Received 12 January 2018

Accepted 2 August 2018

REGULAR PAPER

High level of genetic connectivity in a deep-water reef fish, the *C. microps Caulolatilus microps*

J. R. MCDOWELL AND H. L. BRIGHTMAN

Department of Fisheries Science, Virginia Institute of Marine Science, College of William & Mary, 1375 Greate Road, Gloucester Point, Virginia 23062, U.S.A.

Correspondence

J. R. Mcdowell, Department of Fisheries Science, Virginia Institute of Marine Science, College of William & Mary, 1375 Greate Road, Gloucester Point, Virginia 23062, U.S.A.

Email: McDowell@vims.edu

Funding information

This study was supported by Award 20150601 from the Mid-Atlantic Fisheries Management

Council and NOAA. This paper is contribution No. 3754 of the Virginia Institute of Marine Science, College of William & Mary.

The life-history characteristics of *C. microps Caulolatilus microps* make them particularly vulnerable to overfishing. Although North Carolina, U.S.A. was previously considered to be the northern extent of the range of *C. microps*, concentrations have recently been discovered in Virginia and Maryland, with reports as far north as Montauk, New York. Attempts to manage the fishery have been hampered by a lack of information about whether the U.S. East Coast includes multiple stocks. To assess the appropriateness of alternate management options, we used 25 variable microsatellite loci and sequencing of the mitochondrial (mt)DNA control region (CR) to evaluate the genetic structure of 490 *C. microps* sampled from across the U.S. East Coast range. Pairwise comparisons of genetic differentiation among collection locations based on both nuclear microsatellite and mtCR sequence data were all low and non significant. No significant autocorrelation was observed across multiple distance classes, consistent with widespread dispers*al.*, Although the assumed sedentary nature of adult *C. microps* suggests population structuring, the genetic data were consistent with the presence of sufficient gene flow to prevent the accumulation of significant genetic differences and supports management of *C. microps* as a single stock along the U.S. East Coast.

KEYWORDS

blueline tilefish, *Caulolatilus microps*, control region, gene flow, microsatellites, population structure

1 | INTRODUCTION

Understanding whether a species is comprises demographically independent groups has important implications for effective management. Stocks are principally structured by rates of birth and death rather than immigration and emigration, thus managing several distinct stocks as a single unit can lead to unintended overexploitation, localized depletion and lower recruitment (Ying *et al.,* 2011, Spies & Punt 2015). Alternately, managing distinct stocks as a unit can lead to overly restrictive regulations on the fishery (Tuckey *et al.*, 2007; Ying *et al.*, 2011; Spies & Punt, 2015). Genetic information is also critical to identification of management measures appropriate for the on-going conservation of species, as overharvest of unrecognized genetically distinct stocks can lead to loss of genetic diversity (Laikre *et al.*, 2005; Allendorf *et al.*, 2014; Pinsky & Palumbi, 2014). The establishment of a baseline estimate of genetic variation also allows monitoring of changes in variation (loss of diversity or changes in how diversity is distributed) in the future that may result from overfishing (Schwartz *et al.*, 2007).

Blueline tilefish *Caulolatilus microps* Goode & Bean 1878, family Malacanthidae, also known as grey tilefish (Goode & Bean, 1878), is a bottom dweller found at depths of 75–250m. *Caulolatilus microps* have historically been reported to occur along the continental shelf from

Virginia, U.S.A. (38° N) to the Campeche Banks of Mexico (19° N), occupying the same habitat as groupers and snappers (Dooley, 1978). Although North Carolina was previously considered to be the northern extent of the range of *C. microps*, concentrations have recently been discovered in Virginia and Maryland, with reports as far north as Montauk, New York (41° N). A SEDAR 50 data workshop (SouthEast Data, Assessment and Review,), which compiled several fishery-dependent and fishery-independent data sources, determined that *C. microps* are continuously distributed from the Gulf of Mexico to the Mid-Atlantic Bight, which runs from North Carolina to Massachusetts (Farmer & Klibansky, 2016).

Many of the life-history characteristics of *C. microps* make them particularly vulnerable to overfishing. Like other species of tilefish, *C. microps* are long-lived and grow slowly, with an estimated lifespan of up to 43 years and a maximum fork length (L_F) of 900 mm (SEDAR 32, 2013; Kolmos *et al.*, 2016). However, *C. microps* reach maturity as early as age 2 years (Harris *et al.*, 2004; Kolmos *et al.*, 2016) and are sexually dimorphic, with males reaching a larger maximum size than females (Harris *et al.*, 2004). A study by Harris *et al. (*2004) off the coast of North and South Carolina found that *C. microps* are batch-spawners. Spawning occurs in the evening from February–November with a peak in May and there is a positive relationship between L_F and fecundity in females (Harris *et al.*, 2004; Kolmos *et al.*, 2016). *Caulolatilus microps* eggs are pelagic and have been genetically identified from fish eggs collected along the north-eastern U.S. continental shelf between Cape Hatteras, North Carolina (NC) and Nova Scotia, Canada (45° N), by the Northeast Fisheries Science Center (NEFSC) Ecosystem

Monitoring (EcoMon) programme (Lewis *et al.*, 2016). Data concerning larval duration or dispersal is limited.

The South Atlantic Fishery Management Council (SAFMC; www.safmc.net) manages *C. microps* as part of the deep-water snapper–grouper complex (the snapper-grouper fishery management unit). The south Atlantic management area extends from the North Carolina–Virginia border to Key West, Florida (24^{o} 302N) within the U.S. exclusive economic zone (EEZ). The Gulf of Mexico Fishery Management Council (GMFMC; www.gulfcouncil.org) manages *C. microps* in the U.S. portion of the Gulf of Mexico. Historically, *C. microps* were not a managed species in the mid-Atlantic region. In response to the development of an unregulated fishery off the coast of Virginia, in 2007 the Virginia Marine Resources Commission (VMRC; www.mrc.virginia.gov) enacted regulations on *C. microps* including a recreational landing and possession limit of 7 fish day^{-1} and a commercial possession limit of 300 pounds (*c.* 135 kg) whole weight or 273 pounds (*c.* 125 kg) gutted weight. These same regulations were later adopted for Maryland state waters. The fishery for *C. microps* was not regulated north of Maryland and no federal regulations north of North Carolina were in place until the enactment of emergency management measures in 2015. This lack of regulation became problematic due to a substantial increase in commercial and recreational landings in the U.S. mid-Atlantic in recent years. Commercial landings in this region averaged 4.8 t between 2005 and 2013 but following implementation of a reduction in the commercial catch limit in the South Atlantic region by the SAFMC in 2014, the commercial landings of *C. microps* in the mid-Atlantic skyrocketed to 97.8

t. The unregulated northern fishery combined with the lack of information about the number of stocks and the vulnerable (*k*-selected) life history of *C. microps* raised concern about the sustainability of the resource and led the Mid-Atlantic Fishery Management Council (MAFMC; www.mafms.org) to adopt the same regulations as Virginia in U.S. waters north of the latitude of the Virginia–North Carolina border as an emergency measure until long-term management measures can be implemented.

Caulolatilus microps were formally assessed in the Southeast region in 2013 and were found to be overfished with overfishing occurring (SEDAR 32, 2013). It was suggested that overfishing had been on-going over most of the assessment period and it was noted that there has been considerable uncertainty in biomass estimates since the mid-2000s (SEDAR 32, 2013). The assessment considered *C. microps* to be a single coastwide stock, although data from north of Cape Hatteras, NC was limited. Attempts to manage the fishery have been hindered by the lack of information about whether the U.S. East Coast includes multiple stocks. To address this question, 25 variable microsatellite loci were developed and used in conjunction with sequencing of the mitochondrial (mt)DNA control region (CR) to evaluate the genetic structure of *C. microps* sampled across the U.S. East Coast range. Data from 503 individuals was used to assess the appropriateness of alternate management options; specifically, whether genetic evidence is consistent with the presence of a single coastwide stock, multiple discrete stocks, or if data are consistent with a pattern of isolation by distance (IBD). This is the first genetic study of *C.*

microps and information from this study will be directly applicable to the conservation and management of the species.

2 | MATERIALS AND METHODS

2.1 | Microsatellite marker development and optimization

High molecular weight DNA from a *C. microps* captured off the coast of Virginia (VA) was used to create and template a 400 bp insert genomic library using the Ion Xpress Fragment Library Kit and the Ion PGM Hi-Q View OT2 Kit (ThermoFisher Scientific; www.thermofisher.com). The resulting fragments were sequenced using a PGM Hi-Q Sequencing Kit on an Ion Torrent PGM sequencer using an Ion 318 chip (ThermoFisher Scientific). The FastQC software (Andrews, 2010) integrated into the Galaxy Project platform (Giardine *et al.,* 2005, Goecks *et al.,* 2010, Blankenberg *et al.,* 2010) was used to assess the quality of the resulting sequences and filters integrated into the Galaxy platform were used to remove sequencing artefacts and to filter out sequences below 50 bp. Sequences were trimmed to exclude positions 1–9 and all bases over 400 bp and filtered by quality to exclude those in which 50% of the sequence length had a quality < 20 (base call accuracy < 99%) and the remaining data were exported as a FASTA-formatted file. Exported sequence files were searched for the presence of perfect tetranucleotide repeats using the MSATCOMMANDER 1.0.8 software (Faircloth, 2008). For identified loci, primers were

designed using the Primer3 software (Koressaar & Remm, 2007; Untergasser *et al.*, 2012).

All newly identified primer-pairs were initially assessed and optimized using gradient PCR on a Bio-Rad C1000 thermal cycler (Bio-Rad; www.bio-rad.com) using standard protocols. Each 5 µl PCR reaction contained 1x PCR buffer (Qiagen; www.qiagen.com), 1.5 mM $MgCl_2$, 200 µM of each deoxynucleotide triphosphate (dNTP), 0.125 µM the forward primer, 0.125 of the reverse primer, 0.5 unit of Taq polymerase (Qiagen) and 0.5 µl genomic DNA. Four samples, two collected off VA and two collected off South Carolina (SC) were used for testing primers for reliable amplification of a product of the predicted length. Samples were amplified with an initial denaturation temperature of 94º C for 3 min, followed by 35 cycles at 94º C for 1 min, 48–65º C for 1 min, 72º C for 1 min, with a final elongation step at 72º C for 7 min. Amplified products were visualized to confirm the presence of a single amplification product of correct size by agarose gel electrophoresis (1.5% w/v), stained with Gel Red (Biotium; www.biotium.com) and viewed under a UV light source. Markers found to reliably amplify DNA samples from both VA and SC were further evaluated using a panel of 8 samples each from VA and SC to assess amplification consistency, levels of polymorphism and conformance to the expectations of Hardy-Weinberg equilibrium (HWE). PCR reactions were carried out as above except for the addition of a T3-labeled fluorescent probe (either FAM, VIC, NED, PET). The resulting fluorescently labelled PCR products were separated on an ABI 3130*xl* Prism Genetic Analyzer (Thermofisher–Applied Biosystems) with a GeneScan 500-Liz size standard (Applied

Biosystems). The chromatic peaks for each microsatellite locus were sized using GENEMARKER AFLP/Genotyping 1.75 (SoftGenetics; www.softgenetics.com).

2.2 | Sample collection and dna isolation

Geographic sampling spanned the distributional range of *C. microps* along the U.S. East Coast from New York (NY) to the southern Florida Keys (FL). Sampling occurred between April 2015 and April 2016, with the exception of 20 samples from SC that were collected in 2012–2014. Sampling locations were: NY, Delaware (DE), VA, North Carolina North of Cape Hatteras (NCN), North Carolina South of Cape Hatteras (NCS), SC and FL. A limited number of samples (n = 15) collected off the west coast of Florida (FLW) in the Gulf of Mexico in 2011 were also included for comparison (Figure 1). Cooperating commercial fishermen collected samples for use in this study as part of the 2015 *C. microps* cooperative-with-industry data collection project funded by the National Marine Fisheries Service (NMFS; www.fisheries.noaa.gove). Samples were obtained using three gear types: long bottom longline, short bottom longline and vertical hook and line. Additional samples from Virginia were collected during a recreational charter. All fin clips were stored in ethanol until DNA could be extracted and pertinent collection information (date, length, depth, location, vessel *etc*.) was recorded (Kellison, 2016). DNA was extracted from archived tissue samples using either the DNeasy Tissue Kit (Qiagen) or the Quick-DNA Universal kit (Zymo Research; www.zymoresearch.com) with minor modifications.

Briefly, 2–3 mm fin clip sub-samples were incubated in lysis buffer (Longmire *et al.*, 1997) for 2 h at room temperature to facilitate removal of residual ethanol prior extraction following the manufacturers protocol. All DNA samples were quantified using a NanoDrop 2000 (Thermo Scientific) and stored at –20 °C until analysis.

2.3 | Microsatellite marker analysis

Following optimization, primer pairs were multiplexed into panels using the Type-it microsatellite PCR kit (Qiagen) and alleles were sized as for marker development and optimization (Table 1). Each multiplex reaction contained 1x Type-it Multiplex PCR Master Mix, 1x Q-solution, 0.05 µM of the forward primer, 0.2 µM of the reverse primer, 0.2 µM of the fluorescent dye, 0.5 µl genomic DNA and water to a final volume of 6 µl. Amplifications were performed with an initial denaturation temperature of 95º C for 5 min, followed by 28 cycles at 95º C for 30 s, annealing for 90 s at the temperature indicated in Table 1, extension at 72º C for 30 s and a final extension step at 60º C for 30 min. The resulting fluorescently labelled PCR products were separated on an ABI 3130*xl* Prism Genetic Analyzer (Applied Biosystems) with a GeneScan 500-Liz size standard (Applied Biosystems). The chromatic peaks for each microsatellite locus were sized using the GENEMARKER AFLP/Genotyping software. To control for errors, approximately 20% of samples were amplified and sized twice for all loci as above and results were compared.

After alleles had been sized for each locus, the MICRO-CHECKER 2.2.3 software (Van Oosterhout *et al.*, 2004) was used to check for the presence of null alleles and evidence of scoring errors using 10 000 permutations of the data and a 95% confidence interv*al.*, The GENEPOP 007 software package (Rousset, 2008) was used to test for deviations of genotype distributions from HWE expectations (F_{IS}, exact tests; Guo *et al.,* 1992). Summary statistics including: number of alleles (N_A), effective number of alleles (A_E), observed (H_O) and expected heterozygosity (H_E, gene diversity), unbiased expected heterozygosity (H_{Eu}) and allele frequencies were calculating using GENALEX 6.5 (Peakall & Smouse, 2012). Allelic richness per locus and across all samples was calculated in FSTAT 2.9.3.2 using the rarefaction approach of Hurlbert, 1971 and Petit *et al.*, 1998, which trims unequal sample sizes to a standardized sample size less than or equal to the smallest sample size across populations. Since small sample sizes can cause errors in estimates of allelic richness and the FLW samples was much smaller than samples from other geographic locations, allelic richness was also estimated using extrapolation to a sample size of 50 following the methods of Colwell *et al.*, 2004 using the ARES software (Van Loon *et al.*, 2007) in R 3.1.2 (www.r-project.org) by L. Roser (www.github.com/leandroroser/Ares_1.2-4).

To evaluate whether the data were indicative of the presence of population structure, Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) was used to estimate Weir & Cockerham's (1984) unbiased estimator of Wrights *F*-statistics and to conduct an analysis of molecular variance (AMOVA) based on several alternate geographic groupings of collection locations (Excoffier *et*

al., 1992). Significance was assessed *via* 10 000 permutations of the data. Multivariate analysis was used to explore variability among samples without consideration of geographic collection location. Principal component analysis (PCA) was performed using the adegenet 1.4-2 package in R (Jombart, 2008; Jombart *et al.,* 2009). The scalegen function of adegenet was used to replace missing data with mean allele frequencies for the PCA. The Bayesian clustering algorithm implemented in Structure 2.3.4 (Pritchard *et al.,* 2000, Falush *et al.,* 2007, Hubisz *et al.,* 2009) was used to test whether samples comprised distinct genetic groups. Simulations were performed using an admixture model of ancestry, location prior and a burn-in of 200 000 followed by 1 000 000 Markov chain Monte-Carlo iterations. Scenarios with $K = 1$ to 7 were evaluated with 20 replicates of each K. The most likely number of clusters was determined using the DeltaK method (Evanno *et al.*, 2005) implemented in StructureHarvester 0.6.94 (Earl & vonHoldt, 2012). Replicate simulations of each K scenario evaluated in Structure were summarized using CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and resulting barplots were visualized using DISTRUCT 1.1 (Rosenberg 2004).

Mantel tests were used to test for a significant correlation between genetic and geographic distance among samples from different geographic areas indicative of an IBD (stepping stone) pattern of gene flow (Wright, 1943) along the U.S. East Coast in GenAlEx. Geographic distance was regressed against both F_{ST} and linearized F_{ST}; the results shown are for F_{ST}. To further test the hypothesis of restricted dispersal, we looked for positive and negative correlations of individual genotypes with space using the autocorrelation coefficient (r) (Smouse

& Peakall, 1999; Peakall *et al.*, 2003; Smouse *et al.*, 2008). Autocorrelation analysis was performed for multiple distance class sizes to compare individuals sampled from the same locality (distance class 0–50 km) to those collected at a range of distances spanning the entire dataset from 0–3100 km. For all analyses, geographic distances were calculated as the shortest distance between sampling areas along the continental shelf. The significance of *r* was assessed by permuting individuals among geographic locations using 10 000 iterations of the data and 95% C.I. were generated using 1000 bootstrap trials.

2.4 | Mitochondrial DNA

The mitochondrial (mt)DNA primers of Nohara *et al. (*2010) were used to amplify and sequence a 489 bp segment of the mtDNA CR from a subset of all collected samples (Table 2). Briefly, each 25 µl reaction contained 1x PCR Buffer (Qiagen) 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.125 µM of each primer, 0.5 unit of Taq polymerase (Qiagen) and 1 µl genomic DNA. Amplifications were performed with an initial denaturation temperature of 94º C for 2 min, followed by 35 cycles at 94º C for 1 min, 52º C for 1 min, 72º C for 2 min, with a final elongation step at 72º C for 5 min. Aliquots of amplified products were sized against a DNA ladder of known size using horizontal gel electrophoresis (1.5% w/v agarose), stained with GelRed (Biotium) and visualized under a UV light source to confirm the presence of a single amplification product of correct size. Amplification products were purified using a QIAquick

PCR purification kit (Qiagen) following the manufacturers protocol and subsequently quantified using a NanoDrop 2000 (www.nanodrop.org) prior to storage at –20 °C. Purified PCR products were bi-directionally sequenced using the Big Dye Terminator Cycle sequencing kit (Applied Biosystems) with the original amplification primers and 0.25 the recommended concentration of Big Dye. Sequencing reaction products were precipitated using ethanol–sodium acetate to remove unincorporated nucleotides and resuspended in 16 µl of Hi-Di formamide (Applied Biosystems) and 10 µl of each cleaned reaction were electrophoresed on an ABI 3130*xl* Prism genetic analyzer (Applied Biosystems). The resulting forward and reverse sequences were imported into SEQUENCHER 5.1 (Gene Codes Corporation; www.genecodes.com) for trimming of low-quality sequence and creation of consensus sequences. Consensus sequences were aligned in MACVECTOR 12.5.1 (MacVector; www.macvector.com) and exported as a FASTA file.

Base frequencies, number of variable sites and number of parsimony informative sites among sequences were calculated in DNASP v5 (Librado & Rozas, 2009). The FABOX software (Villesen, 2007) was used to collapse sequences into haplotypes and create input files for the Arlequin 3.5.2.2 software package (Excoffier & Lischer, 2010). The evolutionary model that best described the sequence data, the Jukes-Cantor Model (JC) (Jukes & Cantor, 1969) was identified using the model selection method implemented in PAUP* 4.0a (Swofford, 2000, 2003) based on the corrected Akaike information criterion and this model was used to perform the subsequent analysis. Arlequin was used to generate descriptive statistics including mean number of pairwise differences (*k*), haplotype diversity (*H*) and nucleotide sequence diversity (*À*). Arlequin was also

used for population pairwise $_{/ST}$ and AMOVA (Excoffier *et al.,* 1992). Statistical significance was assessed based on 10 000 permutations of the data. The nearest-neighbour statistic (S_{NN}) (Hudson, 2000) was used to assess how often nearest-neighbours in sequence space were found in the same locality and significance was assessed based on 10 000 permutations of the data in DNASP. The S_{NN} statistic was calculated keeping all collection areas separate and, because Cape Hatteras is a well-known biogeographic barrier for other fishes, by grouping samples collected north and south of Cape Hatteras. The POPART software (Leigh & Bryant, 2015) was used to reconstruct and visualize genealogical relationships among sequences using the minimum spanning network algorithm of Bandelt *et al. (*1999). To test whether data were consistent with past demographic expansion (population growth), Fu's F_S (Fu 1997) was calculated in Arlequin. In all cases, statistical significance was assessed using 10 000 bootstrap replicates.

3 | RESULTS

3.1 | Microsatellite marker development

High throughput sequencing of a *C. microps* DNA sample on an Ion Torrent PGM resulted in approximately 4.8 million DNA sequences ranging in length from 25–587 bp with an average Phred quality score of 30. Filtering using the software programs integrated into the Galaxy Project platform resulted in the retention of 4.7 million high-quality DNA sequences. The

retained sequences were subsequently queried for the presence of perfect tetranucleotide repeat loci, resulting in the identification over 8000 microsatellite loci. Of the loci identified, primers were designed for *c.* 1500 loci and 65 primer pairs were ordered and tested for amplification of a product of predicted length from *C. microps* DNA samples from VA and SC. Of the original 65 primer pairs tested, 26 loci amplified reliably across test samples. Further testing of samples taken from VA and SC (20 samples from each location) indicated that these loci were in conformance to HWE expectations. One of the primers failed to amplify successfully across multiple samples and was excluded from further analysis. The remaining 25 loci were combined into 8 multiplex marker panels for genotyping (Table 1).

3.2 | Microsatellite analysis

In total, 503 samples were analysed across 25 polymorphic microsatellite loci; 488 samples from U.S. East Coast range from NY to the southern Florida Keys (FL) and 15 from FLW in the Gulf of Mexico (Figure 1). Detailed capture information is available upon request. All subsequent analyses were carried out both including and excluding these loci. Results were consistent regardless of whether these loci were included, but all results presented henceforth are based on the 23 loci that were in HWE.

All loci were polymorphic, with the number of alleles ranging from 6 at CM1787993 and CM1787993 to 21 at CM90501 (Table 1). The mean effective number of alleles across loci (A_E)

ranged from 4.43 in FLW to 4.49 in NY and mean expected heterozygosity (H_E) ranged from 0.72 in FLW to 0.74 in NY (Table 3). Allelic richness per sample was even across all sample collections regardless of which methods were used. Using the rarefaction method and a samples size of 15, the size of the FLW collection, allelic richness values ranged from 6.57 alleles per locus in the NCN collection to 6.83 in the FLW sample. Using the extrapolation method and a sample size of 50 individuals, estimates ranged from 8.67 in FL to 9.06 in NCS and the 95% confidence intervals were overlapping among all locations. Mean expected heterozygosity (gene diversity) across loci ranged from 0.721 in FLW to 0.739 in NY.

All pairwise F_{ST} values between collection locations were low; the largest calculated value was 0.003 between the FLW and the NCN sample taken North of Cape Hatteras. No pairwise comparisons were significantly different from zero based on 10 000 permutations of the data (± = 0.05; Table 4). An AMOVA using multiple alternate groupings of sampling locations, including pooling of collections into two groups NCN and NCS showed no significant genetic variance due to variation among any tested groups (Supporting Information Table S1). Likewise, a PCA did not indicate the presence of any discrete clusters that would suggest the presence of multiple populations, the first two principal components explained 1.49% and 1.45% of the data and the scree plot had eigenvalues that were flat across the plot indicating a lack of discriminatory alleles (Figure 2). All Structure simulations were consistent with a single genetic group; iterations with $K = 1$ were associated with the highest log-likelihood values.

Mantel tests of the sample collections along the U.S. East Coast based on geographic distance along the shelf *v*. F_{ST} were not significant ($R_{xy} = –0.016$, $P > 0.05$) indicating no detectable isolation IBD, based on 10 000 permutations of the data. When the FLW was included in the analysis, the correlation remained insignificant ($R_{xy} = –0.720$, $P > 0.05$). Likewise, no significant values were observed for the autocorrelation analysis across multiple distance classes. The null hypothesis that genotypes are randomly distributed in space could not be rejected; calculated *r*-values fell within the bounds of the 95% C.I. for all distance classes.

3.3 | MTDNA analysis

In total, 188 control region sequences were examined across a subset of samples from all geographic locations (GenBank Accession Numbers (MH460681 - MH460752). All sequences were edited to a final length of 407 bp, resulting in 72 haplotypes with 55 variable and 31 parsimony informative sites including 58 substitutions, 51 of which were transitions and 7 of which were transversions. A total of four indels were observed. The most common haplotype, haplotype 9, was recovered 39 times (20.7% of sequences) and was recovered in all locations with the exception of the FLW sample (Table S2). The second most common haplotype was recovered 9 times (4.8% of all samples) and was recovered in all locations except Delaware (DE), but FLW and DE had the smallest number of samples sequenced (8 and 13 respectively). Haplotype diversity (*H*) was 0.94 across all samples and was high in all geographic samples

ranging from 0.89 in samples from NCN to 1.0 in DE (Table 2). The mean number of pairwise differences between sequences (k) across all samples was 3.1 and ranged from 2.4 in NCN to 3.9 in SC. Likewise, there were few differences among haplotypes; nucleotide diversity ($\grave{A}$) was low both across all samples (0.008) and within samples from each geographic collection location (0.006 in NCN to 0.010 in SC), indicating that there were very few differences among haplotypes (Table 2).

Population pairwise D_{ST} values based on the Jukes–Cantor distance between most pairs of sample collections examined was 0, but ranged to 0.017 between NCN and FLW. All values were non-significant based on 10 000 permutations of the data (± = 0.05; Table 4). The S_{NN} statistic was non-significant (S_{NN} = 0.15660, $P > 0.05$), suggesting that the nearest neighbours in sequence space are not significantly associated in geographic space. Likewise, a minimum spanning network showed no evidence that haplotypes were restricted to particular sampling locations (Figure 3). As with the analysis of the microsatellite data, an AMOVA (Excoffier *et al.,* 1992) using multiple alternate groupings of sampling locations showed no significant genetic variance due to variation among any grouping scheme (Supporting Information Table S1). Tests aimed at elucidating the demographic history of *C. microps* were consistent with a model of population expansion. Calculated values of Fu's F_S were negative and statistically significant for all geographic collection locations (Table 2).

4 | DISCUSSION

The focus of this study was the delineation of stock structure in *C. microps* to provide information in support of fisheries management, including stock assessment and conservation. Although the assumed sedentary nature of adult *C. microps* suggests population structuring, no evidence of distinct genetic stocks along the U.S. East Coast range from New York to the southern Florida Keys was observed using either nuclear microsatellite loci or sequencing of the mtDNA control region. Rather, data were consistent with sufficient gene flow to prevent the accumulation of significant genetic differences.

Marine fishes have opportunities for dispersal during egg, larval and adult stages and understanding which phase drives observed structure is difficult when life-history data are sparse. Assuming that a sedentary adult phase is indicative of population structure fails to account for the influence of opportunities for mixing at other life stages. Alternatively, assuming a high level of connectivity due to long pelagic larval duration, migratory ability or lack of physical barriers ignores the possibility that site fidelity (Bonanomi *et al.*, 2016), larval retention (Swearer *et al.*, 2002; Teske *et al.*, 2016), habitat (Natoli *et al.*, 2005) or oceanographic boundaries (Galarza *et al.*, 2009; Pascual *et al.*, 2017) play a role in maintenance of genetic structure. Thus, fishes occupying similar habitats with ostensibly equivalent life history characteristics, for example slow-growing, late to mature, long-lived species occupying deep-water habitats such as snappers, groupers, jacks, grunts, porgies and tilefishes have been found to have vastly different levels of genetic structuring. Ball *et al. (*2007) sampled red porgy *Pagrus*

pagrus (L. 1758) across their U.S. western Atlantic Ocean range and found a lack of detectable genetic structure, suggesting that gene flow is sufficient to prevent the accumulation of genetic differences. This observed genetic homogeneity was attributed to the 30 day pelagic larval stage and occasional adult movement. A study of red snapper *Lutjanus campechanus* (Poey 1860) sampled from the south-eastern U.S. waters of the Atlantic Ocean and Gulf of Mexico (GOM) found evidence that gene flow between the GOM and Atlantic regions was restricted and the results suggested variable levels of connectivity within regions across time (Hollenbeck *et al.,* 2015). Tringali & Higham (2007) sampled vermillion snapper *Rhomboplites aurorubens* (Cuvier 1829) from 9 locations off the south-eastern U.S. Atlantic coast and GOM conformed to an IBD model, suggesting restricted gene flow (dispersal) both between the GOM and Atlantic Ocean and between the eastern and western GOM. The observed pattern was attributed to the capacity for moderate levels of larval transport. At the other end of the spectrum, black sea bass *Centropristis striata* (L. 1758) sampled from the GOM to Massachusetts were found to comprise distinct subspecies between GOM and Atlantic Ocean while the Atlantic coast harboured two genetically differentiated populations separated by Cape Hatteras, NC (Roy *et al.*, 2012; Mccartney *et al.*, 2013). These differences were attributed to phylogeographic barriers to gene flow. Hogfish *Lachnolaimus maximus* (Walbaum 1702) sampled by Seyoum *et al.* (2015) from the south-eastern U.S. range from the Carolinas to the GOM exhibited significant differences among samples from the Carolinas and the east and west coasts of Florida. These differences were also attributed to phylogeographic breaks since no significant pattern of IBD was observed

either among all regions or within a sampling region. For gray snapper *Lutjanus griseus* (L. 1758) Gold *et al.* (2009) found the pattern was potentially more complex. Samples from five localities in the northern GOM and one locality on the Atlantic coast of Florida were inconclusive, but suggested a pattern of IBD interrupted by barriers to gene flow at greater distances. The observed genetic structure was attributed to the fact that larvae and juveniles use shallow inshore estuaries and mangroves as nursery grounds.

Although *C. microps* are assumed to be sedentary bottom dwellers as adults, they have a protracted spawning season (February–November; Harris *et al.*, 2004; Kolmos *et al.*, 2016) and release pelagic eggs in batches along the north-eastern U.S. continental shelf, which presents ample opportunities for mixing. In addition, evaluation of the reproductive state of 1281 female specimens collected from New Jersey to Florida confirmed that a high fraction of mature females were present in all sampling areas (Kolmos *et al.*, 2016).

Pelagic larvae present another opportunity for mixing across a wide geographic range. There is a lack of data concerning pelagic larval duration (PLD) in *C. microps*, but to assess the potential of dispersal of eggs and larvae to mediate connectivity Farmer & Klibansky (2016) plotted the catch position of females estimated to be within 48 h of spawning, based on histology against observations of currents from surface drifter buoys deployed by the Global Drifter Program (GDP; www.aoml.noaa.gov/phod/dac/index.php) and the South Carolina Department of Natural Resources (G. Sedberry, pers. comm.). They concluded that the Loop Current and Gulf Stream provide a mechanism for transport from the GOM coast of Florida to the U.S. East Coast

and northward and counter current eddies probably mediate north to south movement. Taken together, the data on spawning period and current mediated dispersal of eggs and larvae is consistent with sufficient gene flow to maintain genetic homogeneity in blueline tilefish.

Other deep-water reef fishes found to have a high level of genetic homogeneity, including *P. pagrus* (Ball *et al.,* 2007) and *L. campechanus* (Hollenbeck *et al.,* 2015), are also highly fecund, long-lived batch-spawners with protracted spawning seasons; November–May in *P. pagrus* porgy (Manooch, 1976; Daniel, 2003) and May–October in *L. campechanus* (White & Palmer, 2004). The high level of connectivity in these fishes has also been ascribed to the increased opportunities for mixing afforded by the extended spawning season. Other factors, including the presence of relatively warm stable temperatures year round along the shelf of the U.S. Middle Atlantic Bight (9–14° C; Grimes *et al.*, 1986) at depths of *c.* 100–300 m.

It has been suggested that the presence of *C. microps* as well as other deep-water reef associated species north of Cape Hatteras may be attributable to a range expansion (Nye *et al.*, 2009; Møller *et al.*, 2010). However, if *C. microps* north of Cape Hatteras represent a range expansion, a concomitant reduction in genetic diversity at the edge of the northward expansion and genetic heterogeneity would be predicted (Eckert *et al.*, 2008). On the contrary, values of allelic richness were remarkably stable across the sampling locations in this study and there was no evidence of population structure, IBD or significant autocorrelation across any distance class examined, consistent with widespread dispersal., In addition, values of Fu's *F* were all significantly negative, consistent with demographic expansion in all populations examined.

Taken together, this suggests that *C. microps* are not recent colonizers north of Cape Hatteras. Indeed, the presence of *C. microps* was noted as far north as Cape Henry, VA, as early as 1937 (Firth, 1937).

The results of this study are in agreement with the suggestion of Ross *et al. (*2016) that a lack of appropriate sampling due to the difficulty of conducting surveys in the U.S. Middle Atlantic Bight because of the depth and topography has caused these species to be poorly documented. This is further exacerbated by the fact that, although annual large-scale underwater video surveys are conducted along the continental shelf and upper slope as part of the Southeast Reef Fish Survey, these surveys end at Cape Hatteras (Bacheler *et al.*, 2016). In addition, the absence of a commercial *C. microps* code in the 1990s, combined with relatively low observer coverage, probably contributed to the lack of recorded *C. microps* bycatch in the observer data (Nitschke & Miller, 2016).

The results of this study are consistent with the suggestion that there is sufficient gene flow to prevent the accumulation of genetic differences in *C. microps* and supports management of *C. microps* as a single stock along the U.S. East Coast. Although this study included a limited number of samples from the GOM (FLW), the sample size was not sufficient to draw definitive conclusions concerning the level of connectivity between the GOM and Atlantic regions. Future studies should incorporate samples from across the GOM range of *C. microps* and should include comparison of eastern and western GOM samples to address the full extent of mixing. Future studies should also include samples from Georgia south to the Florida Keys, which were not

represented in this study. Finally, traditional or genetic capture–mark–recapture, or acoustic tagging methods should be used to test the assumption that adult *C. microps* are sedentary.

ACKNOWLEDGEMENTS

We thank T. Kellison for coordination of sample collection. We thank J. Didden, S. Musick, S. Feller, D. Portnoy and T. Darden for contributing additional samples. We thank two anonymous reviewers for their valuable input.

REFERENCES

Allendorf, F. W., Berry, O. & Ryman, N. (2014) So Long to Genetic Diversity and Thanks for All the Fish. *Molecular Ecology*. 2014, 23–25, doi:10.1111/mec.12574

Andrews, S. (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data. Cambridge: Babraham Institute. Available online at: www.bioinformatics.babraham.ac.uk/projects/fastqc

Bacheler, N. M., Schobernd, Z. H., Berrane, D. J., Schobernd, C. M., Mitchell, W. A., Teer, B. Z., Gregalis, K. C. & Glasgow, D. M. (2016) Spatial Distribution of Reef Fish Species along the Southeast Us Atlantic Coast Inferred from Underwater Video Survey Data. *PLoS ONE* **11**. doi.org/10.1371/journ*al.*pone.0162653

Ball, A. O., Beal, M. G., Chapman, R. W. & Sedberry, G. R. (2007) Population Structure of Red Porgy, *Pagrus pagrus*, in the Atlantic Ocean. *Marine Biology* **150**, 1321–1332.

Bandelt, H.-J., Peter Forster and A. R. & ¨hl. (1999) Median-Joining Networks for Inferring Intraspeci c Phylogenies. *Molecular Biology* **16**, 37–48.

Blankenberg, D., Kuster, G. Von, Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A. & Taylor, J. (2010) Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. *Current Protocols in Molecular Biology*. 2010, doi:10.1002/0471142727.mb1910s89.

Bonanomi, S., Overgaard Therkildsen, N., Retzel, A., Berg Hedeholm, R., Waever Pedersen, M., Meldrup, D., Pampoulie, C., Hemmer-Hansen, J., Grønkjaer, P. & Nielsen, E. E. (2016) Historical DNA Documents Long Distance Natal Homing in Marine Fish. *Molecular Ecology* 25: 2727–2734. doi:10.1111/mec.13580

Colwell, R. K., Chang, X. M. & Chang, J. (2004) Interpolating, Extrapolating and Comparing Incidence-Based Species Accumulation Curves. *Ecology* **85**, 2717–2727.

Dooley, J. K. (1978) Systematics and Biology of the Tilefishes (Perciformes:Branchiostegidae and Malacanthidae) with Descriptions of Two New Species. *NOAA Technical Report NMFS Circular* 411. Washington, D.C.: U.S. Department of Cmmerce. Available at www. spo.nmfs.noaa.gov/Circulars/CIRC411.pdf

Earl, D. A. & vonHoldt, B. M. (2012) Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. *Conservation*

Genetics Resources **4**, 359–361.

Eckert, C. G., Samis, K. E. & Lougheed, S. C. (2008) Genetic Variation across Species' Geographical Ranges: the Central-Marginal Hypothesis and beyond. *Molecular Ecology*. 2008, 1170–1188, doi:10.1111/j.1365-294X.2007.03659.x.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the Number of Clusters of Individuals Using the Software Structure: a Simulation Study. *Molecular Ecology* **14**, 2611–2620.

Excoffier, L. & Lischer, H. E. L. (2010) Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–567.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992) Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: application to Human Mitochondrial DNA Restriction Data. *Genetics* **131**, 479–491.

Faircloth, B. C. (2008) MSATCOMMANDER: Detection of Microsatellite Repeat Arrays and Automated, Locus-Specific Primer Design. *Molecular Ecology Resources* **8**, 92–94.

Falush, D., Stephens, M. & Pritchard, J. K. (2007) Inference of Population Structure Using Multilocus Genotype Data: dominant Markers and Null Alleles. *Molecular Ecology Notes* **7**, 574–578.

Farmer, N. & Klibansky, N. (2016) Distribution of *C. microps* (*Caulolatilus microps*) in the U.S. EEZ from Fishery-Dependent and Fishery-Independent Data Collections. *SEDAR50-DW11. SEDAR, North Charleston, SC.*

Firth, A. F. E. (1937) Recent Records Extending the Range of Caulolatilus Microps North of Florida. *Copeia* **1937**, 189.

Fu, Y. X. (1997) Statistical Tests of Neutrality of Mutations against Population Growth, Hitchhiking and Background Selection. *Genetics* **147**, 915–925.

Galarza, J. A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G. F. & Rico, C. (2009) The Influence of Oceanographic Fronts and Early-Life-History Traits on Connectivity among Littoral Fish Species. *Proceedings of the National Academy of Sciences of the U.S.A.* **106**, 1473–1478.

Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., *et al.* (2005) Galaxy: A Platform for Interactive Large-Scale Genome Analysis. *Genome Research* **15**, 1451–1455.

Goecks, J., Nekrutenko, A. & Taylor, J. (2010) Galaxy: A Comprehensive Approach for Supporting Accessible, Reproducible and Transparent Computational Research in the Life Sciences. *Genome biology* **11**, R86.

Gold, J. R., Saillant, E., Ebelt, N. D. & Lem, S. (2009) Conservation Genetics of Gray Snapper (*Lutjanus griseus*) in U.S. Waters of the Northern Gulf of Mexico and Western Atlantic Ocean. *Copeia* **2009**, 277–286.

Goode, G. B. & Bean, T. H. (1878) Description of *Caulolatilus microps*, a New Species of Fish from the Gulf Coast of Florida. *Proc. U.S. Natl. Mus* **1**, 42–45.

Grimes, C.B., Able, K.W., Jones, R.S. (1986) Tilefish, Lopholatilus Chamaeleonticeps, habitat,

behavior and community structure in Mid-Atlantic and southern New England waters. *Environmental Biology of Fishes* **15**, 273.

Guo, S. W., Guo, S. W., Thompson, E. a & Thompson, E. a. (1992) Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics* **48**, 361–372.

Harris, P. J., Wyanski, D. M. & Powers Mikell, P. T. (2004) Age, Growth and Reproductive Biology of *C. microps* along the Southeastern Coast of the United States, 1982–1999. *Transactions of the American Fisheries Society* **133**, 1190–1204.

Hollenbeck, C. M., Portnoy, D. S., Saillant, E. & Gold, J. R. (2015) Population structure of red snapper (*Lutjanus campechanus*) in U.S. waters of the western Atlantic Ocean and the northeastern Gulf of Mexico. *Fisheries Research* **172**, 17–25.

Hubisz, M. J., Falush, D., Stephens, M. & Pritchard, J. K. (2009) Inferring Weak Population Structure with the Assistance of Sample Group Information. *Molecular Ecology Resources* **9**, 1322–1332.

Hudson, R. R. (2000) A New Statistic for Detecting Genetic Differentiation. *Genetics* **155**, 2011–2014.

Hurlbert, S. H. (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577–586.

Jakobsson, M. & Rosenberg, N. A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806.

Jombart, T. (2008) Adegenet: A R Package for the Multivariate Analysis of Genetic Markers. *Bioinformatics (Oxford, England)* **24**, 1403–1405.

Jombart, T., Pontier, D. & Dufour, A.-B. (2009) Genetic Markers in the Playground of Multivariate Analysis. *Heredity* **102**, 330–341.

Jukes, T. H. & Cantor, C. R. (1969) Evolution of Protein Molecules. *Mammalian Protein Metabolism* 21–123.

Kellison, T. (2016) Summary of the 2015 blueline tilefish Cooperative-with-Industry Data Collection Project SEDAR50-DW02. North Charleston, SC: SEDAR. Available at www.sedarweb.org/docs/wpapers/S50_DW02_Kellison_2015CoopDCProg_6.17.2016.pdf

Kolmos, K. J., Falk, S., Wyanski, D. M. & Schmidtke, M. A. (2016) A Preliminary Assessment of Reproductive Parameters for blueline tilefish in Atlantic Waters from Virginia to Florida SEDAR50-DW03. North Charleston, SC: *S*EDAR. Available at www.sedarweb.org/docs/wpapers/S50_DW03_Kolmos.etal_BLTRepro_7.15.2016_updated_withnote.pdf

Koressaar, T. & Remm, M. (2007) Enhancements and Modifications of Primer Design Program Primer3. *Bioinformatics* **23**, 1289–1291.

Laikre, L., Palm, S. & Ryman, N. (2005) Genetic population structure of fishes: implications for coastal zone management. *Ambio* **34**, 111–119.

Leigh, J. W. & Bryant, D. (2015) Popart: full-feature software for haplotype network

construction. *Methods in Ecology and Evolution* **6**, 1110–1116.

Lewis, L. A., Richardson, D. E., Zakharov, E. V. & Hanner, R. (2016) Integrating DNA barcoding of fish eggs into ichthyoplankton monitoring programs. *Fishery Bulletin* **114**, 153–165.

Librado, P. & Rozas, J. (2009) DnaSP v5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics* **25**, 1451–1452.

Longmire, J. L., Maltbie, M. & Baker, R. J. (1997) Use Of 'lysis Buffer' in DNA Isolation and Its Implication for Museum Collections. *Occasional Papers of the Museum of Texas Tech University*. 1997, 1–3.

Loon, E. E. van, Cleary, D. F. R. & Fauvelot, C. (2007) ARES: software to compare allelic richness between uneven samples. *Molecular Ecology Notes* **7**, 579–582.

Manooch, C. S. (1976) Reproductive Cycle, Fecundity and Sex Ratios of the Red Porgy, *Pagrus pagrus* Linnaeus (Pisces: Sparidae) in North Carolina Waters. *Fishery Bulletin* **74**, 775–781.

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.

Møller, P. R., Nielsen, J. G., Knudsen, S. W., Poulsen, J. Y., Sünksen, K. & Jørgensen, O. A. (2010) A Checklist of the Fish Fauna of Greenland Waters. *Zootaxa* 1–84.

Natoli, A., Birkun, A., Aguilar, A., Lopez, A. & Hoelzel, A. R. (2005) Habitat Structure and the

Dispersal of Male and Female Bottlenose Dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society B* **272**, 1217–1226.

Nohara, K., Takeuchi, H., Tsuzaki, T., Suzuki, N., Tominaga, O. & Seikai, T. (2010) Genetic Variability and Stock Structure of Red Tilefish *Branchiostegus japonicus* Inferred from mtDNA Sequence Analysis. *Fisheries Science* **76**, 75–81.

Nye, J. A., Link, J. S., Hare, J. A. & Overholtz, W. J. (2009) Changing Spatial Distribution of Fish Stocks in Relation to Climate and Population Size on the Northeast United States Continental Shelf. *Marine Ecology Progress Series* **393**, 111–129.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538.

Pascual, M., Rives, B., Schunter, C. & Macpherson, E. (2017) Impact of life history traits on gene flow: a multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *Plos One* **12**, e0176419.

Peakall, R. & Smouse, P. E. (2012) GenALEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research-an Update. *Bioinformatics* **28**, 2537–2539.

Peakall, R., Ruibal, M. & Lindenmayer, D. B. (2003) Spatial autocorrelation analysis offers new insights into gene flow in the australian bush rat, *Rattus fuscipes*. *Evolution* **57**, 1182–1195.

Petit, R. J., Mousadik, A. E. L. & Pons, O. (1998) Identifying Populations for Conservation on

the Basis of Genetic Markers. *Conservation Biology* **12**, 844–855.

Pinsky, M. L. & Palumbi, S. R. (2014) Meta-Analysis Reveals Lower Genetic Diversity in Overfished Populations. *Molecular Ecology* **23**, 29–39.

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

Rice, W.R. (1989) Analyzing tables of statistical test. *Evolution*, **43**: 223–225.

Rosenberg, N. A. (2004) DISTRUCT: A Program for the Graphical Display of Population Structure. *Molecular Ecology Notes* **4**, 137–138.

Ross, S. W., Rhode, M., Viada, S. T. & Mather, R. (2016) Fish Species Associated with Shipwreck and Natural Hard-Bottom Habitats from the Middle to Outer Continental Shelf of the Middle Atlantic Bight near Norfolk Canyon. *Fishery Bulletin* **114**, 45–57.

Rousset, F. (2008) GENEPOP'007: A Complete Re-Implementation of the GENEPOP Software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–106.

Roy, E. M., Quattro, J. M., W, G. T. & Greig, T. W. (2012) Genetic Management of Black Sea Bass: Influence of Biogeographic Barriers on Population Structure. *Marine and Coastal Fisheries: Dynamics, Management and Ecosystem Science* **4**, 391–402.

Schwartz, M. K., Luikart, G. & Waples, R. S. (2007) Genetic Monitoring as a Promising Tool for Conservation and Management. *Trends in ecology & evolution* **22**, 25–33.

SEDAR 32. (2013) SEDAR 32 Stock Assessment Report South Atlantic Blueline Tilefish. North Charleston, SC: SEDAR. Available at http://sedarweb.org/docs/sar/S32_SA-

BLT_SAR_FINAL_11.26.2013.pdf

Seyoum, S., Collins, A. B., Puchulutegui, C., McBride, R. S. & Tringali, M. D. (2015) Genetically Determined Population Structure of Hogfish (Labridae: *Lachnolaimus maximus*) in the Southeastern United States. *Fishery Bulletin* **113**, 442–455.

Smouse, P. E. & Peakall, R. (1999) Spatial Autocorrelation Analysis of Individual Multiallele and Multilocus Genetic Structure. *Heredity* **82 (Pt 5)**, 561–573.

Smouse, P. E., Peakall, R. & Gonzales, E. (2008) A Heterogeneity Test for Fine-Scale Genetic Structure. *Molecular Ecology* **17**, 3389–3400.

Spies, I. & Punt, A. E. (2015) The Utility of genetics in marine fisheries management: a simulation study based on Pacific cod off Alaska. *Canadian Journal of Fisheries and Aquatic Sciences* **72**, 1415–1432.

Swearer, S. E., Shima, J. S., Hellberg, M. E., Thorrold, S. R., Jones, G. P., Robertson, D. R., Morgan, S. G., Selkoe, K. A., Ruiz, G. M. & Warner, R. R. (2002) Evidence of Self-Recruitment in Demersal Marine Populations. *Bulletin of Marine Science*. 2002, 251–271.

Swofford, D. L. (2000) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)' , Version 4.* Sunderland, MA: Sinauer Associates.

Swofford, D. L. (2003) PAUP*: Phylogenetic Analysis Using Parsimony, Version 4.0b10. Sunderland, MA: Sinauer Associates.

Teske, P. R., Sandoval-Castillo, J., van Sebille, E., Waters, J. & Beheregaray, L. B. (2016)

Oceanography Promotes Self-Recruitment in a Planktonic Larval Disperser. *Scientific Reports* **6**, 34205.

Tringali, M. D. & Higham, M. (2007) Isolation-by-Distance Gene Flow among Vermilion Snapper (*Rhomboplites aurorubens* Cuvier 1829) from the Gulf of Mexico and Southeastern United States. *Gulf of Mexico Science* **25**, 2–14.

Tuckey, T., Yochum, N., Hoenig, J., Lucy, J. & Cimino, J. (2007) Evaluating localized *vs*. large-scale management: the example of tautog in Virginia. *Fisheries* **32**, 21–28.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. & Rozen, S. G. (2012) Primer3-New Capabilities and Interfaces. *Nucleic Acids Research* **40**. doi.org/10.1093/nar/gks596.

Villesen, P. (2007) FaBox: An Online Toolbox for FASTA Sequences. *Molecular Ecology Notes* **7**, 965–968.

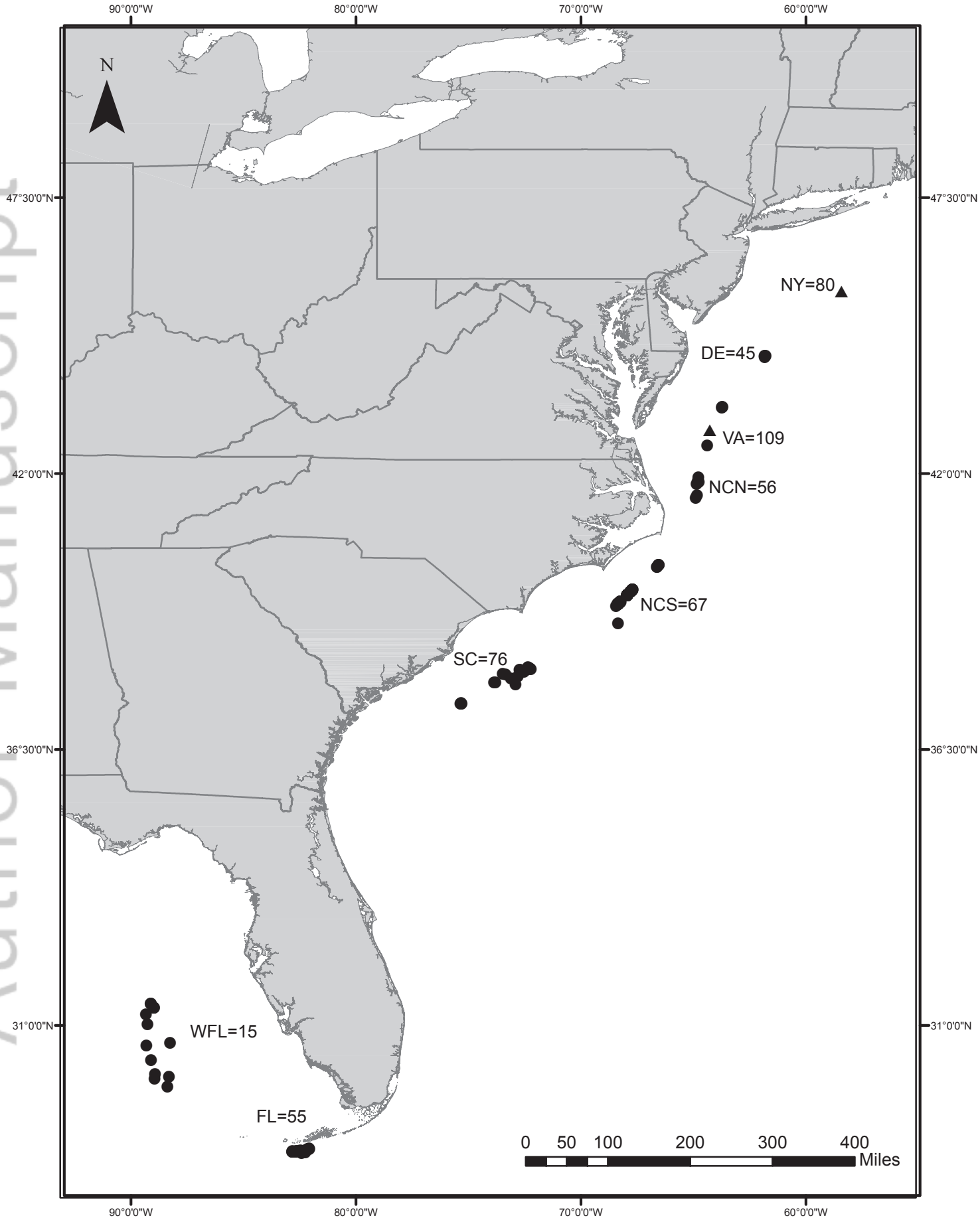
Weir, B. S. & Cockerham, C. C. (1984) Estimating *F*-Statistics for the Analysis of Population Structure. *Evolution* **38**, 1358–1370.

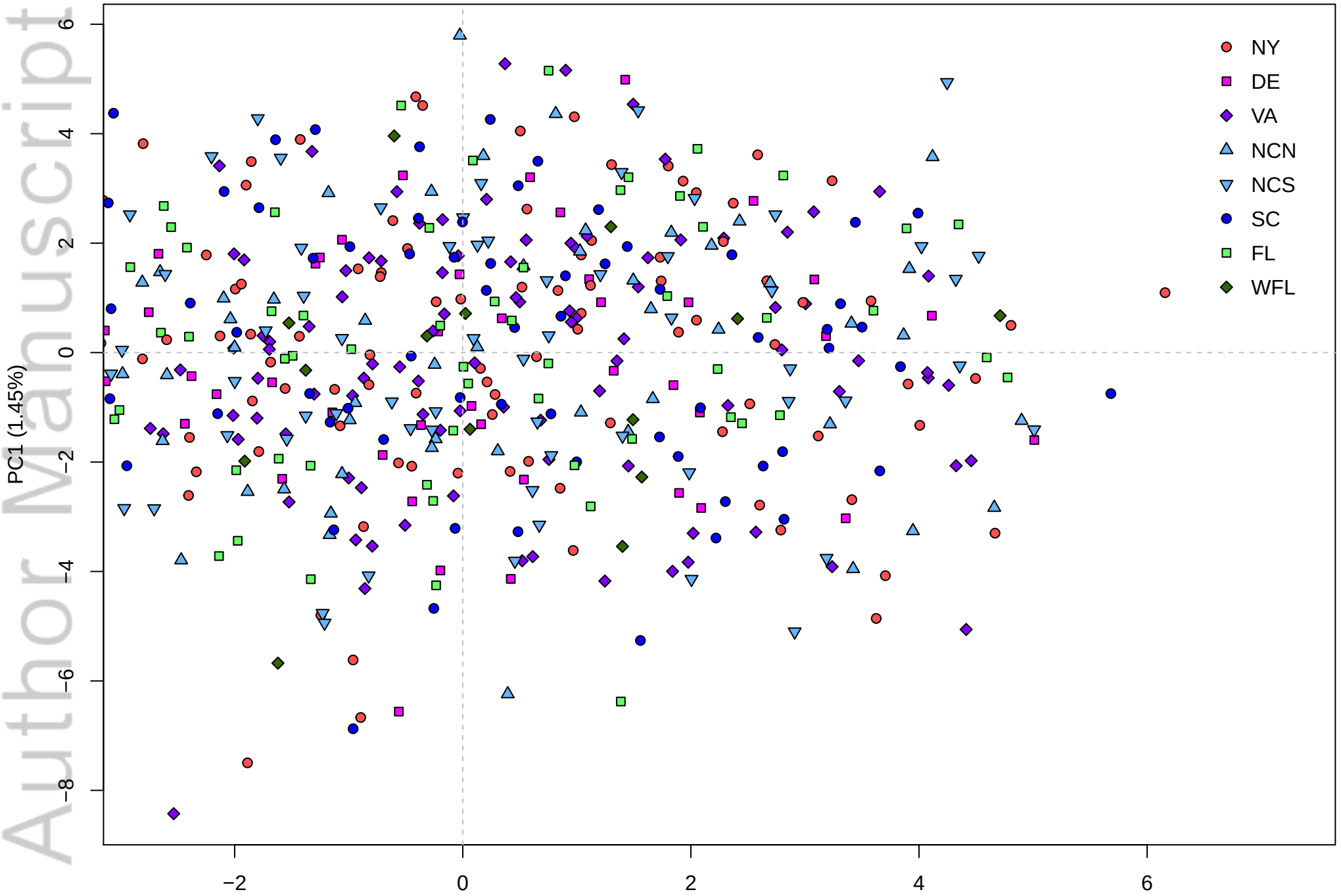
White, D. B. & Palmer, S. M. (2004) Age, Growth and Reproduction of the Red Snapper, Lutjanus Campechanus, from the Atlantic Waters of the Southeastern U.S. *Bulletin of Marine Science* **75**, 335–360.

Wright, S. (1943) Isolation by Distance. *Genetics* **28**, 114–138.

Ying, Y., Chen, Y., Lin, L., Gao, T. & Quinn, T. (2011) Risks of Ignoring Fish Population Spatial Structure in Fisheries Management. *Canadian Journal of Fisheries and Aquatic*

Sciences **68**, 2101–2120.





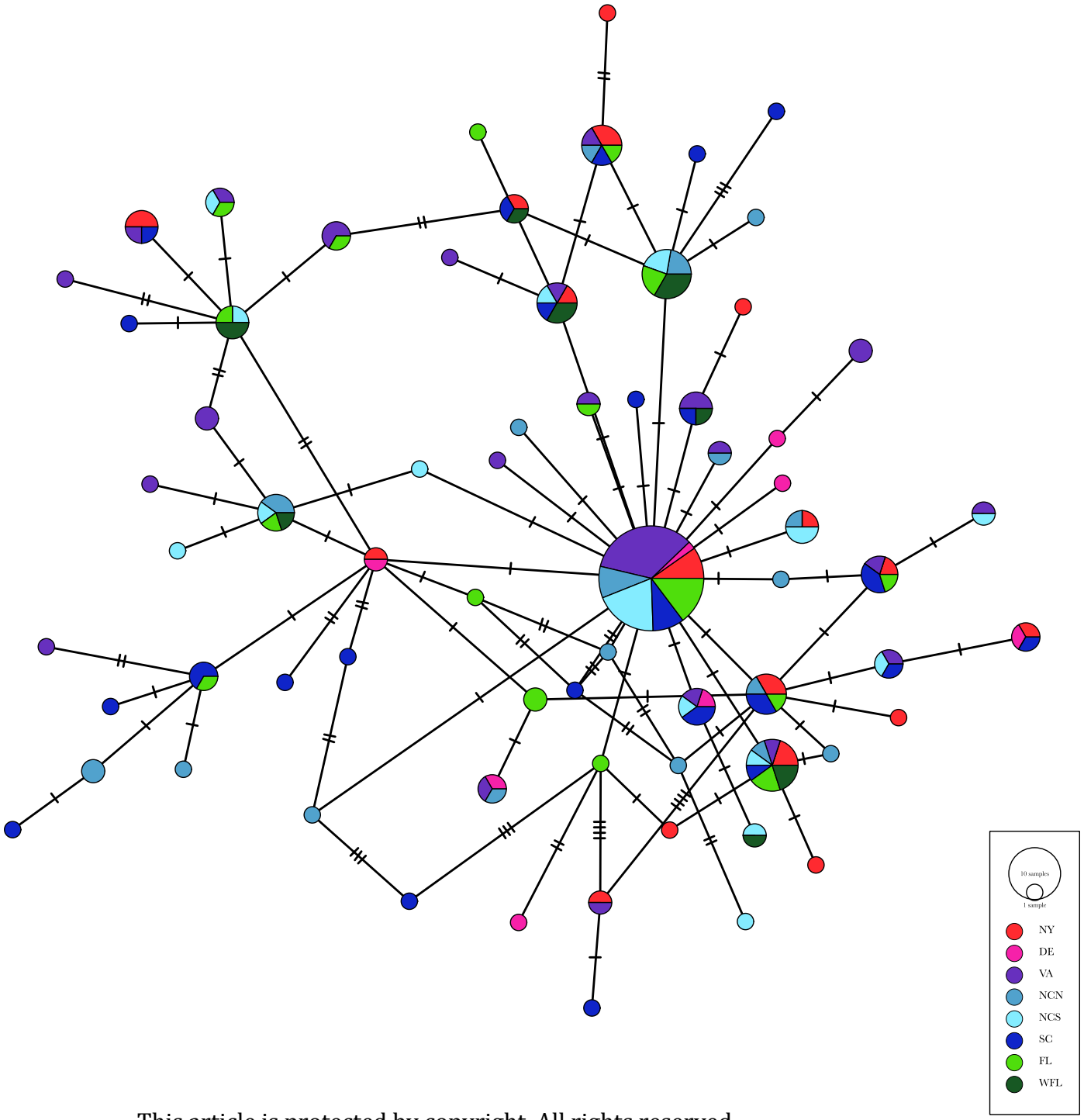


Figure Captions

FIGURE 1 Sampling locations for *Caulolatilus microps*: ●, closed circles indicate known capture latitude and longitude; ▲, sample collector reported an approximate location or statistical area. Sample size given beside each state or regional abbreviation: NY, New York; DE, Delaware; VA, Virginia; NCN, North Carolina north of Cape Hatteras; NCS, North Carolina south of Cape Hatteras; SC, South Carolina; FL, Florida Keys; WFL, western Florida.

Typesetter

1 Delete 002w and 002N from lat and long values.

2 Replace compass arrow with simple latin cross.

3 Replace NCN and NCS with NCN and NCS (small closing caps).

FIGURE 2 Two-dimensional PCA plots displaying principal component axes one and two for *Caulolatilus microps*. Points are colored by sampling location. NY, New York; DE, Delaware; VA, Virginia; NCN, North Carolina north of Cape Hatteras; NCS, North Carolina south of Cape Hatteras; SC, South Carolina; FL, Florida Keys; WFL, western Florida.

Typesetter

1 Replace NCN and NCS with NCN and NCS (small closing caps).

FIGURE 3 Minimum spanning network of the relationship among *Caulolatilus microps*

mtDNA control region haplotypes. #, The number of base pair differences between haplotypes. NY, New York; DE, Delaware; VA, Virginia; NCN, North Carolina north of Cape Hatteras; NCS, North Carolina south of Cape Hatteras; SC, South Carolina; FL, Florida Keys; WFL, western Florida.

Typesetter

1 Replace NCN and NCS with NCN and NCS (small closing caps).

TABLES

TABLE 1 Microsatellite primers for *Caulolatilus microps*

Locus	Forward 5' to 3'	Reverse 5' to 3'	Tail	Dye	T_A (°C)	Repeat motif	Group	Allele Size Range	Number of Alleles
CM-2316467	GGCCTACACCCATGCAAAC	TTTCTCCAACCGCAATGTCG	T7	VIC	48	$(AGAT)_{10}$	1	216-256	10
CM-2352680	CTCCCTGTTCTAATGACCGC	TTTGCCACCAACCTTCTCTG	M13F	NED	48	$(AGAT)_{12}$	1	154-186	8
CM-2492523	GTGAGAAGAAACCAAACACTCC	CCACTGTAACTGCTATACCTGG	M13R	PET	48	$(AGGC)_{10}$	1	170-209	10
CM-54794	TTGAAATGCTGCTATGACAACC	GGTCGGCTGCACTATTTCAG	T3	FAM	48	$(AGAT)_{10}$	1	193-224	10
CM-1787993	AGCTGGAACCAAGAAATGCG	GAAACGCTCGGGTTACATCG	T7	FAM	49	$(AGAT)_{10}$	2	241-261	6
CM-4718692	AAGAGGACCAGGAACGAGTG	CTATTTCAACGCCGGTGGAG	M13F	NED	49	$(AGAT)_{16}$	2	170-221	13
CM-931277	TCACCACGACTGCCACTG	GACAGGCTCACACATACTGC	T3	VIC	49	$(AGAT)_{14}$	2	139-200	unk
CM-1827829	AGAAGAGCAAGACAGGAGCC	TGGCTCCATTCACCAATAATGC	M13F	NED	54.6	$(AGAT)_{11}$	3	172-207	10
CM-2660427	GACAGCCTTATTCCTCGCAG	TACCGGACTGCTGCTATGAC	M13R	PET	54.6	$(AGAT)_{10}$	3	158-201	11
CM-1065459	GACACGCGGTTCTCCAAAG	TCCTTGGGCCATGTTGTAGG	T3	VIC	54.6	$(AGAT)_{11}$	3	278-333	14
CM-459957	TGCTCTCTGTCCCGAGTATTC	AGAGGGATCGAAAGCTGAGTC	T3	NED	55	$(ACAG)_{13}$	4	163-241	16
CM-2212380	AGGGAGGAGGTTTCACACTG	ATCACTGCCATCTTCCCACC	TailA	FAM	55	$(AATG)_{11}$	4	187-214	8

CM-2149957	GTGTATGAGACCCAGAGCCC	TTGGAGTCCGGCTGTCTATG	T7	VIC	55	$(ATCC)_{11}$	4	242-275	10
CM-2374475	CTGTTGCTTCTAGTCTCTGGC	GATCGGCTCCTCTCCCAC	T7	FAM	58.7	$(AGAT)_{11}$	5	115-200	18
CM-4591723	TGTTGGTTTCTGCCTTGGAG	AGACATGCAGATAGACGGAGAC	M13F	NED	58.7	$(AGAT)_{17}$	5	140-198	14
CM-2404273	TACCGGACTGCTGCTATGAC	GGAAGAGCTGCATTTCTCCG	M13R	PET	58.7	$(AGAT)_{10}$	5	126-152	7
CM-310413	CCTCCTGCACTGTTTCCTG	ACCTGAATTTCCCTCGGTACC	T3	VIC	58.7	$(AGAT)_{10}$	5	104-149	12
CM-1191685	TGTCCAGCGCATCAATAAAGG	CTCGCCATGTCACAGTGTTG	T7	FAM	62	$(AGAT)_{22}$	6	129-154	6
CM-2186404	ACCCGTGGATAAGCGGTATAG	GTGAGTGTCAATCAGGGAAGG	M13R	PET	62	$(ATCC)_{13}$	6	306-359	13
CM-732781	GGATACCTACATTTCCCTCAGG	AGTCTGTCTACATTGTCATCGC	T3	VIC	62	$(AGAT)_{10}$	6	154-183	8
CM-90501	ACCACCTCACATCTGACCAC	TGTCATCTCATCTCAGCCAAG	T7	FAM	65	$(AGAT)_{12}$	7	241-324	21
CM-1080088	GTCCAACAGCAGTCCTTGAAG	CACACAGAAAGGCGGGATTC	T3	NED	65	$(AGAT)_{18}$	7	206-270	16
CM-1009046	TCCAAACTGTTTCCAAGGGC	TGTTCCCAGACATGTGTAGC	M13R	PET	63.5	$(AGAT)_{10}$	8	205-244	11
CM-764003	TGTACAACCTCACCGCCTAG	GGTTAATAACAGCCCAGGAGC	T3	FAM	63.5	$(AGAT)_{12}$	8	126-177	12
CM-4391826	GGTCGATGGTAGAAGCACAC	CGCTCCCATCCATCAACAAC	TailB	VIC	63.5	$(AGAT)_{18}$	8	83-161	unk

Tail, The commercially available sequences with the exception of TailA and TailB, which are from Blacket *et al.*, 2012; T_A, annealing temperature; Repeat motif, repeat motif and number of repeats in the sample used for discovery; Group, multiplex group. T7, TAATACGACTCACTATAGGG; M13F, GGTAAAACGACGGCCAGT; M13R, GGCAGGAAACAGCTATGAC; T3, AATTAACCCTCACTAAAGGG; TailA, GCCTCCCTCGCGCCA; TailB, GCCTTGCCAGCCCGC

TABLE 2 Summary statistics for mitochondrial DNA control region data for *Caulolatilus microps*

Sample	*n*	*K*	*À*	*H*	*F*	*P*
All combined	188	3.099 ± 1.618	0.008 ± 0.004	0.943 ± 0.013	–26.24	0.000
New York	24	3.409 ± 1.807	0.008 ± 0.005	0.967 ± 0.024	–13.112	0.000
Delaware	8	3.607 ± 2.043	0.009 ± 0.006	1.000 ± 0.062	–4.958	0.003
Virginia	40	2.931 ± 1.569	0.007 ± 0.004	0.894 ± 0.04	–19.816	0.000
North Carolina north	24	2.754 ± 1.511	0.007 ± 0.004	0.967 ± 0.024	–15.451	0.000
North Carolina south	24	2.442 ± 1.370	0.006 ± 0.004	0.891 ± 0.057	–10.239	0.000
South Carolina	31	3.914 ± 2.017	0.010 ± 0.006	0.978 ± 0.015	–20.286	0.000
Florida	24	2.953 ± 1.601	0.007 ± 0.004	0.956 ± 0.031	–14.667	0.000
Western Florida	13	2.692 ± 1.530	0.007 ± 0.004	0.923 ± 0.050	–2.577	0.052

n, Number of samples; K, mean number of pairwise differences; *À*, nucleotide diversity; *H*, haplotype diversity; *F*, Fu's *F*; ; *P*, Probability of significance for Fu's *D* (all probabilities are based on 10 000 permutations of the data).

Table 3. Mean genetic diversity parameters over loci for the 8 geographic samples of *Caulolatilus microps*

Pop	***n***	N_A	A_E	A_{RR}	A_{RE}	H_O	H_E	u H_E	F_{IS}
New York	80	9.65	4.50	6.76	9.03	0.748	0.739	0.744	–0.015
Delaware	45	8.65	4.41	6.68	9.03	0.751	0.734	0.743	–0.025
Virginia	109	10.17	4.40	6.67	8.96	0.738	0.738	0.741	–0.001
North Carolina north	56	8.96	4.32	6.57	8.78	0.727	0.728	0.735	0.003
North Carolina south	67	9.57	4.39	6.69	9.06	0.739	0.735	0.740	–0.010
South Carolina	76	9.39	4.43	6.62	8.83	0.740	0.738	0.743	–0.004
Florida	55	8.74	4.39	6.56	8.67	0.730	0.731	0.737	0.004
Western Florida	15	6.96	4.43	6.83	9.03	0.755	0.721	0.746	–0.051
Total	503	9.01	4.41	6.65	8.93	0.741	0.733	0.741	–0.012

n, Sample Size; N_A, number of alleles; A_E, number of effective alleles; A_{RR}, per locus rarefaction allelic richness; A_{RE}, per locus extrapolated allelic richness; H_O, observed heterozygosity; H_E), expected heterozygosity; uH_E, unbiased expected heterozygosity; F_{IS}, fixation index. Bolded values are out of HWE ($P < 0.001$). Detailed list of locus names is at the bottom of the table.

TABLE 4 Population pairwise F_{ST} values based 23 microsatellite loci (below the daigonal) for (*Caulolatilus microps*) and population pairwise $_{/}{}'_{ST}$ values based on the mitochondrial control (above the diagonal). There were no significant pairwise comparisons based on either class of molecular marker ($P < 0.05$)

	NY	**DE**	**VA**	**NCN**	**NCS**	**SC**	**FL**	**WF**
New York (NY)	*	–0.0014	–0.0005	0.0139	0.0011	–0.0070	–0.0079	–0.0112
Delaware (DE)	–0.0015	*	0.0020	–0.0028	–0.0021	–0.0260	–0.0091	0.0392
Virginia (VA)	–0.0004	–0.0022	*	0.0104	–0.0150	–0.0051	–0.0111	–0.0131
North Carolina north (NCN)	–0.0006	–0.0011	–0.0002	*	–0.0029	–0.0073	–0.0051	0.0163
North Carolina south (NCS)	–0.0001	–0.0020	0.0008	0.0007	*	–0.0040	–0.0044	–0.0045
South Carolina (SC)	–0.0007	–0.0017	–0.0008	0.0004	–0.0009	*	–0.0132	0.00110
Florida (FL)	–0.0007	–0.0022	–0.0006	–0.0008	–0.0005	–0.0000	*	–0.0161
Western Florida (WFL)	–0.0007	–0.0028	0.0004	0.0001	0.0033	–0.0006	0.00023	*