Worldwide phylogeography of the invasive ctenophore *Mnemiopsis leidyi* (Ctenophora) based on nuclear and mitochondrial DNA data

K. M. Bayha · M. H. Chang · C. L. Mariani · J. L. Richardson · D. L. Edwards · T. S. DeBoer · C. Moseley · E. Aksoy · M. B. Decker · P. M. Gaffney · G. R. Harbison · J. H. McDonald · A. Caccone

Received: 8 January 2014 / Accepted: 11 August 2014
© Springer International Publishing Switzerland 2014

Abstract The ctenophore *Mnemiopsis leidyi* is one of the most successful marine bioinvaders on record. Native to the Atlantic coast of the Americas, *M. leidyi* invaded the Black Sea, Caspian and Mediterranean Seas beginning in the late 1980s, followed by the North and Baltic Seas starting in 2006, with major concomitant alterations in pelagic ecology, including fishery collapses in some cases. Using extensive native range sampling (21 sites), along with 11 invasive sites in the Black, Caspian, Mediterranean, North and Baltic Seas, we examined *M. leidyi* worldwide phylogeographic patterns using data from mitochondrial cytochrome b (*cytb*) and six nuclear microsatellite loci. *Cytb* and microsatellite data sets showed different levels of genetic differentiation in the native range. Analyses of *cytb* data revealed considerable genetic differentiation, recovering three major clusters (northwestern Atlantic, Caribbean, and South America) and further divided northwestern Atlantic sampling sites into three groups, separated approximately at Cape Hatteras on the US Atlantic coast and at the Floridian peninsula, separating the Gulf of Mexico and Atlantic coasts. In contrast, microsatellite data only distinguished samples north and south of Cape Hatteras, and suggested considerable gene flow among native samples with clear evidence of isolation by distance. Both

Electronic supplementary material The online version of this article (doi:10.1007/s10530-014-0770-6) contains supplementary material, which is available to authorized users.

K. M. Bayha
Dauphin Island Sea Lab, 101 Bienville Blvd.,
Dauphin Island, AL 36528, USA

K. M. Bayha (✉)
Gulf Coast Research Laboratory, 703 E. Beach Rd.,
Ocean Springs, MS 39564, USA
e-mail: keith.bayha@usm.edu; kmbayha@gmail.com

M. H. Chang · C. L. Mariani · D. L. Edwards · C. Moseley · E. Aksoy · M. B. Decker · A. Caccone
Department of Ecology and Evolutionary Biology, Yale University, PO Box 208106, New Haven, CT 06520, USA

J. L. Richardson
Department of Ecology and Evolutionary Biology,
University of Connecticut, 75 N. Eagleville Rd., Storrs,
CT 06269, USA

T. S. DeBoer
The Ocean Genome Legacy Center of New England Biolabs, Northeastern University, 430 Nahant Road, Nahant, MA 01908, USA

P. M. Gaffney
College of Earth, Ocean and Environment, University of Delaware, 700 Pilottown Rd., Lewes, DE 19958, USA

G. R. Harbison
Woods Hole Oceanographic Institute, 266 Woods Hole Rd., Woods Hole, MA 02543, USA

J. H. McDonald
Department of Biological Sciences, University of Delaware, Wolf Hall, Newark, DE 19716, USA

Published online: 31 August 2014
cytb and microsatellite data sets indicated that the northern invaders (North/Baltic Seas) originated from north of Cape Hatteras, with cytb data pointing to Delaware and north. Microsatellite data indicated a source for the southern invaders (Black, Caspian and Mediterranean Seas) to be south of Cape Hatteras, while cytb data narrowed the source location to the Gulf of Mexico region. Both cytb and microsatellite data sets suggested that the southern invasion was associated with genetic bottlenecks while evidence was equivocal for the northern invasion. By increasing the native range spatial sampling, our dataset was able to sufficiently characterize patterns and levels of genetic differentiation in the native range of *M. leidyi* and identify likely biogeographic boundaries, allowing for the most complete characterization of *M. leidyi*’s invasion histories and most realistic estimates of its source region(s) to date.

**Keywords**  Marine bioinvasions · Ctenophore · Microsatellites · Cytochrome b · Genetic diversity

**Introduction**

Marine bioinvasions have increased notably over the past century, especially along major trade routes and shipping corridors, as global ship traffic has increased (Carlton and Geller 1993; Ruiz et al. 2000). In recipient regions, invaders are an increasing threat to native biodiversity (Bax et al. 2003) and overall ecosystem functioning (Crooks 2002) and inflict losses in the billions of dollars per year (Pimental et al. 2005). One of the main culprits for the global dispersal of marine organisms has been ballast water of ocean going vessels (Carlton and Geller 1993; Wonham et al. 2001). Adequate understanding of invasion histories is vital to mitigating the effects of established invaders, as well as to predicting and preventing future invasions, whether this be through the identification of vectors, source regions and possible future invasion success, or through the simple identification of ongoing invasions (Mack et al. 2000; Kolar and Lodge 2002; Simberloff 2005; Lodge et al. 2006; Wilson et al. 2009).

The use of molecular genetic techniques has been instrumental in clarifying invasion histories of many marine organisms by genetically comparing native and invasive populations. Such techniques have been employed to identify cryptic invasions and determine if multiple cryptic species are present (Geller et al. 1997; Dawson and Jacobs 2001; Holland et al. 2004; Schuchert 2014), to determine invader source region(s) and likely invasion vector(s) (Brown and Stepien 2008; Pineda et al. 2011), to evaluate whether or not multiple invasions have occurred (Brown and Stepien 2008; Audzijonyte et al. 2009), to assess the degree and impacts of any founder events due to small invasive populations (Holland 2001; Hamner et al. 2007) and to identify the presence and effects of hybridization between native and invasive species (Yara et al. 2010; Muhlfeld et al. 2014).

Since the conclusions of any thorough invasion genetics study are based upon the genetic comparison of invasive animals to genetic structure in the native range, these studies can be quite susceptible to sampling regimes. Muirhead et al. (2008) empirically demonstrated that extensive sampling in the native and invasive ranges, both geographically by number of individuals, is extremely important since insufficient sampling can produce ambiguous results when the native range is poorly sampled. For an organism with multiple genetically isolated populations, the correct source region(s) might not be sampled if few geographic sites are examined, whereas the same genetic data could pinpoint an erroneous definitive source region in the case of a species with widely dispersed populations. Where studies have included extensive sampling of native ranges (Stepien et al. 2002; Darling et al. 2008), estimations of source region(s) and general analyses of invasion attributes are more precise, though logistics, expense, and the rarity or patchiness of the organism can make this difficult.

The ctenophore *Mnemiopsis leidyi* is one of the most successful marine bioinvaders on record. *M. leidyi* is native to estuarine and coastal waters from Massachusetts in the USA to the southern coast of Argentina (GESAMP 1997) and, likely as a result of a high reproductive rate (Reeve et al. 1989) and predatory capabilities (Kremer 1979; Purcell and Decker 2005), can exert significant predation pressure on zooplankton populations (Feigenbaum and Kelly 1984; Purcell and Decker 2005) and economically impact important fish and shellfish (Purcell et al. 1991, 1994).

Over the past 30 years, *M. leidyi* has invaded multiple regions of Eurasia with concomitant ecological and economic alterations (Vinogradov and
Shushkina 1992; GESAMP 1997; Purcell et al. 2001). First observed in the Black Sea in 1982 (Pereladov 1988), *M. leidyi* had rapidly expanded throughout the Black Sea and Sea of Azov by 1987 (Vinogradov et al. 1989), and subsequently was found in various waters of the Mediterranean basin (Shiganova 1993; Kideys and Niermann 1993, 1994; Shiganova 1997) and the Caspian Sea (Ivanov et al. 2000). By the 2000s, *M. leidyi* spanned the entirety of the Mediterranean Sea (Fuentes et al. 2009; Boero et al. 2009; Galil et al. 2009; Shiganova and Malej 2009; Fuentes et al. 2010). It was first found in northern Europe in 2006, both in the North Sea (Faasse and Bayha 2006; Hansson 2006; Boersma et al. 2007) and the Baltic Sea (Javidpour et al. 2006), subsequently spreading to other areas of both seas, as well as adjacent water bodies (Oliveira 2007; Tendal et al. 2007; Janas and Zgrundo 2007; Van Ginderdeuren et al. 2012; Antajan et al. 2014).

Four genetic studies have examined the invasion history of *Mnemiopsis* in Eurasia and all suggest the occurrence of multiple invasions. Reusch et al. (2010) employed microsatellite markers to determine the *M. leidyi* source region(s), while Ghabooli et al. (2011) examined sequence data from the nuclear internal transcribed spacer (ITS) regions (ITS). Both studies suggest that the *M. leidyi* samples from the Black/Caspian/Mediterranean Sea and the North/Baltic Sea areas derive from separate invasions. Reusch et al. (2010) and Ghabooli et al. (2011) concluded that the Gulf of Mexico was the source for the Black Sea invasion while the northern *M. leidyi* European populations derived from samples at the northern extreme of the ctenophore’s range, in New England. A third study using microsatellite data (Bolte et al. 2013) indicated that the Mediterranean Sea regions were likely inoculated with animals from the Black Sea. Recently, Ghabooli et al. (2013) examined variation in both mitochondrial cytochrome c oxidase I (COI) and nuclear ITS regions, indicating the possibility that Mediterranean populations may have been the result of independent invasion(s) in addition to introduction from the Black Sea. In most of these studies, the native *M. leidyi* range was not extensively sampled, as the Reusch et al. (2010) samples were all located in US waters (three in the Gulf of Mexico and one in New England) and Ghabooli et al. (2011, 2013) included three sites along the US Atlantic coast, one in the Gulf of Mexico, and one along the Argentinian coast in South America.

Given that *M. leidyi* is widely distributed latitudinally, inhabiting waters from 41°N to 44°S (GESAMP 1997), reconstructing its invasion history should rely on a sampling regime in the native range that takes into account the geographic and ecological diversity of this species. The native range of *M. leidyi* encompasses eight different large marine ecosystems (LME) (Sherman 2008) divided by biogeographic fronts known to be able to separate marine populations, such as Cape Hatteras (Avise et al. 1987; Graves et al. 1992; Jones and Quattro 1999) and the Floridian peninsula (Avise 2000). However, the ability of biogeographic fronts to separate populations is dependent upon an organism’s dispersal potential and thermal tolerance (Jones and Quattro 1999; Avise 2000), both of which are likely extensive in *Mnemiopsis*. Since *M. leidyi* is holoplanktonic, a simultaneous hermaphrodite and eurytopic, inhabiting an extremely wide range of temperatures and salinities (GESAMP 1997), its dispersal potential and environmental tolerances may overcome some or most of these fronts. Thus, extensive geographic sampling of the native range may be necessary in order to adequately uncover native range population structure that could lead to a better understanding of the source region(s) of the invasive populations.

In this paper, we have increased the breadth of earlier genetic analyses performed on *M. leidyi* by analyzing the genetic variation of this species at 32 sampling sites within both the native and invasive ranges. We used both nuclear (six microsatellite loci) and mitochondrial (mtDNA) sequences (a fragment of cytb) to take advantage of the complementary insights that these two different types of genetic markers may offer (Garrick et al. 2010). Here, we describe the spatial patterns of genetic diversity and the demographic dynamics of both native and invasive populations to gain insights into the patterns and levels of genetic differentiation among native samples and to reconstruct colonization routes and underlying evolutionary processes.

**Materials and methods**

**Sample collection**

*Mnemiopsis leidyi* ctenophores were collected from a total of 32 geographic sites (21 in the native range and 11 in the invasive range) (Table 1; Fig. 1). For two
Table 1 Sampling locations for *M. leidyi* collections, with diversity indices for *cytb* sequence data (N, sample size; \(N_h\), number of haplotypes; \(\theta\), haplotype diversity; \(\pi\), nucleotide diversity) and data from 6 microsatellite loci (N, sample size; \(A_n\), allelic richness; \(H_e\), expected heterozygosity; \(H_0\), observed heterozygosity; \(F_{IS}\), inbreeding coefficients)

<table>
<thead>
<tr>
<th>Location</th>
<th>Site code</th>
<th>Year</th>
<th>Lat.</th>
<th>Long.</th>
<th>mtDNA</th>
<th>Microsatellites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>(N_h)</td>
<td>(\theta)</td>
<td>(\pi)</td>
</tr>
<tr>
<td><strong>Native populations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woods Hole, MA*</td>
<td>WH99</td>
<td>1999</td>
<td>41.53N</td>
<td>70.67W</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Woods Hole, MA*</td>
<td>WH10</td>
<td>2010</td>
<td>41.53N</td>
<td>70.67W</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Woods Hole, MA*</td>
<td>WH11</td>
<td>2011</td>
<td>41.53N</td>
<td>70.67W</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Narragansett Bay, RI*</td>
<td>NRI</td>
<td>2001</td>
<td>41.49N</td>
<td>71.42W</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Niantic Bay, CT*</td>
<td>NCT</td>
<td>1999</td>
<td>41.31N</td>
<td>72.20W</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Long Island Sound, CT*</td>
<td>LIS</td>
<td>2010</td>
<td>41.25N</td>
<td>72.90W</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Rehoboth Bay, DE*</td>
<td>RBD</td>
<td>1998</td>
<td>38.69N</td>
<td>75.08W</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Rhode River, MD*</td>
<td>RRM</td>
<td>1998</td>
<td>38.89N</td>
<td>76.54W</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>N. Chesapeake Bay, MD*</td>
<td>CBN</td>
<td>2010</td>
<td>38.52N</td>
<td>76.40W</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>S. Chesapeake Bay, MD*</td>
<td>CBS</td>
<td>2010</td>
<td>37.73N</td>
<td>76.20W</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Open Atlantic</td>
<td>ATL</td>
<td>2002</td>
<td>39.00N</td>
<td>69.00W</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>N. Pamlico Sound, NC*</td>
<td>PAM</td>
<td>1999</td>
<td>35.91N</td>
<td>75.66W</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>Charleston, SC*</td>
<td>CHS</td>
<td>1999</td>
<td>32.78N</td>
<td>79.95W</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Fort Pierce, FL*</td>
<td>FPF</td>
<td>1999</td>
<td>27.46N</td>
<td>80.31W</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Miami Beach, FL*</td>
<td>MIA</td>
<td>1999</td>
<td>25.81N</td>
<td>80.12W</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Tampa Bay, FL*</td>
<td>TAM</td>
<td>1999</td>
<td>27.90N</td>
<td>82.49W</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Mobile Bay, AL*</td>
<td>MOB</td>
<td>1999</td>
<td>30.57N</td>
<td>88.09W</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Dauphin Island, AL*</td>
<td>DIA</td>
<td>2011</td>
<td>30.26N</td>
<td>88.08W</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>Biloxi, MS*</td>
<td>BMS</td>
<td>1999</td>
<td>30.39N</td>
<td>88.88W</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>Port Aransas, TX*</td>
<td>PAT</td>
<td>2001</td>
<td>27.84N</td>
<td>97.05W</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Belize City, Belize</td>
<td>BEZ</td>
<td>2009</td>
<td>17.50N</td>
<td>88.11W</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>Rio de Janeiro, Brazil</td>
<td>RIO98</td>
<td>1998</td>
<td>22.95S</td>
<td>43.17W</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Rio de Janeiro, Brazil</td>
<td>RIO99</td>
<td>1999</td>
<td>22.94S</td>
<td>43.16W</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Mar del Plata, Argentina</td>
<td>ARG</td>
<td>1999</td>
<td>38.04S</td>
<td>57.53W</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
### Table 1 continued

<table>
<thead>
<tr>
<th>Location</th>
<th>Site code</th>
<th>Year</th>
<th>Year of initial record</th>
<th>Lat.</th>
<th>Long.</th>
<th>N</th>
<th>N_h</th>
<th>θ</th>
<th>π</th>
<th>N</th>
<th>A_n</th>
<th>H_e</th>
<th>H_0</th>
<th>F_{IS}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive populations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denia, Spain</td>
<td>DNS</td>
<td>2010</td>
<td>2009&lt;sup&gt;a&lt;/sup&gt;</td>
<td>38.85N</td>
<td>0.10E</td>
<td>18</td>
<td>2</td>
<td>0.1111</td>
<td>0.0003</td>
<td>28</td>
<td>7.67</td>
<td>16.75</td>
<td>9.67</td>
<td>0.42</td>
</tr>
<tr>
<td>Marseille, France</td>
<td>MRF</td>
<td>2007</td>
<td>2006&lt;sup&gt;b&lt;/sup&gt;</td>
<td>43.40N</td>
<td>5.07E</td>
<td>6</td>
<td>2</td>
<td>0.5333</td>
<td>0.0015</td>
<td>5</td>
<td>4.67</td>
<td>2.97</td>
<td>2.00</td>
<td>0.35</td>
</tr>
<tr>
<td>Lerici, Italy</td>
<td>ITL</td>
<td>2011</td>
<td>2006&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>44.10N</td>
<td>9.82E</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>38</td>
<td>9.83</td>
<td>24.09</td>
<td>18.5</td>
<td>0.2</td>
</tr>
<tr>
<td>N. Evoikos Gulf, Greece</td>
<td>NEG</td>
<td>2003</td>
<td>1990&lt;sup&gt;d&lt;/sup&gt;</td>
<td>38.12N</td>
<td>23.02E</td>
<td>12</td>
<td>3</td>
<td>0.3455</td>
<td>0.0010</td>
<td>27</td>
<td>11.17</td>
<td>17.91</td>
<td>10.33</td>
<td>0.43</td>
</tr>
<tr>
<td>Limnos Island, Greece</td>
<td>LIG</td>
<td>2004</td>
<td>1991&lt;sup&gt;d&lt;/sup&gt;</td>
<td>40.54N</td>
<td>25.42E</td>
<td>15</td>
<td>5</td>
<td>0.6286</td>
<td>0.0028</td>
<td>20</td>
<td>11.17</td>
<td>13.96</td>
<td>9.83</td>
<td>0.34</td>
</tr>
<tr>
<td>Istanbul, Turkey</td>
<td>IST</td>
<td>1999</td>
<td>1989&lt;sup&gt;d&lt;/sup&gt;</td>
<td>41.02N</td>
<td>28.98E</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3</td>
<td>3.67</td>
<td>2.32</td>
<td>2.17</td>
<td>0.14</td>
</tr>
<tr>
<td>Southwest Black Sea</td>
<td>SWBS</td>
<td>2001</td>
<td>1986&lt;sup&gt;e&lt;/sup&gt;</td>
<td>41.99N</td>
<td>29.58E</td>
<td>7</td>
<td>3</td>
<td>0.667</td>
<td>0.0034</td>
<td>6</td>
<td>5.67</td>
<td>4.14</td>
<td>3.00</td>
<td>0.31</td>
</tr>
<tr>
<td>Gelendzhik Bay, Russia</td>
<td>GBR</td>
<td>2002</td>
<td>1982&lt;sup&gt;f&lt;/sup&gt;</td>
<td>44.58N</td>
<td>38.04E</td>
<td>4</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>4</td>
<td>4.17</td>
<td>2.72</td>
<td>2.67</td>
<td>0.02</td>
</tr>
<tr>
<td>Caspian Sea, Azerbaijan</td>
<td>CAS</td>
<td>2001</td>
<td>1997&lt;sup&gt;g&lt;/sup&gt;</td>
<td>40.33N</td>
<td>49.90E</td>
<td>7</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>8</td>
<td>3.33</td>
<td>4.02</td>
<td>3.83</td>
<td>0.05</td>
</tr>
<tr>
<td>Kiel, Germany</td>
<td>KBG</td>
<td>2007</td>
<td>2006&lt;sup&gt;b&lt;/sup&gt;</td>
<td>54.35N</td>
<td>10.16E</td>
<td>11</td>
<td>5</td>
<td>0.7818</td>
<td>0.0033</td>
<td>10</td>
<td>8.67</td>
<td>7.97</td>
<td>6.83</td>
<td>0.24</td>
</tr>
<tr>
<td>Netherlands</td>
<td>NTH</td>
<td>2007</td>
<td>2005&lt;sup&gt;i&lt;/sup&gt;</td>
<td>52.48N</td>
<td>4.56E</td>
<td>9</td>
<td>5</td>
<td>0.8056</td>
<td>0.0039</td>
<td>9</td>
<td>6.67</td>
<td>6.71</td>
<td>5.67</td>
<td>0.21</td>
</tr>
</tbody>
</table>

An asterisk identifies sites from US coastal waters for which abbreviations are included in the location column

<sup>a</sup> Fuentes et al. (2009)
<sup>b</sup> Shiganova and Malej (2009)
<sup>c</sup> Boero et al. (2009)
<sup>d</sup> Shiganova et al. (2001)
<sup>e</sup> Konsulov (1990)
<sup>f</sup> Pereladov (1988)
<sup>g</sup> Ivanov et al. (2000)
<sup>b</sup> Javidpour et al. (2006)
<sup>i</sup> Faasse and Bayha (2006)
collections sites [Woods Hole (USA) and Rio de Janeiro (Brazil)], samples were collected over multiple years, while the Chesapeake Bay (USA) was sampled over multiple years from hydrographically distinct but nearby regions [Rhode River (USA), and N. and S. Chesapeake Bay (USA)]. Specimens were collected dockside using dip nets or dipping cups, by SCUBA divers or snorkelers using plastic containers, or by various plankton nets. All animals were either frozen on dry ice and kept at −20 °C or preserved in 70–95 % ethanol or DMSO/salt buffer (Dawson et al. 1998). Samples of frozen animals were subsequently lyophilized. For both lyophilized tissue and preserved tissue, DNA was extracted either via a CTAB/chloroform method (Ausubel et al. 1989) or using the Qiagen DNAeasy Blood and Tissue Kit (Qiagen Inc.). DNA extracts were stored at −4 °C. For samples from Greece (LIG and NEG), DNA extractions were carried out using a standard phenol/chloroform extraction method (Sambrook et al. 1989).

Mitochondrial cytochrome b sequence data

A 361-base pair portion of the mitochondrial cytochrome b region (cytb) was amplified using primers KMBMT-80 and KMBMT-116 (Bayha 2005). Polymerase Chain Reaction (PCR) conditions consisted of 120 s (s) at 94 °C, followed by 38 cycles of 94 °C for 45 s, 52 °C for 60 s and 72 °C for 90 s, with a final cycle of 600 s at 72 °C. Successful amplifications were evaluated by running PCR products on 2 % agarose gels. PCR products were purified either using the AMPure magnetic plate PCR clean-up protocol (Agencourt Bioscience Corporation) or by treating with EXOSAP-it (USB Corporation). Cleaned PCR products were cycle sequenced using Primer

---

**Fig. 1** Sampling locations for *M. leidyi* and haplogroup frequencies for cytb. Sampling site codes refer to Table 1. Map A shows a close-up of the locations of sampling sites in the vicinity of US coastal waters. An asterisk denotes an idealized representation of Gulf Stream flow that has been hypothesized to drive a biogeographic front at Cape Hatteras (Avise et al. 1987; Graves et al. 1992; Jones and Quattro 1999). Map B shows sampling sites for invasive samples in Eurasian waters. Circles represent individual sampling sites/times and are colored according to the proportion of individuals containing cytb haplogroups (groups of related haplotypes), as defined in Fig. 2. Stars represent sampling sites included in the microsatellite but not in the cytb analysis.
KMBMT-80 and KMBMT-118 (Bayha 2005) with Big Dye Terminator Version 3.1 (Applied Biosystems, Inc., ABI) on an ABI 377 or Spectremedix SCE2410DNA (Spectremedix, L.L.C.) sequencer. Alternatively, both PCR clean up and cycle sequencing were performed by Beckman-Coulter Genomics (Beverly, MA). Sequences were assembled using Lasergene SeqMan Pro v. 9.11 (DNASTar, Inc.), with electropherograms checked visually and poorly resolved terminal regions removed. Sequences were compared to GENBANK nucleotide database using BLASTX (Altschul et al. 1997) to confirm identity of the region, aligned using ClustalX v2.0 (Larkin et al. 2007), and checked for errors by eye based on amino acid sequence, translated using the Mold, Protozoan and Coelenterate Mitochondrial Code (Translation Code 4).

The number of haplotypes, haplotype diversity ($\theta$) and nucleotide diversity ($\pi$), were computed using Arlequin v. 3.5.1.3 (Excoffier and Lischer 2010). A statistical parsimony network was generated using TCS 1.2.1 (Clement et al. 2000), in order to examine phylogenetic similarities of cytb haplotypes and assess relationship with their geographical distribution and indicate possible source region(s) of invaders. The network was redrawn using Adobe Illustrator CS5 v. 15.0.2 (Adobe, Inc.). The genetic divergence among haplotypes was determined using the Kimura 2-parameter model in MEGA 5.2.2 (Tamura et al. 2011). In order to evaluate genetic differences between pairs of populations, pairwise $\Phi_{st}$ values were calculated in Arlequin 3.5.1.3 (10,000 permutations), with significance levels adjusted using the false discovery rate method (Benjamini and Hochberg 1995). A multidimensional scaling (MDS) analysis was carried out in the program SYSTAT v.11 (SYSTAT Software Inc.), using the matrix of average pairwise uncorrected $p$ distances, to graphically represent patterns of genetic differentiation among sampling sites and indicate population clusters and genetic relatedness among native and invasive sampling sites. $p$-distance is preferred over $\Phi_{st}$ when distances are small (Nei and Kumar 2000). The hierarchical distribution of genetic variation among sampling sites was tested using Analysis of Molecular Variance (AMOVA) in Arlequin v. 3.5.1.3 (Excoffier and Lischer 2010), using 10,000 random permutations. Three AMOVA tests were run: (A) native sites only, (B) invasive sites only, (C) subset of paired native and invasive sites [“Northern” (native sites from North of Cape Hatteras and Northern Europe sites) vs. “Southern” (native sites south of Cape Hatteras and Southern Europe sites)].

Genetic diversity indices and neutrality tests were calculated to look for evidence of genetic bottlenecks, which often accompany biological invasions (Allendorf and Lundquist 2003). Average haplotype diversity values ($\theta$) were compared between invasive regions and their likeliest source regions based on phylogeographic and population genetic analyses using Welch’s $t$ tests. As the frequency distribution of DNA sequence haplotypes can carry signal of growth or decline, which may be indicative of genetic bottlenecks, we also evaluated evidence for past demographic changes by calculating Tajima’s $D$ (Tajima 1989), Fu’s $F_S$ (Fu 1997) and $R_2$ (Ramos-Onsins and Rozas 2002) in DNAsp v.5 (Librado and Rozas 2009). We used all three estimators in order to be comprehensive, though simulations have shown that $R_2$, a metric based on segregating site frequency distribution, is superior to $F_S$ when sample sizes are small (Ramos-Onsins and Rozas 2002). Deviation from the null hypothesis of constant size was evaluated by comparing observed values against distributions simulated via neutral coalescence ($2.0 \times 10^3$ replicates), with significance assessed at the 0.05 level (lower tail). To examine genetic changes associated with invasions, we also compared average values of these estimators between invasive and putative source regions via Welch’s $t$ tests. Analyses of the distribution of pairwise sequence differences for the cytb data (mismatch distributions) were carried out for each population in Arlequin 3.5.1.3 (Excoffier and Lischer 2010), using Harpending’s raggedness index (Rogers and Harpending 1992) to test the empirical mismatch distributions for significant deviation from a model of demographic growth (10,000 permutations). As this method assumes demographic growth (Rogers and Harpending 1992), the null hypothesis provides an opportunity to assess strength of evidence for long-term stability in population size, which is typically characterized by a multimodal or ragged mismatch distribution.

Nuclear microsatellite data

DNA extractions were PCR amplified at six species-specific, highly polymorphic microsatellite loci developed previously (Reusch et al. 2010), using the same
amplification parameters. Amplified fragments were run on an Applied Biosystems 3730xl Genetic Analyzer, with Genescan Liz-500 size standard. Allele peak sizes were checked by eye using Genemarker v2.4.0 (SoftGenetics LLC). Alleles were then binned with Tandem v1.09 (Matschiner and Salzburger 2009), using default settings from raw allele size information for consistency (Idury and Cardon 1997).

Microsatellite data were screened by locality for homozygote excess, stuttering effects and null alleles using MICROCHECKER (van Oosterhout et al. 2004). Further screening of microsatellite loci was undertaken using exact tests of Hardy–Weinberg Equilibrium (HWE) and tests of linkage disequilibrium performed in GENEPOP v4.0.10 (Raymond and Rousset 1995). Markov chain parameters were set at 10,000 dememorizations, 1,000 batches and 10,000 iterations per batch. Locus and locality specific estimates of microsatellite allele frequencies were generated using GenAlex v6.4 (Peakall and Smouse 2006). Arlequin v3.5 (Excoffer and Lischer 2010) was used to calculate allelic richness (A), observed (H_o) and expected (H_e) heterozygosity of populations. Welch’s two-sample t tests were used to compare allelic richness and heterozygosity between native and invasive sites and between a few samples for which temporal collections from the same sampling location were available [Woods Hole (USA) and Rio de Janeiro (Brazil)], after confirming that the data follow a normal distribution using a Shapiro–Wilk normality test. All comparisons were undertaken using mean location estimates of allelic richness and heterozygosity in R (R-Core-Team 2013).

We calculated the Weir and Cockerham (1984) estimate of $F_{IS}$ using FSTAT (Goudet 1995) to calculate site specific inbreeding coefficients ($F_{is}$) and used the same program to calculate overall genetic differentiation among different localities and between each set of localities by estimating pairwise $F_{ST}$ values and conducting pairwise tests of differentiation with significance levels adjusted using the false discovery rate method (Benjamini and Hochberg 1995). We also used an analysis of molecular variance (AMOVA), as implemented in Arlequin v3.5 (Excoffer and Lischer 2010), running the following tests: (A) native sites only, (B) invasive sites only, (C) subset of paired native and invasive sites [“Northern” (native sites from North of Cape Hatteras and Northern Europe sites) vs. “Southern” (native sites south of Cape Hatteras to Belize and Southern Europe sites)]. Temporal samples from the same sampling sites that were not genetically distinct were pooled in subsequent analyses.

We tested for isolation by distance (IBD) in native and invasive samples using Mantel tests of matrix correlation with 10,000 permutations within the vegan package of R (Oksanen et al. 2013) using pairwise genetic ($F_{st}$ values) and geographic distances between sampling sites (i.e. distance between sites along the coastline). Mantel tests account for the non-independence of pairwise matrix elements and assess the significance of correlations by permutation.

To evaluate patterns of genetic differentiation among samples from the native range we used Bayesian clustering as implemented in the program STRUCTURE (Pritchard et al. 2000). We tested the most appropriate value for $K$, the number of clusters, for $K = X–Y$. After a burn-in period of 100,000 we ran 1,000,000 iterations. Each parameter set was replicated 40 times. We used STRUCTURE HARVESTER (Earl 2012) to process the results across replicates, and the compiled log likelihood and $\Delta K$ (Evanno et al. 2005) values to determine the most supported value of $K$ for our dataset. CLUMPP and DISTRUCT were used to visualize the results across replicated runs from STRUCTURE (Rosenberg 2004; Jakobsson and Rosenberg 2007).

We tested for the presence of founder effects by conducting bottleneck analyses on sites from the invasive range with sample sizes $\geq 20$ using BOTTLENECK v1.2.02 (Piry et al. 1999) using both the infinite alleles (IAM) (Maruyama and Fuerst 1985) and the stepwise mutation models (SMM) (Cornuet and Luikart 1996). This approach contrasts loss of rare alleles versus loss of heterozygosity, and can provide insights into size changes that occurred within the past $\sim 10$ generations (Peery et al. 2012). To generate distributions of the difference between expected heterozygosity ($H_E$) and expected heterozygosity under mutation-drift equilibrium ($H_{EQ}$), given the empirical number of alleles per locus per population, we used 1,000 simulation iterations (Piry et al. 1999). The null hypothesis of no significant heterozygosity excess was assessed using Wilcoxon’s sign rank test (one-tailed) at the 0.05 level.

In order to examine the origins of the European invasions, STRUCTURE analyses were run to assign individuals from the invasive sites to source locations.
within the native range. We used as reference only individuals from the native range with very high likelihood to be assigned to one of the two clusters identified by the native-only STRUCTURE analyses, as expressed by Q values ≥0.9 (Q is the probability of assignment to a given cluster and ranges from 0 to 1) (Pritchard et al. 2000). While cluster membership was defined for the native samples, it was left undefined for the invasive ones, so that these samples could be assigned to any of the genetic clusters identified in the native samples. STRUCTURE parameters for these runs were the same as the ones for the native samples analyses. STRUCTURE analyses were also carried out only on the Black, Caspian, and Mediterranean Seas samples to investigate patterns of genetic diversity of this invasion. The availability of only two samples for the other invasion (Baltic/North Seas) precluded this analysis.

We used simple linear regression as implemented in the base package within R (R-Core-Team 2013) to evaluate if there was a statistically significant decline in genetic diversity (allelic diversity and heterozygosity) in the invasive samples along a westward transect from the oldest sightings in Black Sea to the most recent ones in the western Mediterranean. We used the years from the first sighting of *M. leidyi* at each invasive sampling site as a proxy for the time elapsed since *M. leidyi* arrival at that site, a minimal estimate given that the lag-time between the species arrival and establishment. Given that sites differed in sampling size, allelic diversity was scaled to sample sizes using HP-Rare (Kalinowski 2004). The dates of first detection of *M. leidyi* were obtained from published reports and accounts (Table 1).

Results

Mitochondrial DNA

Partial *cytb* sequences were obtained for a total of 383 individuals from 30 collection sites, including 21 in the native range and 9 in invasive ranges. After primer sequences were removed, the aligned data set consisted of 361 base pairs (bp). Of the 67 polymorphic sites, 45 were parsimony informative, with 52 synonymous and 15 non-synonymous, resulting in 15 amino acid substitutions. A total of 74 haplotypes occurred in the *cytb* data set, with sequence divergences ranging from 0.3 to 6.2 %. All *cytb* haplotypes were submitted to NCBI GenBank and can be found under accession numbers KM035326-KM035399.

Hierarchical AMOVA based on mitochondrial *cytb* data for native range *M. leidyi* sampling sites indicated significant population structure, as the majority of genetic variation (77 %; p < 0.001) occurred among sampling sites (Table 2A). Significant native range population structure is borne out in the TCS network (Fig. 2), which recovered three major genetic clusters, one haplogroup (a group of closely related haplotypes) including all US samples (Haplogroup 1), one containing haplotypes from Belize (BEZ) samples (Haplogroup 2), and one with haplotypes from South American sites (Haplogroup 3). The minimum number of mutational steps between haplogroups ranged from 12 to 22 and average sequence divergences (Kimura 2-parameter) between haplogroups ranged from 4.6 % to 5.0 %.

Within Haplogroup 1, the *cytb* data recovered two sub-haplogroups (1A and 1B), defined by the two most common haplotypes (3 and 16), which are separated by a single substitution (Fig. 2). Several additional haplotypes stem from these two, creating the two major sub-haplogroups. In the native range, most (92.9 %) of haplotypes in Sub-haplogroup 1A were only found north of Cape Hatteras (WH to PAM), while the majority of haplotypes from Sub-haplogroup 1B (92.5 %) were found only south of Cape Hatteras along the Atlantic coast and into the Gulf of Mexico (CHS to PAT) (Figs. 1, 2, Table S-1). Samples from the open Atlantic (ATL) contained haplotypes from Sub-haplogroup 1B. Haplogroup 3 (ancestral haplotype of Sub-haplogroup 1A) was encountered most frequently north of Cape Hatteras (WH to PAM), though it was also found infrequently in the Gulf of Mexico (TAM, MOB, BMS) and CHS (Figs. 1, 2; Table S1). Haplotype 16 was found exclusively from south of Cape Hatteras to the Gulf of Mexico (CHS to PAT) (Figs. 1, 2; Table S-1).

Further structure was revealed by pairwise \( \Phi_{st} \) values. When comparing pairs of sites from US, Caribbean and South American regions, pairwise \( \Phi_{st} \) values between sites from different regions ranged from 0.835 to 0.970 (average = 0.895; Table S-2), indicating that US, Caribbean, and South American populations are highly diverged. For US coastal sites, pairwise \( \Phi_{st} \) values within this region generally indicated separation into three groups: north of Cape
Hatteras (WH to PAM), Cape Hatteras to southern Florida (CHS to MIA) and Gulf of Mexico (TAM to PAT). Pairwise $\Phi_{st}$ values comparing Atlantic sites north and south of Cape Hatteras ranged from 0.270 to 0.578 (average $\Phi_{st} = 0.435$), while $\Phi_{st}$ values comparing Atlantic sites south of Cape Hatteras and the Gulf of Mexico ranged from 0.011 to 0.226 (average $\Phi_{st} = 0.109$). All pairwise $\Phi_{st}$ values were significant for comparisons between sites north of Cape Hatteras and any site south of it, while 13 out of 18 pairwise comparisons were significant when comparing Atlantic sites south of Cape Hatteras (CHS to MIA) and Gulf of Mexico sites from the Black, Caspian and Mediterranean basins (Table 1: DNS to CAS) and native sampling sites from the Gulf of Mexico (Table 1: TAM to PAT). “Southern” includes invasive sampling sites from the North Sea (NTH) and the Baltic Sea (KBG) and native sampling sites from north of Cape Hatteras (Table 1: WH to PAM).

Table 2  Analysis of molecular variance (AMOVA) results showing partitioning of genetic variation under three different groupings for cytb

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>$\Phi$-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) All native sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among sites</td>
<td>22</td>
<td>713.315</td>
<td>2.49689</td>
<td>77.48</td>
<td>0.77477***</td>
</tr>
<tr>
<td>Within sites</td>
<td>271</td>
<td>196.708</td>
<td>0.72586</td>
<td>22.52</td>
<td></td>
</tr>
<tr>
<td>(B) All invasive sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among sites</td>
<td>8</td>
<td>19.759</td>
<td>0.22040</td>
<td>39.23</td>
<td>0.39229***</td>
</tr>
<tr>
<td>Within sites</td>
<td>80</td>
<td>27.315</td>
<td>0.34143</td>
<td>60.77</td>
<td></td>
</tr>
<tr>
<td>(C) “Northern” versus “Southern”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among groups</td>
<td>1</td>
<td>60.656</td>
<td>0.35940</td>
<td>36.64</td>
<td>$\Phi_{ct} = 0.36640$</td>
</tr>
<tr>
<td>Among sub-groups within groups</td>
<td>2</td>
<td>3.092</td>
<td>0.01532</td>
<td>1.56</td>
<td>$\Phi_{lk} = 0.02464**$</td>
</tr>
<tr>
<td>Within sites</td>
<td>329</td>
<td>199.430</td>
<td>0.60617</td>
<td>61.80</td>
<td>$\Phi_{st} = 0.38202***$</td>
</tr>
</tbody>
</table>

Asterisks indicate significant $\Phi$-statistics (* $<0.05$, ** $<0.01$, *** $<0.001$). A: All native sites from the Americas ungrouped (Table 1: WH to ARG); B: All invasive sites ungrouped (Table 1: DNS to NTH); C: Native and invasive populations pooled into two groups (“Southern” and “Northern”) based on likeliest source regions from phylogeographic and population genetic evidence. “Southern” includes all invasive sampling sites from the Black, Caspian and Mediterranean basins (Table 1: DNS to CAS) and native sampling sites from the Gulf of Mexico (Table 1: TAM to PAT). “Northern” includes invasive sampling sites from the North Sea (NTH) and the Baltic Sea (KBG) and native sampling sites from north of Cape Hatteras (Table 1: WH to PAM).

Hatteras (WH to PAM), Cape Hatteras to southern Florida (CHS to MIA) and Gulf of Mexico (TAM to PAT). Pairwise $\Phi_{st}$ values comparing Atlantic sites north and south of Cape Hatteras ranged from 0.270 to 0.578 (average $= 0.435$), while $\Phi_{st}$ values comparing Atlantic sites south of Cape Hatteras and the Gulf of Mexico ranged from 0.011 to 0.226 (average $= 0.109$). All pairwise $\Phi_{st}$ values were significant for comparisons between sites north of Cape Hatteras and any site south of it, while 13 out of 18 pairwise comparisons were significant when comparing Atlantic sites south of Cape Hatteras (CHS to MIA) and Gulf of Mexico sites (TAM to PAT) (Table S2). A multidimensional scaling (MDS) plot (Fig. 3) for US coastal sites shows Atlantic sites north of Cape Hatteras (WH to PAM) to cluster together while Atlantic sites south of Cape Hatteras and Gulf of Mexico sites form two differentiated, but adjacent clusters (Fig. 3). The open Atlantic (ATL) sample clusters with samples from the coastal Atlantic sites south of Cape Hatteras.

Hierarchical AMOVA based on mitochondrial cytb data for invasive range M. leidyi sampling sites indicates that a majority of the genetic variation occurs within populations, though a large portion (39 %; $p < 0.001$) occurs among populations (Table 2B). Accordingly, invasive sampling sites appear to occur in two general groups. All samples from northern Europe (NTH and KBG) contain Sub-haplogroup 1A haplotypes (Figs. 1, 2). The greatest overlap in haplotype representation between northern European and native samples was for Haplotypes 1–5, with four of them found only north of Cape Hatteras in the native range (Fig. 1, 2; Table S1). Two haplotypes (72 and 73) were only found in the northern European invasive animals, but were closely related to Haplotype 3 from Sub-haplogroup 1A (separated by 1 substitution in each case). On the other hand, most of the cytb haplotypes for samples from southern Europe (defined as sites in Black/Mediterranean/Caspian Seas) were included in Sub-haplogroup 1B, with the majority (84 %) sharing Haplotype 16 (Figs. 1, 2; Table S1). Haplotype 16 was found in all populations south of Cape Hatteras into the Gulf of Mexico, and the haplotype most commonly shared between southern European invasive animals, but were closely related to Haplotype 3 from Sub-haplogroup 1A (separated by 1 substitution in each case). On the other hand, most of the cytb haplotypes for samples from southern Europe (defined as sites in Black/Mediterranean/Caspian Seas) were included in Sub-haplogroup 1B, with the majority (84 %) sharing Haplotype 16 (Figs. 1, 2; Table S1). Two other haplotypes (36 and 48) were only found in the Mediterranean and in the northern Gulf of Mexico (Fig. 2; Table S1). Five haplotypes (67–71) were unique to the Black and Mediterranean Sea sites, with four of them (68–71) one step from Haplotype 16 in Sub-haplogroup 1B and the other (67) one step from Haplotype 3 in Sub-haplogroup 1A.

Table S2 reports the pairwise $\Phi_{st}$ values of all the invasive samples and their comparisons with the native samples. Comparison of northern Europe sites...
with Atlantic sites north of Cape Hatteras ranged from $-0.082$ and $0.247$ (average $= 0.016$) and were nearly all (90 %) statistically non-significant. Similar results were obtained when comparing pairwise $\Phi_{st}$ values among southern Europe sites and Gulf of Mexico sites, as they ranged from $-0.122$ to $0.115$ (average $= 0.011$) and were nearly all (92 %) non-significant. This suggests a close relationship between northern European sites and Atlantic sites from north of Cape Hatteras (WH to PAM) and between southern European sites and Gulf of Mexico sites. On the other hand, the range of the pairwise $\Phi_{st}$ values for the comparison between southern Europe sites and Atlantic sites south of Cape Hatteras was $-0.091$ to $0.341$ and only 47 % of the comparisons were not statistically significant (Table S2). The MDS plot (Fig. 3) shows that samples from the northern European sites are more similar to those from the Atlantic sites north of Cape Hatteras than to southern Atlantic samples, and the southern European sites group closer to samples from the Gulf of Mexico sites than to those from the Atlantic sites south of Cape Hatteras. Hierarchical AMOVA was highly significant ($p < 0.001$) when grouping northern Europe (NTH

Fig. 2  Statistical parsimony network for *M. leidyi* cytb haplotypes from 21 native and 9 invasive sampling sites using the program TCS 1.21 (Clement et al. 2000). Circles denote individual haplotypes and are proportional to haplotype frequency. Haplotype numbers refer to Table S1. Unlike Fig. 1, in which colors refer to haplogroups, Haplotype circles are colored based on the proportion of individuals from different geographical regions (see color legend). Lines represent mutational steps and black dots denote hypothetical intermediate haplotypes. Haplotypes were grouped into three haplogroups (groups of related haplotypes) defined by a combination of genetic similarity and geographic location (see text), and delimited within a closed black rectangle. Haplogroup 1 was further separated in sub-haplogroups (1A and 1B) defined by their relationships to the two most common haplotypes (3 and 16).
and KBG) samples with north of Cape Hatteras samples and southern Europe samples with Gulf of Mexico samples, as more of the genetic variation occurred among groups (37 %) than within groups (1.6 %) (Table 2C).

Haplotype diversity appeared to be lower in the invasive range than the native range only for samples from southern Europe. Haplotype diversity was generally high in the M. leidyi native range, averaging 0.667 and ranging from 0.3778 to 0.9238 for well-sampled populations (N ≥ 10; Table 1). Northern European samples generally had haplotype diversity values similar to or slightly higher than to those found in samples north of Cape Hatteras (Fig. S1-A), while southern European invasive populations had diversity values significantly lower than those found in the Gulf of Mexico, the putative source region based on phylogeographic and population genetic evidence (Fig. S1-B).

In general, tests of neutrality were not significant for native or invasive populations. However, several native range samples had significant values for Tajima’s D (8 of 21), Fu’s F_s (8 of 21) and R_2 (7 of 22), with most of them concentrated in US waters south of Cape Hatteras, as all sites had significant values for at least one of the measures (Table S3). All of the significant values for Tajima’s D and Fu’s F_s were negative, indicating an excess of low frequency polymorphisms. Conversely, these estimates were rarely significant for sites north of Hatteras, save for Tajima’s D (RRM and FPF), and only one of the invasive sites (LIG) showed a significant result (Tajima’s D) (Table S3).

Average neutrality test values were generally not significantly different between native and invasive sites. Based on Welch’s t test for Tajima’s D (t = 0.3763, df = 9.87, p = 0.72) or R_2 (t = −1.2605, df = 3.101, p = 0.29), average values for northern Europe sites were not significantly different from native sites north of Cape Hatteras. Likewise average values were not significantly different for Tajima’s D (t = −2.0001, df = 4.809, p = 0.104), Fu’s F_s (t = −2.3906, df = 4.721, p = 0.065) or R_2 (t = −2.1045, df = 7.994, p = 0.069), when

Fig. 3 Multidimensional scaling (MDS) plot of average pairwise genetic distances (uncorrected p distance) for cytb data for native M. leidyi sites in the vicinity of the US and Eurasian invasive populations. The Caribbean (BEZ) and South American (ARG, RIO) populations were excluded from these analyses, because their inclusion would have concealed the relationships among the other sites, as they are genetically very distant from all of them. Each sampling site is identified by its sample code (Table 1), a color (black for native, white for invasive), and a symbol representing the different geographic regions listed in the legend. Sample sites GBR, CAS and DNS are so similar on the graph so as to appear as a single symbol. The two dimensions of the plot are labeled dimension 1 and dimension 2.
comparing Gulf of Mexico sites and southern European sites. The only significant comparison was Northern Europe vs. north of Hatteras sites for Fu’s Fs ($t = -5.6866$, $df = 9.327$, $p = 0.0003$), though none of the individual populations had significant Fu’s Fs values.

The results of mismatch analyses can be found in Table S3. For all native and invasive sites, the raggedness index ($r$) was always nonsignificant, consistent with historically stable population sizes.

**Microsatellites**

Table 1 shows estimates of genetic diversity per locus and per sampling site for 525 samples and 6 microsatellite loci. Table S4 reports for each sample estimates of locus by locus genetic diversity, inbreeding coefficients ($F_{IS}$), and the probability of significant deviation from HVE, after sequential Bonferroni corrections for multiple tests (Rice 1989). Tests of linkage disequilibrium were significant for only one pair of loci (L13 and L45.1), which was significant as a result of a single sampling site (NEG). A total of 0.3% of LD tests showed significant values after sequential Bonferroni correction for multiple tests. The mean observed heterozygosity for the native range (0.58) was higher than in the invasive range (0.49), as was the inbreeding coefficient, $F_{IS}$ (0.30 vs. 0.22).

Pairwise $F_{st}$ values indicate structure in the native range of *Mnemiopsis*, with two apparent clusters separated near Cape Hatteras (Fig. 6; Table S2). Sites north of Cape Hatteras (WH to PAM) show pairwise $F_{st}$ values ranging from $-0.019$ to $0.082$ (average = 0.026). $F_{st}$ values for sites south of Cape Hatteras (CHS to ARG) range from $-0.020$ to 0.141 (average = 0.044), though when the more distantly related South American samples (RIO98, RIO99 and ARG) are removed, these values are lower ($-0.02$ to 0.075; average = 0.016). When samples north and south of Cape Hatteras are compared, $F_{st}$ values ranged from 0.017 to 0.246 (average = 0.115), with the vast majority (98.5%) of pairwise comparisons significant (Table S2). A multidimensional scaling (MDS) plot indicated this pattern with a cluster of sites north of Cape Hatteras and dispersed cluster south of Cape Hatteras, with South American sites distantly related to the rest of the native sites (Fig. 6). The open Atlantic (ATL) site appeared most closely related to samples south of Cape Hatteras.

Table S5 reports the results of the AMOVA analysis on both the native samples and the invasive ones (only samples from the Black/Caspian/Mediterranean Seas included). Most of the genetic diversity in both the native and invasive samples is partitioned within rather than between sampling sites (91.12 and 92.18%, respectively).

Figure 4 illustrates the results of the isolation by distance (IBD) analysis carried out using the genetic distances for microsatellite data from Table S2 and geographic distances between sites (determined using Google Earth 7.1.2.2041) and shows a strong IBD signal within the native range, with genetic structure clearly associated with geography in many parts of the range, as indicated by Mantel matrix correlation ($r^2 = 0.43$, $p = 0.0001$).

Bayesian clustering with STRUCTURE revealed that the value of K, the optimal number of genetic clusters given the data, was two when the analyses were run with or without locality priors (Fig. 5a), as the value of $\Delta K$ sharply declines (>90%) for all K values greater than two (data not shown, but see Fig. S2 for alternative K plots). Samples from locations north and south of Cape Hatteras (Fig. 5a) were clustered into two groups (northern and southern cluster, respectively). While there is relatively little admixture
between northern and southern cluster samples at the two extremes of the sampled range, the samples from locations from the mid-Atlantic region from Rehoboth Bay (RBD) to Pamlico Sound (PAM) are genetically admixed (Fig. 5a). The structure analyses for $K = 2$, 3, 4, and $K = 15$, the number of groups separated by significant $F_{st}$ values) are shown in supplementary material (Fig. S2), illustrating the lack of structuring beyond $K = 2$.

Pairwise estimates of $F_{st}$ values within invasive samples from the southern and northern invasion as well as among invasive and native samples are reported in Table S2. Average genetic differentiation between samples from the southern invasion (Black, Caspian, and Mediterranean Seas) ranged from 0 to 0.183 (average = 0.054). The $F_{st}$ value between the two samples from the northern invasion was $-0.007$. Genetic divergence between native and invasive was higher than within native or invasive samples, as average $F_{st}$ values between the native and invasive samples ranged from 0 to 0.323. A multidimensional scaling (MDS) plot of $F_{st}$ values indicated that Northern European sites clustered within samples from north of Cape Hatteras (Fig. 6) and pairwise $F_{st}$ values comparing these two regions ranged from $-0.009$ to 0.067 (average = 0.014), with 68.2% of pairwise comparisons nonsignificant (Table S2). Southern European sites appeared most closely related to sites from south of Cape Hatteras (Fig. 6), with pairwise $F_{st}$ values between these regions ranging from $-0.032$ to 0.323 (average = 0.080) and 33% of comparison nonsignificant, though the numbers were different when the more distantly related South American samples were removed from the analyses ($-0.032$ to 0.18; average = 0.055), with 37% of comparisons nonsignificant (Table S2). Like the samples in the native range, there was also a strong pattern of IBD in the invasive range of $M. leidyi$ within the Mediterranean/Black/Caspian/Mediterranean Seas ($r^2 = 0.480$, $p = 0.0024$; Fig. 4).

Figure 5b, c shows the results of the STRUCTURE analyses on the invasive samples using a reference data set of native individuals with high likelihoods of assignment to one of the two native genetic clusters ($Q > 95\%$), using six microsatellite loci. Overall, all sites showed moderate to high levels of genetic variability; $A_R$ ranged from 3.33 to 17.8, $H_O$ ranged from 0.36 to 0.77, and $F_{is}$ ranged from 0 to 0.44, suggesting that genetic
diversity was highly variable among sampling sites. $H_E$ and $H_O$ estimates were similar, indicating random mating within sites.

Figure S3 shows the results of the one-way $t$ tests comparing microsatellite genetic diversity (allelic diversity and observed heterozygosity) levels between native and invasive sites for both allelic richness and heterozygosity (Fig. S3: A–D). For the invasion(s) of Black, Caspian and Mediterranean Seas (Figs. S3: A, B), the results show significance for allelic richness ($t = 2.8127$, $df = 15.761$, $p$ value $= 0.013$) but not observed heterozygosity ($t = 1.9184$, $df = 10.812$, $p$ value $= 0.082$), possibly because there are several native populations with very low heterozygosity. On the other hand, for the North and Baltic Sea invasion(s), our data indicated no significant differences between native and invasive populations for allelic richness ($t = 1.6413$, $df = 3.682$, $p$ value $= 0.082$) or heterozygosity ($t = 0.6892$, $df = 14.637$, $p$ value $= 0.502$) (Figs. S3: C, D).

Table S6 shows the results of the demographic analyses to assess whether the invasive populations show signs of a genetic bottleneck in the recent past. Analyses were carried out only on the four sampling sites with more than 20 individuals and suggest the occurrence of a bottleneck in all of these sites using one of the two models of microsatellite mutation (Table S4).

Regression analyses between pairwise genetic distances at the six microsatellite loci and geographic distances for the native and southern invasive samples suggested some relationship between time since introduction and genetic diversity, but only in allelic richness. The availability of only two samples representative of the northern invasion prevents a similar analysis for northern native and invasive samples. The results suggest that there is no relationship between heterozygosity and the time since introduction from the southern invasion area ($r^2 = 0.034$, $p = 0.636$). However, there is a marginally significant association between time since introduction and allelic richness, with an increasing number of alleles with increasing time ($r^2 = 0.387$, $p = 0.074$).

**Discussion**

Native range genetic structure

Holoplanktonic organisms such as ctenophores spend their entire life cycles in the plankton, are generally weak swimmers and thus are susceptible to transport...
by the ocean currents they inhabit. The resulting long-distance dispersal should lead to broad biogeographical ranges, with little opportunity for geographical isolation to drive population subdivision and speciation, a view supported by many genetic surveys (Casteleyn et al. 2009; Stopar et al. 2010; Bortolotto et al. 2011). However, cryptic speciation and population genetic structure have been demonstrated in a surprisingly wide range of holoplanktonic organisms, presumably reflecting isolating mechanisms created at the continental barriers or by major ocean currents (Bucklin et al. 2000; Darling et al. 2004; Goetze 2005; Peijnenburg et al. 2006; Ryneanson et al. 2006; Chen and Hare 2011). Here, our data indicate significant population structure in the native range of the geographically widely dispersed ctenophore, *M. leidyi*. Many of these patterns of subdivision are concordant with well-established oceanographic boundaries and biogeographic provinces.

Mitochondrial data indicate three deeply divergent phylogenetic clades, consisting of animals collected from South America (Brazil and Argentina), Caribbean (Belize) and North America (US Gulf of Mexico and Atlantic coasts) (Figs. 1, 2). The separation of Caribbean and South American sites into genetically distinct clades is consistent with differentiation driven by the genetic barrier formed via the flow of sediment-laden freshwater the Amazon River basin, thought to separate the two regions (Briggs 1974; Rocha et al. 2002; Longhurst 2007) and to have historically driven high levels of endemism of organisms along the Atlantic coast of South America (Veron 1995; Rocha 2003). While separation of populations in the Caribbean and Gulf of Mexico might be expected since they occur in different biogeographic provinces (Longhurst 2007), the deep genetic divergence seen here is unexpected, since it is quite possible that oceanographic currents could disperse ctenophores from the Caribbean into the Gulf of Mexico. Water moves from the Caribbean through the Yucatan Strait into the Gulf of Mexico and Atlantic Ocean in the form of the Loop Current, which flows north into the center of the Gulf of Mexico before diverting southeast and connecting with the Gulf Stream (Hurlburt and Thompson 1980). It has been proposed that the Loop Current could be a vector for invasive species from the Caribbean Sea into the Gulf of Mexico (Johnson et al. 2005; Johnson and Perry 2008) and likely transports Caribbean planktonic organisms into the Gulf of Mexico (Gasca et al. 2001; Sammarco et al. 2012). Given that very large *M. leidyi* ctenophores can be found in the open Atlantic, indicating long-term survival in the open ocean (Harbison et al. 1978), one would expect some Caribbean mitochondrial haplotypes to be found in the Gulf of Mexico.

Our data also indicate significant population structure within the North American mitochondrial clade (Figs. 2, 3, 5, 6). Both mitochondrial (Figs. 2, 3) and nuclear data (Figs. 5, 6) show a population break in the vicinity of Cape Hatteras, likely driven by the oceanographic front created by the collision of warm Gulf Stream water flowing north from the Carolinan province and cooler Labrador Current water flowing south from the Virginian province (Olson et al. 1994). Cape Hatteras is a genetic break for a large number of organisms, including mollusks (Baker et al. 2008), bryozoans (McGovern and Hellberg 2003), crustaceans (Kelly et al. 2006; Chen and Hare 2011) and fish (Avise et al. 1987; McCartney et al. 2013), and may be related to thermal gradients and/or an organism’s vagility and ability to disperse against currents (Jones and Quattro 1999). Given that *M. leidyi* is tolerant to an extremely wide range of temperatures (GESAMP 1997), it is most likely that this genetic break is driven by water movement, especially of high velocity Gulf Stream water moving offshore. This is consistent with the presence of individuals in the open Atlantic (ATL) that are genetically similar to ctenophores south of Cape Hatteras (Fig. 1, 2). An additional population break at the Floridian peninsula is indicated by the *cytb* data, as Gulf of Mexico populations tend to cluster together and are separate from Atlantic populations south of Cape Hatteras (Fig. 3). Genetic discontinuities are common at the Floridian peninsula (Saunders et al. 1986; Reeb and Avise 1990; Young et al. 2002; Morrison et al. 2011), typically driven by the effects of past sea level changes and current thermal gradients between temperate waters of the northwestern Florida in the Gulf of Mexico and northeastern Florida along the Atlantic and the subtropical waters farther south (Avise 1992, 2000; Soltis et al. 2006). In addition, *Mnemiopsis* is most typically found in bays, estuaries and near-shore habitats (Kremer 1994; Purcell et al. 2001), which might reduce its ability to extend long distances and result in greater population structure than would be expected for a vagile holoplanktonic organism.
Our findings reinforce the need to use multiple types of molecular markers in population genetic studies, given the different levels of discrimination of genetic structure gained from mitochondrial cytb and microsatellite data sets. While the cytb data set indicated deep genetic divergences among North American, Caribbean and South American populations, with additional structure within North America (Figs. 1, 2, 3; Table S1–S2), the microsatellite data set revealed a consistent population level break only at Cape Hatteras (Fig. 5; Table S2), albeit with significant evidence of isolation by distance within both groups north and south of Cape Hatteras (Fig. 4). These findings are especially interesting, since the largest genetic distances among populations, as indicated by the cytb data (e.g., Caribbean vs. South America vs. North America), were undetected in the microsatellite data set, while the more subtle population break at Cape Hatteras was clear in both mitochondrial and nuclear data.

Discordance between genetic data from mitochondrial and nuclear DNA regions, or differences in the ability to resolve genetic structure, are not rare, especially for marine organisms, with mitochondrial markers often showing greater sensitivity for detecting subtle population subdivision (Peijnenburg et al. 2006; Lukoschek et al. 2008; Larmuseau et al. 2010). This difference in sensitivity has been attributed to lowered effective population size for haploid mitochondrial markers (Birky et al. 1983; Buonaccorsi et al. 2001), as well sex-biased dispersal (Karl et al. 1992). Due to the lack of recombination, differentiating selection on any polymorphism in the mitochondrial genome could also contribute to the greater geographic subdivision of mitochondrial variation (Peijnenburg et al. 2006), while sex-biased dispersal can be discounted for a simultaneous hermaphrodite. It is possible that deep phylogeographic divisions revealed by the mitochondrial sequence data are not mirrored by microsatellites, owing to the lowered maximum genetic divergence for the latter in large populations. Unlike other types of nuclear markers, microsatellites combine a limited range of allele sizes and high mutation rate, which can overpower the diversifying effect of random drift in all but small populations (Nauta and Weissing 1996). We would predict that other types of nuclear markers (e.g., SNPs) would corroborate the phylogeographic structure shown by mtDNA.

Source region(s) of invasive M. leidyi

This study illustrates how the inclusion of multiple types of molecular genetic markers (mitochondrial and nuclear microsatellite), as well as increased geographical sampling of the native range, can improve our understanding of biological invasions. Two earlier molecular studies of the M. leidyi invasion depended on a single marker class, either nuclear microsatellites (Reusch et al. 2010) or the nuclear ITS regions (Ghabooli et al. 2011). While both included comprehensive sampling of the invasive regions, neither included geographically broad sampling of M. leidyi’s native range. Regardless, both indicated separate invasions of northern (Baltic/North Sea) and southern (Black Sea/Mediterranean Sea/Caspian Sea) Europe originating from the northern US Atlantic coast of New England (northern Europe) and the vicinity of the Gulf of Mexico (southern Europe) (Reusch et al. 2010; Ghabooli et al. 2011). Our data are in general agreement with earlier studies, albeit with higher confidence levels resulting from increased knowledge of the geographical boundaries of native range M. leidyi populations.

Several lines of evidence place the source region(s) of all southern Europe invasive M. leidyi (Black Sea/Casian Sea/Mediterranean Sea) in the vicinity of the Gulf of Mexico. Phylogeographic data for cytb indicated that haplotypes found in southern Europe ctenophores are most similar to those found south of Cape Hatteras (Figs. 1, 2). STRUCTURE analysis (Fig. 5) and $F_{st}$ data (Fig. 6; Table S2) for microsatellite data also indicate that southern Europe M. leidyi originated from south of Cape Hatteras. In addition, southern Europe sites cluster more closely to sites from the Gulf of Mexico than to Atlantic sites south of Cape Hatteras in the MDS plot of genetic distance for cytb (Fig. 3). Lastly, while the vast majority of invasive ctenophores carried Haplotype 16 for cytb, which is found in all native sites south of Cape Hatteras to the Gulf of Mexico (southern Europe) originating from the northern US Atlantic coast of New England (northern Europe) and the vicinity of the Gulf of Mexico (southern Europe) (Reusch et al. 2010; Ghabooli et al. 2011), our data are in general agreement with earlier studies, albeit with higher confidence levels resulting from increased knowledge of the geographical boundaries of native range M. leidyi populations.
Several lines of evidence place the source region of northern European *M. leidyi* populations along the US Atlantic coast north of Cape Hatteras, specifically between Massachusetts (WH) to the north and Delaware (RB) to the south, as opposed to the general source region of New England based on a single populations collected from Woods Hole (Reusch et al. 2010) or Narragansett Bay (Ghabooli et al. 2011). Phylogeographic data for *cytb* indicates that haplotypes found in northern Europe ctenophores are most similar to those found north of Cape Hatteras (Figs. 1, 2). Additionally, STRUCTURE analysis (Fig. 5) and *F_{st}* values (Fig. 6; Table S2) for microsatellite data indicate that northern Europe *M. leidyi* originated from north of Cape Hatteras. Likewise, while some northern Europe ctenophores contained a widespread haplotype found in nearly all sites from the northern US Atlantic coast to the northern Gulf of Mexico (Haplotype 3), the other northern Europe haplotypes also found in the native range were only recovered in sites from Cape Hatteras and north (Table S1). Lastly, while the two northern European sites (KBG and NTH) clustered closest to sites north of Cape Hatteras, they also showed an affinity for sites from Delaware (RBD) northward and clustered slightly apart from Chesapeake Bay (RRM, CBS and CBN) and Pamlico Sound (PAM) sites (Fig. 3). This region (Delaware to Massachusetts) contains several busy international ports from which *M. leidyi* could have originated [Philadelphia (PA), New York (NY) and Boston (MA)]. An earlier theory to explain the appearance of *M. leidyi* in northern Europe was transport via natural flow of the Gulf Stream (Oliveira 2007). However, samples collected from the Open Atlantic (ATL), within the likely flow of the Gulf Stream, were genetically similar to ctenophores collected from south of Cape Hatteras and not those collected in northern Europe (Figs. 1, 2, 3, 5, 6).

These data corroborate the findings of earlier genetic studies of the *M. leidyi* invasions, while providing more confident inferences regarding the likely source regions. Moreover, our data indicate the importance of using multiple genetic markers on a well-sampled native range geographic dataset. We examined variation at the six nuclear microsatellite markers from Reusch et al. (2010) and had we not also examined mitochondrial *cytb* data, we would not have been able to determine a precise source region (Figs. 5, 6). While we have come to the same conclusion as Reusch et al. (2010) for the source of the southern Europe invaders (vicinity of the Gulf of Mexico), had that study included any native sites outside of the Gulf of Mexico to the north or south, their data would only support a source region somewhere south of Cape Hatteras as did our data employing the same microsatellite markers (Figs. 5, 6). Ghabooli et al. (2011) also employed a single nuclear marker (ITS regions) and concluded that the southern Europe invaders originated from near the Gulf of Mexico, though no Atlantic sites south of Cape Hatteras were sampled. It is unclear whether or not the ITS data would have exhibited the same degree of resolution as the microsatellites had other samples been included, though it must be noted that the data set from Ghabooli et al. (2011) did indicate that Chesapeake Bay animals were genetically distinct from New England samples (Narragansett Bay). While Ghabooli et al. (2013) used mitochondrial (COI) and nuclear (ITS) data and added a single sampling site south of Cape Hatteras (Morehead, NC) alongside extensive invasive range sampling, the source region was not further clarified. Actually, more COI haplotypes and ITS alleles were shared between invasive regions and the south of Cape Hatteras site than the Gulf of Mexico site. While both Reusch et al. (2010) and Ghabooli et al. (2011) both indicated a source region in New England (Massachusetts/Rhode Island) for northern European *M. leidyi*, our data indicate a wider range for the source region from New England to as far south as Delaware (Figs. 1, 2, 3, 5, 6). Once again, the mitochondrial data set (Figs. 1, 2, 3) adds additional precision to the nuclear data set (Figs. 5, 6), underlining the importance of employing multiple genetic markers.

Our data are rather equivocal in their support of the findings of Bolte et al. (2013) or Ghabooli et al. (2013) in regard to the source(s) of invasive *Mnemiopsis* in the Mediterranean Sea. Bolte et al. (2013) showed that Mediterranean *Mnemiopsis* originated in the Black Sea, while Ghabooli et al. (2013) cited evidence of
additional sources in addition to the Black Sea. Although two cytb haplotypes (36 and 48) were found in the Mediterranean and the Gulf of Mexico and not the Black Sea basin, similar to the findings of Ghabooli et al. (2013), it is equally likely that this is due to sampling issues or reflects a biological phenomenon. Similarly, the microsatellite data do not support the occurrence of two independent invasions in the Mediterranean, as seen from the results of the clustering analyses (Fig. 5) and the IBD pattern (Fig. 4). All in all, we would argue that there is not sufficient evidence here or elsewhere to strongly indicate additional invasive source(s) for the Mediterranean Mnemiopsis, though it is possible that more polymorphic molecular markers might answer this question more effectively.

Genetic diversity in invasive animals

Ctenophores possess many traits that likely make them exceptional bioinvaders, perhaps the most important being the ability to self-fertilize (Bayha and Graham 2014). The ability to self-fertilize likely decreases the propagule pressure, or incoming supply of individuals to a new region, necessary to become established (Bayha and Graham 2014). While not universal (Wares et al. 2005; Roman and Darling 2007), it is generally expected that a bioinvader will experience a decrease in genetic diversity with successful establishment in a new region due to founder effects (Tsutsui et al. 2000; Allendorf and Lundquist 2003) and the ability to self-fertilize and establish with an extremely small number of individuals could intensify this phenomenon.

A general decrease in genetic diversity was observed in the mitochondrial and nuclear data for southern Europe ctenophores, findings consistent with this region being established by a small number of founding individuals. There was a decrease in haplotype diversity in the cytb data set for southern European ctenophores with respect to their native region (Fig. S1). However, demographic parameters for cytb measuring departure from neutrality (e.g., Tajima’s D, Fu’s Fs, etc.), which might indicate recent population expansions, were almost all non-significant for invasive M. leidyi populations. There was a significant decrease in allelic richness for the microsatellite data in invasive regions, as opposed to native regions, though there was not a significant difference in heterozygosity (Fig. S3). In addition, regression analyses showed a significant decrease in allelic richness in the microsatellite data (though not for heterozygosity) as time since first appearance decreases, indicating that allelic diversity has decreased as M. leidyi has spread from the Black Sea into the Caspian Sea and Mediterranean. Lastly, microsatellite data for all southern Europe sites (20 or more individuals) showed significant values for population bottleneck tests under either infinite alleles model (IAM) or stepwise mutation model (SMM) (Table S6). These data are in agreement with the findings of Reusch et al. (2010) and Ghabooli et al. (2011) and indicate that M. leidyi established itself in the Black Sea with a small number of individuals prior to 1982 (Pereladov 1988) and existed in relatively modest numbers until conditions allowed for a large population increase in the late 1980s (Vinogradov et al. 1989).

We found no evidence of decreased genetic diversity in the northern Europe invasion, indicating that a sufficient number of individuals established themselves to overcome any founder effects. Our two sites in northern Europe (KBG and NS) were lightly sampled (average = 10 samples) and estimates of genetic diversity were not significantly lower for mitochondrial (Fig. S1) or nuclear (Figs. S-5C, D) datasets. These findings are in agreement with Reusch et al. (2010) and Ghabooli et al. (2011), which found no evidence of founder effects in the northern Europe ctenophore populations and indicate that the North and Baltic Seas were established by a larger number of introduced ctenophore individuals, either as a single large introduction or several separate introductions.

Conclusions

Molecular data have shown great promise in characterizing invasion histories for a wide range of taxa. Genetic approaches are critical for understanding invasion dynamics in highly invasive species with wide geographic distributions, such as M. leidyi, for which biogeography may provide minimal insights for identifying source region(s) (Reusch et al. 2010; Ghabooli et al. 2011, 2013). Here, our analyses highlight the importance of incorporating multiple classes of molecular markers into studies of phylogeography and population structure, as relying on only
mtDNA or microsatellites would have led to disparate conclusions about the invasion history of *M. leidyi*. By increasing the breadth of markers used and increasing geographical sites sampled in the native range, we have advanced our understanding of *M. leidyi* invasions and more precisely characterized the geographic boundaries of genetic groups and source regions for invasive populations. Future work will aim at using additional microsatellite loci and SNP markers in an effort to more precisely pinpoint invasive origins and the colonization dynamics. This understanding is critical in order to forestall, or at least minimize, future invasions by identifying predominant invasion routes via international shipping.

**Acknowledgments** The authors would like to thank the following people and Institutions for either providing tissue, helping in the collection of tissue or offering lab space or ship time, all of which was vital to the success of this manuscript: W.M. Graham, S.M. Chapotin, E. Christou, Dauphin Island Sea Lab, M. Faasse, V. Fuentes, S. Gunn, F. Hernandez, J. Javidpour, P. Kasapidis, A. Kideys, P. Kremer, H. Mianzan, A. Morandini, A. Moss, U. Niermann, K. Robinson, T. Shiganova, J. Sikou-Frangou, B. Sullivan, M. Telli, D. Thibault-Botha, R. Waggett, Smithsonian Environmental Research Center, and the crews of the R/V Bilim [Middle East Technical University], R/V Knorr [Woods Hole Oceanographic Institute]. PCR primers for cytb were designed based on preliminary sequences gratefully provided by M.Q. Martindale prior to sequencing of the complete mitochondrial genome of *M. leidyi* (Pett et al. 2011). A small part of the cytb dataset was collected in the lab of M.N. Dawson. K.M.B would like to acknowledge the assistance of A. Marsh during the early stages of this project. The study was supported by YIBS (Yale Institute for Biospheric Studies) funding and research grants from Sigma Xi and YSEA (Yale Science and Engineering Association) supporting the work of Yale undergraduates (Chang and Moseley) and a Sigma Xi Grant in Aid of Research (GIAR) to K. Bayha. The publication of this paper is supported by CONISMA, the Italian National Interuniversity Consortium for Marine Sciences, which received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) for the project VECTORS ([http://www.marine-vectors.eu](http://www.marine-vectors.eu)). This paper stems from the International workshop MOLTOOLS (Molecular Tools for Monitoring Marine Invasive Species), held in Lecce, Italy, in September 2012.

**References**


Bayha KM (2005) The molecular systematics and population genetics of four coastal ctenophores and scyphozoan jellyfish of the United States Atlantic and Gulf of Mexico. University of Delaware, Newark


Boero F, Putti M, Trainito E, Prontera E, Piraino S, Shiganova TA (2009) First records of *Mnemiopsis leidyi* (Ctenophora) from the Ligurian, Thrghrhenian and Ionian Seas (Western Mediterranean) and first record of *Phyllorhiza punctata* (Cnidaria) from the Western Mediterranean. Aquat Invasions 4:675–680


Maruyama T, Fuerst PA (1985) Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. Genetics 111:675–689


 Springer
Worldwide phylogeography of the invasive ctenophore


Purcell JE (1994) Predation mortality of Bay Anchovy Anchoa mitchilli eggs and larvae due to scyphomedusae and ctenophores in Chesapeake Bay. Mar Ecol Prog Ser 114:47–58


