


ORIGINAL RESEARCH

RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic

Nicholas W. Jeffery¹  | Claudio DiBacco² | Mallory Van Wyngaarden³ | Lorraine C. Hamilton⁴ | Ryan R. E. Stanley² | Renée Bernier⁵ | Jennifer FitzGerald² | K. Matheson¹ | C. H. McKenzie¹ | Praveen Nadukkalam Ravindran⁶ | Robert Beiko⁶ | Ian R. Bradbury¹

¹Northwest Atlantic Fisheries Centre, Fisheries and Oceans Canada, St. John's, NL, Canada

²Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS, Canada

³Ocean Sciences Center and Biology Department, Memorial University of Newfoundland, St John's, NL, Canada

⁴Aquatic Biotechnology Laboratory, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada

⁵Gulf Fisheries Centre, Fisheries and Oceans Canada, Moncton, New Brunswick, Canada

⁶Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada

Correspondence

Nicholas W. Jeffery, Northwest Atlantic Fisheries Centre, Fisheries and Oceans Canada, St. John's, NL, Canada.
Email: nick.jeffery@dfo-mpo.gc.ca

Funding information

Natural Sciences and Engineering Research Council of Canada

Abstract

Genomic studies of invasive species can reveal both invasive pathways and functional differences underpinning patterns of colonization success. The European green crab (*Carcinus maenas*) was initially introduced to eastern North America nearly 200 years ago where it expanded northwards to eastern Nova Scotia. A subsequent invasion to Nova Scotia from a northern European source allowed further range expansion, providing a unique opportunity to study the invasion genomics of a species with multiple invasions. Here, we use restriction-site-associated DNA sequencing-derived SNPs to explore fine-scale genomewide differentiation between these two invasions. We identified 9137 loci from green crab sampled from 11 locations along eastern North America and compared spatial variation to mitochondrial COI sequence variation used previously to characterize these invasions. Overall spatial divergence among invasions was high (pairwise F_{ST} ~0.001 to 0.15) and spread across many loci, with a mean F_{ST} ~0.052 and 52% of loci examined characterized by F_{ST} values >0.05. The majority of the most divergent loci (i.e., outliers, ~1.2%) displayed latitudinal clines in allele frequency highlighting extensive genomic divergence among the invasions. Discriminant analysis of principal components (both neutral and outlier loci) clearly resolved the two invasions spatially and was highly correlated with mitochondrial divergence. Our results reveal extensive cryptic intraspecific genomic diversity associated with differing patterns of colonization success and demonstrates clear utility for genomic approaches to delineating the distribution and colonization success of aquatic invasive species.

KEYWORDS

Carcinus maenas, COI, European green crab, population structure, restriction-site-associated DNA sequencing

1 | INTRODUCTION

Global invasions of plants and animals are occurring at an increasing rate accelerated by human impacts and environmental change

(Levings, Kieser, Jamieson, & Dudas, 2002; Mooney & Cleland, 2001; Ricciardi, 2006). Invasions can have significant ecological and evolutionary consequences at the species (McDonald, Parchman, Bower, Hubert, & Rahel, 2008), community (Matheson et al., 2016; Robinson

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

& Dickerson, 1987), and ecosystem levels (Hänel & Chown, 1998). Ecological impacts have been well studied in coastal marine habitats, which represent some of the most heavily invaded ecosystems due to human-mediated introductions from ship ballast and hull fouling, as well as the aquarium trade and aquaculture activities (Grosholz, 2002). Ecological impacts of aquatic invaders range from large-scale reductions in native biodiversity (Molnar, Gamboa, Revenga, & Spalding, 2008) to the local destruction of marine habitats such as eelgrass or shellfish beds (Garbary, Miller, Williams, & Seymour, 2014; Matheson et al., 2016; McClenachan, O'Connor, & Reynolds, 2015), which can include commercially important species (Floyd & Williams, 2004). In contrast to ecological impacts, evolutionary dynamics of marine invasions are less well studied including both the genomic basis for invasiveness (Lee, 2002) and evolutionary consequences of invasion (Chown et al., 2015; Cristescu, 2015). Increasingly, it is becoming apparent that cryptic diversity can contribute to variation in invasion success and ultimately to impacts on marine and freshwater ecosystems (e.g., Folino-Rorem, Darling, & D'Ausilio, 2009; Lee, 1999; Sherman et al., 2016). A better understanding of the presence of intraspecific diversity within invasive species and how it correlates with variation in invasion success is necessary to successfully manage invaders and mitigate ecological and evolutionary impacts.

An evolutionary perspective of marine invasion dynamics and of cryptic diversity within invaders has been limited by the huge interspecific diversity among invasive taxa and a lack of genomic resources. The ability to study "invasion genomics" using relatively new tools such as restriction-site-associated DNA sequencing (RAD-seq), RNA-seq genotyping, or transcriptomics allows the exploration of evolutionary history, connectivity, range expansion rates, and invasion patterns with markers from across the entire genome without the requirement of an existing whole-genome sequence (Chown et al., 2015; Rius & Darling, 2014; Sherman et al., 2016; Viard, David, & Darling, 2016). Such genomewide sampling techniques can be used to identify so-called invasion genes or marker associations with environmental factors (Chown et al., 2015; Wang et al., 2011) and can be useful in understanding patterns and rates of expansion (e.g., Krehenwinkel & Tautz, 2013), as well as the roles of neutral and adaptive pressures in invasion dynamics (e.g., Tepolt & Palumbi, 2015). Genomewide markers are particularly useful for marine invasive species whose adult and/or larval phases may be transported by coastal currents or by human-mediated vectors, and thus, invasion genomics can greatly facilitate our understanding of the processes behind the initial invasion and subsequent impacts at a scale not attained with typical genetic markers.

European green crab (*Carcinus maenas*) are a notorious invasive species now present on every continent with the exception of Antarctica (Roman, 2006). Green crab have a prolonged larval phase and high tolerance of variable environmental conditions allowing for further expansion throughout their introduced ranges, including both the Pacific and Atlantic coasts of North America (Roman, 2006). In the northwest Atlantic, green crab likely experienced multiple cryptic introductions, having invaded the eastern USA and Canada at least twice from genetically distinct northern and southern European populations in 1817 and again in the 1980s (Darling, Tsai, Blakeslee, & Roman,

2014; Roman, 2006). This species' range is continuing to expand, having recently been discovered in Placentia Bay, Newfoundland, in 2007, with a likely introduction c. 2002 from a Scotian Shelf population (Blakeslee et al., 2010). Green crab population structure has been investigated previously using traditional genetic markers, including mitochondrial DNA (16S, Geller, Walton, Grosholz, & Ruiz, 1997; COI, Darling, Bagley, Roman, Tepolt, & Geller, 2008; Roman & Palumbi, 2004; Roman, 2006) and microsatellites (Darling, 2011; Darling et al., 2008, 2014; Pascoal et al., 2009; Tepolt, Bagley, Geller, & Blum 2006). In their native range, green crab show distinct broadscale structuring of island populations (i.e., Faeroe Islands and Iceland) and northern and southern European populations (Roman & Palumbi, 2004) using COI sequence data and fine-scale (~450 km) structuring has been reported using microsatellite loci (Pascoal et al., 2009). In North America, spatial structuring is driven by both range expansion (Pringle, Blakeslee, Byers, & Roman, 2011) and secondary contact among invasions (Blakeslee et al., 2010; Darling et al., 2014; Tepolt & Palumbi, 2015). However, a genomewide comparison (i.e., RAD-seq) of fine-scale spatial structuring on the east coast has been absent and is warranted to compare these independent invasions and characterize the full genetic diversity of populations along this part of their invasive range.

Our overall objective was to evaluate population structuring using genomewide SNPs obtained by RAD-seq for green crab from 11 locations in eastern North America to explore the genomic extent of differences between colonization events. Specifically, our goals were to (i) quantify genomewide differentiation among samples encompassing both the invasions in eastern North America from northern and southern European source populations, (ii) describe the spatial distribution of these two invasions using a large number of genomewide SNPs, and (iii) compare spatial variation at both highly divergent and selectively neutral SNPs, as well as the mitochondrial cytochrome oxidase I (COI) gene to better understand the evolutionary dynamics associated with the recent invasion and northwards expansion of the species. We predict (i) widespread genomic differentiation among these independent invasions based on existing work and the geographic isolation of the source populations and (ii) that larger numbers of loci used here will reveal finer-scale population structure than has previously detected. We build directly on previous studies (e.g., Blakeslee et al., 2010; Darling et al., 2008, 2014; Tepolt & Palumbi, 2015) by comparing neutral and divergent SNPs to COI sequence data between the two invasions. The datasets developed here will allow for monitoring future range expansion of populations at both the southern and northern limits of the current range and also permit the formulation of conservation strategies that are region- and population-specific.

2 | MATERIALS AND METHODS

2.1 | Specimen collection and DNA extraction

Specimens were collected from 11 locations in eastern Canada and northeastern USA (Table 1, Figures 1 and 2). All samples were collected in 2011 with the exception of New Hampshire, which was sampled in 2013. Tissue samples were preserved in AllProtect (Qiagen,

TABLE 1 List of sampling locations, including latitude and longitude, for green crab obtained for this study. The location code is provided, along with the number of individuals per location for which we obtained a full SNP panel and COI sequence. All locations were sampled in 2011, with the exception of NWH which was sampled in 2013

Location name	Location code	Latitude	Longitude	nSNPs	nCOI
St. George's Bay, NL	SGB	48.492	-58.658	22	22
Placentia Bay, NL	PLB	47.102	-54.184	22	36
Baie de Bassin, QC	BDB	47.473	-61.738	22	22
Sydney Harbour, NS	SYH	46.141	-60.202	22	22
Mabou, NS	MBO	46.069	-61.391	22	22
Brudenell River, PE	BRN	46.192	-62.588	22	22
Cole Harbour, NS	CLH	44.658	-63.442	22	23
Kejimikujik, NS	KJI	43.840	-64.805	21	27
Campobello Island, NB	CBI	44.888	-66.918	22	22
New Hampshire	NWH	43.04	-70.71	22	21
Tuckerton, NJ	TKT	39.588	-74.288	22	21



FIGURE 1 The green crab (*Carcinus maenas*) is native to Europe and has invaded eastern North America at least twice independently over the past 200 years. The range of this species has since increased both northwards and southwards, being found from Virginia to southern Newfoundland

Toronto, ON, Canada) or 80% ethanol. DNA was isolated from the tissue samples using phenol: chloroform extraction or NucleoMag 96 Tissue (Macherey-Nagel, Bethlehem, PA, USA) following the manufacturer's protocol, including RNase A (Qiagen) treatment. All DNA samples were quantified using the Qubit dsDNA HS Assay Kit (Life Technologies, Burlington, ON, Canada) with assays read on a Qubit v2.0 (Life Technologies) or the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies) with assays read on a FLUOStar OPTIMA fluorescence plate reader (BMG Labtech, Ortenberg, Germany).

2.2 | mtDNA analysis

A 502-base pair fragment of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene was amplified using M13F (-21)/M13R (-27) tagged version of the primers described in Roman and Palumbi (2004) (M13F-CmaCOIF 5' TGTAACACGACGGCCAGTGTCTGAGCTGG CATAGTAGG 3' and M13R-CmaCOIR 5' CAGGAAACAGCTATGA CGAATGAGGTGTTAGATTTCG 3') (Integrated DNA Technologies,

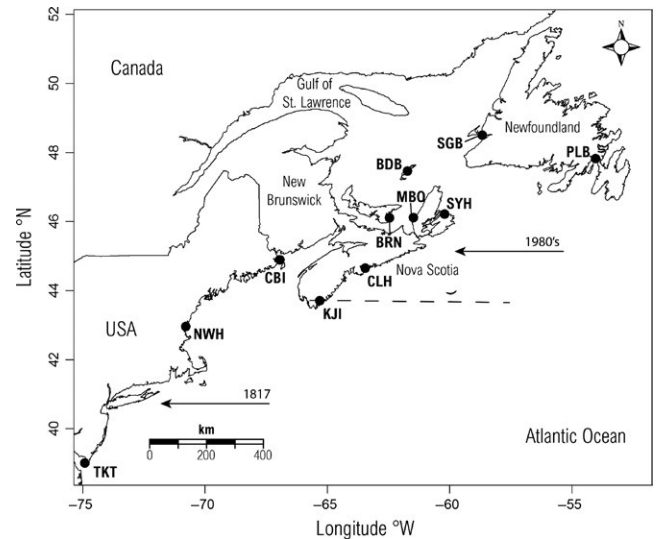


FIGURE 2 Map of eastern North America showing the 11 green crab sampling locations in the USA and Canada. The arrows represent the independent invasions of green crab into North America, and the dashed line indicates the approximate northern limit of the initial southern invasion to the New York area (1817) prior to the recent introduction of green crab to the Halifax region (1980s) (Audet et al., 2003)

Coralville, IA, USA). PCRs were carried out in 50 μ l volumes comprised of 100 ng DNA, 1 \times AmpliTaq 360 Buffer (Life Technologies), 1.5 mM MgCl₂, 0.2 mM each nucleotide (dNTP), 0.5 μ M each primer, and 0.025 U AmpliTaq 360 (Life Technologies). Thermal cycling conditions consisted of 94°C for 3 min; 30 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s; 1 cycle of 72°C for 10 min. To verify amplification, 5 μ l of the PCR product was electrophoresed on a 1% agarose gel, visualized using SYBR Safe (Life Technologies) and documented using a Gel Logic 200 (Kodak). Forty microliters of the remaining PCR product was purified using QIAquick PCR purification kit (Qiagen). The purified PCR products were quantified as described above and sequenced in both directions using M13F (-21) and M13R (-27) and BigDye Terminator v3.1 cycle sequencing kit (Life Technologies) per the manufacturer's instructions with the following modification, and 0.5 μ l of the BigDye Terminator v3.1 Ready Reaction Mix (Life Technologies)

and 3.8 μl of the 5 \times sequencing buffer (Life Technologies) was used per 20 μl reaction. Sequencing products were separated by size on a 36-cm array containing POP7 polymer on an AB3130xl genetic analyzer (Life Technologies). The sequencing results from both directions were assembled using Vector NTI Advance v11 (Life Technologies). For identification, the resulting sequence for each individual was then aligned to reference sequences, haplotypes H1–H6 and H8–H10 from Darling et al. (2008) (Genbank accession numbers FJ159008–FJ15016) using Vector NTI Advance v11 (Life Technologies).

2.3 | RAD sequencing

RAD-seq libraries were prepared as described by Etter, Preston, Bassham, Cresko, and Johnson (2011b; see also Etter, Bassham, Hohenlohe, Johnson, & Cresko, 2011a with modifications). All samples were normalized to 25 ng/ μl . The DNA quality for all samples was verified by agarose gel electrophoresis of 100 ng of extracted DNA on a 0.7% agarose gel (DNA was visualized and documented as described above). Only DNA samples with a visible high molecular weight band on the agarose gel and that were positive for an M13F-CmaCOIF/M13R-CmaCOIR amplification product (see above) were used for library preparation. Each library was comprised of DNA samples from 22 individuals (1 μg DNA per individual) from the same geographic location with each individual sample having a different in-line barcode in the P1 adapter. The P1 adapter in-line barcodes ranged from 5 to 9 bp in length, and the sequences were chosen to ensure equal distribution of all nucleotides at each base position (including those that overlap with the restriction site) and maximizing the edit distance (Faircloth & Glenn, 2012). Based on edittags analysis (Faircloth & Glenn, 2012), the variable-length barcodes had an edit distance ranging from 2 to 8 and a modal edit distance of 6. Gel size selection performed after sonication and PCR amplification was done on a Pippin Prep (Sage Science, Beverly, MA, USA) using the 2% agarose gel cassette (Sage Science) and size selection range of 300–500 bp. PCR amplification was done using Q5 Hot Start Master Mix (NEB, Whitby, ON, Canada) for all libraries. Amplification cycles for all libraries were 98°C for 30 s; 14 cycles of 98°C for 30 s, 65°C for 30 s, 72°C for 30 s; 1 cycle of 72°C for 5 min. All libraries were sequenced on a HiSeq 2000 (Illumina) as 100-bp paired end sequences with one library per lane. Sequencing was performed at McGill University and Génome Québec Innovation Centre, Montréal, Canada.

2.4 | Bioinformatics

RAD data for 242 individuals were cleaned and demultiplexed using *process_radtags* in Stacks v.1.21 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). We used *ustacks* for de novo loci formation with a minimum depth of 5 to create a stack and a maximum distance of 4 between stacks. One individual from Kejimikujik (KJI) was removed due to low sequence coverage, and so, we retained 241 individuals overall. We then used *cstacks* to construct a locus catalog based on sequence identity with 1 mismatch allowed between stacks. We used settings of $M = 4$ as the maximum distance between two loci

to merge and $n = 1$ to allow two catalog loci to merge in case they are differently fixed versions of the same locus. All RAD tags were 80 bp in length. These were then filtered in the Stacks *populations* module to include only RAD tags present in each population and in $\geq 75\%$ of individuals and a minor allele frequency (MAF) greater than 5%. This dataset was filtered for missing data and deviations from Hardy–Weinberg equilibrium using PLINK (Purcell et al., 2007), where individuals missing $>25\%$ of loci and loci missing $>5\%$ of genotypes were removed. Finally, only one SNP per RAD tag was retained to remove the effects of physical linkage. The resulting panel was converted for subsequent analyses using PGDSpider v2.0.8.3 (Lischer & Excoffier, 2012) and the R package *genepopedit* (Stanley, Jeffery, Wringe, DiBacco, & Bradbury, 2017).

2.5 | Population statistics and detecting divergent loci

Multiple approaches were used to identify loci putatively experiencing adaptive selection (outliers). We first employed BayeScan v2.1 (Foll & Gaggiotti, 2008) with 50,000 burn-in and 100,000 iterations, with prior odds set to 10. Second, we used a nonhierarchical island model in Arlequin v.3.5.2.1 (Excoffier & Lischer, 2010) to detect outlier loci using 200,000 simulations and 100 demes simulated with default expected heterozygosity settings. Third, we used the hierarchical island model in Arlequin with the same settings as the nonhierarchical model and 10 simulated groups to further test for outlier loci. While these methods have been shown to produce higher numbers of false positive outliers in some demographic scenarios, they also show increased power and perform similarly to other methods such as FLK (Bonhomme et al., 2010) under recent demographic expansion models (Lotterhos & Whitlock, 2014). Only loci identified as outliers by all three methods were considered to reduce the number of false positives, and our putatively neutral panel was comprised only of SNPs that were not detected as outliers by any of these methods (Figure 3). Outlier loci were subset from the full RAD panel using the *subset_genepop* function in *genepopedit* (Stanley et al., 2017). Observed (H_o) and expected (H_e) heterozygosities were calculated in the package *hierfstat* (Goudet, 2014) in R (R Core Team 2015) for each population within each SNP panel, as well as the average H_o and H_e for each panel. Locus-specific heterozygosities and F_{ST} values were computed in *hierfstat* and *diveRsity* (Keenan, McGinnity, Cross, Crozier, & Prodöhl, 2013), respectively, with 1,000 bootstrap replicates. Heatmaps of linkage disequilibrium r^2 values and standardized allele frequencies, where the highest major allele frequency per SNP among populations is given a value of 1.0 and the lowest major allele frequency a value of 0, for outlier and neutral SNPs, were created using custom R scripts.

2.6 | Population structure

Spatial structuring in neutral and outlier SNPs was explored using several methods. We conducted a discriminant analysis of principal components (DAPC) using *adegenet* 2.0.0 (Jombart, 2008; Jombart & Ahmed, 2011) in R to first find the number of clusters to best

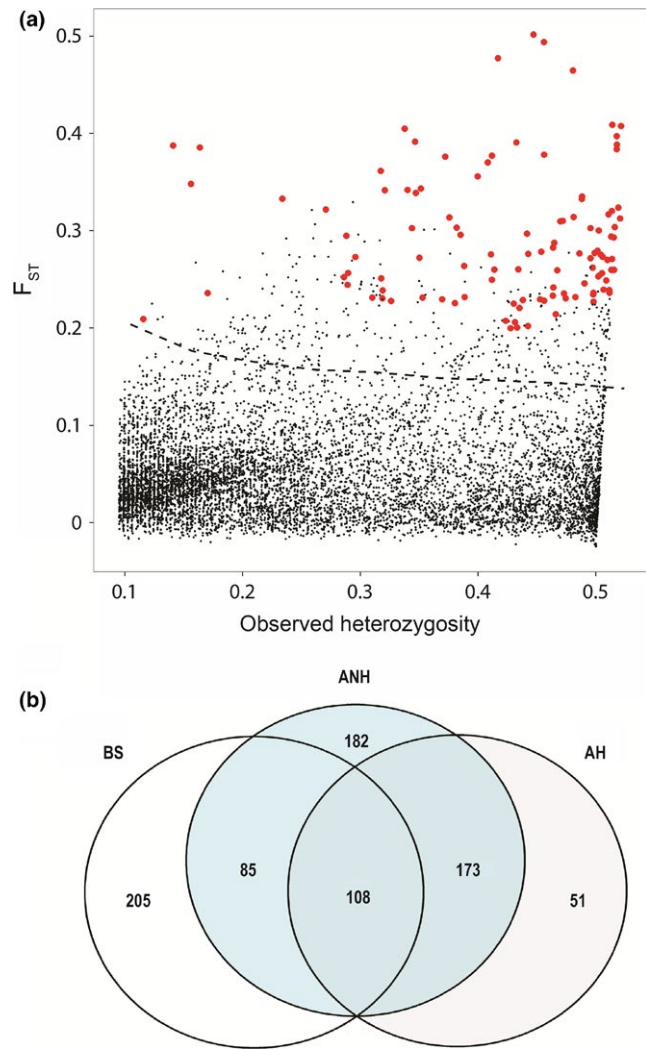


FIGURE 3 Outliers detected in the current study. (a) Locus-specific F_{ST} values against heterozygosity from the nonhierarchical island model implemented in Arlequin. Loci common to all three detection methods are colored in red. The dashed line indicates the 1% quantile for outlier detection which contains all 108 outlier loci. (b) The overlap of outlier loci among BayeScan (BS) and Arlequin's nonhierarchical (ANH) and hierarchical (AH) island models. Note that the area of overlap between methods is not proportional to the number of SNPs in common

represent our 11 populations using the *find.clusters* function. All principal components were retained at this step. This function defines the best number of clusters using a Bayesian information criterion (BIC). We then conducted the DAPC using the defined clusters as prior information and retained the first 50 principal components to determine the best value of K for our data. The DAPC was conducted on the full SNP panel, as well as on the neutral and outlier datasets separately. Second, we used principal components analysis on the outlier and neutral SNP datasets to account for intrapopulation variance and compare with the results of our DAPC. Finally, Cavalli-Sforza and Edwards' (1967) chord distances were calculated from each panel of SNP data in POPULATIONS 1.2.33 (Langella, 2012). Neighbor-joining (NJ) trees

based on chord distances were constructed with 1,000 bootstrap replicates for all three SNP panels, and these were midpoint-rooted and visualized in FigTree v1.4 (<http://tree.bio.ed.ac.uk/software/figtree/>). Geographic visualizations of F_{ST} similarities were built using GenGIS version 2.5 (PMID 23922841). Maps used in GenGIS were obtained from the Natural Earth project (<http://www.naturalearthdata.com>).

Spatial structuring in the SNP and COI datasets were also directly compared. We used Arlequin v3.5.2 (Excoffier & Lischer, 2010) to calculate pairwise F_{ST} values between each sampling location for the COI data and each of the three SNP panels. We then used a Mantel test with 1,000 permutations in the R package *ade4* (Dray & Dufour, 2007) to compare pairwise F_{ST} values between the COI data and the full SNP panel. Additionally, we conducted a Mantel test with 1,000 permutations on linearized pairwise F_{ST} (Slatkin, 1995) values and geographic distance in *ade4* for the full SNP panel and the COI dataset as a test of isolation by distance. Least-cost distances between each pair of sampling locations were estimated in *marmap* (Pante & Benoit, 2013) implemented in R. An analysis of molecular variance (AMOVA) with 1,000 permutations was conducted for each dataset in Arlequin based on our optimal $K = 2$, where each sampled population was considered part of either a northern or southern group.

3 | RESULTS

Following initial filtering in Stacks, we obtained a catalog of 46,067 SNPs per individual crab. Additional filtering in Stacks for data present in 75% of individuals and a MAF of 0.05 retained 15,847 SNPs. A final filtering step in PLINK for missing genotypes yielded our full SNP panel of 9,137 loci resulting in a full SNP panel for 241 individual crabs sampled across the 11 sites. Overall observed heterozygosity (H_o) for the full SNP panel (Figure S1A) was similar to that expected (H_e) (0.254 versus 0.256, respectively). Average H_o for the neutral panel (0.254) was lower than that of the most divergent SNPs (0.290, see below). The global unbiased F_{ST} for the full SNP panel was 0.052. Locus-specific F_{ST} values ranged from 0.00001 to 0.551, and heterozygosities ranged from 0.095 to 0.564 for the full SNP panel (Figure S1B). More than half (52%) of the loci exhibited an $F_{ST} > 0.05$.

3.1 | Outlier SNP detection

We utilized three detection methods to obtain our outlier SNP panel. BayeScan detected 398 SNPs putatively under divergent selection (~4.4% of all loci). The nonhierarchical island model based method in Arlequin detected 548 SNPs under selection, while the hierarchical island model based method detected 332 SNPs; 281 loci were common to both of the models implemented in Arlequin. Among all three methods of outlier detection, there were 108 (1.2% of all SNPs) putatively adaptive loci (Figure 3). Outlier loci-specific F_{ST} values ranged from 0.199 to 0.551 with a mean F_{ST} of 0.291 relative to 0.043 for neutral SNPs (Figure S2). Our final neutral panel of SNPs was comprised of 8,326 loci after removing any outliers identified by all of the three outlier detection methods.

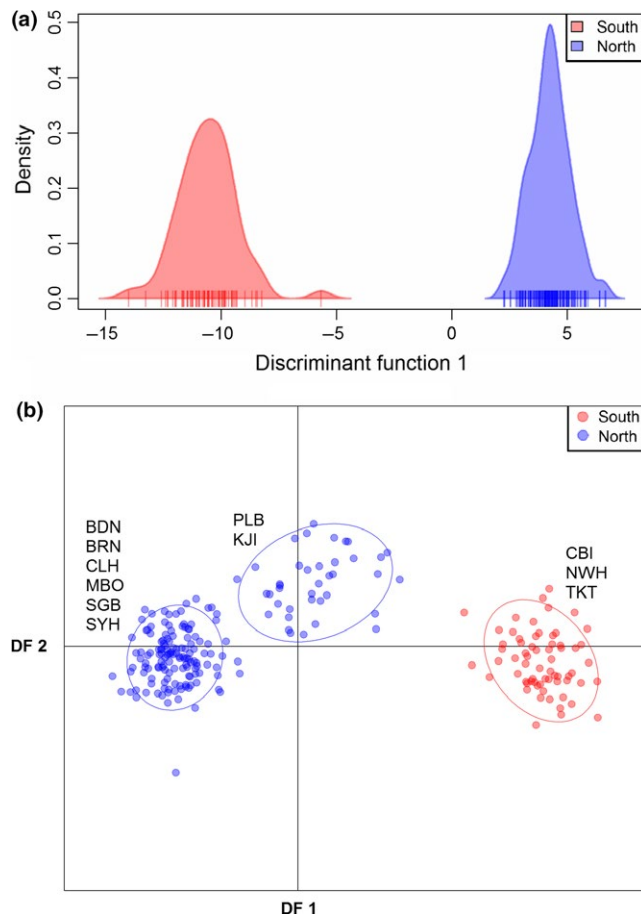


FIGURE 4 Discriminant analyses of principal components for the (a) neutral SNP panel where $K = 2$, displaying the northern and southern populations along a single discriminant function, and (b) outlier SNP panel where all 11 sites are separated into $K = 3$ clusters spatially distributed along the first two discriminant functions. DAPC on the full SNP panel showed the same results as the neutral panel

3.2 | Population structure

DAPC on the full SNP panel indicated the presence of two clusters based on the Bayesian information criterion ($K_{stat} = 1695.06$, Figure 4a). The DAPC had a proportion of conserved variance of 0.442 explained by a single discriminant function for the full RAD panel. Both clusters, representing northern and southern groups, had 100% assignment proportions. Similar results were obtained with the neutral SNP panel where again $K = 2$ was best supported ($K_{stat} = 1666.30$), with a proportion of conserved variance of 0.779 explained by a single discriminant function and 100% correct assignment for both groups. In contrast, DAPC of outlier SNPs suggested a range of clusters from $K = 2$ to 3 ($K_{stat} = 692.69$ – 693.31 , respectively) and had a proportion of conserved variance of 0.925 (Figure 4b). DAPC analyses based on $K = 2$ and $K = 3$ showed the same geographic partitioning as the full and neutral panels, but with KJI and PLB (Cluster 2) intermediate along discriminant function 1 to the pure northern and southern clusters (Clusters 1 and 3; Figure 4b). Principal component analyses on outlier and neutral SNPs revealed the same overall population structure as the DAPC, with clear separation of

the northern and southern populations along the first principal component for both the neutral and outlier locus datasets (Figure S3).

Neighbor-joining dendrograms of chord distances showed high support for northern and southern populations for each SNP panel. Bootstrap support for each cluster using the outlier SNP set was 100, with PLB clustering with the southern population. The neutral SNP panel had bootstrap support values of 52–100, but PLB clustered with the northern population in this case (Figures 5 and S4). Bootstrap support for the northern populations was 100 when excluding PLB in all cases. The chord distances for the outlier tree was twice the magnitude of the full or neutral SNP panels.

Standardized allele frequencies of the outlier loci and neutral loci revealed a clear latitudinal cline overall with a sharp change in allele frequency occurring between CBI and KJI (Figures 6a and S5). Interestingly, PLB shared more alleles with southern sites than northern sites (Figures 6a and S5). Linkage disequilibrium r^2 values among outliers ranged from 4.75×10^{-7} to 0.92 following removal of loci located on the same RAD tag with r^2 values of 1.0 (Figure 6b).

3.3 | Marker comparisons

Haplotype distributions displayed a clear transition from dominance of haplotype H01 in the southern population to dominance of haplotypes H04 to H06 in the northern population (Figure 7). However, approximately 25% of the crabs collected at KJI and PLB were of the southern haplotype. This geographic split mirrors that exhibited by the spatial clustering analysis of the RAD-seq data above. Pairwise F_{ST} values were significantly different between northern and southern populations for all datasets (Table 2, Figure S6). F_{ST} values ranged from 0 to 0.49 for the mitochondrial data and from 0.001 to 0.15 for the full SNP panel. Pairwise F_{ST} values for the outlier SNPs were similar in range to the COI data, ranging from 0.005 to 0.58, while values for the neutral SNPs were similar to the full SNP panel, ranging from 0.001 to 0.14. Mantel tests on linearized pairwise F_{ST} values (Slatkin, 1995) versus least-cost geographic distances between sample sites showed a significant positive correlation for the full SNP panel ($r = .624$, $p = .004$) and the COI data ($r = .60$, $p = .008$; Figure 8a). Mantel tests comparing pairwise F_{ST} values for COI to each SNP panel were highly significant ($r = .93$, outlier SNPs; $r = .95$, neutral/full SNPs, $p < .001$; Figure 8b). AMOVA showed significant genetic differentiation between northern and southern populations when using the neutral SNP panel ($F_{CT} = 0.058$, $p = .004$; Table 3a) and the outlier SNPs ($F_{CT} = 0.389$, $p = .001$; Table 3b). The outlier SNPs explained substantially more of the variation among groups (38.86% versus 5.76% for the neutral SNPs). Consistent with the SNPs, AMOVA based on the COI variation showed significant genetic differentiation between northern and southern populations ($F_{CT} = 0.197$, $p = .002$; Table 3c), with 19.65% of the variation among groups.

4 | DISCUSSION

The use of genomic approaches to study invasive species demographics and evolutionary history has become increasingly common with

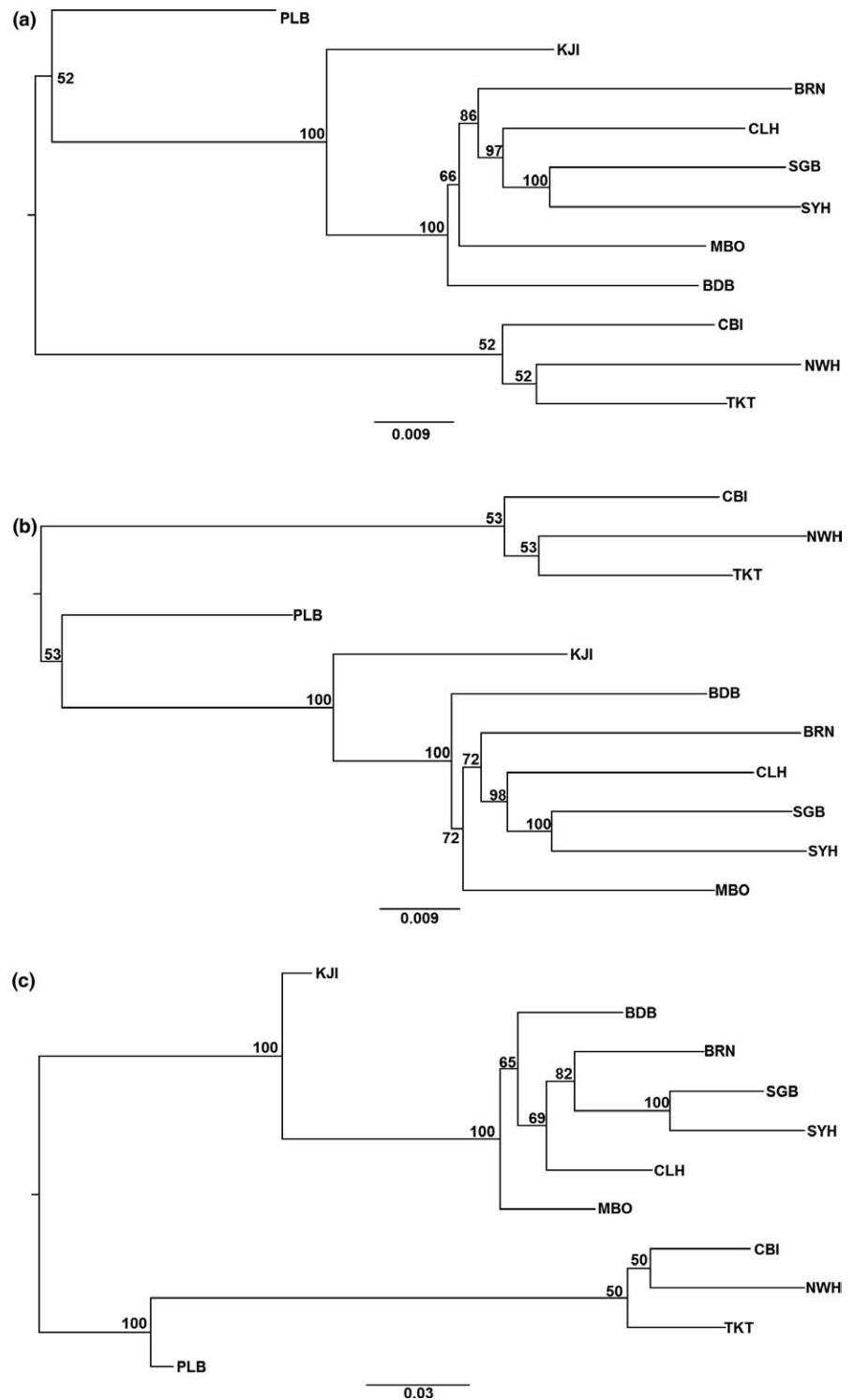


FIGURE 5 Neighbor-joining trees based on Cavalli-Sforza and Edwards (1967) chord distances for three panels of SNPs: (a) the full SNP panel, (b) putatively neutral SNPs, and (c) putatively adaptive SNPs. Values at the nodes represent bootstrap values from 1,000 pseudoreplicates

high-throughput sequencing technology, building upon previous studies that have used traditional genetic markers such as mitochondrial DNA and microsatellites (Cristescu, 2015). We used RAD-seq to generate a panel of genomewide SNPs to examine fine-scale population structuring in invasive European green crab along the east coast of North America. Our results indicate unusually high levels of genomewide differentiation among invasions consistent with observed differences in northern invasion success. We observed strong evidence

for two genetically and geographically distinct populations of green crab spatially segregated in eastern North America corresponding to two independent invasions (Darling et al., 2008, 2014), which show clear latitudinal clines in RAD-seq-derived SNP allele frequency. As observed in other species (e.g., Atlantic salmon (*Salmo salar*), Freemo, O'Reilly, Berg, Lien, & Boulding, 2011; Moore et al., 2014; sea scallop (*Placopecten magellanicus*), Van Wyngaarden et al., 2017), our most divergent SNPs (i.e., outliers) explained more of the spatial variance

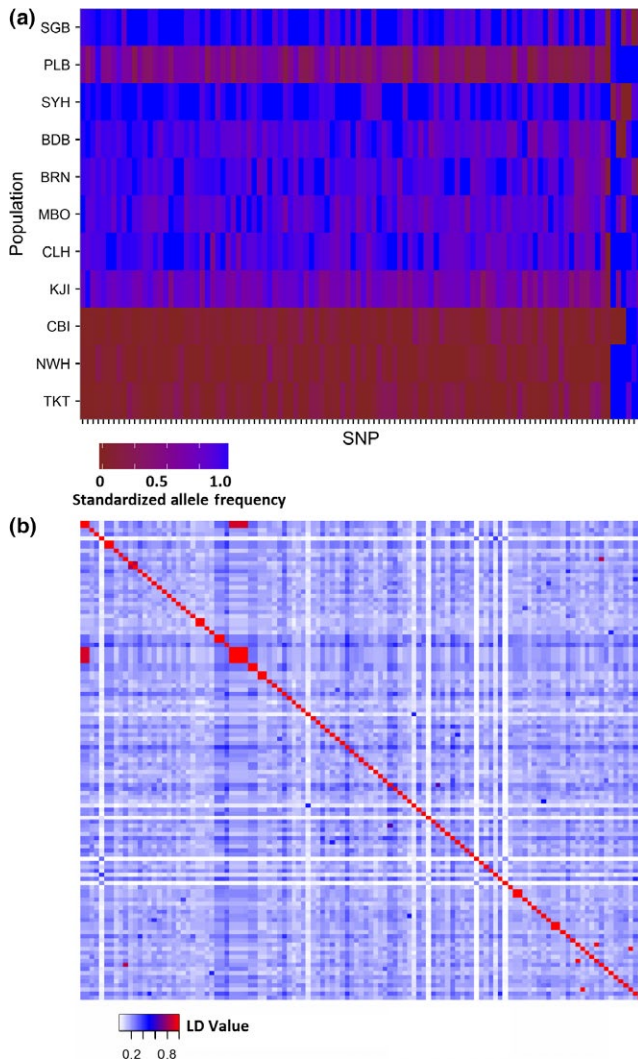


FIGURE 6 (a) Spatial clines in standardized allele frequency at each outlier locus for the 11 sampling sites sorted by latitude. (b) Heat map of linkage disequilibrium r^2 values for each outlier locus. Those loci with an r^2 of 1.0 were located on the same RAD tag, and so only, one locus per RAD tag was retained

in allele frequency than either background (neutral) SNPs or COI data. Interestingly, genomewide divergence was so prominent that background genomic variation (i.e., neutral SNPs) explained an order of magnitude more of the variation between northern and southern groups than observed in native species in the region (e.g., sea scallop, Van Wyngaarden et al., 2017), and locus-specific F_{ST} values were an order of magnitude larger across the genome than observed in native species [e.g., American lobster, *Homarus americanus*, Benestan et al. (2015) or sea scallop]. This work builds directly on previous studies describing the spatial variation of these two invasions by providing enhanced genomic and geographic resolution of the differences. The extent and magnitude of differentiation across the genome was unexpectedly high for a marine species, which suggests long-standing and deep divergence exists between the independent invasions which has been maintained in the northern and southern European source populations.

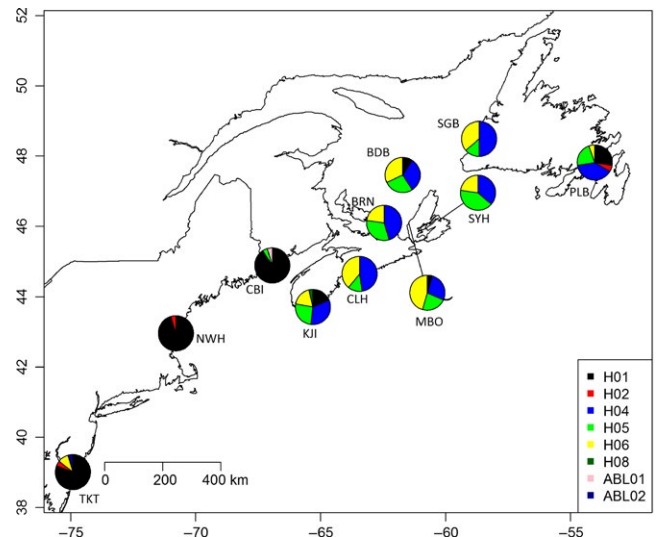


FIGURE 7 Mitochondrial DNA (COI) haplotypes mapped onto each location, showing a transition in haplotype frequency from the southern green crab population to the northern population starting in Nova Scotia, Canada

4.1 | Overall spatial population structure

Our results clearly reveal high levels of genomewide differentiation separating the two invasions into northern and southern groups. This is consistent with the findings of Tepolt and Palumbi (2015) who suggest secondary contact as the dominant structuring force in the introduced range. The presence of two distinct groups of green crab has been consistently reported previously using mtDNA (Roman, 2006), microsatellites (Darling et al., 2014), and RNA-seq SNPs (Tepolt & Palumbi, 2015). This broad structuring of the two independent invasions in turn mirrors the spatial structuring of the native European range and is likely maintained by environmental differences between the northern and southern regions (Roman, 2006). Our spatial analysis did reveal that two locations (PLB and KJI) appear intermediate between the northern and southern clusters which may be indicative of admixture and hybridization (Darling et al., 2014; Tepolt & Palumbi, 2015). One of these locations, Placentia Bay, is characterized by heavy shipping traffic, and the introduction of a previously admixed population from the Scotian Shelf is likely being maintained at this site (Blakeslee et al., 2010). The other location, Kejimikujik, represents the area where the two invasions (north and south) are coming into contact which appears to be the likely source of observed hybridization (Darling et al., 2014). Previous studies suggest the possible southward movement of this contact zone over time (Darling et al., 2014; Pringle et al., 2011) so the temporal stability of the patterns observed here remains to be seen.

The presence of these two groups was strongly supported by RAD-seq loci distributed across the genome. Of interest is the fact that the small panel of highly differentiated loci (i.e., outliers, ~1.2% of loci) detected the same level of population structuring as the full panel of SNPs, reinforcing the use of relatively few “diagnostic” markers for fine geographic scale population identification. Similar conclusions were

TABLE 2 Pairwise F_{ST} values for each population for the COI data (below black diagonal) and the full SNP panel (above black diagonal) ordered by latitude from south to north. Negative F_{ST} values were set to zero. Values in bold are significant at $p < .01$, and shaded cells represent north versus south pairwise comparisons

	TKT	NWH	CBI	KJI	CLH	MBO	BRN	SYH	BDB	SGB	PLB
TKT		0.00353	0.0021	0.07212	0.10226	0.13414	0.10207	0.11152	0.11436	0.12526	0.05022
NWH	0		0.00345	0.08214	0.1135	0.14767	0.11491	0.12421	0.12712	0.13787	0.05717
CBI	0	0.00264		0.07215	0.10227	0.13382	0.10289	0.11355	0.11554	0.12629	0.05061
KJI	0.19868	0.25674	0.23006		0.00732	0.02052	0.00788	0.01139	0.01605	0.01852	0.02758
CLH	0.31817	0.38925	0.36155	0		0.00429	0.00087	0.00151	0.00608	0.00403	0.03393
MBO	0.34707	0.42742	0.40582	0.01056	0.01821		0.00518	0.00346	0.00986	0.00131	0.0547
BRN	0.38332	0.44831	0.42356	0	0	0.02062		0.00359	0.00718	0.00701	0.03851
SYH	0.42181	0.48735	0.46673	0.01194	0.01843	0.00722	0		0.00824	0.00325	0.04136
BDB	0.32257	0.39699	0.36872	0	0	0	0	0		0.01077	0.04542
SGB	0.25572	0.32073	0.29568	0	0	0.00856	0	0.01187	0		0.0494
PLB	0.22428	0.27265	0.2275	0.01207	0.01468	0.11221	0.04152	0.09609	0.05281	0.01381	

reached by Moore et al. (2014) and Candy et al. (2015) for Atlantic salmon and Pacific eulachon (*Thaleichthys pacificus*), respectively, where outlier SNPs showed the same population structure as neutral SNPs. Observations of similar spatial patterns in both outlier and neutral SNPs supports conclusions of significant genomewide divergence of the groups observed. For instance, both Moore et al. (2014) and Candy et al. (2015) focus on anadromous fish species where gene flow may be low and genomewide divergence likely. Similarly, in green crab, the two independent invasions likely would, at least initially, have low levels of gene flow as they appear to originate from distant locations in Europe (Roman, 2006). Our observations of genomewide divergence support a hypothesis of significant and long-standing isolation between the two invasions. Ultimately, isolating mechanisms in both the native or introduced range, the temporal stability of these patterns, and the role of hybridization and selection in structuring these two invasions in North America will require further study and is beyond the scope of this work.

4.2 | Genetic marker comparisons

Previous studies have used COI sequence variation to study green crab population structure and to trace the invasion history (Darling, 2011; Darling et al., 2008, 2014; Roman, 2006). Overall, our SNP data were similar to COI in terms of resolving the current spatial structure of the invasions and suggests that patterns of divergence found in the mitochondria are widespread across the nuclear genome. We identified fewer COI haplotypes relative to previous studies in this system (Darling et al., 2008, 2014) and less evidence of northern haplotypes in the southern population. Nearly 25% of crabs at both KJI and PLB were of the southern haplotype, suggesting anthropogenic introduction from a mixed or hybridized Scotian Shelf population source to PLB (Blakeslee et al., 2010), and secondary contact between northern and southern populations at KJI where the southwards expansion of the more recent invasion met the northern limit of the initial invasion (Darling et al., 2014). Overall, the COI data were equally able to detect

isolation by distance and population structuring via pairwise F_{ST} values among sampling sites compared to the SNPs. The SNP panel clearly revealed this divergence was widespread across the genome and the outlier SNPs were better able to detect genetic differences between the northern and southern groups explaining significantly greater components of the spatial variance.

4.3 | Management and conservation implications

The data presented here are consistent with previous descriptions of green crab invasion history (Darling et al., 2014; Tepolt & Palumbi, 2015), suggesting two groups with southern Nova Scotia acting as the (potentially transient) geographic divide between northern and southern populations. This location is consistent with an initial observation by Darling et al. (2014) who documented an admixture zone between northern and southern populations between New Brunswick and Nova Scotia. Knowledge of invasion histories is necessary to predict secondary spread, which ultimately will determine the impact of an invasive species, particularly those where introduction and spread are mediated by anthropogenic vectors (Cristescu, 2015). A better understanding of the population structuring and environmental factors that sustain this structuring within this invasive species will allow us to monitor and potentially predict range expansion, which in turn helps to focus the development of preemptive and targeted monitoring and mitigation strategies in Atlantic Canada. Green crab have the potential to impact commercial shellfish fisheries (e.g., McClenachan et al., 2015) and lead to regional reduction of eelgrass beds (Garbary et al., 2014; Matheson et al., 2016; Neckles, 2015), and so management strategies need to be implemented to minimize their environmental impact. In fact, it may be necessary to consider these crabs as two distinct ecotypes across their invasive range due to their potential adaptations to different thermal regimes and behavioral differences (Rosson et al., 2012) to implement more effective management strategies. Green crab show different thermal tolerances between lineages (Tepolt & Somero, 2014) and adaptive differences in the native range

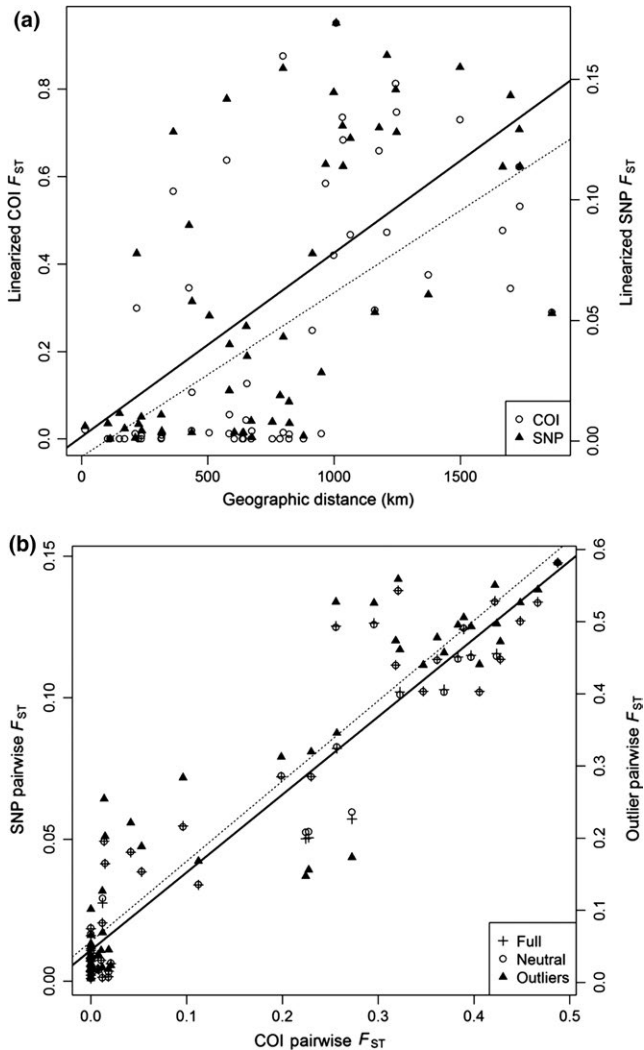


FIGURE 8 (a) Isolation by distance relationships of linearized pairwise F_{ST} values versus least-cost geographic distances between each site. Correlations were significant for both COI using a Mantel test with 1,000 permutations ($r = .60$, $p = .008$) and the SNP panel ($r = .624$, $p = .004$). The solid line represents the SNP correlation, while the dotted line represents the COI correlation. (b) Comparison of spatial structuring at all three SNP panels and mtDNA data. Mantel tests with 1,000 permutations showed strong ($r = .93$ – $.95$) and significant ($p < .001$) correlations for each dataset. The solid line indicates the full and neutral SNP correlation, while the dotted line indicates the outlier relationship

likely facilitated the invasion success and subsequent range expansion in North America (Roman, 2006). However, to date green crab in Atlantic Canada have been treated as a single species and population for risk assessment and habitat suitability modeling (Therriault, Herborg, Locke, & McKindsey, 2008). As evidence suggests there are indeed two ecotypes of green crab in Atlantic Canada, this will necessarily influence future policy and mitigation strategies that account for high and low risk areas for each ecotype that have likely adapted to different temperature regimes. Predictions of habitat suitability and the potential for range expansion would in this case be best evaluated for each ecotype independently.

TABLE 3 Analysis of molecular variance (AMOVA) for the (a) neutral SNP panel, (b) outlier SNP panel, and (c) COI data after categorizing each sampling location into a northern or southern population based on discriminant analysis of principal components

Source of variation	Variance components	Percentage of variation
(a)		
Among groups	70.52	5.76
Among populations within groups	23.30	1.90
Within populations	1129.57	92.33
$F_{ST} = 0.0767$; $F_{SC} = 0.0202$; $F_{CT} = 0.0576$		
(b)		
Among groups	11.51	38.86
Among populations within groups	1.16	3.90
Within populations	16.96	57.24
$F_{ST} = 0.428$; $F_{SC} = 0.043$; $F_{CT} = 0.389$		
(c)		
Among groups	0.28	19.65
Among populations within groups	0.057	3.98
Within populations	1.10	76.37
$F_{ST} = 0.236$; $F_{SC} = 0.049$; $F_{CT} = 0.197$		

4.4 | Summary

The recent green crab range expansion in eastern North America appears to be driven by a second invasion from the northern region of their native range in Europe (Roman, 2006). Our results show that genetic differences between the northern and southern populations of green crab in eastern North America are extensive and genomewide and present strong support for two genetically distinct groups of crabs that were the result of two separate invasions. These crabs may continue to expand northwards and southwards, and potentially hybridize at contact zones or regions high in shipping traffic. The genetic differences observed here are so distinctive that they warrant the division of green crab into two separate ecotypes in the northwest Atlantic, yet the precise latitudinal boundary between ecotypes has yet to be determined and may be transient. Further characterization of the outlier loci based on a whole-genome sequence and investigation of introgression between the two populations of green crab studied here will facilitate continued monitoring and future exploration.

ACKNOWLEDGMENTS

The authors wish to thank staff at Fisheries and Oceans Canada for their assistance with sample collection, in particular Erica Watson, Sophie Boudreau, Lorne Penny, Marie-Helene Therriault, Dawn Sephton, Roland Hagan, Paul Jivoff, and April Blakeslee. They also wish to thank the staff of the Aquatic Biotechnology Lab at the Bedford Institute of Oceanography for assistance with DNA extractions and RAD library preparation.

Funding was provided by the Genomic Resource Development Initiative, Fisheries and Oceans Canada to IRB and CDB. This work was supported by a National Sciences and Engineering Research Council of Canada (NSERC) visiting fellowship awarded to NWJ.

DATA ACCESSIBILITY

DNA sequences: All raw COI mtDNA sequence data will be uploaded to NCBI and accession numbers provided. The RAD-seq genotypes used in this study as well as all distance matrices and geographic least-cost distances will be deposited in Dryad prior to final acceptance. All corresponding sample codes and locations are in Table 1 and Figure 2. Raw RAD-seq reads are available from the NCBI Sequence Read Archive BioProject PRJNA377723 at <https://www.ncbi.nlm.nih.gov/Traces/study/?acc=SRP102198>. COI sequences are available from GenBank (Accession numbers KY800915 - KY801176). Genotype data, including filtered RAD-seq data and aligned COI sequences are available from the Dryad Digital Repository at <http://dx.doi.org/10.5061/dryad.619t0>.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

IRB, R. Beiko, and CDB conceived of the study. JF, KM, CHM, R. Bernier, and CDB collected specimens for use in this study. LCH conducted DNA extraction and RADseq library preparation. NWJ, MVR, RES, PNR, R. Beiko conducted bioinformatics and statistical analyses of the data. All authors aided in the interpretation of the data and initial study design. All authors contributed to writing, revising, and approving the final draft of the manuscript.

REFERENCES

- Audet, D., Davis, D. S., Miron, G., Moriyasu, M., Benhalima, K., & Campbell, R. (2003). Geographical expansion of a nonindigenous crab, *Carcinus maenas* (L.), along the Nova Scotian shore into the southeastern Gulf of St. Lawrence, Canada. *Journal of Shellfish Research*, 22, 255–262.
- Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, 24, 3299–3315.
- Blakeslee, A. M. H., McKenzie, C. H., Darling, J. A., Byers, J. E., Pringle, J. M., & Roman, J. (2010). A hitchhiker's guide to the Maritimes: Anthropogenic transport facilitates long-distance dispersal of an invasive marine crab to Newfoundland. *Biodiversity Research*, 16, 879–891.
- Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., & SanCristobal, M. (2010). Detecting selection in population trees: The Lewontin and Krakauer test extended. *Genetics*, 186, 241–262.
- Candy, J. R., Campbell, N. R., Grinnell, M. H., Beacham, T. D., Larson, W. A., & Narum, S. R. (2015). Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. *Molecular Ecology Resources*, 15, 1421–1434.
- Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140.
- Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics*, 19, 233–257.
- Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications*, 8, 23–46.
- Cristescu, M. E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology*, 24, 2212–2225.
- Darling, J. A. (2011). Interspecific hybridization and mitochondrial introgression in invasive *Carcinus* shore crabs. *PLoS ONE*, 6, e17828.
- Darling, J. A., Bagley, M. J., Roman, J., Tepolt, C. K., & Geller, J. B. (2008). Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Molecular Ecology*, 17, 4992–5007.
- Darling, J. A., Tsai, Y. H. E., Blakeslee, A. M. H., & Roman, J. (2014). Are genes faster than crabs? Mitochondrial introgression exceeds larval dispersal during population expansion of the invasive crab *Carcinus maenas*. *Royal Society Open Science*, 1, 140202.
- Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22, 1–20.
- Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A., & Cresko, W. A. (2011a). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In V. Orgogozo, & M. V. Rockman (Eds.), *Molecular methods for evolutionary genetics* (pp. 157–178). Totowa: Humana Press.
- Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011b). Local De Novo assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE*, 6, e18561.
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567.
- Faircloth, B. C., & Glenn, T. C. (2012). Not all sequence tags are created equal: Designing and validating sequence identification tags robust to indels. *PLoS ONE*, 7, e42543.
- Floyd, T., & Williams, J. (2004). Impact of green crab (*Carcinus maenas* L.) predation on a population of soft-shell clams (*Mya arenaria* L.) in the southern Gulf of St. Lawrence. *Journal of Shellfish Research*, 23, 457–463.
- Folino-Rorem, N. C., Darling, J. A., & D'Ausilio, C. A. (2009). Genetic analysis reveals multiple cryptic invasive species of the hydrozoan genus *Cordylophora*. *Biological Invasions*, 11, 1869–1882.
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.
- Freemo, H., O'Reilly, P., Berg, P. R., Lien, S., & Boulding, E. G. (2011). Outlier SNPs show more genetic structure between two Bay of Fundy meta-populations of Atlantic salmon than do neutral SNPs. *Molecular Ecology Resources*, 11, 254–267.
- Garbary, D. J., Miller, A. G., Williams, J., & Seymour, N. R. (2014). Drastic decline of an extensive eelgrass bed in Nova Scotia due to the activity of the invasive green crab (*Carcinus maenas*). *Marine Biology*, 161, 3–15.
- Geller, J. B., Walton, E. D., Grosholz, E. D., & Ruiz, G. M. (1997). Cryptic invasions of the crab *Carcinus* detected by molecular phylogeography. *Molecular Ecology*, 6, 901–906.
- Goudet, J. (2014). hierFSTat: Estimation and tests of hierarchical F-statistics. R package version 0.04-14. <http://CRAN.R-project.org/package=hierFSTat>.
- Grosholz, E. (2002). Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology and Evolution*, 17, 22–27.
- Hänel, C., & Chown, S. L. (1998). The impacts of a small, alien invertebrate on a sub-Antarctic terrestrial ecosystem: *Limnophyes minimus* (Diptera, Chironomidae) at Marion Island. *Polar Biology*, 20, 99–106.
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403–1405. doi:10.1093/bioinformatics/btn129
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071. doi:10.1093/bioinformatics/btr521

- Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). *diversity*: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4, 782–788.
- Krehenwinkel, H., & Tautz, D. (2013). Northern range expansion of European populations of the wasp spider *Argiope bruennichi* is associated with global warming-correlated genetic admixture and population-specific temperature adaptations. *Molecular Ecology*, 22, 2232–2248.
- Langella, O. (2012). Populations 1.2.33. Available at [https://launchpad.net/~olivier-langella/+archive/ubuntu/ppa].
- Lee, C. E. (1999). Rapid and repeated invasions of fresh water by the copepod *Eurytemora affinis*. *Evolution*, 53, 1423–1434.
- Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, 17, 386–391.
- Levings, C., Kieser, D., Jamieson, G., & Dudas, S. (2002). Marine and estuarine alien species in the Strait of Georgia, BC. In *Alien invaders in Canada's waters, wetlands, and forests*. Natural Resources Canada, pp. 111–131.
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28, 298–299.
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, 23, 2178–2192.
- Matheson, K., McKenzie, C., Gregory, R., Robichaud, D. A., Bradbury, I. R., Snelgrove, P. V. R., & Rose, G. A. (2016). Linking eelgrass decline and impacts on associated fish communities to European green crab *Carcinus maenas* invasion. *Marine Ecology Progress Series*, 548, 31–45.
- McClenachan, L., O'Connor, G., & Reynolds, T. (2015). Adaptive capacity of co-management systems in the face of environmental change: The soft-shell clam fishery and invasive green crab in Maine. *Marine Policy*, 52, 26–32.
- McDonald, D. B., Parchman, T. L., Bower, M. R., Hubert, W. A., & Rahel, F. J. (2008). An introduced and a native vertebrate hybridize to form a genetic bridge to a second native species. *Proceedings of the National Academy of Sciences USA*, 105, 10837–10842.
- Molnar, J. L., Gamboa, R. L., Revenga, C., & Spalding, M. D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, 6, 485–492.
- Mooney, H. A., & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences USA*, 98, 5446–5451.
- Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., ... Bernatchez, L. (2014). Conservation genomics of anadromous Atlantic salmon across its North American range: Outlier loci identify the same patterns of population structure as neutral loci. *Molecular Ecology*, 23, 5680–5697.
- Neckles, H. A. (2015). Loss of eelgrass in Casco Bay, Maine, linked to green crab disturbance. *Northeast Naturalist*, 22, 478–500.
- Pante, E., & Benoit, S.-B. (2013). Marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE*, 8, e73051.
- Pascoal, S., Creer, S., Taylor, M. I., Queiroga, H., Carvalho, G., & Mendes, S. (2009). Development and application of microsatellites in *Carcinus maenas*: Genetic differentiation between northern and central Portuguese populations. *PLoS ONE*, 4, e7268.
- Pringle, J. M., Blakeslee, A. M. H., Byers, J. E., & Roman, J. (2011). Asymmetric dispersal allows an upstream region to control population structure throughout a species' range. *Proceedings of the National Academy of Sciences USA*, 108, 15288–15293.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81, 559–575.
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/
- Ricciardi, A. (2006). Patterns of invasion in the Laurentian Great Lakes in relation to changes in vector activity. *Diversity and Distributions*, 12, 425–433.
- Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution*, 29, 233–242.
- Robinson, J. F., & Dickerson, J. E. (1987). Does invasion sequence affect community structure? *Ecology*, 68, 587–595.
- Roman, J. (2006). Diluting the founder effect: Cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society B*, 273, 2453–2459.
- Roman, J., & Palumbi, S. R. (2004). A global invader at home: Population structure of the green crab, *Carcinus maenas*, in Europe. *Molecular Ecology*, 13, 2891–2898.
- Rosson, M. A., Barrett, T. J., Quijon, P. A., Snelgrove, P. V. R., McKenzie, C. H., & Locke, A. (2012). Regional differences in foraging behavior and morphology of invasive green crab (*Carcinus maenas*) populations in Atlantic Canada. *Biological Invasions*, 14, 659–669.
- Sherman, C. D. H., Lotterhos, K. E., Richardson, M. F., Tepolt, C. K., Rollins, L. A., Palumbi, S. R., & Miller, A. D. (2016). What are we missing about marine invasions? Filling in the gaps with evolutionary genomics. *Marine Biology*, 163, 198.
- Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457–462.
- Stanley, R. E., Jeffery, N. W., Wringe, B. F., DiBacco, C., & Bradbury, I. R. (2017). genepedit: A simple and flexible tool for manipulating multi-locus molecular data in R. *Molecular Ecology Resources*, 17, 12–18. doi:10.1111/1755-0998.12569
- Tepolt, C. K., Bagley, M. J., Geller, J. B., & Blum, M. J. (2006). Characterization of microsatellite loci in the European green crab (*Carcinus maenas*). *Molecular Ecology Notes*, 6, 343–345.
- Tepolt, C. K., & Palumbi, S. R. (2015). Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Molecular Ecology*, 24, 4145–4158.
- Tepolt, C. K., & Somero, G. N. (2014). Master of all trades: Thermal acclimation and adaptation of cardiac function in a broadly distributed marine invasive species, the European green crab, *Carcinus maenas*. *The Journal of Experimental Biology*, 217, 1129–1138.
- Therriault, T. W., Herborg, L. M., Locke, A., & McKindsey, C. W. (2008). *Risk assessment for European green crab (Carcinus maenas) in Canadian waters*. Canada: Canadian Science Advisory Secretariat, Fisheries and Oceans.
- Van Wyngaarden, M., Snelgrove, P. V., DiBacco, C., Hamilton, L. C., Rodriguez-Ezpeleta, N., Jeffery, N. W., ... Bradbury, I. R. (2017). Identifying patterns of dispersal, connectivity, and selection in the sea scallop, *Placopecten magellanicus*, using RADseq-derived SNPs. *Evolutionary Applications*, 10, 102–117. doi:10.1111/eva.12432
- Viard, F., David, P., & Darling, J. A. (2016). Marine invasions enter the genomic era: Three lessons from the past, and the way forward. *Current Zoology*, 62, 629–642. doi:10.1093/cz/zow053
- Wang, X. W., Luan, J. B., Li, J. M., Su, Y. L., Xia, J., & Liu, S. S. (2011). Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics*, 12, 458.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Jeffery NW, DiBacco C, Van Wyngaarden M, et al. RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. *Ecol Evol*. 2017;7:2513–2524. https://doi.org/10.1002/ece3.2872