Conserv Genet
DOI 10.1007/s10592-014-0563-6

RESEARCH ARTICLE

Regional differentiation among populations of the Diamondback terrapin (*Malaclemys terrapin*)

Kristen M. Hart · Margaret E. Hunter · Tim L. King

Received: 14 August 2013 / Accepted: 7 January 2014

Abstract The Diamondback terrapin *(Malaclemys terrapin*) is a brackish-water turtle species whose populations have been fragmented due to anthropogenic activity such as development of coastal habitat and entrapment in commercial blue crab (*Callinectes sapidus*) fishing gear. Genetic analyses can improve conservation efforts for the long-term protection of the species. We used microsatellite DNA analysis to investigate levels of gene flow among and genetic variability within 21 geographically separate collections of the species distributed from Massachusetts to Texas. Quantified levels of genetic variability (allelic diversity, genotypic frequencies, and heterozygosity) revealed three zones of genetic discontinuity, resulting in four discrete populations: Northeast Atlantic, Coastal Mid-Atlantic, Florida and Texas/Louisiana. The average number of alleles and expected heterozygosity for the four genetic clusters were $N_A = 6.54$ and $H_E = 0.050$, respectively. However, the geographic boundaries of the populations did not correspond to accepted terrapin subspecies limits. Our results illuminate not only the need to sample terrapins in additional sites, specifically in the southeast, but also the necessity for allowing uninterrupted gene flow among population groupings to preserve current levels of genetic diversity.

Keywords Diamondback terrapin · *Malaclemys* · Population genetics · Microsatellite · Management

Introduction

The Diamondback terrapin (*Malaclemys terrapin*) inhabits brackish waters within coastal salt marsh and mangrove habitat along the Atlantic and Gulf Coasts of the United States. The range of the terrapin extends from Cape Cod, Massachusetts (MA) to southern Texas (TX), covering ~5,000 km of shoreline (Fig. 1). Ecological studies of this sexually-dimorphic turtle species have shown that philopatry is common in terrapins throughout their range (Auger 1989; Gibbons et al. 2001; Tucker et al. 2001; Hart 2005), with females showing nesting site-fidelity over many seasons (Auger 1989; Roosenburg 1996; Mitro 2003). Current anthropogenic disturbances to terrapins and their preferred coastal habitat predispose them to population fragmentation. Threats to these aquatic turtles include bycatch in traps for blue crabs (*Callinectes sapidus*) (Davis 1942; Bishop 1983; Roosenburg et al. 1997; Dorcas et al. 2007; Hart and McIvor 2008), direct harvest (Russell Burke, pers. comm.), roadkill (Wood and Herlands 1997), and reduced access to suitable coastal habitat.

Terrapin by-catch in crab pots constitutes the major threat for the species (Seigel and Gibbons 1995; Gibbons et al. 2001; Butler et al. 2006; Dorcas et al. 2007; Hart and Crowder 2011). The mortality rate for diamondback terrapins in crab pots ranges from 10–78 % (Bishop 1983;

Electronic supplementary material The online version of this article (doi:10.1007/s10592-014-0563-6) contains supplementary material, which is available to authorized users.

K. M. Hart (✉)
Southeast Ecological Science Center, US Geological Survey, 3205 College Avenue, Davie, FL 33314, USA
e-mail: kristen_hart@usgs.gov

M. E. Hunter
Southeast Ecological Science Center, US Geological Survey, 7920 NW 71 St, Gainesville, FL 32653, USA

T. L. King
Leetown Science Center, US Geological Survey, 11649 Leetown Road, Kearneysville, WV 25430, USA

Published online: 23 January 2014

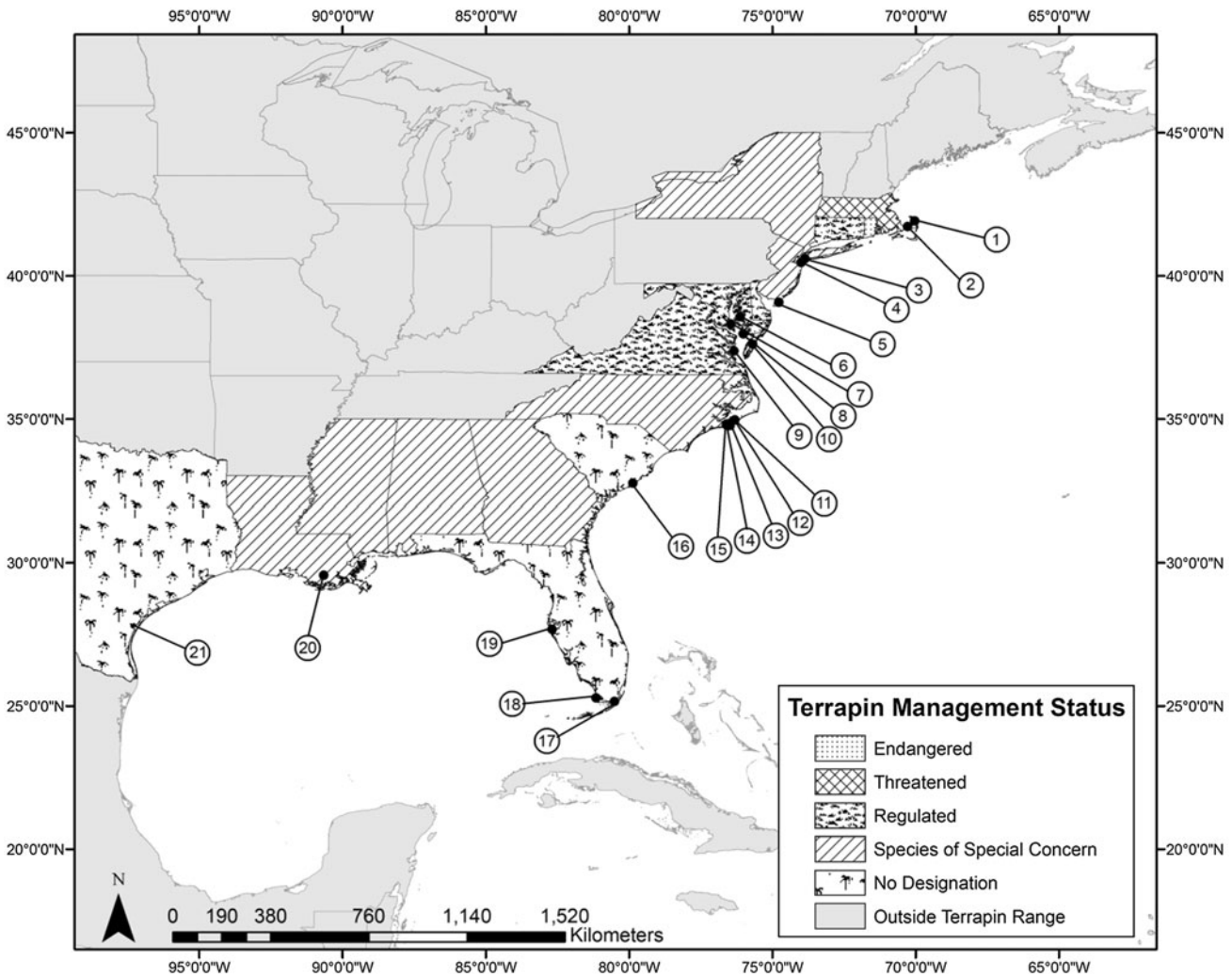

Fig. 1 Diamondback terrapin (*Malaclemys terrapin*) management status by state across the species' Atlantic and Gulf coastal range (see Table 1 for site names for locations 1–21)

Roosenburg et al. 1997; Roosenburg 2004), depending on the time of year and body size of individuals. Crab pots set near resident terrapin populations may directly threaten the viability of the population. Considerable mortality may also stem from terrapin capture in "ghost" or abandoned pots, which are common in this fishery. Grosse et al. (2009) recently reported 94 dead terrapins in one ghost crab pot in Georgia (GA), and earlier Bishop (1983) reported 29 terrapin carcasses in one ghost pot in South Carolina (SC), highlighting the potential impacts of such fishing gear on by-caught species.

Recently, many state-level management agencies within the range of the terrapin have designated them as a species of special concern (Fig. 1), and agencies in MA and Rhode Island (RI) have gone further to declare their populations to be threatened and endangered, respectively; however, overall population estimates and trends are lacking for the species. New Jersey (NJ) and Maryland (MD) now require certain bycatch reduction devices (BRD) on crab traps (Roosenburg et al. 2003 and references therein) to mitigate incidental terrapin mortality in this particular fishing gear. However, other populations of terrapins are likely experiencing similar threats and may be in need of additional protections.

For over 60 years, seven terrapin subspecies have been recognized (Carr 1952; Ernst et al. 1994), primarily based on phenotypic differences. Subspecies divisions occur at the northern portion of North Carolina (NC), the GA/Florida (FL) border, south Florida (with breaks at FL Bay and western Everglades), and the Alabama (AL)/FL border (Fig. 2). Although abundant biological and ecological data for terrapins from particular locations have been collected over the past two decades (Gibbons et al. 2001; Dorcas et al. 2007; Hart and McIvor 2008), this information has not been synthesized across the geographic range to illuminate effective management and conservation plans. To plan and implement biologically sound conservation

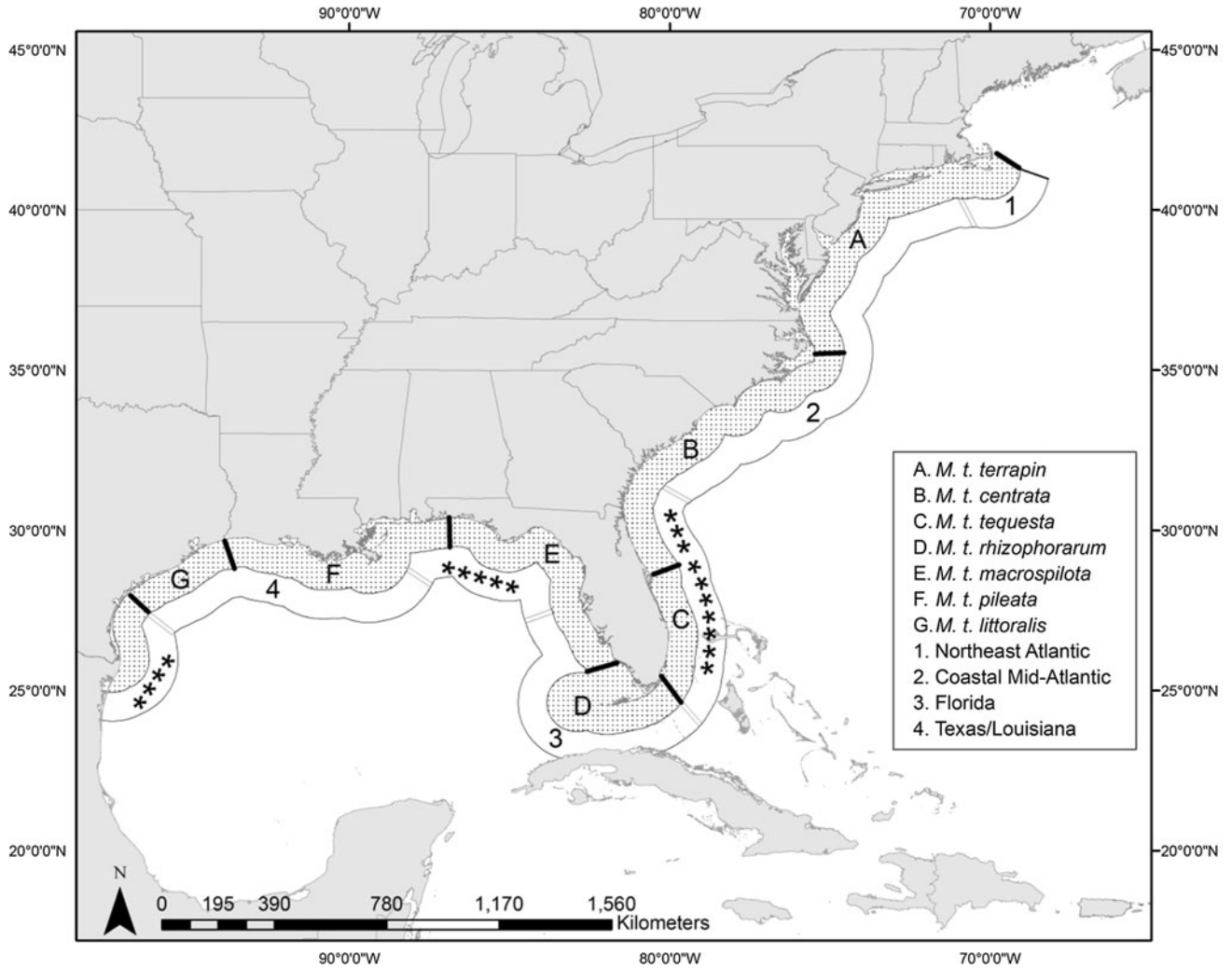

Fig. 2 Diamondback terrapin (*Malaclemys terrapin*) subspecies (A–G *dotted area*, see Ernst et al. 1994) designations and new regional genetic groupings (1–4; *this study*). A "*" indicates locations with no samples

programs for this aquatic turtle, an understanding of the patterns of genetic diversity and the evolutionary relationships among geographical populations is essential.

Clear genetic population definitions for terrapins have been lacking. Previous attempts at genetic stock identification in terrapins (Lamb and Avise 1992) indicated an Atlantic/Gulf disjunction in mtDNA phylogeny. Two more recent regional genetic studies using microsatellites found little genetic differentiation among terrapins at small geographic scales (Hauswaldt and Glen 2005 in SC; Sheridan et al. 2010 in NJ). King and Julian (2004) tested microsatellite markers developed from the bog turtle (*Glyptemys muhlenbergii*) for their potential to amplify across other Emydid turtle species; 23 of 27 bog turtle markers successfully amplified in terrapins, with 16 showing moderate to high polymorphism.

Here we evaluate the population genetic structure and delineate discrete genetic clusters for Diamondback terrapins across its range using 12 moderately to highly polymorphic loci (King and Julian 2004). In addition, we test whether genetic groups matched current morphologically-determined subspecies designations.

Methods

Sample collections and DNA isolation

Turtles were sampled from 21 collection sites for a total of 997 animals from 10 states (Table 1; Fig. 1). In addition to our own collection efforts, terrapin samples were provided by volunteers in various states throughout the species' range. Blood samples were preserved on FTA cards (Whatman, Inc., Clifton, NJ, USA), a cellulose-based absorptive paper, and stored at room-temperature. DNA was isolated using Puregene DNA extraction kits (Gentra

Table 1 General collection localities, abbreviations, and sample size for diamondback terrapins sampled from MA to TX. To eliminate biases from small sample sizes, we only used collections with sample sizes ≥9 individuals

Site #	Site abbreviation	General collection Site	Sample size	N_A	E_A	I	H_O	H_E
1	MA1	Wellfleet, MA	58	3.833	1.895	0.729	0.400	0.407
2	MA2	Sandy Neck, Barnstable, MA	19	3.583	2.282	0.865	0.465	0.481
Genetic cluster			NE ATL	4.250	2.032	0.798	0.416	0.439
3	NY	Jamaica Bay, Hudson River, NY	31	6.500	4.042	1.368	0.610	0.630
4	NJ1	Sandy Hook, NJ	20	6.667	4.058	1.381	0.612	0.633
5	NJ2	Cape May County, NJ	29	6.500	4.024	1.383	0.648	0.646
6	MD1	Marshy Creek, Kent Island, MD	64	7.583	4.672	1.495	0.666	0.669
7	MD2	Patuxent River, MD	63	7.250	4.271	1.421	0.663	0.644
8	MD3	Tylerton, Smith Island, MD	64	7.833	4.584	1.494	0.659	0.666
9	VA1	Mobjack Bay, VA	45	6.667	4.378	1.439	0.661	0.670
10	VA2	Wachapreague, VA	38	6.833	4.269	1.440	0.682	0.660
11	NC1	Cedar Island, NC	44	7.500	4.626	1.517	0.715	0.688
12	NC2	Davis, NC	120	8.083	4.567	1.504	0.657	0.678
13	NC3	Davis Island, NC	16	5.917	4.117	1.414	0.708	0.675
14	NC4	Jarrett Bay, NC	42	7.083	4.888	1.510	0.691	0.686
15	NC5	North River, NC	42	7.917	4.922	1.559	0.658	0.693
16	SC	Charleston Harbor, SC	48	8.333	5.227	1.602	0.686	0.697
Genetic cluster			CMA	9.667	4.958	1.562	0.666	0.687
17	FL1	Nest Key, FL	13	3.250	2.140	0.648	0.322	0.320
18	FL2	Big Sable Creek, Everglades, FL (Sites A,B,C)	187	5.750	2.536	0.830	0.361	0.384
19	FL3	Tampa/St. Petersburg, FL	9	4.000	2.737	0.925	0.472	0.451
Genetic cluster			FL	6.500	2.572	0.858	0.363	0.388
20	LA	Houma, LA	31	5.417	2.871	1.042	0.472	0.495
21	TX	Nueces Bay, TX	14	3.750	2.526	0.909	0.438	0.462
Genetic cluster			TX/LA	5.750	2.848	1.051	0.459	0.495
		Total	997					
			Average	6.202	3.792	1.261	0.583	0.587
			All samples	10.417	5.030	1.587	0.573	0.693
			All clusters	6.542	3.103	1.067	0.476	0.502

Characterization of the 12 microsatellite loci implemented on terrapins collected from 21 geographic areas (above) and the four genetically differentiated regional groups *NE ATL* Northeast Atlantic (Massachusetts), *CMA* Coastal Mid-Atlantic, *FL* Florida, and *TX/LA* Texas and Louisiana. The average values are provided for the 21 populations and four genetic clusters, along with the averages over all samples. Abbreviations are as follows: N_A Average number of alleles per locus, and E_A average effective number of alleles, *I* information index, and H_O observed and H_E expected heterozygosity. Site No. matches that in Fig. 1 and 2

Systems, Inc., Minneapolis, MN, USA) and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

Microsatellite analysis

Twelve polymorphic microsatellite DNA loci developed for bog turtles (*Glyptemys muhlenbergi*) by King and Julian (2004) were screened in all turtles collected. Polymerase chain reaction (PCR) conditions were optimized with genomic DNA from eight terrapins collected from MA to NC. The 12 loci amplified in multiplexed PCR reactions were: *Gmu*A18, *Gmu*B08, *Gmu*B67, *Gmu*B91, *Gmu*D21, *Gmu*D55, *Gmu*D62, *Gmu*D87, *Gmu*D90, *Gmu*D93, *Gmu*D114, and *Gmu*D121 (Supplemental Table 1; King and Julian (2004).

PCR reactions

Each PCR consisted of 100–200 ng of genomic DNA, 0.875× PCR buffer (59mMTris-HCl, pH 8.3; 15 mM$(NH_4)_2SO_4$; 9 mM ß-mercaptoethanol; 6mMEDTA), 3.75 mM $MgCl_2$, 0.31 mM dNTPs, 0.15–0.25 µM of forward and reverse primers, and 0.4 U AmpliTaq, brought up to a total volume of 20 µl with deionized water. Each forward primer was 5′ modified with FAM, NED, or HEX fluorescent labels (Applied Biosystems). The following amplification

conditions were used: 94 °C for 2 min.; 35 cycles of 94 °C denaturation for 45 s, 56 °C annealing for 45 s, 72 °C extension for 2 min.; final extension of 72 °C for 10 min. Thermal cycling was performed in an MJ DNA Engine PTC 200 (MJ Research, Watertown, MA).

Fragment analysis and allelic designations

Fragment analysis and allelic designations followed techniques described in King et al. (2006). Capillary electrophoresis was conducted on an ABI Prism^{TM} 3100 Genetic Analyzer using GeneScan-500 ROX size standard (Applied Biosystems). Fragment size data was generated using GeneScan software version 3.7 (Applied Biosystems). Genotyper software version 3.6 (Applied Biosystems) was used to score, bin, and assign genotypes for each individual.

Statistical analyses

Our objectives were to identify groups of populations isolated by zones of genetic discontinuity.

Observed genotype frequencies were tested for consistency with Hardy-Weinberg equilibrium using GenAlEx 6.5 (Peakall and Smouse 2006). Linkage equilibrium expectations with randomization tests was implemented by GENEPOP 4.0 using the randomization method of Raymond and Rousset (1995) for all pairs of loci within collections to test for the presence of admixture. Sequential Bonferroni adjustments (Rice 1989) were used to determine significance for these and all other multiple tests. Genecap (Wilberg and Dreher 2004) calculated the unbiased probability of identity ($P_{(ID)}$), which is the probability that two individuals drawn at random from a population will have the same genotype at the assessed loci (Paetkau and Strobeck 1994) and a related more conservative statistic for calculating $P_{(ID)}$ among siblings ($P_{(ID)sib}$; Evett and Weir 1998). The program additionally searched for duplicate genotypes. We used Bottleneck 1.2.02 to evaluate heterozygote excess under the sign test and mutation-drift equilibrium under the allele frequency distribution test (Piry et al. 1999).

The genetic diversity was estimated by the observed (H_O) and expected heterozygosity (H_E), information index (I), average number of alleles per locus (N_A), and average effective number of alleles (E_A) using GenAlEx 6.5 (Table 1). The software LDNe (Waples and Do 2008) was used to estimate effective population sizes (N_E) for the four genetic groupings with 95 % confidence intervals (CI) following the bias-corrected method of Waples (2006). The single point estimate method removes the downward bias associated with the true N_E being greater than the sample size used to estimate it (Waples 2006). Micro-Checker (Van Oosterhout et al. 2004) was used to identify loci with evidence of null alleles.

To assess overall genetic differentiation at the population level, we used GenAlEx 6.5 to calculate F_{ST} via Analysis of molecular variance (AMOVA). Results were combined over loci using Fisher's method (Sokal and Rohlf 1994). Differences between each collection pair were summarized by pairwise genetic distance using the chord distance (D_c) of Cavalli-Sforza and Edwards (1967). An unrooted phylogenetic tree was fit to the D_c matrix using the neighbor-joining (NJ) algorithm with 1,000 randomizations, and TreeView (Page 1996) was used to visualize the tree. The statistical significance of the correlation between genetic and geographic distance matrices, or isolation by distance, was assessed with a Mantel randomization test performed with GenAlEx 6.5 with 999 permutations comparing pairwise genetic distance and marsh distance (in meters) (Peakall and Smouse 2006). An AMOVA was used to partition genetic variation within and among terrapin collections and individuals with 999 permutations (Peakall and Smouse 2006).

Several techniques were used to describe genetic relationships among the geographic populations sampled throughout the species' range. The program Structure 2.3.3 (Pritchard et al. 2000) was used to delineate zones of genetic discontinuity isolating regional groupings of populations. We used the model-based clustering method of Structure to infer population structure among collections and probabilistically assign all individuals to distinct clusters (k). For the 21 collections, $k = 1$ to $k = 10$ clusters were considered using a burn-in of 10,000 followed by 50,000 iterations and 20 independent runs for each k using the admixture model. Because complex migration and gene flow patterns likely exist among disjunct genetic groups, a sequential method of inferring k was used; first, the uppermost hierarchical level (i.e., $n = 21$ collections) of population structure was identified at $k = 2$. After the initial run, we performed a subsequent analysis of each cluster to identify within-subset structure using the Locprior model (Hubisz et al. 2009). Each cluster was analyzed separately, for $k = 1$ to $k = 10$ using a burn-in of 10,000 followed by 50,000 iterations and 20 independent runs for each k using the admixture model. The Locprior setting is not biased towards detecting structure when it is not present (Hubisz et al. 2009). As described in Evanno et al. (2005), the optimal number of clusters was identified using Δk. For all runs, individual assignment success to the cluster of origin was recorded as the highest likelihood of assignment (q) and the percentage of individuals in a cluster with $q \geq 0.90$ (Pritchard et al. 2000).

Results

We determined genotypes at 12 microsatellite DNA loci for terrapins sampled from 21 locations (see Table 1 for

listings and abbreviations; Fig. 1) encompassing the geographic range of the species; all 997 individuals had unique multilocus genotypes. Randomization tests showed that genotypes for most collections surveyed were consistent with Hardy-Weinberg expectations. A total of 252 locus-by-collection comparisons revealed that eight (3.2 %) comparisons deviated from expectations after Bonferroni adjustment for multiple tests at $\alpha = 0.05$ ($p < 0.001$). These were NC1 at *Gmu*DD87, NC2 at *Gmu*D93, LA at *Gmu*D93, NJ1 at *Gmu*D93, FL2 at *Gmu*D21, *Gmu*D62, *Gmu*B67, and *Gmu*D114. No linkage disequilibrium was detected in 1,386 comparisons ($p < 0.00009$). The loci produced an unbiased $P_{(ID)}$ estimate of 3.0E − 13 and a $P_{(ID)sib}$ estimate of 3.3E − 05, suggesting that unique individuals can be confidently identified across the region (Supplemental Table 2).

Bottleneck signatures were found for 21 populations under the infinite alleles model (IAM, $p \leq 0.04$), three populations under the stepwise mutation model (SMM, $p \leq 0.03$), and three populations under the two-phase model (TPM, $p \leq 0.04$; Supplemental Table 3). The allele frequency distribution test remained in a normal L-shape distribution for all but two populations, perhaps due to a bottleneck occurring too recently to detect (Cristescu et al. 2010). Shifted mode distributions occurred in TX and FL3; however the FL3 sample size was too small for accurate results.

We observed 125 alleles across the 12 loci, ranging from three alleles at *Gmu*D21 to 17 alleles at *Gmu*D87 (Table 1). The mean number of alleles per locus was lowest at the extremes of the species' range in MA2 (3.6) and FL1 (3.3), and greatest in the centrally located collections from NC2 (8.1) and SC (8.3; Table 1). Observed mean heterozygosity ranged from 32.2 % in terrapins from FL1 to 71.5 % in terrapins from NC1 (Table 1). Similar to the mean number of alleles, observed heterozygosity was lowest in collections in the states at the extremes of the range (MA, FL, LA, and TX; 0.32–0.47), and relatively uniform (61–72 %) in the remainder of collections (Table 1). The N_E for the 21 populations is provided on Supplemental Table 4. Negative values indicate infinity for either the point estimate or the confidence limits. Estimates of infinity are returned when the signal in the genetic data can be attributed entirely to sampling error, rather than genetic drift, which is the case for a very large population or when the population sample contains too little information (Waples and Do 2008). None of the 12 loci was found to have evidence of null alleles due to a heterozygote deficiency.

Microsatellite analysis of F_{ST} at 12 loci identified a wide range of values, 0.001 (VA1 and NC3; $p \geq 0.393$) to 0.491 (MA2 and FL2; $p \geq 0.001$), and revealed the presence of measurable genetic structure in many pairwise comparisons between groups (Supplemental Table 5, below diagonal). The average F_{ST} was 0.161 across all 21 collections, which reflects great genetic differentiation according to Wright (1978). F_{ST} values within coastal Carolina ranged from 0.002–0.010 (average = 0.005), indicating little genetic differentiation between subpopulations within that region (Supplemental Table 5). Similarly, F_{ST} values within FL ranged from 0.037–0.079 (average = 0.059), again indicating little genetic differentiation among sites within this metapopulation. The highest F_{ST} values were observed between sites and delineated groupings at the opposite extremes of the range.

Analysis of chord distance values (D_C), between all collections was consistent with the F_{ST} values (Supplemental Table 5, above diagonal) and revealed that the greatest genetic distance was between the two collections representing the extremes of the Atlantic range MA2 and FL1 (0.267). The lowest genetic distances were observed between NC2 and NC5 (0.008). The unrooted Neighbor-Joining phenogram depicts the underlying structure of the chord distance matrix (Fig. 3). This graphical depiction clearly illustrates the high degree of regional genetic differentiation among the terrapin collections. A Mantel test comparing pairwise genetic distance and marsh distance matrices among terrapin collections identified a positive, statistically significant correlation ($r = 0.632$; $p < 0.001$). This finding conforms to the spatial pattern expected under an isolation by distance model of population structure. Grouping samples into 21 populations showed that 16.1 % of genetic variation was distributed among the populations and 84.9 % was within populations (Table 2a).

Regional analysis

Among the 997 individuals, a total of four regional genetic clusters were identified in three sequential hierarchical STRUCTURE analyses; Northeast Atlantic (NE ATL; two locations in MA), Coastal Mid-Atlantic (CMA; containing collections from six states, NY to SC), FL, and TX/LA. At the uppermost hierarchical level, two clusters were identified with STRUCTURE (Fig. 4; Supplemental Fig. 1). Cluster A was composed of seven collection states, MA-SC (Table 1) with $q = 98.6$ % correct assignment of each terrapin to its cluster of origin based on q values. Cluster B, had 98.2 % correct assignment to FL, and TX/LA samples were admixed between the two clusters (CMA, $q = 42.5$ %, FL, $q = 57.5$ %). To determine the degree of admixture the TX/LA samples were analyzed with both cluster A and B. The cluster A analysis resulted in $k = 3$ subclusters (Fig. 4). Subcluster 1 had 98.2 % correct assignment to CMA, subcluster 2 had correct assignment 95.1 % to NE ATL, and subcluster 3 had 98.3 % correct assignment to TX/LA. The cluster B analysis resulted in $k = 2$ subclusters corresponding to subclusters 3 and 4 (Fig. 4). Again subcluster 3,

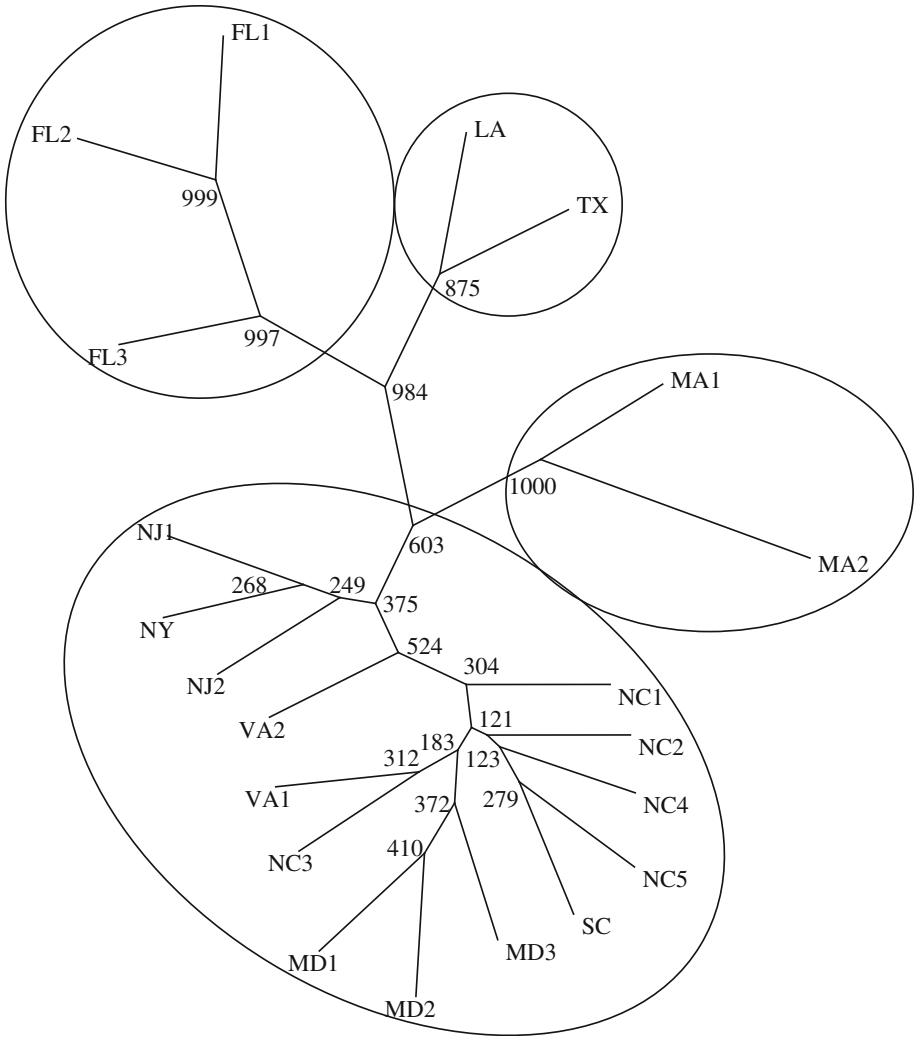

Fig. 3 Unrooted neighbor-joining phenogram depicting genetic (chord) distance (D_c) among all 22 *M. terrapin* collections. Boot strap estimates after 1,000 randomizations are denoted at nodes. Regional clades are grouped by ellipses, inserted after visual inspection of phenogram distances

TX/LA had high correct assignment values ($q = 99.1$ %) and subcluster 4 had 97.7 % correct assignment to FL. Of the FL samples collected on the Gulf Coast, $n = 3$ had partial assignment to TX/LA ($q = 50.7$, 88.1 and 89.5 %), suggesting a relationship to that region or similar allele frequencies.

When the four regional clusters were tested in 98 comparisons for Hardy-Weinberg equilibrium, seven comparisons were significant; these were TX/LA at *Gmu*D62 and *Gmu*D93 and FL at *Gmu*D21, *Gmu*D62, *Gmu*B67, *Gmu*D90, and *Gmu*D114. Linkage disequilibrium occurred in the four regional clusters in six out of 264 comparisons (2.27 %), after correction for multiple tests (overall $\alpha = 0.05$, $p < 0.0009$).

The unbiased $P_{(ID)}$ and $P_{(ID)sib}$ estimate for the four regional clusters suggest that unique individuals can be confidently identified across the region (Supplemental Table 2). Bottleneck signatures were found for FL under the SMM ($p \leq 0.0001$) and TPM ($p \leq 0.0100$), CMA under the IAM ($p \leq 0.0100$) and TX/LA under the SMM ($p \leq 0.0010$). The allele frequency distribution test remained in a normal L-shape distribution, perhaps due to the bottleneck occurring too recently to detect (Cristescu et al. 2010).

The average number of alleles and expected heterozygosity for the four genetic clusters were $N_A = 6.54$ and $H_E = 0.050$, respectively. Values for N_E were CMA $N_E = 72.4$ (CI 44.7–141.6); NE ATL $N_E = 1{,}809.1$ (CI 1,191–3,468.3); TX/LA $N_E = 324.3$ (CI 99.9–∞); FL $N_E = 482.5$ (CI 255.7–1,997.7) for allele frequencies $\geq$0.02 (see Supplemental Table 4). Cluster CMA had the highest heterozygosity and number of alleles, whereas cluster FL had the lowest heterozygosity and cluster NE ATL had the lowest number of alleles (Table 2).

The F_{ST} values for the four regional clusters were significant and ranged from 0.151 (CMA and NE ATL) ($p \geq 0.001$) to 0.463 (FL and NE ATL) ($p \geq 0.001$). The F_{ST} values for TX/LA were 0.165 with CMA, 0.215 with FL and 0.307 with NE ATL. The FL and CMA F_{ST} value was 0.254. The average F_{ST} across four regional clusters

Table 2 Results of analysis of molecular variance

	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)
2a.					
	Among populations	20	1,329.619	0.687	16
	Within populations	1,973	7,037.990	3.567	84
	Total	1,993	8,367.609	4.254	100
2b.					
	Among regions	3	1,147.418	1.097	23
	Among populations	17	182.201	0.088	2
	Within populations	1,973	7,037.990	3.567	75
	Total	1,993	8,367.609	4.752	100
2c.					
	Among subspecies	5	1,016.230	0.662	15
	Among populations	15	313.389	0.201	5
	Within populations	1,973	7,037.990	3.567	81
	Total	1,993	8,367.609	4.430	100

(a) Among the 21 populations and among individuals within the 21 populations, (b) among the four regional clusters, among the 21 populations within the four regional clusters, and among individuals within the 21 populations and, (c) among the seven putative subspecies designations, among the 21 populations within the seven subspecies, and among individuals within the 21 populations (Excoffier et al. 1992)

was 0.238. Overall, low genetic distances were observed between populations within the four regional groupings (Supplemental Table 5).

Quantitative estimates of molecular variance indicated significant genetic population structure, with the greatest amount due to variation among populations. When collections were grouped into the four regional clusters by the hierarchical STRUCTURE analysis, 23.09 % of genetic variation was distributed among regions (Table 2b). The high level of variation observed among populations supports the appropriateness of these regional groupings. Less differentiation was found between the previously proposed subspecies designations (Table 2c). Additionally, no support was provided for the subspecies designations by the hierarchical STRUCTURE analysis.

Discussion

We detected Diamondback terrapin genetic differentiation and population structure at all hierarchical levels. However, this structuring was weak at the local level, a result that supports work by Hauswaldt and Glen (2005) and Sheridan et al. (2010). Our range-wide analyses with multiple collections of terrapins represents the largest genetic dataset for the species, and provides further delineation of terrapin populations than that provided by Lamb and Avise (1992); Atlantic and Gulf populations are more structured than these previous findings suggest.

Although seven subspecies have been recognized for terrapins for over 60 years, our data support four groupings of populations, delineated by three zones of genetic discontinuity. Further, the levels of differentiation we observed are similar to those described as Management Units (MUs; Palsbøll et al. 2006) in other species. Our four regional genetic groupings (NE ATL (Massachusetts), CMA (NY to SC), FL, and TX/LA) in some cases encompass geographic regions different from previous subspecies ranges. The NE ATL grouping is geographically smaller than the extent of the *M. t. terrapin* range. The CMA grouping is larger than the extent of the *M. t. centrata* range. The FL grouping extends around the southwest coast of the state so is larger than previous *M. t. rhizophorarum* range (note no samples analyzed from *M. t.*

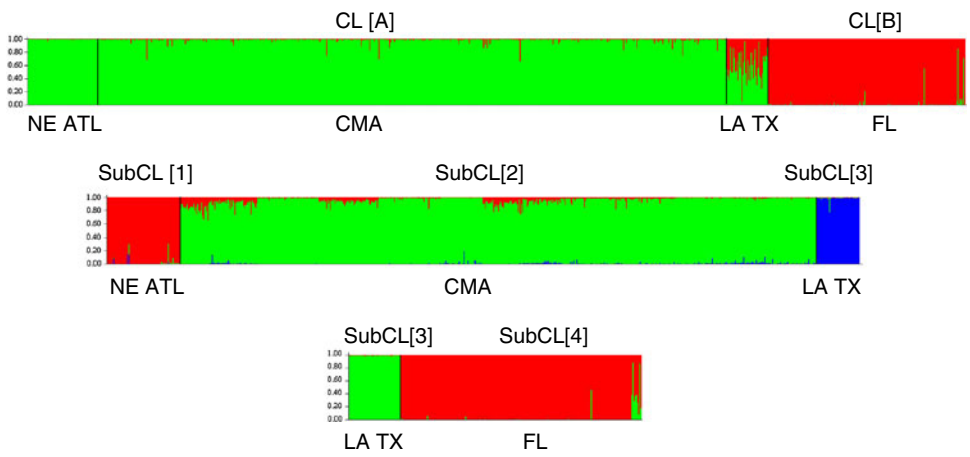

Fig. 4 Results of Bayesian clustering analysis using STRUCTURE. The proportion of membership for $k = 2$ in each plot. Individual clusters were run independently in hierarchical analyses resulting in a series of Subclusters (SubCL). *NE ATL* Northeast Atlantic (Massachusetts), *CMA* Coastal Mid-Atlantic (NY to SC), *FL* Florida, *TX/LA* Texas and Louisiana

tequesta). While the TX/LA grouping includes portions of both *M. t. pileata* and *M. t. littoralis* ranges, no samples have been included from northwest Florida [Gulf coast]) or southern TX. In summary, the data we have presented here do not support the subspecies definitions.

The microsatellite loci screened in this study are robust for surveys of neutral genetic variation among populations of terrapins. In this survey of terrapins collected throughout their range, we identified 125 alleles at 12 loci, comparable to other studies of terrapin genetics (Hauswaldt and Glen 2003, 2005; Sheridan et al. 2010). Overall levels of genetic diversity and variation were sufficiently high to detect phylogeographic structuring at the range-wide and regional level with limited population structuring at the local level. The results were consistent across multiple measures of population differentiation, including heterozygosity values, pairwise genetic distances (chord, D_C), F_{ST} (and associated estimates of gene exchange), Bayesian clustering analyses, and maximum likelihood assignment tests.

Evidence of bottlenecks was detected in each regional grouping with the exception of NE ATL. The bottleneck identified in TX/LA was under the SMM, FL was under the SMM and TPM, and the CMA bottleneck fell under the IAM. Typically, microsatellites are thought to follow SMM (Piry et al. 1999) or TPM (Di Rienzo et al. 1994), however mutation patterns in a number of species better fit the IAM (Cristescu et al. 2010). The detection of bottlenecks in each of the four regional groupings suggests that there has been a severe population contraction (100-fold) over the past several generations (Luikart and Cornuet 1998).

The calculated effective population sizes (N_e) were relatively low for the large regional groupings, especially if local populations (e.g., in CMA) become isolated or fragmented. It has been suggested that at a minimum 50 genetically effective breeders ($\sim$500 individuals) are needed to prevent inbreeding depression for short-term population survival (Wright 1951; Frankham et al. 2002). Yet, other research has also suggested that population levels in the upper hundreds to thousands are needed to maintain evolutionary potential (Traill et al. 2010). The populations with large or infinite N_E confidence intervals may need additional data to provide accurate estimates.

In general, any measurable genetic structure will reflect the balance between genetic drift and selection, which will function to create differences between populations, or homogenize the gene pool through migration (Mockford et al. 2005). The effects of genetic drift are strongest in small populations—the fewer individuals in the population, the more genetic drift affect the population in terms of loss of genetic variation. This occurs because over time, at random, there will be a generation in which certain rare alleles by chance will not get passed on to the next generation. Because terrapins exist in what seems to be small, patchy subpopulations within regional metapopulation structures, effects of size on smaller subpopulations need to be accounted for in future analyses of overall population persistence (Hart 2005). Since genetic drift tends to decrease the potential for adaptation, terrapin subpopulations with small or declining numbers may be more likely to go extinct or to be less resilient to environmental change and further population subdivision. However, even low levels of gene flow can help to counteract drift. Our analysis revealed measurable gene flow (as reflected in low F_{ST} values) among terrapin subpopulations that could counteract problems associated with drift (i.e. inbreeding depression), at least for the near term. Nonetheless, if anthropogenic disturbances increase habitat and population fragmentation, sufficient gene flow may not be prevalent in the future.

Contemporary population fragmentation poses challenges to our understanding of how wildlife populations function and may have been structured in the past prior to anthropogenic impacts. Direct threats as well as landscape-level changes in climate and environmental drivers continue to affect terrapin populations throughout their coastal range. Effective gene flow across sites that are separated by distances further than individual terrapins would be expected to travel (i.e., kms) could be hampered if dispersing and mating groups of terrapins become further disconnected due to changes in coastal land use or morphology. Long-term landscape planning to ensure habitat connectivity could encourage movement within and among regions to allow for the low levels of diversity to increase through admixture.

Conservation implications

Our study is the first to identify biologically meaningful units of management for terrapins and this new knowledge could be incorporated into management plans to interpret trends in data collected from terrapins at individual sites. Terrapins exist as at least four distinct regional groups throughout their range. The observed levels of differentiation among the four regional groups has been previously used to delineate units of management in other species (see Waples et al. 2008).

Although terrapin mark-recapture studies report high site fidelity, we detected a signal for significant gene flow among many sites at local and regional scales. In general, management efforts should be focused at the regional level and data from individual sites within regional groupings should be synthesized to decipher trends for the different regional groupings. Further sampling in GA, the northern part of Atlantic FL, as well as south FL and throughout the Gulf coast could help to resolve boundaries of regional groupings in the southeast US. Samples collected from

ghost crab pots could be useful for this additional genetic sampling effort.

Microsatellites are useful for several conservation problems, one of which is assignment to population of origin or matching trace samples to individual populations or collection sites (Clapham and Waerebeek 2007; Baker 2008). Previously, this technique has been used to track samples of caviar (Rejón et al. 2009), whale (Baker et al. 2007, 1996), and turtle (Roman and Bowen 2000) sold in markets to their populations of origin. Future work testing the ability of this suite of microsatellite markers to discriminate among terrapin samples collected in markets would be valuable, as live terrapins are currently sold in several US seafood markets (R. Burke, personal communication). *A priori* hypotheses regarding assignment success for those samples could rely on estimates derived in this study.

Managers are often forced to set goals based upon limited knowledge of threats, movement patterns, habitat requirements, or distribution and abundance. Although it is especially difficult to define the boundaries of aquatic animal populations, doing so is the first step in defining proper units for conservation and management. The integration of this new genetic information with previously identified microsatellite DNA variation (Hauswaldt and Glen 2005, 2003; Sheridan et al. 2010), mitochondrial DNA variation (Lamb and Avise 1992), and available ecological data will help to develop an effective management strategy for terrapins. Moreover, the proposed regional groupings for terrapins establish a meaningful scale at which to focus synthesis and conservation efforts.

Acknowledgments We thank C. Young, S. Julian, M. Eackles, B. Lubinski, and R. Johnson for assistance in the lab, and A. Sartain with GIS mapping. Many volunteers helped to collect samples for our collection. Funding was provided by the USGS, Biological Resources Division, Status and Trends program, USGS Priority Ecosystem Studies program, and Oak Foundation. All collection efforts in NC followed Duke University Institutional Animal Care and Use Committee Protocols (Protocols A120-02-01 and A120-05-04), and all research in the Florida Everglades was conducted according to guidelines in permits EVER-2001-SCI-0067 and EVER-2002-SCI-0092. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Auger PJ (1989) Sex ratio and nesting behavior in a population of *Malaclemys terrapin* displaying temperature-dependent sex-determination. Ph.D. dissertation, Tufts University

Baker CS (2008) A truer measure of the market: the molecular ecology of fisheries and wildlife trade. Mol Ecol 17:3985–3998

Baker CS, Cipriano F, Palumb SR (1996) Molecular genetic identification of whale and dolphin products from commercial markets in Korea and Japan. Mol Ecol 5:671–685

Baker CS, Cooke JG, Lavery S et al (2007) Estimating the number of whales entering trade using DNA profiling and capture-recapture analysis of market products. Mol Ecol 16:2617–2626

Bishop JM (1983) Incidental capture of diamondback terrapin by crab pots. Estuaries 6:426–430

Butler JA, Seigel RA, Mealey BK (2006) *Malaclemys terrapin*—Diamondback terrapin. In: Meylan P (ed) Biology and conservation of Florida turtles, vol 3. Chelonian Research Monographs, Meylan, pp 279–295

Carr A (1952) Handbook of turtle: the turtles of the United States, Canada, and Baja California. Cornell University Press, Ithaca, p 542

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evolution 21:550–570

Clapham P, Van Waerebeek K (2007) Bushmeat and bycatch: the sum of the parts. Mol Ecol 16:2607–2609

Cristescu R, Sherwin WB, Handasyde K, Cahill V, Cooper DW (2010) Detecting bottlenecks using BOTTLENECK 1.2.02 in wild populations: the importance of the microsatellite structure. Conserv Genet 11:1043–1049

Davis CC (1942) A study of the crab pot as a fishing gear. Publication No. 53. Chesapeake Biological Laboratory, Solomons

Di Rienzo A, Peterson AC, Garza JC et al (1994) Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166–3170

Dorcas ME, Wilson JD, Gibbons JW (2007) Crab trapping causes population decline and demographic changes in diamondback terrapin over two decades. Biol Conserv 137:334–340

Ernst CH, Lovich JE, Barbour RW (1994) Turtles of the United States and Canada. Smithsonian Institution Press, Washington, p 576

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Evett IW, Weir BS (1998) Interpreting DNA evidence: statistical genetics for forensic scientists. Sinauer Associates Inc, Sunderland

Excoffier L, Souse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotopyes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge

Gibbons JW, Lovich JE, Tucker AD, Fitzsimmons NN, Greene JL (2001) Demographic and ecological factors affecting conservation and management of diamondback terrapins (*Malaclemys terrapin*) in South Carolina. Chelonian Conserv Biol 4:66–74

Grosse AM, van Dijk JD, Holcomb KL, Maerz JC (2009) Diamondback Terrapin mortality in crab pots in a Georgia tidal marsh. Chelonian Conserv Biol 8:98–100

Hart KM (2005) Population biology of diamondback terrapins (*Malaclemys terrapin)*: Defining and reducing threats across their range. Dissertation. Duke University, Durham

Hart KM, Crowder LB (2011) Mitigating bycatch of diamondback terrapins in crab pots. J Wildl Manag 75(2):264–272. doi:10.1002/jwmg.49

Hart KM, McIvor CC (2008) Demography and ecology of mangrove diamondback terrapins in a wilderness area of Everglades National Park, Florida, USA. Copeia 2008:200–208

Hauswaldt JS, Glen TC (2003) Microsatellite DNA loci from the diamondback terrapin (*Malaclemys terrapin*). Mol Ecol News 3:174–176

Hauswaldt JS, Glen TC (2005) Population genetics of the diamondback terrapin (*Malaclemys terrapin*). Mol Ecol 14:723–732

Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332

King TL, Julian SE (2004) Conservation of microsatellite DNA flanking sequence across 13 Emydid genera assayed with novel bog turtle (*Glyptemys muhlenbergii*) loci. Conserv Genet 5(5): 719–725

King TL, Eackles MS, Spidle AP, Brockmann HJ (2006) Regional differentiation and sex-biased dispersal among populations of the horseshoe crab (*Limulus polyphemus*). Transactions of the American Fisheries Society

Lamb T, Avise JC (1992) Molecular and population genetic aspects of mitochondrial DNA variability in the diamondback terrapin, *Malaclemys terrapin*. J Hered 83:262–269

Luikart G, Cornuet JM (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv Biol 12:228–237

Mitro MG (2003) Demography and viability analysis of a diamondback terrapin population. Can J Zool 81:716–726

Mockford SW, McEachern L, Herman TB, Snyder M, Wright JM (2005) Population genetic structure of a disjunct population of Blanding's turtle (*Emydoidea blandingii*) in Nova Scotia, Canada. Biol Conserv 123:373–380

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. Mol Ecol 3:489–495

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358

Palsbøll PJ, Bérubé M, Allendorf FW (2006) Identification of management units using population genetic data. Trends Ecol Evol 22(1):11–16

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90:502–503

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raymond M, Rousset R (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rejón MR, Robles F, de la Herrán R, Garrido-Ramos M, Rejón CR (2009) Identification of sturgeon caviar using DNA Markers. Fish and fisheries series, 1, volume 29. Biol Conserv Sustain Develop Sturgeons III:299–319

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Roman J, Bowen BW (2000) Mock turtle syndrome: genetic identification of turtle meat purchases in the United States. Anim Conserv 3:61–65

Roosenburg WM (1996) Maternal condition and nest site choice: an alternate for the maintenance of environmental sex determination. Amer Zool 36:157–168

Roosenburg WM (2004) The impact of crab pot fisheries on terrapin (*Malaclemys terrapin*) populations: where are we and where do we need to go? In: Swarth C, Roosenburg WM, Kiviat E (eds) Conservation and ecology of turtles of the Mid-Atlantic region: a symposium. Biblomania Salt Lake City, Utah, pp 23–30

Roosenburg WM, Cresko W, Modesitte M, Robbins MB (1997) Diamondback terrapin (*Malaclemys terrapin*) mortality in crab pots. Conserv Biol 5:1166–1172

Roosenburg WM, Allman PE, Fruh BJ (2003) Diamondback terrapin nesting on the Poplar Island environmental restoration project. US National Oceanic and Atmospheric Administration. Coastal Services Center. Proceedings of the 13th Biennial Coastal Zone Conference, Baltimore, MD, July 13–17, 2003. NOAA/CS/ 20322-CD. CD-ROM. Charleston, SC: NOAA Coastal Services Center

Seigel RA, Gibbons JW (1995) Workshop on the ecology, status, and management of the diamondback terrapin (*Malaclemys terrapin*) Savannah River Ecology Laboratory, 2 August 1994: final results and recommendations. Chelonian Conserv Biol 1: 241–243

Sheridan CM, Spotila JR, Bien WF, Avery HW (2010) Sex-biased dispersal and natal philopatry in the diamondback terrapin, *Malaclemys terrapin*. Mol Ecol 19:5497–5510

Sokal RR, Rohlf FJ (1994) Biometry: the principles and practice of statistics in biological research, 3rd edn. W.H. Freeman and Co, New York

Traill L, Brook B, Frankham R, Bradsaw C (2010) Pragmatic population viability targets in a rapidly changing world. Biol Conserv 143:28–34

Tucker AD, Gibbons JW, Greene JL (2001) Estimates of adult survival and migration for diamondback terrapins: conservation insight from local extirpation within a metapopulation. Can J Zool 79:2199–2209

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4: 535–538

Waples RS (2006) A bias correction for estimates of effective population size based on linage disequilibrium at unlinked gene loci. Conserv Genet 7:167–184

Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. Mol Ecol Resour 8:753–756

Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: how can we do it better? Fish Fish 9:423–449

Wilberg MJ, Dreher BP (2004) GENECAP: a program for analysis of multilocus genotype data for non-invasive sampling and capture-recapture population estimation. Mol Ecol Notes 4:783–785

Wood RC, Herlands R (1997) Turtles and tires: the impact of road kills on northern diamondback terrapin, *Malaclemys terrapin terrapin*, populations on the Cape May peninsula, southern New Jersey. In: Van Abbema J (ed) Proceedings: conservation, restoration, and management of tortoises and turtles—an international conference. New York Turtle and Tortoise Society, New York, pp 46–53

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–354

Wright S (1978) The relation of livestock breeding to theories of evolution. J Anim Sci 46:1192–1200