## Title

Single zooids, multiple loci: independent colonisations revealed by population genomics of a global invader

## Running title

Population genomics of a global invader

## Authors

Maria Casso[1,2], Xavier Turon[1,§], Marta Pascual[2,§,*]

[1] Center for Advanced Studies of Blanes (CEAB, CSIC), Blanes, Spain.

[2] Department of Genetics, Microbiology and Statistics, and IRBio, University of Barcelona, Spain.

* Corresponding author (93 403 48 50; martapascual@ub.edu)

[§] XT and MP should be considered joint senior authors

XT ORCID: 0000-0002-9229-5541

MP ORCID: 0000-0002-6189-0612

## Abstract

Assessing genomic diversity and population structure of non-indigenous species is crucial to develop adequate management strategies. However, in species with scarce material for DNA extraction, applying genomic techniques can be a difficult task. Here we set a protocol for small DNA samples combining whole genome amplification (WGA) and genotyping-by-sequencing (GBS). This protocol was applied to the worldwide invasive colonial ascidian *Didemnum vexillum* using a single zooid per colony. WGA-GBS performance was tested using half zooids, providing empirical demonstration for genotyping reliability. We analysed 296 individuals from 12 localities worldwide including native and the main invaded areas. Polymorphic loci datasets generated by locality, area and globally, identified genetic differentiation at all levels. The two groups found in Japan, the native area, matched Cytochrome Oxidase I clades and were strongly differentiated at the genomic level suggesting reproductive isolation. Our genomic analyses confirmed that only one clade spread worldwide. We also detected some clones, always within the same locality. Genetic diversity was high in both the introduced and in the native area. Three independent colonisation events determined the global distribution of the species, although population pairwise comparisons within each introduced genetic cluster were significant. Human-mediated transportation seems to drive the distribution pattern of this species worldwide and regionally, as there is a lack of isolation by distance within introduced areas. Diverse and well differentiated populations point to a high expansion potential of this worrisome worldwide invader.



## Keywords

Ascidians, *Didemnum vexillum*, genotyping-by-sequencing, population genomics, invasive species, whole genome amplification.

## INTRODUCTION

Non-indigenous species (NIS) are a major cause of ecological change in marine communities and one of the major threats to global biodiversity (Dukes and Mooney 1999; Bax et al. 2003; McGeoch et al. 2010; Darling et al. 2017). Marine biological invasions are due to anthropogenic vectors such as transportation-related pathways (e.g. ballast waters, fouling, interoceanic canals) and commerce of living organisms (e.g. aquaculture and aquarium trade) (Padilla and Williams 2004; Molnar et al. 2008; Galil et al. 2015).

Diverse methodologies have been developed to identify invasion processes since their understanding is fundamental for implementing adequate management strategies. Population genetic approaches have proved very valuable to assess invasion processes (Viard and Comtet 2015); in particular, the relationship between populations from widely separated invaded areas (Rius et al. 2012). Molecular tools are of great utility for NIS research and management in marine systems (Darling et al. 2017). However, information provided by molecular markers relies strongly on the type and number of markers used (Adrion et al. 2014). Nowadays, several molecular techniques and bioinformatic pipelines allow to perform population genomic studies in non-model species (Chown et al. 2015; Torkamaneh et al. 2016). These approaches generate hundreds to thousands of markers which remarkably increase the accuracy of population genetic studies (Tepolt and Palumbi 2015; Tepolt 2015; Gagnaire et al. 2015). The genome sizes of most species are too large for whole genome assessment and a reduction of the fraction being analysed is necessary. Restriction-enzyme associated DNA techniques allow this reduction, with the resulting fragments being distributed throughout the genome. These technologies, coupled with high throughput sequencing, increase the potential for identifying demographic and adaptive processes of ecologically relevant species (Hudson 2008; Carreras et al. 2017). To date, however, only a few works have applied population genomics to assess invasion processes in marine organisms (Bouchemousse et al. 2016; Jeffery et al. 2017; Gagnaire et al. 2018; Pérez-Portela et al. 2018). Population genomic approaches applied to marine invasive species may allow a deep understanding of

82 invasion processes, revealing unknown population structure, invasion patterns and
83 adaptation (Rius et al. 2015; Viard et al. 2016; Forsström et al. 2017).
84
85 Among marine invasive organisms, ascidians include several species of major
86 concern, some of which have worldwide colonisation ranges (Lambert and Lambert 2003;
87 Shenkar and Swalla 2011; López-Legentil et al. 2015; Zhan et al. 2015). They are diploid
88 and sessile marine filter feeders with characteristics that make them successful
89 colonisers: rapid growth rate and short time to maturity (Rius et al. 2009; Pineda et al.
90 2013; Casso et al. 2018), physical and chemical defences (Stoecker 1980; López-Legentil
91 et al. 2006), ability to reproduce asexually by fragmentation, and large number of short-
92 lived non-feeding planktonic larvae (Lambert 2001). Moreover, colonial ascidians can form
93 chimeras by allogeneic fusions, generating larger colonies with zooids of different
94 genotypes, which can be an advantage for an invasive species (Ben-Shlomo 2017).
95 Invasive ascidians are well-established on artificial substrates (Tyrrell and Byers 2007)
96 which makes them very common in marinas, aquaculture facilities and other human-built
97 structures (Airoldi et al. 2015). They can also be introduced in natural environments where
98 they can overgrow epibenthic communities and cover large areas of the seafloor (Bullard
99 et al. 2007). Thus, these biofouling species can cause an important negative ecological
100 and socio-economic impact (Aldred and Clare 2014).
101
102 *Didemnum vexillum* Kott, 2002 is a colonial ascidian thought to be native from Japan
103 (Lambert 2009; Stefaniak et al. 2012) which has colonised most of the world temperate
104 regions: New Zealand (Kott 2002), East and West coast of North America (Bullard et al.
105 2007; Lambert 2009; Cohen et al. 2011), Atlantic and Mediterranean coasts of Europe
106 (Minchin and Sides 2006; Gittenberger 2007; Griffith et al. 2009; Lambert 2009; El Nagar
107 et al. 2010; Tagliapietra et al. 2012; Ordóñez et al. 2015), and SE Russia (Zvyagintsev et
108 al. 2016). The chronology of its spread is difficult to determine as the taxonomy of the
109 genus is challenging and it has been historically misidentified (Lambert 2009). The most
110 likely large scale vector leading to primary introductions is trans-oceanic transport by
111 commercial shipping and aquaculture-related transfers, while at small scales, leisure craft

112 and regional commercial shipping are more probably contributing to secondary spread
113 within the region of introduction (Lambert 2009; Bishop et al. 2015). This species
114 overgrows commercial bivalves and artificial substrates such as boat hulls and marinas'
115 structures, so it has high economic impact (Bullard et al. 2007). Furthermore, it can spill
116 over natural communities (Mercer et al. 2009), and it has been reported to cause great
117 harm in some fishing grounds (Bullard et al. 2007; Valentine et al. 2007; Kaplan et al.
118 2018).
119
120 Phylogeographic studies of *D. vexillum* using Cytochrome Oxidase Subunit I (COI)
121 gene have shown the presence of two main mitochondrial clades, A and B, of which only
122 clade A has been detected in introduced localities (Stefaniak et al. 2012). However,
123 precise relationships among different areas of introduction could not be determined due
124 to low variability of this marker (Stefaniak et al. 2012; Ordóñez et al. 2015). Microsatellite
125 markers have been developed for this species and used to assess diversity in an invaded
126 locality (Abbott et al. 2011). These markers were also recently applied to identify
127 allogeneic fusion between individuals, a process suggested to play an important role in
128 invasion success and influence genetic data interpretation (Fidler et al. 2018; Watts et al.
129 2019). A microsatellite locus and two presumably nuclear sequence fragments have been
130 found to be useful to assess differentiation among some populations (Hess et al. 2009),
131 but never tested globally. Multimarker approaches applied to *D. vexillum* may provide a
132 more accurate assessment of its population structure and the processes involved in the
133 success of this global invasive species. In particular, genotyping-by-sequencing (GBS)
134 (Elshire et al. 2011) stands out by its simplicity as there is no fragment size selection, and
135 few enzymatic and purification steps are required, being thus time- and cost-efficient. This
136 technique requires an amount of DNA not always available, but whole genome
137 amplification (WGA) protocols are specifically designed to overcome this problem (Dean
138 et al. 2002). In our study, WGA was used to obtain high quality DNA from single zooids,
139 instead of sampling pools of zooids or colony fragments, to avoid contamination, problems
140 related to DNA degradation, and genetic heterogeneity due to chimeric colonies
141 (Rinkevich and Fidler 2014; Watts et al. 2019).

The objective of the present work is to determine patterns of genetic structure, and to assess the connectivity between populations and the colonisation history of *D. vexillum* using population genomics. More specifically, we (i) evaluated the reliability of several WGA-GBS protocols for correct and reproducible genotyping, (ii) compared genomic differentiation between COI clades of *Didemnum vexillum*, (iii) estimated diversity in native and introduced localities and (iv) assessed genetic relationships between and within the main areas where the species has been introduced. We expect that combining GBS and WGA will provide a genome-wide panel of markers for accurate inference of genetic patterns. Moreover, if mitochondrial clades correspond to different evolutionary units, we anticipate finding a strong divergence also at the nuclear level. With our multimarker approach, we expect to be able to accurately identify the population structure of the species in the native and invaded areas and obtain meaningful insights into the colonisation processes of this important worldwide invader. This study will further contribute to set guidelines for genomic analyses in organisms with scarce genetic material and without reference genome.

# METHODS

## Sampling

A total of 291 colonies of *Didemnum vexillum* from 12 locations were sampled around the world during 2015 and 2016 (Fig. 1, Table 1). Colonies were sampled in marinas and other artificial substrates and they were at least 2 m apart to avoid pseudoreplication (Smith et al. 2012). The samples were coded by geographic region as defined in a previous study (Stefaniak et al. 2012). Mediterranean localities, not included in that previous work, were grouped into the European geographic region. Samples were preserved in 96% ethanol.

## COI analysis

In order to determine the COI haplogroup of our Japanese samples, DNA extractions were carried out from thoraxes of five zooids for each of the 41 colonies from Aomori and Sugashima using the REDExtract-N-Amp Tissue kit (Sigma-Aldrich), following manufacturer's recommendations. A fragment of about 600 bp was amplified and sequenced using primers designed in a previous work (Stefaniak et al. 2012). PCR amplification was done in 20 µL total reaction volume with 10 µL of REDExtract-N-Amp PCR reaction mix (Sigma-Aldrich), 0.8 µL (10mM) of each primer, 6.4 µL of ultra-pure water (Sigma-Aldrich) and 2 µL of DNA at a concentration of ca. 5ng/µL. A total of 30 cycles with an annealing temperature of 50ºC were performed in a S1000 Thermal Cycler (BioRad). Sequencing was carried out at Macrogen facilities (Netherlands). Sequences were edited and analysed with Mega7 (Kumar et al. 2016). Our sequences were aligned and trimmed to 444 positions to avoid low quality bases. Additional sequences from GenBank of *D. vexillum* COI haplotypes coded as H1 to H23 (Stefaniak et al. 2012; Smith et al. 2012) (acc. JF738057 to JF738069 and JQ663509 to JQ663517) were included for phylogenetic analysis. Identical sequences, as a result of trimming, were collapsed (i.e. H1.2 = H1 and H2; H5.6 = H5 and H6; H10.21 = H10 and H21) and a neighbour-joining tree of the resulting haplotypes built using the proportion of nucleotide differences (p-distances).

## DNA extraction and sequencing for genomic analysis

For each colony, DNA extraction was performed from the thorax of a single zooid, to avoid contamination due to digestive content and to avoid genetic heterogeneity, as colonies of *Didemnum vexillum* can fuse forming chimeric colonies (Rinkevich and Fidler 2014; Fidler et al. 2018). DNA from each thorax was obtained by whole genome amplification (WGA) with REPLI-g® Single Cell kit (Qiagen) after showing the best performance on a preliminary test (Appendix S1). Additionally, to assess potential genotyping error from WGA-GBS combined protocol, single thoraxes from eight colonies of Ebro Delta were cut in two halves and extracted separately. Overall, 307 samples (16 half thoraxes and 291 thoraxes) were independently extracted and amplified following the

manufacturer's protocol except for a reduction on the amount of polymerase used (1.5 µL).

WGA DNA samples were sent to the National Centre for Genetic Analysis (CNAG, Barcelona). Separate libraries were constructed for each sample. DNA was digested with PstI restriction enzyme (see Appendix S1 for justification of enzyme selection) and adaptors ligated with a single-step PCR. The libraries of approximately 80 individuals were pooled and paired-end sequenced (2x125 bp fragments) in an Illumina HiSeq 2500 platform. Individuals from the same locality were distributed in different lanes.

**Loci datasets construction**

Demultiplexed raw sequences from 16 half-thoraxes on one hand and the 291 individuals of the global dataset on the other hand were filtered and analysed following the pipeline described in Appendix S2. We used the GIbPSs toolkit (Hapke and Thiele 2016) as it can deal with paired ends, and performs better than other programs in terms of computation time and number of shared loci (Hapke and Thiele 2016). We chose the option to define alleles as haplotypes, combining all variable positions of each locus, instead of using each single nucleotide polymorphism (SNP) separately to increase the statistical power of the analyses (RYMAN et al. 2006).

For the half-thoraxes dataset, tables of haplotypic genotypes and sequencing depths were exported from GIbPSs and read into R (R Core Team 2018). The haplotypic genotypes table was used to obtain the Percentage of Shared Genotypes (PSG, i.e. percentage of identical genotypes among shared loci). PSG values together with the table of sequencing depths were used to determine WGA-GBS reliability. For the global dataset, PSG values were also obtained from the table of haplotypic genotypes and used to detect clones. The PSG between samples with the same genotype is expected to be significantly higher than between samples with different genotype, even between siblings. However, PSG values may vary depending on sequencing errors and depth bias.

We kept only one colony from each of the clones detected and ran again the same pipeline to obtain the final global loci dataset which was exported to a genepop file format from GIbPSs and read into R using the 'adegenet' package (Jombart 2008; Jombart and Ahmed 2011). Loci in Hardy-Weinberg disequilibrium in more than 60% of localities (Benestan et al. 2015; Carreras et al. 2017) were detected using the R package 'pegas' (Paradis 2010) and removed. Using the final global loci dataset, we expected to find some genetic structure in the species, at least between COI clades and/or between native and introduced localities. Thus, we decided to first analyse the population structure, build new loci datasets for each identified genetic cluster, and then perform the relevant genetic analyses within each dataset. For comparison with the results obtained using alleles as haplotypes, we ran again the pipeline defining alleles with only the first SNP at each locus, using the sequences of the main genetic cluster.

**Data analyses**

Pairwise genetic distances among individuals were calculated using the Prevosti distance in the R package 'poppr' (Kamvar et al. 2014, 2015). We chose the Prevosti distance as it is a model-free method (Grünwald et al. 2017) which calculates the absolute differences between two individuals at all loci. We corrected this distance by the exact number of loci shared by each pair to avoid bias due to missing data. With the resulting genetic distances, Principal Coordinates Analysis (PCoA) were run using the R package 'ade4' (Dray and Dufour 2007) and plotted with 'ggplot2' (Wickham 2009). To determine the optimal number of genetic groups (K), twenty runs were carried out using Structure v2.3.4 (Pritchard et al. 2000) for each K ranging from one to 16. Burn-in period was set to 50,000 MCMC steps and runs to 500,000 steps to achieve consistent results. The most likely number of genetic groups was inferred by the *ad hoc* statistic ΔK (Evanno et al. 2005) using Structure Harvester v0.6.94 (Earl and vonHoldt 2012). The 20 runs of Structure for the best K values were averaged using Clumpp vs1.1.2 (Jakobsson and Rosenberg 2007) and plotted with the 'ggplot2' R package (Wickham 2009).

Locality pairwise distances ($F_{ST}$) were calculated with 'hierfstat' (Goudet and

Jombart 2015). The significances of these distances were tested by 999 permutations and p-values corrected for multiple comparisons using the Benjamini-Yekutieli FDR correction (White et al. 2019). Mantel tests were performed between $F_{ST}$ and geographic distances using 'ade4' (Dray and Dufour 2007). An AMOVA analysis was performed for the introduced populations, using as hierarchical levels the genetic groups detected, the localities, and the individuals. This analysis was performed with Arlequin v. 3.5. (Excoffier and Lischer 2010).

Heterozygosities, inbreeding coefficients and allelic richness were estimated with 'diveRsity' (Keenan et al. 2013). The number of singletons (alleles found only once in the dataset) and unique alleles (alleles found only in one locality) were obtained from genepop files using 'adegenet' (Jombart 2008; Jombart and Ahmed 2011).

# RESULTS

### COI analysis

Among the 41 analysed individuals from the two Japanese localities we found 10 haplotypes, of which four had not been previously described (accession numbers XXX, pending). The neighbor-joining tree showed that all individuals from Aomori (25) and seven individuals from Sugashima had clade A haplotypes and the other 9 individuals from Sugashima had clade B haplotypes (Fig. 2). The four new haplotypes (Accession numbers MN071417-20) were all of individuals from Sugashima, two of clade A and two of clade B (Fig. 2). In Sugashima, we found 0.97% nucleotide differences within clade A, 0.78% within clade B and 3.75% among clades.

### WGA-GBS performance

On average, 2,765,845 raw reads were obtained for each of the 16 half-thoraxes, and 2,201,744 reads remained after the sequence quality filtering stage of the pipeline (Appendix S2). Overall, 51,296 loci were found from which only 2,981 loci were kept after

the loci filtering stage. The mean Percentage of Shared Genotypes (PSG) (±SE) between samples with the same genotype, in this case half-thoraxes from the same individual, was higher (N=8, 93.24% ±0.53) than among samples from different individuals (N=112, 41.65% ±0.17), and differed significantly (Mann-Whitney-Wilcoxon test; $W$ = 896; p-value < 0.0001) (Fig. 3). When comparing each pair of half thoraxes (Appendix S3), the mean sequencing depths of non-shared-genotype loci were lower (16.5±2.4) than those of shared-genotype loci (112.1±2.5). Most non-shared-genotype loci had two different alleles in each pair of half-thoraxes, which is compatible with sequencing depth bias (i.e., undetected heterozygotes), while only 0.59% of the total loci had three different alleles in a given pair, attributable to sequencing errors, and none had four different alleles.

**Clone identification**

When analysing the sequences of the 291 individuals, 1,975,862 raw reads were obtained on average per individual and one individual from Brest was removed due to low number of reads (less than 20% of the average). After the sequence filtering stage, the number of reads for the 290 retained individuals averaged 1,620,457 per individual. A total of 194,064 loci were identified of which only 473 loci were kept since they were successfully genotyped in more than 70% of individuals. Thirteen pairs of individuals shared 99.71%±0.09 (mean±SE) of the genotypes, and were thus identified as clones. The higher PSG than in the half-thoraxes dataset is partly due to the lower number of loci retained and their higher mean sequencing depth. In all 13 cases, both individuals of the pair belonged to the same locality: 1 pair from Aomori, 2 from New Zealand, 2 from Alaska, 3 from Seattle, 1 from Richmond, 1 from Sausalito, 1 from Portsmouth and 2 from Ebro Delta. Only one individual per clone-pair was retained to redo the global loci dataset.

**Nuclear divergence between mitochondrial clades**

The new loci identification performed on the retained 277 individuals after removing clones found 462 loci after filtering, hereafter the global dataset (Table 2). However, the nine individuals from Sugashima with clade B COI haplotypes had a lot of missing data and only had between 168 and 189 loci of the global dataset. Filtering was applied to loci

but not to individuals and, consequently, loci (not individuals) with more than 30% of missing data were removed. Thus, the low number of shared nuclear loci between both mitochondrial clades could be the result of the large divergence between them. Two genetically differentiated groups of samples were observed with the first axis of the PCoA (Fig. 4). One group included all individuals with mitochondrial clade B haplotypes and the other included all individuals with mitochondrial clade A haplotypes from Aomori and Sugashima and also all individuals from the introduced localities. We therefore assume that all introduced individuals would present mitochondrial clade A haplotypes, the only one recorded so far outside Japan (Stefaniak et al. 2012; Ordóñez et al. 2015).

For a deeper comparison between mitochondrial clades A and B, a new loci identification was performed using only the sequences of the 16 individuals from Sugashima, the only location including both clades in our sampling. With this dataset, 595 loci were identified (Table 2) and the average corrected-Prevosti distance within clade A was 0.107 (SE±0.002), within clade B was 0.183 (SE±0.003) and between clades was 0.340 (SE±0.001). The resulting PCoA clearly separated both clades (Appendix S4). Moreover, the most likely number of clusters identified was K = 2 (ΔK = 5,267.8), with individual assignment probabilities of 100% to belong to one or the other group (results not shown).

**Differentiation within clades**

Due to the strong differentiation between genetic groups, the 277 individuals from the global dataset were then split to run new loci identifications for each group. Hereafter, we will refer to these datasets as clade B (9 individuals from Sugashima) and clade A (the other 268 individuals) loci datasets. The clade B dataset (2,536 loci retained, Table 2) showed high heterozygosities (He=0.518, Ho = 0.387) and the $F_{IS}$ value was significant ($F_{IS}$ = 0.247, $p<0.05$).

The analysis performed using the sequences of the 268 clade A individuals retained 577 loci (Table 2). In the clade A dataset, the most likely number of clusters identified was

K = 3 (ΔK = 1,193.8), with the highest peak, and K = 13 (ΔK = 28.2) with a lower peak (Appendix S5). Both the clustering of the posterior probabilities of each individual with K = 3 (Fig. 5a) and the PCoA of the 268 clade A individuals (Fig. 6a) showed three separate genetic groups for the introduced localities: New Zealand plus East North America (NZ-ENA), West North America (WNA) and Europe (EUR). Sugashima presented a central position in the PCoA and showed membership probabilities split in each of the three clusters for K = 3. Aomori clustered with the NZ-ENA group (Fig. 5), which is consistent with the PCoA where these three localities were in overlapping positions along the first axis (Fig. 6a). For comparison with these results, we performed the same analyses using the dataset with only the first SNP of each locus. The most likely number of clusters identified was again K = 3 (ΔK = 87.4). The PCoA also showed the same three separate genetic groups. However, these groups were less well defined in both analyses and in the PCoA the individuals of the different localities within the same group were more overlapped (Appendix S6).

The clade A loci dataset was further split into the three subsets of introduced populations revealed by the clustering analysis (NZ-ENA, WNA and EUR) and independent loci datasets were constructed (Table 2). In each of the PCoAs performed separately for the three introduced genetic groups (Fig. 6b, 6c and 6d), localities were well differentiated from each other. Consistently with the PCoAs of these three groups, the clustering for K = 13 (Fig. 5b) recovered almost every locality as belonging mostly to a single genetic cluster, with two exceptions. In New Zealand, all individuals had assignment probabilities of ca. 50% to two genetic clusters, one of them being the main genetic cluster in Woods Hole individuals. In Aomori, 15 individuals were assigned to a genetic group while another 9 had mixed membership probabilities with a second genetic group. These two groups were only found with high probability in Aomori. Due to the high differentiation between localities of the clade A, more specific analyses were conducted using this dataset and, additionally, new loci datasets were constructed for each locality, including a loci dataset with the only seven clade A individuals from Sugashima. Using the clade A loci dataset, pairwise corrected-Prevosti distances between individuals from the same

locality averaged 0.112 (SE±2.52e-04), between individuals from the same genetic group but different localities averaged 0.131 (SE±1.20e-04), and between individuals from different genetic groups averaged 0.143 (SE±6.92e-05). The differences were significant (Kruskal-Wallis chi-squared = 10912.58; p-value < 0.001) between each of these three types of comparisons (Dunn test, all p-values < 0.001).

Pairwise $F_{ST}$ values were calculated between localities using the clade A loci dataset, and all of them were significant after FDR correction (Appendix S7). The lowest genetic differentiation was between Brest and Ebro Delta while Seattle was the locality with the highest differentiation when compared to the other localities. Interestingly, Sausalito was more differentiated from the other locality of San Francisco Bay than from Alaska, which is not expected according to geographic distance. The correlation between $F_{ST}$ and geographic distances was not significant within the NWA populations (Mantel test; r = 0.337; p-value = 0.295). Likewise, Brest was more differentiated from Portsmouth than from Delta, despite the latter is geographically more distant. Accordingly, no isolation by distance was detected among European localities (Mantel test; r = 0.396; p-value = 0.252). A hierarchical $F_{ST}$ analysis (AMOVA) revealed significant genetic differentiation between the three groups identified among the introduced populations (NZ-ENA, WNA and EUR), as well as among populations within groups and among and within individuals, the latter explaining the highest percentage of variation (Appendix S8).

Among the locality-specific analyses, Sugashima had the highest number of polymorphic loci (2,942), although the number of polymorphic loci per introduced locality was high (2,046±117, mean±SE) (Table 3). The highest number of alleles per locus was found in Aomori (3.32). Interestingly, as the number of individuals per locality decreased, the number of shared loci increased but the mean number of alleles per locus decreased. In the clade A analysis, the highest allelic richness was also in Aomori (2.08) and mean allelic richness in introduced localities ranged between 1.55 in Seattle and 1.91 in Brest (Table 3). Expected heterozygosities were higher than observed ones, in all cases, and those from the locality-specific analyses were always higher (mean of 40% higher) than

those obtained using clade A loci dataset (Table 3), suggesting that more variable loci are dropped when filtering loci of a dataset with higher number of individuals. The highest expected heterozygosities were found in Sugashima and Aomori in the locality-specific analysis (Table 3). In both clade A and locality-specific analyses, $F_{IS}$ values were all positive and significant except for Sugashima clade A and Seattle.

Using the clade A loci dataset, the number of singletons (alleles found only once in the dataset) and number of unique alleles (alleles found only in one locality) varied from 10 to 121 and from 11 to 193 respectively (Table 3). The mean number of unique alleles in Japanese populations was 114.5 and in the other localities was 45.2. Interestingly, the proportion of unique alleles that were not singletons was higher in Aomori (54%) and Sugashima (50%) than in the other localities (which ranged between 0% and 39%). The number of singletons and unique alleles in the Ebro Delta locality was very high (121 and 126, respectively). However, half of the singletons were found in only two individuals, which also showed the highest number of missing loci within the clade A loci dataset (20.62% and 14.21%).

## DISCUSSION

We present here the first population genomic study on a colonial ascidian, the invasive pest *Didemnum vexillum*. Our results provide empirical evidence for validating the combination of WGA and GBS in population genomic studies of samples with low DNA content. Our assays on half thoraxes showed high genotyping reliability, with discrepancies only in loci with low sequencing depth, indicating that sequencing depth is critical for correct genotype assessment. We showed that the colonisation process of *Didemnum vexillum* comprises three main independent introductions with clear genetic footprints: to New Zealand plus East North America, to West North America and to Europe. These findings showcase how a genomic approach can provide a deeper understanding of invasion processes than traditional population genetic techniques (Clark

et al. 2010; Tepolt and Palumbi 2015). Population genomics would be particularly useful in cases where neither the historical records nor the use of less sensitive genetic markers could disentangle the introduction history of a given species (Carlton and Cohen 2003; Pineda et al. 2011; Adrion et al. 2014).

Previous population genomic studies on ascidians focused only on large solitary forms, such as *Ciona* spp. (Bouchemousse et al. 2016; Lin et al. 2017) and *Pyura chilensis* (Segovia et al. 2017). Colonial ascidians have reduced zooid size, thus many zooids must be pooled to get enough DNA for a GBS approach. However, this increases the probability of sample contamination, and it is highly time consuming. Furthermore, this group of ascidians can form chimeras (Rinkevich and Weissman 1987; Rinkevich 2005) which can be a problem when genotyping a sample as it may contain two or more genotypes (Watts et al. 2019). Thus, to avoid genetic heterogeneity, a single zooid should be used (Pérez-Portela et al. 2013). Recently, a few studies started to use WGA on non-model species to increase the amount of DNA obtained (Blair et al. 2015; Grogan et al. 2016; Onyango et al. 2016) but, to our knowledge, this method had not been previously applied to marine invertebrates. Choosing the best WGA kit is not trivial, as performance can vary substantially, and differences in economic cost are not negligible. Although several comparative studies did not find substantial variation when using different WGA commercial kits (Pinard et al. 2006; Han et al. 2012; Blair et al. 2015; Ivanov et al. 2018), we found significant differences in yield and reliability using three different kits. We selected REPLI-g Single Cell kit (Qiagen) based on amplification success (100%), yield (36.58 µg per sample) and genotyping reliability (94% out of 2,981 loci) with the few failing loci due mainly to low sequencing depth.

The reliability of the technique combining WGA-GBS has been proved using the Percentage of Shared Genotypes (PSG). This variable was very useful to identify samples with the same or different genotype, so it can be applied not only to identify clones but also chimeric colonies, in a fast and easy way. Thus, for a more efficient workflow, this method of sample filtering could be incorporated into pipelines for calling SNPs or allele

variants. We found PSG values greater than 98% between clones from the same locality, which is higher than the 93% found between half-thoraxes from the same individual. This increase can be explained because the mismatches observed between half-thoraxes were mostly due to loci with low sequencing depth, while less than 1% of the total loci had mismatches attributable to other errors. Loci with low sequencing depth were more prone to be eliminated during the filtering process as more individuals are included in the datasets. It is therefore expectable that the larger global dataset has less mismatching loci than the restricted half-thoraxes dataset. Thus, the PSG threshold should be defined *a posteriori*, as it depends on the number of samples and loci in each dataset. The clonality found is likely due to colony fragmentation and propagule reattachment, that can disperse asexually *D. vexillum* (Morris and Carman 2012; Reinhardt et al. 2012; Stefaniak and Whitlatch 2014), as has been reported in other groups of colonial or modular invertebrates (Wulff 1991; Baums et al. 2006; Calderón et al. 2007; Kürn et al. 2011).

For the clade A dataset, we found more than 150,000 loci before filtering, of which ca. two thirds were polymorphic, with an average length of 170 bp, which represents approximately 5% of genome coverage, assuming a genome size of ca. 540 Mb (Velandia-Huerto et al. 2016). Thus we analysed a similar proportion of the genome as in other GBS studies (Carreras et al. 2017), indicating that WGA is not reducing the fraction of the genome being assayed. A large proportion of the initial loci was eliminated during the filtering procedure, mostly because they were not shared by at least 70% of individuals. The mean sequencing depth per locus with our selected restriction enzyme, PstI, was large (Appendices S1 and S3), suggesting that this drop cannot be explained by an inadequate number of reads per individual. Highly variable regions have an increased probability of changes in the restriction site, resulting in null alleles. This implies that the larger is the sample, the smaller is the number of shared loci. Thus, the construction of different loci datasets with subsets of individuals is crucial to optimize the information at different genetic groups (i.e. population, region and species). This hierarchical and multiple dataset approach is especially important in species with large distribution range, and strong differentiation between mitochondrial clades, regions and

localities like *D. vexillum*. It is also important to use a pipeline considering haplotypic variants at each locus, including all variable positions, not only individual SNP loci, as shown by our comparison with the results of the clade A loci datasets using both a single SNP per locus and whole haplotypes. In this way, using alleles as haplotypes we can take full advantage of the large amount of genetic information obtained by GBS.

Previous works on *D. vexillum* based on COI showed the presence of two main mitochondrial clades, A and B, being clade A the only found in invasive populations (Stefaniak et al. 2012; Ordóñez et al. 2015). However, when sequencing the nuclear marker *tho2*, differentiation between clades was not observed (Stefaniak et al. 2012). This was taken as evidence that the two clades do not correspond to separate species. On the contrary, our genome-wide analyses showed clear divergence at the nuclear level and high genetic diversity within clades, suggesting reproductive isolation even in sympatric samples and a potential on-going speciation process. Each clade should be studied in more detail at the biological and phylogenetic levels to determine in which stage of the speciation process these two clades can be placed at present (De Queiroz 2007). In instances of cryptic speciation among ascidians, contrasting invasiveness of the different genetic clades has been found, e.g., *Botryllus schlosseri* complex (Bock et al. 2012) or *Diplosoma listerianum* complex (Pérez-Portela et al. 2013). In our study, clade A was the only found in introduced populations, a result that may be explained either by specific characteristics (i.e. habitat preferences) that increase invasion success in clade A or just by chance with only clade A colonies transported outside the native range (Stefaniak et al. 2012). Since three independent colonisation events have shaped the present day distribution in the introduced area, most probably a combination of different evolutionary processes is responsible for clade A invasion success.

In marine invasive species, similar genetic diversity has been frequently found in introduced and native populations, which has been interpreted as a consequence of a strong propagule pressure or admixture processes, complicating the identification of the invasion processes (Rius et al. 2015). Unravelling the native region and introduction

processes of *D. vexillum* clade A proved difficult as historical records are confusing and with species name changes (Griffith et al. 2009). However, including a sample collected in 1926 but misidentified as another species, these authors could determine that the first record of *D. vexillum* was actually in Japan. The species is common and widespread in Japan, mostly on artificial structures, but also in natural communities, thus it was suggested that the Pacific is likely the origin of *D. vexillum* (Lambert 2009). Further studies using COI (Stefaniak et al. 2009, 2012) also strongly indicated that Japan is located within the native range of *D. vexillum*, based on comparisons of genetic diversity and allelic richness. A decrease of 41% of COI haplotype diversity in non-Japanese vs Japanese populations was found within clade A (Stefaniak et al. 2012). However, our multilocus nuclear dataset revealed a milder reduction in gene diversity (He, 12%) and in the number of alleles (9%) within the non-Japanese localities as compared to Sugashima and Aomori. Yet, the highest values of both genetic variability (He=0.298) and allelic richness (2.08) were found in a Japanese locality (Aomori). A clearer indication of ancestral area is seen with the number of unique alleles, which were 2.5 times more abundant in the Japanese populations than in the other localities. In particular, the proportion of unique alleles that were not singletons were 3.2 times higher in Japanese localities. This tendency is also observed in other invasive species where the frequency of unique alleles is clearly reduced in the introduced area despite low diversity reduction (Adrion et al. 2014). Therefore, our results agree with the NW Pacific being the native area of this species and the origin of its worldwide expansion. Why the reduction in genetic diversity was more marked with a mitochondrial gene can relate to the haploidy and lower effective population size of mitochondrial DNA (Ballard and Whitlock 2004). The weak reduction in nuclear DNA diversity may indicate that introduced populations retain an important proportion of genetic variability upon which selective forces can act and explain why the species survives in diverse environmental conditions in the introduced area. A more exhaustive sampling of the native range is needed to trace back sources and identify if admixture and selection had occurred.

The PCoA and the STRUCTURE analysis performed using clade A loci dataset point

561 to three areas independently colonised: Europe (Portsmouth, Brest, Ebro Delta and
562 Venice), West North America (Alaska, Seattle, Richmond and Sausalito) and another
563 formed by two widely separated regions, East North America (Woods Hole) and New
564 Zealand (Nelson). AMOVA results further confirmed a significant genetic differentiation
565 between these three groups of populations. The related colonisation of New Zealand and
566 East North America, although geographically disjunct, can be explained by long-distance
567 dispersal mediated by maritime traffic or shellfish culture. Although our results do not show
568 higher diversity in Woods Hole (ENA) than in Nelson (NZ), the direction of this spread is
569 most likely from East North America to New Zealand based on historical records. The first
570 observations of *D. vexillum* in New Zealand were in 2001, while the first verified record of
571 the species in the eastern coast of North America is a specimen collected in 1993, and
572 photographic evidence exists at least from the eighties, indicating the likely presence of
573 the species back in the seventies (Lambert 2009). Within the independently colonised
574 areas of Europe and West North America, the genetic and geographic distances are not
575 related. Clearly, anthropogenic vectors such as maritime traffic and aquaculture trade are
576 needed to explain these patterns as shown in similar case studies (Zhan et al. 2010;
577 Hudson et al. 2016). In the case of the Ebro Delta, which is an aquaculture facility, oyster
578 juveniles (seeds) are purchased in West France hatcheries by farmers (pers. comm.).
579 Thus, that area may be a source for different populations in Europe, with present-day
580 gene flow due to recurrent introductions linked to shellfish culture activities.
581
582 Overall, this temperate species seems to be able to thrive from cold waters like
583 Alaska, with temperatures from 4 to 15ºC (Cohen et al. 2011), to warmer environments
584 like the Ebro Delta, 8 to 28ºC (Casso et al. 2018). It can also colonise environments with
585 extreme cold and warm temperatures over the year as the Venetian Lagoon, 0 to >30ºC
586 (Tagliapietra et al. 2012). All this suggests that *D. vexillum* clade A has either a high
587 plasticity in its thermal tolerance ranges, or fast adaptive capability to temperature
588 stresses. Epigenetic changes have been suggested to have a role in surviving
589 environmental stresses in this species (Hawes et al. 2018). All these features, coupled
590 with plasticity in reproductive and growth cycles (Ordóñez et al. 2015), make *D. vexillum*

591 a highly successful invader – but see (Stefaniak 2017) – and an excellent model to test
592 ecophenotypic variation and adaptation. Other future approaches using genome-wide
593 markers should focus on ascertaining processes that can foster the invasive ability of the
594 species, such as the presence of chimeric colonies within a given population (Smith et al.
595 2012; Fidler et al. 2018; Watts et al. 2019), asexual reproduction and multiple paternity.
596 Moreover, the study of the symbionts of *D. vexillum* can also provide insights about its
597 adaptive potential, as ascidians harbour diverse microbiome communities (Erwin et al.
598 2014), that can have a role in adaptation (Evans et al. 2017).
599
600 In conclusion, our study demonstrated that three main introduction events have
601 shaped the present-day structure of *D. vexillum* in temperate waters of the world. Passive
602 transportation, such as aquaculture trade and shipping, seems to drive the distribution
603 pattern of this species worldwide and regionally, since geographically close localities can
604 be more differentiated than distant ones. Sympatric individuals from the two COI clades
605 were strongly differentiated at the genomic level suggesting reproductive isolation and
606 either an on-going speciation process or even the coexistence in sympatry of two cryptic
607 species. The construction of different loci datasets is crucial to test hypotheses at several
608 levels when a hierarchical structure is present. Finally, we empirically show the feasibility
609 and reliability of combining WGA and GBS in population genomics studies when DNA
610 yield of the samples is limited. This approach will set the basis for genomic analyses of
611 small sized organisms or colonial forms with small zooids, as well as studies where only
612 trace material (scales, hairs...) is available, as is often the case for elusive and/or
613 endangered species.
614
615 ## Acknowledgements
616 We are deeply indebted to all people that contributed providing samples for this
617 study: Gretchen Lambert, Ian Davidson, Mike Page, Judith Pederson, and Frédérique
618 Viard. We specially thank Margarita Fernández, Carles Bori, Davide Tagliapietra, Marco
619 Sigovini and Irene Guarneri for their help during sampling in Ebro Delta and Venice; Marc
620 Rius for help while sampling in UK, and Gaku Kumano for assistance with the Aomori

sampling. This research was funded by the projects CHALLENGEN and PopCOmics (CTM2013-48163 and CTM2017-88080, MCIU/AEI/FEDER/UE) from the Spanish Government. MC was funded by a predoctoral FPI contract of the Spanish Government. This is a contribution from the Consolidated Research Group “Benthic Biology and Ecology” SGR2017-1120 (Catalan Government).

## REFERENCES

Abbott CL, Ebert D, Tabata A, Therriault TW (2011) Twelve microsatellite markers in the invasive tunicate, *Didemnum vexillum*, isolated from low genome coverage 454 pyrosequencing reads. Conserv Genet Resour 3:79–81. doi: 10.1007/s12686-010-9294-2

Adrion JR, Kousathanas A, Pascual M, et al (2014) *Drosophila suzukii*: The genetic footprint of a recent, worldwide invasion. Mol Biol Evol 31:3148–3163. doi: 10.1093/molbev/msu246

Airoldi L, Turon X, Perkol-Finkel S, Rius M (2015) Corridors for aliens but not for natives: effects of marine urban sprawl at a regional scale. Divers Distrib 21:755–768. doi: 10.1111/ddi.12301

Aldred N, Clare AS (2014) Mini-review: Impact and dynamics of surface fouling by solitary and compound ascidians. Biofouling 30:259–270. doi: 10.1080/08927014.2013.866653

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13:729–44. doi: 10.1046/j.1365-294X.2003.02063.x

Baums IB, Miller MW, Hellberg ME (2006) Geographic variation in clonal structure in a reef-building Caribbean coral, *Acropora palmata*. Ecol Monogr 76:503–519

Bax N, Williamson A, Aguero M, et al (2003) Marine invasive alien species: a threat to global biodiversity. Mar Policy 27:313–323. doi: 10.1016/S0308-597X(03)00041-1

Ben-Shlomo R (2017) Invasiveness, chimerism and genetic diversity. Mol Ecol 26:6502–6509. doi: 10.1111/mec.14364

Benestan L, Gosselin T, Perrier C, et al (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). Mol Ecol 24:3299–3315. doi: 10.1111/mec.13245

Bishop JDD, Wood CA, Lévêque L, et al (2015) Repeated rapid assessment surveys reveal contrasting trends in occupancy of marinas by non-indigenous species on opposite sides of the western English Channel. Mar Pollut Bull 95:699–706. doi: 10.1016/J.MARPOLBUL.2014.11.043

Blair C, Campbell CR, Yoder AD (2015) Assessing the utility of whole genome amplified DNA for next-generation molecular ecology. Mol Ecol Resour 15:1079–1090. doi:

10.1111/1755-0998.12376
Bock DG, Macisaac HJ, Cristescu ME (2012) Multilocus genetic analyses differentiate between widespread and spatially restricted cryptic species in a model ascidian. Proc R Soc B Biol Sci 279:2377–2385. doi: 10.1098/rspb.2011.2610
Bouchemousse S, Liautard-Haag C, Bierne N, Viard F (2016) Distinguishing contemporary hybridization from past introgression with postgenomic ancestry-informative SNPs in strongly differentiated *Ciona* species. Mol Ecol 25:5527–5542. doi: 10.1111/mec.13854
Bullard SG, Lambert G, Carman MR, et al (2007) The colonial ascidian *Didemnum* sp. A: Current distribution, basic biology and potential threat to marine communities of the northeast and west coasts of North America. J Exp Mar Bio Ecol 342:99–108. doi: 10.1016/j.jembe.2006.10.020
Calderón I, Ortega N, Duran S, et al (2007) Finding the relevant scale: clonality and genetic structure in a marine invertebrate (*Crambe crambe*, Porifera). Mol Ecol 16:1799–1810. doi: 10.1111/j.1365-294X.2007.03276.x
Carlton JT, Cohen AN (2003) Episodic global disperal in shallow water marine organisms: the case history of the European shore crabs *Carcinus maenas* and *C. aestuarii*. J Biogeogr 30:1809–1820
Carreras C, Ordóñez V, Zane L, et al (2017) Population genomics of an endemic Mediterranean fish: differentiation by fine scale dispersal and adaptation. Sci Rep 7:43417. doi: 10.1038/srep43417
Casso M, Navarro M, Ordóñez V, et al (2018) Seasonal patterns of settlement and growth of introduced and native ascidians in bivalve cultures in the Ebro Delta (NE Iberian Peninsula). Reg Stud Mar Sci 23:12–22. doi: 10.1016/j.rsma.2017.11.002
Chown SL, Hodgins KA, Griffin PC, et al (2015) Biological invasions, climate change and genomics. Evol Appl 8:23–46. doi: 10.1111/eva.12234
Clark MS, Tanguy A, Jollivet D, et al (2010) Populations and Pathways: Genomic Approaches to Understanding Population Structure and Environmental Adaptation Melody. In: J C, K T-R, C B, Viard F (eds) Introduction to Marine Genomics. Advances in Marine Genomics, vol 1. Springer, Dordrecht, pp 73–118
Cohen CS, McCann L, Davis T, et al (2011) Discovery and significance of the colonial tunicate *Didemnum vexillum* in Alaska. Aquat Invasions 6:263–271. doi: 10.3391/ai.2011.6.3.03
Darling JA, Galil BS, Carvalho GR, et al (2017) Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. Mar Policy 85:54–64. doi: 10.1016/j.marpol.2017.08.014
De Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56:879–86. doi: 10.1080/10635150701701083
Dean FB, Hosono S, Fang L, et al (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci 99:5261–5266
Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20
Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? Trends Ecol Evol 14:135–139. doi: 10.1016/S0169-5347(98)01554-7
Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for

706 visualizing STRUCTURE output and implementing the Evanno method. Conserv
707 Genet Resour 4:359–361. doi: 10.1007/s12686-011-9548-7
708 El Nagar A, Huys R, Bishop JDD (2010) Widespread occurrence of the Southern
709 Hemisphere ascidian *Corella eumyota* Traustedt, 1882 on the Atlantic coast of
710 Iberia. Aquat Invasions 5:169–173. doi: 10.3391/ai.2010.5.2.06
711 Elshire RJ, Glaubitz JC, Sun Q, et al (2011) A robust, simple genotyping-by-sequencing
712 (GBS) approach for high diversity species. PLoS One 6:e19379. doi:
713 10.1371/journal.pone.0019379
714 Erwin PM, Pineda MC, Webster N, et al (2014) Down under the tunic: bacterial
715 biodiversity hotspots and widespread ammonia-oxidizing archaea in coral reef
716 ascidians. ISME J 8:575–588. doi: 10.1038/ismej.2013.188
717 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals
718 using the software structure: a simulation study. Mol Ecol 14:2611–2620. doi:
719 10.1111/j.1365-294X.2005.02553.x
720 Evans JS, Erwin PM, Shenkar N, López-Legentil S (2017) Introduced ascidians harbor
721 highly diverse and host-specific symbiotic microbial assemblages. Sci Rep 7:1–11.
722 doi: 10.1038/s41598-017-11441-4
723 Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to
724 perform population genetics analyses under Linux and Windows. Mol Ecol Resour
725 10:564–567. doi: 10.1111/j.1755-0998.2010.02847.x
726 Fidler AE, Bacq-Labreuil A, Rachmilovitz E, Rinkevich B (2018) Efficient dispersal and
727 substrate acquisition traits in a marine invasive species via transient chimerism and
728 colony mobility. PeerJ 6:e5006. doi: 10.7717/peerj.5006
729 Forsström T, Ahmad F, Vasemägi A (2017) Invasion genomics: genotyping-by-
730 sequencing approach reveals regional genetic structure and signatures of temporal
731 selection in an introduced mud crab. Mar Biol 164:186. doi: 10.1007/s00227-017-
732 3210-1
733 Gagnaire PA, Broquet T, Aurelle D, et al (2015) Using neutral, selected, and hitchhiker
734 loci to assess connectivity of marine populations in the genomic era. Evol. Appl.
735 8:769–786
736 Gagnaire PA, Lamy JB, Cornette F, et al (2018) Analysis of genome-wide differentiation
737 between native and introduced populations of the cupped oysters *Crassostrea gigas*
738 and *Crassostrea angulata*. Genome Biol Evol 10:2518–2534. doi:
739 10.1093/gbe/evy194
740 Galil BS, Boero F, Campbell ML, et al (2015) 'Double trouble': the expansion of the Suez
741 Canal and marine bioinvasions in the Mediterranean Sea. Biol Invasions 17:973–
742 976. doi: 10.1007/s10530-014-0778-y
743 Gittenberger A (2007) Recent population expansions of non-native ascidians in The
744 Netherlands. J Exp Mar Bio Ecol 342:122–126. doi: 10.1016/j.jembe.2006.10.022
745 Goudet J, Jombart T (2015) hierfstat: Estimation and Tests of Hierarchical F-Statistics. R
746 Packag version 004-22
747 Griffith K, Mowat S, Holt RHF, et al (2009) First records in Great Britain of the invasive
748 colonial ascidian *Didemnum vexillum* Kott, 2002. Aquat Invasions 4:581–590. doi:
749 10.3391/ai.2009.4.4.3
750 Grogan KE, McGinnis GJ, Sauther ML, et al (2016) Next-generation genotyping of

hypervariable loci in many individuals of a non-model species: technical and theoretical implications. BMC Genomics 17:204. doi: 10.1186/s12864-016-2503-y
Grünwald NJ, Everhart SE, Knaus BJ, Kamvar ZN (2017) Best Practices for Population Genetic Analyses. Phytopathology 107:1000–1010. doi: 10.1094/PHYTO-12-16-0425-RVW
Han T, Chang C-W, Kwekel JC, et al (2012) Characterization of whole genome amplified (WGA) DNA for use in genotyping assay development. BMC Genomics 13:217. doi: 10.1186/1471-2164-13-217
Hapke A, Thiele D (2016) GIbPSs: a toolkit for fast and accurate analyses of genotyping-by-sequencing data without a reference genome. Mol Ecol Resour 16:979–990. doi: 10.1111/1755-0998.12510
Hawes NA, Tremblay LA, Pochon X, et al (2018) Effects of temperature and salinity stress on DNA methylation in a highly invasive marine invertebrate, the colonial ascidian *Didemnum vexillum*. PeerJ 6:e5003. doi: 10.7717/peerj.5003
Hess JE, Swalla BJ, Moran P (2009) New molecular markers to genetically differentiate populations of *Didemnum vexillum* (Kott, 2002)-an invasive ascidian species. Aquat Invasions 4:299–310. doi: 10.3391/ai.2009.4.2.1
Hudson J, Viard F, Roby C, Rius M (2016) Anthropogenic transport of species across native ranges: unpredictable genetic and evolutionary consequences. Biol Lett 12:20160620. doi: 10.1098/rsbl.2016.0620
Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Mol Ecol Resour 8:3–17. doi: 10.1111/j.1471-8286.2007.02019.x
Ivanov V, Lee KM, Mutanen M (2018) Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. Mol Ecol 27:1681–1695. doi: 10.1111/mec.14564
Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806. doi: 10.1093/bioinformatics/btm233
Jeffery NW, DiBacco C, Van Wyngaarden M, et al (2017) RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. Ecol Evol 7:2513–2524. doi: 10.1002/ece3.2872
Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. doi: 10.1093/bioinformatics/btn129
Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. doi: 10.1093/bioinformatics/btr521
Kamvar ZN, Brooks JC, Grünwald NJ (2015) Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front Genet 6:208. doi: 10.3389/fgene.2015.00208
Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr : an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. doi: 10.7717/peerj.281
Kaplan KA, Hart DR, Hopkins K, et al (2018) Invasive tunicate restructures invertebrate community on fishing grounds and a large protected area on Georges Bank. Biol Invasions 20:87–103. doi: 10.1007/s10530-017-1517-y

796 Keenan K, McGinnity P, Cross TF, et al (2013) diveRsity: An R package for the
797 estimation and exploration of population genetics parameters and their associated
798 errors. Methods Ecol Evol 4:782–788. doi: 10.1111/2041-210X.12067
799 Kott P (2002) A complex didemnid ascidian from Whangamata, New Zealand. J Mar Biol
800 Assoc United Kingdom 82:625–628. doi: 10.1017/S0025315402005970
801 Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics
802 Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. doi:
803 10.1093/molbev/msw054
804 Kürn U, Rendulic S, Tiozzo S, Lauzon RJ (2011) Asexual propagation and regeneration
805 in colonial ascidians. Biol Bull 221:43–61
806 Lambert CC, Lambert G (2003) Persistence and differential distribution of nonindigenous
807 ascidians in harbors of the Southern California Bight. Mar Ecol Prog Ser 259:145–
808 161
809 Lambert G (2001) A global overview of ascidian introductions and their possible impact
810 on the endemic fauna. In: H. S, H. Y, C.C. L (eds) The Biology of Ascidians.
811 Springer, Tokyo, pp 249–257
812 Lambert G (2009) Adventures of a sea squirt sleuth: unraveling the identity of
813 *Didemnum vexillum*, a global ascidian invader. Aquat Invasions 4:5–28. doi:
814 10.3391/ai.2009.4.1.2
815 Lin Y, Chen Y, Yi C, et al (2017) Genetic signatures of natural selection in a model
816 invasive ascidian. Sci Rep 7:44080. doi: 10.1038/srep44080
817 López-Legentil S, Legentil ML, Erwin PM, Turon X (2015) Harbor networks as
818 introduction gateways: contrasting distribution patterns of native and introduced
819 ascidians. Biol Invasions 17:1623–1638. doi: 10.1007/s10530-014-0821-z
820 López-Legentil S, Turon X, Schupp P (2006) Chemical and physical defenses against
821 predators in *Cystodytes* (Ascidiacea). J Exp Mar Bio Ecol 332:27–36. doi:
822 10.1016/J.JEMBE.2005.11.002
823 McGeoch MA, Butchart SHM, Spear D, et al (2010) Global indicators of biological
824 invasion: Species numbers, biodiversity impact and policy responses. Divers Distrib
825 16:95–108. doi: 10.1111/j.1472-4642.2009.00633.x
826 Mercer JM, Whitlatch RB, Osman RW (2009) Potential effects of the invasive colonial
827 ascidian (*Didemnum vexillum* Kott, 2002) on pebble-cobble bottom habitats in Long
828 Island Sound, USA. Aquat Invasions 4:133–142. doi: 10.3391/ai
829 Minchin D, Sides E (2006) Appearance of a cryptogenic tunicate, a *Didemnum* sp.
830 fouling marina pontoons and leisure craft in Ireland. Aquat Invasions 1:143–147.
831 doi: 10.3391/ai.2006.1.3.8
832 Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of
833 invasive species to marine biodiversity. Front Ecol Environ 6:485–492. doi:
834 10.1890/070064
835 Morris JA, Carman MR (2012) Fragment reattachment, reproductive status, and health
836 indicators of the invasive colonial tunicate *Didemnum vexillum* with implications for
837 dispersal. Biol Invasions 14:2133–2140. doi: 10.1007/s10530-012-0219-8
838 Onyango MG, Aitken NC, Jack C, et al (2016) Genotyping of whole genome amplified
839 reduced representation libraries reveals a cryptic population of *Culicoides*
840 *brevitarsis* in the Northern Territory, Australia. BMC Genomics 17:769. doi:

841 10.1186/s12864-016-3124-1
842 Ordóñez V, Pascual M, Fernández-Tejedor M, et al (2015) Ongoing expansion of the
843 worldwide invader *Didemnum vexillum* (Ascidiacea) in the Mediterranean Sea: high
844 plasticity of its biological cycle promotes establishment in warm waters. Biol
845 Invasions 17:2075–2085. doi: 10.1007/s10530-015-0861-z
846 Padilla DK, Williams SL (2004) Beyond ballast water: aquarium and ornamental trades
847 as sources of invasive species in aquatic ecosystems. Front Ecol Environ 2:131–
848 138. doi: 10.1890/1540-9295(2004)002[0131:BBWAAO]2.0.CO;2
849 Paradis E (2010) pegas: an R package for population genetics with an integrated-
850 modular approach. Bioinformatics 26:419–420. doi: 10.1093/bioinformatics/btp696
851 Pérez-Portela R, Arranz V, Rius M, Turon X (2013) Cryptic speciation or global spread?
852 the case of a cosmopolitan marine invertebrate with limited dispersal capabilities.
853 Sci Rep 3:3197. doi: 10.1038/srep03197
854 Pérez-Portela R, Bumford A, Coffman B, et al (2018) Genetic homogeneity of the
855 invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on
856 Single Nucleotide Polymorphisms. Sci Rep 8:5062. doi: 10.1038/s41598-018-
857 23339-w
858 Pinard R, De Winter A, Sarkis GJ, et al (2006) Assessment of whole genome
859 amplification-induced bias through high-throughput, massively parallel whole
860 genome sequencing. BMC Genomics 7:216. doi: 10.1186/1471-2164-7-216
861 Pineda MC, López-Legentil S, Turon X (2013) Year-round reproduction in a seasonal
862 sea: biological cycle of the introduced ascidian *Styela plicata* in the Western
863 Mediterranean. Mar Biol 160:221–230. doi: 10.1007/s00227-012-2082-7
864 Pineda MC, López-Legentil S, Turon X (2011) The whereabouts of an ancient wanderer:
865 global phylogeography of the solitary ascidian *Styela plicata*. PLoS One 6:e25495.
866 doi: 10.1371/journal.pone.0025495
867 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
868 multilocus genotype data. Genetics 155:945–959
869 R Core Team (2018) A Language and Environment for Statistical Computing
870 Reinhardt JF, Gallagher KL, Stefaniak L, et al (2012) Material properties of *Didemnum*
871 *vexillum* and prediction of tendril fragmentation. Mar Biol 159:2875–2884. doi:
872 10.1007/s00227-012-2048-9
873 Rinkevich B (2005) Natural chimerism in colonial urochordates. J Exp Mar Bio Ecol
874 322:93–109. doi: 10.1016/j.jembe.2005.02.020
875 Rinkevich B, Fidler AE (2014) Initiating laboratory culturing of the invasive ascidian
876 *Didemnum vexillum*. Manag Biol Invasions 5:55–62. doi: 10.3391/mbi.2014.5.1.05
877 Rinkevich B, Weissman IL (1987) A long-term study on fused subclones in the ascidian
878 *Botryllus schlosseri*: the resorption phenomenon (Protochordata: Tunicata). J Zool
879 213:717–733. doi: 10.1111/j.1469-7998.1987.tb03736.x
880 Rius M, Pineda MC, Turon X (2009) Population dynamics and life cycle of the introduced
881 ascidian *Microcosmus squamiger* in the Mediterranean Sea. Biol Invasions
882 11:2181–2194. doi: 10.1007/s10530-008-9375-2
883 Rius M, Turon X, Bernardi G, et al (2015) Marine invasion genetics: from spatio-
884 temporal patterns to evolutionary outcomes. Biol Invasions 17:869–885. doi:
885 10.1007/s10530-014-0792-0

886 Rius M, Turon X, Ordóñez V, Pascual M (2012) Tracking invasion histories in the sea:
887 Facing complex scenarios using multilocus data. PLoS One 7:e35815. doi:
888 10.1371/journal.pone.0035815
889 RYMAN N, PALM S, ANDRÉ C, et al (2006) Power for detecting genetic divergence:
890 differences between statistical methods and marker loci. Mol Ecol 15:2031–2045.
891 doi: 10.1111/j.1365-294X.2006.02839.x
892 Segovia NI, Gallardo-Escárate C, Poulin E, Haye PA (2017) Lineage divergence, local
893 adaptation across a biogeographic break and artificial transport, shape the genetic
894 structure in the ascidian *Pyura chilensis*. Sci Rep 7:44559. doi: 10.1038/srep44559
895 Shenkar N, Swalla BJ (2011) Global Diversity of Ascidiacea. PLoS One 6:e20657. doi:
896 10.1371/journal.pone.0020657
897 Smith KF, Stefaniak L, Saito Y, et al (2012) Increased inter-colony fusion rates are
898 associated with reduced COI haplotype diversity in an invasive colonial ascidian
899 *Didemnum vexillum*. PLoS One 7:e30473. doi: 10.1371/journal.pone.0030473
900 Stefaniak L (2017) Mechanisms for invasion success by *Didemnum vexillum* (Chordata:
901 Ascidiacea): observations versus manipulations. Biol Invasions 19:1213–1225. doi:
902 10.1007/s10530-016-1317-9
903 Stefaniak L, Lambert G, Gittenberger A, et al (2009) Genetic conspecificity of the
904 worldwide populations of *Didemnum vexillum* Kott,2002. Aquat Invasions 4:29–44.
905 doi: 10.3391/ai
906 Stefaniak L, Whitlatch RB (2014) Life history attributes of a global invader: factors
907 contributing to the invasion potential of *Didemnum vexillum*. Aquat Biol 21:221–229.
908 doi: 10.3354/ab00591
909 Stefaniak L, Zhang H, Gittenberger A, et al (2012) Determining the native region of the
910 putatively invasive ascidian *Didemnum vexillum* Kott, 2002. J Exp Mar Bio Ecol
911 422–423:64–71. doi: 10.1016/j.jembe.2012.04.012
912 Stoecker D (1980) Relationships between chemical defense and ecology in benthic
913 ascidians. Mar Ecol Prog Ser 3:257–265
914 Tagliapietra D, Keppel E, Sigovini M, Lambert G (2012) First record of the colonial
915 ascidian *Didemnum vexillum* Kott, 2002 in the Mediterranean: Lagoon of Venice
916 (Italy). BioInvasions Rec 1:247–254. doi: 10.3391/bir.2012.1.4.02
917 Tepolt CK (2015) Adaptation in marine invasion: a genetic perspective. Biol Invasions
918 17:887–903. doi: 10.1007/s10530-014-0825-8
919 Tepolt CK, Palumbi SR (2015) Transcriptome sequencing reveals both neutral and
920 adaptive genome dynamics in a marine invader. Mol Ecol 24:4145–4158. doi:
921 10.1111/mec.13294
922 Torkamaneh D, Laroche J, Belzile F (2016) Genome-wide SNP calling from Genotyping
923 by Sequencing (GBS) data: a comparison of seven pipelines and two sequencing
924 technologies. PLoS One 11:e0161333. doi: 10.1371/journal.pone.0161333
925 Tyrrell MC, Byers JE (2007) Do artificial substrates favor nonindigenous fouling species
926 over native species? J Exp Mar Bio Ecol 342:54–60. doi:
927 10.1016/J.JEMBE.2006.10.014
928 Valentine PC, Carman MR, Blackwood DS, Heffron EJ (2007) Ecological observations
929 on the colonial ascidian *Didemnum* sp. in a New England tide pool habitat. J Exp
930 Mar Bio Ecol 342:109–121. doi: 10.1016/J.JEMBE.2006.10.021

Velandia-Huerto CA, Gittenberger A, Brown FD, et al (2016) Automated detection of ncRNAs in the draft genome sequence of a colonial tunicate: the carpet sea squirt *Didemnum vexillum*. BMC Genomics 17:691. doi: 10.1186/s12864-016-2934-5

Viard F, Comtet T (2015) Applications of DNA-based Methods for the Study of Biological Invasions. In: Canning-Clode J (ed) Biological Invasions in Changing Ecosystems: Vectors, Ecological Impacts, Management and Predictions. De Gruyter Open, Berlin, pp 411–435

Viard F, David P, Darling JA (2016) Marine invasions enter the genomic era: Three lessons from the past, and the way forward. Curr Zool 62:629–642. doi: 10.1093/cz/zow053

Watts AM, Hopkins GA, Goldstien SJ (2019) Chimerism and population dieback alter genetic inference related to invasion pathways and connectivity of biofouling populations on artificial substrata. Ecol Evol 1–16. doi: 10.1002/ece3.4817

White T, van der Ende J, Nichols TE (2019) Beyond Bonferroni revisited: concerns over inflated false positive research findings in the fields of conservation genetics, biology, and medicine. Conserv Genet. doi: 10.1007/s10592-019-01178-0

Wickham H (2009) ggplot2 : Elegant graphics for data analysis. Springer-Verlag, New York

Wulff JL (1991) Asexual fragmentation, genotype success, and population dynamics of erect branching sponges. J Exp Mar Bio Ecol 149:227–247

Zhan A, Briski E, Bock DG, et al (2015) Ascidians as models for studying invasion success. Mar. Biol. 162:2449–2470

Zhan A, MacIsaac HJ, Cristescu ME (2010) Invasion genetics of the *Ciona intestinalis* species complex: From regional endemism to global homogeneity. Mol Ecol 19:4678–4694. doi: 10.1111/j.1365-294X.2010.04837.x

Zvyagintsev AY, Sanamyan KE, Turanov S V., Kartavtsev YF (2016) Colonial ascidian *Didemnum vexillum* Kott, 2002 is an alien species in Peter the Great Bay (Sea of Japan). Russ J Biol Invasions 7:237–246. doi: 10.1134/S2075111716030140

## Data accessibility

Raw reads from all individuals, including information of location of all samples, can be found at NCBI SRA Bioproject PRJNA555829. The genotypic data used in the analyses of global, clade A and clade B individuals is available in Appendices S8 to S10. All loci datasets will be available upon request to the authors.

## Author Contributions

XT and MP designed research. MC and XT collected samples. MC performed

laboratory work, ran the bioinformatics pipeline, and wrote the first draft of the manuscript. All authors contributed to analyses, discussed and interpreted results, and revised the manuscript.

## Tables

Table 1. Region (Jap: Japan; NZ: New Zealand; ENA: East North America; WNA: West North America; EUR: Europe), location, code and number of *D. vexillum* individuals analysed.

| Region | Location (country) | Codes | Coordinates | N individuals |
|---|---|---|---|---|
| Jap | Aomori (Japan) | Aom | 40.900 N, 140.853 E | 25 |
| Jap | Sugashima (Japan) | Sug | 34.480 N, 136.881 E | 16 |
| NZ | Nelson (New Zealand) | NZ | 41.263 S, 173.267 E | 25 |
| ENA | Woods Hole (USA) | Woo | 41.772 N, 70.490 W | 25 |
| WNA | Sitka (USA) | Ala | 57.045 N, 135.371 W | 25 |
| WNA | Seattle (USA) | Sea | 47.398 N, 122.330 W | 25 |
| WNA | Richmond (USA) | Ric | 37.913 N, 122.349 W | 25 |
| WNA | Sausalito (USA) | Sau | 37.859 N, 122.480 W | 25 |
| EUR | Portsmouth (UK) | Por | 50.799 N, 1.119 W | 25 |
| EUR | Brest (France) | Bre | 48.374 N, 4.498 W | 25 |
| EUR | Ebro Delta (Spain) | Del | 40.776 N, 0.737 E | 25 |
| EUR | Venice (Italy) | Ven | 45.436 N, 12.379 E | 25 |

Table 2: Number of individuals (N), loci before filters (Loci BF), loci after filters (Loci AF), alleles, and the mean, minimum and maximum (range) number of alleles per locus, for each subset.

| Subset | N | Loci BF | Loci AF | Alleles | Mean A/L (range) |
|---|---|---|---|---|---|
| Global | 277 | 191,015 | 462 | 2,170 | 4.70 (2 - 15) |
| Sug A&B | 16 | 79,990 | 595 | 3,047 | 5.12 (2 - 13) |
| Clade A | 268 | 155,052 | 577 | 2,163 | 3.75 (2 - 13) |
| Clade B | 9 | 56,413 | 2,536 | 9,537 | 3.76 (2 - 10) |
| NZ-ENA | 48 | 56,181 | 1,642 | 5,270 | 3.21 (2 - 12) |
| WNA | 93 | 76,509 | 1,051 | 3,361 | 3.20 (2 - 14) |
| EUR | 97 | 116,795 | 933 | 3,188 | 3.42 (2 - 14) |

Footnotes: Global: all individuals of this study; Sug A&B: individuals from Sugashima of both clades; Clade A: individuals of the clade A; Clade B: individuals of the clade B; NZ-ENA: individuals from Nelson and Woods Hole; WNA: West North America individuals; and EUR: European individuals.

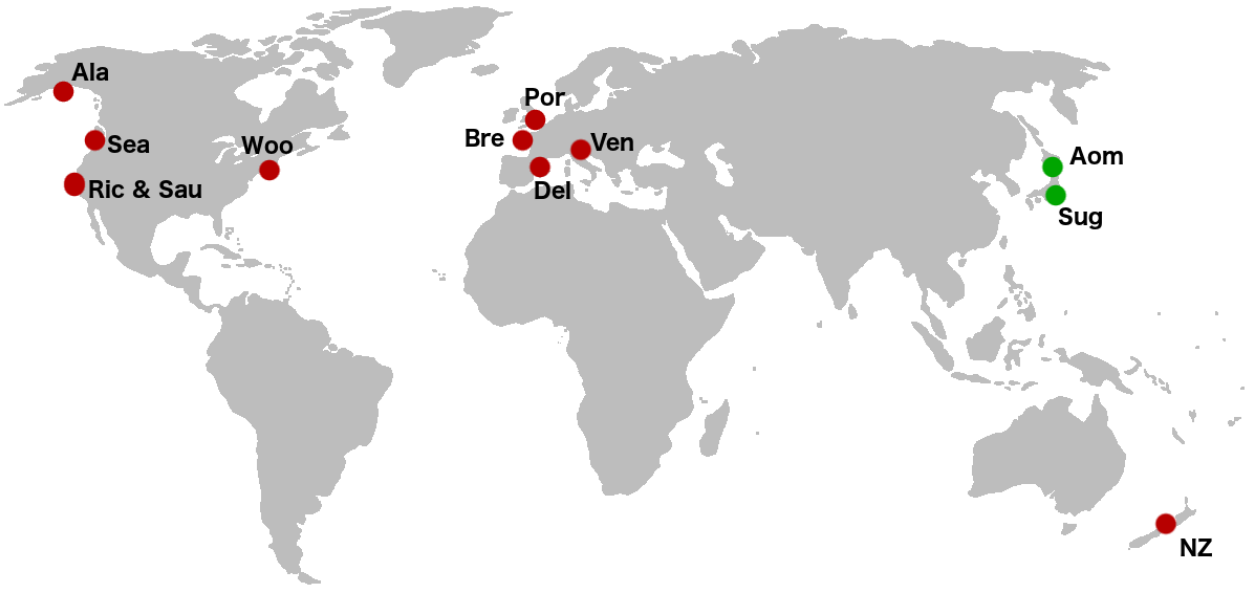
990 Table 3: Number of individuals after filtering for each locality (N). Number of loci (L) and
991 alleles per locus (A/L), expected and observed heterozygosity (He, Ho), and inbreeding
992 coefficients ($F_{IS}$) of the locality-specific loci datasets. Allele-richness (Ar), singletons (S),
993 unique alleles (UA), expected and observed heterozygosities (He, Ho) and inbreeding
994 coefficients ($F_{IS}$) obtained with the clade A loci dataset with 577 loci. * indicate significant
995 $F_{IS}$ values.
996

| | | Locality-specific analyses | | | | | Clade A analysis (with 577 loci) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Code** | **N** | **L** | **A/L** | **He** | **Ho** | **FIS** | **Ar** | **S** | **UA** | **He** | **Ho** | **FIS** |
| Aom | 24 | 2162 | 3.32 | 0.382 | 0.321 | 0.147* | 2.08 | 89 | 193 | 0.298 | 0.243 | 0.162* |
| SugA | 7 | 2942 | 2.71 | 0.412 | 0.384 | 0.050 | 1.86 | 18 | 36 | 0.236 | 0.228 | 0.024 |
| NZ | 23 | 2312 | 2.90 | 0.356 | 0.307 | 0.112* | 1.74 | 22 | 36 | 0.228 | 0.191 | 0.151* |
| Woo | 25 | 1995 | 2.92 | 0.357 | 0.314 | 0.103* | 1.80 | 28 | 44 | 0.247 | 0.220 | 0.107* |
| Ala | 23 | 2522 | 2.89 | 0.361 | 0.317 | 0.098* | 1.73 | 10 | 12 | 0.229 | 0.199 | 0.107* |
| Sea | 22 | 2382 | 2.73 | 0.328 | 0.322 | 0.008 | 1.55 | 15 | 15 | 0.186 | 0.185 | 0.005 |
| Ric | 24 | 1510 | 2.86 | 0.336 | 0.304 | 0.078* | 1.79 | 76 | 88 | 0.243 | 0.214 | 0.097* |
| Sau | 24 | 1805 | 2.94 | 0.335 | 0.302 | 0.079* | 1.71 | 11 | 11 | 0.222 | 0.192 | 0.113* |
| Por | 24 | 2470 | 2.91 | 0.348 | 0.316 | 0.078* | 1.75 | 16 | 23 | 0.238 | 0.220 | 0.080* |
| Bre | 24 | 1604 | 3.05 | 0.352 | 0.309 | 0.108* | 1.85 | 62 | 68 | 0.250 | 0.219 | 0.108* |
| Del | 23 | 1754 | 2.95 | 0.335 | 0.292 | 0.098* | 1.91 | 121 | 126 | 0.264 | 0.231 | 0.103* |
| Ven | 25 | 2101 | 2.94 | 0.349 | 0.303 | 0.109* | 1.76 | 25 | 29 | 0.236 | 0.199 | 0.148* |

997

998 **Figures**

999




1000

1001 Figure 1. Distribution map of the species. Sampled localities are indicated with codes (as

1002 in Table 1), in green dots (native) and red dots (introduced).

1003

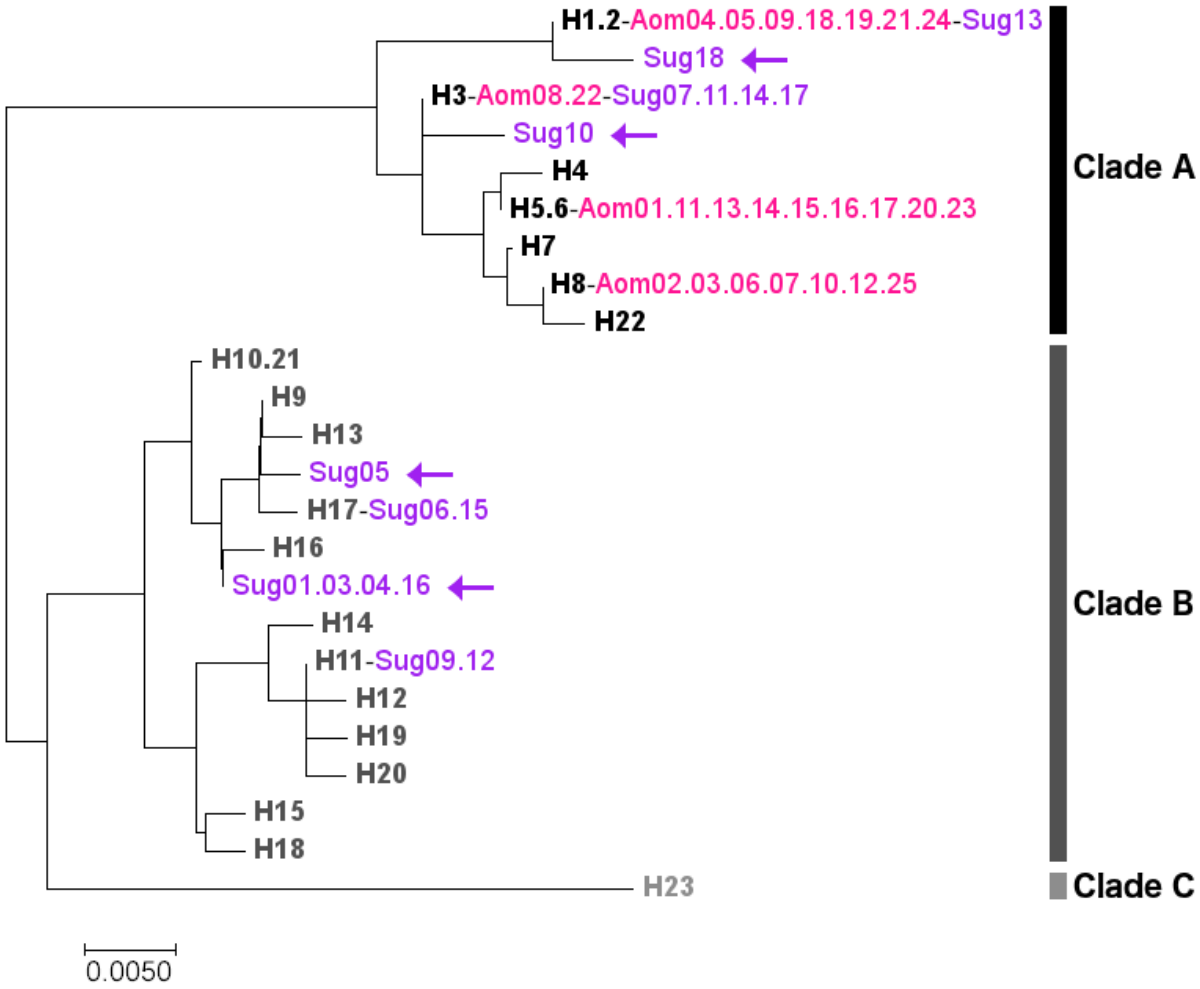
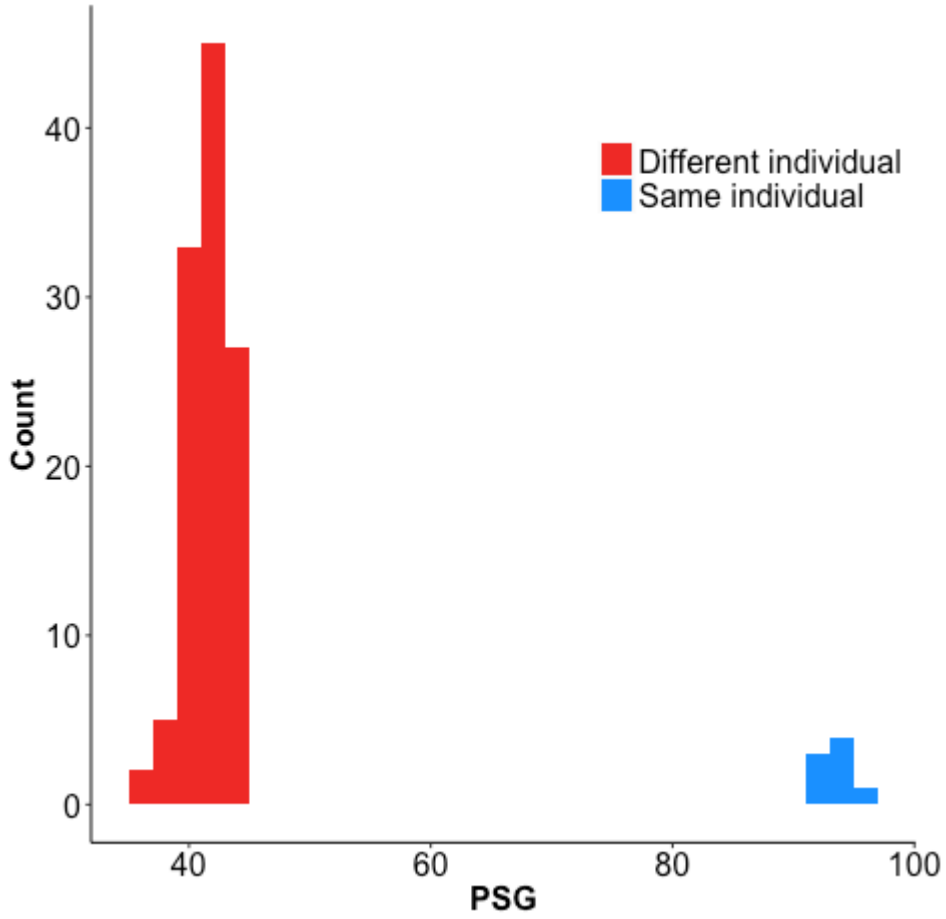



Figure 2: Neighbor-joining tree of the COI sequences from Aomori and Sugashima as well as previously described haplotypes (H1 to H23 in Smith et al 2012 and Stefaniak et al 2012). Clade A haplotypes are in black, those from clade B are in dark grey and that from clade C in light grey. Samples from Aomori (Aom01 to Aom25) are in pink. Samples from Sugashima (Sug01 to Sug18) are in purple. Multiple names in one tip correspond to individuals with the same trimmed sequence. Arrows indicate new haplotypes.




1012

1013 Figure 3: Frequency distribution of the Percentage of Shared Genotypes (PSG) between
1014 half thoraxes from the same individual (in blue) and from different individuals (in red), for
1015 the 2,981 retained loci.
1016

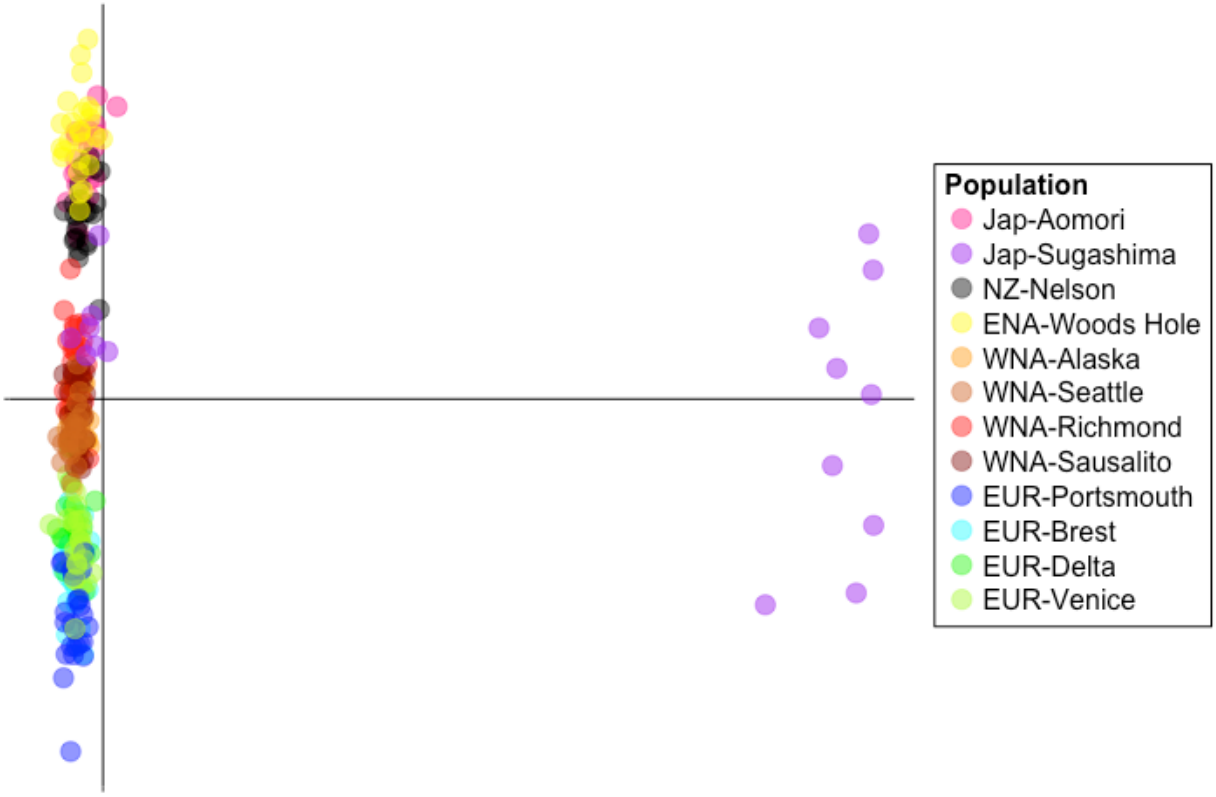
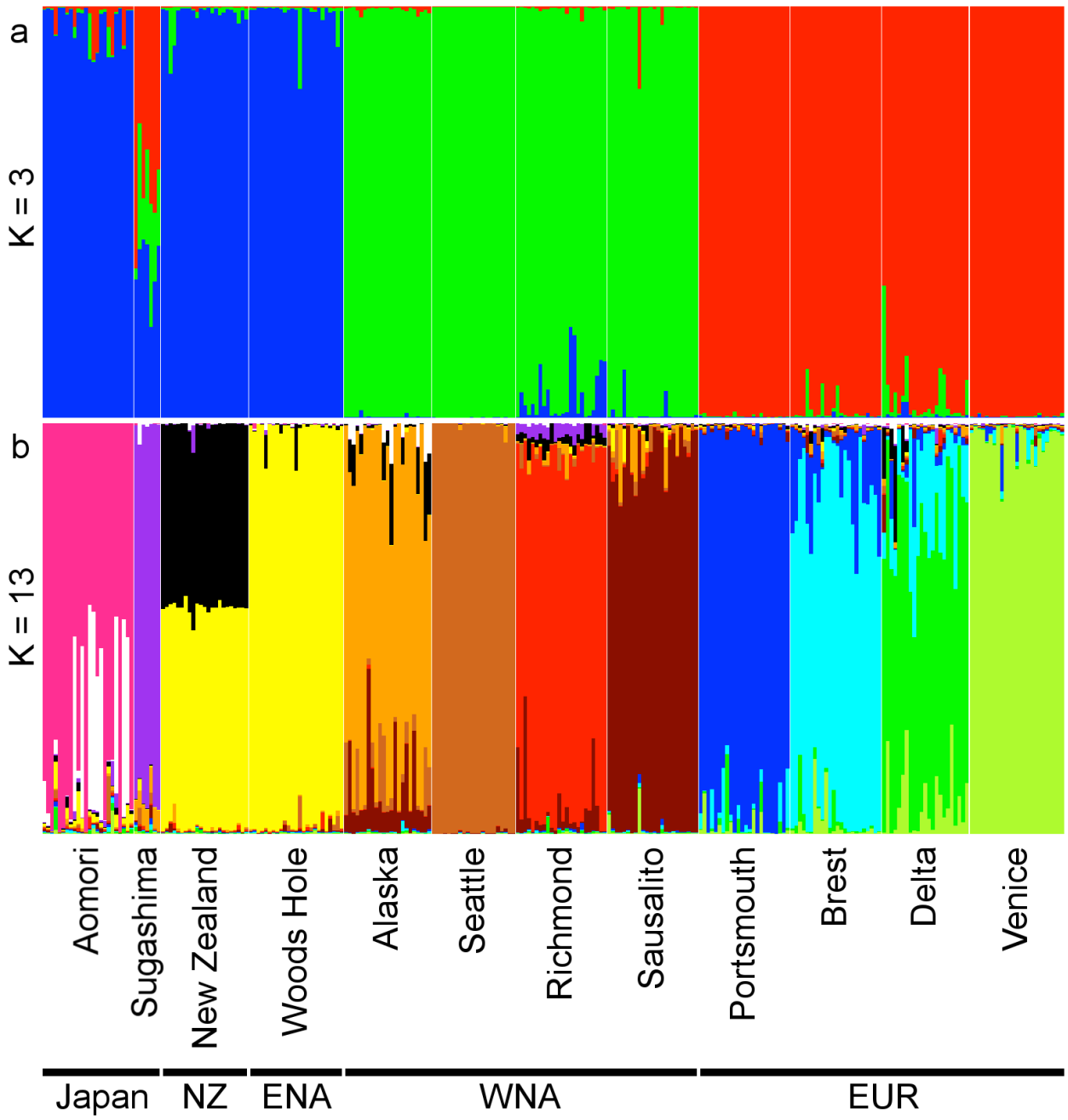



Figure 4: PCoA of the global loci dataset including 462 loci. Separated dots in the right end of the first axis correspond to the nine individuals from Sugashima with clade B haplotypes. The names of the localities in the legend include their geographic sampling region: Jap: Japan, NZ: New Zealand, ENA: East North America, WNA: West North America, and EUR: Europe. Percentages of variance explained by the first and second axis are 34.04% and 4.42%, respectively.




1025

1026 Figure 5: Posterior probabilities of individual assignment to the most probable number of
1027 clusters using the clade A dataset: a) K = 3; and b) K = 13. Geographic regions are
1028 indicated: Japan, NZ: New Zealand, ENA: East North America, WNA: West North
1029 America, and EUR: Europe.

1030

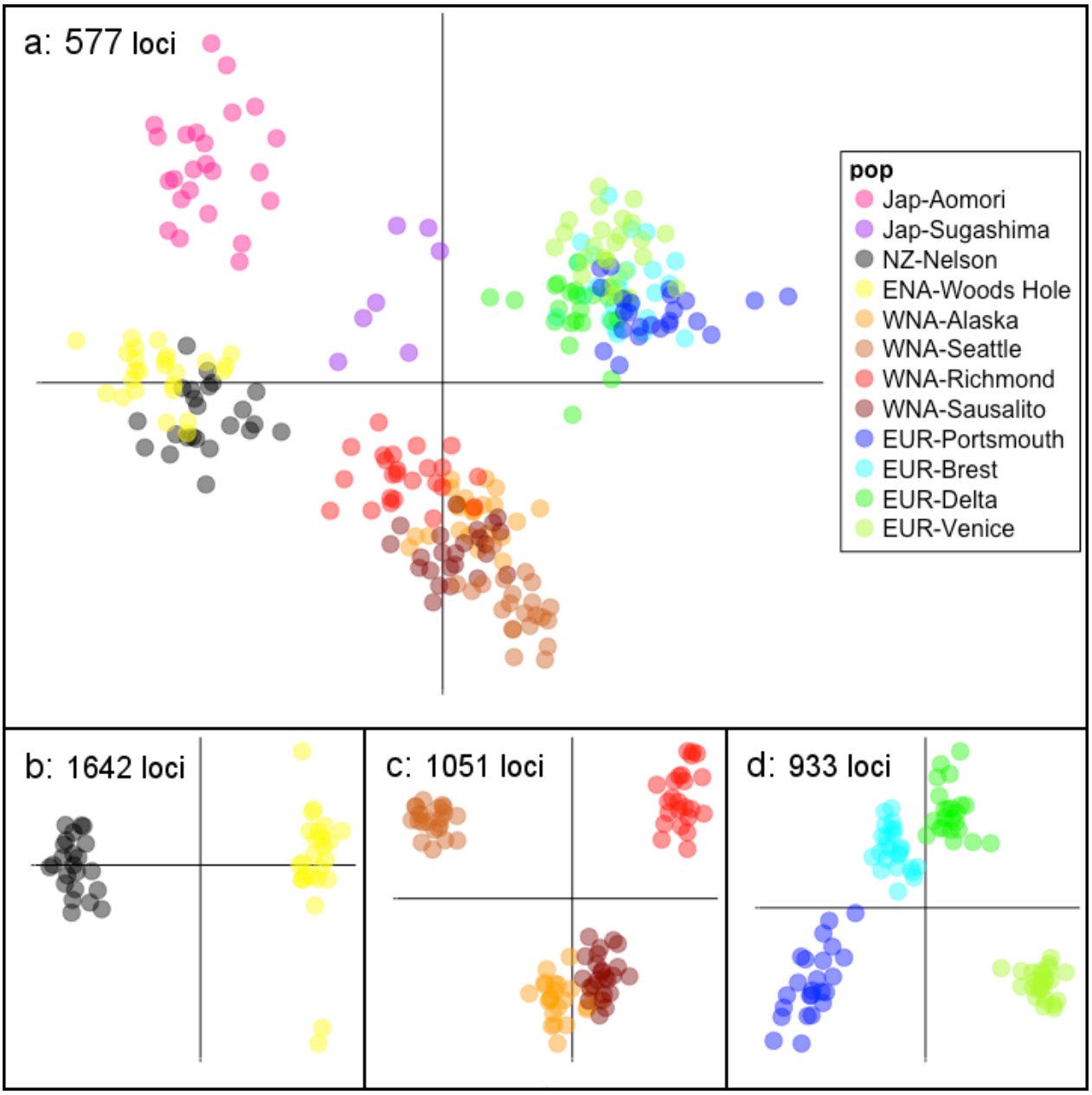



Figure 6: PCoA of a) clade A loci dataset, b) NZ-ENA c) WNA and d) EUR. The names of the localities in the legend include their geographic sampling region: Jap: Japan, NZ: New Zealand, ENA: East North America, WNA: West North America, and EUR: Europe. Percentages of variance explained by the first and second axis are respectively a) 7.48% and 6.13%, b) 22.09% and 3.65%, c) 14.59% and 7.84%, and d) 9.78% and 7.00%.