Phylogeography of the snake pipefish, *Entelurus aequoreus* (Family: Syngnathidae) in the northeastern Atlantic Ocean

Ines Braga Goncalves¹, Luca Cornetti², Abraham S. Couperus³, Cindy J.G. van Damme³, Kenyon B. Mobley⁴

¹School of Biological Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK
²Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland
³Wageningen Marine Research (IMARES), Postbus 68, 1970 AB Ijmuiden, The Netherlands
⁴Department of Evolutionary Ecology, Max Planck Institute for Evolutionary Biology, August-Thienemann Str. 2, 24306, Plön, Germany

Corresponding author: Kenyon Mobley
Email: mobley@evolbio.mpg.de
Phone: +49 4522.763.347

Running title: Phylogeography of the snake pipefish
ABSTRACT

The snake pipefish, *Entelurus aequoreus*, is a widespread marine species occurring in pelagic and coastal environments in the northeastern Atlantic Ocean. Recently, the snake pipefish underwent a short-lived, yet substantial, increase in abundance and range expansion into arctic waters. However, little is known about the species’ population structure or if different ecotypes contributed to this outbreak. Specimens (n=178) were sampled from 25 locations from six regions spanning 1.9 million km². A fragment of the mitochondrial cytochrome *b* gene and control region was used to assess population structure and genetic diversity. Both loci showed high haplotype diversity (Hd) and low nucleotide diversity (π) over all sampled locations. A genetic signature of population expansion was evident through mismatch distributions and tests for recent population expansion (Fu’s Fₛ, Tajima’s D, and R²). Effective population size analyses (Bayesian Skyline Plot) suggest an ancient expansion (50-100 thousand years ago). However, we found neither significant population differentiation (AMOVA) among regions, nor evidence of genetically distinct ecotypes. This lack of structure is likely due to a pelagic life style, fast development and long distance dispersal aided by ocean currents. Our work highlights the need for further research to better understand the recent outbreak and how this species may respond to future environmental challenges.

Keywords: Bayesian Skyline Plot, control region, cytochrome *b*, fish, life history, mitochondrial DNA, pelagic, population increase, population structure, range expansion
Introduction

The introduction of molecular techniques to the study of species distributions has greatly improved our knowledge regarding species’ evolutionary history. Current genetic tools allow scientists to assess historical species distribution patterns, gene flow between populations, the identification of source populations, routes and patterns of invasion, and timing of range shifts, expansions as well as contractions (Hewitt, 1999; Schmitt, 2007; Moran & Alexander, 2014). For example, the assessment of levels of gene flow and genetic diversity between geographically separate populations is of particular interest for conservation practises because it can facilitate the identification of genetically independent populations or units, and guide conservation efforts (e.g., Bernard, Feldheim, Heithaus et al., 2016). These studies highlight the importance of linking genetic data with species historical and geographical information to make biologically relevant interpretations. The discipline of phylogeography does exactly that by looking at geographical patterns of genetic diversity across populations or species over time (Avise, 2000).

In this study, we investigate the phylogeography of the snake pipefish, *Entelurus aequoreus* L. 1758. This species is a member of the family Syngnathidae, the pipefishes, seahorses and seadragons, a group characterized by its unique form of reproduction, male pregnancy (Dawson, 1986). Life history traits appear to shape the past and present geographical distributions of syngnathids (Mobley, Small & Jones, 2011). Syngnathids tend to be poor swimmers with small fins and armoured bodies, and most are strict habitat specialists, relying heavily on crypsis for survival (Vincent, Berglund & Ahnesjö, 1995). All species in this family produce free-living juveniles (Hasse, 1974; Mi, Kornienko & Drozdov, 1998; Monteiro, Almada & Vieira, 2003; Wetzel & Wourms, 2004) with short (Wilson & Vincent, 1998; Planas, Blanco, Chamorro et al., 2012) or entirely absent pelagic dispersal phases (Silva, Monteiro, Almada et al., 2006; Silva, Monteiro, Vieira et al., 2006). Having a
short pelagic dispersal phase is known to significantly limit dispersal potential (Grantham, Eckert & Shanks, 2003). Among adults, limited seasonal vertical migrations can occur, with individuals of some species coming into warmer shallow waters for mating, and returning to deeper waters at the end of the breeding season or during brooding (Vincent et al., 1995; Monteiro, Berglund, Vieira et al., 2006).

Previous studies on syngnathids have generally revealed relatively high genetic diversity and high population structuring, indicative of large effective population sizes and low dispersal ability (reviewed in Mobley et al., 2011). Most northern-hemisphere temperate syngnathid species show evidence of population expansions towards northern regions after the end of the last glaciation period (circa 20 thousand years before present (ka); Woodall, Koldewey, Santos et al., 2009; Mobley, Small, Jue et al., 2010; Wilson & Eigenmann Veraguth, 2010; Woodall, Koldewey, Boehm et al., 2015) following similar northerly expansion of suitable habitat, i.e., seagrass meadows (Olsen, Stam, Coyer et al., 2004). Yet, their strong habitat dependency and limited dispersal capability is also reflected in their current geographical distributions that generally show high levels of population differentiation on relatively small geographical scales (e.g. Lourie, Green & Vincent, 2005; Wilson, Stiller & Rouse, 2016; Stiller, Wilson, Donnellan et al., 2017) and strong indications of limited dispersal (e.g. Chenoweth, Hughes & Connolly, 2002; reviewed in Mobley et al., 2011; Wilson & Orr, 2011). Exceptions to these patterns have been interpreted as a result of recent colonization events (Nickel & Cursons, 2012), assumed to occur via rafting, where individuals drift with floating marine vegetation on ocean currents (Teske, Hamilton, Palsbøll et al., 2005; Fedrizzi, Stiassny, Boehm et al., 2015).

From a phylogeographical perspective, the snake pipefish is an interesting species to study because of its wide geographic distribution, recent range expansion into polar waters, and dramatic fluctuations in abundance. Historically, this species inhabits a vast geographical
range in the temperate northeastern Atlantic, spanning from the European continental shelf to the west, southern Norway and Iceland in the north, to the Azores in the south and the Baltic Sea to the east (Dawson, 1986). Normally, the species is encountered throughout its range at low densities (Harris, Beare, Toresen et al., 2007). However, between 2003 and 2007 the snake pipefish reappeared in large numbers in coastal areas where it had for decades been considered rare (e.g. northern Wadden Sea, Polte & Buschbaum, 2008), including brackish estuarine waters (e.g. the Severn Estuary, Henderson & Bird, 2010). During this time, the snake pipefish was caught in numbers several orders of magnitude higher than in catches prior to 2003 (Kloppmann & Ulleweit, 2007; van Damme & Couperus, 2008). The snake pipefish also underwent a geographical range expansion into the Barents and Greenland Seas by 2005 (Harris et al., 2007; Rusyaev, Dolgov & Karamushko, 2007) and the first ever records of occurrence in Svalbard were reported by August 2006, representing a 15° latitudinal expansion northwards (approximately 1650km, Fleischer, Schaber & Piepenburg, 2007). After 2007, populations of the snake pipefish declined dramatically and returned to low levels of abundance throughout its geographic range (Heath, Neat, Pinnegar et al., 2012).

There is anecdotal evidence that this species has undergone mass mortality events off the European continental shelf in the Atlantic Ocean in 1885, 1887, and in the North Sea in 1911, although no satisfactory explanation for these events exists (Brongersma-Sanders, 1957). It is possible that these mass mortality events are an indication of previous population increases in snake pipefish abundance although data to corroborate this link are currently lacking.

The snake pipefish is also an interesting species to investigate phylogeographically because of its unique life history traits. Unlike most other syngnathids that are predominantly associated with benthic habitats, the snake pipefish is described primarily as an oceanic species displaying a pelagic lifestyle and is found in both coastal and oceanic waters to depths up to 100m (Dawson, 1986; Kloppmann & Ulleweit, 2007). Because this species does
not require benthic habitats to reproduce they may breed in open waters and offspring as well as adults may be transported and mixed by ocean currents.

Finally, the potential for cryptic species to contribute to the temporary population increase and range expansion of the snake pipefish has not yet been addressed. Previously, two species have been proposed for *E. aequoreus* over the years based on differences in body size, colour, position of the dorsal fin and in number of rays (Yarrel, 1839; Moreau, 1881; Fries, Ekström & Sundevall, 1895; Holt & Byrne, 1906; Dunker, 1915). However, these two species are currently not recognized (Dawson, 1986). Additional lines of investigation have suggested coastal and oceanic habitat-specific phenotypes, or ‘ecotypes’, based on morphology (van Damme & Couperus, 2008) or timing of breeding between coastal benthic populations (summer) and oceanic pelagic populations (spring) (Kloppmann & Ulleweit, 2007). Thus, the potential for ecotypes or cryptic species to exist within *E. aequoreus* needs to be resolved with molecular markers.

In this study, we investigate phylogeographic patterns in the snake pipefish throughout its contemporary distribution. Our specific goals are to assess population structure and historical patterns of population expansion over the geographical range of the snake pipefish. Further, we investigate whether molecular data support proposed coastal benthic and pelagic ecotypes or cryptic species. To achieve these goals, we used mitochondrial DNA (mtDNA, cytochrome *b* and control region) markers to investigate genetic differentiation and genetic variation in snake pipefish from 25 locations among six geographical regions spanning over 1.9 million km² of their range in the Northeastern Atlantic Ocean.

**Materials and methods**

**Collections**
A total of 237 snake pipefish were collected from 25 locations in the Northeastern Atlantic Ocean between 2003 and 2010 using a variety of capture methods (Table 1, Fig. 1, Supplementary file S1). We assigned samples to regions based on conventional naming schemes of local water bodies or coastal areas. These regions correspond to regions defined by the OSPAR Commission (Region I: Arctic Waters = Norwegian Sea; Region II: Greater North Sea = North Sea; Region III Celtic Seas = Continental Shelf; Region IV Bay of Biscay and Iberian Coast = Spanish Coast) with the exception that Skagerrak/Kattegat were analysed separately from the North Sea due to the potential influence of the Baltic Sea, and the French Coast was analysed separately as it lies on the boundary of North Sea and Celtic Seas regions.

DNA was extracted from a small piece of tail tissue (~1cm) using a Qiagen DNeasy kit from live, frozen or EtOH preserved fish. A portion of the mitochondrial cytochrome b gene and control region was amplified. We used primers L14725 (Pääbo, Thomas, Whitfield et al., 1991) and H15926 (Wilson, Vincent, Ahnesjö et al., 2001) to amplify the cytochrome b locus, and primers L15926 (Kocher, K., Meyer et al., 1989) and H16498 (Meyer, Kocher, Basasibwaki et al., 1990) to amplify the control region. Fragments were amplified via polymerase chain reaction (PCR) in 30µl volumes containing 3µl of 10X buffer, 1.8 µl of dNTPs (10µM of each dNTP), 3 µl of MgCl₂ (50mM), 1.5 µl of each primer (10µM), and 0.5 µl of Taq (5 units/µl; InviTaq, Stratech Biomedical, Birkenfeld, DE). The PCR thermal profile consisted of an initial denaturation at 95°C (2 min), followed by 35 cycles of 94°C (30 sec), reannealing temperature (30 sec), 72°C (90 sec), and a final extension at 72°C for 10 min. Reannealing temperature for cytochrome b was 48°C and control region was 56°C. PCR products were purified before sequencing with Illustra™ ExoStar (GE Healthcare, Buckinghamshire, UK) using 5µl of PCR product and 2µl of ExoStar. PCR products were
sequenced with the forward and reverse primers using an ABI 3100 DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

Contigs were created using forward and reverse sequences using CODONCODE ALIGNER v. 5.1.5 (Codon Code Corporation, Centerville, MA, USA) for each individual and aligned for each locus using MUSCLE (Edgar, 2004). Sequences were verified by eye and trimmed. Unique haplotype sequences were identified using ‘DNA to haplotype collapse and converter’ in FaBox v. 1.41 (http://birc.au.dk/fabox) and deposited in GenBank (accession numbers #: cytochrome b, KY857646 - KY857823; control region, KY965149 - KY965308).

Genetic analyses

Relationships between mitochondrial haplotypes were analysed for each locus independently to assess whether population structuring exists among the six regions combined over all time periods. Standard haplotype (h) and nucleotide (π) diversity statistics (Nei, 1987) were calculated for each region using DNASP 5.10.01 (Librado & Rozas, 2009). Mismatch distributions were investigated in both loci independently, and evidence for recent population expansion was tested using Tajima’s $D$ test (Tajima, 1989), Fu’s $F_s$ test (Fu, 1997) and the Ramos-Onsins and Rozas’s $R_2$ statistic (Ramos-Onsins & Rozas, 2002), recommended for small sample sizes (Ramírez-Soriano, Ramos-Onsins, Rozas et al., 2008). Tajima’s $D$, $F_s$ and $R_2$ were calculated using the total number of mutations, excluding sites with alignment gaps or missing data, and significance was ascertained using coalescent simulations with 1000 replicates as implemented by DNASP. Population expansion under the constant size growth model was used to estimate $R_2$.

To visualize the relationship between mitochondrial haplotypes from different regions, a haplotype network was constructed using HAPLOVIEWER (Salzburger, Ewing &
von Haeseler, 2011) based on maximum likelihood trees drawn in DNAML in PHYLIP v3.695 (Felsenstein, 1989) for both loci independently. To test for significant population subdivision among individual collections, we conducted pairwise \( \Phi_{ST} \) tests for cytochrome \( b \) and control region sequences using ARLEQUIN v3.5.2.1 (Excoffier, Smouse & Quattro, 1992; Excoffier, Laval & Schneider, 2005). Significance for pairwise differences was ascertained using an exact test with 100,000 permutations. An Analysis of Molecular Variance (AMOVA) was used to test the proportion of genetic differentiation within and between regions using ARLEQUIN for both loci independently pooled across years. AMOVA was also used to test whether there is support for genetic differentiation between pelagic and coastal benthic populations. Only locations with a minimum of five sequenced individuals were included in pairwise \( \Phi_{ST} \) and AMOVA analyses, except for the three Norwegian Sea individuals that were included in the cytochrome \( b \) dataset since they represented a unique location. Significance of AMOVAs was determined using 99,999 permutations as implemented in ARLEQUIN.

We analysed the potential for fluctuations in effective population size using Bayesian Skyline Plot (BSP, Drummond, Rambaut, Shapiro et al., 2005), a coalescent-based method implemented in BEAST 1.8.4 (Drummond & Rambaut, 2007). We first estimated the best-fit model of nucleotide substitution for each gene using JModelTest 2.0 (Darriba, Taboada, Doallo et al., 2012) on all samples. We selected the best-fit models according to the Bayesian Information Criterion, which were HKY+I+G and HKY+I for cytochrome \( b \) and the control region, respectively. We then ran the BSP analysis on the concatenation of the two mitochondrial loci using samples that successfully amplified both cytochrome \( b \) and the control region (\( n=140 \)) using specific substitution models and mutation rates for each partition. We chose a mutation rate of \( 1 \times 10^{-8} \) substitutions per nucleotide site per year for cytochrome \( b \) and \( 5 \times 10^{-8} \) substitutions per nucleotide site per year for the control region.
based on recommendations by Bowen, Muss, Rocha et al. (2006) and Mobley et al. (2010) assuming a constant mutation rate. We assumed a generation time of one generation per year based on unimodal distributions of body size in juveniles caught in plankton tows (Kirby, Johns & Lindley, 2006). Adults collected within a year show a bimodal distribution in body size but this may be accounted for by sexual dimorphism in body length (van Damme & Couperus, 2008) also suggesting one reproductive cycle per year.

The BSP analysis consisted of $1 \times 10^8$ generations; parameters were sampled every 10,000 generations of which 10% was discarded as burn-in. In order to check the analysis performance (i.e., the convergence of parameters by visually checking the effective sample size (ESS>200) values), we used TRACER 1.6 (Rambaut, Suchard, Xie et al., 2014).

Ethical statement

The present study was conducted in accordance to local and European Union law and no permit was needed for collecting fish. The study did not involve any endangered or protected species.

Results

Molecular diversity

We resolved 900 bp of the cytochrome b gene from 178 snake pipefish from a large portion of their range. Cytochrome b yielded 94 unique haplotypes ($H_d = 0.967 \pm 0.006$ S.D.) and 82 polymorphic sites ($\pi = 0.0050 \pm 0.0002$ S.D., Table 2).

Sequence analyses of 385 bp of the control region locus in 160 snake pipefish revealed variation in a dinucleotide microsatellite repeat in the control region at site 287.
Most sequences contained [TA]_{10} but [TA]_{9} occurred at a low frequency (0.16) and was present in all geographical regions analysed except the Norwegian Sea samples that failed to amplify at this locus (Table 1). Inclusion of a mitochondrial microsatellite is problematic for several reasons (Lunt, Whipple & Hyman, 1998). For example, microsatellites generally have much higher mutation rates (in terms of the gain or loss of a repeat) than nucleotide substitutions, and mutation rates in general are unknown for mitochondrial microsatellites (Sia, Butler, Dominska et al., 2000) and for this microsatellite in particular. Moreover, we did not determine if heteroplasmy, or the coexistence of nonidentical mtDNA molecules in the same individual, is occurring in this species. Heteroplasmy is common in species that have microsatellites in mtDNA in the AT-rich or control region causing difficulties for population genetics analyses (Lunt et al., 1998; Mayer & Kerth, 2005). Additionally, length variation in the microsatellite can be considered a gap or missing data and thus may be inappropriate for some analyses (Yang & Rannala, 2012). Therefore, we excluded the variable repeat from all analyses, and we resolved 383 bps of the control region resulting in 63 unique haplotypes (H_{d} = 0.893 ± 0.015 S.D.) and 43 polymorphic sites (π = 0.0054 ± 0.0003 S.D., Table 2).

**Population structure**

A maximum likelihood haplotype network constructed for cytochrome b showed four major (n ≥ 10) haplotypes (Fig. 2a). However, these haplotypes were comprised of representative individuals from most regions such that no genetic clustering could be discerned. For the control region, a maximum likelihood haplotype network also showed four major haplotypes and individuals from all populations were represented in these major haplotypes (Fig. 3). Overall, haplotype networks showed a star-like topology with high numbers of low-frequency mutations representative of a recent population expansion.
Pairwise $\phi_{ST}$ values ranged from -0.119 to 0.212 with a mean of -0.009 for cytochrome $b$, and ranged from -0.290 to 0.242 with a mean of -0.030 for the control region (Appendix S1). No significant differences between pairs of collections were detected after Bonferroni adjustment to correct for multiple comparisons (Rice, 1989). An AMOVA among collections within regions pooled across years indicated no significant population structuring ($\phi_{ST}$) for either cytochrome $b$ or the control region (Table 3, Appendix S1). Finally, no evidence for genetically distinct ecotypes or cryptic species was found when investigating individuals captured in nearshore benthic habitats versus pelagic captures (Table 3). In all AMOVA comparisons, the variance explained by among groups and among collections within groups was negligible in comparison to the variance within collections (Table 3).

**Population expansion**

Significant values of Fu’s $F_s$, Tajima’s $D$, and Ramos-Onsins and Rozas’s $R_2$ found in cytochrome $b$ sequences, with the exception of the Spanish coast collection, support a scenario of recent population expansion when samples were pooled within locations and sampling times (Table 2). The control region, on the other hand, showed significant $F_s$ values, but only Continental shelf and French Coast collections showed significant departures from neutrality in Tajima’s $D$ and $R_2$ tests (Table 2). When collections were pooled across all location and sampling times, both the cytochrome $b$ and control region showed highly significant $F_s$, $D$ and $R_2$ values indicating recent population expansion was apparent (Table 2). Mismatch distributions were unimodal and failed to reject the hypothesis of the sudden expansion model (Appendix S1).

Bayesian Skyline Plot analysis with concatenated cytochrome $b$ and control region sequences revealed that snake pipefish experienced a historical effective population size expansion starting about 100ka and obtaining current effective population sizes around 50ka.
Individual effective population size analyses showed expansions approximately 125ka for cytochrome $b$ and 40ka for control region sequences (Appendix S1).

Discussion

Our study provides insight into the phylogeographic history of the snake pipefish, a widely distributed syngnathid species in the northeastern Atlantic with a pelagic lifestyle. Results from molecular analyses did not reveal any clear patterns in population structure by regions, despite relatively high haplotype diversity estimates. However, we did uncover a signature of a Pleistocene population expansion in the northeastern Atlantic approximately 50-100ka. We also did not find any evidence for genetically distinct coastal or pelagic ecotypes or cryptic species, indicating that differences in phenotype are likely due to differences in ecological conditions. Taken together, our results point to a large population of snake pipefish throughout its contemporary northeastern Atlantic distribution and that such a high degree of homogenization is likely the result of a combination of specialized/unique life history traits and mixing by oceanic currents.

Historical phylogeography

The history of the northeastern Atlantic is one of extreme climatic changes, with multiple glaciation cycles until the late Pleistocene (last interglacial period ~125 ka, last glacial maximum ~20ka, Mokeddem, McManus & Oppo, 2014). During this period, many marine species were deeply impacted by glacial activity, which caused large reductions and/or shifts in suitable habitat (Mäkinen & Merilä, 2008) occasionally leading to precipitous population declines (Almada, Pereira, Robalo et al., 2008; Boehme, Thompson, Fedak et al., 2012). The subsequent recolonization of the northeastern Atlantic oft times leads to complex genetic signatures of glacial refugia, range expansions and bottlenecks (e.g. Coyer, Peters, Stam et al., 2003; Gysels, Hellemans, Pampoulie et al., 2004; Luttikhuizen, Campos, van
The current study suggests that the snake pipefish underwent a population expansion in the northeastern Atlantic Ocean approximately 50-100ka during the Pleistocene. This scenario is supported by a star-like network topology, mismatch distribution analyses and an increase in effective population size indicating a recent expansion and/or a short evolutionary history of the species in the northeastern Atlantic Ocean (Grant & Bowen, 1998). Several other species show limited geographic partitioning and a sudden population expansion much earlier than the last glacial maximum similar to the snake pipefish. For example, the time of population expansion in the snake pipefish as estimated by BSP is similar to that for Atlantic Cod (*Gadus morhua*) which shows population expansion ~60ka in the northeastern Atlantic based on mismatch distributions of the mitogenome (Carr & Marshall, 2008). Other species that show signatures of population expansion predating the last glacial maximum include pelagic migrating species such as Atlantic bluefin tuna *Thunnus thynnus* (Bremer, Viñas, Mejuto et al., 2005) and the European anchovy, *Engraulis encrasicolus* (Silva, Horne & Castilho, 2014). In contrast, other sympatric nearshore species of syngnathids closely associated with seagrass habitats show more recent recolonization of the northeastern Atlantic (15-36ka) based on coalescence analysis (Wilson & Eigenmann Veraguth, 2010). Thus, it appears that pelagic species show a reduced influence of Pleistocene glacial cycles in the northeastern Atlantic.

Our estimates of expansion time are likely conservative, given that mutation rates are time dependent and are higher in younger lineages (Ho, Phillips, Cooper *et al.*, 2005; Ho, Lanfear, Bromham *et al.*, 2011; Crandall, Sbrocco, DeBoer *et al.*, 2012; Grant, 2015). It is therefore possible that time since expansion is more recent than estimated from BSP. It is also probable that the differences in mutation rates chosen for cytochrome *b* and the control region
contribute to the discrepancy between individual locus estimates of expansion time and may have led to a more gradual expansion time using the concatenated loci as compared to individual loci (Appendix S1). Fluctuations in drift, in population size, selection, and reproductive skew may also influence divergence times (Burridge, Craw, Fletcher et al., 2008; Grant, 2015; Niwa, Nashida & Yanagimoto, 2016). Finally, results of the BSP do not show recent (i.e., 10ka) fluctuations in effective population size although the probability of mtDNA markers capturing these phenomena is extremely unlikely.

**Population structure**

Population structuring over a large geographical range was not evident in this species from $\phi_{ST}$ and AMOVA results. These results mirror those reported for several Atlantic fish species where high mobility and/or high dispersal potential are highlighted as causes for little or no population sub-structuring (Nesbo, Rueness, Iversen et al., 2000; Dannewitz, Maes, Johansson et al., 2005; Carr & Marshall, 2008; Limborg, Hanel, Debes et al., 2012). Our results suggest that despite having relatively low swimming ability, the snake pipefish is capable of dispersing over long distances, likely aided by a pelagic lifestyle and faster development rates relative to other northeastern Atlantic syngnathid species (Braga Goncalves, Ahnesjö & Kvarnemo, 2016). A similar pattern has been described in the non-migratory, demersal marine fish, *Sebastes schegelii*, where the association of larvae and juveniles with rafting precludes population genetic differentiation throughout its geographical range (Zhang, Yanagimoto, Zhang et al., 2016). These results provide a stark contrast with several species of syngnathid that show population substructure over large spatial scales using mitochondrial markers (Lourie & Vincent, 2004; Lourie et al., 2005; Teske et al., 2005; Mobley et al., 2010; Wilson & Eigenmann Veraguth, 2010; Wilson et al., 2016).

Greater resolution of population sub-structuring in snake pipefish could potentially be provided by additional nuclear markers. However, attempts to design microsatellite markers
for this species have not yet yielded sufficient numbers of polymorphic microsatellites to conduct such analyses (KB Mobley, I Braga Goncalves, unpublished data). Future studies that benefit from combined mtDNA and nuclear markers may reveal additional insights into the phylogeography of this species including resolution of the time since population expansion and estimate gene flow within the species.

**Range expansion**

Although a significant number of snake pipefish were caught in polar regions during the population increase in 2003-2007, they may have originated in the southern Norwegian Sea and drifted to these locations in ocean currents (Nesbo *et al.*, 2000; Luttikhuizen *et al.*, 2008). In our study, we obtained only a few samples in polar waters from the Norwegian Sea in 2008. Haplotypes obtained from these samples were not distinct from southern populations suggesting that they are derived from the same large Atlantic gene pool. Due to the low abundance of snake pipefish after the population increase in north Atlantic waters (Heath *et al.*, 2012), future surveys should investigate whether or not these fish continue to inhabit polar regions representing a true range expansion or whether these were just transient individuals that rafted northward on ocean currents during the 2003-2007 outbreak.

**Cryptic species**

Previous studies have documented phenotypic differences between snake pipefish collected in coastal and oceanic areas (Holt & Byrne, 1906; Zhang *et al.*, 2016). Based on these differences, two species have been proposed previously, *E. aequoreus* found in oceanic waters and *E. anguineus* found in inshore habitats (Yarrel, 1839; Moreau, 1881), although *E. anguineus* is not currently recognized as a valid species (Dawson, 1985). These oceanic and coastal ‘ecotypes’ were recently described and a third potential intermediate ecotype that shares phenotypic similarities with both coastal and oceanic forms has been proposed (van Damme & Couperus, 2008). Despite differences in coloration and body condition, van
Damme and Couperus (2008) found no differences in ring and fin ray counts between perceived ecotypes in either sex and therefore concluded that all specimens belong to a single species. Based on mitochondrial haplotype frequencies, we also find no support that oceanic pelagic and benthic coastal ecotypes of snake pipefish collected in these habitats are different species. Instead, the snake pipefish appears to form a single species in the northeastern Atlantic and phenotypic plasticity in response to local ecological conditions encountered is the most probable explanation for the presence of multiple ecotypes.

**Conservation concerns**

According to the International Union for Conservation and Nature red list, *E. aequoreus* is evaluated as least concern (IUCN, 2017). Despite this listing, the sudden and substantial population increase in snake pipefish, although short lived, had a dramatic effect on the ecology of the northeastern Atlantic during the outbreak, and therefore warrants special mention. The significant increase in the numbers of snake pipefish in European waters was paralleled by a similar increase in the number of pipefish fed by parents to seabird nestlings in several species around the UK, Norway, Iceland, and the Faroe Islands (Luttikhuizen et al., 2008; Anderson, Evans, Potts et al., 2014). The hard exoskeleton and relatively low nutritional content of the snake pipefish (Harris, Newell, Daunt et al., 2008) make them unsuitable for consumption by nestlings and adults alike, and their increased use as food items was associated with seabird breeding failures in the UK (Mavor, Parsons, Heubeck et al., 2005; Mavor, Parsons, Heubeck et al., 2006; Luttikhuizen et al., 2008).

Despite the ecological significance of this species, the cause of the recent population increase of the snake pipefish still remains unknown, as does its subsequent decline (Heath et al., 2012). Several non-mutually exclusive hypotheses have been put forward to explain its
population increase and expansion, namely: 1) a rise in surface seawater temperatures promoting longer breeding seasons and higher recruitment (Kirby et al., 2006; Gremillet & Boulinier, 2009; Neumann, Ehrich & Kroncke, 2009; Anderson et al., 2014), 2) a by-product of tracking changing and/or shifting plankton communities (van Damme & Couperus, 2008), 3) promoted by the establishment of invasive algal species such as the Japanese seaweed, Sargassum muticum, that increased the amount of suitable habitat for successful reproduction in the coastal regions (Gysels et al., 2004) and, 4) a result of decreasing population size of interspecific competitors such as Lesser Sandeels (Ammodytes marinus) due to fishing and climate change (Heath et al., 2012; Anderson et al., 2014). Yet, there is no conclusive evidence to explain the sudden expansion.

The increase in snake pipefish potentially started in one source spot on the continental shelf and dispersed over the entire northeastern Atlantic. Interestingly, the increase and expansion in the pelagic environment itself can be seen as a strategy of the species to colonize new suitable habitats. However, expansions of the same magnitude appear to be very rare: there are indications that a similar population increase took place at the end of the nineteenth century (Brongersma-Sanders, 1957; van Damme & Couperus, 2008). Climate change is expected to change ocean currents which, together with higher sea surface temperatures, will affect larval import, export and recruitment, leading to faster development, shorter larval stages and dispersal into new habitats (Kendall, Poti & Karnauskas, 2016). Improving our understanding of snake pipefish reproductive biology, habitat, feeding ecology, and dispersal potential is a critical next step to help pinpoint how current and future changes in climate and in prey distributions in the northeastern Atlantic may potentiate further population increases, which in turn may affect community structure.

Conclusions
Understanding species’ current distribution patterns and historical demography is a fundamental goal in evolutionary biology. Our study contributes to this goal by investigating phylogeographic patterns in a species that has undergone a sudden population increase and range expansion that had a negative impact on the ecosystem of the northeastern Atlantic. Although the cause for the recent population increase and range expansion and contraction are still unknown, the phylogeographic patterns uncovered in our study demonstrate that the snake pipefish represents a single large population with no evidence of population substructuring. This result, in contrast to all other syngnathids studied to date, may be explained by the pelagic lifestyle and poor swimming capabilities of the species, allowing individuals to be transported long distances by ocean currents. Our study adds to the understanding of this ecologically important species and future studies should incorporate a wider range of genetic markers to investigate population demographics, particularly concerning the recent population increase.

Supplementary Material

Additional Supporting Information may be found in the online version of this article:

Appendix S1: Population pairwise $\Phi_{ST}$ estimates, mismatch distributions, and individual Bayesian skyline plots for cytochrome $b$ and control region loci.

Acknowledgements

We would like to thank Anders Berglund, Zach Cress, Loren Mendoza, Laurent Leveque, Ian Probert and the staff at the Roscoff Marine Station Centre, France and for assistance with field collections. We would also like to thank Adam Jones and Clay Small at Texas A&M University, Tammes Menne at the Zoological Museum, University of Copenhagen and Tony Wilson and Alexandra Wegmann at the University of Zürich for access to preserved
specimens. Sophie Bodenstein, Thoomka Brüning, Withe Derner, Nadja Gogrefe, Sybille Liedtke, and Ines Schlultz assisted with DNA sequencing. We thank Folmer Bokma at Umeå University and the Max Planck Society for technical support. We thank two anonymous reviewers for helpful comments on the manuscript. This work was supported by a transnational access grant to KM and IBG from the Association of European Marine Biological Laboratories (ASSEMBLE), for fieldwork at Roscoff Marine Station, France and Lundgrenska Stiftelserna and Helge Axelsson Johnsons stiftelse for financial support to IBG.

Data Accessibility

All cytochrome b and control region sequences will be available in Genbank and supporting sample information and haplotype lists will be deposited in Dryad.

References


Carr SM, Marshall HD. 2008. Intraspecific phylogeographic genomics from multiple complete mtDNA Genomes in Atlantic cod (*Gadus morhua*): Origins of the


Mobley KB, Small CM, Jones AG. 2011. The genetics and genomics of Syngnathidae (pipefishes, seahorses and seadragons). *Journal of Fish Biology* **78**: 1624-1646.

Mobley KB, Small CM, Jue NK, Jones AG. 2010. Population structure of the dusky pipefish (*Syngnathus florideus*) from the Atlantic and Gulf of Mexico, as revealed by mitochondrial DNA and microsatellite analyses. *Journal of Biogeography* **37**: 1363-1377.


Polte P, Buschbaum C. **2008.** Native pipefish *Entelurus aequoreus* are promoted by the introduced seaweed *Sargassum muticum* in the northern Wadden Sea, North Sea. *Aquatic Biology* **3:** 11-18.


Table 1. *Entelurus aequoreus* collections. Region of collection, number of individuals collected (n), year collected, location, collection abbreviation (AB), habitat, latitude, longitude, the number of individuals amplifiable at the cytochrome *b* locus (*n*<sub>cytb</sub>) and control region locus (*n*<sub>CR</sub>), and presence of the microsatellite repeat [TA]<sub>9</sub> in the control region are listed.

<table>
<thead>
<tr>
<th>Region</th>
<th>n</th>
<th>Year</th>
<th>Location</th>
<th>AB</th>
<th>Habitat</th>
<th>Latitude</th>
<th>Longitude</th>
<th><em>n</em>&lt;sub&gt;cytb&lt;/sub&gt;</th>
<th><em>n</em>&lt;sub&gt;CR&lt;/sub&gt;</th>
<th>[TA]&lt;sub&gt;9&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continental Shelf</td>
<td>5</td>
<td>2005</td>
<td>Atlantic Ocean</td>
<td>CS1</td>
<td>pelagic</td>
<td>51°45'0.61&quot;N</td>
<td>11°45'0.61&quot;W</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2005</td>
<td>Atlantic Ocean</td>
<td>CS2</td>
<td>pelagic</td>
<td>56°10'2.39&quot;N</td>
<td>9°47'4.81&quot;W</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2005</td>
<td>Atlantic Ocean</td>
<td>CS3</td>
<td>pelagic</td>
<td>50°45'0.00&quot;N</td>
<td>11° 7'0.01&quot;W</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>2005</td>
<td>Atlantic Ocean</td>
<td>CS4</td>
<td>pelagic</td>
<td>51°15'1.80&quot;N</td>
<td>13°42'1.80&quot;W</td>
<td>14</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2007</td>
<td>Atlantic Ocean</td>
<td>CS5</td>
<td>pelagic</td>
<td>49°8'15.99&quot;N</td>
<td>10°22'48&quot;W</td>
<td>10</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>2010</td>
<td>Atlantic Ocean</td>
<td>CS6</td>
<td>pelagic</td>
<td>51°52'42.60&quot;N</td>
<td>13° 8'3.01&quot;W</td>
<td>22</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>Spanish Coast</td>
<td>10</td>
<td>2007</td>
<td>Galacia, ES</td>
<td>SC</td>
<td>coastal</td>
<td>42°15'N</td>
<td>8°52'W</td>
<td>10</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>North Sea</td>
<td>3</td>
<td>2005</td>
<td>North Sea</td>
<td>NS01</td>
<td>pelagic</td>
<td>57°48'34.99&quot;N</td>
<td>0°52'50.70&quot;W</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2005</td>
<td>North Sea</td>
<td>NS02</td>
<td>pelagic</td>
<td>57°52'7.14&quot;N</td>
<td>3°14'54.06&quot;W</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2005</td>
<td>North Sea</td>
<td>NS03</td>
<td>pelagic</td>
<td>57°44'55.03&quot;N</td>
<td>1°21'28.26&quot;W</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2005</td>
<td>North Sea</td>
<td>NS04</td>
<td>pelagic</td>
<td>58°10'10.99&quot;N</td>
<td>0°33'42.23&quot;W</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2005</td>
<td>North Sea</td>
<td>NS05</td>
<td>pelagic</td>
<td>58°10'37.27&quot;N</td>
<td>3°10'12.00&quot;W</td>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2007</td>
<td>North Sea</td>
<td>NS06</td>
<td>pelagic</td>
<td>56°21'38.99&quot;N</td>
<td>2° 4'58.19&quot;W</td>
<td>8</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2007</td>
<td>North Sea</td>
<td>NS07</td>
<td>pelagic</td>
<td>56° 7'34.21&quot;N</td>
<td>3°28'3.00&quot;E</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2007</td>
<td>North Sea</td>
<td>NS08</td>
<td>pelagic</td>
<td>53°28'53.40&quot;N</td>
<td>0°54'54.00&quot;E</td>
<td>7</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2007</td>
<td>North Sea</td>
<td>NS09</td>
<td>pelagic</td>
<td>55°36'16.20&quot;N</td>
<td>2°46'31.19&quot;E</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2007</td>
<td>North Sea</td>
<td>NS10</td>
<td>pelagic</td>
<td>55°53'26.41&quot;N</td>
<td>4°15'45.61&quot;E</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2007</td>
<td>North Sea</td>
<td>NS11</td>
<td>pelagic</td>
<td>56°45'29.41&quot;N</td>
<td>1°33'30.60&quot;W</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2007</td>
<td>North Sea</td>
<td>NS12</td>
<td>pelagic</td>
<td>55°23'28.79&quot;N</td>
<td>1°34'34.79&quot;E</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Skagerrak/Kattegat</td>
<td>2</td>
<td>2003</td>
<td>Kølpen/Deget, DK</td>
<td>SK1</td>
<td>coastal</td>
<td>57°27'17.66&quot;N</td>
<td>10°35'53.61&quot;E</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2004</td>
<td>Dronningmølle, DK</td>
<td>SK2</td>
<td>coastal</td>
<td>56° 6'6.07&quot;N</td>
<td>12°24'34.77&quot;E</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2005</td>
<td>Gåsö, SE</td>
<td>SK3</td>
<td>coastal</td>
<td>58°14'23.32&quot;N</td>
<td>11°22'44.86&quot;E</td>
<td>7</td>
<td>0</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Year</td>
<td>Area</td>
<td>Code</td>
<td>Depth</td>
<td>Lat/Long</td>
<td>Continent</td>
<td>IM</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>------------------</td>
<td>------</td>
<td>--------</td>
<td>------------------</td>
<td>-----------</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian Sea</td>
<td>2006</td>
<td>Gåsö, SE</td>
<td>SK4</td>
<td>coastal</td>
<td>58°14'23.32&quot;N</td>
<td>11°22'44.86&quot;E</td>
<td>20</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008</td>
<td>Norwegian Sea</td>
<td>NOR1</td>
<td>pelagic</td>
<td>68°15'40.02&quot;N</td>
<td>4° 7'45.00&quot;E</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian Sea</td>
<td>2008</td>
<td>Norwegian Sea</td>
<td>NOR2</td>
<td>pelagic</td>
<td>68°15'24.66&quot;N</td>
<td>0°31'45.36&quot;W</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French Coast</td>
<td>2010</td>
<td>Bay de Roscoff, FR</td>
<td>FC</td>
<td>coastal</td>
<td>48°42'13.75&quot;N</td>
<td>3°55'11.67&quot;W</td>
<td>21</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 A subset of 20 individuals was extracted. 2 A subset of 35 individuals was extracted.
Table 2. Genetic diversity indices and tests of neutrality pooled across years in regions for Cytochrome b (900 bp) and control region (383 bp with microsatellite removed) loci. Number of individuals sequenced (n), number of unique haplotypes (H), haplotype diversity (Hd, ± S.D.), nucleotide diversity (π, ± S.D.), Tajima’s D (D), Fu’s Fs (Fs), and Ramos-Onsins and Rozas’s R2 (R2) test are given. Asterisks denote significant departures from neutrality: P<0.05 = *, P<0.01 = **, P<0.001 ***.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cytochrome b</th>
<th></th>
<th></th>
<th>Control region</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>H</td>
<td>Hd</td>
<td>π</td>
<td>D</td>
<td>Fs</td>
</tr>
<tr>
<td>Continental Shelf</td>
<td>59</td>
<td>41</td>
<td>0.972 (0.011)</td>
<td>0.0049 (0.0004)</td>
<td>-1.97**</td>
<td>-41.267***</td>
</tr>
<tr>
<td>Spanish Coast</td>
<td>10</td>
<td>7</td>
<td>0.911 (0.077)</td>
<td>0.0059 (0.0008)</td>
<td>-0.53</td>
<td>-0.89</td>
</tr>
<tr>
<td>North Sea</td>
<td>55</td>
<td>36</td>
<td>0.962 (0.015)</td>
<td>0.0048 (0.0004)</td>
<td>-1.82**</td>
<td>-31.99***</td>
</tr>
<tr>
<td>Skagerrak/Kattegat</td>
<td>30</td>
<td>22</td>
<td>0.968 (0.019)</td>
<td>0.0049 (0.0005)</td>
<td>-1.50*</td>
<td>-14.72***</td>
</tr>
<tr>
<td>Norwegian Sea</td>
<td>3</td>
<td>3</td>
<td>1.000 (0.272)</td>
<td>0.0030 (0.0009)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>French Coast</td>
<td>21</td>
<td>19</td>
<td>0.986 (0.022)</td>
<td>0.0055 (0.0006)</td>
<td>-1.48*</td>
<td>-14.32***</td>
</tr>
<tr>
<td>Pooled</td>
<td>178</td>
<td>94</td>
<td>0.967 (0.006)</td>
<td>0.0050 (0.0002)</td>
<td>-2.22***</td>
<td>-133.75***</td>
</tr>
</tbody>
</table>
Table 3. Analysis of molecular variance (AMOVA) of *Entelurus aequoreus* based on 900 bp of mtDNA cytochrome b (Cytb) sequence and 383 bp of mtDNA control region (CR) sequence (microsatellite at position 287 excluded). The percent variation (% var), and P value are listed for among groups (\( \phi_{CT} \)), among collections within groups (\( \phi_{SC} \)) and within collections \( \phi_{ST} \). AMOVAs were conducted with all collections within regions (regions) and all collections pooled within habitat (coastal vs. pelagic).

<table>
<thead>
<tr>
<th></th>
<th>Among groups % var</th>
<th>Among groups ( \phi_{CT} )</th>
<th>Among groups P value</th>
<th>Among collections within groups % var</th>
<th>Among collections within groups ( \phi_{SC} )</th>
<th>Among collections within groups P value</th>
<th>Within collections ( \phi_{ST} )</th>
<th>Within collections P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All populations</td>
<td>0.11</td>
<td>0.0011</td>
<td>0.4534</td>
<td>-0.62</td>
<td>-0.0062</td>
<td>0.5535</td>
<td>-0.0051</td>
<td>0.6189</td>
</tr>
<tr>
<td></td>
<td>3.76</td>
<td>0.0376</td>
<td>0.0603</td>
<td>-3.11</td>
<td>-0.0324</td>
<td>0.9151</td>
<td>0.0065</td>
<td>0.4495</td>
</tr>
<tr>
<td>Habitat</td>
<td>0.52</td>
<td>-0.0052</td>
<td>0.7534</td>
<td>-0.28</td>
<td>-0.0028</td>
<td>0.4951</td>
<td>-0.0080</td>
<td>0.6165</td>
</tr>
<tr>
<td></td>
<td>2.82</td>
<td>0.0282</td>
<td>0.0937</td>
<td>-1.48</td>
<td>-0.0152</td>
<td>0.8135</td>
<td>0.0135</td>
<td>0.4472</td>
</tr>
</tbody>
</table>
Figure Legends

Fig. 1. Sampling locations for *Entelurus aequoreus* in the northeastern Atlantic Ocean. Predominant ocean currents are shown in grey (after OSPAR, 2010). See Table 1 for sample information and abbreviations.

Fig. 2. Maximum likelihood networks for *Entelurus aequoreus* pooled over years for a) mtDNA *cytochrome b* sequences (900bp) and b) mtDNA control region sequences with microsatellite at position 287 removed (383bp). Each circle represents a haplotype and its size is proportional to its total frequency. Colors correspond to regions (see Table 1 and text for descriptions). Black crossbars represent a single nucleotide mutation and filled in circles represent reconstructed haplotypes not sampled.

Fig. 3. Bayesian Skyline Plot (BSP) showing changes in effective population size through time (thousand years before present, ka). Dashed line represents the median posterior estimate of the effective population size. The grey area delimited by continuous black lines shows the 95% highest posterior density limits.