MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 4678–4694 doi: 10.1111/j.1365-294X.2010.04837.x

Invasion genetics of the *Ciona intestinalis* species complex: from regional endemism to global homogeneity

AIBIN ZHAN, HUGH J. MACISAAC and MELANIA E. CRISTESCU

Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Avenue, Windsor, Ontario, Canada N9B 3P4

Abstract

Determining the degree of population connectivity and investigating factors driving genetic exchange at various geographical scales are essential to understanding population dynamics and spread potential of invasive species. Here, we explore these issues in the highly invasive vase tunicate, *Ciona intestinalis*, a species whose invasion history has been obscured by its poorly understood taxonomy and population genetics. Recent phylogenetic and comparative genomic studies suggest that *C. intestinalis* is a cryptic species complex consisting of at least three species. We reconstructed phylogenies based on both mitochondrial (cytochrome c oxidase subunit 3—NADH dehydrogenase subunit 1 region and NADH dehydrogenase subunit 4 gene) and nuclear (internal transcribed spacer 1) sequences, results of which support four major phylogroups corresponding to the previously reported spA, spB and *Ciona* spp. (spC) as well as an undescribed cryptic species (spD). While spC and spD remain restricted to their native ranges in the Mediterranean Sea and Black Sea, respectively, the highly invasive species (spA and spB) have disjunct global distributions. Despite extensive interspecific divergences, we identified low phylogeographical structure within these two invasive species. Haplotype network analyses revealed comparatively limited mutation steps among haplotypes within each species. Population genetic analyses based on two mtDNA fragments and eight unlinked microsatellites illustrated relatively low population differentiation and high population connectivity at both regional and continental scales in the two invasive species. Human-mediated dispersal coupled with a high potential for natural dispersal is probably responsible for the observed genetic homogeneity.

Keywords: ascidian, biological invasion, cryptic speciation, gene flow, genetic differentiation, population connectivity

Received 13 June 2010; revision received 16 August 2010; accepted 22 August 2010

Introduction

Understanding the scale and mechanisms of population connectivity is critical for risk assessment and design of appropriate management strategies for invasive species (Hampton *et al.* 2004). Population connectivity in aquatic invasive species can be influenced by natural and human-mediated pathways of propagule dispersal, species' life histories and by variation in environmental and community composition across geographical scales (Lee 2002; Darling & Folino-Rorem 2009; Jesse *et al.* 2009; Goldstien *et al.* 2010; Sorte *et al.* 2010). For marine species, particularly those with free-swimming larval stages, water currents may offer an effective method of long distance dispersal in 'open' environments lacking of physical barriers (Palumbi 1992). However, genetic studies show large disparities between the expected and realized dispersal (Palumbi 2003; Taylor & Hellberg 2003). As free-swimming larvae could be easily introduced by ballast water exchange, and sessile adults can be transplanted by vessel and aquaculture hull fouling, human activities can greatly facilitate the introduction and spread of aquatic invasive species (Ricciardi 2006;

Correspondence: Aibin Zhan, Fax: 519 971 3616;
E-mail: zhanaibin@hotmail.com

Lambert 2007). In addition, many marine species show a large effective population size and high fecundity (i.e. large propagule pools), which may lead to the wide dispersal of a large number of propagules by water currents and/or human activities.

Recent population genetic studies show various levels of population connectivity among marine invasive species with various life history characteristics. While some species such as the European green crab *Carcinus maenas* and the colonial tunicate *Symplegma rubra* exhibit high degrees of population connectivity at various geographical scales (Dias *et al.* 2006; Tepolt *et al.* 2009), others such as the colonial tunicate *Botryllus schlosseri* show high genetic differentiation among different populations even at fine geographical scales (Ben-Shlomo *et al.* 2006), suggesting limited population connectivity. Even within the same species, the level of population connectivity varies at comparative geographical scales. For example, the solitary tunicate *Styela clava* exhibits a complicated scenario of population connectivity (Dupont *et al.* 2009). While migration was observed between some geographically distant populations, limited gene flow was detected at very small scales, suggesting various dispersal mechanisms in range expansion and occupancy by *S. clava* (Dupont *et al.* 2009). Moreover, the field of invasion biology has been lacking a good model system that would allow comprehensive ecological genetic studies at various geographical scales and comparative analyses on closely related species with various degrees of invasion capacities (but see Winkler *et al.* 2008).

The vase tunicate, *Ciona intestinalis*, a cold-water or temperate hermaphroditic solitary ascidian, has become a model system for developmental and evolutionary studies (Kano 2007; Dehal *et al.* 2002). However, its invasion history has been heavily obscured by its ambiguous taxonomic status. Recent studies have demonstrated that *C. intestinalis* consists of morphologically cryptic but genetically distinct species (Suzuki *et al.* 2005; Caputi *et al.* 2007; Iannelli *et al.* 2007; Nydam & Harrison 2007, 2010). Two of the species, the Northeast Pacific/Mediterranean type (spA) and the Northwest Atlantic type (spB), inhabit largely disjoint geographical regions worldwide but overlap in the English Channel (Caputi *et al.* 2007). In contrast, a third cryptic species, *Ciona* spp. (hereafter referred to as spC), was found only in one location of the Mediterranean Sea (Nydam & Harrison 2007, 2010). Because of the ambiguous taxonomy, worldwide distribution and cryptic invasions, the native range of *C. intestinalis* species complex remains highly debated (Therriault & Herborg 2008).

During the past century, *C. intestinalis* and its congener *Ciona savignyi* have invaded benthic communities throughout the temperate zone and have been extending geographical ranges even along tropical coasts (Kott 1952; Monniot & Monniot 1994; Lambert & Lambert 2003; McDonald 2004). In most invaded areas, *C. intestinalis* is a very competitive species that rapidly covers nearly 100% of the available substratum and excludes almost all of the native species in a short period of time (Ramsay *et al.* 2008). However, because of the ambiguous taxonomy within the genus *Ciona*, the invasion history remains uncertain for many areas. In North America, *C. intestinalis* was first detected in the benthic invertebrate community on the west coast in the early 1910s (Huntsman 1912). However, it is probably that specimens of *C. savignyi*, an Asian native congener, were misidentified at that time (Lambert 2003). *C. intestinalis* was probably introduced to the west coast around 1930s (Lambert & Lambert 1998; Blum *et al.* 2006). On the east coast of North America, only the bays in Prince Edward Island have been clearly demonstrated as the invasive distribution range, because this area has been newly invaded since 2004 (Ramsay *et al.* 2008). For the other regions on the east coast of North America, extensive debates have not resolved the origin of *C. intestinalis*, leading to a cryptogenic status (Therriault & Herborg 2008). Since the late 1960s, new populations of *C. intestinalis* were detected along the coasts of South America, Australia, Asia, New Zealand and South Africa (Monniot & Monniot 1994; Seo & Lee 2009; Rocha *et al.* 2009).

Generally, human activities such as ballast water exchange, vessel hull fouling and aquaculture transfers have been considered responsible for the recent range expansion of *C. intestinalis* (Carver *et al.* 2003; Lambert 2007). However, the large effective population size (Dehal *et al.* 2002), broadcast-spawning nature and the relatively long free larval phase of 1–5 days compared with other tunicates (Carver *et al.* 2003) indicate that *Ciona* might have a high potential for natural dispersal. The interaction between natural and human-mediated dispersal at various geographical scales remains largely unexplored. Moreover, there is little genetic information about the population genetic structure and the population connectivity at both large and regional geographical scales among cryptic species of the *C. intestinalis* species complex with various invasive abilities.

Here, we investigate the evolutionary relationships and invasion history of *C. intestinalis* species complex using phylogenetic and comparative phylogeographical analyses based on multiple mitochondrial and nuclear markers. Additionally, we measure population connectivity of highly invasive species at both continental and regional scales using population genetic analyses based on two mtDNA fragments and eight unlinked polymorphic microsatellite markers.

Materials and methods

Samples collection

Our sampling design aimed to maximize the geographical coverage of *Ciona intestinalis* species complex for phylogenetic analyses and to provide intensive sampling of representative sites for population genetic studies. In total, 25 sites were selected along the coasts of five continents, of which ten were located in Europe, ten in North America, two in Asia, one in Africa and two in Oceania (Fig. 1, Table 1). For population genetic analyses, 12 representative sites—with more than 21 individuals per population—were selected (Fig. 1, Table 1). A total of 515 specimens were collected and stored in 95% ethanol at 4 °C and subsequently analysed for genetic characterization (Table 1). With the exception of two individuals morphologically identified as *Ciona roulei* by Nydam & Harrison (2007), all other specimens were analysed using a 'blind' approach with no prior morphological identification to the species level, because of the lack of reliable diagnostic characters.

DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from ~50 mg of tissue according to the proteinase K method (Waters *et al.* 2000). Mitochondrial DNA (mtDNA) fragments including cytochrome c oxidase subunit 3—NADH dehydrogenase subunit 1 (COX3-ND1), NADH dehydrogenase subunit 4 gene (ND4) and nuclear internal transcribed spacer 1 (ITS1) fragment were amplified using four primer pairs (Table 2). Polymerase chain reaction (PCR) amplifications were performed according to the protocol described by Iannelli *et al.* (2007) with locus-specific annealing temperature (Table 2). Sequencing was performed using forward primers and BigDye Terminator 3.1 sequencing chemistry on an ABI 3130XL automated sequencer. All mtDNA sequences that contained ambiguous sites were subsequently sequenced with the reverse primers. Nuclear fragments that contained double nucleotide calls were cloned using a pGEM T-easy vector system II (Promega) to verify the sequence of both alleles in heterozygous individuals.

Phylogenetic analyses

All sequences were aligned using CODONCODE ALIGNER version 2.0.6 (CodonCode). Subsequently, alignments were inspected and edited manually. Bayesian inference (BI) and neighbour-joining (NJ) phylogenetic analyses were conducted on both the combined mtDNA and the nuclear ITS1 alignment using *Ciona savignyi* as outgroup (Nydam & Harrison 2007). To better resolve the phylogenetic relationships among members of the cryptic species complex, two to four individuals from each major clade were selected to construct phylogenies based on the combined mtDNA (COX3-ND1 and ND4) and nuclear (ITS1) fragments. All molecular evolution model parameters were estimated using MRMODELTEST version 3.7 (Posada & Crandall 1998) under the Akaike information criterion (AIC). Based on the results from MRMODELTEST, the models K81uf+I+G and TVM+G were selected for the combined mtDNA data and the nuclear ITS1, respectively. The BI analysis was conducted using MRBAYES version 3.2 (Ronquist & Huelsenbeck 2003). Trees were sampled every 100 generations for 2 million generations, and the first 25% of all the trees sampled before convergence were discarded as burn-in. NJ phylogenetic analyses were performed using MEGA version 4.0 (Tamura *et al.* 2007) based on nucleotide distances corrected using the Tamura–Nei model (Tamura & Nei 1993). Clade support was

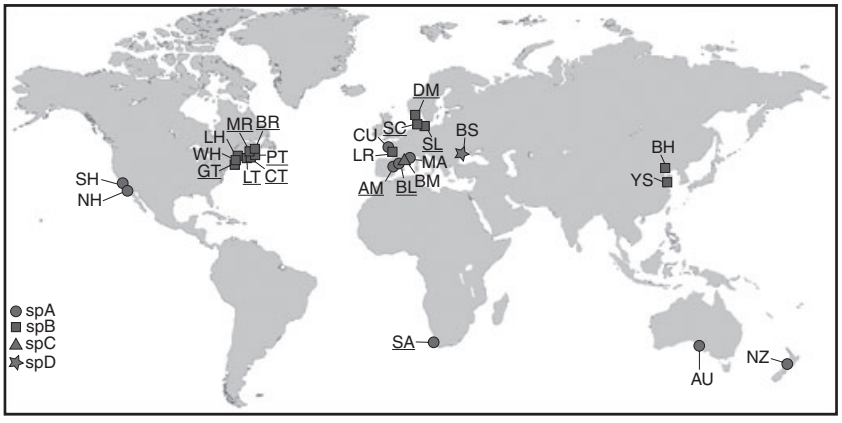

Fig. 1 Distribution of sampling sites for *Ciona intestinalis* species complex. Site IDs as per Table 1. The populations used for population genetic analyses were underlined.

Table 1 Collection sites, and mitochondrial (mtDNA)/nuclear (ITS1) diversity for *Ciona intestinalis* species complex

Sample name	Region/state and country	mtDNA					ITS1				
		N	*n*	Haplotype code	*h*	π	*N*	*n*	Allele code	N_{SSR}	Species
AM	Arenys de Mar, Spain	52	7	Ma01–Ma07	0.405	0.0004	5	2	Na01, Na02	48	spA
BL	Blanes, Spain	26	3	Ma01-Ma03	0.195	0.0002	4	2	Na01, Na02	26	spA
BR	Brudenell River, PE, Canada	30	6	Mb04, Mb20–Mb24	0.692	0.0057	5	3	Nb05–Nb07	30	spB
CT	Chester, NS, Canada	28	7	Mb04, Mb20, Mb21, Mb27–Mb30	0.743	0.0033	3	2	Nb05, Nb06	25	spB
DM	Limfjorden, Denmark	49	7	Mb01–Mb07	0.675	0.0030	5	2	Nb01, Nb02	54	spB
GT	Groton, CT, USA	48	11	Mb04, Mb08, Mb17, Mb20, Mb34–Mb40	0.896	0.0004	4	2	Nb05, Nb06	57	spB
LT	Port La Tour, NS, Canada	29	6	Mb04, Mb08, Mb11, Mb17, Mb31, Mb32	0.889	0.0057	5	5	Nb01, Nb03–Nb05, Nb09	29	spB
MR	Murray River, PE, Canada	30	5	Mb04, Mb20, Mb21, Mb25, Mb26	0.670	0.0029	5	3	Nb03, Nb05, Nb06	30	spB
PT	Point Tupper, NS, Canada	21	7	Mb04, Mb11, Mb17, Mb20, Mb21, Mb28, Mb33	0.752	0.0072	6	5	Nb01, Nb04–Nb06, Nb08	21	spB
SA	Cape town, South Africa	28	1	Ma01	—	—	5	2	Na01, Na02	28	spA
SL	Salzhaff, Germany	31	12	Mb04, Mb06, Mb09–Mb18	0.847	0.0026	4	1	Nb04	31	spB
SC	Schleimünde, Germany	30	6	Mb02, Mb04, Mb06, Mb11, Mb18, Mb19	0.717	0.0054	4	3	Nb04-Nb06	30	spB
AU	Port Lincoln, SA, Australia	6	4	Ma07–Ma10	0.800	0.0013	2	1	Na01	—	spA
BH	Bohai Sea, Liaoning, China	6	3	Mb43–Mb45	0.318	0.0013	2	2	Nb05, Nb06	—	spB
BM	Banyuls-sur-Mer, France	4	3	Mr01, Mr02, Mc01	1.000	0.0062	4	3	Nbr03, Nc01-02	—	*C. roulei* spC
BS	Black Sea, Romania	11	4	Md01–Md04	0.821	0.0063	5	3	Nd01-Nd03	—	spD
CU	Concarneau, France	1	1	Ma08	—	—	1	2	Na01, Na02	—	spA
LH	Little Harbor, NH, USA	1	1	Mb42	—	—	1	2	Nb05, Nb06	—	spB
LR	La Rochelle, France	1	1	Mb08	—	—	1	1	Nb03	—	spB
MA	Marseille, France	1	1	Ma01	—	—	1	1	Na01	—	spA
NH	Newport Harbor, CA, USA	1	1	Ma13	—	—	1	1	Na01	—	spA
SH	Santa Barbara Harbor, CA, USA	1	1	Ma13	—	—	1	1	Na01	—	spA
WH	Woods Hole, MA, USA	3	3	Mb04, Mb27, Mb41	1.000	0.0741	1	1	Nb03	—	spB
YS	Yellow Sea, Shandong, China	8	3	Mb43, Mb44, Mb46	0.530	0.0013	2	2	Nb05, Nb06	—	spB

All populations were used for phylogenetic analyses, while populations with population size >21 individuals (names bolded) were also used for population genetic analyses. The first letter in the haplotype/allele name denotes mtDNA and nuclear markers, respectively, and the second letter indicates species affiliation: a, spA; b, spB; c, spC; d, spD; r, *C. roulei*; and br, allele shared between spB and *C. roulei*. *N*, number of individuals tested for different molecular markers: mtDNA, ITS1, and microsatellite (SSR); *n*, number of haplotypes/alleles; *h*, haplotypic diversity; π, nucleotide diversity.

Table 2 Molecular markers and their corresponding primers used in this study

Locus name	Primer sequences (5′–3′)	T_a	Repeats	Source	Marker location
COX3-ND1	TX3F: GAGTGTGCKATTTGGTATTGAC TN1R: ATYTGAGCYACTCCTCGAATTC	56	—	Iannelli *et al.* (2007)	mtDNA
ND4	ND4F: CACTTYTCWTCRTGRAGTRGG ND4R: TRACWGARCTAAAKGCYACAAT	56	—	This study	mtDNA
ITS1	ITS1F1: CCGATTGAATGGTTTAGTGAGG ITS1F2: CATTTAGAGGAAGTAAAAGTCG ITS1R1: GTGATCCACCGCCAAGAGCC ITS1R2: AGGCCGCAATTCGCGTTCAG	56	—	This study	Nuclear
Cin-1	F: AACGGTCCTGTACTGCCAAT R: CGAGCTGAAACAGAAGGACTG	55	$(AT)_{28}$	Andreakis *et al.* (2007)	10p:976100
Cin-10	F: TGGATGTCCTGCTGCTACTG R: TAAGCGAACACGGGGTGTAT	55	$(TGAA)_8(TGGA)_7$ $(TGAA)_8(TGGA)_6$ $(TGAA)_{28}$	Andreakis *et al.* (2007)	9q: 66626
Cin-12	F: AATGTCGGTCCAAAGTAGGT R: CTGCCAACAGACTTCTACAGC	55	$(CTT)_{20}$	This study	3p: 3627203
Cin-13	F: ATTCCAGGTTCTTCATCCAT R: ACTATTGAACAACCAAGCAG	55	$(TA)_{31}$	This study	9q: 2589767
Cin-14	F: AGAAGACGACCGACATAGAC R: AGTCAGGGCTAGTATTTACAC	50	$(TG)_{22}$	This study	12p: 702886
Cin-15	F: ATCGCCAATGATTCCTAAAC R: GTTACAAACAAGCCACATAC	50	$(GAAT)_{18}$	This study	13q: 187235
Cin-16	F: GGTCTTTCATGTTGTATCCT R: GTATGGCACCCAACCAAATG	55	$(AT)_{21}$	This study	14p: 2988021
Cin-17	F: TCGCTTCCTTCTCCATAGTT R: ACCGCCACGAGTAAATAAGT	54	$(GAAT)_{11}$	This study	14q: 1367071

T_a, optimal annealing temperature; Marker location for microsatellites shows chromosome number and starting nucleotide of the microsatellite DNA, based on the whole genome project (http://www.jgi.doe.gov/).

estimated using bootstrap analysis with 1000 replicates. Relationships between mtDNA haplotypes were further examined using a statistical parsimony haplotype network generated at the 95% connection limit with TCS version1.21 (Clement *et al.* 2000).

Microsatellite DNA markers and genotyping

A total of eight unlinked microsatellites were used to conduct population genetic analyses (Table 2). Six markers were developed in this study based on the *C. intestinalis* spA genome sequence (http://www.jgi.doe.gov/), and two were developed by Andreakis *et al.* (2007). All markers were screened for polymorphism, evaluated for cross utility among all species and screened for high success amplification rate (>95%) to avoid null alleles. PCRs were performed in a volume of 12.5 μL containing 1 U *Taq* DNA polymerase (Qiagen), 0.1 μM forward primer, 0.2 μM reverse primer, 0.2 μM fluorescent dye, 2 μM Mg^{2+}, 100 μM each dNTP and 25–100 ng DNA template. The thermal regime for PCR reactions was 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, locus-specific annealing temperature (Table 2) for 30 s, 72 °C for 30 s and a final extension at 72 °C for 5 min. Amplified fragments were separated on an ABI 3130XL automated sequencer with the GeneScan™-500 LIZ™ internal size standard and scored using GENEMAPPER® version 4.0 (Applied Biosystems).

Population genetic analyses

Genetic diversity was calculated based on two mtDNA fragments and eight microsatellites. Intrapopulation genetic diversity for mtDNA was characterized by the standard indices of haplotype diversity (h) and nucleotide diversity (π) using DNASP version 5.00.07 (Rozas *et al.* 2003). Microsatellite-based intrapopulation diversity was evaluated by the number of alleles (A), allelic frequency (F), allelic richness (A_r), as well as the observed heterozygosity (H_O) and the expected

heterozygosity (H_E) using F_{STAT} version 2.9.3.2 (Goudet 2001). The presence of null alleles was tested with the software MICRO-CHECKER version 2.2.0 (van Oosterhout *et al.* 2004). Markov chain method (Guo & Thompson 1992) was employed to estimate the probability of significant deviation from Hardy–Weinberg equilibrium (HWE) and genotypic linkage disequilibrium using GENEPOP version 3.4 (Raymond & Rousset 1995). Significance criteria were adjusted for the number of simultaneous tests using sequential Bonferroni corrections (Rice 1989).

Population differentiation was determined by F_{ST} (θ estimator, Weir & Cockerham 1984) based on microsatellite markers for all population pairs using F_{STAT}. Additionally, the population differentiation was also assessed by Φ_{ST} based on mtDNA with Tamura & Nei (TrN) substitution model using ARLEQUIN version 3.1 (Excoffier *et al.* 2005). A total of 10^4 permutations were performed to allow for significance after sequential Bonferroni corrections. Additionally, difference in allelic richness (A_r) and expected heterozygosity (H_E) among populations was tested using nonparametric statistics (Mann–Whitney *U* test) with loci as replicates in SPSS version 12 (SPSS Inc., Chicago, IL, USA). To assess hierarchical population genetic structure, an analysis of molecular variance (AMOVA) was conducted using both mtDNA and microsatellite data based on 10^4 random permutations with software ARLEQUIN. Sampling sites were classified into continental groups based on geographical distributions. Permutation tests were performed at three hierarchical levels: among groups (across continents), among sites within groups (across populations) and among individuals within sites (within populations).

To examine the most likely number of distinct genetic clusters (*K*) for spA and spB, a Bayesian, Markov Chain Monte Carlo-based approach implemented in STRUCTURE version 2.1 (Pritchard *et al.* 2000) was used to perform clustering for partitioning all individuals into an estimated number of subpopulations (*K*). We assessed likelihoods for models with the number of clusters ranging from $K = 1$ to the total number of populations (three for spA and nine for spB). Ten independent runs were performed for each specified *K*-value to verify convergence. For each run, we used 10^5 generations discarded as burn-in followed by an additional 10^6 generations. The measure of ΔK (Evanno *et al.* 2005) was used to infer the number of biologically relevant clusters. In addition, a nonequilibrium Bayesian method was applied to estimate recent migration rate (i.e. over the last several generations) using BAYESASS version 1.3 (Wilson & Rannala 2003). We ran a total of 3×10^6 Markov chain Monte Carlo (MCMC) iterations, discarding the first 10^6 iterations as burn-in. Samples were collected every 2000 iterations to infer posterior probability distributions of parameters of interest.

Dispersal pattern on the east coast of North America, isolation by distance, was examined by testing the correlation between genetic distance ($F_{ST}/(1-F_{ST})$ for microsatellite data and $\Phi_{ST}/(1-\Phi_{ST})$ for mtDNA data) and geographical distance using the Mantel test with 10^4 permutations implemented in GENEPOP.

Results

Sequence diversity and phylogenetic pattern

Of the 515 individuals examined by PCR, 450 (87%) were successfully amplified for both mtDNA fragments. The 633- bp alignment for COX3-ND1 region did not include gaps for the coding regions (COX3, trnK and ND1). However, higher level of polymorphism including insertions, deletions and nucleotide substitutions was observed in noncoding regions. Considering the gaps in the noncoding regions, we identified 34 haplotypes (GenBank accession nos. HM036361–HM036394) and 108 variable sites, of which 101 were parsimony informative. The 591- bp alignment for ND4 gene included 23 haplotypes (GenBank accession nos. HM036395–HM036417). The alignment contained 99 variable sites, of which 89 were informative based on the parsimony criterion. The combined mitochondrial fragments resulted in a total of 66 haplotypes (Table 1). A representative subset of 78 individuals from each major clade and/or sampling site was examined at the ITS1 locus (Table 1). A total of 17 alleles (GenBank accession nos. HM045478–HM045494) were detected across all species (Table 1).

Both BI and NJ phylogenetic analyses based on combined mtDNA sequences highly supported four major phylogroups, corresponding to the previously reported species (spA, spB and spC) plus a previously undescribed species, spD (Fig. 2a). Individuals morphologically identified as *Ciona roulei* were nested within the largest clade of spB (Fig. 2a). The fourth group, corresponding to the undescribed cryptic spD, was comprised of four haplotypes derived from 11 individuals collected from Black Sea. Interspecific Tamura–Nei corrected distances (11.1–18.4%) were much higher than intraspecific distances (<1.0%). Little phylogeographical structure was identified within the major phylogroups.

Nuclear phylogenies using both BI and NJ methods revealed largely congruent phylogenies with those based on mtDNA (Fig. 2b). The phylogenies based on combined mitochondrial and nuclear sequences (COX3-ND1, ND4 and ITS1) better resolved the relationships among different cryptic species. Interestingly, the sister species relationship between the highly invasive species

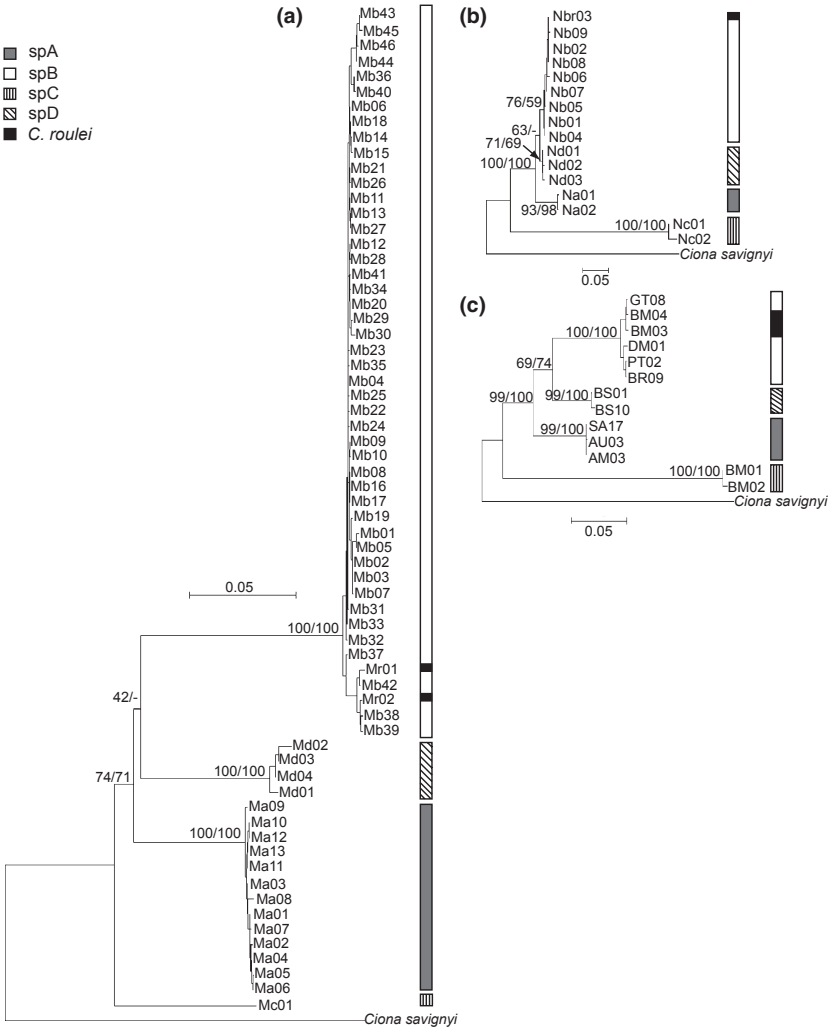

Fig. 2 Neighbour-joining trees based on (a) combined cytochrome c oxidase subunit 3—NADH dehydrogenase subunit 1 (COX3-ND1) and NADH dehydrogenase subunit 4 (ND4) haplotypes and (b) ribosomal internal transcribed spacer 1 (ITS1) alleles. Phylogenetic relationship among cryptic species was resolved by combining mtDNA and ITS1 sequences for two to four individuals from each cryptic species (c). Haplotype/allele names are as in Table 1. Bootstrap values for NJ reconstructions and posterior probabilities for Bayesian inferences (in percentage) are shown at major nodes.

(spA and spB) was not supported. Instead, spB and spD were shown to be most close to each other and then clustered with spA and spC (Fig. 2c).

The TCS network analyses based on mtDNA haplotypes for spA and spB show only several mutation steps within each cryptic species. The TCS networks for both species show star-shape patterns with a dominant haplotype with very high frequency at the centre. Derivatives with lower frequencies were connected to the dominant haplotypes by several mutation steps (Fig. 3). Generally, high-frequency haplotypes were not locality specific, and almost all high-frequency haplotypes for both species had widespread geographical distribution (Table 1, Fig. 3).

Intrapopulation genetic diversity

Overall, spA had a smaller number of haplotypes per population than spB (3.7 vs. 7.4), as well as lower haplotype diversity (0.2000 vs. 0.7646) and nucleotide diversity (0.0002 vs. 0.0040) at mtDNA loci (Table 1). All eight microsatellite DNA loci were polymorphic across all sampling sites for both spA and spB (Table S1, Supporting Information). Generally, measures of microsatellite diversity were relatively high in both invasive species. Allelic richness (A_r) ranged from 5.1 to 5.9 (average 5.4) for spA and from 5.5 to 8.0 (average 6.8) for spB. Extremely high expected heterozygosity (H_E) values were observed for all populations, ranging from 0.6547 to 0.6865 for spA and from 0.7701 to 0.8710 for spB (Table S1, Supporting Information). Deviations from HWE were observed at multiple loci and sampling locations. All of the deviated cases (60 of 103) showed significant heterozygote deficiency ($P < 0.001$) (Table S1, Supporting Information). After sequential Bonferroni corrections, no linkage disequilibrium was detected between any locus pairs in spA, and only three of 252 locus pairs were significant in spB.

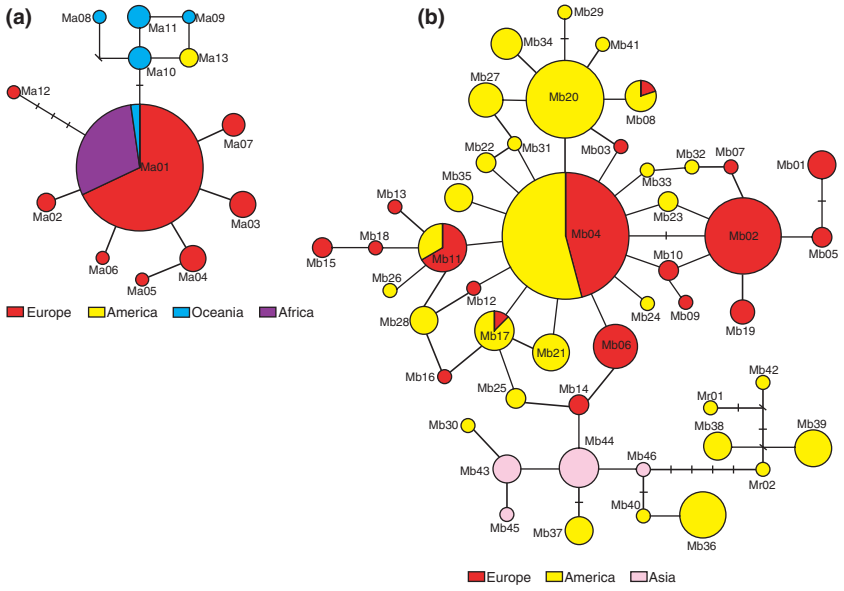

Fig. 3 The combined cytochrome c oxidase subunit 3—NADH dehydrogenase subunit 1 region (COX3-ND1) and NADH dehydrogenase subunit 4 gene (ND4) haplotype network generated with TCS for *Ciona intestinalis* spA (a) and spB (b). Haplotype names as per Table 1. Sampled haplotypes are indicated by circles and missing or unsampled haplotypes are indicated by hashes. Haplotypes are painted according to continents from which the sample was collected. Circle size is proportional to the observed haplotype frequency.

Interpopulation genetic differentiation

Genetic diversity indices H_E and A_r were not statistically significant in all comparisons among different populations, including distantly isolated populations within each species (Table 3). Global F_{ST} values for spA and spB were 0.0395 and 0.0520, respectively, dramatically lower than the average values (0.2354) observed between the two invasive species (Table 3). Generally, relatively low genetic differentiation was observed between population pairs within two invasive species (Table 3; Table S2, Supporting Information). F_{ST} values varied from 0.0353 to 0.0543 for spA and from 0.0015 to 0.1489 for spB, while Φ_{ST} ranged from −0.0014 to 0.0180 for spA and from 0.0063 to 0.4760 for spB. Despite the large geographical isolation between the populations of spA, genetic differentiation was minimal and not statistically significant (Table 3; Table S2, Supporting Information). In contrast, a higher level of differentiation was detected between populations for spB (Table 3; Table S2, Supporting Information). Overall, genetic differentiation among European populations was higher than that among North American populations for spB (Table 3; Table S2, Supporting Information). For example, the highest values of F_{ST} were detected between population SL from Baltic Sea and other European and North American populations, while the values from the remaining population pairs were lower than 0.1 (Table 3). However, F_{ST} values between European and North American populations were similar to those among European populations (0.0798 vs. 0.0902). Moreover, genetic differentiation between some European and North American populations were extremely low, for example F_{ST} = 0.0097 for the DM (from Limfjorden, Denmark)—CT (from Chester, Canada) population pair (Table 3).

Hierarchical multilocus AMOVA using microsatellite data for both cryptic species show that more than 93% of variation was attributed to the among individuals within sites component and only a small portion to the among groups (across continents) and among populations within groups components (Table 4). Standard AMOVA based on mtDNA data provided different proportions but yielded the same pattern of variation among components (Table 4).

High population connectivity

Genetic clustering analyses did not show apparent substructure among three different populations of spA. Analyses on spB yielded the highest ΔK for two genetic clusters, but no continent-specific cluster was detected (Fig. 4a). Under two clusters model for the whole data set, two populations from Europe (SC and SL) and four populations from North America (BR, MR, LT and PT) were assigned to two different clusters with the averaged membership coefficients (Q) greater than 85% (Fig. 4a). Individuals from the remaining populations

Table 3 Estimates of population genetic differentiation (pairwise F_{ST}) based on microsatellite markers (below diagonal) and *P*-values for the exact test of difference in allelic richness (above diagonal out of brackets) and expected heterozygosity (above diagonal in brackets) using a nonparametric test (Mann–Whitney *U* test) for both invasive species *Ciona intestinalis* spA and spB

Species	Pop	AM	BL	SA	DM	SC	SL	BR	MR	LT	PT	CT	GT
spA	AM	*****	0.62 (0.76)	0.61 (0.79)	*0.59 (0.180)*	*0.16 (0.81)*	*0.63 (0.28)*	*0.64 (0.14)*	*0.68 (0.16)*	*0.64 (0.31)*	*0.55 (0.13)*	*0.30 (0.09)*	*0.16 (0.06)*
	BL	0.0353	*****	0.98 (0.96)	*0.15 (0.12)*	*0.01 (0.01)*	*0.19 (0.25)*	*0.15 (0.07)*	*0.19 (0.09)*	*0.62 (0.16)*	*0.11 (0.06)*	*0.02 (0.03)*	*0.01 (0.01)*
	SA	0.0359	0.0543	*****	*0.14 (0.07)*	*0.01 (0.00)*	*0.17 (0.18)*	*0.13 (0.04)*	*0.17 (0.05)*	*0.61 (0.14)*	*0.10 (0.02)*	*0.02 (0.01)*	*0.00 (0.00)*
spB	DM	***0.2583***	***0.2165***	***0.2306***	*****	0.18 (0.16)	0.94 (0.68)	0.85 (0.89)	0.81 (0.97)	0.32 (0.53)	0.96 (0.84)	0.45 (0.56)	0.15 (0.33)
	SC	***0.2340***	***0.1933***	***0.2053***	**0.0767**	*****	0.18 (0.09)	0.08 (0.13)	0.09 (0.12)	0.11 (0.08)	0.15 (0.09)	0.36 (0.24)	0.91 (0.42)
	SL	***0.2803***	***0.2474***	***0.2563***	**0.0970**	**0.0968**	*****	0.93 (0.57)	0.88 (0.64)	0.24 (0.84)	0.90 (0.51)	0.43 (0.32)	0.16 (0.19)
	BR	***0.2630***	***0.2225***	***0.2366***	0.0633	**0.0825**	**0.1399**	*****	0.94 (0.92)	0.11 (0.26)	0.78 (0.95)	0.23 (0.60)	0.07 (0.33)
	MR	***0.2671***	***0.2303***	***0.2426***	**0.0722**	**0.0892**	**0.1489**	0.0015	*****	0.26 (0.44)	0.75 (0.86)	0.23 (0.54)	0.06 (0.30)
	LT	***0.2613***	***0.2215***	***0.2337***	**0.0412**	**0.0830**	**0.1214**	0.0341	0.0328	*****	0.06 (0.16)	0.18 (0.07)	0.16 (0.12)
	PT	***0.2619***	***0.2208***	***0.2343***	0.0354	**0.0746**	**0.1202**	0.0199	0.0325	0.0636	*****	0.41 (0.60)	0.10 (0.28)
	CT	***0.2538***	***0.2121***	***0.2266***	0.0097	0.0538	0.0672	0.0255	0.0381	0.0526	0.0085	*****	0.29 (0.61)
	GT	***0.2368***	***0.1993***	***0.2105***	0.0231	0.0448	0.0781	0.0481	0.0538	**0.0627**	0.0286	−0.0039	*****

The values between different species were represented in italic, and bold numbers indicate statistical significance after sequential Bonferroni corrections.

(DM from Europe, and CT and GT from North America) did not appear to assign to a single cluster (Fig. 4a). When clustering analyses were run independently on European and North American populations, analyses still support two clusters for the North American populations (Fig. 4c). However, hierarchical structure was detected in European populations: populations DM and SC were assigned to one cluster, while population SL was assigned to the other cluster. When a group comprising populations DM and SC was examined subsequently, these two populations were assigned to two different clusters with the averaged *Q* larger than 90% (Fig. 4b). Runs for *K* values larger than three for each European population did not yield more clusters, suggesting no further significant substructure.

Migration rate as calculated with BAYESASS indicated a high level of current migration between continents in both invasive species (Table 5). Interestingly, the immigration and emigration rates were not equal. When the data were pooled based on continents, the European group showed very high emigration rate but very low or no immigration for both species. For example, emigration rate from Europe to South Africa is greater than 0.30 but immigration lower than 0.01 for spA. Similarly, emigration rate from Europe to North America is larger than 0.10 but immigration smaller than 0.03 for spB. When the analyses were performed at the population level, the high recent migration rate between European and North American populations was also very well evidenced (Table 5). High self-recruitment rate was observed for spB in Europe (Table 5), which is consistent with the relatively high pairwise F_{ST} values and hierarchical genetic structure (Table 3, Fig. 4). In contrast, high recent migration rates were observed between the North American populations (Table 5). Migration occurred not only between neighbouring populations but also between distant populations such as the BR-PT population pair (Table 5).

Isolation by distance analyses based on both mtDNA and microsatellite DNA show no correlation ($P > 0.05$) between genetic and geographical distances, suggesting no pattern of isolation by distance.

Discussion

Invasion history in Ciona intestinalis *species complex*

Understanding the evolutionary and ecological causes responsible for the rapid spread of invasive species represents one of the main challenges for conservation biologists and environmental managers. Genome sequences of *Ciona intestinalis* and its relative *Ciona savignyi*

Table 4 Results of the analysis of molecular variance (AMOVA) for two invasive species *Ciona intestinalis* spA and spB. Sites were separated into different continental groups based on geographical distribution

Cryptic species	Source of variation	Sum of square	Variance components	Percentage variation	*P* value
spA	mtDNA				
	Among groups	0.19	−0.03	−20.74	0.01
	Among populations within groups	2.12	0.04	28.08	0.03
	Among individuals within populations	15.18	0.15	92.67	0.40
	Total	17.48	0.16		
	Microsatellite				
	Among groups	3.43	0.01	1.17	0.00
	Among populations within groups	2.50	0.03	2.59	0.00
	Among individuals within populations	187.32	1.00	96.24	0.30
	Total	193.26	1.04		
spB	mtDNA				
	Among groups	5.71	0.02	4.97	0.00
	Among populations within groups	19.14	0.06	13.46	0.00
	Among individuals within populations	106.02	0.38	81.57	0.04
	Total	130.87	0.46		
	Microsatellite				
	Among groups	22.82	0.03	0.96	0.00
	Among populations within groups	98.14	0.20	5.75	0.00
	Among individuals within populations	1614.86	3.29	93.29	0.13
	Total	1735.82	3.52		

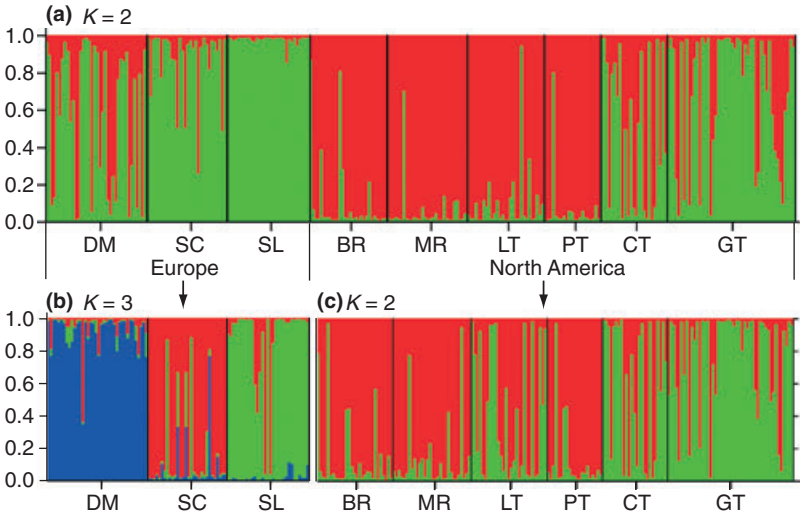

Fig. 4 Bayesian clustering of *Ciona intestinalis* spB genotypes performed in STRUCTURE for (a) all populations ($K = 2$), (b) European populations ($K = 3$) and (c) North American populations ($K = 2$). Each genotype is represented by a thin vertical line, with proportional membership in different clusters. Bold vertical lines separate collection sites, with site IDs indicated below the plot.

provide valuable tools for answering broad evolutionary questions related not only to development and evolution of chordates but also to invasion biology. However, the unsettled taxonomic status of the genus *Ciona* greatly obscures the invasion history of the group. While some researchers have proposed that *C. intestinalis* and *C. savignyi* be the only valid shallow water species recognized in the genus, and all other congeners should be made synonyms to *C. intestinalis*, others consider *Ciona roulei* and *Ciona edwardsi* to be valid species (see Kano 2007 and references therein). Recent molecular studies documented cryptic speciation within *C. intestinalis*, as well as the polyphyletic nature of *C. roulei* nested within spB (Caputi *et al.* 2007; Iannelli *et al.* 2007; Nydam & Harrison 2007, 2010).

Our results derived from both phylogenetic and population genetic analyses clearly support that *C. intestinalis* is a species complex. Three previously reported cryptic species (spA, spB and spC) were highly resolved in our phylogenetic reconstructions, and an

Table 5 Immigration rates (means with associated 95% confidence interval; all standard deviations <0.05) among populations of *Ciona intestinalis* spA (A) and spB (B) as estimated by BAYESASS. Source localities are given in rows, recipient localities in columns

		Europe		Africa
A	Pop	AM	BL	SA
	AM	**0.99** (**0.97–0.99**)	0.30 (0.24–0.33)	0.31 (0.26–0.33)
	BL	0.00 (3.7e-5–1.6e-2)	**0.68** (**0.66–0.72**)	0.01 (4.0e-4–4.2e-2)
	SA	0.00 (2.0e-5–1.6e-2)	0.02 (4.0e-4–5.5e-2)	**0.68** (**0.66–0.72**)

		Europe			North America					
B	Pop	DM	SC	SL	BR	MR	LT	PT	CT	GT
	DM	**0.99** (**0.97–0.99**)	0.00 (3.2e-13–1.9e-2)	0.00 (1.3e-13–1.2e-2)	0.00 (4.8e-12–1.5e-2)	0.00 (3.8e-7–2.6e-2)	0.01 (4.0e-7–3.6e-2)	0.01 (1.0e-7–3.4e-2)	0.09 (0.03–0.17)	0.08 (0.04–0.13)
	SC	0.00 (7.5e-13–1.1e-2)	**0.98** (**0.94–0.99**)	0.00 (2.6e-13–1.2e-2)	0.00 (1.5e-12–1.3e-2)	0.00 (7.4e-7–2.5e-2)	0.00 (1.3e-7–1.4e-2)	0.01 (1.8e-7–3.7e-2)	0.04 (2.7e-3–0.12)	0.06 (0.02–0.17)
	SL	0.00 (4.6e-14–9.8e-3)	0.00 (2.9e-13–2.7e-2)	**0.99** (**0.96–0.99**)	0.00 (3.1e-12–1.6e-2)	0.00 (1.3e-7–2.2e-2)	0.00 (5.3e-7–2.7e-2)	0.01 (1.9e-7–3.9e-2)	0.06 (0.02–0.12)	0.11 (0.07–0.17)
	BR	0.00 (3.8e-13–9.2e-3)	0.00 (8.1e-12–1.9e-2)	0.00 (9.4e-13–1.0e-2)	**0.99** (**0.96–0.99**)	0.29 (0.25–0.32)	0.29 (0.24–0.33)	0.28 (0.22–0.32)	0.11 (0.04–0.19)	0.06 (0.02–0.11)
	MR	0.00 (3.2e-14–8.7e-3)	0.00 (2.4e-12–1.6e-2)	0.00 (3.6e-12–1.1e-2)	0.00 (3.3e-12–1.2e-2)	**0.68** (**0.67–0.70**)	0.00 (1.6e-7–2.3e-2)	0.01 (5.1e-7–3.4e-2)	0.01 (4.6e-5–2.7e-2)	0.00 (3.4e-5–0.02)
	LT	0.00 (3.0e-13–1.0e-2)	0.00 (9.3e-13–1.5e-2)	0.00 (2.6e-12–9.6e-3)	0.00 (2.7e-13–1.4e-2)	0.00 (2.9e-7–2.1e-2)	**0.69** (**0.67–0.70**)	0.10 (0.08–0.14)	0.01 (2.5e-5–3.3e-2)	0.00 (4.7e-5–0.02)
	PT	0.00 (2.7e-13–9.7e-3)	0.00 (1.9e-12–1.4e-2)	0.00 (8.4e-13–1.4e-2)	0.00 (3.2e-13–1.1e-2)	0.00 (1.3e-7–2.3e-2)	0.00 (1.6e-7–2.4e-2)	**0.58** (**0.56–0.62**)	0.01 (2.8e-5–2.6e-2)	0.00 (3.2e-5–0.02)
	CT	0.00 (8.1e-13–1.1e-2)	0.00 (1.1e-12–1.5e-2)	0.00 (4.2e-13–1.2e-2)	0.00 (3.7e-12–1.3e-2)	0.00 (1.5e-7–2.4e-2)	0.00 (3.5e-7–2.1e-2)	0.01 (3.2e-7–3.4e-2)	**0.65** (**0.64–0.70**)	0.00 (1.1e-5–0.02)
	GT	0.00 (1.7e-12–7.7e-3)	0.00 (2.0e-12–1.3e-2)	0.00 (2.8e-12–1.1e-2)	0.00 (6.2e-12–1.4e-2)	0.00 (4.1e-7–2.5e-2)	0.00 (6.3e-7–2.4e-2)	0.01 (3.2e-7–3.2e-2)	0.01 (2.5e-5–0.03)	**0.67** (**0.67–0.69**)

Values along the diagonal are self-recruitment rates for each locality (bold). The immigration rates between spA and spB were not considered.

undescribed cryptic species (spD) was strongly supported by both mtDNA and nuclear ITS1 (Fig. 2). Phylogenetic analyses grouped all the Black Sea isolates into one well-supported clade, highly divergent from the other clades (Fig. 2). Although our phylogenies did not include the geographically restricted and rare *C. edwardsi*, described only from Mediterranean Sea (Harant & Vernieres 1933), individuals sampled from the low salinity area in the northwestern Black Sea (*Phyllophora* field) probably represents a Pontic endemic and not *C. edwardsi*. Moreover, the high level of endemism in the Mediterranean Sea and Black Sea regions, coupled with the high level of nucleotide sequence divergence between the major clades (11.1–18.4%) which corresponds to about 5.5–12.0 Ma divergence time (1.5–2.5% per Myr, Martin & Palumbi 1993), suggests that *C. intestinalis* might have radiated sometime during the late Miocene in the brackish Sarmatian (10.5–8.2 Ma) or Pontian (8.2–6.3 Ma) Seas. Both paleo-basins fostered intensive radiation in many other brackish water or estuarine invertebrate groups such as crustaceans (Cristescu *et al.* 2003).

With the exception of the specimens that were morphologically identified as *C. roulei*, phylogenies based on both mtDNA and nuclear markers provide very clear species boundary (Fig. 2). Previous reciprocal fertilization trials using allopatric individuals of the two species, spA and spB, resulted in failures in fertilization or infertile hybrids with defective gametogenesis (Caputi *et al.* 2007). The results suggest that strong pre- and/or postzygotic mechanisms of reproductive isolation among cryptic species might explain why genetic hybrids were not detected in the wild (Caputi *et al.* 2007).

Based on phylogeographical analyses, species spA and spB are widely distributed worldwide. These two species occur on different coasts: while spA invaded the west coast of North America, coasts along Australia, New Zealand and South Africa, spB occupies the east coast of North America as well as the coasts of the Bohai Sea and the Yellow Sea. East Asian coasts were considered to be invaded by spA (Caputi *et al.* 2007); however, our results indicate that coastal areas of the Bohai Sea and the Yellow Sea were invaded by spB rather than by spA. The other two species, spC and spD, remain restricted to Black Sea and Mediterranean Sea, respectively.

Genetic diversity within populations

Overall, we detected high intraspecific genetic diversity within the two highly invasive species (Table S1, Supporting Information). The expected heterozygosity (H_E = 0.6457–0.6865 for spA and 0.7753–0.8710 for spB) detected in this study is comparable to the H_E reported in other *C. intestinalis* studies (e.g. H_E = 0.73, Sordino *et al.* 2008) but higher than those in other solitary tunicates such as *Styela clava* (H_E = 0.48–0.63, Dupont *et al.* 2009, 2010; Goldstien *et al.* 2010). The high level of H_E is consistent with the unusual high level of genetic diversity reported in *Ciona* species. For example, high genetic diversity was revealed in an analysis of the whole genome data for *C. savignyi*. The single nuclear polymorphism ratio of 4.5% (Vinson *et al.* 2005; Small *et al.* 2007) is considered the highest in the animal kingdom (Sauvage *et al.* 2007). A very high degree of polymorphism was also observed in *C. intestinalis* (Procaccini *et al.* 2000; Kano 2007; Caputi *et al.* 2008). The large effective population size and high mutation rate are probably the major causes of the high level of intrapopulation diversity (Kimura 1983; Hedgecock *et al.* 2004). However, Small *et al.* (2007) suggested that extreme polymorphism in *Ciona* is probably associated with large effective population size rather than an elevated mutation rate.

Both invasive species show a massive deviation from HWE, and all deviations, representing 58.3% of all cases analysed, were attributed to heterozygote deficit (Table S1, Supporting Information). Similar or higher ratio of departure from HWE was previously reported in *C. intestinalis* (Sordino *et al.* 2008) and in other ascidians including the solitary tunicate *S. clava* (Dupont *et al.* 2010) and colonial tunicate *Botryllus schlosseri* (Ben-Shlomo *et al.* 2006). The departure resulting from heterozygote deficit might be explained by recurrent inbreeding (i.e. mating among kin and/or self-fertilization), subpopulation structure (i.e. Wahlund effect) and/or null alleles. Most populations in this study were sampled from enclosed or semi-enclosed waters such as marinas, where self-recruitment is dominant and reproduction between related individuals might be more common than in open habitats (Petersen & Svane 1995; Dupont *et al.* 2009). Moreover, settlement studies have demonstrated that *C. intestinalis* larvae generally settle very close to the adult individuals (Petersen & Svane 1995; Howes *et al.* 2007), increasing the possibility of breeding with related individuals. Additionally, fertilization studies in *C. intestinalis* indicate that self-sterility is not absolute (Rosati & Santis 1978; Kawamura *et al.* 1987); about 15–20% of individuals can self-fertilize (Rosati & Santis 1978; Kawamura *et al.* 1987). Moreover, subpopulation structure (Fig. 4) and high connectivity among populations (Table 5) suggest that a temporal and/or spatial Wahlund effect is probably another cause for the massive heterozygote deficit. Null allele has been recognized to be one important cause for heterozygote deficiency in many marine species

(e.g. Hedgecock *et al.* 2004; Zhan *et al.* 2007). In the present study, MICRO-CHECKER suggested significant occurrence of null alleles in most heterozygote deficit cases (Table S1, Supporting Information). Null alleles in *C. intestinalis* could arise as a result of a highly polymorphic genome as mentioned previously. However, high successful PCR amplification rate (>90% across all populations) derived from strict locus selection suggests that the frequencies of null alleles at the selected loci should be much lower than those estimated by MICRO-CHECKER.

Interpopulation genetic differentiation

All analyses including pairwise F_{ST} (Table 3) and STRUCTURE (Fig. 4) show no significant genetic differentiation among the populations of spA. Similarly, Caputi *et al.* (2007) found a very high level of genetic homogeneity and lack of geographical differentiation across a wide geographical area covering European waters, the west coast of North America and Asian coasts. The absence of geographical differentiation could be the outcome of recent population expansion (Caputi *et al.* 2007) coupled with the unobstructed gene flow among populations (Table 5).

By contrast, we observed different scenarios of population differentiation in Europe and North America for spB. Populations of spB from Europe exhibit a relatively high interpopulation genetic differentiation, which is supported by high pairwise F_{ST} values (Table 3) and Φ_{ST} values (Table S2, supporting information), hierarchical genetic structure (Fig. 4) and limited gene flow (Table 5). The spatial genetic structure of spB in Europe was also evidenced by other population-based analyses (Caputi *et al.* 2007). However, spB in North America is generally represented by low genetic differentiation (Table 3; Table S2, Supporting Information) and high population connectivity (Table 5). Different levels of population differentiation were also observed even at local geographical scales in other marine invasive species including solitary tunicates such as *S. clava* (Dupont *et al.* 2009). When comparing genetic differentiation and migration rate, we observed consistent results in most population pairs, i.e. migration rates appear inversely correlated with pairwise F_{ST} and Φ_{ST} values (Table S2, Supporting Information; Table 3; Table 5). However, several exceptions were observed, for example, pairwise F_{ST} value was as low as 0.0097 between two populations DM (Limfjorden, Denmark) and CT (Chester, NS, Canada), but no migration was detected. The highly invasive nature of *C. intestinalis* suggests that recent population separation is probably the reason for such disparities.

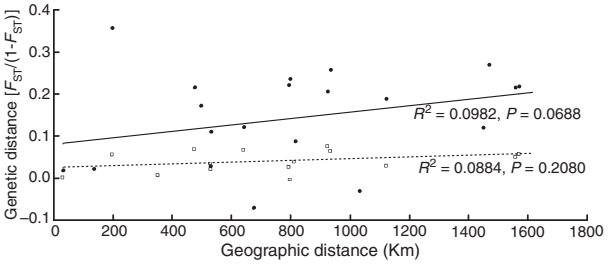

Fig. 5 Correlation between geographical distance in kilometres (x-axis) and genetic distance given as $F_{ST}/(1-F_{ST})$ for microsatellite data and $\Phi_{ST}/(1-\Phi_{ST})$ for mitochondrial data (y-axis). Solid circles and open squares represent data derived from mitochondrial and microsatellite markers, respectively.

High population connectivity

Generally, we observed a high level of connectivity at both continental and regional scales. The high population connectivity hypothesis is supported by low F_{ST} values (Table 3), nonsignificant differences in genetic diversity indices between populations (Table 3), no pattern of isolation by distance (Fig. 5) and a high level of recent gene flow (Table 5). High gene flow among populations may dilute effects of genetic bottlenecks and decrease genetic differentiation among populations (Nei *et al.* 1975).

At the intercontinental scale, some studies show intercontinental natural dispersal does occur but is insufficient to homogenize genetic variation among continents (e.g. Waters *et al.* 2000). Strong geographical patterns have been observed in species with similar biological characteristics to *C. intestinalis,* such as sea urchins (genus *Echinometra*; Palumbi *et al.* 1997). However, we observed relatively low pairwise F_{ST} values (Table 3) and high recent gene flow (Table 5) between European and North American populations. The low intercontinental genetic differentiation was also observed in other highly invasive solitary ascidians such as *S. clava* (Dupont *et al.* 2010). Although some ascidian adults can occasionally be dispersed attached to drifting logs, weeds and algal blades, this dispersal method appears to have a limited range and is not expected to be sufficient to homogenize genetic variation at an intercontinental scale (Dupont *et al.* 2010). The most likely cause for such a genetic pattern is the high level of human-mediated gene flow, and the most possible vectors are international shipping (Lambert 2007) and/or aquaculture transfer (Dijkstra *et al.* 2007). The effective population size of *C. intestinalis* is extremely large (Dehal *et al.* 2002). Furthermore, *C. intestinalis* can produce gametes continually as long as temperatures are suitable (Carver *et al.* 2003). Each mature individual can potentially spawn once daily over the spawning period and release

approximately 500 eggs per day (Carver *et al.* 2003). After 1–5 days of a free-swimming, nonfeeding larval stage, the tadpole larvae metamorphose and then settle down to benthic hard substrates as sessile adults (Carver *et al.* 2003). Consequently, the large effective population size coupled with other spawning-related characteristics can result in extremely large propagule pools that can be introduced to distant locations in discharged ballast water for larvae and/or gametes or by adults fouled to vessel hull structures or shells of bivalves. In fact, large numbers of propagules have been observed inside ballast tanks of cargo vessels (NOAA report, http://www.glerl.noaa.gov/pubs/brochures/ballast.pdf).

At a regional scale, high population connectivity in *C. intestinalis* might result from human-mediated and natural dispersal. The most direct evidence supporting human-mediated dispersal comes from gene flow analyses (Table 5). On the east coast of North America, larvae advected by currents should mainly migrate from north to south, given the dominant marine currents in this zone (Kenchington *et al.* 2006). However, gene flow between populations does not thoroughly support this hypothesis. Instead, results suggest long distance migration between geographically distant populations (Table 5). Additionally, almost all populations in this study were sampled from enclosed waters. Usually, marine currents and retention gyres associated with banks or embayments in bays or gulfs can form significant genetic barriers for marine species with lifestyle of a free-swimming larval stage and sedentary adult (Bilton *et al.* 2002; Zhan *et al.* 2009). However, we observed a high level of population connectivity among both European and North American populations. Possible reasons for such population connectivity include aquaculture transfers and shipping, which are responsible for regional spread of other species such as the clubbed tunicate *S. clava* (Dijkstra *et al.* 2007).

Compared to other tunicates, natural dispersal of *C. intestinalis* at regional and local scales is expected to be high because of its life history characteristics, especially a large effective population size, long spawning time and a long free-swimming larval phase. Marine species with similar characteristics usually exhibit very low geographical differentiation (Palumbi 1992; Nielsen & Kenchington 2001; Zhan *et al.* 2009). Although local recruitment has been observed in *C. intestinalis* (Howes *et al.* 2007), regional natural dispersal has been confirmed (Svane & Havenhand 1993). A high level of natural dispersal was found in other species with similar biological characteristics. For example, Tremblay *et al.* (1994) modelled the scallop larvae (*Placopecten magellanicus*) distribution and suggest populations in the vicinity of Georges (Canada) and Georges (US) can mix in a single generation, although they acknowledge that self-recruitment is possible for both areas. The high potential of natural dispersal of *C. intestinalis* might partially be responsible for the high population connectivity at regional scales.

Conclusions

Both phylogenetic and population genetic analyses support that *Ciona intestinalis* is a species complex consisting of at least four highly divergent species. Interestingly, two of them (spA and spB) are invasive and have a global distribution, whereas the other two (spC and spD) remain restricted to their native ranges. Phylogenetic and population genetic analyses of the two invasive species provide evidence of high population connectivity among locations. Genetic analyses support the hypothesis that continuous human-mediated gene flow, coupled with high potential for natural dispersal, is responsible for high population connectivity at both regional and continental scales. Given high apparent connectivity of populations in Europe and North America and economic harm that results from establishment of new populations in areas supporting aquaculture, greater attention must be paid to identifying and managing pathways responsible for ascidian introductions.

Acknowledgements

We thank G. Arsenault, Z. Bao, T. Begun, T. Bolton, J. Cheng, J. Davidson, M. Ekins, R. Graille, P. Chevaldonne, C. Griffiths, C. Hiebenthal, J. Hill, J. Hu, X. Huang, R. Karez, P. Kjerulf, A. Lacoursière-Roussel, C. Lejeusne, B. Ligondes, C. Lyons, K. Maczassek, J. Nicolas, M. Nydam, A. Ramsay, J. Reimhardt, M. Ruis-Viladomiu, H. Soennichsen, A. Teaca, X. Turon, B. Vercaemer, M. Wahl, Z. Zhou and F. Zuberer for assistance with sample collection. A. Abisola helped with DNA extractions. D. Bock (University of Windsor), J. Darling (US Environmental Protection Agency), T. Duda (University of Michigan), M. Ekins (Queensland Museum) and anonymous reviewers provided valuable comments. This work was supported by research grants from Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Aquatic Invasive Network (CAISN) to MEC and HJM, an DFO Invasive Species Research Chair to HJM and an Ontario Ministry of Research and Innovation, Early Researcher Award to MEC.

References

Andreakis N, Caputi L, Sordino P (2007) Characterization of highly polymorphic nuclear microsatellite loci from the ascidian *Ciona intestinalis*. *Molecular Ecology Notes*, **7**, 610–612.

Ben-Shlomo R, Paz G, Rinkevich B (2006) Postglacial-period and recent invasions shape the population genetics of Botryllid Ascidians along European Atlantic coasts. *Ecosystems*, **9**, 1118–1127.

Bilton DT, Paula J, Bishop JDD (2002) Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuarine, Coastal and shelf Science*, **55**, 937–952.

Blum JC, Chang AL, Liljesthröm M, Schenk ME, Steinberg MK, Ruiz GM (2006) The non-native solitary ascidian *Ciona intestinalis* (L.) depresses species richness. *Journal of Experimental Marine Biology and Ecology*, **342**, 5–14.

Caputi L, Andreakis N, Mastrototaro F, Cirino P, Vassillo M, Sordino P (2007) Cryptic speciation in a model invertebrate chordate. *Proceedings of the National Academy of Sciences USA*, **104**, 9364–9369.

Caputi L, Borra M, Andreakis N, Biffali E, Sordino P (2008) SNPs and Hox gene mapping in *Ciona intestinalis*. *BMC Genomics*, **9**, 39.

Carver CE, Chisholm A, Mallet AL (2003) Strategies to mitigate the impact of *Ciona intestinalis* (L.) biofouling on shellfish production. *Journal of Shellfish Research*, **22**, 621–631.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cristescu ME, Hebert PD, Onciu TM (2003) Phylogeography of Ponto-Caspian crustaceans: a benthic-planktonic comparison. *Molecular Ecology*, **12**, 985–996.

Darling JA, Folino-Rorem NC (2009) Genetic analysis across different spatial scales reveals multiple dispersal mechanisms for the invasive hydrozoan *Cordylophora* in the Great Lakes. *Molecular Ecology*, **18**, 4827–4840.

Dehal P, Satou Y, Campbell RK *et al.* (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*, **298**, 2157–2167.

Dias GM, Duarte LFL, Solferini VN (2006) Low genetic differentiation between isolated populations of the colonial ascidian *Symplegma rubra* Monniot, C. 1972. *Marine Biology*, **148**, 807–815.

Dijkstra J, Harris LG, Westerman E (2007) Distribution and long-term temporal patterns of four invasive colonial ascidians in the Gulf of Maine. *Journal of Experimental Marine Biology and Ecology*, **342**, 61–68.

Dupont L, Viard F, Dowell MJ, Wood C, Bishop JDD (2009) Fine- and regional-scale genetic structure of the exotic ascidian *Styela clava* (Tunicata) in southwest England, 50 years after its introduction. *Molecular Ecology*, **18**, 442–453.

Dupont L, Viard F, Davis MH, Nishikawa T, Bishop JDD (2010) Pathways of spread of the introduced ascidian *Styela clava* (Tunicata) in Northern Europe, as revealed by microsatellite markers. *Biological Invasions*, **12**, 2707–2721.

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

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetic data analysis. *Evolutionary Bioinformatics Online*, **2005**, 47–50.

Goldstien SJ, Schiel DR, Gemmell NJ (2010) Regional connectivity and coastal expansion: differentiating pre-border and post-border vectors for the invasive tunicate *Styela clava*. *Molecular Ecology*, **19**, 874–885.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 396 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hampton JO, Spencer PBS, Alpers DL, Twigg LE, Woolnough AP, Doust J (2004) Molecular techniques, wildlife management and the important of genetic population structure and dispersal: a case study with feral pigs. *Journal of Applied Ecology*, **41**, 735–743.

Harant H, Vernieres P (1933) Tuniciers. *Faune de France*, **27**, 1–93.

Hedgecock D, Li G, Hubert S, Bucklin K, Ribes V (2004) Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *Journal of Shellfish Research*, **23**, 379–385.

Howes S, Herbinger CM, Darnell P, Vercaemer B (2007) Spatial and temporal patterns of recruitment of the tunicate *Ciona intestinalis* on a mussel farm in Nova Scotia, Canada. *Journal of Experimental Marine Biology and Ecology*, **342**, 85–92.

Huntsman AG (1912) Ascidians from the coasts of Canada. *Transactions of Canadian Institutes*, **9**, 111–148.

Iannelli F, Pesole G, Sordino P, Gissi C (2007) Mitogenomics reveals two cryptic species in *Ciona intestinalis*. *Trends in Genetics*, **23**, 419–422.

Jesse R, Pfenninger M, Fratini S, Scalici M, Streit B, Schubart CD (2009) Disjunct distribution of the Mediterranean freshwater crab *Potamon fluviatile* – natural expansion or human introduction? *Biological Invasions*, **11**, 2209–2221.

Kano S (2007) Initial stage of genetic mapping in *Ciona intestinalis*. *Developmental Dynamics*, **236**, 1768–1781.

Kawamura K, Fujita H, Nakauchi M (1987) Cytological characterization of self incompatibility in gametes of the Ascidian, *Ciona intestinalis*. *Development, Growth & Differentiation*, **29**, 627–642.

Kenchington EL, Patwary MU, Zouros E, Bird CJ (2006) Genetic differentiation in relation to marine landscape in a broadcast-pawning bivalve mollusc (*Placopecten magellanicus*). *Molecular Ecology*, **15**, 1781–1796.

Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Kott P (1952) The ascidians of Australia. *Australian Journal of Marine and Freshwater Research*, **3**, 206–333.

Lambert G (2003) New records of ascidians from the NE Pacific: a new species of *Tridemnum*, range extension and redescription of *Aplidiopsis pannosum* (Ritter, 1899) including its larva, and several non-indigenous species. *Zoosystema*, **25**, 665–679.

Lambert G (2007) Invasive sea squirts: a growing global problem. *Journal of Experimental Marine Biology and Ecology*, **342**, 3–4.

Lambert CC, Lambert G (1998) Non-indigenous ascidians in southern California harbors and marinas. *Marine Biology*, **130**, 675–688.

Lambert CC, Lambert G (2003) Persistence and differential distribution of nonindigenous ascidians in harbors of the Southern California Bight. *Marine Ecology Progress Series*, **259**, 145–161.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386–391.

Martin AP, Palumbi SR (1993) Body size, metabolic-rate, generation time, and the molecular clock. *Proceedings of the*

National Academy of Sciences of the United States of America, **90**, 4087–4091.

McDonald C (2004) The invasive pest species *Ciona intestinalis* (Linnaeus, 1767) reported in a harbour in southern Western Australia. *Marine Pollution Bulletin*, **49**, 868–870.

Monniot C, Monniot F (1994) Additions to the Inventory of Eastern Tropical Atlantic Ascidians; Arrival of Cosmopolitan Species. *Bulletin of Marine Science*, **54**, 71–93.

Nei M, Maruyama T, Chakraborty R (1975) Bottleneck effect and genetic-variability in populations. *Evolution*, **29**, 1–10.

Nielsen EE, Kenchington E (2001) Prioritising marine fish and shellfish populations for conservation: a useful concept? *Fish and Fisheries*, **2**, 328–343.

Nydam ML, Harrison RG (2007) Genealogical relationships within and among shallow-water *Ciona* species (Ascidiacea). *Marine Biology*, **151**, 1839–1847.

Nydam ML, Harrison RG (2010) Polymorphism and divergence within the ascidian genus Ciona. *Molecular Phylogenetics and Evolution*, **56**, 718–726.

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

Palumbi SR (1992) Marine speciation on a small planet. *Trends in Ecology & Evolution*, **7**, 114–118.

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Application*, **13**, s146–s158.

Palumbi SR, Grabowsky G, Duda T, Geyer L, Tachino N (1997) Speciation and population genetic structure in tropical sea urchins. *Evolution*, **51**, 1506–1517.

Petersen JK, Svane I (1995) Larval dispersal in the ascidian *Ciona intestinalis* (L.). Evidence for a closed population. *Journal of Experimental Marine Biology and Ecology*, **186**, 89–102.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

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

Procaccini G, Pischetola M, Lauro RD (2000) Isolation and characterization of microsatellite loci in the ascidian *Ciona intestinalis* (L.). *Molecular Ecology*, **9**, 1924–1926.

Ramsay A, Davidson J, Landry T, Arsenault G (2008) Process of invasiveness among exotic tunicates in Prince Edward Island, Canada. *Biological Invasions*, **10**, 1311–1316.

Raymond M, Rousset F (1995) genepop (version 1.2): population genetic software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Ricciardi A (2006) Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distribution*, **12**, 425–433.

Rice RW (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rocha RM, Kremer LP, Baptista MS, Metri R (2009) Bivalve cultures provide habitat for exotic tunicates in southern Brazil. *Aquatic Invasions*, **4**, 195–205.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosati F, Santis RD (1978) Studies on fertilization in the ascidians. I. Self-sterility and specific recognition between gametes of Ciona intestinalis. *Experimental Cell Research*, **112**, 111–119.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Sauvage C, Nierne N, Lapègue S, Boudry P (2007) Single nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene*, **406**, 13–22.

Seo KS, Lee Y (2009) A First Assessment of Invasive Marine Species on Chinese and Korean Coasts. *Biological Invasions in Marine Ecosystems*, **204**, 577–585.

Small KS, Brudno M, Hill MM, Sidow A (2007) Extreme genomic variation in a natural population. *Proceedings of the National Academy of Sciences USA*, **104**, 5698–5703.

Sordino P, Andreakis N, Brown ER *et al*. (2008) Natural variation of model mutant phenotypes in *Ciona intestinalis*. *PLoS ONE*, **3**, e2344.

Sorte CJ, Williams SL, Carlton JT (2010) Marine range shifts and species introductions: comparative spread rates and community impacts. *Global Ecology and Biogeography*, **19**, 303–316.

Suzuki M, Nishikawa T, Bird A (2005) Genomic approaches reveal unexpected genetic divergence within *Ciona intestinalis*. *Journal of Molecular Evolution*, **61**, 627–635.

Svane I, Havenhand JN (1993) Spawning and dispersal in *Ciona intestinalis* (L.). *Marine Ecology*, **14**, 53–66.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science*, **299**, 107–109.

Tepolt CK, Darling JA, Bagley MJ, Geller JB, Blum MJ, Grosholz ED (2009) European green crabs (*Carcinus maenas*) in the northeastern Pacific: genetic evidence for high population connectivity and current-mediated expansion from a single introduced source population. *Diversity and Distributions*, **15**, 997–1009.

Therriault TW, Herborg L (2008) A qualitative biological risk assessment for vase tunicate *Ciona intestinalis* in Canadian waters: using expert knowledge. *ICES Journal of Marine Science*, **65**, 781–787.

Tremblay MJ, Loder JW, Werner FE, Naimie CE, Page FH, Sinclair MM (1994) Drift of sea scallop larvae *Placopecten magellanicus* on Georges Bank: a model study of the roles of mean advection, larval behaviour and larval origin. *Deep Sea Research*, **41**, 7–29.

Vinson JP, Jaffe DB, O'Neill K *et al*. (2005) Assembly of polymorphic genomes: algorithms and application to *Ciona savignyi*. *Genome Research*, **15**, 1127–1135.

Waters JM, Dijkstra LH, Wallis GP (2000) Biogeography of a southern hemisphere freshwater fish: how important is marine dispersal? *Molecular Ecology*, **9**, 1815–1821.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.

Winkler G, Dodson JJ, Lee CE (2008) Heterogeneity within the native range: population genetic analyses of sympatric invasive and noninvasive clades of the freshwater invading copepod *Eurytemora affinis*. *Molecular Ecology*, **17**, 415–430.

Zhan A, Bao Z, Hui M *et al.* (2007) Inheritance pattern of EST-SSRs in self-fertilized larvae of the bay scallop *Argopecten irrandians*. *Annales Zoologici Fennici*, **44**, 259–268.

Zhan A, Hu J, Hu X *et al.* (2009) Fine-scale population genetic structure of Zhikong scallop (*Chlamys farreri*): do local marine currents drive geographical differentiation? *Marine Biotechnology*, **11**, 223–235.

AZ is interested in conservation and evolutionary genetics of aquatic animals. HJM studies vectors of biological invasions. MEC investigates the evolutionary causes and consequences of aquatic invasions with special emphasis on the genetics of habitat transitions.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Genetic diversity at eight microsatellite loci for two highly invasive *Ciona intestinalis* species, spA and spB

Table S2 Estimates of population genetic differentiation (pairwise Φ_{ST}) based on combined cytochrome c oxidase subunit 3—NADH dehydrogenase subunit 1 region (COX3-ND1) and NADH dehydrogenase subunit 4 gene (ND4) for *Ciona intestinalis* spA and spB. Bold numbers indicate statistical significance after sequential Bonferroni corrections

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.