Vol. 565: 95–111, 2017
doi: 10.3354/meps12043

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published February 17

Misidentification of megalopae as a potential source of error in studies of population genetics and ecology of the blue crab *Callinectes sapidus*

Timothy J. Sullivan*, Joseph E. Neigel

Department of Biology, University of Louisiana at Lafayette, Lafayette, LA 70503, USA

ABSTRACT: Inaccuracy in taxonomic identification is an unknown but potentially important source of error in studies of planktonic larval ecology and evolution. We address the misidentification of blue crab (*Callinectes sapidus*) megalopae (post-larvae) as a source of error in investigations of genetic variation and factors influencing settlement. *Callinectes* spp. megalopae were sampled monthly in spring and summer from the water column at 2 locations on the Texas (USA) coast and identified by 16S mitochondrial sequences. Most of the megalopae could be assigned to *C. sapidus* (62%), *C. similis* (36%), *C. rathbunae* (1.5%), or *C. danae* (0.12%), while 5 (0.8%) were ambiguously grouped with both *C. similis* and *C. danae*. Previously used morphological characters (rostrum length, carapace length, and their ratio) were not diagnostic. Species composition differed between locations and among monthly samples. A recurring seasonal pattern in species composition was discerned, with ~95% *C. similis* in April shifting to ~95% *C. sapidus* by May/June, and variable proportions in August. This pattern strongly parallels changes in allozyme allele frequencies previously reported for blue crab megalopae at the same locations. Models selected by the Akaike information criterion indicated lunar phase, temperature, salinity, storms, and wind stress components all affecting megalopal abundance. The importance and sign of these factors differed between species. Temperature, the most important factor for each species analyzed separately, was not important when species were combined. This study demonstrates that misidentification of larvae could create the appearance of temporal genetic variation, inflate estimates of abundance, and obscure factors influencing settlement.

KEY WORDS: Larval ecology · Larval invertebrates · Larval settlement · Gulf of Mexico · *Callinectes* spp. · 16S rRNA gene · Barcoding · Blue crab

INTRODUCTION

The life histories of many aquatic animal species include a planktonic larval phase. From their release until settlement, the movements and fates of planktonic larvae influence spatial and temporal patterns of genetic variation, population dynamics, and community structure (Hempel 1965, Johnson & Black 1982, 1984, Caley et al. 1996, Pechenik 1999, Gilg et al. 2003, Cowen & Sponaugle 2009, Selkoe & Toonen 2011, Treml et al. 2012). However, progress in understanding the dynamics of the larval phase has been hindered by the conflicting requirements of identifying individual larvae to species while sampling sufficient numbers to determine their distributions in time and space. Incomplete taxonomic coverage, lack of diagnostic characters, cryptic species, and morphological plasticity have plagued investigations of the larvae of bivalves (Elderkin et al. 2016, Weigelt et al. 2016), corals (Forsman et al. 2015, Knowlton & Leray 2015), crustaceans (Witt et al. 2006, Raupach et al. 2015, Brandão et al. 2016, Katouzian et al. 2016), and fishes (Pereira et al. 2013, Hubert et al. 2015), among other groups.

*Corresponding author: tsulli1988@gmail.com

Larvae are typically too small to examine without a microscope, and characters that readily identify adults of a species are of no use for identifying their larvae (Anger 2006, Pardo et al. 2009). Many descriptions of larval morphology are based on cultured larvae spawned by females of known species identity (Anger 2001, Pardo et al. 2009). However, this approach is labor-intensive and limited by suitable culturing methods (Anger 2001). Furthermore, larval characters can be highly plastic (Anger 2006, Pardo et al. 2009), exhibiting environmental, geographic, and seasonal variation to the extent that characters that are diagnostic under one set of conditions may fail to be so under other conditions. Often it is either simply accepted that some level of misidentification is unavoidable or else the larval forms of related species are lumped together. However, such compromises leave important questions unanswered, such as how the distributions, behavior, and settlement patterns of larvae differ among related species and whether common patterns can be reliably inferred when the larvae of different species are not distinguished. Fortunately, molecular methods of species identification have advanced to the point where they may provide an alternative to morphology for the routine identification of planktonic larvae. Gene sequences represent characters that differentiate species, are constant across life stages and environmental conditions, and are easily scored by methods that have become relatively inexpensive.

The blue crab *Callinectes sapidus* is an ecologically and economically important species that has been the subject of >1000 peer-reviewed research publications (as of October 2016; http://apps.webofknowledge.com). In the Gulf of Mexico, 8 species of *Callinectes* have been reported: *C. sapidus*, *C. similis*, *C. rathbunae*, *C. danae*, *C. bocourti*, *C. larvatus*, *C. ornatus*, and *C. exasperatus* (Felder 2009). However, only 2, *C. sapidus* and *C. similis*, are common in the northern Gulf of Mexico. The majority of studies of blue crab settlement and population genetics conducted in the northern Gulf of Mexico or western Atlantic have relied on morphology to distinguish the megalopae of *C. sapidus* from those of *C. similis* (see Tables 1 & 2). Unfortunately, distinguishing the megalopae of these 2 species is 'extremely difficult based on available descriptions and keys' (Ogburn et al. 2011, p. 107). Bookhout & Costlow (1977, p. 704) compared the megalopae of lab-raised *C. similis* with those of *C. sapidus*, and found '… diagnostic differences in lengths of parts of the claw of the first leg and setation of pleopods …' However, scoring characters of this nature requires lengthy examination of each specimen and in some cases dissection (Anger 2001), hindering their widespread use. Furthermore, the extent to which pleopod setation varies in the field is not well known, and this character has been shown to be unreliable for these species (Ogburn et al. 2011). Two morphological characters used recently (rostrum length [RL] and total carapace length [TCL]) have been found to vary considerably with season and temperature, and their values for *C. sapidus* and *C. similis* can overlap, even within collections. The ratio of RL to TCL, which corrects for size, was found to better distinguish these species (Ogburn et al. 2011).

There are puzzling inconsistencies in the literature on settlement in blue crabs (Table 1). Some studies report that the megalopae of *C. similis* as well as *C. sapidus* settle in substantial numbers at sites in the northern Gulf of Mexico (Stuck & Perry 1981, Perry et al. 1995, Rabalais et al. 1995), while others report that the megalopae of *C. similis* are in such low numbers that they can be ignored (Kordos & Burton 1993, Grey et al. 2015). Interestingly, the studies that found only *C. sapidus* in high numbers were also those that combined morphological identification with allozyme (Kordos & Burton 1993) or DNA markers (Grey et al. 2015). Inconsistency also characterizes studies of the factors that drive variation in blue crab settlement rates (Table 1). Among 20 studies on Atlantic and Gulf of Mexico coasts (Table 1), there is variation in the numbers of studies that found lunar factors to be important (10 of 20), season (5 of 20), hydrodynamic factors (9 of 20), wind (9 of 20), temperature (4 of 20), salinity (5 of 20), or the presence of storms (4 of 20). However, as noted by Grey et al. (2015), these studies differed in which factors were analyzed as well as the statistical methods used to test for their importance. Thus, it is unclear whether the drivers of blue crab settlement vary as much as might be inferred from the literature.

Surveys of genetic marker variation in blue crabs have been generally consistent in their findings, with one notable exception. Most surveys found little or no significant geographic variation, even over 100s of km (Table 2). This is expected for a species with planktonic larvae that remain offshore for a month or longer before moving inshore to settle (Costlow & Bookhout 1959). Similarly, while significant between-year changes in marker frequencies have been detected in blue crab adults and megalopae, they are typically small in magnitude (Yednock & Neigel 2014). However, one survey of variation at 3 allozyme loci in blue crabs from bays along the Texas coast (Kordos & Burton 1993) found significant differences in allele frequencies over distances of <100 km. This

Table 1. Blue crab larval ecology studies, including: whether *Callinectes* spp. identification (ID) was attempted (no. of *Callinectes* spp. encountered), area studied (ATL = Atlantic; GOM = Gulf of Mexico coasts), factors investigated (for lunar phase: P = period, Q= quarter; Flux = sea level flux, Anomaly = sea level anomaly, Day-night = day-night cycle, Temp = temperature, Sal = salinity, CDOM = colored dissolved organic matter, Speed = wind speed, Dir = wind direction, Stress = wind stress, Storm = presence of storms), and which factors influenced settlement. NR = not reported. Date of final literature search: December 19, 2016

Study	Spp. ID attempted (no. encountered)	ID method	Area	Years sampled	Factors investigated	Factors influencing settlement
Stuck & Perry (1981)	Yes (2)	Stuck & Perry (1982)	GOM	1974–1979	Season	Season
Goodrich et al. (1989)	NR (NR)	NR	ATL	1985–1987	Lunar (P), Storm	Lunar (P), Storms
Van Montfrans et al. (1990)	NR (NR)	NR	ATL	1985–1986	Lunar (Q), Flux	Lunar (Q), Flux
Boylan & Wenner (1993)	Yes (2)	NR	ATL	1987–1988	Lunar (Q), Flux, Temp, Sal, Speed & Dir	Lunar (Q), Flux, Temp, Sal, Speed & Dir
Jones & Epifanio (1995)	Yes (2)	Costlow & Bookhout (1959)	ATL	1989–1992	Lunar (Q), Spring tide, Anomaly, Speed	Anomaly, Speed
Mense et al. (1995)	Yes (NR)	NR	ATL	1990–1992	Lunar (Q), Speed & Dir, Temp, Sal	Lunar (Q), Dir
Perry et al. (1995)	Yes (2)	NR	GOM	1991–1992	Lunar (Q), Flux, Temp, Sal, Speed & Dir	Temp, Sal
Van Montfrans et al. (1995)	NR (NR)	NR	ATL	1989–1992	Lunar (Q)	Lunar (Q)
Rabalais et al. (1995)	Yes (2)	NR	GOM	1990–1992	Lunar (Q), Flux, Temp, Sal, Dir	Lunar (Q)
Morgan et al. (1996)	Yes (2)	Stuck & Perry (1982)	GOM	1990–1991	Lunar (P), Flux, Temp, Speed & Dir	Speed, Dir
Hasek & Rabalais (2001)	Yes (NR)	NR	GOM	1990–1991	Lunar (Q), Declination, Flux, Max. sea level, Temp, Sal, Speed & Dir	Lunar (Q), Declination, Max. sea level
Spitzer et al. (2003)	Yes (2)	Stuck & Perry (1982)	GOM	1997–1998	Lunar (Q), Flux, Temp, Sal, Speed & Dir, Max. speed	Speed, Dir
Forward et al. (2004)	NR (NR)	NR	ATL	1993–2002	Season, Lunar (P), Max. night sea level, Speed & Dir	Season, Lunar (P), Max. night sea level
Ogburn et al. (2009)	Yes (NR)	Costlow & Bookhout (1959)	ATL	2004–2006	Stress, Tidal range, Duration of flood tide, Water level, Temp, Storms	Flood tides, Onshore winds, Storms, Downwelling
Bishop et al. (2010)	Yes (NR)	NR	ATL	2005	Season, Max. night sea level, Max. sea level, Temp, Sal, Speed & Dir	Season, Max. night sea level, Temp, Sal, Speed, Dir
Eggleston et al. (2010)	Yes (2)	Stuck & Perry (1982)	ATL	1996–2005	Lunar (Q), Flux, Night flood tide, Speed & Dir, Storm days	Speed, Dir, Storm days
Ogburn et al. (2012)	Yes (2)	Ogburn et al. (2011)	ATL	1993–2009	Season, Lunar (Q), Night flood tide, Tidal range, Speed & Dir	Season, Night flood tide, Speed & Dir, Storms
Grey et al. (2015)	Yes (2)	Ogburn et al. (2011)	GOM	2011–2012	Season, Lunar (P), Flux, Max. sea level, Speed & Dir	Season, Lunar (P)
Biermann et al. (2016)	Yes (NR)	Costlow & Bookhout (1959)	ATL	2005–2006	Sal, Temp, CDOM, Current velocity, Light, Tides, Day-night	Day-night, Tides, Sal
Present study	Yes (4)	16S sequencing	GOM	2013–2015	% Lunar illumination, Temp, Sal, N/S & E/W wind stress components, Storms	% Lunar illumination, Temp, Sal, Storms

Table 2. Blue crab *Callinectes sapidus* population genetic study designs and results from Atlantic and Gulf of Mexico populations. Structure = inferred population structure, Temporal = presence of temporal variance in allele frequencies, RFLP = restriction fragment length polymorphism, NR = not investigated or reported in the original study. Date of final literature search: December 19, 2016

Study	Area	Larvae included	ID method	Marker type (no. used)	Structure	Temporal
Cole & Morgan (1978)	Maryland, USA	No	NR	Allozyme (20)	None	NR
Kordos & Burton (1993)	Texas, USA	Yes	Costlow & Bookhout (1959)	Allozyme (3)	All stages (all loci)	All stages, all loci
McMillen-Jackson et al. (1994)	Atlantic and Gulf of Mexico	No	NR	Allozyme (19)	Adults/juveniles (3 loci)	Adults/juveniles (1 locus)
Berthelemy-Okazaki & Okazaki (1997)	Alabama, Louisiana, and Texas, USA	No	NR	Allozyme (9)	None	NR
McMillen-Jackson & Bert (2004)	Atlantic and Gulf of Mexico	No	NR	RFLP (5)	Low in adults/juveniles	NR
Yednock & Neigel (2014)	Louisiana, USA	Yes	16S sequencing	Nuclear sequences (4)	Adults/juveniles	Adults/juveniles, not megalopae
Lacerda et al. (2016)	Brazil	No	NR	Nuclear microsatellite (9)	None	NR

survey also found extreme between-month shifts in allele frequencies for megalopae sampled from the water column, e.g. the frequency of an esterase allele increased from <0.1 to 1.0 over just 2 mo (Kordos & Burton 1993). Multiple explanations have been proposed to account for this study's anomalous levels of spatial and temporal variation, including larvae arriving in variable proportions from genetically divergent source populations and selection of varying strength and direction (Kordos & Burton 1993). The possibility that the appearance of temporal variation was caused by unintentional lumping of the megalopae of *C. sapidus* and *C. similis* was discounted because several characters were examined and appeared to confirm that nearly all the megalopae belonged to *C. sapidus*. These characters included: (1) the number of long simple setae on the distal end of epipodite of the 3rd maxilliped for megalopae (Bookhout & Costlow 1977), (2) the number of frontal teeth on the carapace of early juveniles (Williams 1984), and (3) GOT (glutamate-oxaloacetate transaminase) and EST (esterase) allozyme electrophoretic patterns in comparison with those of 2 presumptive early juveniles of *C. similis*.

Here, we assess the suitability of the mitochondrial 16S sequence to distinguish the megalopae of *C. sapidus* from *C. similis* and other species of *Callinectes*. We apply this approach to estimate abundances of each species in the water column and evaluate the potential for misidentification or intentional lumping of species to produce erroneous conclusions about genetic heterogeneity and factors influencing abundance. We also discuss the problem of judging whether misidentification has influenced the findings of particular studies of blue crab genetics and settlement, and the need for more detailed reporting and greater standardization of methods used for identification.

MATERIALS AND METHODS

Collection of megalopae

Megalopae were collected between August 2013 and August 2015 (Table 3) at 4–6 wk intervals during summer months, from 2 locations near Galveston, TX (GLV) and Freeport, TX (FPT; Fig. 1). Collections were made using a 253 µm plankton net towed in ~1 m of water in the surf zone parallel to shore between 05:00 and 09:00 h. The time of year, locations, and method of collection were chosen for comparison with the findings of Kordos & Burton (1993). Tows were standardized at ~200 m and performed in triplicate. Measurements of temperature and salinity were taken with a

Table 3. Collection information for *Callinectes* spp. megalopae. *C. sapidus* abundance = monthly combined abundance adjusted for proportion of individuals sequenced as *C. sapidus*, and *C. similis* abundance = monthly combined abundance adjusted for proportion of individuals sequenced as *C. similis*

Location	Collection date (m/d/y)	Salinity (ppt)	Temperature (°C)	Combined abundance (megalopae per tow)	*C. sapidus* abundance (megalopae per tow)	*C. similis* abundance (megalopae per tow)
Galveston	8/4/2013	30	30	13	0	12
	4/11/2014	20	20	91	0	91
	5/22/2014	16	26	39	18	14
	6/25/2014	24	27	21	20	1
	7/25/2014	28	30	17	11	4
	4/18/2015	19	24	60	0	54
	5/16/2015	17	28	37	37	0
	6/14/2015	24	29	142	122	9
	7/11/2015	30	29	33	30	4
	8/8/2.015	35	30	34	23	10
	Mean	–	–	49	28	20
Freeport	8/2/2013	29	31	18	0	16
	4/13/2014	19	19	32	0	32
	5/23/2014	18	25	48	44	5
	6/24/2014	23	28	32	32	0
	7/24/2014	26	31	40	39	0
	4/17/2015	23	20	13	0	13
	5/15/2015	12	25	152	147	5
	6/13/2015	14	30	528	528	0
	7/10/2015	18	34	33	31	2
	8/7/2015	29	32	241	222	19
	Mean	–	–	114	104	9

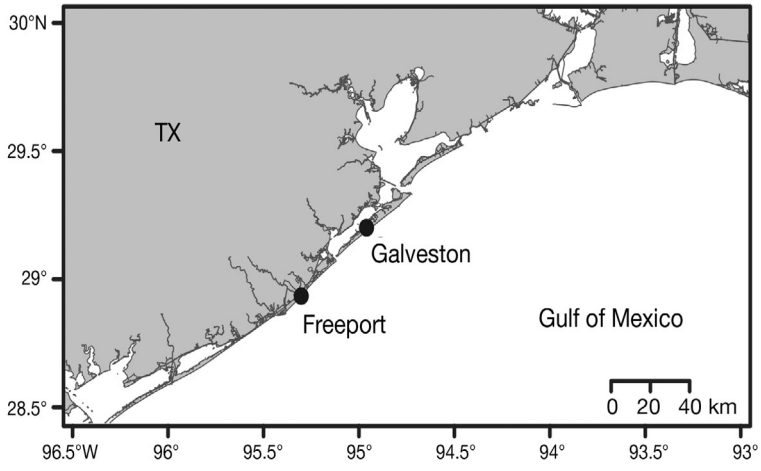

Fig. 1. Megalopal sampling locations along the Texas (USA) coast of the Gulf of Mexico

thermometer and refractometer at the time of collection. Plankton samples were washed from the net onto a 200 µm filter with ambient seawater. All megalopae of *Callinectes* spp. were immediately removed and preserved in pre-chilled 95% ethanol at 4°C.

Photography and morphological measurements

Megalopae collected in 2015 were photographed with a Nikon SMZ18 research stereo microscope. Care was taken to ensure that the dorsal surface of each individual was horizontal before images were taken. From these images, measurements of RL and TCL were made using ImageJ (http://imagej.nih.gov). RL, TCL, and the ratio RL:TCL were evaluated for identification of megalopae, following Ogburn et al. (2011). Statistical significance of differences were tested using an ANOVA model in R 3.2.0, and distributions by both species and month were visualized with box-and-whisker plots. Lastly, we used our measurements of percent RL:TCL for megalopae of *C. sapidus* and *C. similis* identified by DNA sequencing to estimate the percentage of megalopae that would have been correctly identified, incorrectly identified, or considered ambiguous in each month of sampling by comparing our measured values to those in Ogburn et al. (2011, their Fig. 1).

DNA extraction, amplification, and sequencing

Genomic DNA was extracted from whole ethanol-preserved megalopae using NucleoSpin 96 tissue kits (Macherey-Nagel) with an epMotion 5075 TMX liquid handling workstation (Eppendorf) following the manufacturer's protocol. DNA concentrations were determined with a NanoDrop spectrophotometer (Thermo Scientific). Species identification was based on the portion of the mitochondrial large subunit ribosomal RNA gene (16S) amplified with the universal primers 16SAR and 16SBR (Palumbi 1991). PCR reactions were in 15 µl with 1× AmpliTaq Gold PCR buffer (Applied Biosystems), 2.5 mM $MgCl_2$, 1 mM dNTPs, 1.2 µM of both forward and reverse primers, 0.6 units of AmpliTaq Gold (Applied Biosystems), 20 ng of DNA, and Milli-Q water. PCR conditions are as follows: 95°C for 5 min, then 40 cycles of: 96°C for 15 s, 56°C for 30 s, 72°C for 45 s, and lastly 72°C for 10 min. PCR products were electrophoresed on 0.7% agarose gels with 0.05% ethidium bromide and visualized on a Molecular Imager Gel Doc XR system (Bio-Rad). Amplicons were prepared for sequencing by modification with exonuclease I (New England Biolabs) and Antarctic phosphatase in a reaction with 0.1 µl (20 U μl^{-1}) exonuclease I, 0.3 µl (5 U μl^{-1}) phosphatase, and 6.6 µl Milli-Q filtered water at 37°C for 1 h 15 min. Enzymes were then inactivated by incubation at 95°C for 5 min. Cycle sequencing reactions were performed in reactions of 10 µl volume with 4.5 µl Milli-Q filtered water, 2.5 µl (5×) sequencing buffer (0.4 M Tris-HCl, pH 9, 10 µM $MgCl_2$), 2 µl primer (0.8 µM), and 0.5 µl BigDye Terminator v.1.1 (Applied Biosystems). The cycle sequencing profile followed Platt et al. (2007). Products were purified by ethanol precipitation, rehydrated in 15 µl Hi-Di formamide (Applied Biosystems), and denatured at 95°C for 5 min. Sequencing was performed on an ABI 310 genetic analyzer (Applied Biosystems); base calls were made with KB Basecaller in Applied Biosystems Sequencing Analysis software v.5.2.

Sequence-based identification

The mitochondrial *cytochrome c oxidase subunit I* (*COI*) gene is the standard for metazoan DNA barcoding (Hebert et al. 2003). Yet *COI* sequences are publicly available for only 3 species within the genus *Callinectes*. The 16S gene is an attractive alternative for *Callinectes* and many other crustacean taxa because it has been a standard for crustacean systematics (Schubart et al. 2000). Consequently, sequences of the mitochondrial 16S gene for the 13 American species of *Callinectes* and 2 outgroup species, *Arenaeus cribrarius* and *Charybdis hellerii*, were downloaded from GenBank to be used as reference sequences for DNA sequence-based identifications in this study (Table S1 in the Supplement at www.int-res.com/articles/suppl/m565p095_supp.pdf).

All 16S sequences for *Callinectes* currently in GenBank (release 214) were used, except for 2: accession numbers U75267 and J298169. U75267 is attributed to a specimen of *C. sapidus* (Geller et al. 1997) for which no photographs or vouchers are available (J. Geller pers. comm.). It has been convincingly argued that this sequence is unlikely to represent *C. sapidus* (Schubart et al. 2001), and it is considered by systematists who have worked on *Callinectes* to be misidentified (Schubart et al. 2001, D. Felder pers. comm., J. Schubart pers. comm.). J298169, identified as representing *C. bocourti*, was not used because the sequence is only 295 bp in length, while all other reference sequences were at least 469 bp in length. The remaining reference sequences were aligned with MUSCLE (Edgar 2004). MEGA 7 (Kumar et al. 2016) was used to construct a neighbor-joining (NJ) tree from the untrimmed alignment with Tamura-Nei distances (Tamura & Nei 1993) and a gamma distribution of rate variation with a parameter of 0.134 (see next paragraph).

The combined set of megalopal and reference sequences was aligned with MUSCLE with the recommended default parameters and trimmed to a length of 401 bp. A Perl script was used to identify all unique haplotype sequences (Table S1 in the Supplement), and these 34 haplotypes were re-aligned in MUSCLE. The Perl script MrAIC (Nylander 2004) in combination with PhyML 3.0 (Guindon & Gascuel 2003) was used to determine the best model of sequence evolution and to fit a gamma distribution with 5 rate categories to the distribution of rates across sites. MEGA 7 (Kumar et al. 2016) was used to construct a NJ tree of the re-aligned haplotype sequences with the Tamura-Nei distances and a gamma parameter of 0.134 as determined by PhyML 3.0 and MrAIC. A Perl script was used to generate statistics for pairwise p-distances (uncorrected percent differences) within and between species for species of *Callinectes*. Megalopae were assigned to taxa by either of 2 criteria: (1) an exact match to a reference sequence, or (2) placement within the clade of haplotypes in which all reference sequences represented *C. sapidus*.

Analysis and models of factors influencing abundance

Regression analysis was used to investigate the effects of 6 variables on the abundance of megalopae collected from the water column at locations where settlement was occurring. Separate models were developed for the megalopae of *C. sapidus*, *C. similis*, and the 2 species combined. The independent variables, all previously reported to explain variation in rates of blue crab settlement, were percent collection illumination, water temperature at the time of collection, salinity at the time of collection, the occurrence of tropical storms (based on National Hurricane Center Designations following the Saffir-Simpson scale for 1 wk pre- or post- collection; www.nhc.noaa.gov), the north/south component of wind stress, and the east/west component of wind stress. Wind stress components were calculated from the equation for wind stress from Large & Pond (1981), with wind speed and direction downloaded from the National Data Buoy Program (www.ndbc.noaa.gov), and averaged over the week preceding collection. The MuMIn package in R 3.2.0 was used for model selection and multimodal inference (Burnham & Anderson 2003) based on corrected Akaike information criterion (AICc) values. Estimates of abundance (mean megalopae per tow) were $\log_{10}(x + 1)$-transformed to stabilize levels of variance and allow for zero values.

RESULTS

16S sequences

A total of 49 sequences of the mitochondrial 16S gene representing the 13 American species of *Callinectes* were downloaded from GenBank, including 10 sequences for *C. sapidus* and 2 for *C. similis*. A NJ tree (Fig. 2) for these sequences was mostly consistent with previous findings based on morphological and molecular evidence (Robles et al. 2007) and showed that *C. sapidus* and *C. similis* belong to distinct groups within *Callinectes*, and these groups are clearly differentiated by 16S sequences. In our alignment of reference and megalopal sequences, all representatives of the 2 groups differed by at least 19 nucleotide substitutions, and all *C. sapidus* sequences differed from those of *C. similis* by 26 or 27 nucleotide substitutions. Also consistent with previous findings (Schubart et al. 2001, Robles et al. 2007), sequences representing specimens identified as *C. affinis*, *C. bocourti*, and *C. maracaiboensis* were intermingled in a '*bocourti* group'. The lack of separation among these species has led to proposals that *C. maracaiboensis* (Schubart et al. 2001, Robles et al. 2007) and possibly *C. affinis* (Robles et al. 2007) should be synonymized with *C. bocourti*. Our tree also places the sequence representing *C. rathbunae* within the *bocourti* group, although this sequence differs by 4–6 substitutions from others in the group. For the sequences representing all species of *Callinectes*, the distributions of pairwise p-distances (uncorrected percent differences) within species overlapped considerably with the distribution between species. However, there was no overlap for p-distances within *C. sapidus* versus p-distances between *C. sapidus* and other species (Fig. S1 in the Supplement).

We determined 16S sequences for 639 megalopae, 316 from GLV and 323 from FPT. Of these, 614 (96%) were exact matches to reference sequences, including 376 that matched haplotypes of *C. sapidus*, 227 that matched haplotypes of *C. similis*, 10 that matched the sequence for *C. rathbunae*, and 1 that matched a haplotype of *C. danae* (GenBank: AJ298184.1). The 25 that were not exact matches to reference haplotypes comprised 6 distinct haplotypes; 4 of these haplotypes (20 megalopae) were nested within the group of reference haplotypes for *C. sapidus*, and so were assigned to *C. sapidus* (Fig. 3). The other 2 haplotypes (5 megalopae) were placed in sister relationships to the group of haplotypes representing *C. similis* and *C. danae* (Fig. 3), and so could not be assigned to a single species on the basis of sequence alone. These unassigned megalopae were removed from the data set prior to subsequent analysis.

Seasonal patterns in species composition

In samples from both GLV and FPT, the monthly composition of *Callinectes* species was variable and followed a consistent seasonal pattern. At GLV, nearly all of the megalopae sampled in April of 2014 and 2015 were identified as *C. similis*, but in May, most of those sampled were identified as *C. sapidus*, and in June, the percentage of *C. sapidus* reached ~95% (Fig. 4A). In late summer, the proportion of megalopae identified as *C. similis* increased, in 1 year (2013) outnumbering *C. sapidus* (Fig. 4A). A similar pattern was observed at FPT. In April of 2014 and 2015, all the megalopae were identified as *C. similis* (Fig. 4B), but in May of both years, nearly

all (~95%) were identified as *C. sapidus*, and this preponderance of *C. sapidus* continued through June and July in both years (Fig. 4B). In August 2015, the majority of megalopae sampled were identified as *C. sapidus*, while in August 2013, none of the sampled megalopae were identified as *C. sapidus*; most were identified as *C. similis* along with a few as *C. rathbunae* (Fig. 4B). At both locations, megalopae identified as *C. rathbunae* were found only in July or August. The single megalopa identified as *C. danae* was found at GLV in July. Megalopae that could not be identified because their 16S sequences were similar to those of both *C. similis* and *C. danae* (haplotypes H32 and H34) were found at GLV in April and August, and at FPT in April and May.

Correspondence between morphological and sequence-based identification

Among the megalopae collected in 2015 and identified by 16S sequences, 285 individuals from 3 species (Fig. 5) were sufficiently intact for the measurement of 2 characters that have been used to distinguish *C. sapidus* from *C. similis*: RL and TCL. In our samples, RL and the ratio RL:TCL significantly differed among megalopae identified as *C. sapidus*, *C. similis*, and *C. rathbunae* (RL: $F = 8.6$, df = 2,282, $p < 0.001$, Fig. 6A; RL:TCL ratio: $F = 9.3$, df = 2,282, $p < 0.001$, Fig. 6B), but differences in TCL were not significant ($F = 2.3$, df = 2,282, $p = 0.099$, Fig. 6C). Tukey's difference of

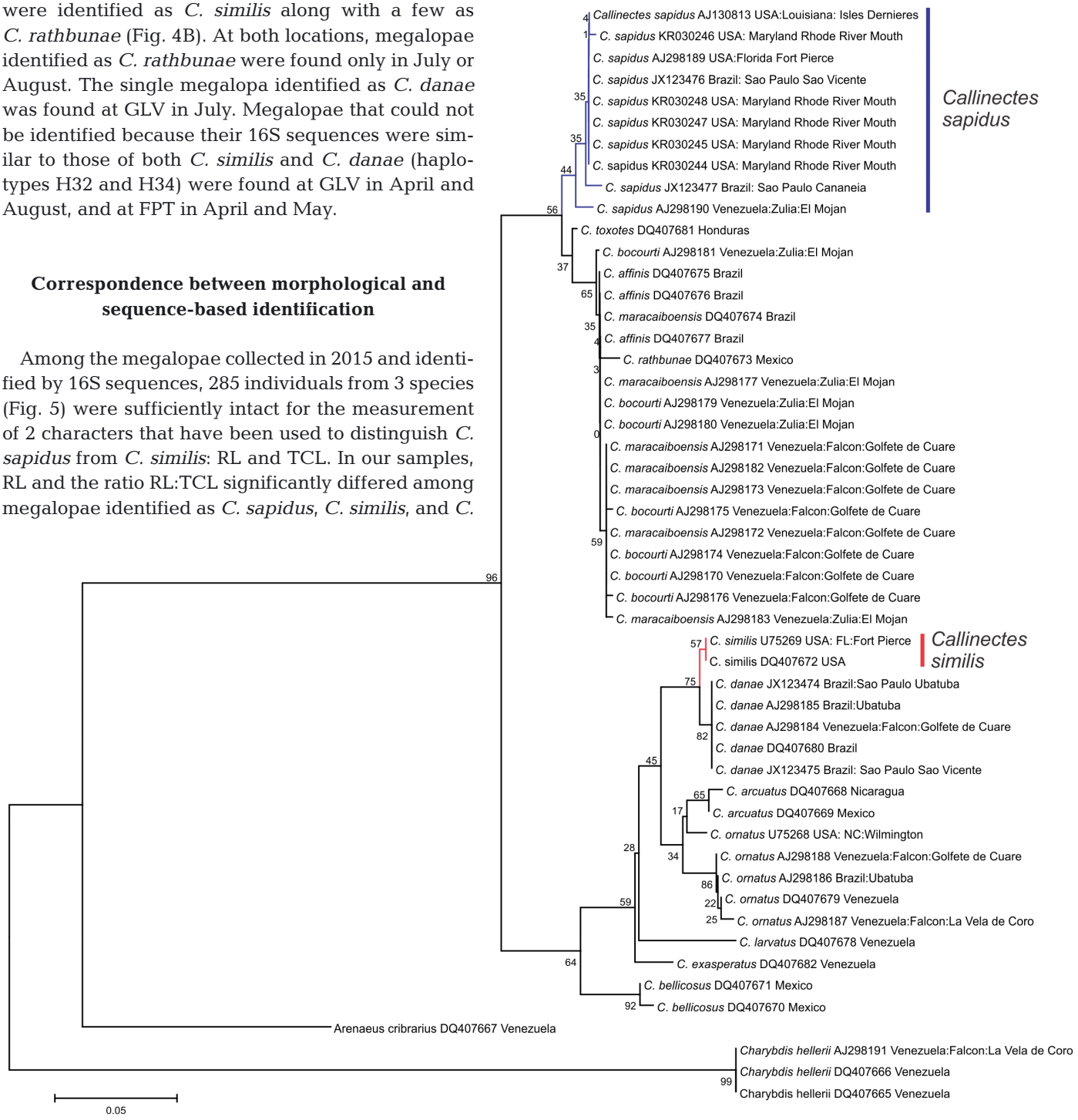

Fig. 2. Bootstrapped neighbor-joining tree of 16S reference sequences for American species of *Callinectes* with species name, GenBank accession no., and location of collection (where available). Blue = sequences for *C. sapidus*, red = sequences for *C. similis*

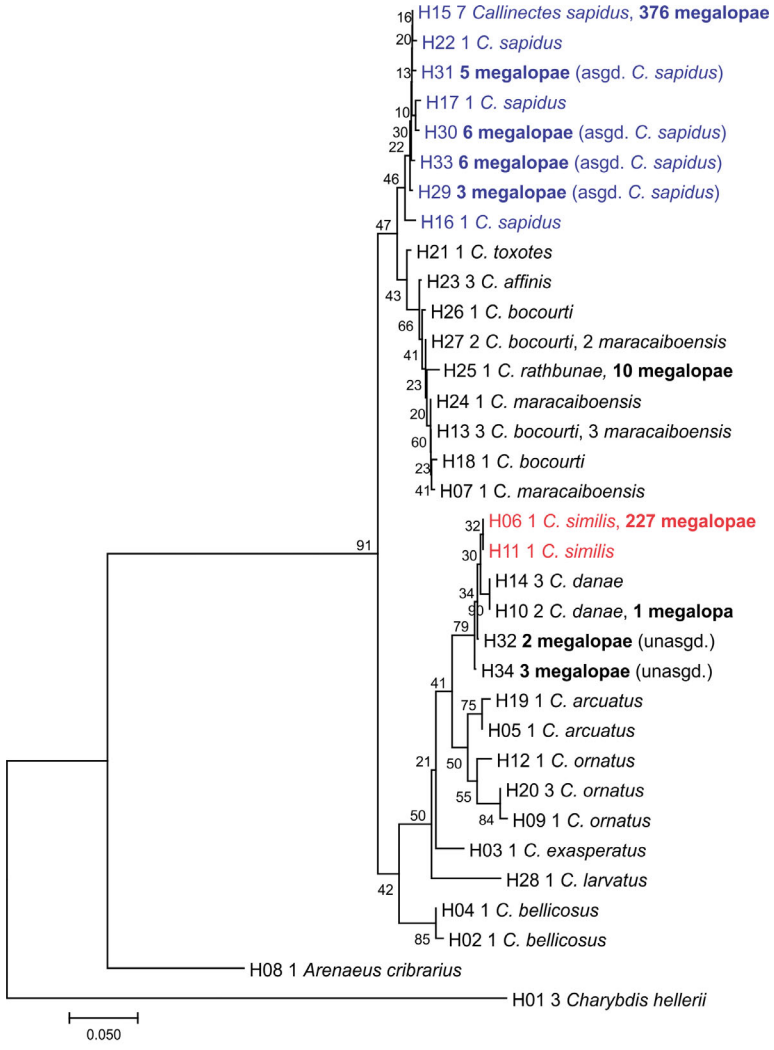

Fig. 3. Bootstrapped neighbor-joining tree for haplotypes of 16S sequences representing a combined 401 bp alignment of *Callinectes* spp. reference sequences. Each numbered haplotype designation is followed by the number of reference sequences from each species that shared the haplotype and the number of megalopae with that haplotype in bold. Asgd. = assigned to that species, unasgd. = not assigned to any species, blue = haplotypes for *C. sapidus*, red = haplotypes for *C. similis*

means post hoc test showed *C. similis* was on average larger than *C. sapidus* by 0.05 mm for RL and 1.33% for RL:TCL, both of which were significant ($p < 0.001$). Overall, these findings were similar to those presented for *C. sapidus* and *C. similis* in Ogburn et al. (2011), with mean values of RL and RL:TCL for *C. similis* appearing larger than for *C. sapidus*. However, in our samples, the ranges of measured values for the 2 species were greater than previously reported, both within months (Fig. S2A–C in the Supplement) and across all months (Fig. 6A–C). RL and RL:TCL did not significantly differ between *C. rathbunae* and either *C. sapidus* or *C. similis* (Fig. 6A–C; Fig. S2A–C in the Supplement).

We used megalopae identified by DNA sequencing to investigate the accuracy of morphological identification based on the ranges of percent RL:TCL for *C. sapidus* and *C. similis* reported by Ogburn et al. (2011). For megalopae we collected between May and August 2015, our Table 4 shows that 72.2% were assigned to the correct species based on percent RL:TCL, 12.4% were incorrectly assigned to a different species, and 15.4% fell outside the range of values reported by Ogburn et al. (2011). It should be noted that there could be differences in morphology associated with the locations where megalopae were collected by Ogburn et al. (2011) in Mississippi and South Carolina and the locations in Texas where our collections were made.

Effects of environmental factors on abundance

The abundance of megalopae in the water column, estimated as catch-per-standard-tow of a plankton net, varied over time and was correlated with environmental factors previously found to explain variation in blue crab settlement rates (Table 1). For all *Callinectes* species combined, abundance ranged from 13 to 528 megalopae per tow (Table 3), for *C. sapidus* it ranged from 0 to 528, and for *C. similis* from 0 to 91 (Table 3). *Callinectes* spp. abundance and the abundance of *C. sapidus* was higher at the FPT location, but abundance of *C. similis* was higher at GLV (Table 3).

There were 4 models of the effects of environmental factors on combined (both *C. sapidus* and *C. similis*) abundance within 2 ΔAICc of the best model (Table 5). They included a combination of salinity, percent lunar illumination, the east/west wind stress component, and the presence of tropical storms, with no 1st-order interactions (Table 5). The most important parameter was the presence of tropical storms, which occurred in all 4 models (Table 5). The results of our averaged final model for combined abundance was:

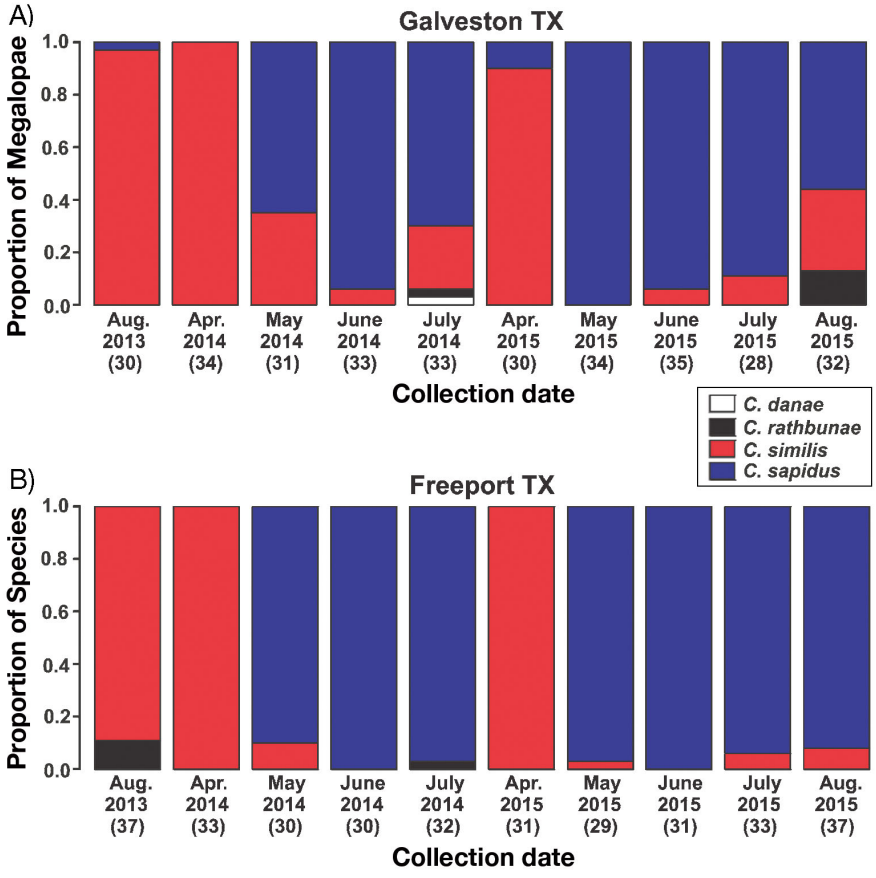

Fig. 4. Species composition of *Callinectes* spp. megalopae collected from water column at sites where settlement was occurring near (A) Galveston and (B) Freeport. Values in parentheses indicate number of megalopae identified by sequencing

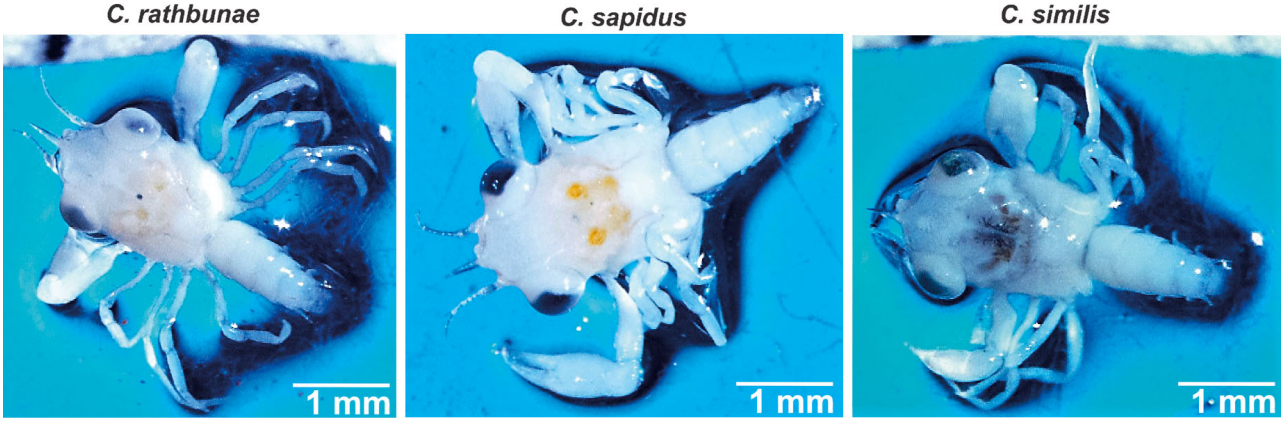

Fig. 5. Photomicrographs of megalopae identified as *Callinectes rathbunae*, *C. sapidus*, and *C. similis* by 16S sequences

$$\text{Log(Combined abundance + 1)} = 3.64 + 1.76 \times \text{Storms} - 10.89 \times \text{Week 1 east/west wind stress component} + 0.70 \times \text{Percent lunar illumination} - 0.04 \times \text{Collection salinity} \quad (1)$$

Abundance was unusually high in June 2015, coinciding with the passage of tropical storm Bill. Abundance also increased with higher percent lunar illumination and decreased with higher salinity. Lastly, our results showed that abundance was higher when wind stress was from the east (onshore; Table 5).

There were 4 models of the effects of environmental factors on the abundance of *C. sapidus* megalopae within 2 ΔAICc of the best model (Table 5). They contained a combination of collection salinity, collection

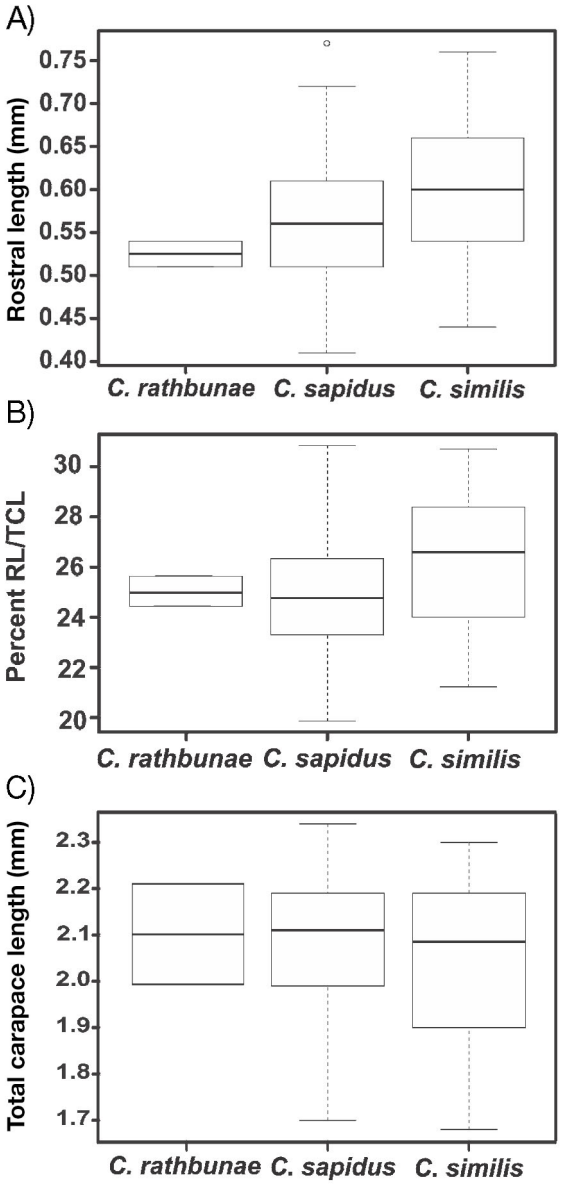

Fig. 6. Box-and-whisker plots of (A) rostrum length, (B) ratio of rostrum length to total carapace length (RL:TCL) as a percentage, and (C) total carapace length for megalopae collected in 2015 and identified as *Callinectes rathbunae*, *C. sapidus* and *C. similis* by 16S sequences. Whiskers show 1.5× the interquartile range

temperature, percent lunar illumination, and the presence of tropical storms, with no 1st-order interactions (Table 5). The most important parameter was collection temperature, which occurred in all 4 of the best models (Table 5). The AIC-weighted average model for abundance of *C. sapidus* was:

$$\text{Log}(C.\ sapidus\ \text{abundance} + 1) = -2.91 - 0.14 \times \text{Collection salinity} + 0.30 \times \text{Collection temperature} + 2.10 \times \text{Storm} + 1.41 \times \text{Percent lunar illumination} \quad (2)$$

Table 4. Comparison of morphological identification and 16S sequencing for *Callinectes* spp. megalopae collected in May–Aug 2015, including total number of individuals (N), the number correctly (N_C) and incorrectly (N_I) identified, and the number unable to be classified (N_U) based on monthly *C. sapidus* and *C. similis* percent rostrum length to total carapace length ratio (RL:TCL) variability presented in Ogburn et al. (2011). Percentages shown are of the total N for that row

Month	N	N_C	N_I	N_U
May	64	51 (80%)	4 (6%)	9 (14%)
Jun	66	46 (70%)	9 (14%)	11 (17%)
Jul	60	41 (68%)	9 (15%)	10 (17%)
Aug	44	31 (70%)	7 (16%)	6 (14%)
Total	234	169 (72.2%)	29 (12.4%)	36 (15.4%)

Abundance increased with higher temperature, percent lunar illumination, and tropical storm passage, and decreased with higher salinity (Table 5).

There were 4 models of the effects of environmental factors on the abundance of *C. similis* megalopae within 2 ΔAICc of the best model (Table 5). The most important parameter was temperature, included in all 4 of the best models (Table 5). The AIC-weighted average model for abundance of *C. similis* was:

$$\text{Log}(C.\ similis\ \text{abundance} + 1) = 5.99 + 0.09 \times \text{Collection salinity} - 0.19 \times \text{Collection temperature} - 13.34 \times \text{Week 1 east/west wind component} \quad (3)$$

Abundance increased with higher salinity and decreased with higher temperature. Abundance was also higher when wind stress was from the east (Table 5).

There were striking differences in the importance and sign of the effects of particular environmental factors between models for combined abundance of megalopae and those for the abundance of each species considered separately (Table 6). Higher salinity was associated with decreased combined abundance and abundance of *C. sapidus*, but with increased abundance of *C. similis* (Table 6). Percent lunar illumination and the presence of tropical storms were associated with higher combined abundance and abundance of *C. sapidus*, but had no effect on the abundance of *C. similis* (Table 6). For models of combined and *C. similis* abundance, wind stress from the east increased abundance, but did not affect the abundance of *C. sapidus*. Lastly, higher temperatures increased abundance of *C. sapidus* and decreased abundance of *C. similis* but did not affect their combined abundance (Table 6).

Table 5. Model results of relationships between *Callinectes sapidus*, *C. similis*, and combined abundance (megalopae per tow) and environmental, physical, and lunar characteristics. Log LK = log likelihood, AICc = corrected Akaike information criterion, Percent moon = percent lunar illumination, Storm = occurrence of tropical storms. Models presented are less than 2 ΔAICc from the best model

	Model	Log LK	AICc	ΔAICc	Weight
Combined	Abundance ~ 1 + Storm + Week 1 east/west wind stress component	–19.54	49.74	0.00	0.40
	Abundance ~ 1 + Percent moon + Storm + Week 1 east/west wind stress component	–18.23	50.74	1.00	0.24
	Abundance ~ 1 + Collection salinity + Storm + Week 1 east/west wind stress component	–18.51	51.31	1.57	0.18
	Abundance ~ 1 + Percent moon + Storm	–20.35	51.37	1.63	0.18
C. sapidus	Abundance ~ 1 + Collection salinity + Collection temperature	–34.58	79.83	0.00	0.35
	Abundance ~ 1 + Collection salinity + Collection temperature + Storm	–33.02	80.33	0.50	0.27
	Abundance ~ 1 + Collection salinity + Collection temperature + Percent moon	–33.11	80.51	0.68	0.25
	Abundance ~ 1 + Collection temperature + Storm	–35.58	81.82	1.99	0.13
C. similis	Abundance ~ 1 + Collection salinity + Collection temperature	–29.09	68.85	0.00	0.31
	Abundance ~ 1 + Collection salinity + Collection temperature + Week 1 east/west wind stress component	–27.31	68.90	0.05	0.30
	Abundance ~ 1 + Collection temperature	–30.79	69.07	0.22	0.27
	Abundance ~ 1 + Collection temperature + Week 1 east/west wind stress component	–30.00	70.67	1.81	0.12

DISCUSSION

Callinectes spp. along Texas coast

In this study, we used 16S mitochondrial DNA sequences to identify megalopae of *Callinectes* to species level and showed that this approach is likely to be more accurate than is currently practical with morphology. We identified the megalopae of at least 3 species of *Callinectes* along the Texas coast, and found that the megalopae of *C. sapidus* and *C. similis* were, at different times, highly abundant and followed distinct seasonal patterns. Greater abundance of *C. sapidus* was associated with higher water temperatures; the reverse was true for *C. similis*. The abundances of the 2 species were also affected differently by salinity, wind, storms, and lunar phase. Although this study was not intended as a thorough investigation of factors influencing blue crab settlement, it did reveal that failure to distinguish the megalopae of *C. sapidus* from *C. similis* could lead to erroneous conclusions about their effects.

Suitability of mitochondrial 16S sequence for DNA barcoding

The basic premise of DNA barcoding is that sequences from the same species are more similar to each other than to sequences from different species (Hebert et al. 2004). However, this is not always the case due to instances of true paraphyly (Neigel & Avise 1986), apparent paraphyly (Funk & Omland 2003), or errors due to pseudogenes (Schneider-Broussard & Neigel 1997, Williams & Knowlton 2001). Phylogenetic relationships among 16S reference sequences for American species of *Callinectes* include multiple instances of paraphyly within the '*bocourti* group' of *C. bocourti*, *C. maracaiboensis*, and *C. affinis* (Fig. 2). Synonymizing these taxa, as has been proposed (Schubart et al. 2001, Robles et al. 2007), would resolve these apparent cases of paraphyly. No other paraphyletic relationships are apparent among the reference sequences representing the American *Callinectes*. However, this may simply reflect the small number of sequences that are available (Fig. 2), as small sample sizes often fail to reveal paraphyletic relationships (Meyer & Paulay 2005).

Our results do suggest that the 16S gene can be used to distinguish *C. sapidus* from other species of *Callinectes* in the northern Gulf of Mexico. Pairwise p-distances among the 10 reference sequences for *C. sapidus* were all less than those between *C. sapidus* and the other American species of *Callinectes* (Fig. S1 in the Supplement). *C. similis* and *C. ornatus* are also common in the Gulf of Mexico (Felder 2009), but both of these species are in a different subclade of *Callinectes* than *C. sapidus*, making it unlikely that true paraphyly would

Table 6. Influence of each factor on *Callinectes sapidus*, *C. similis*, and overall abundance in our final models. Storm = occurrence of tropical storms, No = factor not included in the final model, '+' indicates positive influence, '−' is negative influence

Factor	Overall	*C. sapidus*	*C. similis*
Collection temperature	No	+	−
Collection salinity	−	−	+
Percent lunar illumination	+	+	No
Storm	+	+	No
Week 1 east/west wind stress component	−	No	−

occur among their sequences. *C. toxotes* is in the same subclade as *C. sapidus*, but its range is confined to the eastern Pacific. The species of the *bocourti* group and *C. rathbunae* are also within the same subclade as *C. sapidus*. In our NJ tree of 16S reference sequences, these species are together in a group that is clearly separated (although with low bootstrap support) from the group of *C. sapidus* sequences (Fig. 2). We therefore found no indication of paraphyly between sequences of *C. sapidus* and other species that occur in the Gulf of Mexico.

16S sequences for the majority of the megalopae we sampled matched the most common haplotype of *C. sapidus*, shared by 7 out of 10 reference sequences for that species (Fig. 3). Sequences from an additional 20 megalopae represented 4 new haplotypes nested within the reference sequences for *C. sapidus* (Fig. 3). Most of the remaining megalopae matched 1 of the 2 reference haplotypes for *C. similis* (Fig. 3). However, the reference sequences for *C. similis* are very similar to those of *C. danae.* Furthermore, the sequence for 1 megalopa matched 1 of the 2 haplotypes represented among the 4 reference sequences for *C. danae*, and among 5 other megalopae, there were 2 new haplotypes (H32 and H34) that are in ambiguous sister relationships to the group consisting of sequences from *C. danae* and *C. similis* (Fig. 3). It could be argued that most if not all of these '*danae* or *similis*' sequences belong to *C. similis*, because *C. similis* is common on the coast of Texas, while the reported range of *C. danae* does not extend as far north as our sampling locations. However, on the basis of 16S sequences alone, we cannot be as confident of these identifications as we are for the megalopae identified as *C. sapidus*. Furthermore, regardless of whether haplotypes H32 and H34 represent *C. similis* or *C. danae*, their placement on the NJ tree implies paraphyletic relationships among the 16S sequences for these 2 species (Fig. 3).

We also found 10 megalopae with a 16S sequence that is an exact match to the 1 reference sequence for *C. rathbunae*. The reference sequence of *C. rathbunae* differs from other *Callinectes* reference sequences by at least 4 nucleotide substitutions. However, with only 1 sequence, we would not be able to detect paraphyly between *C. rathbunae* and other species. Nevertheless, this sequences is very different from those of *C. sapidus* (14–15 substitutions), *C. similis* (25 substitutions), or *C. danae* (26–30 substitutions), which indicates that even if we cannot be confident that it represents *C. rathbunae*, there is likely a 4th species of *Callinectes* represented among the megalopae we sampled. Overall, it appears that identifications based on 16S sequences were likely to be correct for all the megalopae assigned to *C. sapidus* and the majority of megalopae assigned to *C. similis*.

Misidentification as a cause of apparent genetic heterogeneity

Our finding of high proportions (up to 100%) of megalopae identified as *C. similis*, especially in early and late summer (Fig. 4), contrasts sharply with the findings of Kordos & Burton (1993), who reported that they detected no megalopae of *C. similis* in their samples. One possibility is that the abundance of megalopae of *C. similis* dramatically increased between the 2 decades when their collections were made (1990–1992) and our collections (2013–2015). However, *C. similis* megalopae were found to be present in seasonally high abundances in late spring/early summer from the coastal waters of Mississippi by Stuck & Perry (1981), who sampled 10–15 yr prior to the analysis conducted by Kordos & Burton (1993). Another possibility is that megalopae of *C. similis* were present in the samples of Kordos & Burton (1993), but misidentified as *C. sapidus*. The character they used to distinguish megalopae of the 2 species was the number of distal setae on the epipodite of the 3rd maxilliped: 21 for *C. similis* (Bookhout & Costlow 1977) and 14 for *C. sapidus* (Costlow & Bookhout 1959). However, morphological characters in both species, including setation (Ogburn et al. 2011), vary geographically and seasonally, and the diagnostic values of these characters were determined for megalopae hatched from ovigerous females from the Atlantic coast and cultured in the laboratory. In a later study by Ogburn et al. (2011), megalopae collected from the Gulf of Mexico (Mississippi) were reared in the laboratory until they could be identi-

fied. For these field-collected megalopae, the number of distal setae in *C. similis* ranged from 13 to 21, and overlapped with the range of values in *C. sapidus* (12–20). Kordos & Burton (1993) also reported that among several hundred megalopae reared to the 1st crab stage, only 2 were identified as *C. similis*. However, distinguishing 1st crabs and early juveniles of *C. sapidus* from those of *C. similis* is also problematic, and the character they used, the number of frontal teeth on the carapace, is based on descriptions of specimens from the Atlantic coast (Williams 1984). Finally, the 2 juveniles raised from megalopae that they recognized as clearly morphologically distinct and so identified as *C. similis* could have belonged to a 3rd species of *Callinectes*, perhaps *C. danae* or *C. rathbunae*.

The possibility that Kordos & Burton (1993) unintentionally lumped the megalopae of *C. similis* with those of *C. sapidus* would explain the dramatic changes in allele frequencies they observed at all 3 of the allozyme loci they surveyed. The timing and magnitude of the shifts in species composition that we observed are similar to those of the allele frequency shifts they reported (compare our Fig. 4 to Kordos & Burton 1993, their Fig. 2). However, lumping of species cannot explain all of the anomalous phenomena reported in Kordos & Burton (1993), such as the unusually high levels of geographic differentiation among adult populations along the Texas coast, which contrasts with the findings of other surveys of genetic variation in blue crabs (McMillen-Jackson et al. 1994, Berthelemy-Okazaki & Okazaki 1997, McMillen-Jackson & Bert 2004, Yednock & Neigel 2014).

Previous studies of the usefulness of setation (Ogburn et al. 2011), along with our investigation of the characters RL and TCL, suggest that there do not appear to be any known morphological characters that distinguish the megalopae of *C. sapidus* from *C. similis* with high reliability. Molecular characters offer a viable alternative for the identification of *C. sapidus* in the northern Gulf of Mexico, where closely related species are rare or absent. However, the prospects for accurately distinguishing among the megalopae of other species of *Callinectes* with 16S sequences are less certain. Our findings recall previously recognized limitations of DNA barcoding with a single mitochondrial gene sequence (e.g. Meyer & Paulay 2005). A general weakness of approaches based solely on mitochondrial sequences is the nonequivalence of gene trees and species trees, concerning which it has been argued that independently recombining nuclear sequences should play a greater role in crustacean systematics (Neigel & Mahon 2007, Mahon & Neigel 2008).

Importance of species-level identification in ecological analysis

Our analysis of the effects of environmental factors on the abundance of megalopae in the water column demonstrates the potential for misidentification to lead to erroneous conclusions about larval ecology. The abundances of *C. sapidus* and *C. similis* followed distinct seasonal patterns (Fig. 4), and were affected differently, even oppositely, by specific environmental factors (Tables 5 & 6). Most tellingly, the effect of temperature on abundance, the most important factor for both species considered individually, was not detected when the abundances of the 2 species were combined (Table 6). Thus, lumping species may not only fail to reveal differences among species, but may also obscure factors that affect multiple species strongly but in different ways.

The methods we used for identification should work for megalopae collected from substrates on which they have settled as well as from the water column. Our collection methods and sampling locations were chosen to follow the work of Kordos & Burton (1993), and so our data reflects the abundance and composition of megalopae in the water column. Although settlement rates are clearly constrained by water-column abundance, they also depend on hydrodynamics and megalopal behavior and so cannot be simply equated (Pawlik et al. 1991, Pawlik 1992, Raimondi & Aileen 2000, Crimaldi et al. 2002, Pernet et al. 2003, Bolle et al. 2009). However, we are confident that at least some of the megalopae we sampled were in the process of settling; during collection, megalopae were actively grasping the plankton net, tow ropes, and legs of the collector. Furthermore, the seasonal changes in species abundance we observed were so extreme (changes in relative abundance from 0 to 100%) that there were clearly times when water-column abundance determined the species composition of settlers.

It is difficult to judge which of the apparent inconsistencies in studies of blue crab settlement conducted over the past 35 yr (Tables 1 & 2) can be attributed to misidentification. Some inconsistency is expected from differences in sampling design and statistical analysis. Furthermore, settlement of blue crabs is episodic and exhibits high interannual variability, so that capturing the entirety of factors that influence settlement would require sampling at fre-

quent intervals for multiple years (e.g. Ogburn et al. 2009). Nevertheless, our limited survey of megalopal abundance along the coast of Texas demonstrates that assumptions made about species composition (e.g. all megalopae belong to *C. sapidus*) could be correct for some locations and times but very wrong for others (Fig. 4). Unfortunately, crucial methodological details that might support such assumptions, such as how many megalopae from each sample were identified, which characters were used for identification, and how much variation or overlap was found in those characters are often omitted from publications. The prospect of routine identification of large numbers of individuals using DNA sequence characters is becoming more plausible as the cost of DNA sequencing (e.g. Shokralla et al. 2015) and other methods of high-throughput genotyping (e.g. Cornwell et al. 2016) continues to decline. Furthermore, it will not always be necessary to perform DNA-based (or morphological) identification of every individual in a sample; estimating the proportions of target species for subsamples of each sample should be sufficient for many purposes. The accuracy of these estimated proportions will depend on the sizes of the subsamples, with confidence intervals that can be determined by the method of Sison & Glaz (1995).

Our findings, which suggest that previous studies may have failed to distinguish the megalopae of *C. sapidus* from those of other species, have important implications for marine management and conservation. Fisheries models for the relationship between adult stocks and recruitment (e.g. Lipcius & Stockhausen 2002) should be based upon and tested with accurate estimates of the abundances of early life stages, as should assessment of the impacts of disturbances on settlement (e.g. Grey et al. 2015). For *C. sapidus*, both misidentification and intentional lumping of related species are likely to result in overestimates of the abundance of *C. sapidus*, which is often assumed to be the most abundant if not the only species of *Callinectes* represented among settling megalopae in the northern Gulf of Mexico. Thus, failure to accurately distinguish the early life stages of *C. sapidus* from those of related species introduces systematic biases in estimates of settlement and recruitment.

Data accessibility. Data files and megalopae images are publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC) at: https://data.gulfresearchinitiative.org (DIF: R2.x214.000:0006). Sequence data generated has been submitted to GenBank (www.ncbi.nlm.nih.gov/genbank) under accession numbers KY381582 to KY381587.

Acknowledgements. This research was made possible in part by grants from the Gulf of Mexico Research Initiative (GRI-013 and GoMRI2012-II-523) and by a University of Louisiana at Lafayette Doctoral Fellowship to T.J.S. We thank Rachel Sullivan for assistance collecting samples.

LITERATURE CITED

Anger K (2001) The biology of decapod crustacean larvae. Crustac Issues 14:1–420

Anger K (2006) Contributions of larval biology to crustacean research: a review. Invertebr Reprod Dev 49:175–205

Berthelemy-Okazaki N, Okazaki R (1997) Population genetics of the blue crab *Callinectes sapidus* from the northwestern Gulf of Mexico. Gulf Mex Sci 15:35–38

Biermann JL, North EW, Boicourt WC (2016) The distribution of blue crab (*Callinectes sapidus*) megalopae at the mouths of Chesapeake and Delaware Bays: implications for larval ingress. Estuar Coasts 39:201–217

Bishop TD, Miller HL III, Walker RL, Hurley DH, Menken T, Tilburg CE (2010) Blue crab (*Callinectes sapidus* Rathbun, 1896) settlement at three Georgia (USA) estuarine sites. Estuar Coasts 33:688–698

Bolle LJ, Dickey-Collas M, van Beek JK, Erftemeijer PL, Witte JI, van der Veer HW, Rijnsdorp AD (2009) Variability in transport of fish eggs and larvae. III. Effects of hydrodynamics and larval behaviour on recruitment in plaice. Mar Ecol Prog Ser 390:195–211

Bookhout C, Costlow J (1977) Larval development of *Callinectes similis* reared in the laboratory. Bull Mar Sci 27: 704–728

Boylan JM, Wenner EL (1993) Settlement of brachyuran megalopae in a South Carolina, USA, estuary. Mar Ecol Prog Ser 97:237–246

Brandão MC, Freire AS, Burton RS (2016) Estimating diversity of crabs (Decapoda: Brachyura) in a no-take marine protected area of the SW Atlantic coast through DNA barcoding of larvae. Syst Biodivers 14:288–302

Burnham KP, Anderson DR (2003) Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media, New York, NY

Caley M, Carr M, Hixon M, Hughes T, Jones G, Menge B (1996) Recruitment and the local dynamics of open marine populations. Annu Rev Ecol Syst 27:477–500

Cole MA, Morgan RP (1978) Genetic variation in two populations of blue crab, *Callinectes sapidus.* Estuaries 1: 202–205

Cornwell BH, Fisher JL, Morgan SG, Neigel JE (2016) Chaotic genetic patchiness without sweepstakes reproduction in the shore crab *Hemigrapsus oregonensis*. Mar Ecol Prog Ser 548:139–152

Costlow JD, Bookhout C (1959) The larval development of *Callinectes sapidus* Rathbun reared in the laboratory. Biol Bull (Woods Hole) 116:373–396

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Annu Rev Mar Sci 1:443–466

Crimaldi JP, Thompson JK, Rosman JH, Lowe RJ, Koseff JR (2002) Hydrodynamics of larval settlement: the influence of turbulent stress events at potential recruitment sites. Limnol Oceanogr 47:1137–1151

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113

Eggleston DB, Reyns NB, Etherington LL, Plaia GR, Xie L

(2010) Tropical storm and environmental forcing on regional blue crab (*Callinectes sapidus*) settlement. Fish Oceanogr 19:89–106

Elderkin CL, Clewing C, Wembo Ndeo O, Albrecht C (2016) Molecular phylogeny and DNA barcoding confirm cryptic species in the African freshwater oyster *Etheria elliptica* Lamarck, 1807 (Bivalvia: Etheriidae). Biol J Linn Soc 118:369–381

Felder DL (2009) Gulf of Mexico origin, waters, and biota: biodiversity. Texas A&M University Press, College Station, TX

Forsman Z, Wellington GM, Fox GE, Toonen RJ (2015) Clues to unraveling the coral species problem: distinguishing species from geographic variation in *Porites* across the Pacific with molecular markers and microskeletal traits. PeerJ 3:e751

Forward R Jr, Cohen J, Irvine R, Lax J and others (2004) Settlement of blue crab *Callinectes sapidus* megalopae in a North Carolina estuary. Mar Ecol Prog Ser 269:237–247

Funk D, Omland K (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences with insights from animal mitochondrial DNA. Annu Rev Ecol Evol Syst 34:397–423

Geller J, Walton E, Grosholz E, Ruiz G (1997) Cryptic invasions of the crab *Carcinus* detected by molecular phylogeography. Mol Ecol 6:901–906

Gilg MR, Hilbish TJ, Bonhomme F (2003) Patterns of larval dispersal and their effect on the maintenance of a blue mussel hybrid zone in southwestern England. Evolution 57:1061–1077

Goodrich DM, van Montfrans J, Orth RJ (1989) Blue crab megalopal influx to Chesapeake Bay: evidence for a wind-driven mechanism. Estuar Coast Shelf Sci 29:247–260

Grey EK, Chiasson SC, Williams HG, Troeger VJ, Taylor CM (2015) Evaluation of blue crab, *Callinectes sapidus*, megalopal settlement and condition during the Deepwater Horizon oil spill. PLOS ONE 10:e0135791

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Hasek BE, Rabalais NN (2001) Settlement patterns of brachyuran megalopae in a Louisiana estuary. Estuaries 24:796–807

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc B 270:313–321

Hebert PD, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLOS Biol 2:e312

Hempel G (1965) On the importance of larval survival for the population dynamics of marine food fish. CalCOFI Rep 10:12–23

Hubert N, Kadarusman, Wibowo A, Busson F and others (2015) DNA barcoding Indonesian freshwater fishes: challenges and prospects. DNA Barcodes 3:144–169

Johnson M, Black R (1982) Chaotic genetic patchiness in an intertidal limpet, *Siphonaria* sp. Mar Biol 70:157–164

Johnson MS, Black R (1984) Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. Evolution 38:1371–1383

Jones M, Epifanio C (1995) Settlement of brachyuran megalopae in Delaware Bay: an analysis of time series data. Mar Ecol Prog Ser 125:67–76

Katouzian AR, Sari A, Macher JN, Weiss M, Saboori A, Leese F, Weigand AM (2016) Drastic underestimation of amphipod biodiversity in the endangered Irano-Anatolian and Caucasus biodiversity hotspots. Sci Rep 6:22507

Knowlton N, Leray M (2015) Exploring coral reefs using the tools of molecular genetics. In: Birkeland C (ed) Coral reefs in the Anthropocene. Springer, Dordrecht, p 117–132

Kordos L, Burton R (1993) Genetic differentiation of Texas Gulf Coast populations of the blue crab *Callinectes sapidus*. Mar Biol 117:227–233

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874

Lacerda ALF, Kersanach R, Cortinhas MCS, Prata PFS and others (2016) High connectivity among blue crab (*Callinectes sapidus*) populations in the western South Atlantic. PLOS ONE 11:e0153124

Large W, Pond S (1981) Open ocean momentum flux measurements in moderate to strong winds. J Phys Oceanogr 11:324–336

Lipcius RN, Stockhausen WT (2002) Concurrent decline of the spawning stock, recruitment, larval abundance, and size of the blue crab *Callinectes sapidus* in Chesapeake Bay. Mar Ecol Prog Ser 226:45–61

Mahon BC, Neigel JE (2008) Utility of arginine kinase for resolution of phylogenetic relationships among brachyuran genera and families. Mol Phylogenet Evol 48: 718–727

McMillen-Jackson A, Bert T (2004) Mitochondrial DNA variation and population genetic structure of the blue crab *Callinectes sapidus* in the eastern United States. Mar Biol 145:769–777

McMillen-Jackson AL, Bert TM, Steele P (1994) Population genetics of the blue crab *Callinectes sapidus*: modest population structuring in a background of high gene flow. Mar Biol 118:53–65

Mense D, Posey M, West T, Kincheloe K (1995) Settlement of brachyuran postlarvae along the North Carolina coast. Bull Mar Sci 57:793–806

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLOS Biol 3:e422

Morgan SG, Zimmer-Faust RK, Heck KL Jr, Coen LD (1996) Population regulation of blue crabs *Callinectes sapidus* in the northern Gulf of Mexico: postlarval supply. Mar Ecol Prog Ser 133:73–88

Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Kailin S, Nevoe E (eds) Evolutionary processes and theory. Academic Press, Cambridge, MA, p 515–534

Neigel J, Mahon B (2007) Molecular approaches in crustacean evolutionary approach. In: Duffy J, Thiel M (eds) Evolutionary ecology of social and sexual systems: crustaceans as model organisms. Oxford University Press Canada, Ontario, p 48–67

Nylander JAA (2004) MrAIC.pl [Perl script]. Program distributed by author. Evolutionary Biology Centre, Uppsala University

Ogburn MB, Diaz H, Forward RB Jr (2009) Mechanisms regulating estuarine ingress of blue crab *Callinectes sapidus* megalopae. Mar Ecol Prog Ser 389:181–192

Ogburn MB, Stuck KC, Heard RW, Wang SY, Forward RB Jr (2011) Seasonal variability in morphology of blue crab, *Callinectes sapidus*, megalopae and early juvenile stage crabs, and distinguishing characteristics among co-occurring Portunidae. J Crustac Biol 31:106–113

Ogburn MB, Hall M, Forward RB Jr (2012) Blue crab (*Callinectes sapidus*) larval settlement in North Carolina: environmental forcing, recruit–stock relationships, and numerical modeling. Fish Oceanogr 21:20–32

Palumbi S (1991) The simple fool's guide to PCR. http://palumbi.stanford.edu/SimpleFoolsMaster.pdf (accessed 25 Jan 2017)

Pardo LM, Ampuero D, Véliz D (2009) Using morphological and molecular tools to identify megalopae larvae collected in the field: the case of sympatric *Cancer* crabs. J Mar Biol Assoc UK 89:481–490

Pawlik JR (1992) Chemical ecology of the settlement of benthic marine invertebrates. Oceanogr Mar Biol Annu Rev 30:273–335

Pawlik JR, Butman A, Starczak VR (1991) Hydrodynamic facilitation of gregarious settlement of a reef-building tube worm. Science 251:421–424

Pechenik JA (1999) On the advantages and disadvantages of larval stages in benthic marine invertebrate life cycles. Mar Ecol Prog Ser 177:269–297

Pereira LH, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genet 14:20

Pernet F, Tremblay R, Bourget E (2003) Settlement success, spatial pattern and behavior of mussel larvae *Mytilus* spp. in experimental 'downwelling' systems of varying velocity and turbulence. Mar Ecol Prog Ser 260:125–140

Perry HM, Eleuterius CK, Trigg CB, Warren JR (1995) Settlement patterns of *Callinectes sapidus* megalopae in Mississippi Sound: 1991, 1992. Bull Mar Sci 57:821–833

Platt AR, Woodhall RW, George AL (2007) Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. Biotechniques 43:58–62

Rabalais NN, Burditt RF Jr, Coen LD, Cole BE and others (1995) Settlement of *Callinectes sapidus* megalopae on artificial collectors in four Gulf of Mexico estuaries. Bull Mar Sci 57:855–876

Raimondi PT, Aileen NCM (2000) The consequences of complex larval behavior in a coral. Ecology 81:3193–3211

Raupach MJ, Barco A, Steinke D, Beermann J and others (2015) The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. PLOS ONE 10:e0139421

Robles R, Schubart CD, Conde JE, Carmona-Suárez C, Alvarez F, Villalobos JL, Felder DL (2007) Molecular phylogeny of the American *Callinectes* Stimpson, 1860 (Brachyura: Portunidae), based on two partial mitochondrial genes. Mar Biol 150:1265–1274

Schneider-Broussard R, Neigel JE (1997) A large-subunit mitochondrial ribosomal DNA sequence translocated to the nuclear genome of two stone crabs (*Menippe*). Mol Biol Evol 14:156–165

Schubart CD, Neigel JE, Felder DL (2000) Use of the mitochondrial 16S rRNA gene for phylogenetic and population studies of Crustacea. Crustac Issues 12:817–830

Schubart CD, Conde JE, Carmona-Suarez C, Robles R, Felder DL (2001) Lack of divergence between 16S mtDNA sequences of the swimming crabs *Callinectes bocourti* and *C. maracaiboensis* (Brachyura: Portunidae) from Venezuela. Fish Bull 99:475–481

Selkoe K, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Mar Ecol Prog Ser 436:291–305

Shokralla S, Porter TM, Gibson JF, Dobosz R and others (2015) Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Sci Rep 5:9687

Sison CP, Glaz J (1995) Simultaneous confidence intervals and sample size determination for multinomial proportions. J Am Stat Assoc 90:366–369

Spitzer PM, Heck KL, Valentine JF (2003) Then and now: a comparison of patterns in blue crab post-larval abundance and post-settlement mortality during the early and late 1990s in the Mobile Bay system. Bull Mar Sci 72:435–452

Stuck K, Perry H (1982) Morphological characteristics of blue crab larvae, *Callinectes sapidus* Rathbun, from the northern Gulf of Mexico. Gulf States Marine Fishery Commission completion report 000-011, Ocean Springs, MS

Stuck KC, Perry HM (1981) Observations on the distribution and seasonality of portunid megalopae in Mississippi coastal waters. Gulf Caribb Res 7:93–95

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10: 512–526

Treml EA, Roberts JJ, Chao Y, Halpin PN, Possingham HP, Riginos C (2012) Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. Integr Comp Biol 52:525–537

Van Montfrans J, Peery CA, Orth RJ (1990) Daily, monthly and annual settlement patterns by *Callinectes sapidus* and *Neopanope sayi* megalopae on artificial collectors deployed in the York River, Virginia: 1985–1988. Bull Mar Sci 46:214–229

Van Montfrans J, Epifanio CE, Knott DM, Lipcius RN and others (1995) Settlement of blue crab postlarvae in Western North Atlantic Estuaries. Bull Mar Sci 57:834–854

Weigelt R, Lippert H, Borges LM, Appelqvist C, Karsten U, Bastrop R (2016) First time DNA barcoding of the common shipworm *Teredo navalis* Linnaeus 1758 (Mollusca: Bivalvia: Teredinidae): molecular-taxonomic investigation and identification of a widespread wood-borer. J Exp Mar Biol Ecol 475:154–162

Williams AB (1984) Shrimps, lobsters, and crabs of the Atlantic coast of the eastern United States, Maine to Florida. Smithsonian Institution Press, Washington, DC

Williams ST, Knowlton N (2001) Mitochondrial pseudogenes are pervasive and often insidious in the snapping shrimp genus *Alpheus*. Mol Biol Evol 18:1484–1493

Witt JD, Threloff DL, Hebert PD (2006) DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation. Mol Ecol 15:3073–3082

Yednock BK, Neigel JE (2014) An investigation of genetic population structure in blue crabs, *Callinectes sapidus*, using nuclear gene sequences. Mar Biol 161:871–886

Editorial responsibility: Steven Morgan, Bodega Bay, California, USA

Submitted: October 24, 2016; Accepted: January 1, 2017
Proofs received from author(s): January 30, 2017