ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Founder effects shape linkage disequilibrium and genomic diversity of a partially clonal invader

Ben A. Flanagan[1,2] | Stacy A. Krueger-Hadfield[1,3] | Courtney J. Murren[1] |
Chris C. Nice[4] | Allan E. Strand[1] | Erik E. Sotka[1]

[1]Department of Biology, College of Charleston, Charleston, SC, USA

[2]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA

[3]Department of Biology, University of Alabama at Birmingham, Birmingham, AL, USA

[4]Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX, USA

Correspondence
Ben A. Flanagan, Allan Hancock Foundation, 3616 Trousdale Pkwy, Los Angeles, CA 90089, USA.
Email: flanaganbena@gmail.com

Erik E. Sotka, Grice Marine Laboratory, 205 Fort Johnson Road, Charleston, SC 29412, USA.
Email: SotkaE@cofc.edu

Funding information
National Science Foundation, Grant/Award Number: BIO-OCE 1357386 and DEB 1050355

Abstract

The genomic variation of an invasive species may be affected by complex demographic histories and evolutionary changes during the invasion. Here, we describe the relative influence of bottlenecks, clonality, and population expansion in determining genomic variability of the widespread red macroalga *Agarophyton vermiculophyllum*. Its introduction from mainland Japan to the estuaries of North America and Europe coincided with shifts from predominantly sexual to partially clonal reproduction and rapid adaptive evolution. A survey of 62,285 SNPs for 351 individuals from 35 populations, aligned to 24 chromosome-length scaffolds indicate that linkage disequilibrium (LD), observed heterozygosity (H_o), Tajima's D, and nucleotide diversity (Pi) were greater among non-native than native populations. Evolutionary simulations indicate LD and Tajima's D were consistent with a severe population bottleneck. Also, the increased rate of clonal reproduction in the non-native range could not have produced the observed patterns by itself but may have magnified the bottleneck effect on LD. Elevated marker diversity in the genetic source populations could have contributed to the increased H_o and Pi observed in the non-native range. We refined the previous invasion source region to a ~50 km section of northeastern Honshu Island. Outlier detection methods failed to reveal any consistently differentiated loci shared among invaded regions, probably because of the complex *A. vermiculophyllum* demographic history. Our results reinforce the importance of demographic history, specifically founder effects, in driving genomic variation of invasive populations, even when localized adaptive evolution and reproductive system shifts are observed.

KEYWORDS
Agarophyton vermiculophyllum, bottleneck, forward evolutionary simulations, haplodiplontic, invasion, linkage disequilibrium

1 | INTRODUCTION

Biological invasions can wreak havoc on native communities by homogenizing environments thereby compromising ecosystem function (Pimentel et al., 2005). As globalization continues to escalate, the frequency of potential and successful biological invasions continues to increase (Seebens et al., 2017). Efforts to understand

why certain species become invasive where others fail has been focused on identifying key phenotypic traits shared among invaders, such as high growth rate, high reproductive output, broad physiological tolerance (Lee, 1999) and flexible reproductive modes (Grotkopp & Rejmanek, 2007; Kolar & Lodge, 2001; Pannell, 2015; Roman & Darling, 2007). While there is widespread evidence that existing phenotypic trait variation may facilitate successful invaders

(Bossdorf et al., 2005), and that traits can rapidly evolve to facilitate invasion establishment and subsequent spread (Lee, 2002; Sotka et al., 2018), we have fewer examples of shifts in genomic-scale variability and diversity that accompany marine invasions. Moreover, for marine and terrestrial invasions, it is unclear what the relative importance of neutral and non-neutral evolutionary processes is in driving resultant genomic changes.

Biological invasion dynamics are commonly characterized as a series of demographic phases or stages, which include population establishment, growth, and expansion (Blackburn et al., 2011; Bourne et al., 2018; Williamson, 1993; Williamson & Fitter, 1996). During each stage, multiple demographic processes cumulatively determine neutral genomic diversity and co-inheritance patterns of non-native populations. For example, during establishment, invading populations can experience founder effects in which only a few individuals become established. The intensity and length of a given founder event and resulting bottleneck will determine the subsequent genomic changes (Pérez-Figueroa et al., 2010; Slatkin, 2008). On the other hand, when a non-native population receives continued propagule pressure, there are no bottlenecks, and instead, genomic diversity is maintained or can even increase when multiple populations are admixed (Dlugosch et al., 2015; Dlugosch & Parker, 2008). Following population establishment and subsequent growth, the expansion phase will increase population size and geographic range resulting in novel genomic variation which in large part depends on new mutations. Consequently, adaptive changes occurring within a few generations of non-native population establishment are probably less reliant on new mutational input and adaptive evolution should mostly depend on standing genetic variation in the newly established non-native population (Maruyama & Fuerst, 1984; Prentis et al., 2008). Successful invasions are the result of complex demographic and evolutionary histories and no single invasion process appears to determine genomic variability of non-native populations (Boheemen et al., 2017) and delineating among these alternative processes based upon empirical data can be difficult. However, recent advances in computing power and genomic-scale in silico simulations have made the process of testing the effects of alternative demographic hypotheses on empirical data more robust and accessible to a wider array of species, with unique life history strategies, reproductive systems, and eco-evolutionary dynamics (Haller & Messer, 2019; Hoban et al., 2012).

The haplodiplontic red seaweed *Agarophyton vermiculophyllum* (Ohmi) Papenfuss (formerly *Gracilaria vermiculophylla*) is native to the northwest Pacific and has invaded temperate estuaries along the coastlines of North America, Europe, and northwestern Africa (Bellorin et al., 2004; Guillemin et al., 2008; Kim et al., 2010; Krueger-Hadfield et al., 2017; Weinberger et al., 2008). Krueger-Hadfield et al. (2017) used a combination of mitochondrial haplotypes and microsatellite genotypes to determine that the vast majority of non-native populations originated from three sites along the northeastern coastline of Japan: Mangoku and Matsushima Bays (Miyagi Prefecture) and Akkeshi Bay (Hokkaido Prefecture). Given that this area of Japan has long been an epicentre for oyster aquaculture and export, they hypothesized that *A. vermiculophyllum* thalli were inadvertently transported with oysters to each coastline in the Northern Hemisphere. The strongest evidence for a severe bottleneck comes from mitochondrial cytochrome c oxidase subunit I (*cox*1) sequencing, in which the vast majority of non-native populations have a single haplotype, although this haplotype is common in the source region of the invasion (haplotype 6; Kim et al., 2010; Krueger-Hadfield et al., 2017). In contrast, there were no regional differences in expected heterozygosity nor allelic richness at 10 microsatellite loci (Krueger-Hadfield et al., 2016, 2017), suggesting that the bottleneck was not severe or was countered by other demographic factors. After establishment, populations rapidly increased in population size and expanded in range throughout the North American west and east coasts and along European and northern African coastlines (Krueger-Hadfield et al., 2017, 2018; Weinberger et al., 2008).

Two other aspects of the *A. vermiculophyllum* invasion have the potential to affect patterns of genomic variation. First, there was a profound ecological shift from hard- to soft-substratum that resulted in diploid dominance as a consequence of high rates of clonal fragmentation in the non-native range (Krueger-Hadfield et al., 2016). This shift in reproductive mode from sexual, native populations to largely asexual, non-native populations fits the expectations of Baker's Law (Krueger-Hadfield et al., 2016) which posits colonization success may be enhanced if an individual is able to undergo uniparental reproduction (Baker, 1955; Baker & Stebbins, 1965). In the *A. vermiculophyllum* invasion, this was confirmed by a four-fold increase in the proportion of repeated microsatellite genotypes observed in non-native populations relative to native populations (average: 0.36 vs. 0.09, respectively; Krueger-Hadfield et al., 2016). In those species capable of self-fertilization or asexual reproduction, shifts in reproductive system (e.g., increases in selfing or clonality) can facilitate an invasion (Baker & Stebbins, 1965; Krueger-Hadfield et al., 2016; Sakai et al., 2001) and has predictable genomic consequences, including reducing effective population size, increasing observed heterozygosity, and increasing linkage disequilibrium (LD) (Balloux et al., 2003; Bazin et al., 2014; Halkett et al., 2005).

Second, Sotka et al. (2018) used laboratory-based common-garden assays and niche modeling to detect an expansion of thermal niches and rapid phenotypic evolution. Non-native populations now occupy warmer environments than did source populations from Japan and also exhibit increased thermal tolerance (Sotka et al., 2018). Such strong selection may yield a selective sweep in which one or more alleles in a subset of the genome increase in frequency and serve to increase population-level differentiation at those sites (Messer & Petrov, 2013). The genomic signals of rapid evolution have not been explored to date in *A. vermiculophyllum*.

Overall, the *A. vermiculophyllum* invasion is the result of a complex evolutionary history where rapid selection, a shift in reproductive mode, strong bottlenecks, and population expansion all potentially molded genomic diversity throughout the non-native range (see Casso et al., 2019; Le Cam et al., 2020 for other marine invasion examples). To assess the relative impact of these different factors,

we sequenced and assembled the first *A. vermiculophyllum* genome and performed individual-based reduced representation genome sequencing for 35 populations across the native and non-native ranges. We, then, characterized shifts in genomic variability and extent of LD between native and non-native regions. We performed forward evolutionary simulations to test how hypothesized reproductive system shifts and demographic changes, specific to the *A. vermiculophyllum* invasion, compare to these empirical data. Because testing hypothesized evolutionary and demographic changes during the invasion requires the identification of the population sources in the native range (Cristescu, 2015; Estoup & Guillemaud, 2010), we used population genomic approaches to further refine the invasion source. Finally, we used an F_{ST} outlier analysis to detect regions of the genome that were particularly differentiated when compared to native genetic source populations due to selection early in the invasion. Using these data and analyses, we explored several predictions: (a) If non-native *A. vermiculophyllum* populations went through a severe bottleneck, we expected the extent of LD to increase rapidly and alleles to be lost to drift. (b) Clonal reproduction may also increase LD, but we expected it to be a gradual process leading to the accruement of long-range LD over time. (c) If population expansion impacted the non-native population genomic diversity, we expected LD to decrease (Slatkin, 1994).

2 | MATERIALS AND METHODS

2.1 | Sample collection

We collected *A. vermiculophyllum* thalli at 35 intertidal sites across the Northern Hemisphere as part of a larger effort to genotype and phenotype native and non-native populations (Krueger-Hadfield et al., 2017; Sotka et al., 2018). This included 11 Japanese sites, five sites along the western coast of North America, 10 sites on the eastern coast of the United States, and nine sites along the European coast (Table S1, Figure S1). At each site, we haphazardly collected 100 thalli separated by at least one meter and determined the reproductive state of each thallus (i.e., male gametophyte, female gametophyte, diploid tetrasporophyte, or nonreproductive) using a dissecting microscope (40×) following Krueger-Hadfield et al. (2016). We then preserved 5–10 cm fragments of each thallus in silica gel as vouchers and for DNA extraction.

A single haploid phenotypically-male thallus was collected at the Fort Johnson mudflat in Charleston, SC, USA (fjn) and grown in culture for over 4 months at the University of Alabama, Birmingham, AL, USA. The thallus was large and split into several 50 ml conical polyethylene vials filled with 45 ml of filtered, natural seawater collected from Charleston Harbor, SC, USA (salinity: 30 ppt) at a temperature of 15°C, and an irradiance of approximately 60 μmol photons ·m^{-2} s^{-1} with a 12:12 light:dark cycle in a Percival growth chamber (Percival). The water was changed two times per week with no nutrient enrichment. The thallus was flash frozen in liquid N$_2$ and shipped overnight to Dovetail Genomics for genome sequencing.

2.2 | Genome sequencing and assembly

High-molecular weight gDNA with mean fragment length = 50 kbp was isolated from the male thallus and 500 ng of gDNA was used to generate an Illumina shotgun sequencing library because sequencing a haploid individual aids in assembly. The sequenced library resulted in 405 M paired-reads (2 × 150 bp) which were then trimmed using Trimmomatic (Bolger et al., 2014). In addition, 500 ng of gDNA was used to generate a proprietary Chicago library. Briefly, to construct the Chicago library, gDNA was reconstituted into chromatin in vitro and digested with *DpnII*, the 5′ overhangs were filled in with biotinylated nucleotides. The free blunt ends were then ligated. After ligation, the crosslinks were reversed and the DNA purified from protein (Putnam et al., 2016). This purified DNA was sheared to 350 bp and used to construct sequencing libraries. Illumina sequencing resulted in 225 M paired-end reads (2 × 150 bp), resulting in 358× coverage of the genome.

A HiC library was also constructed by Dovetail Genomics from the same thallus (Lieberman-Aiden et al., 2009). For this library, chromatin was fixed with formaldehyde in nuclei and then extracted from the tissue. The chromatin was digested with *DpnII*, filled with biotinylated nucleotides, ligated and then sequenced as above. Illumina sequencing resulted in 215 M paired-end reads (2 × 150 bp) resulting in 9000× coverage of the genome. De novo assembly started with the paired shotgun sequencing reads and was performed using *Meraculous* (Chapman et al., 2011) with a kmer size of 109. The de novo assembly and the shotgun, Chicago, and HiRise library sequences were all used as inputs into Dovetail's proprietary HIRISE software pipeline for scaffolding.

Previous studies indicate that gracilarioid species (including *A. vermiculophyllum*) have identical 2C DNA contents of 0.4 pg and chromosome complements of 2*n* = 48 (Kapraun & Freshwater, 2012). These were nearly matched by the Dovetail assembly. The total assembly had 731 scaffolds across 47.8 Mbp. After scaffolds with high matches to mitochondrial and chloroplast genes in GenBank were removed from nuclear assembly, the remaining nuclear scaffolds were 47.6 Mbp. We used R::ape (Paradis et al., 2004) to rename the 24 largest scaffolds by length (scaf01 = 3.56 Mbp; scaf24 = 0.08 Mbp; Table S2). These 24 scaffolds represent 46.2 Mbp, or 96.5% of all data and 96.9% of nuclear DNA (L50/N50 = 10 scaffolds; 2.06 Mbp). We note these are not true "chromosomes" until further sequencing that connects the remaining ~4%; they may be considered as large scaffolds that represent single linkage groups.

2.3 | Genomic library preparation and sequencing

A. vermiculophyllum has an isomorphic haplodiplontic life cycle with free living haploid male gametophytes and female gametophytes and diploid tetrasporophytes. Native populations are composed of ~60% diploid thalli, while non-native populations in soft-substratum habitats are composed of ~90% or greater diploid thalli (Krueger-Hadfield et al., 2016). Because non-native

populations are primarily diploid, we focused our sequencing efforts on only the diploid stage. Further, by only sequencing diploids, we were able to more fully estimate genomic diversity as well as avoid mixed ploidy samples (Krueger-Hadfield & Ryan, 2020). Total gDNA was extracted from approximately 5–10 mg of dried tissue of 560 thalli (n = 16 per site) that were either phenotypically-diploid (i.e., bearing reproductive structures) or when not enough thalli were reproductive from a site, we used nonreproductive thalli. We used the Nucleospin Plant Kit (Macherey-Nagel) and followed the manufacturer's protocol with two exceptions: we performed the lysis step for 1 h at room temperature and eluted in 100 µl of molecular grade water.

We prepared one library of all 560 individuals using protocols of Parchman et al. (2012). Briefly, we digested gDNA with two restriction enzymes, *Eco*RI and *Mse*I, and ligated adaptors containing unique 8–10 bp barcodes to the digested DNA of each individual. The products were, then, PCR amplified in two independent reactions with standard Illumina primers. All amplicons were pooled and shipped to the University of Texas Genomic Sequencing and Analysis Facility, which used Blue Pippen Prep to isolate the 300–500 bp fraction. This fraction was then single-read sequenced with one lane each on Illumina HiSeq 2000 and HiSeq 4000 platforms.

2.4 | Data processing

The two Illumina sequencing lanes yielded 4.4×10^8 short read fragments after removal of PhiX sequences with bowtie2 (Langmead & Salzberg, 2012). Fragments that did not have adapter nor barcode sequences were removed using a custom perl script. All reads were then aligned to the 24 scaffolds using BWA ver. 0.7.15 (Li & Durbin, 2009). We used SAMTOOLS/BCFTOOLS 1.3.1 (Li et al., 2009) to identify variant, multiallelic SNPs and generate genotype likelihoods (i.e., "samtools mpileup -v" and "bcftools call -m" options). We then used "bcftools filter" to remove loci with a minor allele frequency less than or equal to 2%, and confirmed this method was successful using "bcftools stats" and the python script "plot-vcfstats." We used "bedtools multicov" option (Quinlan & Hall, 2010) to assess our coverage per locus and custom R script (R Core Team, 2018) to filter out poorly sequenced individuals (i.e., reads at <20% of all loci) and loci (i.e., 1+ reads at 50% of individuals).

All thalli were simultaneously genotyped at 10 microsatellite loci (see methods of Kollars et al., 2015; Krueger-Hadfield et al., 2016). Individuals that had a single allele at all 10 microsatellite loci were considered haploid (see Kollars et al., 2015; Krueger-Hadfield et al., 2016) and removed (n = 27). Previously, Krueger-Hadfield et al. (2016) reported highly accurate ploidy assignment based on 10 microsatellite loci where no haploids or diploids were misassigned (n = 230). Seventeen individuals were removed because they did not have enough reads to meet minimum sequencing requirements, yielding 516 diploid thalli (Table S1). Particularly in non-native populations, fragmentation occurs resulting in clones sharing the same multilocus genotype (MLG). We used the microsatellites to identify repeated MLGs using *R::RClone* (Bailleul et al., 2016). If the probability that a given number of genetically identical ramets are separate genets (P_{sex} p-value; Parks & Werth, 1993) was greater than 0.05, duplicated MLGs were considered different genets and retained in the data set. If the P_{sex} p-value was smaller than 0.05, the duplicated MLGs were considered as ramets (or clones) of the same genet (Arnaud-Haond et al., 2007). Our analyses focused on populations that had 10 or more thalli after removing duplicate MLGs; this reduced number of populations analysed overall, especially in the non-native range where clonal reproduction is more common. We were left with a set of 62,285 loci and 351 individuals, sequenced to a depth of 19.4 reads per locus, on average (Flanagan et al., 2021). This data set was used for all analyses, except when noted.

2.5 | Population genomic analyses

With few exceptions, we used genotype likelihoods in order to incorporate the uncertainty that arises from errors in sequencing and sampling (Nielsen et al., 2012). Per-locus allele frequencies were generated from genotype likelihoods and compared between native versus non-native populations in a one-way ANOVA in R. The software ANGSD (Korneliussen et al., ,2013, 2014) compiles genotype likelihoods generated by samtools into a beagle-formatted file that served as the input for later analyses.

We estimated LD for each pair of SNPs within a scaffold for 22 populations with greater than 10 individuals using ngsLD (Fox et al., 2019). Between each pair of SNPs, ngsLD estimated LD as r^2. Genetic distances were, then, binned into 0.5 kbp windows and the average r^2 within that window was calculated. This was repeated separately for each of 22 scaffolds and 22 populations. Because the smaller scaffolds (sca23, sca24) had only a few pairwise comparisons within some populations, we used only the first 22 scaffolds for these analyses.

To quantify the decay of LD with genetic distance within each scaffold, we used a nonlinear least squares fit to the decay rate estimator using custom R code (Hill & Weir, 1988). We used two-way ANOVAs to assess the effect of native versus non-native regions, scaffolds, and their interaction on the average (by population) half-decay distance and genetic distance across which $r^2 > 0.1$.

Observed heterozygosity (--freq2), Tajima's D (--TajimaD), and nucleotide diversity (--site-pi) were estimated within each population using VCFTOOLS v. 0.1.15 (Danecek et al., 2011). Our analyses focused on populations that had 10 or more thalli remaining after filtering out ramets.

We assessed population genomic structure using principal components analysis (PCA) (Skotte et al., 2013). PCAngsd uses the beagle-formatted genotype likelihoods to estimate covariance among individuals. We retained the first four PC axes. To assign population sources, we employed a linear discriminant analysis, or LDA, implemented in R::MASS (Venables & Ripley, 2002). We trained the LDA on first four PCs with the Japanese populations and predicted assignment of each non-native thallus. The probabilities

were summed for each population, and visualized using *R::circlize* (Gu et al., 2014).

We implemented maximum-likelihood admixture methods to infer individual ancestry assuming an a priori number of genetic clusters (NGSadmix; Meisner & Albrechtsen, 2018). We generated 10 independent runs of NGSadmix using a broad range of admixing populations (*k* = 2–20; following suggestions of Evanno et al. (2005), we plotted mean and variance in log-likelihood values within a given *k* and delta (lnL). While likelihoods increase with *k* and there is no clear asymptote (Figure S8), there are peaks in delta (lnL) at *k* = 3, 7, 9 and 16. To compare how population structure changes with different *k* clusters, we plotted individual assignment proportions across clusters *k* = 2–16. We also visualized the average contribution of *k* = 7 clusters across the Northern Hemisphere as *k* = 7 was the highest *k* in which nearly all clusters found in non-native populations were also found in the native range.

We used ngsFST (Fumagalli et al., 2013) to calculate a method-of-moments estimator of locus-specific F_{ST} values (Reynolds et al., 1983) for each pairwise set of populations with 10 or more thalli in our final data set (*n* = 22 populations). For each pairwise analysis, we used the suite of loci with calls in greater than 80% of individuals and with a minor allele frequency (MAF) greater than 2%. We also generated pairwise F_{ST} values for non-native populations versus the two native source populations (i.e., thalli from mng and sou combined) that served as ultimate sources for the majority of non-native thalli (see Results). To assess whether genomic responses of independent invasion events differ, we present analysed per-locus F_{ST} data for the three populations with the lowest genome-wide F_{ST} value relative to the native source (i.e., tmb of western North America, osh of eastern North America, and fdm of Europe). We assume that non-native populations with the lowest genome-wide F_{ST} relative to the source populations were the initial point of introduction to that shoreline. Our goal was to use these most closely-related population comparisons to assess the role that selection during or after initial introduction in altering the genome, and to minimize the influence of neutral processes (mainly drift) that may tend to inflate F_{ST} during expansion in the non-native range. The F_{ST} values of SNPs and their 99% distributions were plotted along the 24 scaffolds for visual inspection.

2.6 | Evolutionary simulations

To disentangle the demographic effects influencing the empirically observed patterns of LD, observed heterozygosity, Tajima's D, and nucleotide diversity, we performed individual-based forward evolutionary simulations using SLiM v3.0 (Haller & Messer, 2019). Substantial demographic changes may influence the observed patterns of LD and genomic variation. Previous evidence indicates *A. vermiculophyllum* underwent a shift from sexual reproduction to partial clonality (Krueger-Hadfield et al., 2016), a mitochondrial haplotype bottleneck (Kim et al., 2010; Krueger-Hadfield et al., 2016, 2017), and rapid population expansion (Weinberger et al., 2008) when establishing populations throughout the non-native range.

We performed forward evolutionary simulations with the following framework to emulate hypothesized invasion dynamics. The simulations began with 10,000 generations of a Wright-Fisher population with a census size of 1000 diploid individuals to reach drift-mutation equilibrium, and following the 10,000th generation, the demographic change occurred.

It is important to note these simulations simplify the *A. vermiculophyllum* haplodiplontic life cycle as simulations only included diploids and no cycling between free-living haploid and diploid stages. There are at least three principle reasons why we felt this was reasonable. First, many populations in the non-native range are dominated by the diploid stage; the number of diploids is much higher in the non-native range relative to the native range (81% vs. 58% diploid on average, respectively; Krueger-Hadfield et al., 2016). While we do not know how often fertilization occurs in the non-native range, we know the diploid populations are clonally reproducing via fragmentation. Second, haploids express recessive mutations (i.e., there are no heterozygotes) and could, potentially, facilitate a selective sweep. Moreover, haploid and diploid stages exploit different niches and are probably subject to different selection pressures (Hughes & Otto, 1999; see also Bessho & Otto, 2020), and in *A. vermiculophyllum*, ploidy impacts ecologically relevant phenotypes (Krueger-Hadfield & Ryan, 2020; Lees et al., 2018). Crucially, we have explicitly focused on only neutral evolutionary processes in our simulations, and thus we consider our simulations a conservative estimate of the their genomic impacts.

To mimic a severe population bottleneck, we simulated a single-generation population decline of varying severity (90%, 95%, 99%). Following the single-generation bottleneck, the population size grew in a single step to return to the census size of 1000 individuals. The simulation then proceeded for 200 more generations reflective of the likely *A. vermiculophyllum* invasion dynamics since non-native populations have been established for well less than 200 generations. To understand the genomic consequences of shift in reproductive mode (Krueger-Hadfield et al., 2016), clonality (rate of clonal reproduction; 10%, 50%, 70%, 90%, 99%, 99.5%, 100%) began immediately following the initial 10,000 generations and continued for the remainder of the simulation up to 200 generations. Similarly, population expansion (simple growth per generation; 0.05%, 0.1%, 0.15%) occurred after 10,000 generations and proceeded for the remainder of the simulation up to 200 generations. Based upon the results of the individual demographic effects, we combined the bottleneck and clonality as they fit our empirical observations. For the combination of both bottleneck and clonality, the bottleneck occurred for a single-generation at generation 10,000 and then grew to census size of 1,000. At the same time the bottleneck occurred, clonal reproduction was initiated and proceeded until the end of the simulation.

Further, we aimed to characterize how hypothesized demographic changes interact with crossover interference (Stapley et al., 2017). Positive crossover interference occurs when chromosomes of varying length have different per base pair recombination rates where longer chromosomes have lower per base pair recombination

rates. Here, we performed similar simulations as described above, but with the addition of several recombination rates (1e−7, 1e−8, 1e−9) throughout the entire simulation. Recombination rate was used as a proxy for chromosome length. We ran forward simulations for each recombination rate independently with a 99% bottleneck and fully clonal reproduction or fully sexual reproduction.

Using SLiM v3.0 (Haller & Messer, 2019), we performed 100 replicates of each simulation and exported vcf files. We exported vcf files before, immediately, 10, 25, 50, 100, and 200 generations following demographic change. Using the exported vcf files, we calculated LD (r^2) using VCFTOOLS v. 0.1.15 (Danecek et al., 2011) –geno-r2 call with a 100 kbp sliding window (vcftools --geno-r2 --min-r2 0.001 --ld-window-bp 100,000). Using the r^2 estimation from vcftools we calculated half-decay distance (Hill & Weir, 1988) using a custom R script. Additionally, we estimated observed heterozygosity (--freq2), nucleotide diversity (--site-pi), and Tajima's D (--TajimaD 2000) within each simulation using VCFTOOLS v. 0.1.15 (Danecek et al., 2011). Calculation of H_o and Pi used all variable sites

present across timepoints in a replicate simulation. Using allele frequencies estimated from vcftools we calculated per-locus F_{ST} in R to compare populations before to those after the demographic change (Reynolds et al., 1983). We plot the distribution of each result with boxplots and statistically assess differences with Tukey's HSD.

3 | RESULTS

3.1 | Linkage disequilibrium

Scaffold-wide LD was greater among non-native than among native populations. To visualize this, we plotted LD half-decay curves (Hill & Weir, 1988) along the 22 largest scaffolds and within the 22 populations for which we have 10 or more thalli (Figure 1; Figure S2). LD half-decay distance (W.G. Hill & Weir, 1988) – the physical distance where $r^2 = 0.5$ within a scaffold – is significantly greater among non-native populations than native populations (median = 47.2 and
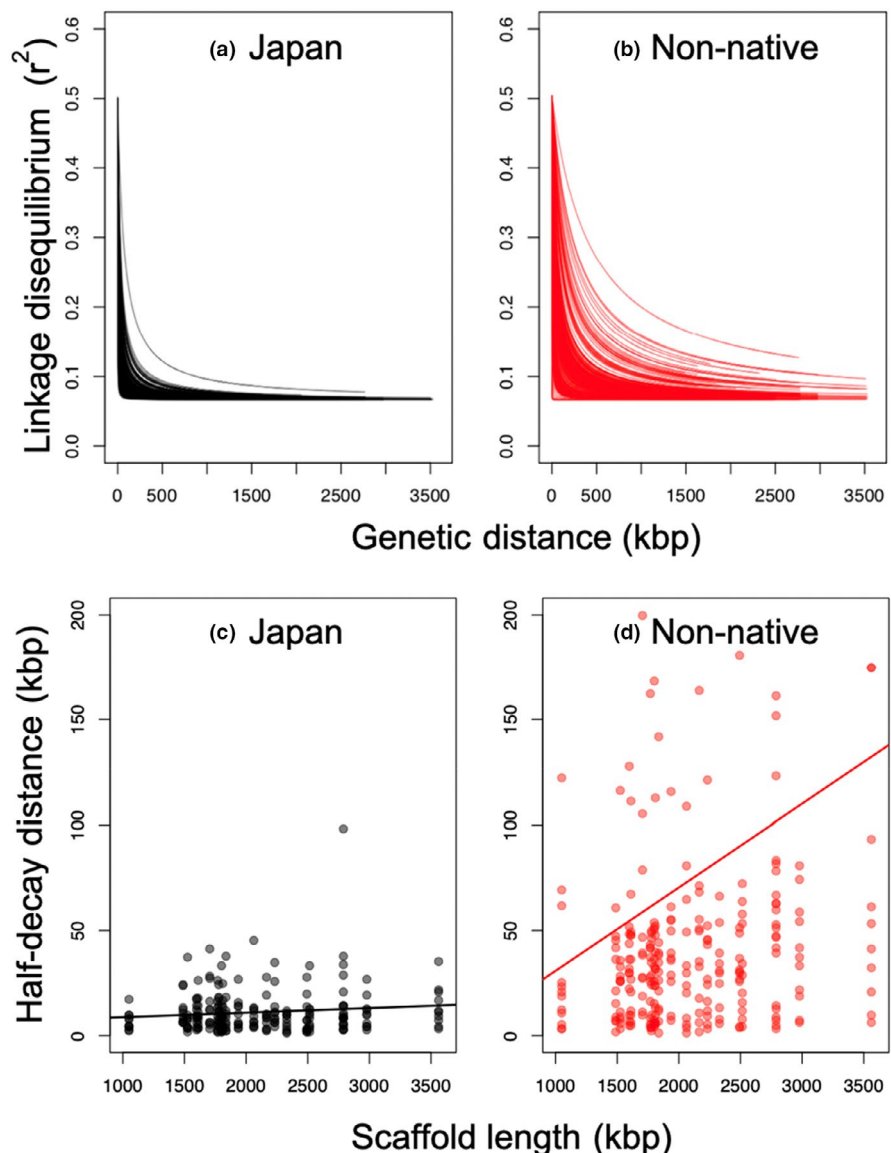

FIGURE 1 Linkage-disequilibrium decay and the correlation between scaffold size and half-decay distance for each scaffold within native Japanese and non-native populations. Genome-wide linkage disequilibrium (r^2) estimated for each scaffold within each population (a, b). Native Japanese populations (black) show lower linkage-disequilibrium than non-native populations. Half-decay distance estimated for each scaffold within each population (c, d). Non-native populations (red) show a strong correlation between half-decay distance and scaffold length. 12 points were removed from Japanese half-decay distance correlations because the half-decay distance was greater than 200 and to aid visualization [Colour figure can be viewed at wileyonlinelibrary.com]

8.5 kbp, respectively; $F_{1,20}$ = 12.9; p = .0018; Table 1; Figures 1a–b and 2a; Figure S2). Additionally, there is a significant effect of scaffold identity and the interaction of region and scaffold on half-decay distance (Table 1).

Evolutionary simulations indicate that elevated LD is principally generated by severe single-generation bottlenecks (>99%; Figure 2b), while weaker bottlenecks (90%, 95%) do not generate elevated LD (Figure S3). Though estimating the generation time of a partially clonal species is challenging, the timing of the invasion (i.e., near mid-20th century; Krueger-Hadfield et al., 2017) suggests a maximum of 50 generations. Our simulations also indicate that high rates of clonal reproduction (90%–100%) by itself (i.e., without bottlenecks, or in combination with 90% and 95% bottlenecks) do not generate elevated LD in the timeframe of this invasion (Figure 2b; Figure S3). Finally, our simulations indicate that a rapid population growth could not generate elevated LD by itself because rapid growth decreases LD over time (Slatkin, 1994) (Figure S4).

There was a tendency for the combination of high clonality (100%) and strong bottleneck (99%) to increase LD relative to simulations with a strong bottleneck only (Figure 2b), therefore we cannot preclude clonal reproduction as a contributor to the observed elevated LD in the non-native range. There was an overlap in the distributions of these simulation types, yet the distributions were statistically distinct (n = 100; Tukey's HSD p < .005 for generations 0, 100 and 200; p = .10 at generation 10).

We note that our 22 largest scaffolds ranged in size from 1.04 to 3.56 Mbp (Table S2; Figure S10). We estimated the effects of positive crossover interference, in which each scaffold experiences a relatively small, but equal number of crossover events during meiosis, yielding lower per base pair recombination rates on longer chromosomes. We detected a strong and significant interaction between native/non-native and scaffold length on half-decay distance (Table 2). In other words, half-decay distance increases with chromosome length in both native and non-native populations, but the increase with scaffold length is far steeper in the non-native region (Table 2; Figure 1c–d). To simulate shifts in per base pair recombination rates (which serves as a proxy for chromosome length), we altered recombination rate in our evolutionary simulations by 100-fold and assessed the consequences of clonality and bottleneck effects on LD. As expected, LD was greater with decreasing recombination rates, and, thus, by extension, among larger scaffolds

(Figures S5–S6). Further, clonality in combination with a bottleneck decreased crossover interference more rapidly than with bottleneck alone which maintained elevated levels of interference through the extent of the simulation (Figure S6).

3.2 | Genomic diversity

We explored genome-wide metrics of genetic diversity across native and non-native populations and within our evolutionary simulations. First, observed heterozygosity (H_o) was greater among non-native populations than among native populations ($F_{1,33}$ = 9.923; p = .003; Figure 2c). Our simulations indicate H_o is largely unaffected by our simulated invasion history (Figure 2d). Interestingly, the mean non-native H_o falls near H_o for the genetic source populations of the invasion, which were higher than all other native populations (Figure 2c). This suggests that the significantly greater H_o in invaded regions are a genetic legacy of the elevated H_o in the source population, which itself is greater than nonsource native populations as estimated with our marker set.

Second, Tajima's D for all populations was significantly greater than zero, except for three native populations (sou, akk were indistinguishable from D = 0; waj was significantly lower than 0). Tajima's D among non-native populations was greater (more positive) than those of native populations ($F_{1,33}$ = 10.822; p = .002; Figure 2e). Evolutionary simulations indicate the severe bottleneck scenario increases Tajima's D (Tajima, 1989) while clonality had no impact on Tajima's D either in isolation or in combination with a bottleneck (Figure 2f; Figure S3C).

Third, the proportion of polymorphic loci (Pi) was significantly greater in non-native populations ($F_{1,33}$ = 4.928; p = 0.033; Figure 2g). Similar to the H_o trend, the mean non-native Pi falls near Pi estimated for the genetic source of the invasion (Figure 2c) and the simulated hypothesized invasion histories show no effect on Pi (Figure 2h; Figure S3D).

3.3 | Population structure in the native range

There was substantial genetic differentiation among native Japanese populations at spatial scales of tens to hundreds of kilometers. The first and second PCs (Figure 3a) separated the native Japanese sites into four regional groups. Sites sampled along the Sea of Japan (shk, waj, hay) and Hokkaido (fut, akk, usu) coastlines were separated from other Japanese regions along the second PC. The second PC separated the sites along the southwestern Japanese coastline (hik, nag) from the remainder of northeastern sites. The northeastern Japanese sites (mou, sou, mng) were separated from all others via a combination of first and second PCs.

This strong, hierarchical, native differentiation in Japan was also reflected in the admixture analyses (Figure 4; Figure S7). At k = 2–12, populations generally split into two major regions: Sea of Japan and Hokkaido (shk, waj, hay, usu, akk, fut) versus

TABLE 1 ANOVA tables testing the effect of region (native vs. non-native populations), scaffold identity (1–22) and their interaction on the half-decay distance (kbp) of linkage disequilibrium

	df	MS	*F*-value	*p*-value
Region	1	4.167	11.6	.003
Scaffold ID	21	1.378	3.8	<.001
Region: scaffold ID	21	0.903	5.5	<.001

n = 22 populations treated as random intercept. Half-decay distance was log-transformed to normalize its distribution.

FIGURE 2 Descriptions of genomic diversity within empirical and simulated data. Empirical data (a, c, e, g) and forward evolutionary simulations (b, d, f, h) of half-decay distance (a, b), observed heterozygosity (c, d), Tajima's D (e, f) and proportion of polymorphic sites (Pi; g, h). The data for forward evolutionary simulations includes a shift to fully clonal reproduction (100% clonality), a severe bottleneck (99% Bottleneck), and a combination of a severe bottleneck and clonal reproduction (99% bottleneck and 100% clonality). Asterisks indicate significance level determined by ANOVA for empirical data (*<0.05, **<0.01). Black circles in c and g signify the primary sources of the invasion (mng, sou) [Colour figure can be viewed at wileyonlinelibrary.com]
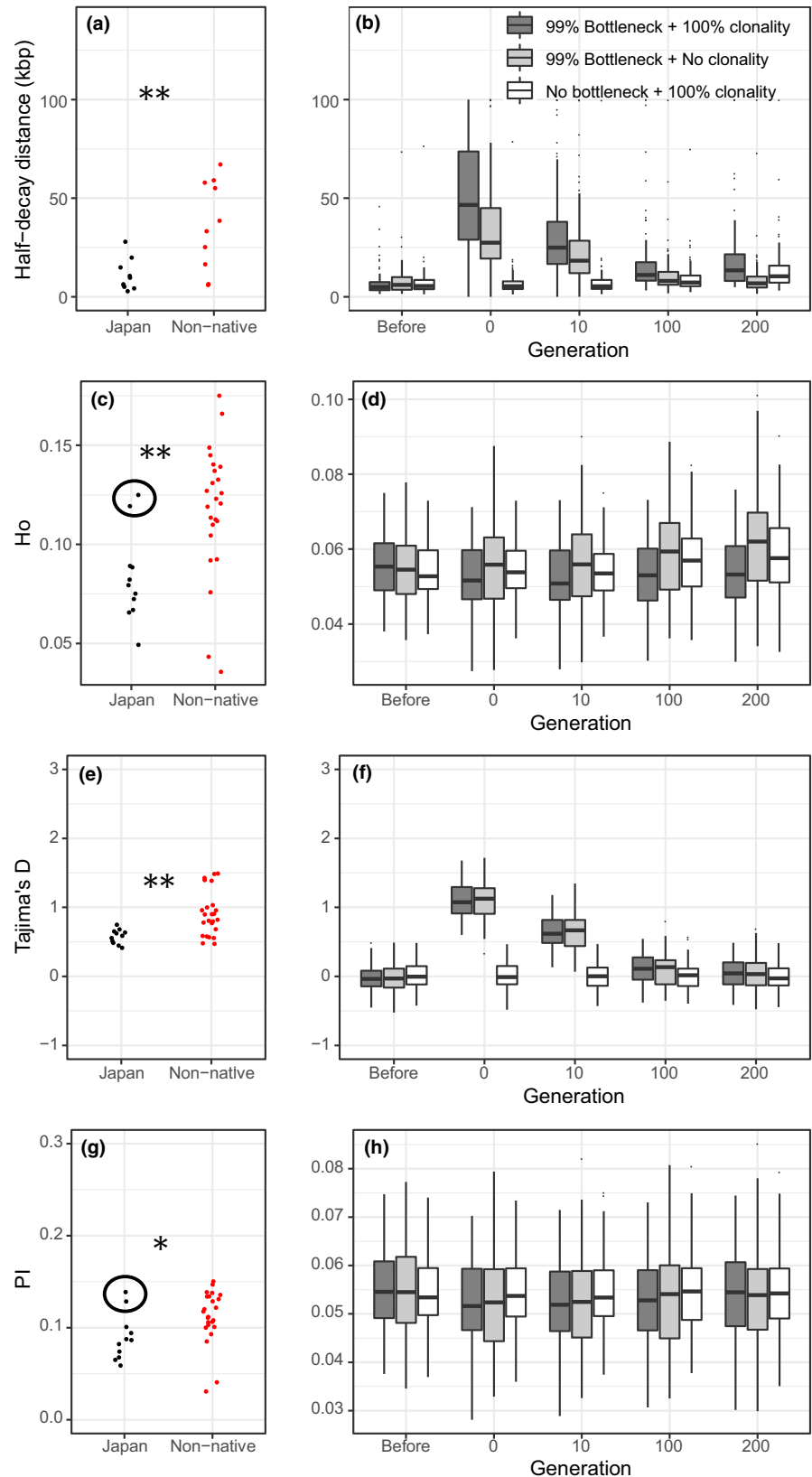
the remaining Pacific coast populations (mng, sou, mou hik, nag). With increasing *k*, population structure emerged within each of these regions. For example, at *k* = 8, there were six genetic clusters (1-hik, nag; 2-mng, sou; 3-shk, waj; 4-hay; 5-usu; 6-akk, fut,

with mou a mix of two clusters that dominate mng/sou and usu). We detect high levels of differentiation throughout native Japan where *k* = 7 (Figure 4a; see Figure S8 for Evanno et al. (2005) analysis of likelihoods). Increasing *k* to up to 16 (Figure S7) did

not change these patterns of population-level differentiation in the native range.

Pairwise F_{ST} values among all Japanese populations were relatively high (mean ± SE; F_{ST} = 0.58 ± 0.02; Figure S9), with the exception of mng vs. sou, which was relatively low (F_{ST} = 0.22), probably because mng and sou were separated by only ~40 km; the closest of any two native sites sampled. In contrast, pairwise F_{ST} values within eastern North America and Europe - two non-native shorelines

TABLE 2 The effect of region (native vs. non-native populations), scaffold size (kbp) and their interaction on half-decay distance

	df	MS	F-value	p-value
Region	1	0.731	1.9	.183
Scaffold length	1	13.41	33.8	<.001
Region: scaffold length	1	4.209	10.6	.001

n = 22 populations treated as random intercept. Half-decay distance was log-transformed to normalize its distribution.

of comparable geographic extent - had far lower levels of F_{ST} (0.24 ± 0.03 and 0.28 ± 0.04, respectively; Figure S9). We could not make a comparison along western North America because only one population (tmb) had greater than 10 MLGs based on P_{sex}.

3.4 | Genetic structure of non-native versus native range

Strong differentiation among native populations provides an opportunity to robustly assess the source populations of the introductions (Geller et al., 2010). Overall, analyses indicate that nearly all non-native regions were sourced from mng/sou, with the exception of the Pacific Northwest populations (moo, ptw, eld) that have a mix of nag, hik, mng and sou, confirming and elaborating on the previous findings of Krueger-Hadfield et al. (2017). We explore each of the non-native regions in turn.

First, populations of the Pacific Northwest (Washington, USA: ptw, eld; and British Columbia, Canada: moo) were distinguished from other
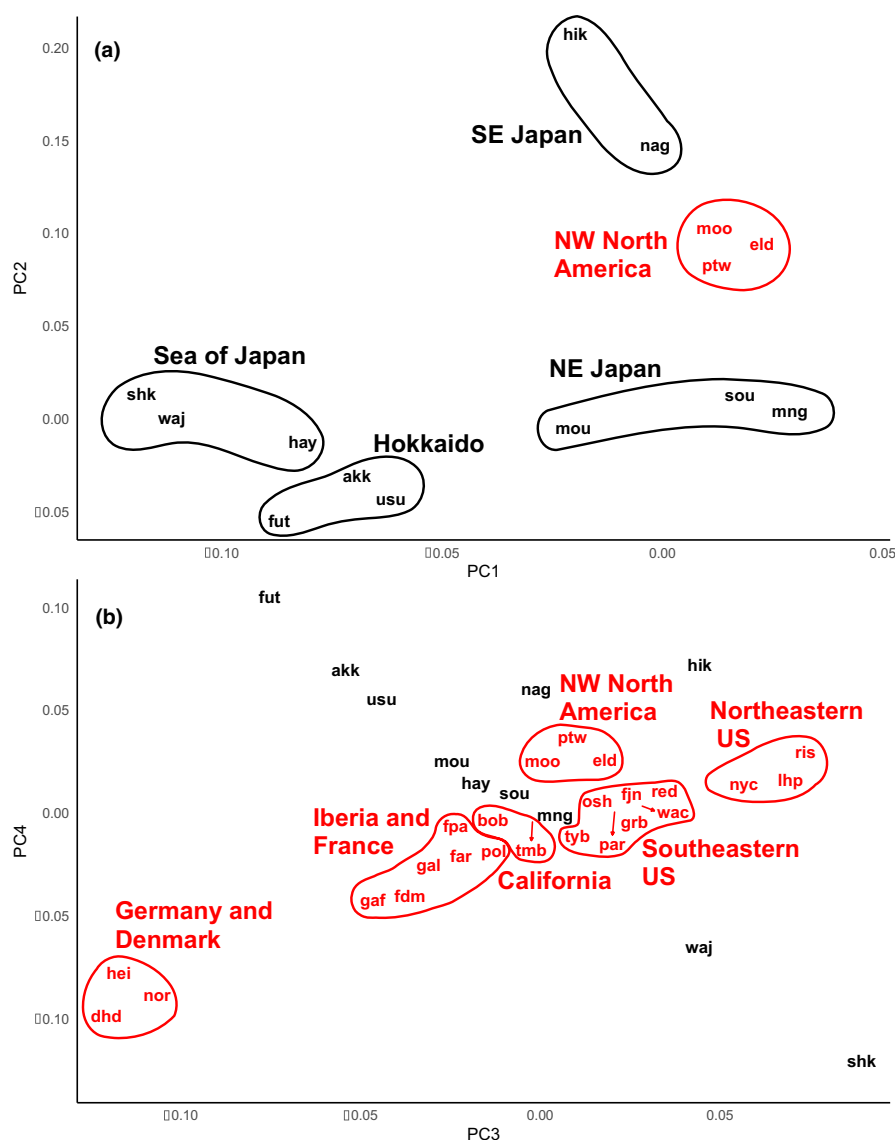

FIGURE 3 A principal components analysis (PCA) of genotype likelihoods. (a) First and second PCs strongly delineate native Japanese (black) and non-native populations (red). (b) Third and Fourth PCs delineate non-native populations. Variance explained by PC1 thru PC4 is 14.9, 7.4, 6.4, 5.8%, respectively. N = 351 unique diploid thalli. Site abbreviations are in Table S1; Site geographic locations visualized in Figure S1 [Colour figure can be viewed at wileyonlinelibrary.com]
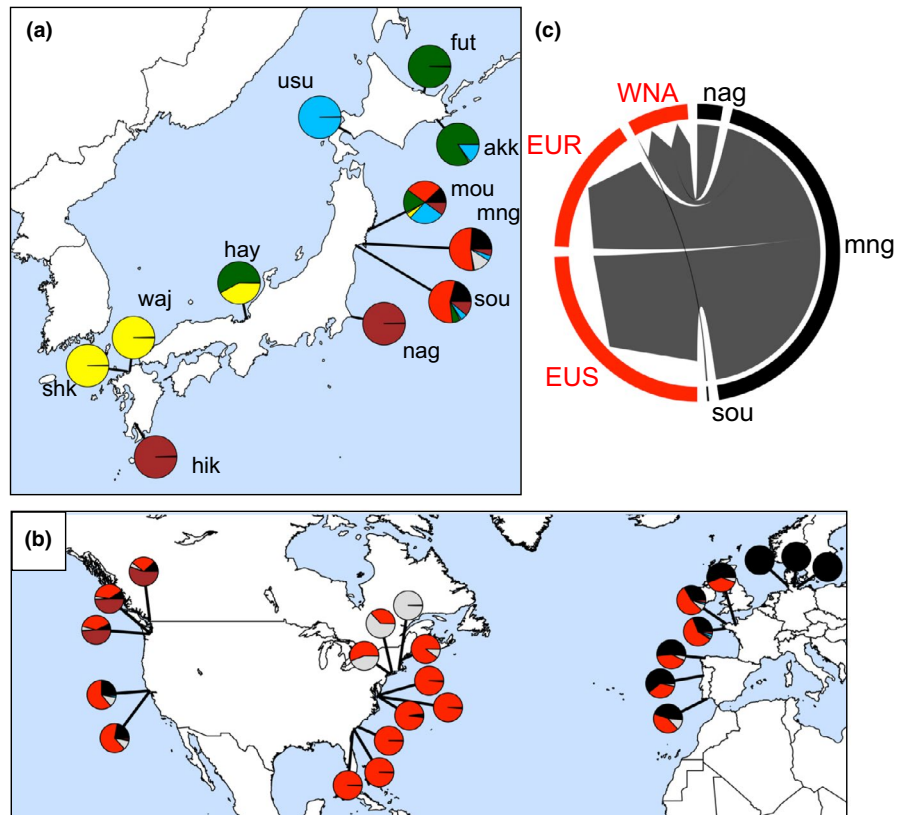
non-native populations in the second PC (Figure 3a), sitting half-way between NW Japan (mng, sou) and SW Japan (nag, hik) samples. An assignment of non-native thalli using PCA and LDA showed that approximately half of thalli in the Pacific Northwest were assigned to nag, and the other half to mng and sou (Figure 4c). Similarly, in the admixture analysis (Figure 4a), the Pacific Northwest contained genetic clusters that dominate mng/sou and nag/hik (Figure 4b).

Second, Californian (tmb, bob) populations are most closely aligned with mng and sou in PCA space (Figure 3), admixture (Figure S7), and pairwise F_{ST} (tmb; Figure S9). There is also some suggestion that mou may have served as a source to California from the admixture analysis, although this effect disappears at higher number of genetic clusters (Figure S7).

Third, there appears to be two separate introductions along the east coast of North America, although mng/sou was the native source populations for both events. Along the third PC (Figure 3b) and in admixture plots (Figure S7), three populations in Long Island Sound (lhp, nyc, ris) are distinct from the rest of the populations (grb in New Hampshire; wac, red, and osh in Virginia; fjn and par in South Carolina, and tyb in Georgia). There is also suggestion of a mix of these two groups in nyc and lhp, particularly in the admixture analysis (Figure S7).

Finally, PCA (Figure 3b) and admixture (Figure 4c) data are consistent with separate introductions to northern Europe (Germany and Denmark: hei, dhd, nor) and the Atlantic coast of Europe (France: pol, fdm, fpa; Spain: gal; Portugal: gaf and far). The ultimate native source of European thalli was most likely mng/sou based on PCA and F_{ST} analyses. The mou population also clustered with non-native

populations as well as nearby mng and sou, indicating some genetic relatedness of this extant population and the historic population(s) of the source material (Figures 3 and 4).

3.5 | F_{ST} outliers

We compared the per-locus F_{ST} values of the two native source populations (mng, sou) versus the most closely-related non-native population on each continental margin (tmb on western North America; osh on eastern United States, and fdm in Europe). The sets of SNPs that are outliers in this analysis (defined as F_{ST} <1%) differ between the three comparisons (Figure 5a,c,e), indicating a complex genomic response to the invasions with no consistently differentiated genomic region. Per-locus F_{ST} across forward evolutionary simulations show highly stochastic response where some replicates exhibit high differentiation between the founding population and the population following the severe bottleneck and shift to clonal reproduction (Figure 5b,d,f).

4 | DISCUSSION

By sequencing and assembling the first *A. vermiculophyllum* genome and comparing empirical genomic data with evolutionary simulations, we determined drift-driven founder effects are the primary forces shaping genomic diversity and LD in the non-native range
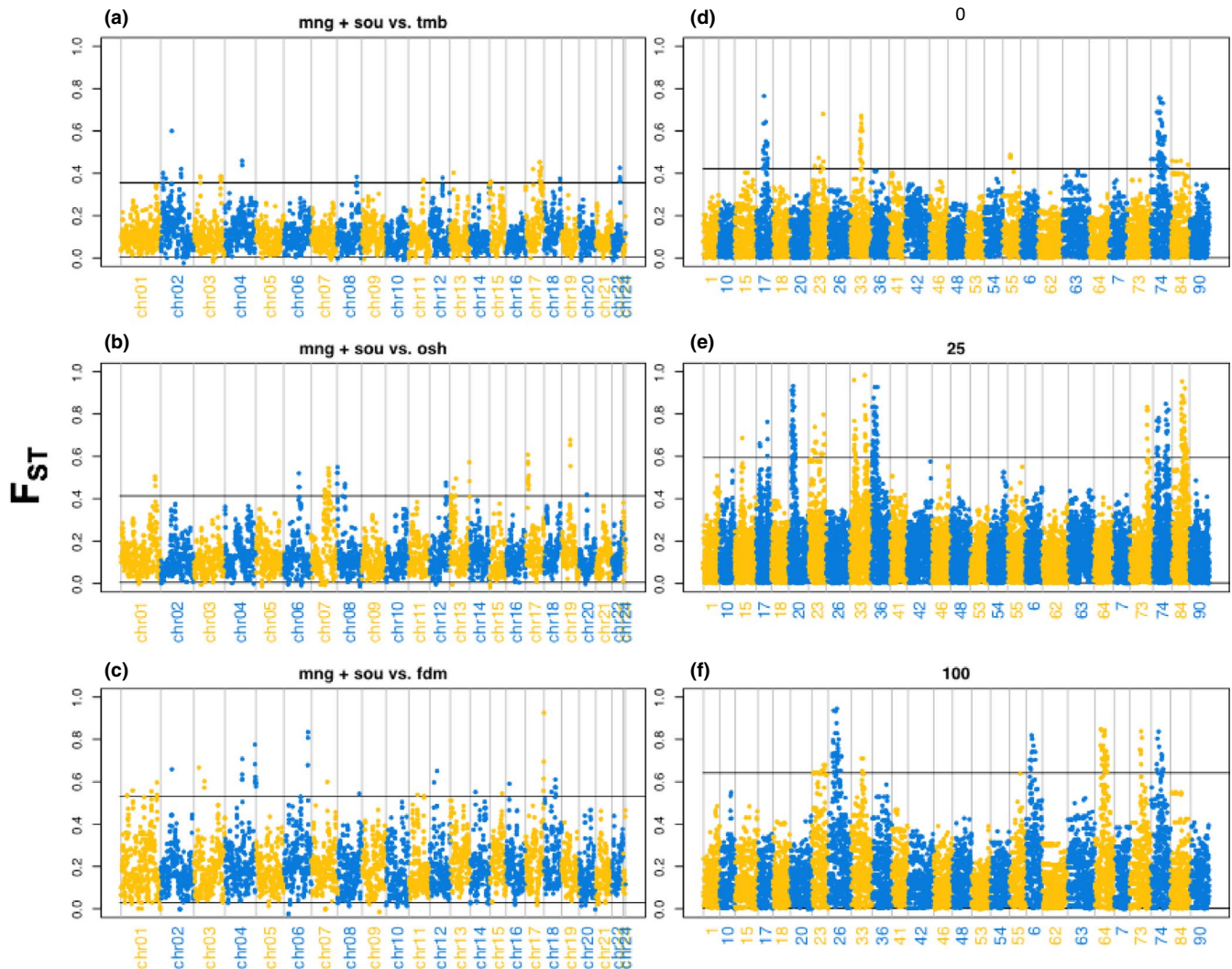
FIGURE 5 Per-locus F_{ST} distribution estimated using empirical and simulation data. Empirical per-locus F_{ST} was estimated between combined native Japanese source populations (mng/sou) and the most closely related non-native population along the west coast of North America (a, tmb), the east coast of the United States (b, osh) and Europe (c, fdm). Simulation F_{ST} values estimated for 25 randomly sampled replicate simulations before and 0 (d), 25 (e) and 100 (f) generations since single-generation bottleneck (99% reduction) and shift to clonal reproduction (100%). Lines indicate the 99% F_{ST} distribution for comparison purposes only. Empirically, no loci were consistently differentiated across all invaded regions and simulations show how neutral processes can generate outliers loci [Colour figure can be viewed at wileyonlinelibrary.com]

(Figure 2). Because of the varied evolutionary processes impacting non-native *A. vermiculophyllum* populations, we discuss the implications of a single-generation bottleneck and altered reproductive mode on evolutionary change. Further, we describe patterns consistent with crossover interference, and how documented rapid adaptive changes may have occurred even though non-native populations have undergone a severe population bottleneck.

4.1 | Relative importance of a single-generation bottleneck and clonal reproduction

Using reduced representation genomic sequencing, the *A. vermiculophyllum* genome and forward evolutionary simulations, we conclude

that increased levels of scaffold-wide LD and elevated Tajima's D in the non-native range is consistent with a strong bottleneck (e.g., a single generation of 99% or greater) with or without high rates of clonal reproduction (Figure 2). While a bottleneck primarily contributed to the elevated LD in the non-native range, we cannot exclude the possibility that a bottleneck in concert with clonal reproduction impacted non-native LD because the simulated combination elevated the LD further than the bottleneck alone. The simulations indicate that weaker bottlenecks of a single-generation (e.g., 95%) could not have generated the elevated LD that we documented, although they could have generated the Tajima's D pattern (Figure S3). We also note that weaker bottlenecks that last longer than a single-generation may also generate elevated levels of LD; we did not simulate these conditions.

In contrast, simulations indicate that increased rates of clonal reproduction in the non-native range by itself could not generate these empirical patterns (Figure 2). Previously, Krueger-Hadfield et al. (2016) reported an increase in the number of microsatellite-based clonal diploids in non-native populations (mean ± SD = 0.36 ± 0.31; range = 0–0.94; n = 35 populations) relative to native populations (mean ± SD = 0.09 ± 0.12; range = 0–0.44; n = 30 populations), consistent with an increase in the frequency of clonal reproduction in the non-native range. However, it seems likely that few to no non-native populations are reproducing clonally at the simulated rate of 100%. There are some populations with unique genotypes, suggesting either enormous population sizes, low levels of clonal fragmentation, or continual input of sexually produced diploids from nearby sexually reproducing populations that are being found with increased sampling intensity in some areas (Krueger-Hadfield & Ryan, 2020; Krueger-Hadfield et al., 2018). Indeed, there are non-native populations with reproductively mature gametophytes and signatures of sexual reproduction (Krueger-Hadfield et al., 2016). Thus, our simulated 100% clonal reproduction probably overestimates true clonal rates. Clonality may have modified the consequences of a bottleneck, but it did not primarily drive observed patterns (see also Navascués et al., 2010; Stoeckel et al., 2021).

While *A. vermiculophyllum* populations show higher rates of clonal reproduction in the non-native range (Krueger-Hadfield et al., 2016), uniparental reproduction through self-fertilization can also facilitate population expansion (Baker, 1955). While this has not yet been fully explored in *A. vermiculophyllum*, it is possible for dioicous haplodiplontic taxa to undergo selfing (see Klekowski, 1969; Krueger-Hadfield, 2020; Valero et al., 2001). Partial selfing in the invasive brown macroalga, Sargassum muticum, probably facilitated the success of non-native population establishment and maintenance (Engelen et al., 2015; Le Cam et al., 2020). The authors posit that uniparental reproduction may help in early-stage establishment by generating novel genotypes through recombination, and then facilitate increases in population size during the later stages of the invasion.

4.2 | Linkage disequilibrium and crossover interference

With the *A. vermiculophyllum* genome, we were able to estimate LD on each scaffold and we observed a dramatic increase in the extent of LD in the non-native range when compared to the native range (Figure 1). Variation in LD allows for the inference of past population dynamics including bouts of selection, mutation, hybridization, and change in population size (Pritchard & Przeworski, 2001). LD is also important to generating genotype-phenotype correlations that help to understand the genetic basis of phenotypic evolution (Lowry et al., 2016). When a population undergoes a bottleneck, fewer genomes exist in subsequent generations resulting in high correlation between variant loci as seen in crop domestication (Ching et al., 2002; Maccaferri et al., 2005) and the invasive weed, *Mikania*

micrantha (Banerjee et al., 2020; Wang et al., 2007). LD will be further increased if the bottleneck severely reduces population size (Wall et al., 2002). In contrast, rapid population expansion decreases LD because the expanding population experiences high levels of recombination with mutational input breaking up coinherited loci because LD depends on recombination rate and population size (Hill, 1981). Therefore, we attribute the dramatic LD increase primarily to the few number of individuals which established non-native populations as corroborated by our simulations.

Clonal reproduction appears to magnify the effect of the bottleneck on LD. Recombination ceases when a population undergoes exclusively clonal reproduction, and as a consequence, LD increases (Awadalla, 2003; Tibayrenc et al., 1991). A hallmark of clonal reproduction is the temporal, but gradual increase in LD with a fully clonal population. The shift to clonal reproduction alone does not immediately elevate LD but instead requires 100–200 generations to accumulate an effect (Figure S3A).

We were also able to document an increase in the strength of the bottleneck effect on LD with scaffold length (Figure 1; Figures S5 and S6). It is likely that this is due to positive crossover interference (though see Stapley et al., 2017 for other mechanisms that may alter recombination rate). During meiosis, chromosomes experience the same absolute number of crossover events per chromosome, and consequently longer chromosomes have lower number of recombination events per base pair relative to shorter chromosomes. We infer positive crossover interference in both the native and non-native range because longer chromosomes have higher half-decay distance. This effect is magnified in the non-native range from a bottleneck. Our simulations with variable mutation rates recapitulate these empirical patterns and that 100% clonality slightly reduces the strong effect of bottleneck (Figures S5 and S6).

4.3 | Genomic diversity

Most genetic variation is lost in a population bottleneck due to drift, yet the rate of loss depends on the strength and timing of the bottleneck (Nei et al., 1975). When the bottleneck size is greater than a single individual, the number of generations at a bottlenecked population size is minimal and subsequent growth rate is high, the loss of heterozygosity will be marginal (Nei et al., 1975). Strong bottlenecks will cause the loss of low frequency alleles (Cornuet & Luikart, 1996). In the *A. vermiculophyllum* invasion, we detect a significant increase in H_o for non-native populations (Figure 2c; see also Krueger-Hadfield et al., 2016 for microsatellites), but when comparing native genetic source populations with the non-native range, there is no difference in H_o (Figure 2c). The empirical results are consistent with all simulated invasion histories as there were limited effects of simulation type on H_o (Figure 2d). Moreover, the loss of low frequency alleles in a bottleneck eliminates rare variants, which results an excess of variants existing at intermediate frequencies causing in a positive Tajima's D (Tajima, 1989) (Figure 2e–f).

A nonmutually exclusive explanation for the higher diversity in the non-native range is that the non-native range was seeded by multiple native populations. Indeed, there is evidence (Figures 2 and 3c) that the northwestern North American populations represent a mix of southeastern (nag) and eastern Japanese populations (mng, sou). For simplicity, we did not simulate multiple invasions, therefore we cannot determine how resultant admixture would impact the empirical genomic diversity estimates when compounded with the evaluated bottleneck and clonality effects.

4.4 | Refinement of invasion source and invasion pathway

Identifying the genetic source of an invasion enables the accurate characterization of evolutionary processes impacting non-native populations (Geller et al., 2010). Based on microsatellites, Krueger-Hadfield et al. (2017) identified the Pacific coastline of northeast Japan as the genetic source of the invasion. Here, by increasing the number of genetic markers, we were able to detect higher levels of differentiation among populations in the native range which enabled further refinement of the genetic source of the invasion (Figures 3 and 4). Our analyses indicate that two populations, Mangoku-Ura (mng) and Soukanzan (sou), in Matashima Bays, acted as the source for the majority of invaded regions in the United States and Europe. These two sites are important areas of oyster aquaculture (see Krueger-Hadfield et al., 2017 for a discussion on the pathway of the oyster and its role in the *A. vermiculophyllum* invasion).

4.5 | Implications for quantitative traits and rapid adaptation

Invasive species exhibit common traits, such as increased growth rates, fecundity, and stress tolerance (van Kleunen et al., ,2010, 2015) all of which are highly polygenic. Severe bottlenecks can increase variation for quantitative traits because founder effects increase additive genetic variance when compared to the ancestral population (Goodnight, 1987, 1988; Whitlock et al., 1993). Consequently, newly established non-native populations undergoing a bottleneck may show rapid evolutionary change due to increased additive genetic variation when the non-native populations experience a high strength of directional selection (Blows & McGuigan, 2015; Falconer & Mackay, 1996).

Non-native *A. vermiculophyllum* populations have expanded into niches that are not found in the native range and rapidly evolved greater tolerance for acute high temperatures and low salinities (Sotka et al., 2018). The rapid evolutionary change in the non-native range may have been a result of a bottleneck-induced increase of additive gentic variation. We also point out that rapid adaptation to the non-native environment may have been influenced by the flexible *A. vermiculophyllum* reproductive mode. Bazin et al. (2014) simulated invasion adaptation dynamics to a novel environments and

found fully asexual reproduction produced lower invasion success while high rates of asexual reproduction, but not fully asexual (i.e., partially clonal), produced the highest invasion success. Bazin et al. (2014) argue that mixed mating increases additive genetic variance relative to populations which are fully clonal, as the latter decreases trait variance due to the Hill-Robertson effect (Felsenstein, 1974; Hill & Robertson, 1966). Most non-native populations are partially clonal as well as probably engaging in a mix of outcrossing and selfing (Krueger-Hadfield et al., 2016, 2017), suggesting that the effect described by Bazin et al. (2014) may have played a role in maintaining adaptive variation.

Alternatively, molecular adaptation may result in loci that are particularly differentiated when compared to native genetic source populations due to selection early in the invasion. This allelic differentiation resulting from selection may be difficult to detect due to nonequilibrium invasion scenarios. We estimated locus-specific F_{ST} for a single population from each invaded coastline and found no shared outliers among the three regions (Figure 5). To determine the impact of our proposed single-generation bottleneck invasion history on locus-specific differentiation, we estimated F_{ST} between pre- and post-bottleneck populations. Many representative simulations exhibited extremely high levels of differentiation in the absence of selection (Figure 5). Thus, interpreting these results can be difficult, but controlling for demographic history may still be possible (Hoban et al., 2016; Lotterhos & Whitlock, 2014). While we attribute the lack of shared outliers among invaded coastlines to complex demography, we cannot eliminate the possibility of context-dependent selection resulting in uniquely differentiated loci driven by differing selective pressures present in each non-native region. Investigating rapid evolutionary changes associated with invasions may benefit from a polygenic model approach to determine adaptive changes (Bourret et al., 2014; Le Corre & Kremer, 2012).

5 | CONCLUSIONS

An increase in LD along with estimates of genomic diversity indicate a single-generation population bottleneck probably occurred when *A. vermiculophyllum* established populations along each of the coastlines of North America and Europe. Empirical results were corroborated by forward evolutionary simulations based on hypothesized demographic changes associated with the invasion. Additionally, we were able to determine that two sites previously identified as important by Krueger-Hadfield et al. (2017) were the ultimate source of the Northern Hemisphere invasion of *A. vermiculophyllum*, corroborating earlier work suggesting oysters as the primary vector for initial invasions. Additionally, we failed to detect loci consistently differentiated between the native source populations and non-native regions, despite previous evidence of rapid phenotypic evolution facilitating the invasion (Sotka et al., 2018). Our simulations suggest that the *A. vermiculophyllum* invasion demography could lead to spurious genomic outlier inference. Further work is needed to determine the genomic architecture of

the rapid evolutionary change which may be beneficial for implementing a strategy to detect polygenic adaptation or a robust null models specific to the *A. vermiculophyllum* invasion history, but also to other invaders more generally. Future work may benefit from implementing a quantitative genetic approach to estimate phenotypic evolution occurring during invasions. This approach may uncover previously overlooked patterns of rapid adaptation occurring in non-native populations.

AUTHOR CONTRIBUTIONS

Ben A. Flanagan, Stacy A. Krueger-Hadfield, and Erik E. Sotka collected samples; Ben A. Flanagan and Stacy A. Krueger-Hadfield extracted DNA; Ben A. Flanagan, Stacy A. Krueger-Hadfield, and Chris C. Nice performed library preparation; Ben A. Flanagan performed all simulations. All authors conceived project, analysed data, and wrote and edited the manuscript.

DATA AVAILABILITY STATEMENT

Microsatellite genotypes were deposited in DRYAD (https://doi.org/10.5061/dryad.fn53k). RAD-seq data, reference genome, and R scripts were deposited to DRYAD at https://doi.org/10.5061/dryad.dv41ns1xc (Flanagan et al., 2021). Raw sequencing reads are available as an NCBI BioProject (PRJNA700770), and the assembled Agarophyton vermiculophyllum genome was submitted to NCBI BioSample (SAMN17860744).

ORCID

Ben A. Flanagan https://orcid.org/0000-0002-0204-6139
Stacy A. Krueger-Hadfield https://orcid.org/0000-0002-7324-7448
Courtney J. Murren https://orcid.org/0000-0003-0361-1790
Erik E. Sotka https://orcid.org/0000-0001-5167-8549

REFERENCES

Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrao, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, *16*(24), 5115–5139. https://doi.org/10.1111/j.1365-294X.2007.03535.x

Awadalla, P. (2003). The evolutionary genomics of pathogen recombination. *Nature Reviews Genetics*, *4*(1), 50–60. https://doi.org/10.1038/nrg964

Bailleul, D., Stoeckel, S., & Arnaud-Haond, S. (2016). RClone: a package to identify multilocus clonal lineages and handle clonal data sets in R. *Methods in Ecology and Evolution*, *7*(8), 966–970. https://doi.org/10.1111/2041-210X.12550

Baker, H. G. (1955). Self-compatibility and establishment after "long-distance" dispersal. *Evolution*, *9*(3), 347–349.

Baker, H. G., & Stebbins, G. L. (1965). Genetics of colonizing species, proceedings. In *International Union of Biological Sciences Symposia on General Biology 1964: Asilomar, Calif.).*

Balloux, F., Lehmann, L., & de Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, *164*(4), 1635–1644.

Banerjee, A. K., Hou, Z., Lin, Y., Lan, W., Tan, F., Xing, F., Li, G., Guo, W., & Huang, Y. (2020). Going with the flow: analysis of population structure reveals high gene flow shaping invasion pattern and inducing range expansion of *Mikania micrantha* in Asia. *Annals of Botany*, *125*(7), 1113–1126. https://doi.org/10.1093/aob/mcaa044

Bazin, É., Mathé-Hubert, H., Facon, B., Carlier, J., & Ravigné, V. (2014). The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. *Biological Invasions*, *16*(4), 875–886. https://doi.org/10.1007/s10530-013-0544-6

Bellorin, A. M., Oliveira, M. C., & Oliveira, E. C. (2004). *Gracilaria vermiculophylla*: A western Pacific species of Gracilariaceae (Rhodophyta) first recorded from the eastern Pacific. *Phycological Research*, *52*(2), 69–79. https://doi.org/10.1111/j.1440-183.2004.00330.x

Bessho, K., & Otto, S. P. (2020). Fixation and effective size in a haploid-diploid population with asexual reproduction. *BioRxiv*. https://doi.org/10.1101/2020.09.17.295618

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R. U., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, *26*(7), 333–339. https://doi.org/10.1016/j.tree.2011.03.023

Blows, M. W., & McGuigan, K. (2015). The distribution of genetic variance across phenotypic space and the response to selection. *Molecular Ecology*, *24*(9), 2056–2072. https://doi.org/10.1111/mec.13023

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W. E., Siemann, E., & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, *144*(1), 1–11. https://doi.org/10.1007/s00442-005-0070-z

Bourne, S. D., Hudson, J., Holman, L. E., & Rius, M. (2018). Marine invasion genomics: Revealing ecological and evolutionary consequences of biological invasions. *Population Genomics: Marine Organisms*, 363–398. https://doi.org/10.1007/13836_2018_21

Bourret, V., Dionne, M., & Bernatchez, L. (2014). Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: Polygenic multilocus analysis surpasses genome scan. *Molecular Ecology*, *23*(18), 4444–4457. https://doi.org/10.1111/mec.12798

Casso, M., Turon, X., & Pascual, M. (2019). Single zooids, multiple loci: Independent colonisations revealed by population genomics of a global invader. *Biological Invasions*, *21*(12), 3575–3592. https://doi.org/10.1007/s10530-019-02069-8

Chapman, J. A., Ho, I., Sunkara, S., Luo, S., Schroth, G. P., & Rokhsar, D. S. (2011). Meraculous: *de novo* genome assembly with short paired-end reads. *PLoS One*, *6*(8), e23501. https://doi.org/10.1371/journal.pone.0023501

Ching, A. D. A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O. S. H., Tingey, S., Morgante, M., & Rafalski, A. J. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, *3*(1), 19. https://doi.org/10.1186/1471-2156-3-19

Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, *144*(4), 2001–2014.

Cristescu, M. E. (2015). Genetic reconstructions of invasion history. *Molecular Ecology*, *24*(9), 2212–2225. https://doi.org/10.1111/mec.13117

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, *24*(9), 2095–2111. https://doi.org/10.1111/mec.13183

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Engelen, A. H., Serebryakova, A., Ang, P., Britton-Simmons, K., Mineur, F., Pedersen, M. F., Arenas, F., Fernandez, C., Steen, H., & Svenson, R. (2015). Circumglobal invasion by the brown seaweed *Sargassum muticum*. In *Oceanography and marine biology: An annual review* (Vol. 53, pp. 81–126). Taylor and Francis.

Estoup, A., & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology*, *19*(19), 4113–4130. https://doi.org/10.1111/j.1365-294X.2010.04773.x

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics. Essex*. Longman Group.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, *78*(2), 737–756.

Flanagan, B. A., Krueger-Hadfield, S. A., Murren, C. J., Nice, C. C., Strand, A. E., & Sotka, E. E. (2021). Founder effects shape linkage disequilibrium and genomic diversity of a partially clonal invader. *Dryad*, Dataset, https://doi.org/10.5061/dryad.dv41ns1xc

Fox, E. A., Wright, A. E., Fumagalli, M., & Vieira, F. G. (2019). NgsLD: Evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics*, *35*(19), 3855–3856. https://doi.org/10.1093/bioinformatics/btz200

Fumagalli, M., Vieira, F. G., Korneliussen, T. S., Linderoth, T., Huerta-Sánchez, E., Albrechtsen, A., & Nielsen, R. (2013). Quantifying population genetic differentiation from next-generation sequencing data. *Genetics*, *195*(3), 979–992. https://doi.org/10.1534/genetics.113.154740

Geller, J. B., Darling, J. A., & Carlton, J. T. (2010). Genetic perspectives on marine biological invasions. *Annual Review of Marine Science*, *2*, 367–393. https://doi.org/10.1146/annurev.marine.010908.163745

Goodnight, C. J. (1987). On the effect of founder events on epistatic genetic variance. *Evolution*, *41*(1), 80–91. https://doi.org/10.1111/j.1558-5646.1987.tb05772.x

Goodnight, C. J. (1988). Epistasis and the effect of founder events on the additive gentic variance. *Evolution*, *42*(3), 441–454. https://doi.org/10.1111/j.1558-5646.1988.tb04151.x

Grotkopp, E., & Rejmanek, M. (2007). High seedling relative growth rate and specific leaf area are traits of invasive species: Phylogenetically independent contrasts of woody angiosperms. *American Journal of Botany*, *94*(4), 526–532. https://doi.org/10.3732/ajb.94.4.526

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics*, *30*(19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Guillemin, M. L., Faugeron, S., Destombe, C., Viard, F., Correa, J. A., & Valero, M. (2008). Genetic variation in wild and cultivated populations of the haploid- diploid red alga *Gracilaria chilensis*: how farming practices favor asexual reproduction and heterozygosity. *Evolution*, *62*(6), 1500–1519. https://doi.org/10.1111/j.1558-5646.2008.00373.x

Halkett, F., Simon, J., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, *20*(4), 194–201. https://doi.org/10.1016/j.tree.2005.01.001

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. *Molecular Biology and Evolution*, *36*(3), 632–637. https://doi.org/10.1093/molbev/msy228

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research*, *38*(3), 209–216. https://doi.org/10.1017/S0016672300020553

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, *8*(3), 269–294. https://doi.org/10.1017/S0016672300010156

Hill, W. G., & Weir, B. S. (1988). Variances and covariances of squared linkage disequilibria in finite populations. *Theoretical Population Biology*, *33*(1), 54–78. https://doi.org/10.1016/0040-5809(88)90004-4

Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: Tools for population and evolutionary genetics. *Nature Reviews Genetics*, *13*(2), 110–122. https://doi.org/10.1038/nrg3130

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*(4), 379–397. https://doi.org/10.1086/688018

Hughes, J. S., & Otto, S. P. (1999). Ecology and the evolution of biphasic life cycles. *The American Naturalist*, *154*(3), 306–320. https://doi.org/10.1086/303241.

Kapraun, D. F., & Freshwater, D. W. (2012). Estimates of nuclear DNA content in red algal lineages. *AoB PLANTS*, *2012*. https://doi.org/10.1093/aobpla/pls005

Kim, S. Y., Weinberger, F., & Boo, S. M. (2010). Genetic data hint at a common donor region for invasive Atlantic and Pacific populations of *Gracilaria vermiculophylla* (Gracilariales, Rhodophyta). *Journal of Phycology*, *46*(6), 1346–1349. https://doi.org/10.1111/j.1529-8817.2010.00905.x

Klekowski, E. J. (1969). Reproductive biology of the Pteridophyta. II. Theoretical considerations. *Botanical Journal of the Linnean Society*, *62*(3), 347–359. https://doi.org/10.1111/j.1095-8339.1969.tb01972.x

Kolar, C. S., & Lodge, D. M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, *16*(4), 199–204. https://doi.org/10.1016/S0169-5347(01)02101-2

Kollars, N. M., Krueger-Hadfield, S. A., Byers, J. E., Greig, T. W., Strand, A. E., Weinberger, F., & Sotka, E. E. (2015). Development and characterization of microsatellite loci for the haploid–diploid red seaweed *Gracilaria vermiculophylla*. *PeerJ*, *3*, e1159. https://doi.org/10.7717/peerj.1159

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Korneliussen, T. S., Moltke, I., Albrechtsen, A., & Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. *BMC Bioinformatics*, *14*(1), 289. https://doi.org/10.1186/1471-2105-14-289

Krueger-Hadfield, S. A. (2020). What's ploidy got to do with it? Understanding the evolutionary ecology of macroalgal invasions necessitates incorporating life cycle complexity. *Evolutionary Applications*, *13*(3), 486–499. https://doi.org/10.1111/eva.12843

Krueger-Hadfield, S. A., Kollars, N. M., Byers, J. E., Greig, T. W., Hammann, M., Murray, D. C., Murren, C. J., Strand, A. E., Terada, R., Weinberger, F., & Sotka, E. E. (2016). Invasion of novel habitats uncouples haplo-diplontic life cycles. *Molecular Ecology*, *25*(16), 3801–3816. https://doi.org/10.1111/mec.13718

Krueger-Hadfield, S. A., Kollars, N. M., Strand, A. E., Byers, J. E., Shainker, S. J., Terada, R., Greig, T. W., Hammann, M., Murray, D. C., Weinberger, F., & Sotka, E. E. (2017). Genetic identification of

source and likely vector of a widespread marine invader. *Ecology and Evolution*, 7(12), 4432–4447. https://doi.org/10.1002/ece3.3001

Krueger-Hadfield, S. A., & Ryan, W. H. (2020). Influence of nutrients on ploidy-specific performance in an invasive, haplodiplontic red macroalga. *Journal of Phycology*, 56(4), 1114–1120. https://doi.org/10.1111/jpy.13011

Krueger-Hadfield, S. A., Stephens, T. A., Ryan, W. H., & Heiser, S. (2018). Everywhere you look, everywhere you go, there's an estuary invaded by the red seaweed *Gracilaria vermiculophylla* (Ohmi) Papenfuss, 1967. *BioInvasions Records*, 7(4), 343–355. https://doi.org/10.3391/bir.2018.7.4.01

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. https://doi.org/10.1038/nmeth.1923

Le Cam, S., Daguin-Thiébaut, C., Bouchemousse, S., Engelen, A. H., Mieszkowska, N., & Viard, F. (2020). A genome-wide investigation of the worldwide invader *Sargassum muticum* shows high success albeit (almost) no genetic diversity. *Evolutionary Applications*, 13(3), 500–514. https://doi.org/10.1111/eva.12837

Le Corre, V., & Kremer, A. (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, 21(7), 1548–1566. https://doi.org/10.1111/j.1365-294X.2012.05479.x

Lee, C. E. (1999). Rapid and repeated invasions of fresh water by the copepod *Eurytemora affinis*. *Evolution*, 53(5), 1423. https://doi.org/10.2307/2640889

Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, 17(8), 386–391. https://doi.org/10.1016/S0169-5347(02)02554-5

Lees, L. E., Krueger-Hadfield, S. A., Clark, A. J., Duermit, E. A., Sotka, E. E., & Murren, C. J. (2018). Nonnative *Gracilaria vermiculophylla* tetrasporophytes are more difficult to debranch and are less nutritious than gametophytes. *Journal of Phycology*, 54(4), 471–482. https://doi.org/10.1111/jpy.12746

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950), 289–293. https://doi.org/10.1126/science.1181369

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23(9), 2178–2192. https://doi.org/10.1111/mec.12725

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2016). Breaking RAD: An evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, 17(2), 142–152. https://doi.org/10.1111/1755-0998.12596

Maccaferri, M., Sanguineti, M. C., Noli, E., & Tuberosa, R. (2005). Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Molecular Breeding*, 15(3), 271–290. https://doi.org/10.1007/s11032-004-7012-z

Maruyama, T., & Fuerst, P. A. (1984). Population bottlenecks and non-equilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics*, 108(3), 745–763.

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, 210(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Messer, P. W., & Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology & Evolution*, 28(11), 659–669. https://doi.org/10.1016/j.tree.2013.08.003

Navascués, M., Stoeckel, S., & Mariette, S. (2010). Genetic diversity and fitness in small populations of partially asexual, self-incompatible plants. *Heredity*, 104(5), 482–492. https://doi.org/10.1038/hdy.2009.159

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29(1), 1. https://doi.org/10.2307/2407137

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency Eestimation from new-generation sequencing data. *PLoS One*, 7(7), e37558. https://doi.org/10.1371/journal.pone.0037558

Pannell, J. R. (2015). Evolution of the mating system in colonizing plants. *Molecular Ecology*, 24(9), 2018–2037. https://doi.org/10.1111/mec.13087

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), 289–290. https://doi.org/10.1093/bioinformatics/btg412

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, 21(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Parks, J. C., & Werth, C. R. (1993). A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, 80(5), 537–544. https://doi.org/10.1002/j.1537-2197.1993.tb13837.x

Pérez-Figueroa, A., García-Pereira, M. J., Saura, M., Rolán-Alvarez, E., & Caballero, A. (2010). Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, 23(10), 2267–2276. https://doi.org/10.1111/j.1420-9101.2010.02093.x

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52(3), 273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002

Prentis, P. J., Wilson, J. R. U., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science*, 13(6), 288–294. https://doi.org/10.1016/j.tplants.2008.03.004

Pritchard, J. K., & Przeworski, M. (2001). Linkage disequilibrium in humans: models and data. *The American Journal of Human Genetics*, 69(1), 1–14. https://doi.org/10.1086/321275

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., Troll, C. J., Fields, A., Hartley, P. D., Sugnet, C. W., Haussler, D., Rokhsar, D. S., & Green, R. E. (2016). Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. *Genome Research*, 26(3), 342–350. https://doi.org/10.1101/gr.193474.115

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team (2018). *R: A Language and Environment for Statistical Computing*.

Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics*, 105(3), 767–779.

Roman, J., & Darling, J. A. (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution*, 22(9), 454–464. https://doi.org/10.1016/j.tree.2007.07.002

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., & Weller, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics*, 32(1), 305–332. https://doi.org/10.1146/annurev.ecolsys.32.081501.114037

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., Bacher, S., Blasius, B., Brundu, G., Capinha, C., Celesti-Grapow, L., Dawson, W., Dullinger, S., Fuentes, N., Jäger, H., ... Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*(1), 14435. https://doi.org/10.1038/ncomms14435

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702. https://doi.org/10.1534/genetics.113.154138

Slatkin, M. (1994). Linkage disequilibrium in growing and stable populations. *Genetics*, *137*(1), 331–336.

Slatkin, M. (2008). Linkage disequilibrium — Understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*(6), 477–485. https://doi.org/10.1038/nrg2361

Sotka, E. E., Baumgardner, A. W., Bippus, P. M., Destombe, C., Duermit, E. A., Endo, H., Flanagan, B. A., Kamiya, M., Lees, L. E., Murren, C. J., Nakaoka, M., Shainker, S. J., Strand, A. E., Terada, R., Valero, M., Weinberger, F., & Krueger-Hadfield, S. A. (2018). Combining niche shift and population genetic analyses predicts rapid phenotypic evolution during invasion. *Evolutionary Applications*, *11*(5), 781–793. https://doi.org/10.1111/eva.12592

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Stoeckel, S., Porro, B., & Arnaud-Haond, S. (2021). The discernible and hidden effects of clonality on the genotypic and genetic states of populations: Improving our estimation of clonal rates. *Molecular Ecology Resources*, *1755–0998*, 13316. https://doi.org/10.1111/1755-0998.13316

Tajima, F. (1989). The effect of change in population size on DNA polymorphism. *Genetics*, *123*(3), 597–601.

Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S. F., Darde, M. L., & Ayala, F. J. (1991). Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proceedings of the National Academy of Sciences*, *88*(12), 5129–5133. https://doi.org/10.1073/pnas.88.12.5129

Valero, M., Engel, C., Billot, C., Kloareg, B., & Destombe, C. (2001). Concept and issues of population genetics in seaweeds. *Cahiers De Biologie Marine*, *42*, 53–62.

Van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Hodgins, K. A., Gauffre, B., Rieseberg, L. H., van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, *26*, 5421–5434. https://doi.org/10.1111/mec.14293

van Kleunen, M., Dawson, W., & Maurel, N. (2015). Characteristics of successful alien plants. *Molecular Ecology*, *24*(9), 1954–1968. https://doi.org/10.1111/mec.13013

van Kleunen, M., Weber, E., & Fischer, M. (2010). A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters*, *13*(2), 235–245. https://doi.org/10.1111/j.1461-0248.2009.01418.x

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S (Fourth)*. Springer.

Wall, J. D., Andolfatto, P., & Przeworski, M. (2002). Testing models of selection and demography in *Drosophila simulans*. *Genetics*, *162*(1), 203–216.

Wang, T., Su, Y., & Chen, G. (2007). Population genetic variation and structure of the invasive weed *Mikania micrantha* in southern China: Consequences of rapid range expansion. *Journal of Heredity*, *99*(1), 22–33. https://doi.org/10.1093/jhered/esm080

Weinberger, F., Buchholz, B., Karez, R., & Wahl, M. (2008). The invasive red alga *Gracilaria vermiculophylla* in the Baltic Sea: Adaptation to brackish water may compensate for light limitation. *Aquatic Biology*, *3*(3), 251–264. https://doi.org/10.3354/ab00083

Whitlock, M. C., Phillips, P. C., & Wade, M. J. (1993). Gene interaction affects the additive genetic variance in subdevided populations with migration and extinction. *Evolution*, *47*(6), 1758–1769. https://doi.org/10.1111/j.1558-5646.1993.tb01267.x

Williamson, M. (1993). Invaders, weeds and the risk from genetically manipulated organisms. *Experientia*, *49*(3), 219–224. https://doi.org/10.1007/BF01923529

Williamson, M., & Fitter, A. (1996). The varying success of invaders. *Ecology*, *77*(6), 1661–1666. https://doi.org/10.2307/2265769

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.