Conservation Genetics
https://doi.org/10.1007/s10592-022-01452-8

RESEARCH ARTICLE

Global distribution of cryptic native, introduced and hybrid lineages in the widespread estuarine amphipod *Ampithoe valida*

Katherine E. Harper[1] · Lauren A. Scheinberg[2,3] · Katharyn E. Boyer[2] · Erik E. Sotka[1]

Received: 24 June 2021 / Accepted: 18 May 2022

Abstract

Biological invasions can pose a severe threat to coastal ecosystems, but are difficult to track due to inaccurate species identifications and cryptic diversity. Here, we clarified the cryptic diversity and introduction history of the marine amphipod *Ampithoe valida* by sequencing a mtDNA locus from 683 individuals and genotyping 10,295 single-nucleotide polymorphisms (SNPs) for 349 individuals from Japan, North America and Argentina. The species complex consists of three cryptic lineages: two native Pacific and one native Atlantic mitochondrial lineage. It is likely that the complex originated in the North Pacific and dispersed to the north Atlantic via a trans-arctic exchange approximately 3 MYA. Non-native *A. valida in* Argentina have both Atlantic mitochondrial and nuclear genotypes, strongly indicating an introduction from eastern North America. In two eastern Pacific estuaries, San Francisco Bay and Humboldt Bay, California, genetic data indicate human-mediated hybridization of Atlantic and Pacific sources, and possible adaptive introgression of mitochondrial loci, nuclear loci, or both. The San Francisco Bay hybrid population periodically undergoes population outbreaks and profoundly damages eelgrass *Zostera marina* thalli via direct consumption, and these ecological impacts have not been documented elsewhere. We speculate that novel combinations of Atlantic and Pacific lineages could play a role in *A. valida*'s unique ecology in San Francisco Bay. Our results reinforce the notion that we can over-estimate the number of non-native invasions when there is cryptic native structure. Moreover, inference of demographic and evolutionary history from mitochondrial loci may be misleading without simultaneous survey of the nuclear genome.

Keywords Introgressive hybridization · Biological invasions · Cryptic diversity · Mitonuclear discordance · Seagrass · *Zostera marina*

Introduction

The rate of human-mediated introductions into coastal ecosystems has risen dramatically over the past few centuries as a result of increased global trade and transport (Ojaveer et al. 2018). The number of successful invaders to some estuaries ranges from dozens to hundreds of known species (Ruiz et al. 1997, 2011; Bax et al. 2003; Molnar et al. 2008). However, due to a relative lack of systematic, biogeographic, and historical data for marine species relative to terrestrial or freshwater species, many species remain cryptogenic as their status as native or introduced is unknown or is unreliable (Carlton 1996; Geller et al. 2010; McGeoch et al. 2010). Correctly identifying native vs. non-native species is important for understanding their community-level impacts, and is a critical first step in any effort to manage or prevent future invaders (Darling et al. 2017).

The widespread presence of morphologically indistinguishable species complexes makes inferring rates and patterns of dispersal difficult, necessitating the use of molecular methods to accurately quantify cryptic diversity (Knowlton 1993). Cryptic species complexes are especially common among geographically widespread species (Geller

✉ Katherine E. Harper
keharper2@gmail.com

✉ Erik E. Sotka
SotkaE@cofc.edu

[1] Department of Biology, College of Charleston, 29412 Charleston, South Carolina, USA

[2] Estuary & Ocean Science Center, Department of Biology, San Francisco State University, 94920 Tiburon, California, USA

[3] California Academy of Sciences, 94118 San Francisco, California, USA

Published online: 11 June 2022

et al. 2010) and small invertebrates with high rates of speciation (Morais and Reichard 2017). Failure to recognize cryptic diversity may result in under- or overestimating the frequency of invasions. Underestimates arise when an introduced cryptic species is mistaken for a related, morphologically similar native species or lineage, while overestimates occur when a native species that is morphologically identical to a species described elsewhere is incorrectly identified as introduced. Underestimates have occurred in a variety of taxonomic groups including hydrozoans, mollusks, bryozoans, seaweeds, decapods, and arthropods, among others (Morais and Reichard 2017). As the default is to classify species as native, overestimates are rare but have been reported for a number of marine species (Rilov and Crooks 2009). For example, the cryptogenic colonial ascidian, *Botryllus schlosseri*, was thought to be introduced to the northwest Atlantic but a population genetic survey found its lineage was native (Yund et al. 2015).

Introductions of cryptic species or lineages may result in the formation of novel hybrids (Morais and Reichard 2017). Introgression of introduced alleles can result in the loss of native genotypes or complete replacement of a native species or population. Recombination of native and introduced parental genotypes can also result in the creation of novel, adaptive genotypes that increase hybrid fitness and/or lead to hybrid speciation (Lee 2002). At times, such hybrids can

Table 1 Dates *Ampithoe valida* was first collected or seen in the Atlantic and Pacific Ocean

Region	Location	Date first reported	Reference
North-Western Atlantic			
Eastern USA	Great Egg Bay, NJ	1873	Smith 1873
South-Western Atlantic			
Argentina	Puerto Madryn	1980	de Pina 1997
Caribbean Sea			
Venezuela	Isla Garrapata & Playa Esteban	1991–2002	Martin and Diaz 2003
North-Eastern Atlantic			
Portugal	Mondego Estuary	1992	Pardal et al. 2000
Portugal	Ria de Averio	1999	Cunha et al. 1999
France	Balaruc-les-Bains (Bassin de Thau)	2000	Faasse 2015
France	Martigues (Etange de Berre)	2003	Faasse 2015
Netherlands	Port near Vlissingen	2014	Faasse 2015
North-Eastern Pacific			
Korea	Cheju Island	1985	Kim and Kim 1987
Japan/China	various	unclear	see Nagata 1960
North-Western Pacific			
Western USA	Tomales Bay, CA	1941	Carlton 1979
Western USA	San Francisco Bay,CA	1941	Carlton 1979
Western USA	Newport Bay, CA	1942	Barnard 1965; Carlton 1979
Western USA	Coos Bay, OR	1950	Barnard 1954
British Columbia	Haida Gwaii	1957	Sloan and Bartier 2004
Western USA	Morro Bay, CA	1965	Carlton 1979
Western USA	Bodega Harbor, CA	1975	Carlton 1979
Western USA	Elkhorn Slough, CA	1988	Wasson et al. 2001
Western USA	Puget Sound, WA	1998	Cohen 1998
Western USA	Willapa Bay, WA	2000	Cohen et al. 2001
Western USA	Humboldt Bay, CA	2000	Boyd et al. 2002
Western USA	Channel Islands Harbor, CA	2001	Cohen et al. 2002
Western USA	Long Beach Harbor, CA	2001	Fairey et al. 2002
Western USA	Oceanside Harbor, CA	2001	Fairey et al. 2002
Western USA	San Diego Bay, CA	2001	Fairey et al. 2002
British Columbia	Prince Rupert Harbor	2005	Lu et al. 2007
British Columbia	Vancouver Harbor	2005	Lu et al. 2007
Western USA	Yaquina Bay	2006	Pilgrim and Darling 2010

Table 2 Amphipod collection localities for genetic and genomic analyses. We provide the number of individuals with the three mtDNA clades (A, B, and C) as defined here and in Pilgrim and Darling 2010, and sample size genotyped for SNPs. The Notes contains either the collector (ES = E. Sotka, CS = C. Sotka, RT = R. Terada, MV = M. Valinas, KH = K. Harper, LS = L. Scheinberg, HE = H. Endo, SKH = S. Krueger-Hadfield, NK = N. Kollars, JP = J. Pocklington, TK = T. Krueger, BK = B. Krueger, BH = B. Hughes) or GenBank numbers

Site	Region	Site.Code	Latitude	Longitude	mtDNA.A	mtDNA.B	mtDNA.C	SNPs	Notes
Argentina, Playa Bonita Beach	Argentina	ARG	-43.37	-65.05	21	0	0	21	MV
Charleston Harbor, SC	Eastern US	NA	32.75	-79.9	7	0	0	NA	GU048371-GU048377
Charleston Harbor, SC	Eastern US	CHS	32.7513	-79.9	22	0	0	22	KH,LS
Chesapeake Bay, Goodwin Islands (VA)	Eastern US	CBV	37.2167	-76.3833	24	0	0	22	LS
Millstone Point, CT	Eastern US	NA	41.3	-72.17	18	0	0	NA	GU048353-GU048370
Sandwich Marina, MA	Eastern US	SND	41.7703841	-70.503639	18	0	0	18	KH
Allan Harbor, RI	Eastern US	ALN	43	-71.412556	20	0	0	20	KH
Shimabira	Japan	SHI	31.42393	130.16317	0	20	0	20	ES,CS,RT
Kaeda	Japan	KAE	31.82067	131.4481	0	8	0	8	ES,CS,RT
Wajiro	Japan	WAJ	33.673563	130.426959	0	2	5	7	ES,CS,RT,SKH
Osaka: Misaka, mouth of Higashi River	Japan	OSA	34.3166667	135.116667	0	3	17	20	LS
Mangoku-ura	Japan	MNG	37	142	0	0	18	20	SKH,CH,HE,JP
Matsushima Bay	Japan	SOU	38.352839	141.059694	0	1	22	23	SKH,CH,HE,JP,ES
Moune Bay	Japan	MOU	38.9012	142.5	0	1	20	21	ES,HE,JP
Elkhorn Slough, CA	Western US	EHS	36.504474	-121.45131	0	0	20	20	SKH, TK, BK, BH
San Francisco Bay, CA	Western US	NA	37.72	-122.28	4	63	0	NA	GU048411-GU048477
San Francisco Bay, CA	Western US	SFB	37.95	-122.5	0	51	0	51	LS
Tomales Bay, CA	Western US	TMB	38.17	-122.91	0	0	17	17	SKH,NK,LS
Tomales Bay, CA	Western US	NA	38.17	-122.91	0	0	31	NA	GU048478-GU048508
Bodega Bay, CA	Western US	NA	38.31	-123.05	0	0	2	NA	JX545458-JX45459
Humboldt Bay, CA	Western US	NA	40.72	-124.24	32	0	1	NA	GU048378-GU048410
Humboldt Bay, CA	Western US	HUM	40.7333333	-124.21667	19	0	4	23	LS
Coos Bay, OR	Western US	NA	43.41	-124.21	0	0	15	NA	GU048179-GU048193
Yaquina Bay, OR	Western US	NA	44.62	-124.02	0	0	9	NA	GU048170-GU048178
Willapa Bay, WA	Western US	NA	46.54	-123.99	1	0	95	NA	GU048257-GU048352
Willapa Bay, WA	Western US	WIL	46.6666667	-124	0	0	26	16	LS
Grays Harbor, WA	Western US	NA	46.95	-124.04	0	0	7	NA	GU048250-GU048256
Puget Sound, WA	Western US	NA	47.94	-122.53	0	0	56	NA	GU048194-GU048249
Ampithoe caddi	Australia	NA	NA	NA	NA	NA	NA	NA	KP316287
Ampithoe dalli	California, US	NA	NA	NA	NA	NA	NA	NA	KP316288
Ampithoe longimana	Rhode Island, US	NA	NA	NA	NA	NA	NA	NA	KP316293
Ampithoe lacertosa	Japan	NA	NA	NA	NA	NA	NA	NA	KP316294
Ampithoe sectimanus	California, US	NA	NA	NA	NA	NA	NA	NA	KP316301

become invasive as a result of hybrid vigor, as evidenced in San Francisco Bay where the introduced cordgrass, *Spartina alterniflora*, introduced to the area in 1975, extensively hybridized with the native cordgrass, *S. foliosa*, ultimately driving out local genotypes via pollen swamping (Ayres et al. 1999; Daehler and Strong 1997; Strong and Ayres 2013). Molecular data can also be used to identify hybrid invasions and assess the extent of introgression of introduced alleles into native populations.

Ampithoe valida is a benthic, tube-building gammarid amphipod with a cosmopolitan distribution (Table 1). Originally described by Smith (1873) from Long Island Sound on the east coast of the United States, it was later observed elsewhere along this coast and along the west coast of the US, including California (Alderman 1986), Oregon and Washington (Barnard 1954). Its distribution also extends to Japan and Korea and has more recently been recorded to occur in Argentina as well as Europe, specifically France and Portugal (Alonso et al. 1995; Pardal et al. 2000; Chapman 2007; Kim 2011; Faasse 2015). *A. valida* is a grazer, typically feeding on a mixture of macroalgae and epiphytic algae that grows on seagrass, but not the seagrass itself (e.g., Douglass et al. 2011). In contrast, this amphipod directly consumes large quantities of *Zostera marina* (eelgrass) leaves and fruits in San Francisco Bay, in addition to algae (Reynolds et al. 2012; Carr and Boyer 2014; Lewis and Boyer 2014). This switch from a mutualistic partner with eelgrass to a pest manifests as extensive damage to natural beds and interference with eelgrass restoration efforts in San Francisco Bay (Lewis and Boyer 2014).

Previously, Pilgrim and Darling (2010) investigated genetic diversity in *A. valida* from the east and west coasts of the US using the mitochondrial cytochrome *c* oxidase I (COI) gene. Their data supported three distinct evolutionary lineages that have been diverging for almost three million years, suggesting both cryptic species-level diversity and multiple introduction. However, the Pilgrim and Darling (2010) analyses were limited in geographic scope relative to the entirety of the *A. valida* described range. Moreover, the analyses were based on a single genetic marker, mitochondrial COI, which has several limitations: it has less power than nuclear genes to accurately resolve some evolutionary relationships due to its high mutation rate, which can overestimate differences between closely related taxa (Funk and Omland 2003, but see Kemppainen et al. 2009), its maternal mode of inheritance can indicate erroneous divergences if the species exhibits sex-biased dispersal, and differential introgression relative to other markers is impossible to detect and impedes discovery of admixture following introduction events (Dupuis et al. 2012).

Here, we investigated cryptic diversity and inferred historic and recent movement of *A. valida* through the pursuit of three objectives. First, phylogenetic analysis of mtDNA sequence data for samples from Japan, Argentina, and the east and west coasts of the US was used to test for confirmation of the existence of a cryptic species complex within *A. valida*. Second, to construct a full picture of *A. valida's* population genetic structure, combined mitochondrial and nuclear DNA data were used to evaluate whether mitonuclear discordance is present in its cryptic lineages. Lastly, combined mitochondrial and nuclear DNA were also used to determine if introgression among cryptic lineages occurred following introduction events.

Materials and methods

Sampling

A total of 359 individuals of amphipods identified morphologically as *A. valida* from 17 populations across its geographic distribution were collected and/or supplied by colleagues from 2014 to May 2016 (Table 2). Individuals were collected by hand from eelgrass beds and from algae (primarily *Ulva lactuca*, *Gracilaria vermiculophylla* and *Fucus vesiculosus)* on docks, mudflats, and sandy/rocky shores. They were immediately euthanized and stored in 95% ethanol until use.

Extractions, PCR amplification, and mtDNA sequencing

Molecular lab work was conducted at the Center for Comparative Genomics (CCG) at the California Academy of Sciences (CAS) and at Grice Marine Laboratory at the College of Charleston. Genomic DNA was extracted from amphipod tissue using either a Qiagen DNeasy spin-column kit following the manufacturer's protocol for animal tissues or a Macherey-Nagel Nucleospin Tissue kit. For individuals that were less than approximately one centimeter in length or severely degraded, the entire individual was used; otherwise only half of the individual was used in the extraction. An approximately 547 bp fragment of the mitochondrial gene, *cytochrome c oxidase subunit 1* (COI) was amplified using standard PCR conditions and forward and reverse custom made primers from Pilgrim and Darling (2010). PCR products were purified using an Exo-sap it protocol and were either sequenced on an Applied Biosystems 3130 Genetic Analyzer or sent to Eurofins Genomics for sequencing.

mtDNA sequence alignment and phylogenetic analyses

Sequences of our 359 individuals were combined with sequences for 342 individuals sampled by Pilgrim and Darling (2010; GU048170-GU048508; Table 2). To serve as outgroups for phylogenetic analyses, we used previously sequenced samples of *Ampithoe caddi* (KP316287), *A. dalli* (KP316288), *A. longimana* (KP316293), *A. lacertosa* (KP316294) and *A. sectimanus* (KP316301; Sotka et al. 2017). All sequences were aligned in Clustal using default conditions. These sequences were trimmed to 508 basepairs and duplicate sequences were removed using a custom R code and R::ape (Paradis and Schliep 2019). We imported the 47 unique *A. valida* and five outgroup sequences into SeaView (Gouy et al. 2010) to construct a maximum likelihood phylogeny using PhyML (Guindon et al. 2010) with 100 bootstrap replicates and a neighbor joining tree. We then used MrBayes (v 3.2.7a) to generate Bayesian phylogenies (Ronquist et al. 2012) with 100,000 replicates, sampled every 100 times. We confirmed the presence of three mitochondrial clades (i.e., A, B, C) as defined in Pilgrim and Darling (2010).

Reduced-representation genomic library preparation and sequencing

We genotyped single nucleotide polymorphisms (SNPs) using a restriction-site associated genotype-by-sequencing following Parchman et al. (2012) on DNA extractions from 350 individuals. Eleven individuals sampled for mtDNA, but not included in the RADseq library did not have a sufficient amount of DNA extraction product to be included. Genomic DNA was digested with two restriction enzymes, EcoRI and MseI. Pooled fragments for each individual were ligated with customized adaptor sequences containing the Illumina adaptors and primer sequences and unique 8–10 bp barcodes to allow for the in silico identification of individuals. DNA fragments were amplified by PCR twice and size selected (300–450 bp) using a Blue Pippin. The final library was sequenced using a single-end 100 bp protocol within a lane of an Illumina Hiseq-4000 sequencer at the University of Texas at Austin Genome Sequencing and Analysis Facility (GSAF).

PhiX sequences present in the raw reads were identified using a *bowtie2* (Langmead and Salzberg 2012) assembler and removed (13% was phiX). Cut sites and adaptors were removed from the remaining reads (256,667,215) and then parsed using barcodes into a single FASTQ file per individual using custom *Perl* scripts. A subset of 15 million raw reads was used to create a de novo assembly using *vsearch* (Rognes et al. 2016). Contigs were assembled with a minimum match percent set to 92% and pruned with a minimum sequence length of 50 bp. To remove potential paralogs, consensus sequences obtained for each contig were assembled de novo to each other with a minimum match percent set to 83%. Parsed reads were then assembled, with a maximum edit distance set to 6 bp, to the pruned contigs coming out of the de novo assembly using *bwa* (Li and Durbin 2009). These alignments were then used to call variant SNP sites with *samtools / bcftools* version 0.1.19 (Li et al. 2009). We analyzed a single SNP per contig to reduce the effects of physical linkage on population genetic parameters and filtered out SNPs that were recorded in fewer than 50% of individuals, had greater than two alleles per individual and possessed a minor allele frequency < 5%. We generated a set of genotype likelihoods that combine the uncertainty generated by sequence coverage, sequencing error, and alignment error. We converted the phred-scale genotype likelihoods (from samtools/bcftools) per SNP-sample combination into probabilities that summed to 1, and then converted these to a single value that ranges from 0 to 2, where 0, 1 and 2 represent the highest probability of a homozygote, heterozygote, and alternative homozygote, respectively.

Population genomic analyses

Admixture analyses were performed with the final, filtered set of 10,295 SNPs using the program Entropy (Gompert et al. 2014). Entropy uses a Bayesian clustering algorithm that incorporates uncertainty in sequencing coverage and uses allele frequencies as a prior. Genotype estimates were used to create starting values for the Bayesian clustering algorithm in Entropy. Entropy runs (five independent chains) were each initiated with a discriminant analysis of principal components function (DAPC; Jombart et al. 2010) and a k-means clustering for numbers of putative clusters (k = 2–8) based on the genotype estimates. The optimal number of clusters was determined by comparing the deviance information criteria (DIC) for each number of clusters. From Entropy, genotype probabilities and admixture proportions were extracted for each respective cluster, k = 2–8. Bar plots of the admixture proportions for all k clusters were created in R (R Core Team 2018).

To visualize geographic structure, posterior genotype probabilities were extracted for all k clusters and principal component analyses performed for each using the *prcomp* function in R.

To provide a view of hybridization across loci, we filtered loci that were highly divergent between Pacific (all Japanese plus Willapa Bay, Tomales Bay and Elkhorn Slough) and Atlantic populations and all populations, including those that appeared using PCA to be composed of hybrids (i.e., Humboldt and San Francisco Bay). We generated genotypes

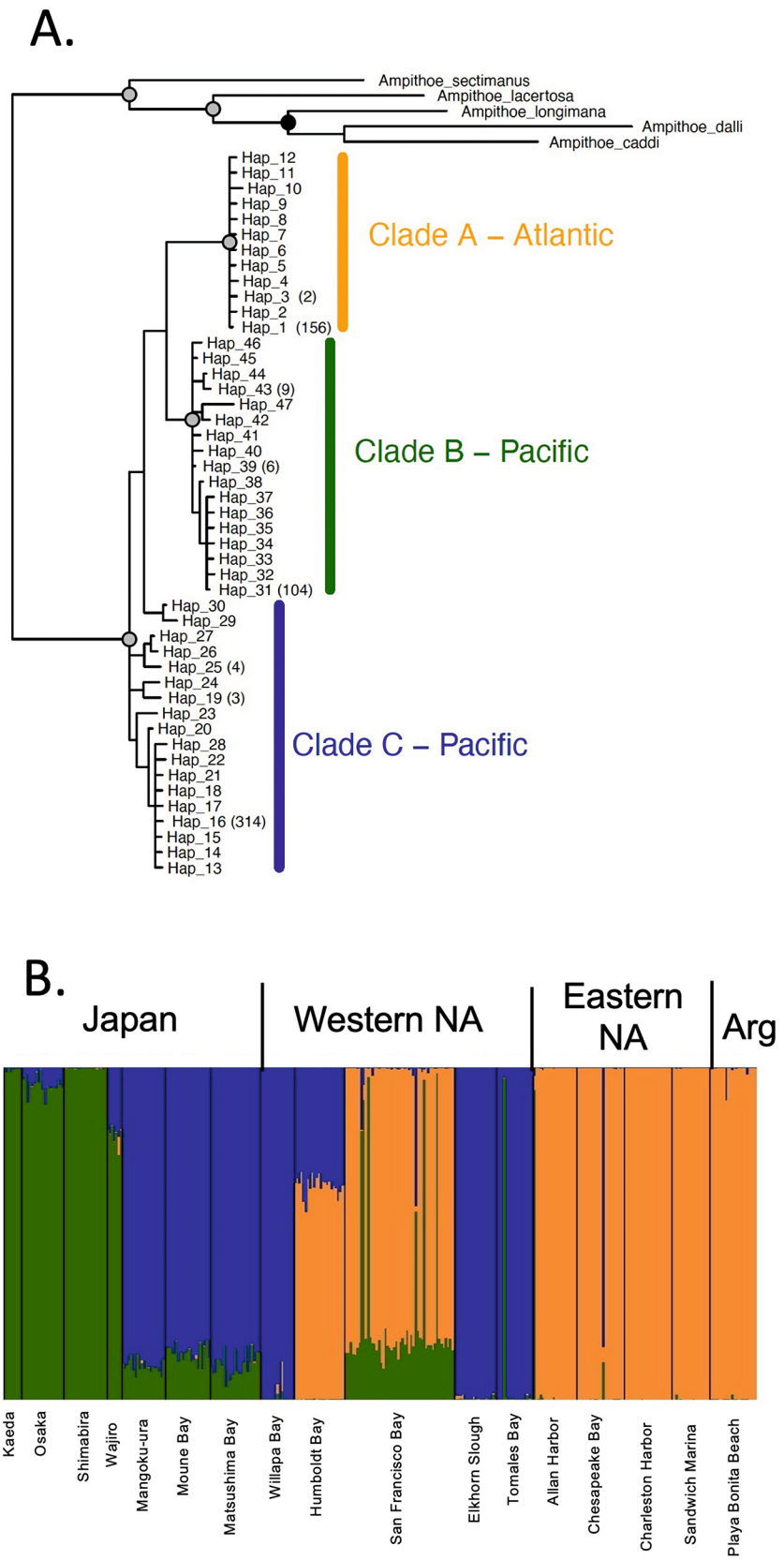

Fig. 1 A) Bayesian phylogeny based on *Ampithoe* species COI haplotypes. Small grey circles indicate nodes with 70% maximum likelihood bootstrap support and black circles indicate 95% posterior probability from Bayesian search. Clade A (orange), B (green) and C (blue) were defined by Pilgrim and Darling 2010. Numbers after each haplotype indicate the number of repeated haplotypes sequenced (B) Admixture proportions based on three putative clusters using SNP genotype likelihoods (k = 3). Each bar represents one individual and is separated into colored segments that represent that individual's likely proportion of membership to a cluster. Black lines separate individuals of different populations. Populations are labeled below the figure with their regional affiliations above (Arg = Argentina)

by rounding the genotype likelihoods for each SNP by individual combination to 0, 1 or 2, estimated mean allele frequency across all Pacific individuals (Japan and western North America) and all eastern Atlantic individuals, and then filtered for those loci that differed in allele frequencies by > 95% (e.g., pPac = 0.99 versus pAtl = 0.03 yields 0.96). These genotypes were then used to construct a heat map showing individual- and locus-specific genotypes. Each block represents the estimated genotype for an individual at a particular locus, with red blocks representing an individual who is homozygous for alleles confined to the Pacific at particular loci, yellow blocks representing individuals who are homozygous for alleles confined to the Atlantic at particular loci, and orange blocks representing individuals who are heterozygous at particular loci. Metadata, genotype likelihoods, the reference denovo scaffolds, R code for the hybrid plot, and a fasta-formatted mtDNA alignment can be found at https://github.com/esotka/AmpithoeValida. Novel mtDNA sequences (Accession ON148466-ON148512) and FASTQ files (BioProject PRJNA825556) have been uploaded to GenBank.

Results

Phylogeography of mitochondrial COI haplotypes

Phylogenetic analyses of 47 unique *A. valida* sequences confirmed that *A. valida* is a monophyletic species comprised of three deep mtDNA lineages (support of monophyly 100% for both maximum likelihood bootstrap and posterior probability; Fig. 1 A). We follow the convention of Pilgrim and Darling (2010) and label the Atlantic lineage as COI-A, and the Pacific lineages as COI-B and COI-C. While Bayesian (Fig. 1 A) and maximum likelihood (not shown) phylogenies indicated that the COI-C clade is basal to the species, these were not well supported statistically. Moreover, while clades A and B appear to be sister taxa, their statistical support was not strong. Of the 660 *A. valida* individuals, we detected 39 unique haplotypes and eight (8) that occurred in more than one individual. Three (3) haplotypes were found in 87% of individuals: Haplotype 1 (Clade A; 156 individuals; e.g., GenBank Accession # GU048370), Haplotype 31 (Clade B; 104 individuals; e.g., GU048414) and Haplotype 16 (Clade C; 314 individuals; e.g., GU048199). The pairwise genetic distance between these three *A. valida* haplotypes was approximately 4.8% (A vs. B), 5.4% (A vs. C) and 4.3% (B vs. C).

COI-C haplotypes were abundant in northern Japan and the west coast of the US (Fig. 2). COI-C was found at 85%, 95%, 71%, 96%, and 100% of individuals from Osaka, Moune Bay, Wajiro, Matsushima Bay, and Mangoku-ura, while all individuals from Tomales Bay and Willapa Bay, and 3% of Humboldt Bay, had COI-C. COI-B was found only in Japan and in a single locality, San Francisco Bay, on the west coast of the US. It comprised 100% of the individuals from Kaeda and Shimabira, 97% of the individuals from San Francisco Bay, and 5%, 5%, 29%, and 4% of individuals in Osaka, Moune Bay, Wajiro, and Matsushima Bay, respectively. COI-A dominated the east coast of the US and a single population in Argentina (Fig. 2). Additionally, the COI-A clade composed 97% of the individuals in Humboldt Bay, but only 3% of those in San Francisco Bay.

Geographic structure of nuclear SNPs

The number of reads linked to 348 individuals after removing phiX sequences and quality filtering was 25.6 million. We removed one Allan Harbor individual from further analyses as there were zero reads. The number of reads per individual averaged 733,304 (median: 768,742; range: 0 to 1,894,020 reads per individual. After selecting a single SNP per contig, a total of 20,745 variable sites were detected; 10,295 sites had a minor allele frequency > 5% and were used for all analyses. Mean number of reads per SNP per individual was 2.2.

The set of 10,295 SNPs was analyzed for genetic admixture using *Entropy* and putative genetic clusters ranging from k = 2 to 8. An admixture model of K = 3 (DIC = 13539170.9) best explained the genetic variation among the set of SNPs (Fig. 1B), and these splits were generally reflected within runs with greater k (i.e., k = 4–8; Figures S1 and S2). For most populations, the geographic distribution of the three SNP clusters broadly mirrored the distribution of the three mtDNA haplotypes (Fig. 2). In Japan, one cluster dominated southern Japan (coincident with dominance by COI-B), a second cluster dominated northern Japan (coincident with dominance by COI-C) and a hybrid zone appeared between these two regions. Along the east coast of the US and Argentina, a third cluster dominated nearly all individuals (coincident with COI-C). Along the west coast of the US, some populations (Willapa Bay, Tomales Bay, and Elkhorn Slough) were dominated by the same cluster and mitochondrial marker (COI-C) dominating northern Japan with very few exceptions.

Two populations (Humboldt Bay and San Francisco Bay) consisted of individuals with mixed genomic ancestry. Individuals in Humboldt Bay were composed of 64% of the genomic cluster that dominates the eastern US and 36% of the genomic cluster that dominates northern Japan and other western US populations. Individuals in San Francisco Bay were composed of 75% of the genetic cluster in the eastern US and 25% with the cluster that dominates southern Japan.

Moreover, within some populations, a few individuals were clearly distinct from local population ancestry. For example, seven individuals from San Francisco Bay clustered with southern Japan. Another individual from San Francisco Bay clustered approximately 50/50 with the two genomic clusters occurring in the Pacific. Two individuals in the Atlantic, one from Allan Harbor and one from Chesapeake Bay, strongly clustered with southern Japan and northern Japan/western US populations, respectively.

Principal component analysis of nuclear SNP data largely confirmed these same patterns. For most populations, the three SNP clusters corresponded almost completely with the three mtDNA haplotype clades (Fig. 3 A). Principal component axis 1, which explained 55% of the variation, separated all eastern US populations and Argentina that are dominated by the COI-A clade from all Japanese and 'non-mixed' west coast US populations that are predominantly composed of either the COI-B or COI-C clade. The only exceptions were the previously mentioned individuals from Allan Harbor and Chesapeake Bay that grouped with COI-C and COI-B individuals, respectively. Principal component axis 2 explained 11% of the variation and separated northern Japan populations that were composed mostly of the COI-C clade from southern Japan populations predominantly composed of the COI-B clade. Major exceptions include some individuals in Osaka and Wajiro that fall into the COI-C clade, but cluster with individuals from southern Japan that fall into the COI-B clade.

Similar to the admixture analysis, PCA revealed that most individuals from Humboldt and San Francisco Bays are intermediate between Atlantic and Pacific regions (Fig. 3 A). Individuals from Humboldt Bay lie almost equidistant between the two groups, while San Francisco Bay individuals lie much closer to individuals from the Atlantic yet do not cluster tightly with them. The seven individuals that clustered with southern Japan under the admixture model also cluster with individuals from southern Japan in the PCA. The one individual that grouped equally with the Pacific clusters lies directly between the two Pacific clusters in PCA space as well validating it as the only clear hybrid between two different Pacific sources. The intermediate placement of most of the individuals from these populations, however, suggests that hybridization between Pacific and Atlantic sources has also occurred.

Another view of the hybridization between Atlantic and Pacific genes in Humboldt and San Francisco Bays was revealed when focusing on 335 nearly-fixed loci (i.e., loci with at least a 95% allele frequency difference) between the east coast of the US and populations in the Pacific (i.e., all Japanese plus Willapa Bay, Tomales Bay and Elkhorn Slough). San Francisco Bay had a highly Atlantic nuclear background, but had the COI-B allele from the Pacific. In contrast, most individuals from Humboldt Bay had a mixed Atlantic/Pacific background while having the COI-C allele that also dominates the Pacific. Argentina was dominated by both Atlantic mitochondrial and nuclear alleles.

To better understand structure within the eastern US, principal component analysis on the entire set of nuclear SNPs was performed for individuals native to the eastern US; i.e., all individuals sampled from along the east US coast with the exception of the two outliers from Allan Harbor and Chesapeake Bay, which were likely recently introduced (N = 80; Fig. 4 A). PC1 delineates most individuals in SC (Charleston) from other eastern US individuals (RI, MA and VA), while PC2 splits the northern populations (RI and MA) from VA.

An analysis of eastern US and Argentine individuals (Fig. 4B) showed no clear pattern, suggesting that Argentine invasion may have originated from multiple eastern US populations. Principal component analysis of 53 individuals with the COI-C haplotype among the native western US populations (i.e., Elkhorn Slough, Tomales Bay, and Willapa Bay; Fig. 4 C) indicated a major split between Elkhorn Slough individuals and the other California populations. An analysis of COI-C lineages from western US and Japanese populations (Fig. 4D) indicates a combination of PC1 and PC2 splits the Japanese and western US populations, and substantial differentiation among populations within these two regions.

Discussion

We confirm Pilgrim and Darling 2010's conclusion that *A. valida* is composed of three cryptic lineages using both COI sequencing and a novel SNP dataset. Our dataset also reveals the source of multiple *A. valida* introductions worldwide, and characterizes the hybridization dynamics of cryptic lineages in California as a consequence of some introductions (see summary in Fig. 5). We treat each of these findings below in turn.

Cryptic species complex

A. valida was originally described in 1873 from the eastern US in Long Island Sound in New Jersey (Smith 1873). Subsequent scholars then described *A. valida* as introduced to the western US, Japan, Europe and South America (Table 1). However, contrary to these interpretations, there are several reasons to believe that *A. valida* species complex arose initially in the North Pacific, and then expanded into the Atlantic. First, the COI-C clade, which occurs exclusively in the north Atlantic, does not appear to be basal in the COI phylogeny, although statistical support for these nodes is not

high. Second, a molecular phylogeny of the family Ampithoidae indicates that *A. valida* resides in a North Pacific clade of *Ampithoe* species (Sotka et al. 2017). Finally, using a molecular clock for crustacean COI sequences of 1.4% per million years, Pilgrim and Darling (2010) suggest that *A. valida* participated in a trans-arctic exchange from the Pacific to the north Atlantic over 3 million years ago. This is consistent with a dispersal event after the opening of the Bering Strait sometime during the Pliocene era, along with a series of other Pacific species that dispersed via a trans-Arctic route Orti et al. 1994; Palumbi and Kessing 1991; Reid et al. 1996; van Oppen et al. 1995; see review by Laakkonen et al. 2021). For example, the opening of the Bering Strait allowed for the initial colonization of blue mussels from the Pacific to the Atlantic Ocean. Repeated periods of glaciation eventually led to the allopatric divergence of *Mytilus edilus* in the Atlantic and *Mytilus trossulus* in the Pacific (Rawson and Harper 2009).

Since arriving to the western North Atlantic, *A. valida* has extended its range from Maine to as far south as Cape Canaveral, Florida (Bousfield 1973; Fox and Bynum 1975). The most southern population sampled from the eastern US in this study (Charleston Harbor) was found to be genetically differentiated from both the Chesapeake Bay and the two New England populations. This genetic split between northern and southern populations along the eastern US coast is consistent with those patterns found for the closely related amphipod *Ampithoe longimana* (Sotka et al. 2003), and a suite of other fish and invertebrates (Wares 2002). It is likely that the Last Glacial Maximum (~ 20 K years ago) forced some northern populations to retreat southward for refuge while others were able to achieve refuge in northern periglacial areas. As a result, periglacial and southern refugia populations genetically diverged from one another over time (Maggs et al. 2008).

Evidence of human-mediated introductions

There is evidence of multiple human-mediated introductions of *A. valida* lineages. For example, the Argentine *A. valida* represents an introduction from the eastern US. Strong mitonuclear concordance in Argentina and the fact that *A. valida* was not discovered there until 1980 (Table 1) suggests that the invasion was recent. The source for the Argentine invasion within the eastern US coast remains unclear, however.

A. valida has recently been discovered in Europe, including Portugal, France, and the Netherlands (Cunha et al.1999; Faase 2015; Pardal et al. 2001), and has also been reported from Korean waters (Kim and Kim 1987). As an aside, we found three South Korean samples on GenBank (JX502956-JX502958) that are within the COI-C clade (analyses not shown). Further sampling of these areas is needed to determine whether lineages are cryptic native or introduced populations.

More complicated introduction histories describe those of California bays of San Francisco and Humboldt. Pilgrim and Darling (2010) posed two hypotheses for the invasion history of *A. valida* in California. First, the Pacific clade COI-C was introduced to San Francisco Bay and spread northward, and then the Atlantic clade COI-A and Pacific clade COI-B replaced COI-C in Humboldt and San Francisco Bay, respectively. Alternatively, they suggested that Pacific clade COI-B was introduced to San Francisco Bay followed by separate cryptic introductions of Atlantic clade COI-A and Pacific clade COI-C to Humboldt Bay and the rest of the west Pacific, respectively.

After our more extensive sampling in the Pacific and sequencing of the nuclear genome, we reject both hypotheses, and propose a more parsimonious scenario. The San Francisco Bay populations are largely a hybrid mix of both southern Japan and eastern US individuals. The placement of San Francisco Bay individuals in PCA space (Fig. 3) suggests the majority of them are genetically similar to individuals in the Atlantic with the exception of five individuals that grouped with individuals in southern Japan and one outlier between southern and northern Japan. The same pattern is seen under the admixture model (Fig. 2 C) with all individuals except six clustering with most of the individuals from the Atlantic. A view of 335 nearly fixed loci between the Atlantic and the Pacific indicates that San Francisco Bay has a highly Atlantic (eastern US) nuclear background, suggesting there has been strong introgression of Atlantic alleles into San Francisco Bay. The Pacific COI-B haplotype dominates San Francisco Bay and remains absent from any other west Pacific populations likely because it was introduced from southern Japan, where it appears to be native.

The Humboldt Bay population is dominated by COI-A, which remains absent everywhere else except the Atlantic. This suggests Humboldt Bay was invaded by an Atlantic COI-A that survives intact. However, the genomes of Humboldt Bay individuals are genetically intermediate between Pacific and Atlantic populations. Analysis of the 335 nearly fixed loci show Humboldt Bay individuals to have an equal mixture of both Atlantic and Pacific genotypes indicating introgression of Atlantic alleles into a Pacific nuclear background, or vice versa. Humboldt Bay exhibits a high proportion of heterozygotes, suggesting either that hybridization between individuals from the Atlantic and the Pacific is recent and ongoing with little to no backcrossing or selection in favor of heterozygosity.

Finally, there is evidence for a rare frequency of introductions to the eastern US. One individual from Chesapeake Bay clusters with northern Pacific populations (Japan and

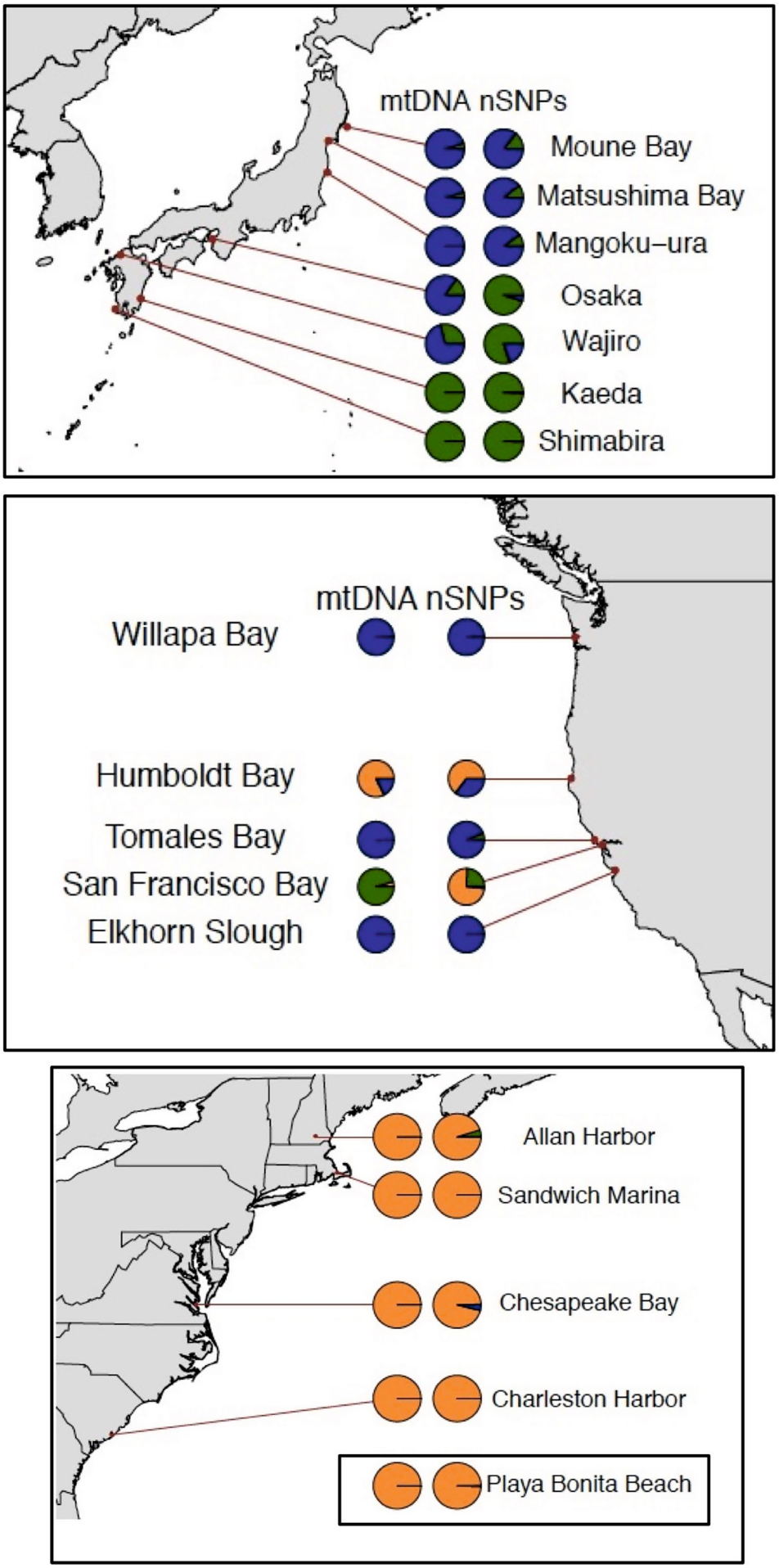

Fig. 2 Mitochondrial COI vs. nuclear SNPs. Pie charts based on mtDNA represent clade frequencies (Orange = Clade A, Green = Clade B, Blue = Clade C) for each population; pie charts based on nuclear SNP data represent the average admixture proportions based on three putative clusters (k = 3) for each population

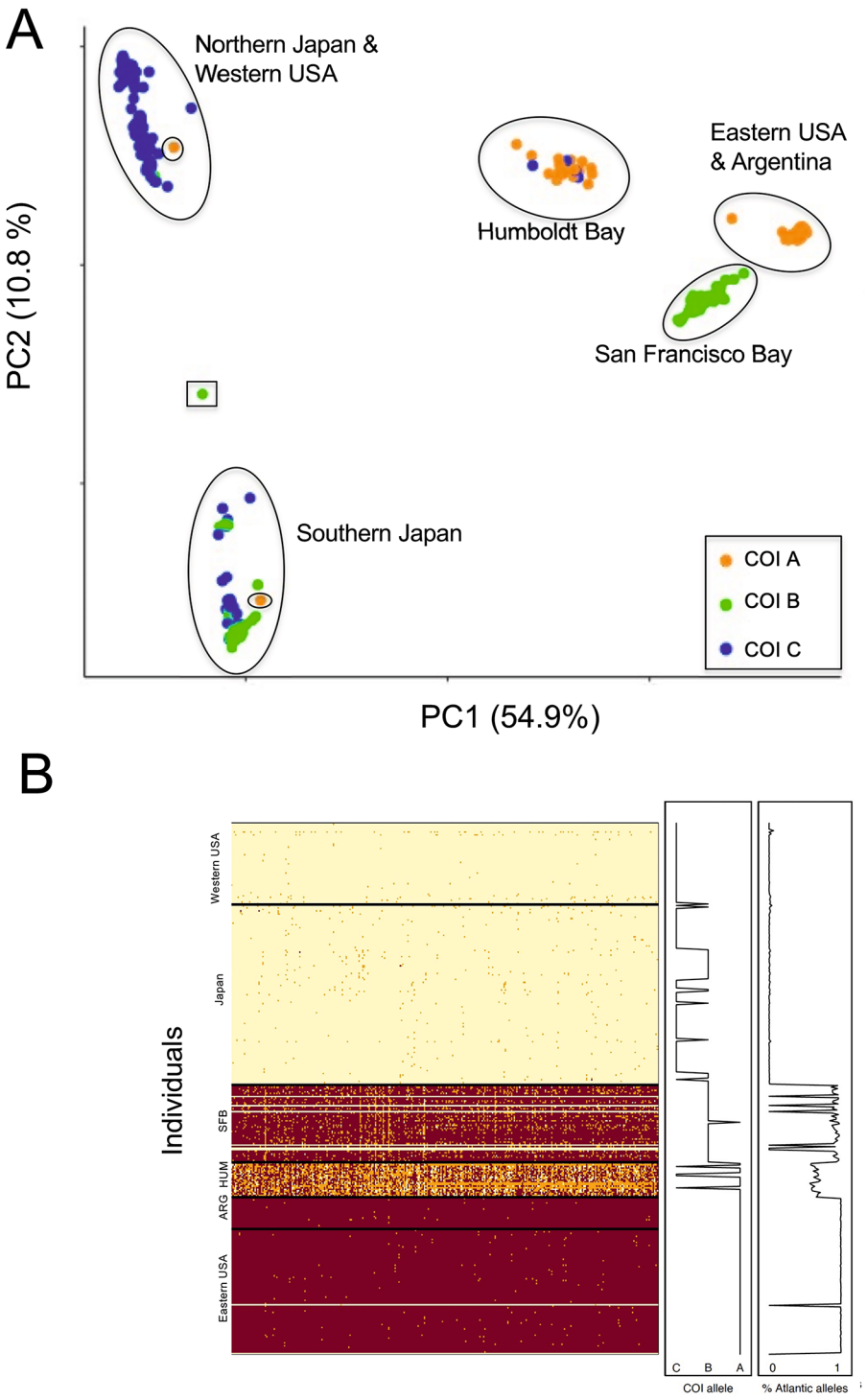

Fig. 3 (A) Principal component analysis (PC1 against PC2) of SNPs. Individuals are colored by their mitochondrial clade assignment (Orange = Clade A, Green = Clade B, Blue = Clade C). Circles of a single individual represent outliers from the Eastern USA and a square of a single individual represents an outlier from San Francisco Bay (SFB). (B) Genotypes of individuals for a subset of 335 loci that are 'fixed' between US Atlantic populations and Pacific (US and Japan) populations. Red blocks represent homozygotes for Pacific alleles, yellow blocks represents homozygotes for Atlantic alleles, and orange blocks represent heterozygotes. The first plot to the right indicates the COI allele assignment for each individual and the second plot indicates the proportion of each individual's genome that is of Atlantic origin. Note the strongest cytonuclear mismatch are those genotypes in SFB that have predominantly Atlantic nuclear background (red) coupled with Pacific mitochondrial clades (COI-B)

western US) while an individual from Allan Harbor clusters with southern Japan. Both individuals have a high proportion of Pacific genotypes suggesting that these introductions are likely very recent.

Evidence for human-mediated hybridization

When previously isolated lineages or species come into contact, a hybrid zone may form (Allendorf et al. 2001). Hybrid vigor in resulting offspring may occur depending on the level of genetic/genomic introgression between

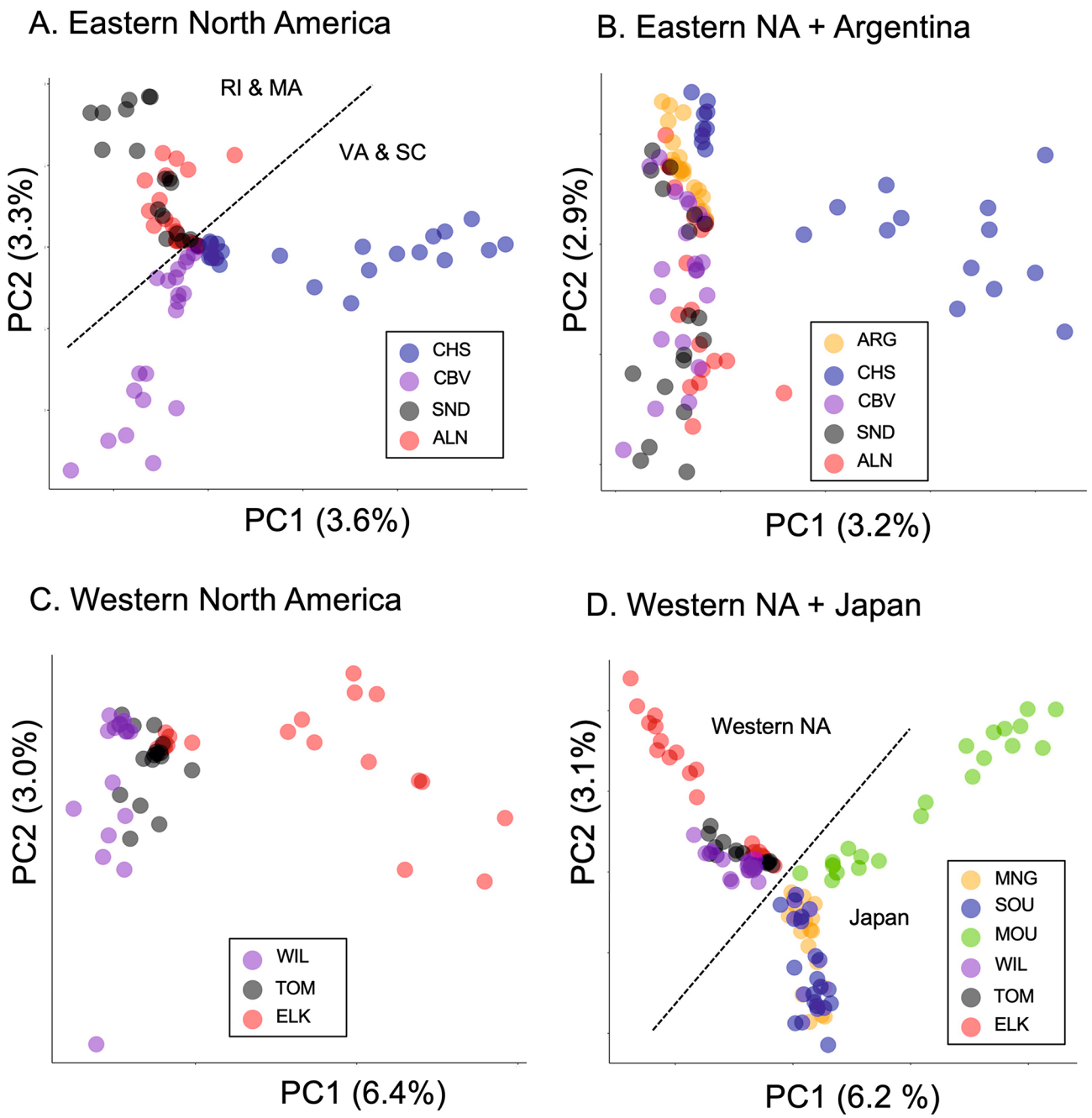

Fig. 4 Principal component analysis based on a K = 3 for individuals (A) within the native eastern US, (B) within the East Atlantic, (C) within the native western US and (D) within the native North Pacific. Colors represent the following populations-for (A & B) red-Allan Harbor, black-Sandwich Marina, blue-Charleston Harbor, purple-Chesapeake Bay; orange-Playa Bonita Beach, Argentina; for (C & D) red-Elkhorn Slough, black-Tomales Bay, purple-Willapa Bay; orange-Mangoku-ura, green-Moune Bay, blue-Matsushima Bay

lineages and the relative adaptive advantage of their different alleles (Arnold 1997; Lee 2002; Mallet 2005). For example, the native California Tiger Salamander and introduced Barred Salamander have recently come back into contact to produce larger, earlier reproducing individuals as a result of adaptive introgression of some introduced alleles (Fitzpatrick and Shaffer 2004; Ryan et al. 2009; Fitzpatrick et al. 2010). In San Francisco Bay, *A. valida* has been found at very high densities on *Zostera marina* with up to a mean of 380 individuals per vegetative shoot and 650 per flowering shoot, or over 20,000 individuals per m^2 (Ayala 2021). It consumes large amounts of eelgrass leaves and fruits even

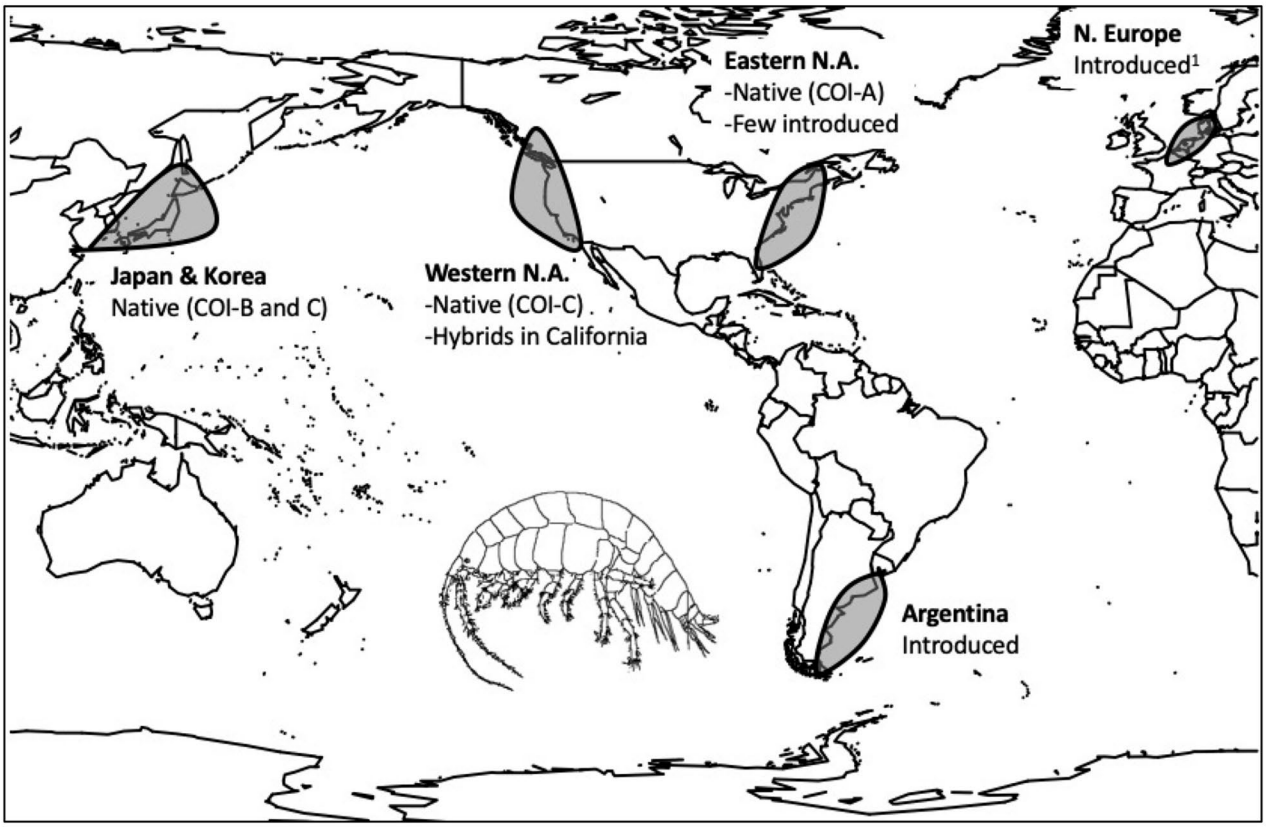

Fig. 5 A summary of our current understanding of cryptic native, introduced and hybrid lineages of *Ampithoe valida* from this study, Pilgrim and Darling (2010) and Faasse (2015)

when epiphytes on eelgrass leaves or palatable macroalgae are plentiful (Reynolds et al. 2012; Carr and Boyer 2014; Lewis and Boyer 2014). *A. valida* outbreaks and damaging levels of direct eelgrass consumption have not been reported in this species before nor, to our knowledge, at any other locality worldwide, suggesting a recent phenomenon unique to San Francisco Bay. We speculate that novel combinations of Atlantic nuclear genotypes and Pacific COI haplotypes and adaptive genomic introgression could play a role in *A. valida's* unusual feeding behavior and perhaps reproduction in San Francisco Bay.

Vectors

A. valida is a direct developer, and a single female can carry approximately 10–20 individuals (Sotka, personal observation). Thus, a relatively small number of pregnant mothers may found new populations. Possible vectors for *A. valida's* recent (i.e., human-mediated) introductions include ship fouling and ballast water, bait boxes, and shipments of oysters, *Spartina* salt marsh plants and seaweeds. Cohen and Carlton (1995) initially proposed ship fouling and ballast water as a possible vector for introductions of *A. valida* from the Atlantic to the Pacific coast of the US, although *A. valida* has now been confirmed native to the Pacific as well. However, ship fouling and ballast water are common vectors of marine introductions and could be responsible for any of *A. valida's* recent introduction events including those back to the Pacific from the Atlantic. The live bait trade is also a potential mechanism for *A. valida's* introductions. The Maine marine baitworm trade, one of the largest suppliers of marine bait for recreational fishing, was found to have a high diversity and abundance of marine organisms including *A. valida* (Fowler et al. 2016). Shipments are mostly made along the east coast of the US, but are also sent to the US west coast and Europe including Italy, France, and Spain, and could account for recent introductions from the eastern US to Europe.

A. valida could have also hitchhiked on other marine organisms that have been moved around by human means. The Pacific or Japanese oyster, *Crassostrea gigas*, was introduced from Japan and Korea to the western US in 1902 (Kincaid 1968; Andrews 1980; Chew 1990) and reintroduced back to Japan in 1980 (Chew 1990). It was then

introduced from the western US to Portugal in 1977 (Chew 1990). The eastern oyster, *Crassostrea virginica*, was also introduced from the east coast of the US to California (Barrett 1963), to Japan in 1968 (Chiba et al. 1989) and to France from 1861 to 1875 (Carlton and Man 1996). While oysters are the most plausible explanation for movement between Japan and the western US, shipments of *Spartina* are a more likely vector for introductions from the eastern US into Europe and Argentina. Seeds of *Spartina alterniflora* were initially introduced from the northeast US to England via ship ballast water in the early nineteenth century, and to France sometime before 1960 (Goulletquer et al. 2002). *Spartina versicolor* was also introduced to France from the northwest Atlantic, specifically into Arachon Bay in 1901 (Goulletquer et al. 2002). As for Argentina, evidence indicates that *Spartina alterniflora* may have been introduced from the eastern US in the 18th or early 19th century by human activity (Bortolus et al. 2015).

Conclusion

The invasion histories of marine species can be misinterpreted when obscured by complexes of geographically separated cryptic species (Rilov and Crooks 2009). A single recognized species might turn out to actually be several separate species or subspecies, each with their own native and introduced ranges. This can lead to an underestimation of the number of introduction events, as in the cases of the decapod *Carcinus maenas* (Roman 2006) and the common reed *Phragmites australis* (Saltonstall 2002). However, in other cases, failure to recognize distinct lineages may result in the misidentification of native lineages, resulting in an overestimation of the number of invasions (e.g., the alpheid shrimp, *Automate branchialis* (Galil et al. 2002), and the octocoral, *Carijoa riisea* (Kahng and Grigg 2005). Similarly, *A. valida* was originally thought to be native to the Atlantic Coast of North America and introduced to all other regions. However, *A. valida* is a cryptic species complex native to the northwest Atlantic, and the eastern and western Pacific. Thus, it is likely that the number of *A. valida* invasions has been overestimated.

The present-day distribution of *A. valida* reflects both natural and recent human-mediated movement and hybridization dynamics in recently invaded habitats. As such, we confirm advice from others (e.g., Viard et al. 2016) that even though mitochondrial loci were useful in identifying within species-level lineages, nuclear SNPs were required to detect hybridization and reconstruct invasion histories.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10592-022-01452-8.

Acknowledgements This research was supported by the National Science Foundation (OCE-1357386 and OCE-1924599). The authors thank those listed in Table 2 for help with sample collection, and especially Stacy Krueger-Hadfield, Macarena Valinas, Glauco Machado, Ben Flanagan. We thank Chris Nice and Katie Bell for facilitating the RADseq analysis and providing barcoding primers, and Jonathon Stillman, Anna Sellas, Allan Strand, Gavin Naylor, and Amy Fowler for thoughtful advice and feedback.

References

Alderman AL (1986) Some new or little known amphipods of California. Univ Calif Publ Zool 41:53–74

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids:Setting conservation guidelines. Trends Ecol Evol 16:613–622

Alonso G, Tablado A, Lopez Gappa J, Magaldi N (1995) Seasonal changes in an intertidal population of the amphipod *Ampithoe valida* Smith, 1873. Oebalia 21: 77–91

Andrews JD (1980) A review of introduction of exotic oysters and biological planning for new importations. Mar Fish Rev 42:1–11

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press

Ayala G (2021) Eelgrass (*Zostera marina*) invertebrate community recovery following an extreme low salinity event. Master's thesis, San Francisco State University, 66 pp

Ayres DR, Garcia-Rossi D, Davis HG, Strong DR (1999) Extent and degree of hybridization between exotic (*Spartina alterniflora*) and native (*S. foliosa*) cordgrass (Poaceae) in California, USA determined by random amplified polymorphic DNA (RAPDs). Mol Ecol 8(7):1179–1186

Barnard JL (1954) Marine Amphipoda of Oregon. Oregon State Monographs. Stud Zool 8:1–103

Barnard JL (1965) Marine amphipoda of the family Ampithoidae from southern California. Proceedings of the United States National Museum

Barrett EM (1963) The California oyster industry. Resources Agency of California, Department of Fish and Game

Bax N, Williamson A, Aguero M, Gonzalez E, Geeves W (2003) Marine invasive alien species: a threat to global biodiversity. Mar Policy 27:313–323

Bortolus A, Carlton JT, Schwindt E (2015) Reimagining South American coasts: unveiling the hidden invasion history of an iconic ecological engineer. Divers Distrib 21(11):1267–1283

Bousfield EL (1973) Shallow-water gammaridean amphipoda of New England. Comstock Publishing Associates

Boyd MJ, Mulligan TJ, Shaughnessy FJ (2002) Non-indigenous marine species of Humboldt Bay, California

Carlton JT (1979) History, biogeography, and ecology of the introduced marine and estuarine invertebrates of the Pacific Coast of North America. Dissertation, University of California Davis

Carlton JT (1996) Biological invasions and cryptogenic species. Ecology 77(6):1653–1655

Carlton JT, Mann R (1996) Transfers and world-wide introductions in the eastern oyster *Crassostrea virginica*. Maryland Sea Grant, College Park, Maryland, pp 691–706

Carr AL, Boyer KE (2014) Variation at multiple trophic levels mediates a novel seagrass-grazer interaction. Mar Ecol Prog Ser 508:117–128

Chapman JW (2007) Gammaridea. In: Carlton J (ed) The Light and Smith Manual: Intertidal Invertebrates from Central California to Oregon. Berkely:University of California Press, pp 545–610

Chew KK (1990) Global bivalve shellfish introductions. World Aquacult 21:9–21

Chiba KEN, Taki Y, Sakai K, Oozeki Y (1989) Present status of aquatic organisms introduced into Japan. Exotic aquatic organisms in Asia, vol 3. Asian Fisheries Society Special Publication, pp 63–70

Cohen AN (1998) A Rapid Assessment Survey of Non-indigenous Species in the Shallow Waters of Puget Sound: Report of the Puget Sound Expedition, September 8–16, 1998. Nearshore Habitat Program, Aquatic Resources Division, Washington State Department of Natural Resource, pp. 1–37

Cohen AN, Berry HD, Mills CE, Milne D, Britton-Simmons K, Wonham MJ, Secord DL, Barkas JA, Bingham B, Bookheim BE, Byers JE (2001) Washington state exotics expedition 2000: a rapid survey of exotic species in the shallow waters of Elliott Bay, Totten and Eld Inlets, and Willapa Bay. Washington State Department of Natural Resources, Olympia WA 46:54

Cohen AN, Harris LH, Bingham BL, Carlton JT, Chapman JW, Lambert CC, Lambert G, Ljubenkov JC, Murray SN, Rao LC, Reardon K (2002) A rapid assessment survey of exotic species in sheltered coastal waters. Sacramento, California. Project report for the Southern California exotics expedition 2000. California Department of Fish and Game/State Water Resources Control Board/National Fish and Wildlife Foundation, pp 1–23

Cohen AN, Carlton JT (1995) Nonindigenous aquatic species in a United States estuary: A case study of the biological invasions of the San Francisco Bay and Delta. U.S. Fish and Wildlife Service, Washington DC

Cunha MR, Sorbe JC, Moreira MH (1999) Spatial and seasonal changes of brackish peracaridan assemblages and their relation to some environmental variables in two tidal channels of the Ria de Aveiro (NW Portugal). Mar Ecol Prog Ser 190:69–87

Daehler C, Strong D (1997) Hybridization between introduced smooth cordgrass (*Spartina alterniflora*; Poaceae) and native California cordgrass (*S. foliosa*) in San Francisco Bay, California, USA. Am J Bot 84:607–611

Darling JA, Galil BS, Carvalho GR, Ruis M, Viard F, Piraino S (2017) Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. Mar Policy 85:54–64. https://doi.org/10.1016/j.marpol.2017.08.014

de Pina GM (1997) Records of intertidal amphipods from the southwest Atlantic, with the description of a new species of *Elasmopus*. J Crustacean Biol 17(4):745–757

Douglass JG, Duffy JE, Canuel EA (2011) Food web structure in a Chesapeake Bay eelgrass bed as determined through gut contents and 13 C and 15 N isotope analysis. Est Coasts 34:701–711

Dupuis JR, Roe AD, Sperling FA (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol Ecol 21(18):4422–4436

Faasse MA (2015) New records of the non-native amphipod *Ampithoe valida* in Europe. Mar Biodivers Records 8:e87

Fairey R, Dunn R, Sigala M, Oliver J (2002) Introduced aquatic species in California's coastal waters: Final Report. California Department of Fish and Game, Sacramento

Fitzpatrick BM, Shaffer HB (2004) Environment-dependent admixture dynamics in a tiger salamander hybrid zone. Evolution 58(6):1282–1293

Fitzpatrick BM, Johnson JR, Kump KD, Smith JJ, Voss RS, Shaffer BH (2010) Rapid spread of invasive genes into a threatened native species. Proc Natl Acad Sci 107(8):3606–3610

Fowler AE, Blakeslee AM, Canning-Clode J, Repetto MF, Phillip AM, Carlton JT, Moser FC, Ruiz GM, Miller AW (2016) Opening Pandora's bait box: a potent vector for biological invasions of live marine species. Divers Distrib 22(1):30–42

Fox RS, Bynum KH (1975) The amphipod crustaceans of North Carolina estuarine waters. Chesap Sci 16(4):223–237

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu Rev Ecol Evol Syst 34(1):397–423

Galil B, Froglia C, Noel P (2002) CIESM Atlas of exotic species in the Mediterranean. Crustaceans. Decapods and stomatopods. CIESM Publishers, Monaco

Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. Annual Rev Mar Sci 2:367–393

Gompert Z, Lucas LK, Buerkle CA, Forister ML, Fordyce JA, Nice CC (2014) Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. Mol Ecol 23:4555–4573

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27(2):221–224

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 59(3):307–321

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11(1):94

Kahng SE, Grigg RW (2005) Impact of an alien octocoral, *Carijoa riisei*, on black corals in Hawaii. Coral Reefs 24(4):556–562

Kemppainen P, Panova M, Hollander J, Johannesson K (2009) Complete lack of mitochondrial divergence between two species of NE Atlantic marine intertidal gastropods. J Evol Biol 22:2000–2011. https://doi.org/10.1111/j.1420-9101.2009.01810.x

Kim YH (2011) Side swimmers I. Arthropoda: Crustacea: Malacostraca: Amphipoda: Corophiidea, Gammaridea. National Institute of Biological Resources: Invertebrate Fauna of Korea 21(8)

Kim HS, Kim CB (1987) Marine gammaridean Amphipoda (Crustacea) of Cheju Island and its adjacent waters, Korea. Anim Syst Evol Divers 3(1):1–23

Kincaid T (1968) The ecology of Willapa Bay, Washington, in relation to the oyster industry. Seattle, WA: Self-published, pp. 84

Knowlton N (1993) Sibling species in the sea. Annu Rev Ecol Syst 24:189–216

Laakkonen HM, Hardman M, Strelkov P, Väinölä R (2021) Cycles of trans-Arctic dispersal and vicariance, and diversification of the amphi-boreal marine fauna. J Evol Biol 34:73–96. https://doi.org/10.1111/jeb.13674

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Lee CE (2002) Evolutionary genetics of invasive species. Trends Ecol Evol 17(8):386–391

Lewis JT, Boyer KE (2014) Grazer functional roles, induced defenses, and indirect interactions: implications for eelgrass restoration in San Francisco Bay. Divers 6(4):751–770

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14):1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecassis G, Durbin R (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079

Lu L, Levings CD, Piercey GE (2007) Preliminary investigation on aquatic invasive species of marine and estuarine macroinvertebrates on floating structures in five British Columbia harbours. Can Manuscr Rep Fisheries Aquat Sci 2814:2–30

Maggs CA, Castilho R, Foltz D, Henzler C, Jolly MT, Kelly J, Olsen J, Perez KE, Stam W, Väinölä R, Viard F (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology 89(11 Suppl):S108–S122

Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20(5):229–237

Martin A, Diaz YJ (2003) The amphipod fauna (Crustacea: Amphipoda) of shallow waters from the northeastern coast of Venezuela. Bol Instituto Esp de Oceanografia 19:327–344

McGeoch MA, Butchart SH, Spear D, Marais E, Kleynhans EJ, Symes A, Chanson J, Hoffmann M (2010) Global indicators of

biological invasion: species numbers, biodiversity impact and policy responses. Divers Distrib 16:95–108
Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. Front Ecol Environ 6:485–492
Morais P, Reichard M(2017) Cryptic invasions: A review.Sci Total Environ613–614
Nagata K (1960) Preliminary notes on benthic gammaridean amphipoda from the *Zostera* region of Mihara Bay, Seto Inland Sea, Japan. Publ Seto Mar Biol Lab 8(1):163–182
Ojaveer H, Galil BS, Carlton JT et al (2018) Historical baselines in marine bioinvasions: Implications for policy and management. PLoS ONE 13:e0202383. https://doi.org/10.1371/journal.pone.0202383
Orti G, Bell MA, Reimchen TE, Meyer A (1994) Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. Evolution 48(3):608–622
Palumbi SR, Kessing BD (1991) Population biology of the trans-arctic exchange: MtDNA sequence similarity between Pacific and Atlantic sea urchins. Evolution 45(8):1790–1805
Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528
Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle CA (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. Mol Ecol 21:2991–3005
Pardal MA, Marques JC, Metelo I, Lillebø AI, Flindt MR (2000) Impact of eutrophication on the life cycle, population dynamics and production of *Ampithoe valida* (Amphipoda) along an estuarine spatial gradient (Mondego Estuary, Portugal). Mar Ecol Prog Ser 196:207–219
Pilgrim EM, Darling JA (2010) Genetic diversity in two introduced biofouling amphipods (*Ampithoe valida* & *Jassa marmorata*) along the Pacific North American coast: investigation into molecular identification and cryptic diversity. Divers Distrib 16(5):827–839
R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Rawson PD, Harper FM (2009) Colonization of the northwest Atlantic by the blue mussel, *Mytilus trossulus* postdates the last glacial maximum. Mar Biol 156(9):1857–1868
Reid DG, Rumbak E, Thomas RH (1996) DNA, morphology and fossils: phylogeny and evolutionary rates of the gastropod genus *Littorina*. Phios Trans R Soc Lond B 351(1342):877–895
Reynolds LK, Carr LA, Boyer KE (2012) A non-native amphipod consumes eelgrass inflorescences in San Francisco Bay. Mar Ecol Prog Ser 451:107–118
Rilov G, Crooks JA (2009) Biological invasions in marine ecosystems. Ecol Stud 204:353
Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open 501 source tool for metagenomics. PeerJ 4:e2584
Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. Proc R Soc London B 273:2453–2459
Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling JA, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model across a large model space. Syst Biol 61:539–542
Ruiz GM, Carlton JT, Grosholz ED, Hines AH (1997) Global invasions of marine and estuarine habitats by non-indigenous species: mechanisms, extent, and consequences. Am Zool 37(6):621–632
Ruiz GM, Fofonoff PW, Steves B, Foss SF, Shiba SN (2011) Marine invasion history and vector analysis of California: a hotspot for western North America. Divers Distrib 17:362–373. https://doi.org/10.1111/j.1472-4642.2011.00742.x
Ryan ME, Johnson JR, Fitzpatrick BM (2009) Invasive hybrid tiger salamander genotypes impact native amphibians. Proc Natl Acad Sci 106(27):11166–11171
Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. Proc Natl Acad Sci 99(4):2445–2449
Sloan NA, Bartier PM (2004) Introduced Marine Species in the Haida Gwaii (Queen Charlotte Islands) Region, British Columbia. Can Field-Nat 118(1):77–84
Smith SI(1873) Crustacea, ex. Isopoda. In: AE. Verrill's "Report upon the invertebrate animals of Vineyard Sound" U.S. Commission of Fish and Fisheries. Pt. 1. Report on the Condition of the Sea Fisheries of the South Coast of New England in 1871 and 1872. 295–778, 4 Fig. 38 pls
Sotka EE, Wares JP, Hay ME (2003) Geographic and genetic variation in feeding preference for chemically defended seaweeds. Evolution 57:2262–2276
Sotka EE, Bell T, Hughes LE, Lowry JK, Poore AG (2017) A molecular phylogeny of marine amphipods in the herbivorous family Ampithoidae. Zool Scr 46(1):85–95
Strong DR, Ayres DR (2013) Ecological and evolutionary misadventures of Spartina. Annu Rev Ecol Evol Syst 44:389–410
Van Oppen MJH, Draisma SGA, Olsen JL, Stam WT (1995) Multiple trans-Arctic passages in the red alga *Phycodrys rubens*: evidence from nuclear rDNA ITS sequences. Mar Biol 123(1):179–188
Viard F, David P, Darling JA (2016) Marine invasions enter the genomic era: three lessons from the past, and the way forward. Curr Zool 62(6):629–642
Wares JP (2002) Community genetics in the Northwestern Atlantic intertidal. Mol Ecol 11:1131–1144
Wasson K, Zabin CJ, Bedinger L, Diaz MC, Pearse JS (2001) Biological invasions of estuaries without international shipping: the importance of intraregional transport. Biol Conserv 102(2):143–153
Yund PO, Collins C, Johnson SL (2015) Evidence of a native northwest atlantic COI haplotype clade in the cryptogenic colonial ascidian *Botryllus schlosseri*. Biol Bull 228(3):201–216