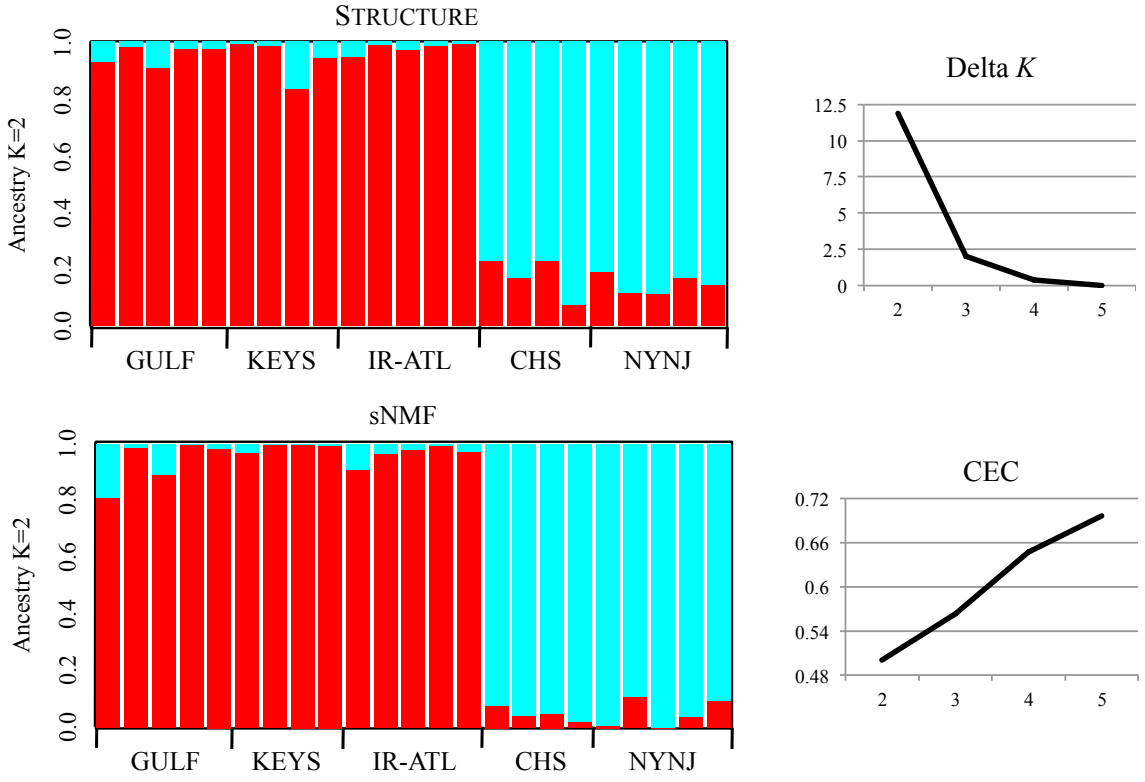
2.1 (a) Additional details on the inference of ancestry coefficients, (b) sNMF and STURCTURE comparison test, (c) isolation-by-distance methods

Appendix 2.1a) Inference of ancestral gene pools: The program sNMF (Frichot et al., 2014) was utilized to estimate individual ancestry and population clustering. Ancestry membership coefficients (K) were estimated to determine subpopulation membership by running 10 replicates of K 2-6 using a cross-entropy criterion (CEC). To select the ancestry coefficient with the highest likelihood, sNMF outputs a cross-entropy score, and the lowest value represents the best-supported K value (K=3) (Figure below).

Cross entropy score

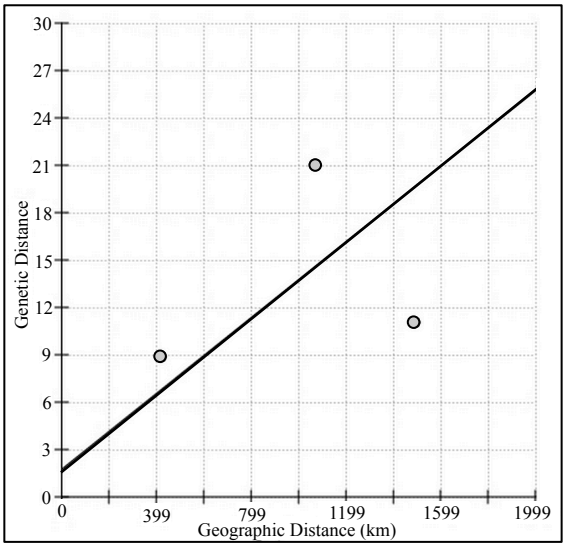


Appendix 2.1b) sNMF and STRUCTURE SNP subset comparison test: As a comparison test between STRUCTURE and sNMF we randomly selected 2000 SNPs from our full dataset. K-values ranging from 1 to 5 were analyzed in both programs with ten independent runs per K. In STRUCTURE each run consisted of 100,000 burn-in and 150,000 MCMC iterations, using the "Admixture Model" and "Correlated Allele Frequency Model" with default settings. This run length followed (Massatti & Knowles, 2014), and as they reported we found that results from longer preliminary runs were not different using more burn-in or MCMC iterations. STRUCTURE HARVESTER (Earl & vonHoldt, 2011) and R were used to visualize the STRUCTURE results, with the most probable K chosen based on Delta-*K* (Evanno et al., 2005). The most probable K was selected based on cross-entropy criterion (CEC) values in sNMF. The CEC value can be interpreted as the inverse of Delta-*K*. Based on this subset of SNPs both programs supported the most probable K of 2. These plots clearly distinguish populations north and south of Cape Hatteras, though the differentiation of the cline visualized in our PCA and full data analysis separating eastern Gulf and southern Atlantic populations was not retained. Each line represents an individual from the following locations: GULF = Gulf of Mexico; KEYS = Florida Keys; IR-ATL = Jacksonville/Indian River Lagoon, FL; CHS = Chesapeake Bay; NYNJ = New York and New Jersey.

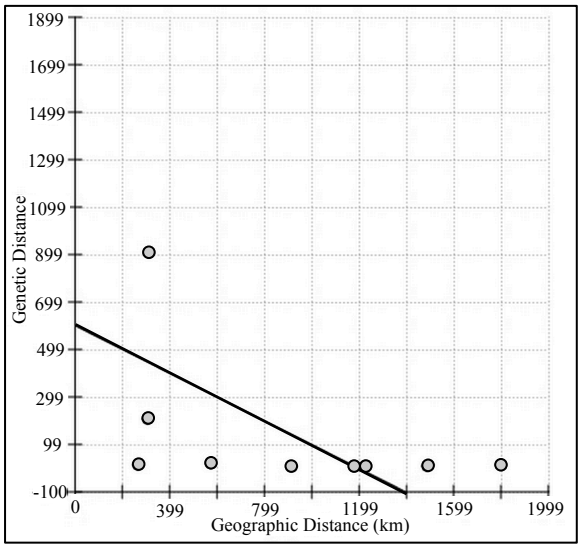


Appendix 2.1c) Isolation-by-distance: A test for isolation-by-distance was conducted using the IBD program (Jensen et al., 2005) by Mantel's test (10,000 randomizations) of FST and shoreline distance (km). Pairwise linearized FST (FST /(1-FST) was calculated in vcftools and coastal distance of coastlines between sampling locations was determined using Google Earth Tools. For the between subpopulations (a) test, as presented in the manuscript, the centroid distance between sampling locations for the Gulf-Keys subpopulation was utilized and results indicated a non-significance correlation between geographic and genetic distance (p = 0.4925). A second test for IDB between the five main sampling locations was also conducted (b) resulting in a non-significant value of p = 0.9026 (not reported in the manuscript).

a)

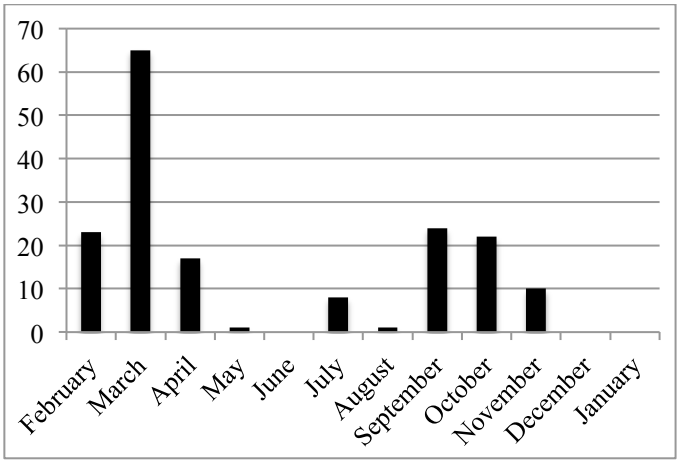


b)

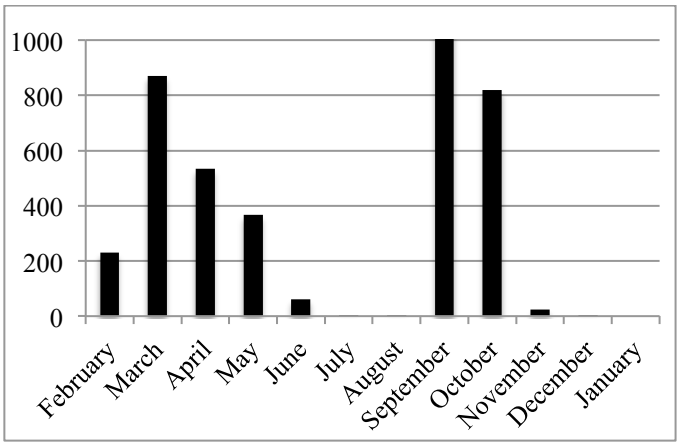


Appendix 2.2: NOAA Long-term bottom trawl survey catch records (1972-2008) for *Hippocampus erectus* and *Syngnathus fuscus* the Mid-Atlantic Bight (i.e., Virginia Province)

Aggregated monthly catch totals from the Virginia province conducted from 1972-2008 for *H. erectus* and *S. fuscus.* Peaks in monthly catch totals of *H. erectus* correspond to inshore-offshore local migration to the intercontinental shelf reported for *S. fuscus*. For additional details on specific migration rates of *S. fuscus* across latitudes see (Lazzari & Able, 1990). >90% of samples were collected 20 km off the coast of Chesapeake Bay to Cape Cod, MA., at a depth 10-20m. Dependent on latitude the majority of both species are found in nearshore zones and estuaries between the months of April-October, with higher latitudes resulting in earlier off-shore catch record abundance.



H. erectus monthly offshore catch totals from 1972-2008.

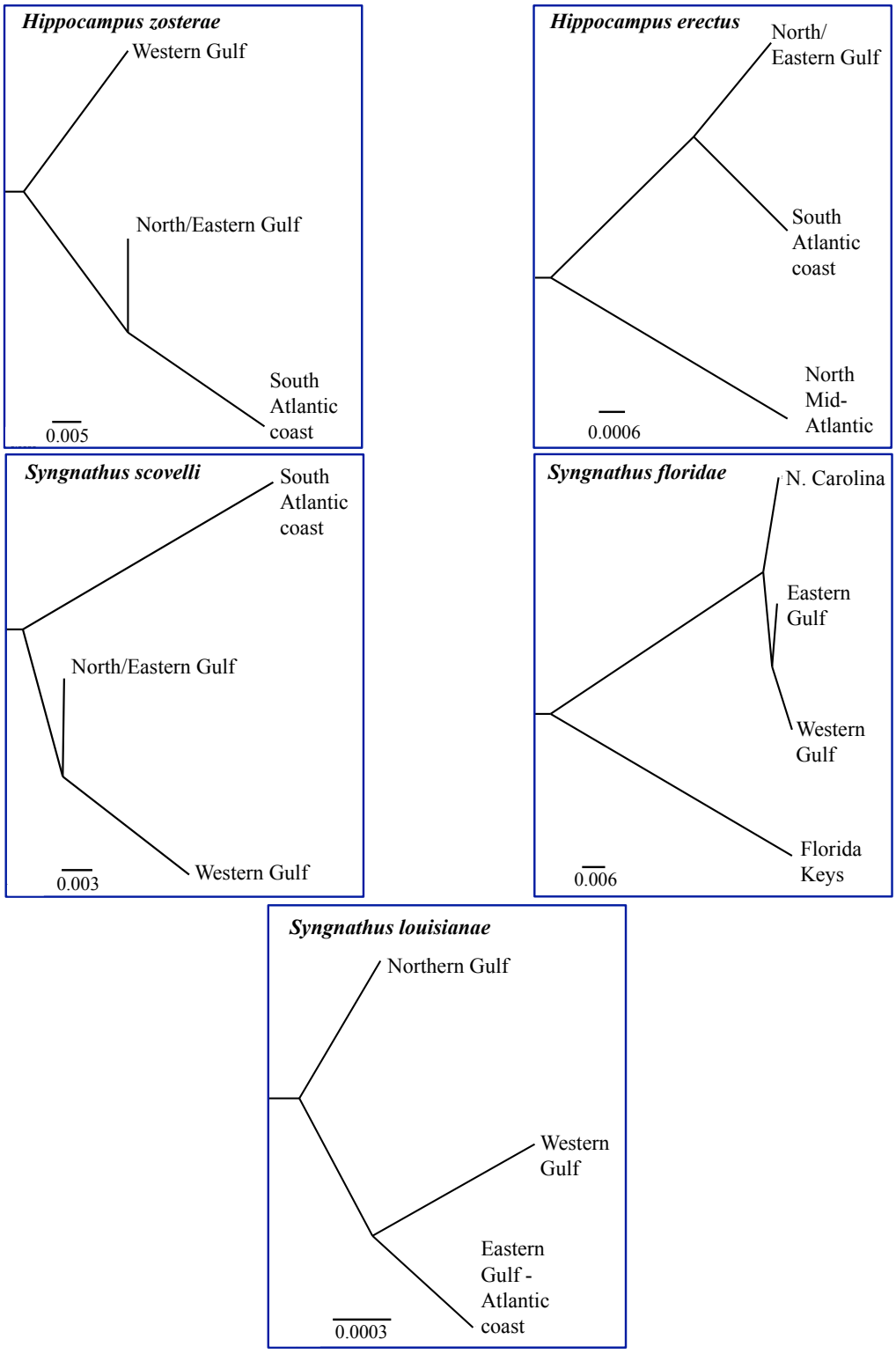


S. fuscus monthly offshore catch totals from 1972-2008.

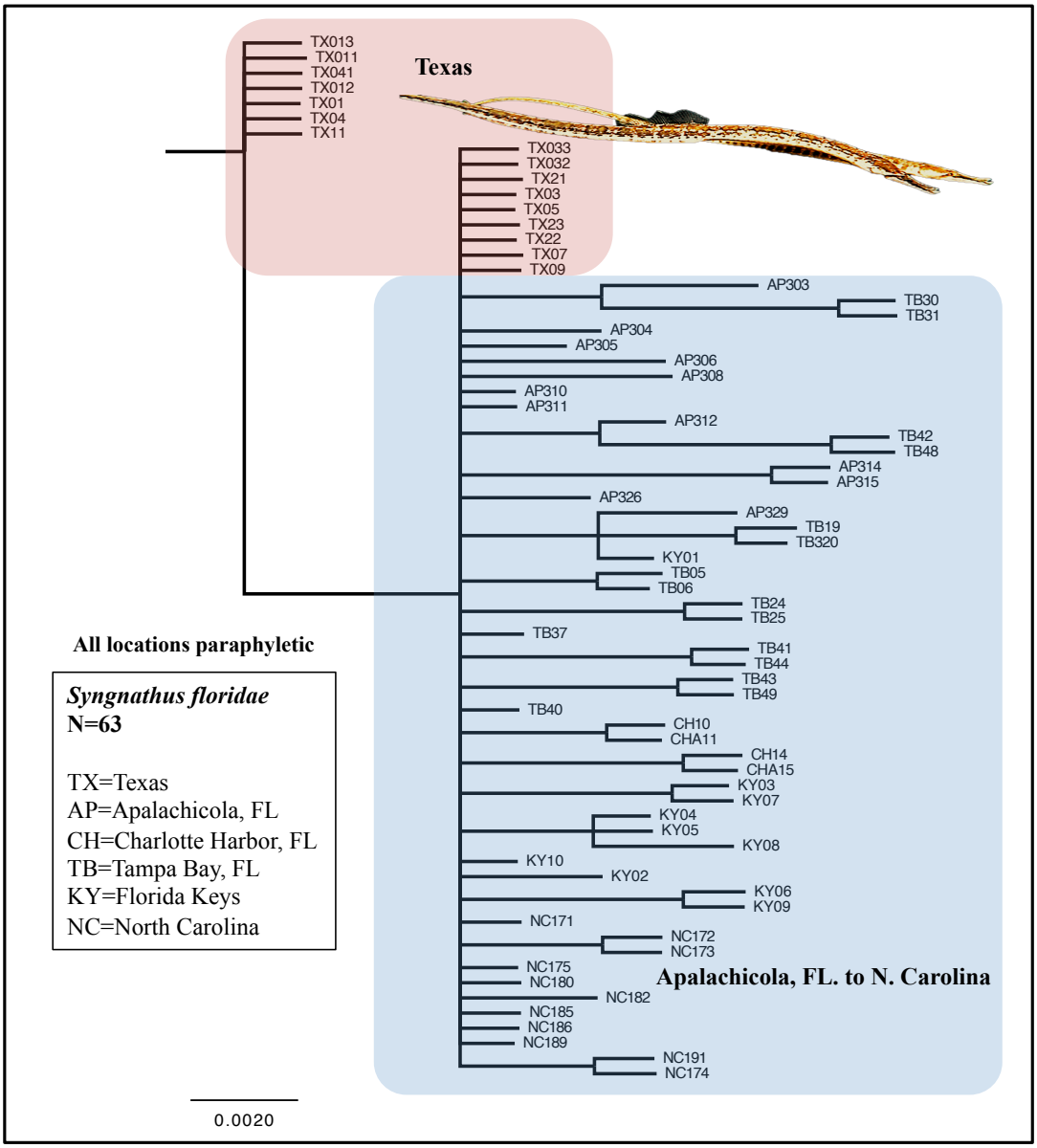
Appendix Table 2.1 Summary of genomic data collected for each individual and sampling location. Shown are the raw read counts per individual from the HiSeq Illumina run. The number of reads after initial processing; 65% RAD tag coverage across individuals. The analyzed reads from homologous loci across individuals at all sites. Raw sequence reads are available at NCBI SRA Project: SRP048776.

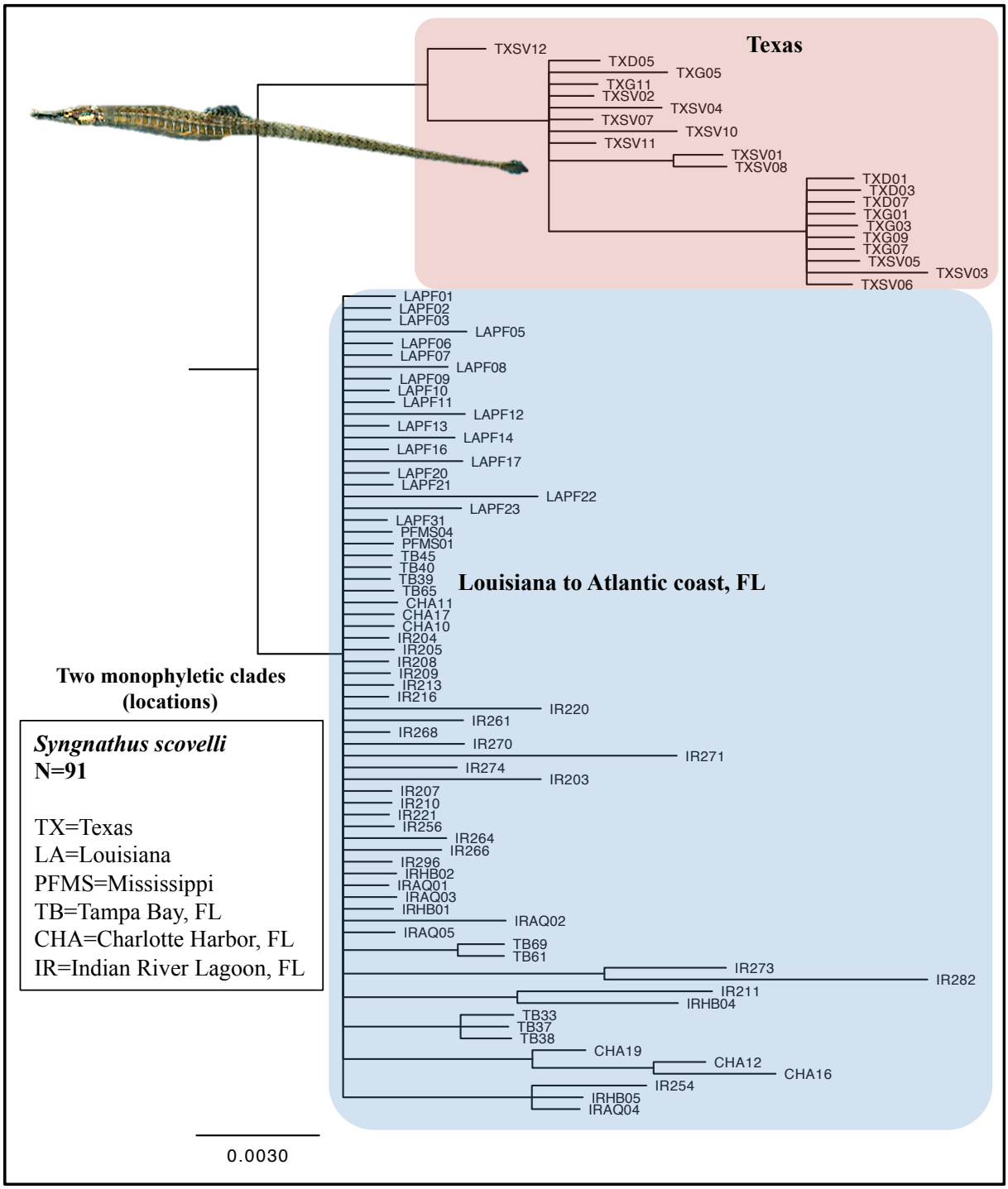
Individual	SRA Accession Number	Location	Total 90bp reads	Reads post initial quality control (65% coverage, q=15)	Analyzed reads (100% coverage, q=20)
HEBGu1	SRX732943	Apalachicola, FL	3,157,866	1,110,944	814,411
HEGu2B	SRX732948	Apalachicola, FL	3,132,578	1,130,041	831,952
HETB1N	SRX732960	Tampa Bay, FL	1,338,653	534,152	398,006
HETB01	SRX732959	Tampa Bay, FL	2,578,857	973,553	712,973
HEBF1	SRX732942	Charlotte Harbor, FL	1,098,313	419,305	313,920
HEKY06	SRX732951	FL Keys	4,736,299	1,799,006	1,311,869
HEKY51	SRX732953	FL Keys	4,504,203	1,667,968	1,223,992
KYHE02	SRX732964	FL Keys	4,198,059	1,618,812	1,187,937
HEKY12	SRX732952	FL Keys	4,459,738	1,656,374	1,210,098
HEIR3	SRX732949	Indian River, FL	2,309,705	898,087	663,094
HEIR5	SRX732950	Indian River, FL	3,426,244	1,314,835	962,582
IR03HE	SRX732961	Indian River, FL	3,904,405	1,505,889	1,102,438
JXIR02HE	SRX732963	Jacksonville, FL	3,262,469	1,258,885	919,853
JXIR01HE	SRX732962	Jacksonville, FL	3,141,316	1,216,565	890,663
HECHS1	SRX732944	Chesapeake Bay	2,704,326	1,071,894	790,781
HECHS3	SRX732945	Chesapeake Bay	4,239,371	1,661,500	1,218,476
HECHS4	SRX732946	Chesapeake Bay	2,370,238	923,860	684,441
HECHS12	SRX732947	Chesapeake Bay	264,862	481,791	358,833
HENJ31	SRX732954	New Jersey	4,071,240	1,519,521	1,111,883
HENY05	SRX732955	New York	3,028,313	1,131,996	829,156
HENY09	SRX732956	New York	3,958,295	1,471,727	1,076,320
HENY20	SRX732957	New York	3,683,443	1,374,731	999,777
HENY29	SRX732958	New York	3,266,915	1,115,816	817,710
Total Reads			**72,835,708**	**27,857,252**	**20,431,165**

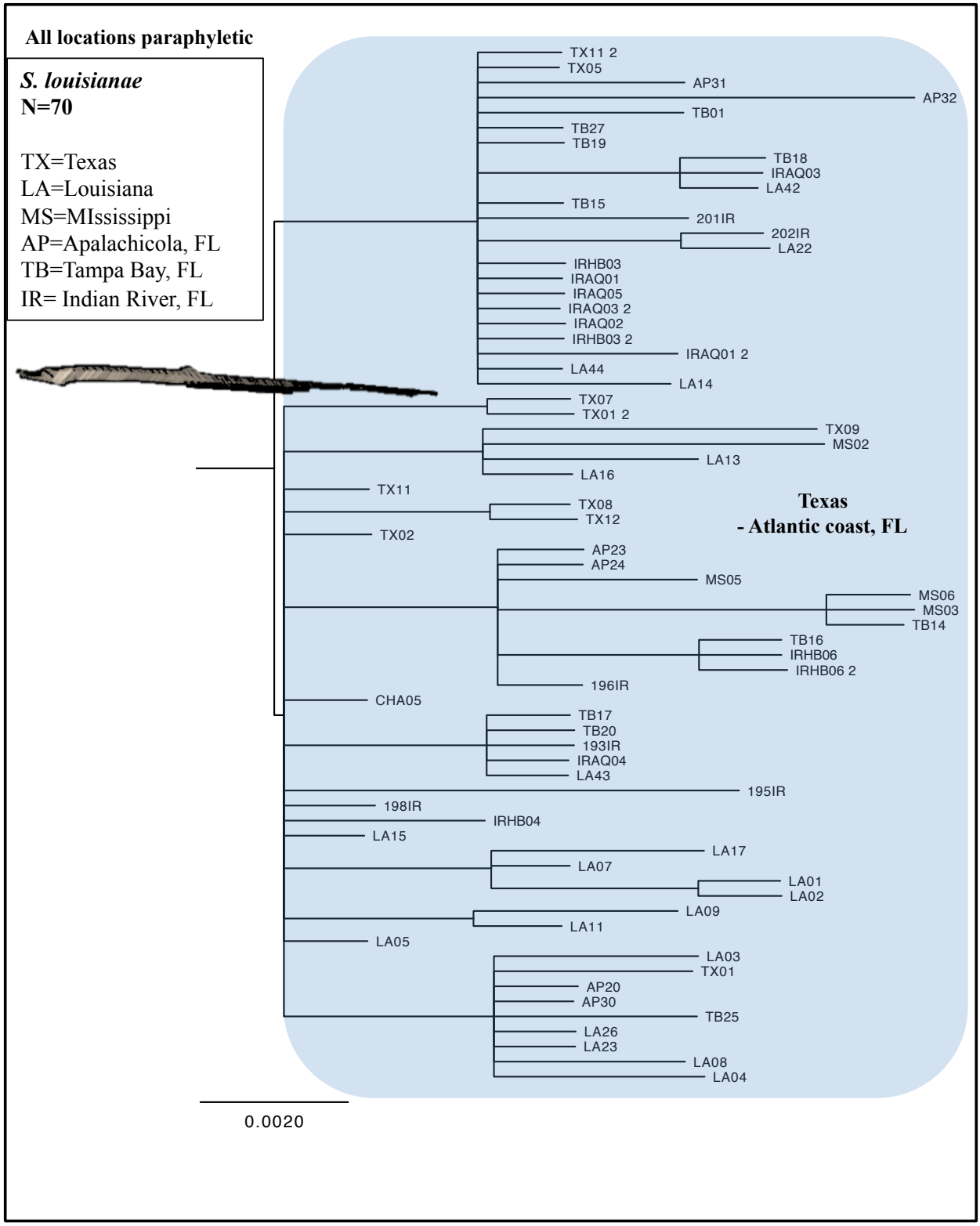
Appendix 3.1 Maximum likelihood *Treemix* graphs. Branch length scaled to the amount of genetic drift between subpopulations. Individuals were grouped based on the concordance between PCA and ancestral gene pool inference (sNMF).

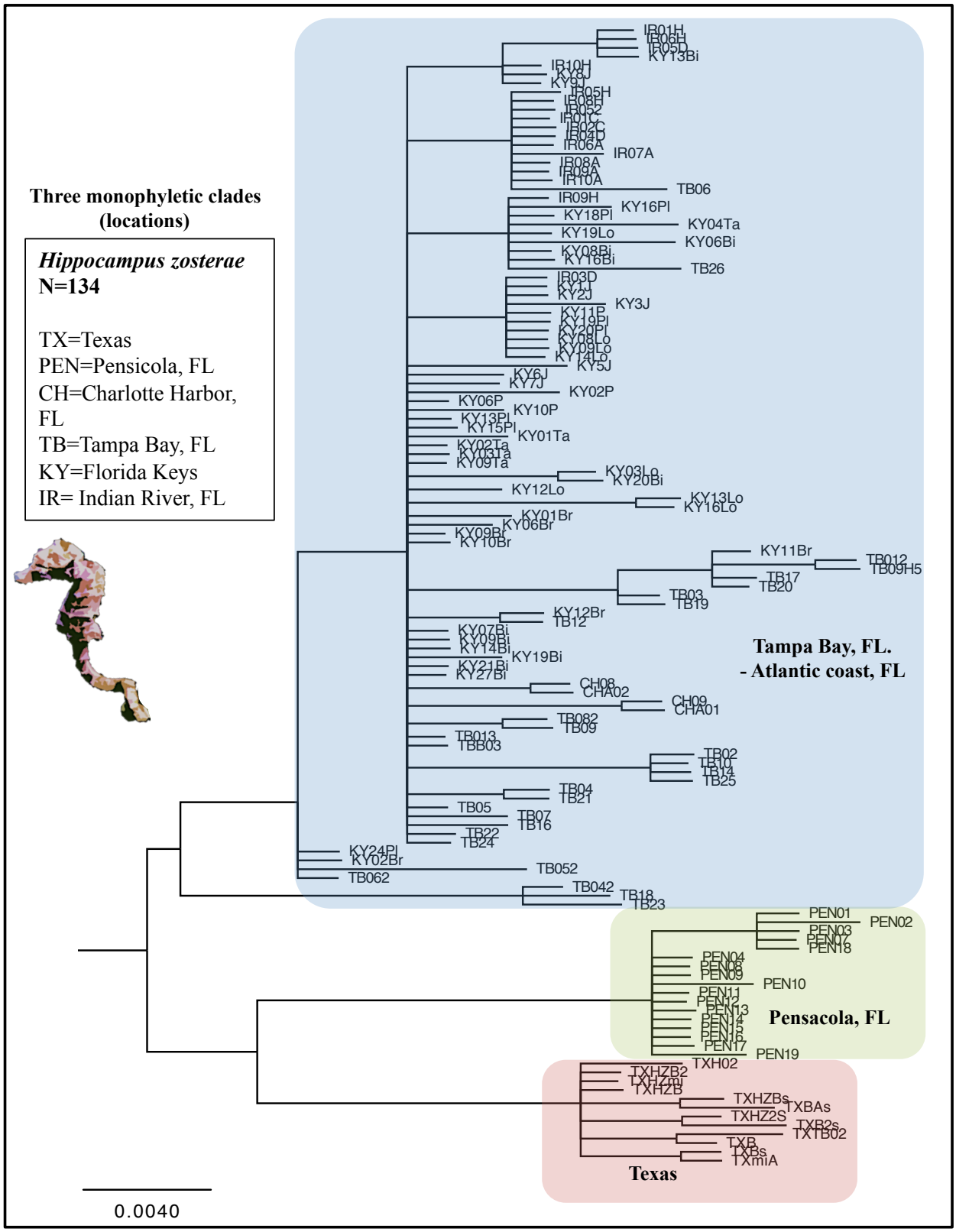


Appendix 3.2 Mitochondrial phylogeny (cytochrome oxidase subunit 1 652bp): Bayesian phylogenies generated using MrBayes v.3.1.2 (Ronquist and Huelsenbeck 2003). Each tree was run for 1,000,000 iterations, sampling every 300 generations, with the first 250,000 generations discarded as burn-in. The GTR model of sequence evolution was selected for all trees based on Bayesian information criteria scores in jModelTest. Maximum likelihood trees were also constructed with PhyML v.3.0 with 1,000 bootstrap replicates (Guindon et al. 2010) and produced concordant topologies (not shown).




Texas
TXSV12
TXD05
TXG05
TXG11
TXSV02
TXSV04
TXSV07
TXSV10
TXSV11
TXSV01
TXSV08
TXD01
TXD03
TXD07
TXG01
TXG03
TXG09
TXG07
TXSV05
TXSV03
TXSV06
LAPF01
LAPF02
LAPF03
LAPF05
LAPF06
LAPF07
LAPF08
LAPF09
LAPF10
LAPF11
LAPF12
LAPF13
LAPF14
LAPF16
LAPF17
LAPF20
LAPF21
LAPF22
LAPF23
LAPF31
PFMS04
PFMS01
TB45
TB40
TB39
TB65
Louisiana to Atlantic coast, FL
CHA11
CHA17
CHA10
IR204
IR205
IR208
IR209
IR213
IR216
IR220
IR261
IR268
IR270
IR271
IR274
IR203
IR207
IR210
IR221
IR256
IR264
IR266
IR296
IRHB02
IRAQ01
IRAQ03
IRHB01
IRAQ02
IRAQ05
TB69
TB61
IR273
IR282
IR211
IRHB04
TB33
TB37
TB38
CHA19
CHA12
CHA16
IR254
IRHB05
IRAQ04
0.0030
Two monophyletic clades (locations)
Syngnathus scovelli
N=91
TX=Texas
LA=Louisiana
PFMS=Mississippi
TB=Tampa Bay, FL
CHA=Charlotte Harbor, FL
IR=Indian River Lagoon, FL

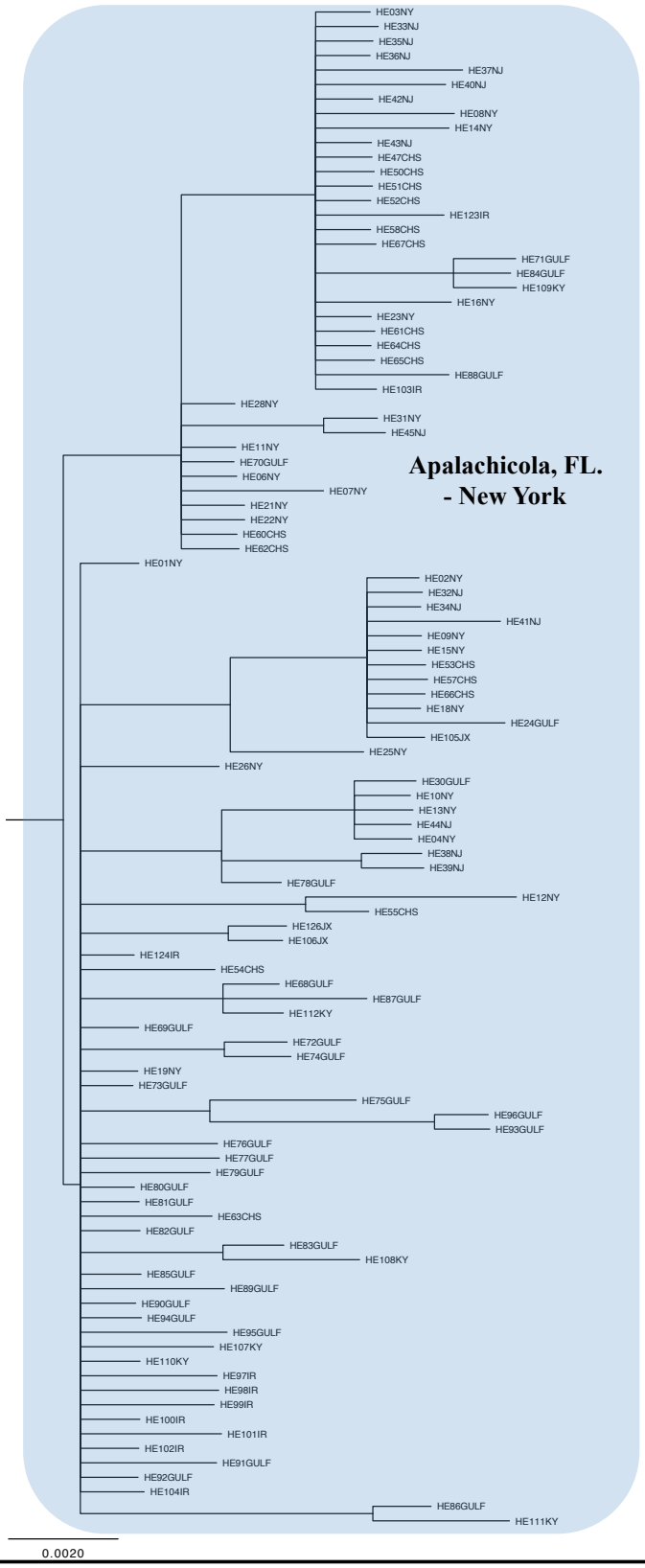

All locations paraphyletic
S. louisianae
N=70
TX=Texas
LA=Louisiana
MS=MIssissippi
AP=Apalachicola, FL
TB=Tampa Bay, FL
IR= Indian River, FL
TX11 2
TX05
AP31
AP32
TB01
TB27
TB19
TB18
IRAQ03
LA42
TB15
201IR
202IR
LA22
IRHB03
IRAQ01
IRAQ05
IRAQ03 2
IRAQ02
IRHB03 2
IRAQ01 2
LA44
LA14
TX07
TX01 2
TX09
MS02
LA13
LA16
TX11
TX08
TX12
TX02
Texas
- Atlantic coast, FL
AP23
AP24
MS05
MS06
MS03
TB14
TB16
IRHB06
IRHB06 2
196IR
CHA05
TB17
TB20
193IR
IRAQ04
LA43
195IR
198IR
IRHB04
LA15
LA17
LA07
LA01
LA02
LA09
LA11
LA05
LA03
TX01
AP20
AP30
TB25
LA26
LA23
LA08
LA04
0.0020


Three monophyletic clades (locations)
Hippocampus zosterae
N=134
TX=Texas
PEN=Pensicola, FL
CH=Charlotte Harbor, FL
TB=Tampa Bay, FL
KY=Florida Keys
IR= Indian River, FL
Tampa Bay, FL.
- Atlantic coast, FL
Pensacola, FL
Texas
0.0040

All locations paraphyletic

Hippocampus erectus
N=115

GULF=Apalachicola – Charlotte Harbor, FL
KY=Florida Keys
IR= Indian River Lagoon, FL
JX=Jacksonville, FL
CHS=Chesapeake Bay
NJ=New Jersey
NY=New York



Appendix Table 3.3 Number of *Sargassum* particles relative to number of particles that reached a recruitment zone in a given year (out of 23,000 released per year from the eastern Gulf of Mexico from March-May). Simulations ran from March through the end of November per year.

Particles allowed to "bounce" from one coastal recruitment zone to the next

Year	Chandleur Sound, LA	Pensacola, FL	Tampa Bay, FL	Indian River Lagoon, FL	Morehead City, NC	Chesapeake Bay
2004	66	5	24	748	551	42
2005	551	47	440	685	507	20
2006	93	4	46	1388	358	23
2007	23	134	1	66	6	0
2008	89	18	13	485	16	47
2009	13	0	0	297	1	0
2010	111	7	228	271	20	4
2011	20	0	0	115	20	1
2012	54	13	4	124	0	0
2013	94	34	0	91	1	0
Average	*111.4*	*26.2*	*75.6*	*427*	*148*	*13.7*

Particles are retained (i.e., "beached") upon contact with the initial coastal recruitment zone

Year	Chandleur Sound, LA	Pensacola, FL	Tampa Bay, FL	Indian River Lagoon, FL	Morehead City, NC	Chesapeake Bay
2004	47	1	25	99	77	14
2005	152	33	137	203	142	9
2006	106	6	48	154	98	22
2007	32	105	1	1	0	0
2008	79	8	8	28	1	4
2009	4	0	0	42	1	0
2010	85	9	244	29	30	0
2011	20	0	0	44	17	0
2012	62	3	1	32	3	0
2013	88	17	0	21	2	0
Average	*67.5*	*18.2*	*46.4*	*65.3*	*37.1*	*4.9*

Appendix 4.1

4.1: Citizen science and student led wildlife trade investigations

Curriculum Guide

TABLE OF CONTENTS

INTRODUCTION

The goal of this protocol is to allow other educators, scientists, and students to replicate the workflow performed by lead author, J.T. Boehm, and his high school and undergraduate student mentees to investigate the seahorse trade through ecommerce sites, traditional medicine markets, and physical curio shops. Cumulative wildlife trade through these combined markets is a large and difficult to monitor enterprise; providing the opportunity for educators and scientists to use the help of citizen scientists and students for investigating where species might be exploited along this commodity pipeline. **Possible avenues of investigation include identifying a surge in domestic species trade that might not be monitored by CITES or other governmental organizations, finding trends in changes of geographic origin of the primary species being identified, locating instances of mislabeling and identifying areas that might need more governmental oversight**.

Molecular techniques involved with using DNA barcoding to identify species are taught to student mentees. These skills are instrumental to future success as a researcher in the life sciences and are applicable to STEM fields in general. The implementation of DNA barcoding as curriculum has already been established (Harris & Bellino, 2013), and the following workflow acts as a separate module that can be used to enhance The Student DNA Barcoding Project created by Harris & Bellino.

MODULE GOALS

Learning Objectives

Research Literacy (RL)

1. Locate printed and online articles from the professional scientific literature
2. Read and interpret data in the professional scientific literature
3. Critically assess relevance and procedures of primary research

Experimental Design (ED)

1. Explore patterns in biological data and trade reports that lead to the development of testable hypotheses
2. Understand the significance of relevant variables, controls, materials, and overall experimental design
3. Develop a research plan involving a family or genus of organisms currently being traded
4. Explain and perform analytical procedures in a family or genus of organisms currently being traded

Data Collection/Analysis (DC)

1. Apply appropriate field and laboratory techniques to sample collection in physical markets and ecommerce
2. Summarize and present data in an appropriate format using statistical software such as Excel
3. Interpret and evaluate data as supporting or refuting the hypothesis proposed
4. Maintain organized lab/field notebook

Sharing Results (SR)

1. Write, present and publicly defend a research proposal/report
2. Summarize key points of research concisely and clearly using presentation software
3. Write a final research proposal/report

Relevance and Impact of Research (RI)

1. Connect primary research to everyday problem solving
2. Evaluate the ethical implications of research and acknowledge limitations of research
3. Develop hypotheses using a whole system approach that incorporates various fields of inquiry. Incorporate ecology, biology, and genetics, for example.
4. Apply critical thinking skills outside of the classroom

Research Objectives (RO)

1. Develop a method for sampling organisms currently being traded
2. Fieldwork and sample collection
3. Extract DNA, Amplify barcode region, check results
4. Analyze DNA sequences using bioinformatics

MODULE OVERVIEW FOR EDUCATORS

Workflow	Essential Question	Learning Objectives	Research Objectives
1. Identifying a group of organisms for trade exploitation investigation	What family or genus of organisms are currently being traded in your area, and are individual species being exploited in this trade?	ED1, ED3, ED4 DC1, DC2, DC3, DC4 SR1, SR3 RI1, RI2, RI3, RI4	RO1, RO2
Utilize DNA barcoding curriculum developed by Harris & Bellino.			
2. Molecular Biology Theory and Techniques	How do we work with DNA and how is it used in wildlife forensics?	DC1, DC2, DC3, DC4	RO3
3. Generating DNA Barcodes	How can DNA barcodes be used to identify a species and how is this useful for wildlife trade?	DC1, DC2, DC3, DC4 RI1, RI2, RI3, RI4	RO3
4. Analyzing DNA Barcodes	How can existing molecular data help us interpret our DNA barcode results?	DC1, DC2, DC3, DC4 RI2, RI3, RI4	RO4
5. Developing new DNA barcoding research questions	How can we generate new research questions based on the findings of DNA barcoding in our wildlife trade research?	RL1, RL2, RL3, RL4 ED1, ED2, ED3, ED4 DC1 SR1, SR2, SR3 RI1, RI2, RI3, RI4	RO1, RO2, RO3, RO4

INSTRUCTIONAL INFORMATION

Identifying, collecting, and sampling a group of organisms in the wildlife trade

Example resources for identifying a group of organisms to study

1. *CITES: Convention on International Trade in Endangered Species of Wild Fauna and Flora*
   The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) is an international agreement between governments. Its aim is to ensure that international trade in specimens of wild animals and plants does not threaten their survival. While many wildlife species in trade are not endangered, the existence of an agreement to ensure the sustainability of the trade is important in order to safeguard these resources for the future. CITES protects more than 30,000 species of animals and plants, whether they are traded as live specimens, fur coats, or dried herbs. Students can join forums, read reports, and look through the dozens of visuals provided by CITES in order to identify potential avenues of investigation. http://cites.org/eng
2. *The IUCN Redlist of Endangered Species*
   The IUCN Red List of Threatened Species™ provides taxonomic, conservation status and distribution information on plants, fungi and animals that have been globally evaluated using the IUCN Red List Categories and Criteria. This system is designed to determine the relative risk of extinction, and the main purpose of the IUCN Red List is to catalogue and highlight those plants and animals that are facing a higher risk of global extinction (i.e. those listed as **Critically Endangered**, **Endangered** and **Vulnerable**).
   http://www.iucnredlist.org/about/introduction
3. The website, http://www.healthmap.org/wildlifetrade/, provides a customizable visualization of worldwide reports on interceptions of illegally traded wildlife and wildlife products.
4. TRAFFIC, the wildlife trade monitoring network, is the leading non-governmental organization working globally on trade in wild animals and plants in the context of both biodiversity conservation and sustainable development. http://www.traffic.org/overview/

Fieldwork and sample collection

Where to find samples

1. Walk through local traditional medicine markets and pay attention to the most common dried specimens you see listed for sale for medicinal purposes or look for anything that looks interesting or out of place.
2. Search online marketplaces like ebay. Use search terms like, "dried (your organism name)" or "real (your organism name)"
3. For marine organisms, websites like http://www.shellhorizons.com/products.asp?category=13 and http://www.shelloutlet.com/sealife.html sell preserved marine animals.
4. Use Internet searches to find physical curio shops in your local area. Often, souvenir shops with ocean themes, commonly in coastal areas, will carry different types of dried animals. These animals can range from starfish, to pufferfish to baby alligators!
5. Many unusual species can often be found in food markets, particularly in large cities that contain "Chinatowns."

Measurements, identification, and subsampling for genetics

How to process samples once they have been collected

1. Once you have collected several samples of your organism of interest, it is important to try to identify the species using morphological or physical measurements. Using Google Scholar, http://scholar.google.com, or Mendeley, http://www.mendeley.com/dashboard/, is a good way to get access to scientific articles that have dichotomous keys or identification protocols in place for your organism of interest.
2. Use these same scientific articles to find what measurements are generally taken for each individual collected from your organism of interest. For example, with small rodents, scientists measure tail length, ear length, and hind paw length. You will want to make sure that you properly measure and document your samples. Digital photographs are also a good idea.
3. For genetic sub-sampling, follow the DNA extraction guide in whatever kit you will be using. Depending on the state of your samples, you might need to modify procedures slightly. For example, often with dried specimens, the outside might be covered in a form of wax. To access DNA, a small incision should be made and dried tissue from inside the organism should be used for DNA extraction. If your organism is found in a food market it may be fresh and should be frozen until tissue samples are taken, or preserved in a fixative such as ethanol (70-95%). Some species will be dried and preserved in salt, therefore similar to specimens covered in a wax-like coating, it is important to try to take a tissue sample from within the species where DNA may be less degraded. The preservative substances may have the potential to inhibit downstream DNA amplification.
4. At this point, following The Student DNA Barcoding Project curriculum (Harris & Bellino, 2013) (http://www.sciencemag.org/content/342/6165/1462/suppl/DC1) will bring you from tissue sample to DNA sequence and species identification. A brief illustration of tissue sampling from seahorses is also included in the "Seahorse specific protocol" section (p. 10-11).
5. There are also other DNA barcoding protocols useful for novice molecular biologists. The Urban Barcode Project is an additional resource that describes the process of DNA extraction,

PCR amplification and sequencing with a user-friendly, illustrated, step-by-step. http://www.urbanbarcodeproject.org/files/using-dna-barcodes.pdf

6. Once DNA sequences have been obtained, it is important to identify the species collected. This is accomplished by using public databases. The National Center for Biotechnology Information (NCBI) curates genetic sequences submitted by scientists all over the world and contains many millions of DNA sequences. Their search program BLAST will take your DNA sequence generated from your sample and match it to the closest related species. http://www.ncbi.nlm.nih.gov/BLAST
7. The Barcode of Life Database (BOLD) is also a terrific resource for wildlife forensic identification using the CO1 gene, and species are added regularly. For example, in 2010 the BOLD database contained roughly 6560 fish species. As of August 2014, BOLD now contains 14,359 species of fishes. http://www.boldsystems.org/index.php/databases http://www.boldsystems.org/index.php/TaxBrowser_Home http://www.boldsystems.org/index.php/IDS_OpenIdEngine

Seahorse specific information and protocol

Once you have collected your specimens, follow the instructions above in the "Measurements, identification, and subsampling for genetics" section. Fortunately, If you are interested in the study of seahorses (*Hippocampus*), Project Seahorse developed "The Guide to the Identification of Seahorses" (Lourie et al., 2004), which is an invaluable resource. This guide contains illustrations and outlines the morphology and life history information for 33 species.

Many species of seahorses have a high degree of morphological plasticity, which led to species being over-split into various synonyms (i.e., new species that are not taxonomically valid or "true" species). These splits were generally based on morphological traits such as color and cirri (i.e., skin filaments that help species blend in with algal habitat). However, many species of seahorses can change these traits to blend into their surroundings, and therefore they are not appropriate indicators of species type. To correct for this problem, Lourie et al. (2004) used a combination of DNA sequencing and morphology to consolidate many species of seahorses that had been over-split. However, since publication several new species have been identified. Most of these are rare pygmy species that would not be found in trade market samples. Some new species have also been confirmed through genetic analysis, such as *H. patagonicus* from South America (Boehm et al. 2013). Therefore a new guide is being produced, however with the exception of *H. patagonicus*, it is unlikely that any species common to the dried seahorse trade would not be in the current version.

Download The Guide to the Identification of Seahorses:

http://www.traffic.org/species-reports/traffic_species_fish29.pdf

You can find key morphological terms and information for identifying species on pages 7-12. In the Appendix (p.93-103) there is also a seahorse identification data sheet that provides information on how to use a species checklist, with photographs of over 20 dried seahorse species. Prior to barcoding to try identifying the species you've collected using this checklist. Once you've gone through the procedure of molecular identification you can see if your results match with the species you identified through morphology.

Because seahorses are under protection and the sustainable harvest size is currently recommended as 10 cm, it is important to first measure your specimens before taking your DNA samples. Figure 1 below shows a picture of two seahorses. One is straight, and the other has a curled tail. One way to measure the "height" of the seahorse is to create a flexible measuring device. You can use a string, or flexible wire. Hold it along the length of a ruler and mark the measuring device to every ¼ centimeter. (Note: If you want to be extra meticulous you can mark off each millimeter). Place the seahorse on its lateral side and measure the height from the tip of the tail (posterior) to the top of the coronet (anterior). If the tail is curled use your flexible measuring device to follow the length of the tail to the coronet.

As previously outlined above, for genetic sub-sampling, follow the DNA extraction guide in whatever kit you will be using. Depending on the state of your samples, you might need to modify procedures slightly. For example, dried specimens might be covered in a form of wax or gloss to preserve the specimen or give them a shiny appearance. These materials can create issues with downstream processing, and may inhibit PCR amplification. To access DNA, a small incision should be made on one side of the seahorse. Because seahorses have a bilateral symmetry (the same morphology on each side), the side that has not been sampled will still retain its appearance. After removing part of the tail, dried tissue from inside the organism (the tissue below the outer skin) should be used for DNA extraction, and added to the DNA extraction tube. Figure 2 below demonstrates a step-by-step example.

Figure 1

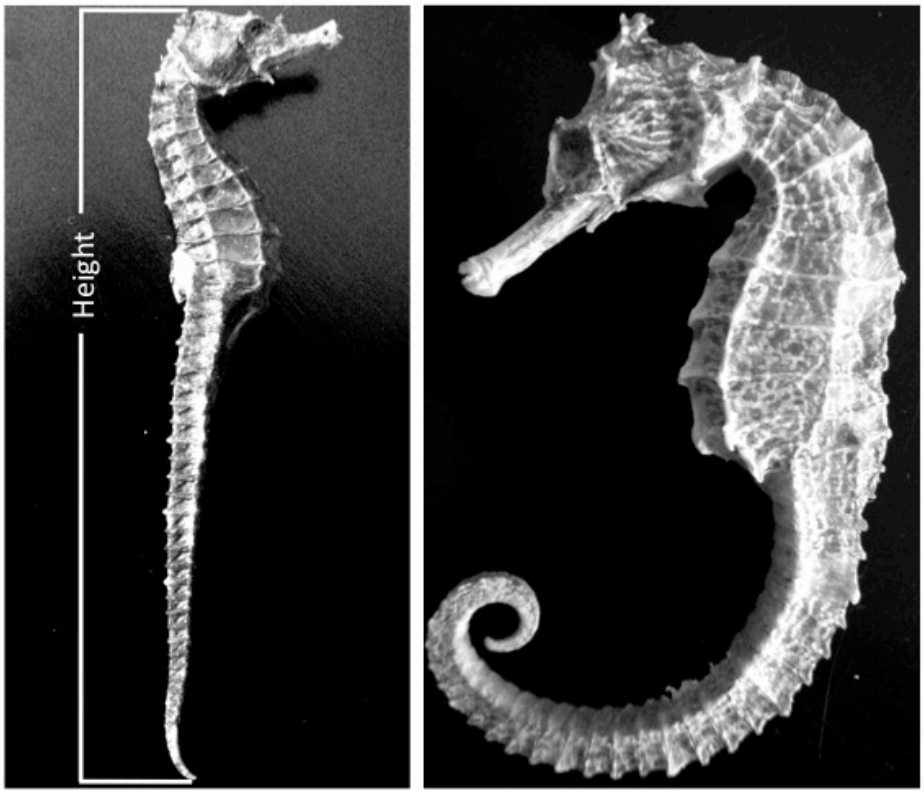
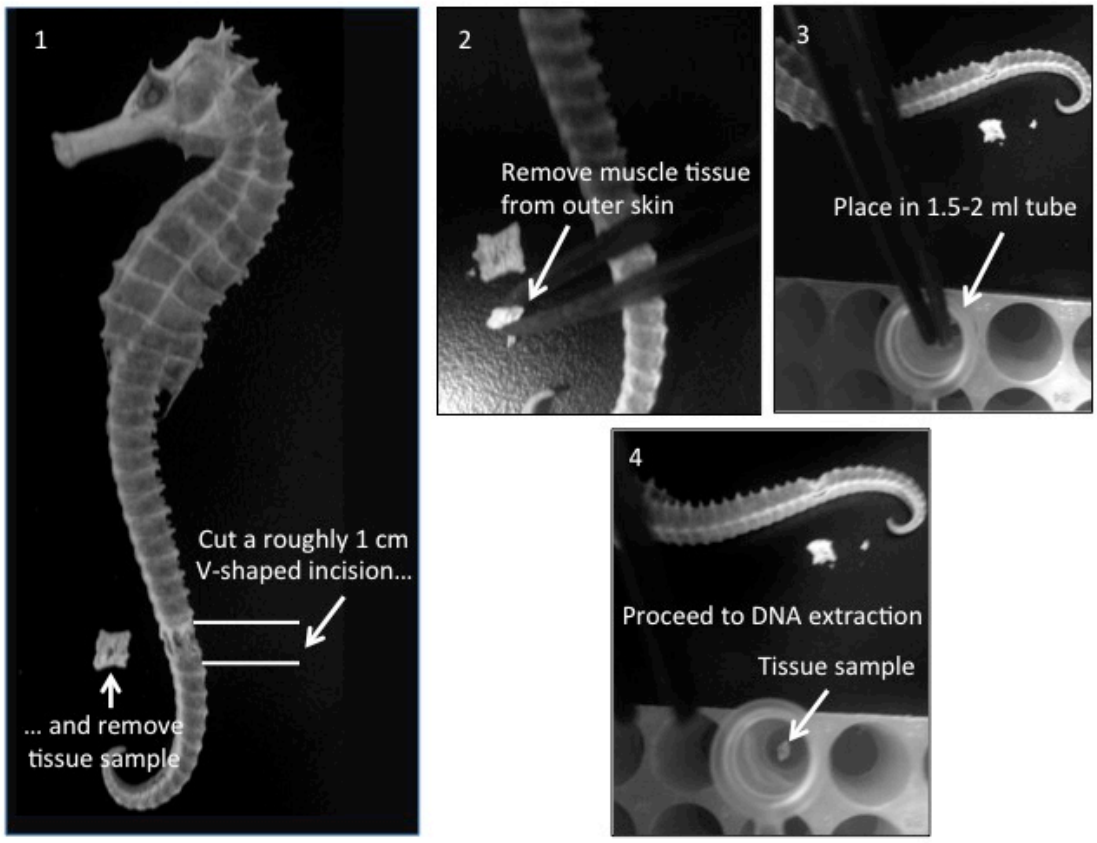


Figure 2

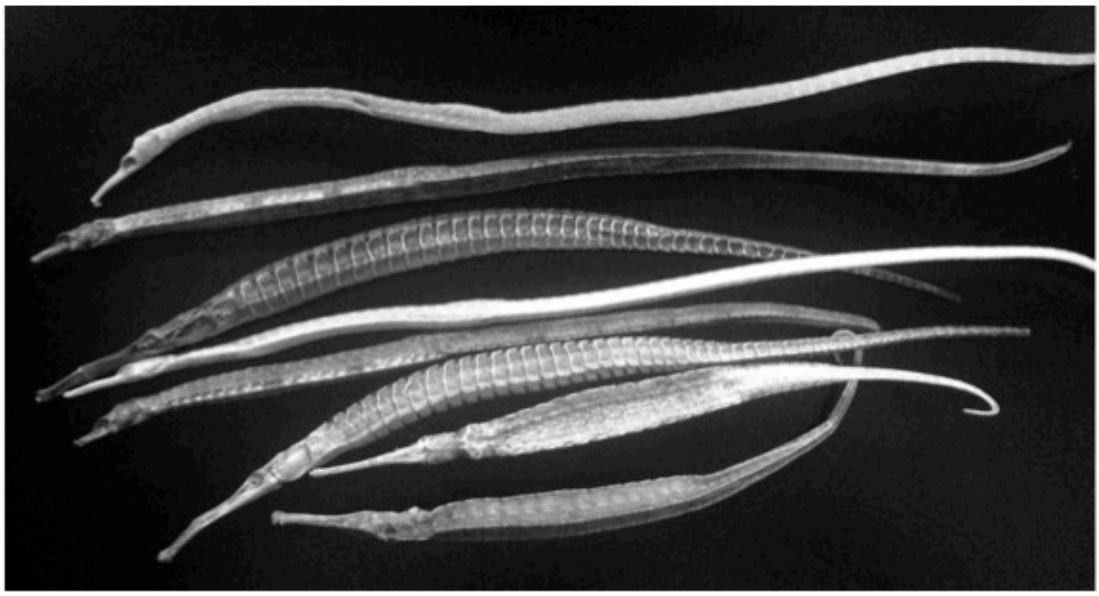


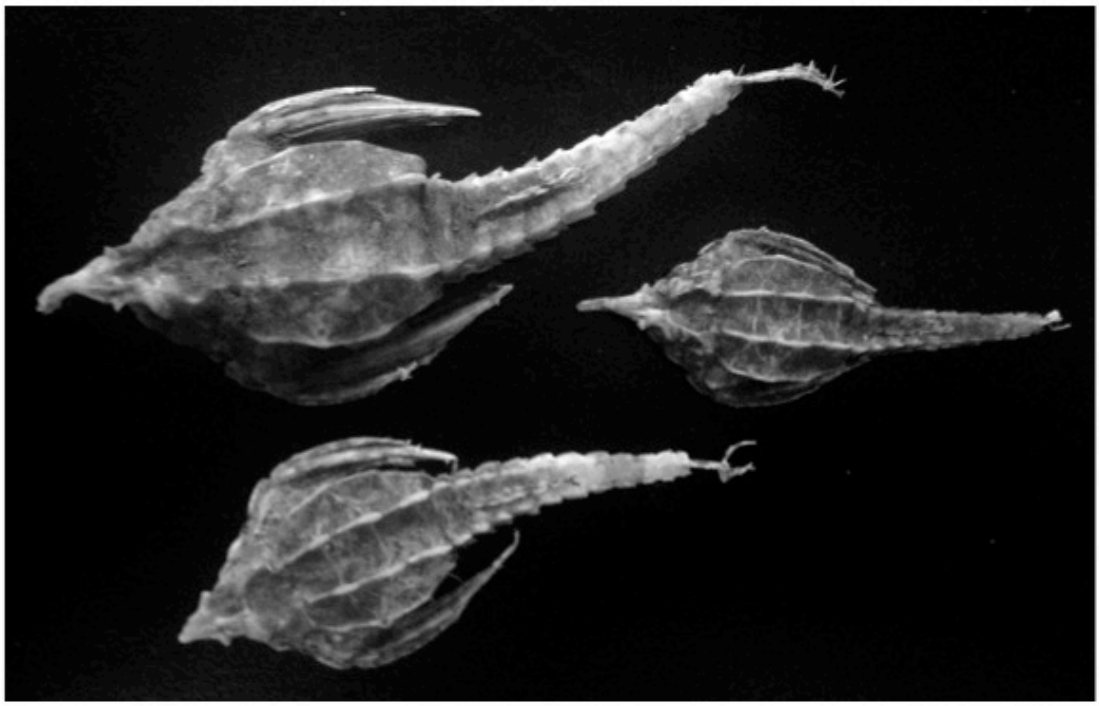
Other possible wildlife products: *Synganthidae and Pegasidae*:

Another avenue of research may be with one of the many species of Syngnathidae (pipefishes, pipehorses, seadragons) and their relatives Pegasidea (seamoths). Many of these species share similar life history traits (i.e., low population densities and habitat specificity) and habitat loss makes them vulnerable to overexploitation (Pajaro et al., 1997; Martin-Smith & Vincent, 2006). These species also serve similar functions in traditional medicine as *Hippocampus spp*. and are often combined for medicinal treatment (Martin-Smith & Vincent, 2006). The trade of Pegasidae fishes, pipefishes, and pipehorses has intensified, reaching millions of specimens per year. Examples from trade records show that Hong Kong imported 12.3 metric tonnes of pipefishes and pipehorses in 1998 and re-exported at least 1.9 tonnes to China and Taiwan. Taiwan alone imported more than 16.6 tonnes in 1993 and 10.3 tonnes in 1994 (Project Seahorse, 2014). Like with seahorses, the majority of these fish enter the TCM market as incidental bycatch from trawl fisheries leading to incomplete descriptions of imports, exports, and trade patterns (Vincent et al., 2011b). Consequently, it is relatively unknown which species are favored or collected in the global trade of Syngnathidae and Pegasidae species for traditional medicine use. Estimates of the number of oceanic pipefish, pipehorses and ghost pipefish species are >200, while Pegasidae are composed of five species (from two genera) (Nelson, 2006).

Thanks to the work of Project Seahorse (www.seahorse.fisheries.ubc.ca) the morphology of most *Hippocampus* species are well documented. However, the morphological identification of pipefish for species identification is often difficult, especially if the region of origin is unknown. Kuiter (2000) and Dawson (1985) are both comprehensive guides for detailed information regarding Syngnathidae species (Dawson, 1985; Kuiter, 2000). Below are some pictures of species that may be encountered in traditional Chinese medicine stores.

Pipefishes (picture 1) and Pegasidae (picture 2)





Appendix 4.2

Table 4.2. Information on collection sites and store type, specimens, and descriptions. Specified descriptions refer to terms used by sellers.

Sites	Store or shipping location	Individuals purchased	Store type	Description
Site 1	Florida	12	eBay	"Tiny" seahorse
Site 2	Florida	12	eBay	"Tiny"seahorse
Site 3	Florida	24	eBay	Dwarf seahorse
Site 4	Florida	10	eBay	Dried seahorse
Site 5	Florida	14	eBay	Dried seahorse
Site 6	Florida	24	eBay	“Pygmy” seahorse
Site 7	Texas	25	eBay	Real seahorse
Site 8	Texas	75*	eBay	Dwarf seahorse
Site 9	Florida	80**	Independent site	Dried seahorses
Site 10	Florida	24	Independent site	Dried seahorse
Store 1	Texas	17	Souvenir store	Dwarf seahorse
Store 2	Texas	10	Souvenir store	
Store 3	Texas	8	Souvenir store	
Store 4	Texas	9	Souvenir store	
Store 7	Manhattan	20	TCM	
Store 8	Manhattan	14	TCM	
Store 9	Manhattan	18	TCM	
Store 10	Philidelphia	10	TCM	
Store 11	Philidelphia	8	TCM	
Store 12	Brooklyn	10	TCM	
Store 13	Brookyn	12	TCM	
Store 14	Brooklyn	10	TCM	
Store 15	Queens	12	TCM	
Store 11	Queens	12	TCM	
Store 12	Queens	21	TCM	

Store 13	Queens	13	TCM	
Store 14	Queens	8	TCM	
Confiscated	FWL	20	N/A	

* Although additional large "lots" of *H. zostera* were available, after receiving one large "lot" (N=75) from an eBay source, we only purchased samples that could be bought in smaller quantities to reduce biasing the overall representation of any one species in our samples.

** Samples from "Site 9" were composed of 4 size classes of 20 individuals each. The photograph on the website showed several large seahorse species, however when we received the shipment all samples consisted of *H. kuda*.

BIBLIOGRAPHY

Able K. & Fahay M. (1998) *Ecology of Estuarine Fishes: Temperate Waters of the Western North Atlantic*. John Hopkins University Press.

Adams A.M. & Hudson R.R. (2004) Maximum-likelihood estimation of demographic parameters using the frequency spectrum of unlinked single-nucleotide polymorphisms. *Genetics*, **168**, 1699–712.

Aharon P. (2003) Meltwater flooding events in the Gulf of Mexico revisited: Implications for rapid climate changes during the last deglaciation. *Paleoceanography*, **18**.

Alacs E.A., Georges A., FitzSimmons N.N., & Robertson J. (2010) DNA detective: a review of molecular approaches to wildlife forensics. *Forensic science, medicine, and pathology*, **6**, 180–94.

Alexander D.H., Novembre J., & Lange K. (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome research*, **19**, 1655–64.

Alves R.R.N. & Rosa I.M.L. (2007) Biodiversity, traditional medicine and public health: where do they meet? *Journal of ethnobiology and ethnomedicine*, **3**, 14.

Anderson J., Karel W., & Mione A. (2012) Population structure and evolutionary history of southern flounder in the Gulf of Mexico and western Atlantic Ocean. *Transactions of the American Fisheries Society*, **141**, 46-55.

Anderson J., Wallace D., & Simms A. (2014) Variable response of coastal environments of the northwestern Gulf of Mexico to sea-level rise and climate change: Implications for future change. *Marine Geology*, **352**, 348-366.

Avila S.P., Marques Da Silva C., Schiebel R., Cecca F., Backeljau T., & De Frias Martins A.M. (2009) How did they get here? The biogeography of the marine molluscs of the Azores. *Bulletin de la Societe Geologique de France*, **180**, 295–307.

Avise J.C. (1992) Molecular Population Structure and the Biogeographic History of a Regional Fauna: A Case History with Lessons for Conservation Biology. *Oikos*, **63**, 62.

Avise J.C. (2000) *Phylogeography: The history and formation of species*. Harvard University Press, Cambridge.

Avise J.C. (2008) Phylogeography: retrospect and prospect. *Journal of Biogeography*, **36**, 3–15.

Avise J.C. & Walker D. (1998) Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **265**, 457–463.

Ayre D., Minchinton T., & Perrin C. (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Molecular ecology*, .

Baird N. a, Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z. a, Selker E.U., Cresko W. a, & Johnson E. a (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one*, **3**, e3376.

Baum J., Meeuwig J., & Vincent A. (2003) Bycatch of lined seahorses (*Hippocampus erectus*) in a Gulf of Mexico shrimp trawl fishery. *Fishery Bulletin*, **101**, 721–731.

Baum J.K. & Vincent A.C.J. (2005) Magnitude and inferred impacts of the seahorse trade in Latin America. *Environmental Conservation*, **32**, 305.

Beaumont M.A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology Evolution and Systematics*, **41**, 379–406.

Beaumont M.A., Zhang W., & Balding D.J. (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Bellwood D. & Wainwright P. (2002) The history and biogeography of fishes on coral reefs. Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem. Academic Press.

Bernatchez L. & Wilson C. (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, .

Blum M., Misner T., & Collins E. (2001) Middle Holocene sea-level rise and highstand at+ 2 m, central Texas coast. *Journal of Sedimentary Research*, **71**, 581-588.

Blumenthal J.M., Abrei-Grobois F.A., Austin T.J., Broderick A.C., Bruford M.W., Coyne M.S., Ebanks-Petrie G., Formia A., Meylan P.A., Meylan A.B., & Godley B.J. (2009) Turtle groups or turtle soup: dispersal patterns of hawkbill turtles in the Caribbean. *Molecular Ecology*, **18**, 4841–4853.

Boehm J.T., Waldman J.R., Robinson J.D., & Hickerson M.J. (*in press*) Population genomics reveals seahorses (Hippocampus erectus) of the western mid-Atlantic coast to be residents rather than vagrants. *PloS one*.

Boehm J.T., Woodall L., Teske P.R., Lourie S. A., Baldwin C., Waldman J., & Hickerson M. (2013) Marine dispersal and barriers drive Atlantic seahorse diversification. *Journal of Biogeography*, **40**, 1839–1849.

Bortone S., Hastings P., & Collard S. (1977) The pelagic-Sargassum ichthyofauna of the eastern Gulf of Mexico. *Northeast Gulf Science*, **1**, 60-67.

Bouckaert R. (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, **26**, 1372–1373.

Bratton J.F., Colman S.M., Thieler E.R., & Seal R.R. (2002) Birth of the modern Chesapeake Bay estuary between 7.4 and 8.2 ka and implications for global sea-level rise. *Geo-Marine Letters*, **22**, 188–197.

Briggs J.C. (1970) A faunal history of the North Atlantic ocean. *Systematic Zoology*, **19**, 19–34.

Briggs J.C. (1974) *Marine zoogeography*. McGraw Hill, New York.

Briggs J.C. (1995) *Global Biogeography*. Elsevier Science, Arnoldsville, GA, USA.

Briggs J.C. & Bowen B.W. (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography*, **39**, 12–30.

Briggs J.C. & Bowen B.W. (2013) Marine shelf habitat : biogeography and evolution. *Journal of Biogeography* **40**, 1–13.

Briggs P.T. & Waldman J.R. (2002) Annotated list of fishes reported from the marine waters of New York. *Northeastern Naturalist*, **9**, 47–80.

Bruckner A., Field J., & Davies N. (2005) The Proceedings of the international workshop on CITES implementation for seahorse conservation and trade. *NOAA Technical Memorandum NMFS-OPR*

Carstens B.C., Brennan R.S., Chua V., Duffie C. V, Harvey M.G., Koch R. a, McMahan C.D., Nelson B.J., Newman C.E., Satler J.D., Seeholzer G., Posbic K., Tank D.C., & Sullivan J. (2013) Model selection as a tool for phylogeographic inference: an example from the willow Salix melanopsis. *Molecular ecology*, **22**, 4014–28.

Casazza T.L. & Ross S.W. (2008) Fishes associated with pelagic Sargassum and open water lacking Sargassum in the Gulf Stream off North Carolina. *Fisheries Bulletin*, **106**, 348–363.

Casey S.P., Hall H.J., Stanley H.F., & Vincent A.C.J. (2004) The origin and evolution of seahorses (genus *Hippocampus*): a phylogenetic study using the cytochrome b gene of mitochondrial DNA. *Molecular Phylogenetics and Evolution*, **30**, 261–272.

Catchen J., Bassham S., Wilson T., Currey M., O'Brien C., Yeates Q., & Cresko W. a (2013) The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular ecology*, **22**, 2864–83.

Chambers J.M., Cleveland W.S., & Kleiner B. (1983) *Graphical Methods for Data Analysis (Statistics)*. Chapman and Hall/CRC,

Chassignet E. & Hurlburt H. (2007) The HYCOM (hybrid coordinate ocean model) data assimilative system. *Journal of Marine System*, **65**, 1-4.

Coates A.G., Jackson J.B.C., Collins L.S., Cronin T.M., Dowsett H.J., Bybell L.M., Jung P., & Obando J.A. (1992) Closure of the Isthmus of Panama: the near shore marine record of Costa Rica and west Panama. *Geological Society of America bulletin*, **104**, 814–828.

Collette B.B. & Rutzler K. (1977) Reef fishes over sponge bottoms off the mouth of the Amazon River. *Proceedings, Third International Coral Reef Symposium*, 305–310.

Collins R.A. & Cruickshank R.H. (2013) The seven deadly sins of DNA barcoding. *Molecular ecology resources*, **13**, 969–75.

Corander J., Majander K.K., Cheng L., & Merilä J. (2013) High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. *Molecular ecology*, **22**, 2931–40.

Cornuet J.-M., Ravigné V., & Estoup A. (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC bioinformatics*, **11**, 401.

Cornuet J.-M., Santos F., Beaumont M.A., Robert C.P., Marin J.-M., Balding D.J., Guillemaud T., & Estoup A. (2008) Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Crandall E.D., Treml E. a, & Barber P.H. (2012) Coalescent and biophysical models of stepping-stone gene flow in neritid snails. *Molecular ecology*, **21**, 5579–98.

Cronin T.M., Szabo B.J., Ager T.A., Hazel J.E., & Owens J.P. (1981) Quaternary climates and sea levels of the u.s. Atlantic coastal plain. *Science*, **211**, 233–40.

Cruickshank T.E. & Hahn M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular ecology*, **23**, 3133–57.

Cunningham C. & Collins T. (1998) Beyond area relationships: extinction and recolonization in molecular marine biogeography. Molecular approaches to ecology and evolution. Birkhäuser Basel.

Curran M. (1989) Occurrence of tropical fishes in New England waters. *American Academy of Underwater Sciences*.

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T., McVean G., & Durbin R. (2011) The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, **27**, 2156–8.

Davey J.W., Cezard T., Fuentes-Utrilla P., Eland C., Gharbi K., & Blaxter M.L. (2013) Special features of RAD Sequencing data: implications for genotyping. *Molecular ecology*, **22**, 3151–64.

Davey J.W., Hohenlohe P. a, Etter P.D., Boone J.Q., Catchen J.M., & Blaxter M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature reviews. Genetics*, **12**, 499–510.

Davis, J. (1992) Holocene coastal development on the Florida peninsula. *Special Publication of SEPM*.

Dawson C. (1985) *Indo-Pacific Pipefishes (Red Sea to the Americas)*. Coast Research Laboratory, Ocean Springs, Mississippi, USA.

Dawson C.E. & Vari R.P. (1983) Fishes of the western North Atlantic. Part eight. Order Gasterosteiformes, suborder Syngnathoidei. Syngnathidae (Doryrhamphinae, Syngnathinae, Hippocampinae). *Bulletin of Marine Science*, **33**, 517–518.

Deagle B.E., Jones F.C., Absher D.M., Kingsley D.M., & Reimchen T.E. (2013) Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. *Molecular ecology*, **22**, 1917–32.

Dias T.L., Rosa I., & Baum J.K. (2002) Threatened fishes of the world: *Hippocampus erectus* Perry, 1810 (Syngnathidae). *Environmental Biology of Fishes*, **65**, 326.

Dooley J. (1972) Fishes associated with the pelagic Sargassum complex with a discussion of Sargassum community. *Contributions in Marine Science*. University of Texas Marine Science Institute.

Doukakis P., Pikitch E.K., Rothschild A., DeSalle R., Amato G., & Kolokotronis S.O. (2012) Testing the effectiveness of an international conservation agreement: marketplace forensics and CITES caviar trade regulation. *PloS one*, **7**, e40907.

Drummond A., Ashton B., Buxton S., Cheung M., Cooper A., Duran C., Field M., Heled J., Kearse M., Markowitz S., Moir R., Stones-Havas S., Sturrock S., Thierer T., & Wilson A. (2011) Geneious v5.4. *Available at: http://www.geneious.com*.

Earl D.A. & vonHoldt B.M. (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Edwards S. V (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, **63**, 1–19.

Engelhardt B., Stephens M., & Walsh B. (2010) Analysis of Population Structure: A Unifying Framework and Novel Methods Based on Sparse Factor Analysis. *PLoS genetics*, ***6***, e1001117.

Evanno G., Regnaut S., & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology*, **14**, 2611–20.

Evanson M., Foster S., & Vincent A. (2011a) Tracking the international trade of seahorses (*Hippocampus* species)–the importance of CITES. *Fisheries Centre Research Reports*, ***19***. Vancouver, BC: Fisheries Centre, University of British Columbia.

Excoffier L., Dupanloup I., & Huerta-Sánchez E. (2013) Robust demographic inference from genomic and SNP data. *PLoS genetics*, **9**, e1003905.

Eytan R.I. & Hellberg M.E. (2010) Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution: International Journal of Organic Evolution*, **64**, 3380–3397.

Faurby S. & Barber P.H. (2012) Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Molecular ecology*, **21**, 3419–32.

Felsenstein J. (1981) Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. *Evolution*, **35**, 1229–1242.

Felsenstein J. (2006a) Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? *Molecular Biology and Evolution*, **23**, 691–700.

Figueiredo J., Hoorn C., van der Ven P., & Soares E. (2009) Late Miocene onset of the Amazon River and the Amazon deep-sea fan: evidence from the Foz do Amazonas Basin. *Geology*, **37**, 619–622.

Fine M. (1970) Faunal variation on pelagic *Sargassum*. *Marine Biology*, **7**, 112-122.

Fish, M.P. and Mowbray M.H. (1970) *Sounds of western North Atlantic fishes*. The Johns Hopkins University Press,

Floeter S.R., Rocha L.A., Robertson D.R., Joyeux J.C., Smith-Vaniz W.F., Wirtz P., Edwards A.J., Barreiros J.P., Ferreira C.E.L., Gasparini J.L., Brito A., Falcón J.M., Bowen B.W., & Bernardi G. (2007) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, **35**, 22–47.

Folmer O., Black M., Hoeh W., Lutz R., & Vrijenhoek R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology & Biotechnology*, **3**, 294–299.

Fossette S., Putman N.F., Lohmann K.J., Marsh R., & Hays G.C. (2012) A biologist's guide to assessing ocean currents: a review. *Marine ecology progress series*, **457**, 285–301.

Foster S.J. & Vincent A.C.J. (2004) Life history and ecology of seahorses: implications for conservation and management. *Journal of Fish Biology*, **65**, 1–61.

Foster S.J. & Vincent A.C.J. (2005) Enhancing Sustainability of the International Trade in Seahorses with a Single Minimum Size Limit. *Conservation Biology*, **19**, 1044–1050.

Fraser C.I., Nikula R., & Waters J.M. (2011) Oceanic rafting by a coastal community. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 649–55.

Frichot E., Mathieu F., Trouillon T., Bouchard G., & François O. (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics*, **196**, 973–83.

Gazave E., Chang D., Clark A.G., & Keinan A. (2013) Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. *Genetics*, **195**, 969–78.

Gillespie R.G., Baldwin B.G., Waters J.M., Fraser C.I., Nikula R., & Roderick G.K. (2012) Long-distance dispersal: a framework for hypothesis testing. *Trends in ecology & evolution*, **27**, 47–56.

Ginsburg I. (1937) Review of the seahorses (*Hippocampus*) found on the coasts of the American continents and of Europe. *Proceedings of the United States National Museum*, **83**, 497–594.

Glenn T.C. (2011) Field guide to next-generation DNA sequencers. *Molecular ecology resources*, **11**, 759–69.

Gompert Z., Lucas L.K., Buerkle C.A., Forister M.L., Fordyce J.A., & Nice C.C. (2014) Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular ecology*, **23**, 4555–73.

Gower J., Hu C., Borstad G., & King S. (2006) Ocean Color Satellites Show Extensive Lines of Floating Sargassum in the Gulf of Mexico. **44**, 3619–3625.

Gower J. & King S. (2008) Satellite Images Show the Movement of Floating Sargassum in the Gulf of Mexico and Atlantic Ocean.

Gower J.F.R. & King S. a. (2011) Distribution of floating Sargassum in the Gulf of Mexico and the Atlantic Ocean mapped using MERIS. *International Journal of Remote Sensing*, **32**, 1917–1929.

Grant W.S. & Bowen B.W. (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, **89**, 415–426.

Griffiths A.M., Miller D.D., Egan A., Fox J., Greenfield A., & Mariani S. (2013) DNA barcoding unveils skate (Chondrichthyes: Rajidae) species diversity in "ray" products sold across Ireland and the UK. *PeerJ*, **1**, e129.

Gronau I., Hubisz M.J., Gulko B., Danko C.G., & Siepel A. (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics*, **43**, 1031–4.

Grosberg C. & Cunningham C. (201AD) Genetic structure in the sea: from populations to communities. *Marine Community Ecology* (ed. by M. Bertness, S. Gaines, and M. Hay), pp. 61–84. Sinauer Associates,

Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., & Gascuel O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, **59**, 307–21.

Gunther A. (1870) *Catalogue of Fishes*. British Museum, London.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., & Bustamante C.D. (2009) Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLoS Genetics*, **5**, e1000695.

Hall M.A., Alverson D.L., & Metuzals K.I. (2000) By-Catch: Problems and Solutions. *Marine Pollution Bulletin*, **41**, 204–219.

Hare M.P., Guenther C., & Fagan W.F. (2005) Nonrandom dispersal can steepen marine clines. *Evolution*, **59**, 2509–2517.

Hare M.P., Nunney L., Schwartz M.K., Ruzzante D.E., Burford M., Waples R.S., Ruegg K., & Palstra F. (2011) Understanding and estimating effective population size for practical application in marine species management. *Conservation biology : the journal of the Society for Conservation Biology*, **25**, 438–49.

Harris S.E. & Bellino M. (2013) DNA Barcoding from NYC to Belize. *Science*, **342**, 1462–1463.

Harris S.E., O'Neill R.J., & Munshi-South J. (2014) Transcriptome resources for the white-footed mouse (*Peromyscus leucopus*): new genomic tools for investigating ecologically divergent urban and rural populations. *Molecular ecology resources*, .

Haug G.H. & Tiedemann R. (1998) Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. *Nature*, **393**, 673–676.

Hearn J., Stone G.N., Bunnefeld L., Nicholls J.A., Barton N.H., & Lohse K. (2014) Likelihood-based inference of population history from low coverage de novo genome assemblies. *Molecular Ecology*, **23**, 198–211.

Hebert P.D.N., Cywinska A., Ball S.L., & deWaard J.R. (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 313–21.

Hebert P.D.N., Ratnasingham S., & Jeremy R. (2003b) Barcoding animal life : cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences* 96–99.

Heled J. & Drummond A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–80.

Hellberg M.E. (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 291–310.

Hewitt G.M. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hey J. (2010) Isolation with migration models for more than two populations. *Molecular biology and evolution*, **27**, 905–20.

Hey J. & Nielsen R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747-760.

Hice L.A., Duffy T.A., Munch S.B., & Conover D.O. (2012) Spatial scale and divergent patterns of variation in adapted traits in the ocean. *Ecology letters*, **15**, 568–75.

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F., & Yoder A.D. (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.

Hickerson M.J. & Meyer C.P. (2008) Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical Bayesian approach. *BMC Evolutionary Biology*, **8**, 1–18.

Hickerson M.J., Stahl E.A., & Lessios H.A. (2006) Test for simultaneous divergence using approximate Bayesian computation. *Evolution*, **60**, 2435–2453.

Ho S.Y.W., Lanfear R., Bromham L., Phillips M.J., Soubrier J., Rodrigo A.G., & Cooper A. (2011a) Time-dependent rates of molecular evolution. *Molecular Ecology*, **20**, 3087–3101.

Ho S.Y.W., Lanfear R., Phillips M.J., Barnes I., Thomas J., Kolokotronis S., & Shapiro B. (2011b) Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Systematic biology*, **60**, 366–75.

Hobday A. (2000) Abundance and dispersal of drifting kelp *Macrocystis pyrifera* rafts in the Southern California Bight. *Marine Ecology Progress Series*, **195**, 101-116.

Hoelzel A.R. (2001) Shark fishing in fin soup. *Conservation Genetics*, **2**, 69–72.

Hoffmayer E. & Franks J. (2005) Larval and juvenile fishes associated with pelagic *Sargassum* in the northcentral Gulf of Mexico. *Proceedings of the Gulf and Caribbean Fisheries Institute,* **56**, 259-270.

Hohenlohe P. a, Amish S.J., Catchen J.M., Allendorf F.W., & Luikart G. (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular ecology resources*, **11**, 117–22.

Hohenlohe P. a, Bassham S., Etter P.D., Stiffler N., Johnson E. a, & Cresko W. a (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS genetics*, **6**, e1000862.

Holmes B.H., Steinke D., & Ward R.D. (2009) Identification of shark and ray fins using DNA barcoding. *Fisheries Research*, **95**, 280–288.

Hoorn C., Guerrero J., Sarmiento G.A., & Lorente M.A. (1995) Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology*, **23**, 237–240.

Horne J.B. (2014) Thinking outside the barrier: neutral and adaptive divergence in Indo-Pacific coral reef faunas. *Evolutionary Ecology*, **28**, 991–1002.

Howell P. & Auster P.J. (2012) Phase Shift in an Estuarine Finfish Community Associated with Warming Temperatures. *Marine and Coastal Fisheries*, **4**, 481–495.

Huang W., Takebayashi N., Qi Y., & Hickerson M.J. (2011) MTML-msBayes: approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC bioinformatics*, **12**, 1.

Ilves K.L., Huang W.E.N., Wares J.P., & Hickerson M.J. (2010) Colonization and/or mitochondrial selective sweeps across the North Atlantic intertidal assemblage revealed by multi-taxa approximate Bayesian computation. *Molecular Ecology*, **19**, 4505–4519.

Ingolfsson A. (1995) Floating clumps of seaweed around Iceland: natural microcosms and a means of dispersal for shore fauna. *Marine Biology*, **122**, 13–21.

IUCN (2014) Available at: www.iucnredlist.org.

Jahn A.E. (1976) On the midwater fish faunas of Gulf Stream rings with respect to habitat differences between slope water and northern Sargasso Sea. *Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Woods Hole Oceanographic Institution.*, 1–171.

Jeffreys H. (1961) *Theory of Probability*. Oxford University Press, USA,

Jensen J., AJ B., & Kelley S. (2005) Isolation by distance, web service. *BMC Genetics*, **6**, http://ibdws.sdsu.edu/.

Joyeux J.C., Floeter S.R., Ferreira C.E.L., & Gasparini J.L. (2001) Biogeography of tropical reef fishes: the South Atlantic puzzle. *Journal of Biogeography*, **28**, 831–841.

Kaneps A.G. (1979) Gulf Stream: velocity fluctuations during the late Cenozoic. *Science*, **204**, 297–301.

Keffer T., Martinson D.G., & Corliss B.H. (1988) The position of the Gulf Stream during quarternany glaciations. *Science, News Series*, **241**, 440–442.

Kennett J.P. & Shackleton N.J. (1975) Laurentide ice sheet meltwater recorded in gulf of Mexico deep-sea cores. *Science*, **188**, 147–50.

Kinlan B.P. & Gaines S.D. (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*, **84**, 2007–2020.

Kleckner R.C. & McCleave J.D. (1982) Entry of migrating American eel leptocephali into the Gulf Stream system. *Helgolander Meeresuntersuchungen*, **35**, 329–339.

Knowlton N. & Weigt L.A. (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 2257–2263.

Kuiter R. (2000) *Seahorses, Pipefishes and their relatives: a comprehensive guide to Syngnathiformes*. TCM Publishing, Chorleywood, UK.

Laffoley D., Roe H., Angel M., & Ardron J. (2011) The protection and management of the Sargasso Sea: The golden floating rainforest of the Atlantic Ocean: Summary Science and Supporting Evidence Case.

Langmead B., Trapnell C., Pop M., & Salzberg S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, **10**, R25.

Lazzari M.A. & Able K.W. (1990) Northern pipefish, *Syngnathus fuscus*, occurrences over the Mid-Atlantic Bight continental shelf: evidence of seasonal migration. *Environmental Biology of Fishes*, **27**, 177–185.

Lee T.N., Rooth C., Williams E., McGowan M., Szmant A.F., & Clarke M.E. (1992) Influence of Florida Current, gyres and wind-driven circulation on transport of larvae and recruitment in the Florida Keys coral reefs. *Continental Shelf Research*, **12**, 971–1002.

Lessios H.A. (2008) The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 63–91.

Lessios H.A. & Robertson D.R. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2201–2208.

Lett C., Verley P., Mullon C., & Parada C. (2008) A Lagrangian tool for modelling ichthyoplankton dynamics. *Environmental Modelling & Software*, 23, 1210-1214.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., & Durbin R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, **25**, 2078–9.

Li S. & Jakobsson M. (2012) Estimating demographic parameters from large-scale population genomic data using Approximate Bayesian Computation. *BMC Genetics*, **13**, 22.

Librado P. & Rozas J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Locke D.P., Hillier L.W., Warren W.C., Worley K.C., Nazareth L. V, Muzny D.M., Yang S.-P., Wang Z., Chinwalla A.T., Minx P., Mitreva M., Cook L., Delehaunty K.D., Fronick C., Schmidt H., Fulton L. a, Fulton R.S., Nelson J.O., Magrini V., Pohl C., Graves T. a, Markovic C., Cree A., Dinh H.H., Hume J., Kovar C.L., Fowler G.R., Lunter G., Meader S., Heger A., Ponting C.P., Marques-Bonet T., Alkan C., Chen L., Cheng Z., Kidd J.M., Eichler E.E., White S., Searle S., Vilella A.J., Chen Y., Flicek P., Ma J., Raney B., Suh B., Burhans R., Herrero J., Haussler D., Faria R., Fernando O., Darré F., Farré D., Gazave E., Oliva M., Navarro A., Roberto R., Capozzi O., Archidiacono N., Della Valle G., Purgato S., Rocchi M., Konkel M.K., Walker J. a, Ullmer B., Batzer M. a, Smit A.F. a, Hubley R., Casola C., Schrider D.R., Hahn M.W., Quesada V., Puente X.S., Ordoñez G.R., López-Otín C., Vinar T., Brejova B., Ratan A., Harris R.S., Miller W., Kosiol C., Lawson H. a, Taliwal

V., Martins A.L., Siepel A., Roychoudhury A., Ma X., Degenhardt J., Bustamante C.D., Gutenkunst R.N., Mailund T., Dutheil J.Y., Hobolth A., Schierup M.H., Ryder O. a, Yoshinaga Y., de Jong P.J., Weinstock G.M., Rogers J., Mardis E.R., Gibbs R. a, & Wilson R.K. (2011) Comparative and demographic analysis of orangutan genomes. *Nature*, **469**, 529–33.

Lohse K. & Frantz L.A.F. (2014) Neandertal admixture in Eurasia confirmed by maximum-likelihood analysis of three genomes. *Genetics*, **196**, 1241–51.

Lohse K., Harrison R.J., & Barton N.H. (2011) A general method for calculating likelihoods under the coalescent process. *Genetics*, **189**, 977–87.

Lourie S., Vincent A., & Hall H. (1999) *Seahorses: an identification guide to the world's species and their conservation*. Project Seahorse.

Lourie S.A. (2006) Spatial genetic patterns in the *Hippocampus barbouri* species group (Teleostei : Syngnathidae) across the Coral Triangle. *Proceedings of the 10th International Coral Reef Symposium*, 478–484.

Lourie S.A., Foster S.J., Cooper E.W.T., & Vincent A.C.J. (2004) A guide to the identification of seahorses. *Project Seahorse and TRAFFIC North America: University of British Columbia and World Wildlife Fund*.

Lourie S.A., Green D.M., & Vincent A.C.J. (2005) Dispersal, habitat differences, and comparative phylogeography of Southeast Asian seahorses (Syngnathidae: *Hippocampus*). *Molecular Ecology*, **14**, 1073–1094.

Lourie S.A. & Vincent A.C.J. (2004) A marine fish follows Wallace's Line: the phylogeography of the three-spot seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in Southeast Asia. *Journal of Biogeography*, **31**, 1975–1985.

Lowe W.H. & Allendorf F.W. (2010) What can genetics tell us about population connectivity? *Molecular ecology*, **19**, 3038–51.

Lozier J.D. (2014) Revisiting comparisons of genetic diversity in stable and declining species: assessing genome-wide polymorphism in North American bumble bees using RAD sequencing. *Molecular ecology*, **23**, 788–801.

Ludt W.B. & Rocha L.A. (2014) Shifting seas: the impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. *Journal of Biogeography*, n/a.

Luiz O.J., Madin J.S., Robertson D.R., Rocha L.A., Wirtz P., & Floeter S.R. (2012) Ecological traits influencing range expansion across large oceanic dispersal barriers: insights from tropical Atlantic reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 1033–40.

Luzzatto D.C., Rodrigo S., Maria P.G., & Juan D.A.M. Geographical range of the seahorse *Hippocampus patagonicus* in the Argentine Sea based on the Cytochrome b sequence of mitochondrial DNA. *Journal of Fish Biology (Submitted)*.

Marichamy R., Lipton A.P., Ganapathy A., & Ramalingam J.R. (1993) Large scale exploitation of sea horse (*Hippocampus kuda*) along the Palk Bay coast of Tamil Nadu. *Marine Fisheries Information Service, Technical and Extension Series*.

Marko P.B. (2004) "What"s larvae got to do with it?' Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. *Molecular Ecology*, **13**, 597–611.

Marko P.B. & Hart M.W. (2011a) Retrospective coalescent methods and the reconstruction of metapopulation histories in the sea. *Evolutionary Ecology*, **26**, 291-315.

Marko P.B. & Hart M.W. (2011b) The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, **26**, 448-456.

Marko P.B., Hoffman J.M., Emme S.A., McGovern T.M., Keever C.C., & Nicole Cox L. (2010) The "Expansion-Contraction" model of Pleistocene biogeography: rocky shores suffer a sea change? *Molecular Ecology*, **19**, 146–169.

Martin-Smith K.M. & Vincent A.C.J. (2006) Exploitation and trade of Australian seahorses, pipehorses, sea dragons and pipefishes (Family Syngnathidae). *Oryx*, **40**, 141.

Martinez-Solano I. & Gonzalez E.G. (2008) Patterns of gene flow and source-sink dynamics in high altitude populations of the common toad *Bufo bufo* (Anura: Bufonidae). *Biological Journal of the Linnean Society*, **95**, 824–839.

Masonjones H. & Lewis S. (1996) Courtship Behavior in the Dwarf Seahorse, *Hippocampus zosterae*. *Copeia*, **3**, 634–640.

Massatti R. & Knowles L.L. (2014) Microhabitat differences impact phylogeographic concordance of co-distributed species: Genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution,* **68**, 2833-2846.

McBride R.S. (2014) Managing a Marine Stock Portfolio: Stock Identification, Structure, and Management of 25 Fishery Species along the Atlantic Coast of the United States. *North American Journal of Fisheries Management*, **34**, 710–734.

McCartney M. a., Burton M.L., & Lima T.G. (2013) Mitochondrial DNA differentiation between populations of black sea bass (*Centropristis striata*) across Cape Hatteras, North Carolina (USA). *Journal of Biogeography*, **40**, 1386–1398.

McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., & Brumfield R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular phylogenetics and evolution*, **66**, 526–38.

Meeuwig J.J., Hoang D.H., Ky T.S., Job S.D., & Vincent A.C.J. (2006) Quantifying non-target seahorse fisheries in central Vietnam. *Fisheries Research*, **81**, 149–157.

Meyer C.P., Geller J.B., & Paulay G. (2005) Fine scale ednemism o coral reefs: Archipelagic differentiation in turbid gastropods . *Evolution*, **59**, 113–125.

Blum A. (2003) Middle Holocene Sea-Level and Evolution of the Gulf of Mexico Coast. **53**, 64–77.

Milstein C.B. & Thomas D.L. (1976) Fishes New or Uncommon to the New Jersey Coast. *Chesapeake Science*, **17**, 198.

Mobley K.B., Small C.M., & Jones A.G. (2011) The genetics and genomics of Syngnathidae: pipefishes, seahorses and seadragons. *Journal of fish biology*, **78**, 1624–46.

Mobley K.B., Small C.M., Jue N.K., & Jones A.G. (2010) Population structure of the dusky pipefish (*Syngnathus floridae*) from the Atlantic and Gulf of Mexico, as revealed by mitochondrial DNA and microsatellite analyses. *Journal of Biogeography*, **37**, 1363–1377.

Morgan S.K. & Panes H.M. (2007) Threatened fishes of the world: *Hippocampus spinosissimus* Weber 1913 (Syngnathidae). *Environmental Biology of Fishes*, **82**, 21–22.

Moritz C. (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Morton R., Paine J., & Blum M. (2000) Responses of stable bay-margin and barrier-island systems to Holocene sea-level highstands, western Gulf of Mexico. *Journal of Sedimentary Research*, **70**, 478-490.

Mullins H.T., Gardulski A.F., Wise S.W.. J., & Applegate J. (1987) Middle Miocene oceanographic event in the eastern Gulf of Mexico: Implications for seismic stratigraphic succession and Loop Current/Gulf Stream circulation. *Geological Society of America Bulletin*, **98**, 702–713.

Murugan A., Dhanya S., Rajagopal S., & Balalsubramanian T. (2008) Seahorses and pipefishes of the Tamil Nadu coast. *Current science*, **95**, 253–260.

Muss A., Robertson D.R., Stepien C.A., Wirtz P., & Bowen B.W. (2001) Phylogeography of Ophioblennius: the role of ocean currents and geography in reef fish evolution. *Evolution*, **55**, 561–72.

Myers R.A., & Ottensmeyer C.A. (2005) Extinction risk in marine species. In: Norse EA, Crowder LB (ed) In *Marine Conservation Biology: the Science of Maintaining the Sea's Biodiversity* Island Press, Washington, DC, pp 58-79.

Nelson J. (2006) *Fishes of the world*. Wiley Press.

Nielsen R. & Wakeley J. (2001) Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012) "SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data." *PloS one* 7.7: e37558.

Nosil P., Gompert Z., Farkas T.E., Comeault A.A., Feder J.L., Buerkle C.A., & Parchman T.L. (2012a) Genomic consequences of multiple speciation processes in a stick insect. *Proceedings of The Royal Society B*, **279**, 5058–65.

Novembre J., Johnson T., Bryc K., Kutalik Z., Boyko A.R., Auton A., Indap A., King K.S., Bergmann S., Nelson M.R., Stephens M., & Bustamante C.D. (2008) Genes mirror geography within Europe. *Nature*, **456**, 98–101.

Pajaro M., Vincent A., Buhat D., & Perante N. (1997) The role of seahorse fishers in conservation and management. *Proceedings of the 1st International Symposium in Marine Conservation*, 118–126.

Partridge C., Boettcher A., & Jones A.G. (2012) Population structure of the Gulf pipefish in and around Mobile Bay and the northern Gulf of Mexico. *The Journal of heredity*, **103**, 821–30.

Patterson N., Price A.L., & Reich D. (2006) Population structure and eigenanalysis. *PLoS genetics*, **2**, e190.

Paulay G. & Meyer C. (2002) Diversification in the Tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology*, **42**, 922–934.

Perry A.L., Lunn K.E., & Vincent A.C.J. (2010) Fisheries, large-scale trade, and conservation of seahorses in Malaysia and Thailand. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **20**, 464–475.

Piacentino G.L.M. & Luzzatto D.C. (2004) *Hippocampus patagonicus* sp. nov., nuevo cabllito de mar para la Argentina (Pisces, Syngnathiformes). *Revista del Museo Argentino de Ciencas Naturales*, **6**, 339–349.

Pickrell J.K. & Pritchard J.K. (2012) Inference of population splits and mixtures from genome-wide allele frequency data. **8**, 28.

Pinho C. & Hey J. (2010) Divergence with Gene Flow: Models and Data. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 215–230.

Portnoy D.S. & Gold J.R. (2012) Evidence of multiple vicariance in a marine suture-zone in the Gulf of Mexico. *Journal of Biogeography*, **39**, 1499–1507.

Pringle J.M., Wares J.P., & Building L.S. Going against the flow : Maintenance of alongshore variation in allele frequency in a coastal ocean. *Marine Ecology Progress Series*, **335**, 69–84.

Pritchard J.K., Stephens M., & Donnelly P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945–959.

Project Seahorse. (2014) Available at: seahorse.fisheries.ubc.ca

Putman N., Scanlan M., & Billman E. (2014) An inherited magnetic map guides ocean navigation in juvenile Pacific salmon. *Current Biology*, **24**, 446-450.

Putman N., Scott R., Verley P., Marsh R., & Hays G. (2012a) Natal site and offshore swimming influence fitness and long-distance ocean transport in young sea turtles. *Marine biology*, **159**, 2117-2126.

Putman N.F., Endres C.S., Lohmann C.M.F., & Lohmann K.J. (2011) Longitude perception and bicoordinate magnetic maps in sea turtles. *Current biology : CB*, **21**, 463–6.

Putman N.F. & He R. (2013) Tracking the long-distance dispersal of marine organisms : sensitivity to ocean model resolution Tracking the long-distance dispersal of marine organisms : sensitivity to ocean model resolution. *Journal of the Royal Society Interface*, **10.81**: 20120979.

Putman N.F. & Naro-Maciel E. (2013) Finding the "lost years" in green turtles: insights from ocean circulation models and genetic analysis. *Proceedings. Biological sciences / The Royal Society*, **280**, 20131468.

Putman N.F., Shay T.J., & Lohmann K.J. (2010) Is the geographic distribution of nesting in the Kemp's ridley turtle shaped by the migratory needs of offspring? *Integrative and comparative biology*, **50**, 305–14.

Putman N.F., Verley P., Shay T.J., & Lohmann K.J. (2012b) Simulating transoceanic migrations of young loggerhead sea turtles: merging magnetic navigation behavior with an ocean circulation model. *The Journal of experimental biology*, **215**, 1863–70.

De Queiroz K. (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, **20**, 68–73.

R Development Core Team (2004) R: A Language and Environment for Statistical Computing.

Reynolds J.D., Dulvy N.K., Goodwin N.B., & Hutchings J.A. (2005) Biology of extinction risk in marine fishes. *Proceedings of the Royal Society B: Biological Sciences,* **272**, 2337–44.

Riginos C., Douglas K.E., Jin Y., Shanahan D.F., & Treml E.A. (2011) Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, **34**, 566–575.

Robert C.P., Cornuet J.-M., Marin J.-M., & Pillai N.S. (2011) Lack of confidence in approximate Bayesian computation model choice. *Proceedings of the National Academy of Sciences*, **108**, 15112–7.

Roberts J.J., Best B.D., Dunn D.C., Treml E.A., & Halpin P.N. (2010) Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software*, **25**, 1197–1207.

Robertson D.R., Karg F., Leao R., Moura D., Victor B.C., & Bernardi G. (2006) Mechanisms of speciation and faunal enrichment in Atlantic parrotfishes. *Molecular Phylogenetics and Evolution*, **40**, 795–807.

Robinson J., Bunnefeld L., Hearn J., Stone G., & Hickerson M.J. (2014) ABC inference of multi-population divergence with admixture from un-phased population genomic data. *Molecular ecology*, **23,** 4458-4471.

Rocha L.A. (2003) Patterns of distribution and processes of speciation in Brazilian reef fishes. *Journal of Biogeography*, **30**, 1161–1171.

Rocha L.A., Bass A.L., Robertson D.R., & Bowen B.W. (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). *Molecular Ecology*, **11**, 243–252.

Rocha L.A. & Bowen B.W. (2008) Speciation in coral-reef fish. *Journal of Fish Biology*, 1101–1121.

Rocha L.A., Craig M.T., & Bowen B.W. (2007) Phylogeography and the conservation of coral reef fishes. *Coral Reefs*, **26**, 501–512.

Rocha L.A., Robertson D.R., Rocha C.R., Van Tassell J.L., Craig M.T., & Bowen B.W. (2005a) Recent invasion of the tropical Atlantic by an Indo-Pacific coral reef fish. *Molecular ecology*, **14**, 3921–8.

Rocha L.A., Robertson D.R., Roman J., & Bowen B.W. (2005b) Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 573–579.

Rocha L.A., Rocha C.R., Robertson D.R., & Bowen B.W. (2008) Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. *BMC Evolutionary Biology*, **8**, 157–172.

Roux C., Tsagkogeorga G., Bierne N., & Galtier N. (2013) Crossing the species barrier: genomic hotspots of introgression between two highly divergent Ciona intestinalis species. *Molecular biology and evolution*, **30**, 1574–87.

Sanders J.G., Cribbs J.E., Fienberg H.G., Hulburd G.C., Katz L.S., & Palumbi S.R. (2007) The tip of the tail: molecular identification of seahorses for sale in apothecary shops and curio stores in California. *Conservation Genetics*, **9**, 65–71.

Schwartz J. (2008) Fish tale has DNA hook: students find bad labels. *New York Times*.

Selkoe K.A., Gaggiotti O.E., Bowen B.W., & Toonen R.J. (2014) Emergent patterns of population genetic structure for a coral reef community. *Molecular ecology*, **23**, 3064–79.

Short F., Carruthers T., Dennison W., & Waycott M. (2007) Global seagrass distribution and diversity: A bioregional model. *Journal of Experimental Marine Biology and Ecology*, **350**, 3–20.

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.
Smith-Vaniz W.F., Collette B.B., & Luckhurst B.E. (1999) *Fishes of Bermuda*. American Society of Ichthyologists and Herpetologists, USA.
Sousa V. & Hey J. (2013) Understanding the origin of species with genome-scale data: modelling gene flow. *Nature reviews. Genetics*, **14**, 404–14.
Staaterman E. & Paris C.B. (2013) Modelling larval fish navigation: the way forward. *ICES Journal of Marine Science*, **71**, 918–924.
Staaterman E., Paris C.B., & Helgers J. (2012) Orientation behavior in fish larvae: a missing piece to Hjort's critical period hypothesis. *Journal of theoretical biology*, **304**, 188–96.
Steinke D., Zemlak T.S., & Hebert P.D.N. (2009) Barcoding nemo: DNA-based identifications for the ornamental fish trade. *PloS one*, **4**, e6300.
Strawn K. (1958) Life History of the Pigmy Seahorse, *Hippocampus zosterae* Jordan and Gilbert, at Cedar Key, Florida. *Copeia*, **1**, 16–22.
Tabone W. (2011) Ground Truthing Sargassum in Satellite Imagery: Assessment of Its Effectiveness as an Early Warning System. .
Teixeira R. & Musick J. (2001) Reproduction and food habits of the lined seahorse, *Hippocampus erectus* (Teleostei: Syngnathidae) of Chesapeake Bay, Virginia. *Revista Brasileira de Biologia*, **61**, 79-90.
Teske P., Hamilton H., Palsbøll P., Choo C., Gabr H., Lourie S., Santos M., Sreepada A., Cherry M., & Matthee C. (2005) Molecular evidence for long-distance colonization in an Indo-Pacific seahorse lineage. *Marine Ecology Progress Series*, **286**, 249–260.
Teske P.R., & Beheregaray L.B. (2009) Evolution of seahorses' upright posture was linked to Oligocene expansion of seagrass habitats. *Biology Letters*, **5**, 521–523.
Teske P.R., Cherry M.I., & Matthee C.A. (2003) Population genetics of the endangered Knysna seahorse, Hippocampus capensis. *Molecular Ecology*, **12**, 1703–1715.
Teske P.R., Cherry M.I., & Matthee C.A. (2004) The evolutionary history of seahorses (Syngnathidae: *Hippocampus*): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. *Molecular Phylogenetics and Evolution*, **30**, 273–286.
Teske P.R., Hamilton H., Matthee C.A., & Barker N.P. (2007) Signatures of seaway closures and founder dispersal in the phylogeny of a circumglobally distributed seahorse lineage. *BMC Evolutionary Biology*, **7**, 138–157.
Theil M. & Gutow L. (2005a) The ecology of rafting in the marine environment. I. The floating substrata. *Oceanography and Marine Biology: Annual Review*, **42**, 181–264.
Theil M. & Gutow L. (2005b) The ecology of rafting in the marine environment. II. The rafting organisms and community. *Oceanography and Marine Biology: An Annual Review*, **43**, 279-418.
Treml E., Halpin P.N., Urban D.L., & Pratson L.F. (2007) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology*, **23**, 19–36.
Tyberghein L., Verbruggen H., Pauly K., Troupin C., Mineur F., & De Clerck O. (2012) Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, **21**, 272–281.

Ultsch G. (1989) Ecology and physiology of hibernation and overwintering among freshwater fishes, turtles, and snakes. *Biological Reviews*, **64**, 435–515.
Vandendriessche S. & Messiaen M. (2007) Hiding and feeding in floating seaweed: floating seaweed clumps as possible refuges or feeding grounds for fishes. *Estuarine, Coastal and Shelf Science,* **71**, 691-703.
Vangriesheim A., Bournot-Marec C., & Fontan A.-C. (2003) Flow variability near the Cape Verde frontal zone (subtropical Atlantic Ocean). *Oceanologica Acta*, **26**, 149–159.
Vermeij G.J. (1978) *Biogeography and adaptation: patterns of marine life*. Harvard University Press, Cambridge, MA.
Vincent A.C.J. (1996) The international trade in seahorses. *TRAFFIC International, Network Report*.
Vincent A.C.J., Foster S.J., & Koldewey H.J. (2011a) Conservation and management of seahorses and other Syngnathidae. *Journal of fish biology*, **78**, 1681–724.
Vincent A.C.J., Giles B.G., Czembor C.A., & Foster S.J. (2011b) Trade in seahorses and other Syngnathids in countries outside Asia (1998-2001). *Fisheries Centre Research Report*, **19**, 1–181.
Vincent A.C.J. & Hall H.J. (1996) The threatened status of marine fishes. *Trends in Ecology & Evolution*, **11**, 360–361.
Vincent A.C.J., Meeuwig J.J., Pajaro M.G., & Perante N.C. (2007) Characterizing a small-scale, data-poor, artisanal fishery: Seahorses in the central Philippines. *Fisheries Research*, **86**, 207–215.
Vincent A.C.J., Sadovy de Mitcheson Y.J., Fowler S.L., & Lieberman S. (2013) The role of CITES in the conservation of marine fishes subject to international trade. *Fish and Fisheries*.
Wagner C.E., Keller I., Wittwer S., Selz O.M., Mwaiko S., Greuter L., Sivasundar A., & Seehausen O. (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular ecology*, **22**, 787–98.
Waltari E. & Hickerson M.J. (2012) Late Pleistocene species distribution modelling of North Atlantic intertidal invertebrates. *Journal of Biogeography*, in press.
Wang Y. & Hey J. (2010) Estimating divergence parameters with small samples from a large number of loci. *Genetics*, **184**, 363–79.
Wares J.P. & Cunningham C.W. (2001) Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution; international journal of organic evolution*, **55**, 2455–69.
Wares J.P., Gaines S.D., & Cunningham C.W. (2001) A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution*, **55**, 295–306.
Weersing K. & Toonen R. (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**, 1–12.
Weir B.S. & Cockerhan C.C. (1984) Estimation of gene flow from F-statistics. *Evolution*, **38**, 1358–1370.
Weis J.S. (1968) Fauna Associated with Pelagic *Sargassum* in the Gulf Stream. *American Midland Naturalist*, **80**, 554–558.

Wells R.J.D. & Rooker J.R. (2004) Spatial and temporal patterns of habitat use by fishes associated with *Sargassum* mats in the northwestern Gulf of Mexico. *Bulletin of Marine Science*, **74**, 81–99.

Wenner C.A. & Sedberry G.R. (1989) Species composition, distribution, and relative abundance of fishes in the coastal habitat off the southeastern United States. 49 pp.

Whitlock M.C. & McCauley D.E. (1999) Indirect measures of gene flow and migration: Fst≠1/(4Nm+1). *Heredity*, **82**, 117–125.

Wicklund R., Wilk S., & Ogren L. (1968) Observations on wintering locations of the northern pipefish and spotted seahorse. *Underwater Naturalist*, **5**, 26-28.

Willing E.-M., Dreyer C., & van Oosterhout C. (2012) Estimates of genetic differentiation measured by F(ST) do not necessarily require large sample sizes when using many SNP markers. *PloS one*, **7**, e42649.

Wilson A.B. & Orr J.W. (2011) The evolutionary origins of Syngnathidae: pipefishes and seahorses. *Journal of Fish Biology*, **78**, 1603–1623.

Wood E. (2001) Collection of coral reef fish for aquaria: global trade, conservation issues and management strategies. *Marine Conservation Society*. 80pp.

Woodall L.C., Koldewey H.J., Santos S. V, & Shaw P.W. (2009) First occurrence of the lined seahorse *Hippocampus erectus* in the eastern Atlantic Ocean. *Journal of Fish Biology*, **75**, 1505–1512.

Woodall L.C., Koldewey H.J., & Shaw P.W. (2011) Historical and contemporary population genetic connectivity of the European short-snouted seahorse *Hippocampus hippocampus* and implications for management. *Journal of Fish Biology*, **78**, 1738–1756.

Worm B., Barbier E.B., Beaumont N., Duffy J.E., Folke C., Halpern B.S., Jackson J.B.C., Lotze H.K., Micheli F., Palumbi S.R., Sala E., Selkoe K.A., Stachowicz J.J., & Watson R. (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science (New York, N.Y.)*, **314**, 787–90.

Worm B., Hilborn R., Baum J.K., Branch T.A., Collie J.S., Costello C., Fogarty M.J., Fulton E.A., Hutchings J.A., Jennings S., Jensen O.P., Lotze H.K., Mace P.M., McClanahan T.R., Minto C., Palumbi S.R., Parma A.M., Ricard D., Rosenberg A.A., Watson R., & Zeller D. (2009) Rebuilding global fisheries. *Science*, **325**, 578–85.