DR. ROMAN WENNE (Orcid ID : 0000-0001-6994-1181)

Article type : Original Article

Native and invasive taxa on the Pacific coast of South America: Impacts on aquaculture, traceability and biodiversity of blue mussels (*Mytilus* spp.)

M.A. Larraín[1], M. Zbawicka[2], C. Araneda[3], J. P. A. Gardner[4] and R. Wenne[2]

[1]Universidad de Chile, Facultad de Ciencias Químicas y Farmacéuticas, Departamento de Ciencia de los Alimentos y Tecnología Química, Sergio Livingstone 1007, Independencia, Santiago, Chile. mlarrain@uchile.cl

[2]Institute of Oceanology, Polish Academy of Sciences, Powstańców Warszawy 55, 81-712 Sopot, Poland. mzbawicka@iopan.gda.pl

[3]Universidad de Chile, Facultad de Ciencias Agronómicas, Departamento de Producción Animal, Avda. Santa Rosa 1315, La Pintana, Santiago, Chile. craraned@uchile.cl

[4]Victoria University of Wellington, School of Biological Sciences, P O Box 600, Wellington 6140, New Zealand. jonathan.gardner@vuw.ac.nz

CORRESPONDING AUTHOR

Roman Wenne
Institute of Oceanology
Polish Academy of Sciences
Powstańców Warszawy 55, 81-712
Sopot, Poland
rwenne@iopan.gda.pl

RUNNING TITLE

Blue mussel species identification in Chile

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/eva.12553

Abstract

Gaining new knowledge of the native distributions of species (phylogeography) is more and more difficult in a world affected by anthropogenic disturbance, in particular by species translocations. Increasingly, molecular markers are required to support decisions about the taxonomy of native versus introduced species, and the existence of their hybrids, to answer phylogeographic questions. In many fields, including aquaculture, traceability, and food security, taxonomic and phylogeographic knowledge is key to the successful management and conservation of biodiversity. The Pacific coast of Chile is one of the last regions without a clear and agreed understanding of the taxonomy and systematics of smooth-shelled blue mussels of the genus *Mytilus*. A panel of 49 bi-allelic single nucleotide polymorphisms (SNPs) was genotyped in 338 *Mytilus* individuals collected from nine Chilean and five reference populations. All analyses confirmed the hypothesis that the native Chilean blue mussel is genetically distinct from the reference species *M. edulis*, *M. galloprovincialis* and *M. trossulus*. These results support the hypothesis of a unique evolutionary history of the native Chilean blue mussel on the Pacific coast of South America. It is therefore concluded that the native blue mussel from Chile should be recognised as *M. chilensis* Hupé 1854. We confirmed a recent Mediterranean origin of introduced *M. galloprovincialis* on the coast of Chile. This knowledge advances the understanding of global phylogeography of blue mussels and their bioinvasions, and harmonises taxonomy in the context of aquaculture production, seafood traceability, labelling and trade.

KEYWORDS

Aquaculture – Food security – Invasive species – Conservation genetics – *Mytilus chilensis* – *Mytilus galloprovincialis*

1. INTRODUCTION

Determining the native distributions of species and thereby establishing baseline areas of natural ranges in the absence of anthropogenic disturbance is a key component of phylogeography (Avise, 2000). Such research is now often underpinned by taxonomic classifications and phylogenies that are based on molecular differentiation at one or more marker types (Morin et al., 2004; Pereira et al., 2008). The determination of distributions of native species is not only important in its own right, but is applied in a variety of different management aspects, such as aquaculture, food security, traceability and labelling. Food fraud is now a bigger problem than ever, including dilution, mislabelling, ingredient substitution and tampered with products. This problem is addressed in recent international legislation (FDA 2011; E.U. 2013). Related to aquaculture, marine bioinvasions that may affect cultivated species are now being detected in many production areas and also in remote regions including offshore islands and Antarctica (Lee and Chown, 2007; Shaw et al., 2014; Gardner et al., 2016). It has been suggested that we are no longer in a position to wait for the establishment of native species baselines and that immediate management action is required (Ojaveer et al., 2015). Thus, taxonomy has a key role to play in the protection and sustainable exploitation of species (Mace, 2004; Seddon et al., 2005).

Blue mussels of the genus *Mytilus* exhibit an antitropical distribution (Hilbish et al., 2000). They are an important source of protein for many coastal communities and are farmed commercially in many countries (FAO 2016; Ferreira and Bricker, 2016). Smooth-shelled *Mytilus* species are characterised by a high degree of phenotypic plasticity that has limited power to discriminate amongst the taxa, but they can now be identified with much greater certainty using genetic molecular markers (Zbawicka et al., 2012; Wennerstrom et al. 2013; Zbawicka et al., 2014; Gardner et al., 2016; Wenne et al., 2016; Jilberto et al., 2017). Traditionally, biochemical and molecular characterisation has divided the *Mytilus edulis*

species complex into three sibling species: *M. edulis* (Linnaeus, 1758), *M. trossulus* (Gould, 1850) and *M. galloprovincialis* (Lamarck, 1819) (Koehn, 1991; McDonald et al., 1991). Although the distributions of these taxa have been studied extensively in the Northern hemisphere, in the Southern hemisphere taxon distributions do not always fit the classical "*Mytilus edulis* complex" scheme. Whilst we have a better (but still not completely resolved) understanding of the situation in Australia and New Zealand (Westfall and Gardner, 2010; 2013; Ab Rahim et al., 2016; Gardner et al., 2016), the situation in South America still requires resolution (Gaitán-Espitia et al., 2016; Oyarzún et al., 2016).

The taxonomic status of native smooth-shelled blue mussels inhabiting the coast of Chile has been unclear and disputed for a number of years, with authors advancing different suggestions for its nomenclature, including *M. edulis*-like (Koehn, 1991; McDonald et al., 1991), *M. edulis chilensis* (Toro, 1998), *M. galloprovincialis chilensis* (Cárcamo et al., 2005), *M. galloprovincialis* (Toro et al., 2005), *M. edulis platensis* (Borsa et al., 2012), *M. galloprovincialis* lineage of Southern hemisphere origin (Westfall and Gardner, 2013), *M. planulatus* and *M. platensis* (Astorga et al., 2015). These different taxonomic designations may have arisen from a lack of clarity about which mussel species occur at specific sites, and because the number and type of markers used in each research project are different (Borsa et al., 2012; Larraín et al., 2015). Despite the lack of agreement about the taxonomic status of the native Chilean mussel, the term *M. chilensis* has long been employed on food product labels and in scientific articles (Ouagajjou et al., 2011; Astorga, 2014; Larraín et al., 2014; Araneda et al., 2016; Oyarzún et al., 2016). The name is also used in aquaculture production statistics (FAO 2016) and on good aquaculture practice certifications (GAA 2013). However, as several authors have pointed out, the name has historically had no formal taxonomic standing. This controversy and the history of the discussion are reflected in changes listed in the Word Register of Marine Species (Horton, 2017).

Mytilus species in Chile, both native and introduced, have been studied extensively using nuclear (RFLP-PCR *Me 15-16*, *mac*-1, *Glu*-5, *ITS*) and/or mitochondrial (*16s rDNA* RFLP, *COI*, *COIII*) molecular markers (Toro, 1998; Toro et al., 2005; Gérard et al., 2008; Westfall and Gardner, 2010; Fernandez-Tajes et al., 2011; Borsa et al, 2012; Larrain et al., 2012; Tarifeño et al., 2012; Westfall and Gardner, 2013). The primary problem has been that these markers target different regions of the genome that have different evolutionary rates, giving non-equivalent results when a few are used simultaneously (Rawson et al., 1999; Kijewski et al., 2011; Zbawicka et al., 2012). Whilst the mono-locus approach has the advantage of being easy to apply, it has the drawback that one locus represents a low power approach to delimit a species. A secondary problem arises from the fact that all smooth-shelled *Mytilus* taxa interbreed extensively (Michalek et al., 2016). Such hybridisation may complicate taxonomic resolution by blurring species boundaries, is not reflected in maternally inherited mtDNA markers that are often used for species identification, and also the evolutionary history of extensive interbreeding and introgression is not well represented by only a few nuclear DNA markers. What has been missing until recently is nuclear DNA markers that are specifies-specific and that can resolve ancestry when hybridisation occurs either naturally or as a result of anthropogenic transfer of mussel types to non-native regions. Recent advances in next generation sequencing methods, along with the increase in sequence data deposited in public databases, now permit the development of genome-wide single nucleotide polymorphism markers (SNPs) that may be applied in mussel studies to resolve many of these issues because they cover multiple regions of the genome (Zbawicka et al., 2012; Zbawicka et al., 2014; Saarman and Pogson, 2015; Araneda et al., 2016; Mathiesen et al., 2017).

Blue mussels have long been farmed in many regions of the world (Kijewski et al., 2009; Molinet et al., 2015). Chile is now the world's second largest mussel (family Mytilidae) aquaculture producer (FAO 2016). Production is concentrated in the Gulf of Reloncaví and along the coastline of Chiloé Island (Los Lagos region), and is based on the native blue mussel, nominally *Mytilus chilensis*. However, other species in Chile have also been reported, including *M. galloprovincialis* in the Gulf of Arauco – Bío-Bío region (Daguin and Borsa, 2000; Westfall and Gardner, 2010; Tarifeño et al., 2012; Astorga et al., 2015) and *M. edulis* in the Strait of Magellan – Magallanes region (Oyarzún et al., 2016), although this last species is named as *M. platensis* by Gaitán-Espitia et al. (2016). In addition, at a limited number of locations along the Chilean coast, alleles characteristic of *M. trossulus* (but not *M. trossulus* mussels themselves) have also been reported (Larrain et al., 2012; Oyarzún et al., 2016). Because accurate identification of mussel species produced by aquaculture is necessary for labelling, traceability, food security and marketing purposes (Larrain et al., 2012; 2014) it is important to understand which species is being grown where, which are native mussels, if native and introduced mussels interbreed, and if/how invasive mussels may affect aquaculture.

Hybridisation between *Mytilus* taxa has been reported in all geographic areas where two or more species co-exist (Gardner, 1997) and hybridisation patterns are complicated additionally by the occurrence of two mitochondrial lineages, female and male, and their recombination and introgression (Zbawicka et al., 2003; Filipowicz et al., 2008; Zbawicka et al., 2014). Significantly, a number of studies have now demonstrated that hybridisation of *Mytilus* taxa in aquaculture can cause unwanted or harmful effects to the industry and/or to native mussel populations (Dias et al., 2014; Crego-Prieto et al., 2015; Michalek et al., 2016).

In this paper we describe molecular genetic analyses of blue mussels from the coast of Chile, one of the last biogeographic regions without a clear and universally agreed

understanding of its mytilid taxonomy. Using a 49 SNPs panel against blue mussels from multiple locations in Chile and reference species, we test the hypothesis that the native Chilean blue mussel (*M. chilensis*) is an endemic species, with its own unique evolutionary history in the Southern hemisphere. Our aim is to contribute to the understanding of the taxonomic status of native smooth-shelled blue mussels and to test for the presence of other *Mytilus* taxa in Chile. This knowledge has practical applications by providing tools to solve issues related to aquaculture policy and management, and highlights the potential threats to native mussel populations and aquaculture posed by introduced taxa along the Chilean coast. The SNP panel has direct global uses in seafood labelling, traceability and food security, and regionally by setting the basis for a protected origin designation for native Chilean mussels.

2. MATERIALS AND METHODS

2.1 Sample collection and DNA extraction

Samples (n = 338) of *Mytilus* spp. were collected from nine locations in Chile spanning almost all of the native mussel's distributional range (~2500 km) and from five regions as reference samples: Pacific coast of Canada, Northern Ireland, Italy, New Zealand and Spain (Table 1, Fig. 1). Provisional species identification of each individual was determined using PCR of the non-repetitive region of the polyphenolic adhesive protein gene with the RFLP *Aci*I method with primers Me15-16 (Santaclara et al., 2006) or directly by genotyping the equivalent SNP locus BM151A (Gardner et al., 2016). Reference samples of *M. edulis*, *M. trossulus* and *M. galloprovincialis* were previously genotyped using SNPs (Zbawicka et al., 2012; Zbawicka et al., 2014).

2.2 SNP genotyping

In total, 338 *Mytilus* samples were genotyped using the Sequenom MassARRAY iPLEX genotyping platform (Gabriel et al., 2009). Assays were designed for 79 candidate SNPs, selected from 385 putative SNPs, which were tested on 300 specimens of *Mytilus* collected from geographic regions including Europe, North and South America, and New Zealand. Mussel SNPs were genotyped as previously described and following earlier testing of their reproducibility (Zbawicka et al., 2012; Zbawicka et al., 2014; Gardner et al., 2016).

2.3 Data analysis

Observed (H_O) and expected (H_E) heterozygosities and exact tests of departure from Hardy–Weinberg equilibrium (HWE) were determined using the R package adegenet 3.1-1 (Jombart 2008). Significance was determined by Markov chain Monte Carlo with 100,000 simulations, and the Benjamini and Yekutieli false discovery rate (FDR-BY) was used to correct significance (P) values after multiple testing (Benjamini and Yekutieli, 2011). Genetic differentiation amongst populations was determined using global and pairwise F_{ST} (theta) values (Weir and Cockerham, 1984) and the 95% confidence intervals for F_{ST} were estimated by bootstrapping with 10,000 replicates using the R package diveRsity (Keenan et al., 2013). The F_{ST} distance matrix was used to construct a neighbour-joining (NJ) tree illustrating the genetic relations of all populations, using POPTREEW (Takezaki et al., 2014).

Cluster analysis was performed using three unsupervised methods: (1) Discriminant Analysis of Principal Component (DAPC) performed with adegenet 3.1-1 (Jombart et al., 2010) where the number of clusters (K) was identified using the Bayesian Information Criteria (BIC). DACP variation was plotted using a K=14 in order to match mussel populations with clusters; (2) the non-parametric method implemented in AWclust (Gao and Starmer, 2008) that identifies the number of clusters based on a gap statistic; and (3) the

Bayesian-based method implemented in STRUCTURE with no prior information about the origin of individuals (Pritchard et al., 2000) assuming admixture and allowing for the correlation of allele frequencies between clusters. The tested number of clusters (K) ranged from 1 to the number of sampling locations plus 1. The length of burn-in period and the number of MCMC cycles after burn-in was 1,000,000 iterations each. Six runs were carried out for each K, and we used the Evanno et al. (2005) method to identify the single value of K which captures the uppermost level of structure in Structure Harvester (Earl, 2012).

Genetic assignment was performed to assign or exclude sampled populations as possible origins of individuals, using the frequency criteria of Paetkau et al. (1995) in a self-assignment test with the leave-one-out (LOO) procedure, implemented in GeneClass2.0 (Piry, 2004). In this supervised method, individuals were considered to be correctly assigned to their location of origin if the assignment probability to that group was higher than any other assignment probability to any other group.

To identify loci with high information content for individual assignment to species, different ranking criteria were tested: (1) F_{ST} outlier loci above the upper limit of the 95% confidence interval (CI 0.95) were identified by use of LOSITAN (Antao et al., 2008) with 1,000,000 simulations, a false discovery rate of 0.1, and a subsample size of 50; (2) loci with minor allele frequency (MAF) above 0.1, 0.2, 0.3 and 0.4 (Hess et al., 2011); and (3) the most informative loci selected with backward elimination locus selection (BELS) version 1.0 (Bromaghin, 2008). This program assesses the power of all loci and sequentially eliminates the locus that makes the smallest contribution to individual assignment performance, thereby providing a ranking order for all loci.

3 RESULTS

3.1 SNP markers, genetic diversity and Hardy-Weinberg equilibrium

Initially 79 SNPs were used to genotype 338 *Mytilus* individuals from the 14 locations. Of these, 24 did not provide an acceptable quality score, four were monomorphic in all samples, and two were tri-allelic. These 30 loci were removed from further analysis. The remaining 49 loci were bi-allelic, of which seven loci (BM11A, BM151A, BM202A, BM202B, BM203D, BM2G, BM92B) were monomorphic across all locations but for alternative alleles, eight loci (BM103B, BM10B, BM201B, BM30A, BM30C, BM44B, BM50B, BM75C) were polymorphic only in the *M. trossulus* (VACA) reference population, and three loci (BM32A, BM38B, BM9B) were polymorphic in mussels from all locations (Table S1, S2). MAF by locus ranged from 0.003 (BM75C) to 0.491 (BM101A) with a mean (±SD) across all loci of 0.181 ±0.144. In total, 60% of the SNP loci exhibited MAF values of < 0.2. Considering all populations globally, 19 of 49 loci departed from HWE (Table S1). The average observed heterozygosity (*Ho*) by locus ranged from <0.001 (BM106B, BM11A, BM151A, BM201B, BM202A, BM202B, BM203D, BM2G, BM92B) to 0.434 (BM32A).

3.2 Genetic differentiation amongst populations

Global genetic differentiation amongst the 14 populations was high, F_{ST} = 0.517 (95% CI of 0.505 - 0.523). All loci, with the exception of BM75C, had F_{ST} values significantly different from zero, and seven loci had F_{ST} = 1.000 (Table S1). Only the pairwise F_{ST} values amongst populations within the putative *M. chilensis* group and between the Mediterranean *M. galloprovincialis* sample (ORIT) and the Chilean Cocholgue sample (COCL) were not significantly different from zero according to their 95% CIs (Table S3). All other pairwise F_{ST} values were significantly different from zero, reaching a maximum value of 0.830 (VACA – WENZ). Very large genetic differentiation was observed amongst species

(estimated as the arithmetical mean of pairwise F_{ST} values from Table S3) between native mussels from Chile (QICL, LACL, PICL, ABCL, CBCL, GTCL, IPCL and PUCL locations) and the reference mussels *M. edulis* (LFGB), *M. galloprovincialis* (COCL, ORIT and CAES) and *M. trossulus* (VACA), 0.538, 0.555 and 0.805, respectively.

The NJ tree based on the pairwise F_{ST} distance matrix (Fig. 2) revealed four groups, showing high coincidence with species: (i) *M. trossulus* from Canada (VACA), (ii) *M. edulis* from Northern Ireland (LFGB), (iii) *M. galloprovincialis* including the New Zealand sample (WENZ), the Atlantic (CAES) and Mediterranean (ORIT) individuals and also the Chilean sample from Cocholgue (COCL), and (4) a group with 100% bootstrap support and with short internal branches that included all the other Chilean samples (here after the CLMch-mix group).

3.3 Population genetic structure

The DAPC identified three clusters (K=3) as shown by the elbow in the curve of BIC values versus K (Fig. S1a). In this scenario, all the VACA samples (*M. trossulus*) were clearly separated from the second group formed by the CLMch-mix (*M. chilensis*) and from the third cluster including the LFGB (*M. edulis*), CAES, ORIT, COCL and WENZ (*M. galloprovincialis*) samples. Only a single QICL individual was included in the *M. galloprovincialis – M. edulis* group. Exactly the same clustering results were apparent from the AWclust analysis, with K=3 determined from the gap statistic (Fig. S1b, Table S4). Both methods correctly matched 99.7 % of the individuals (337 of 338 mussels) to the three groups.

The DAPC plot of variation considering all locations (K=14), also revealed the three major groups as previously detected using K=3 (Fig. 3). As clusters are abstract objects that are not necessarily coincident with sampling sites, to draw the plot, the colour and shape of

the symbols used for all individuals included in the cluster were given by the location that contributed highest number of individuals to that cluster. The number of individuals from each location in each of the 14 clusters is presented in Table 2. Seven clusters included exclusively CLMch-mix (*M. chilensis*) individuals represented by red filled circles in the plot; these were grouped together in Fig. 3. All VACA samples (*M. trossulus*) were clearly separated into one cluster, represented by filled black squares. A third group included individuals from WENZ plotted as circles, as well as LFGB (*M. edulis*) individuals identified by squares and CAES, ORIT, COCL (*M. galloprovincialis*) represented as triangles. In this last group the mussels from Spain – Atlantic (CAES) were divided in two clusters, the first one containing only CAES individuals and the second one containing mussels from CAES but also from ORIT (Mediterranean *M. galloprovincialis*) and Northern Ireland (LFGB) individuals (*M. edulis*). The Cocholgue (COCL) mussels were mainly clustered with the ORIT individuals, but because they were few in number in each cluster, no red triangles are shown in Fig. 3. Overall, these results indicate that *Mytilus* in Chile are composed of two groups: (1) a northern group located in the Gulf of Arauco (Bío-Bío region) represented here by the COCL population, presumptively being introduced *M. galloprovincialis* and (2) a southern group of native Chilean blue mussels (*M. chilensis*).

The Bayesian clustering algorithm STRUCTURE identified three clusters (K=3) as capturing the highest level of structure, separating *M. trossulus, M. chilensis* and placing together in a single group the reference *M. edulis* population (LFGB) with the two European reference *M. galloprovincialis* populations of Italy (ORIT) and Spain (CAES), as well as with the populations of Cocholgue in Chile (COCL) and Wellington in New Zealand (WENZ) (Figs. S1c and 4a). For 337 individuals, the Q-values (cluster membership assignment estimates) were greatest in the cluster that coincides with the species, with the exception of one QICL individual that showed a higher Q value in the *M. edulis – M. galloprovincialis*

cluster (Table S4). The maximum Q-values per individual were > 0.8, so they may be considered as group members, with the exception of two individuals, one from QICL and the other from LFGB, that were considered potentially admixed. A second peak in the Delta K plot identified seven clusters (K=7) that separated into 5 groups, *M. trossulus* (VACA), *M. chilensis* individuals (CLMch-mix – composed of three clusters), *M. edulis* (LFGB), *M. galloprovincialis* (ORIT, CAES and COCL), and New Zealand (WENZ) (Figs. S1c and 4b).

For the 49 SNP panel, the LOO method (Piry et al. 2004) correctly assigned 337 of 338 individuals (99.7%) to species. In total, 100% of the COCL, ORIT, CAES and WENZ individuals were assigned to the *M. galloprovincialis* group (Table 3a). As well, 100% of the LFGB and VACA individuals were assigned correctly to the *M. edulis* and *M. trossulus* groups, respectively. Only one individual (0.5 %) from the CLMch-mix group was assigned to the *M. galloprovincialis* group, indicating a low rate of hybridisation (*M. chilensis* × *M. galloprovincialis*).

Assignment success to region of origin was 100% for the VACA and WENZ locations (Table 3b). Not one mussel of the Chilean COCL samples was assigned to the MchCL-All (Chilean) group, but instead all were assigned to the ORIT (90%) Italian (Mediterranean Sea) and the CAES (10%) Spain (Atlantic Ocean) locations. In addition, 7.4% of the LFGB samples were also assigned to CAES, whilst the rest of the LFGB mussels were assigned to the Northern Ireland location. The ORIT and CAES samples were assigned to both locations, with 93.1% and 86.2% of assignment success, respectively, showing the mixture of individuals between these two European locations. The mussels from VACA and WENZ were assigned with 100% accuracy to their sampling locations (Table 3b).

3.4 Highly informative locus panel

The ranking criteria selected different numbers of most informative loci, ranging from 6 to 25 (MAF>0.4 to MAF>0.1, respectively). The panel that performed best at assigning individuals to species was selected using the ranking criterion of MAF>0.2 and included 19 highly informative SNP loci that correctly assigned 336 (99.4%) individuals to species (Tables S5 and 4). Considering this MAF criterion along with F_{ST} outlier loci and results from BELS, three loci (BM106B, BM151A and BM6C) were included in the group of most informative loci (Fig. S2). Two of these three loci have known mRNA functions, with the polyphenolic adhesive foot protein (BM151A) and the elongation factor G (BM6C), whilst the function of BM106B is presently unknown.

4 DISCUSSION

Single nucleotide polymorphisms (SNPs) are powerful markers to monitor organisms at the individual, population and species levels. They have been developed recently for mussels of the genus *Mytilus* to study hybrid zones, introgression and adaptive genetic variation with traceability purposes (Zbawicka et al., 2014; Saarman and Pogson, 2015; Araneda et al., 2016). The availability of a SNP panel now allows for a multi-locus scan of the genome, adding more confidence to *Mytilus* species identification. Because not all SNP loci are equally informative based on their performance answering a specific research question, different ranking criteria have been employed to identify loci with high information content (Hess et al., 2011; Storer et al., 2012). These criteria can be used to create minimum panels that maximise individual assignment success to test hypotheses about species identification.

4.1 The taxonomic status of the native blue mussel in Chile

The question of the taxonomic status of the native Chilean blue mussel goes back more than 150 years. In the present study, the taxonomic priority and also the validity of the descriptions of the two South American species (*M. platensis* on the Atlantic coast, by d'Orbigny in 1846 and *M. chilensis* on the Pacific coast, by Hupé in 1854) are critical in assigning species designations. Based on morphological traits, d'Orbigny (1846) described the native blue mussel collected at Maldonaldo, Uruguay (Rio de la Plata) on the Atlantic coast of South America as *M. platensis*. In total, d'Orbigny described 17 *Mytilus* species from South America from his journey (1826-1833) via the Atlantic to the Pacific coast of South America. Significantly, d'Orbigny did not refer to or describe *M. platensis* (the Atlantic species) from Chile, despite the fact that he sampled in Chile. Subsequently, based on morphology, Hupé (1854) described *M. chilensis* from Concepcion, Chile on the Pacific coast. Thus, in accordance with the rules of taxonomic priority, *M. platensis* d'Orbigny (1846) holds for the native Atlantic smooth-shelled blue mussels, whereas *M. chilensis* Hupé (1854) holds for native Pacific smooth-shelled blue mussels. Interestingly, this situation is often not reflected in current taxonomic websites such as WoRMS and ITIS. Subsequently, but much later on, protein-based (allozyme) and morphometric assessments of blue mussels led McDonald et al. (1991) to conclude that mussels from South America (both coasts), the Falkland Islands and the Kerguelen Islands should be included tentatively in *M. edulis*. More recently, Borsa et al. (2012), in their review of allozyme and morphometric variation of Chilean blue mussels, confirmed that the Southern Hemisphere form of *M. edulis* occurs *"*along the shores from the North Patagonia region of Chile to the southern tip of the South American continent" (p. 52, Borsa et al., 2012) and concluded that native Chilean blue mussels should be assigned subspecific rank and named *M. edulis platensis* d'Orbigny 1846. Numerous authors, using a range of different molecular markers, have reported molecular

genetic differences between the native Chilean blue mussel and reference Northern hemisphere *M. edulis* and/or *M. galloprovincialis* (e.g., Gérard et al., 2008; Westfall & Gardner, 2010; Astorga et al. 2015; Śmietanka & Burzyński, 2017). Regardless of the taxonomic recommendation made by these (and several other) research groups, the common theme is that native *Mytilus* from South America, and often specifically from Chile, are genetically different from other smooth-shelled blue mussels anywhere in the world. These results challenge the interpretation of native Chilean mussels as being *M. edulis*-like and also of being like South American Atlantic mussels.

As our SNPs data reveals, there are pronounced nuclear genetic differences between the native blue mussels of Chile and all other mussels that we tested, including reference Northern hemisphere *M. edulis* (e.g., as employed by McDonald et al., 1991 and also by Borsa et al., 2012). Our analyses confirmed the hypothesis that the native Chilean blue mussel is genetically distinct from reference *M. edulis, M. galloprovincialis* and *M. trossulus*, to the extent that this mussel now warrants recognition as a separate and geographically isolated taxon within the genus *Mytilus.* We therefore conclude, based on taxonomic priority as discussed above, that the genetically distinct native blue mussel from Chile should be recognised as *M. chilensis* Hupé 1854. In addition, the SNPs results support the hypothesis that the native Chilean blue mussel has a unique evolutionary history in the Southern hemisphere. Similarly, Śmietanka and Burzyński (2017), who analysed the complete sequence of the female mitogenome of the native Chilean blue mussel, reported that the genetic distance between *M. edulis* and *M. galloprovincialis* (~2.5%) was half the distance that separates *M. chilensis* from either of these two species (5%), a result that indicates that the native Chilean mussel is a separate taxon within the genus *Mytilus.*

The use of SNPs in the present study has revealed that *M. trossulus* is the most differentiated of the taxa examined here, consistent with the suggestion that it is the oldest (ancestral) form (Kafanov, 1987; Vermeij, 1992; Chichvarkhin et al. 2000). Whilst the SNPs are able to differentiate amongst reference *M. edulis* (LFGB), *M. galloprovincialis* (Northern hemisphere – CAES, ORIT) and *M. galloprovincialis* (Southern hemisphere – WENZ), these three distinct taxa form a group more similar to one another than any of these are to *M. chilensis*. This highlights the power of a multi-locus approach in comparison with a single marker approach, when multiple loci are scattered throughout the genome to more accurately reflect lineage and/or species evolutionary histories. As noted elsewhere (e.g., Zbawicka et al., 2012; 2014; Gardner et al., 2016) the application of SNPs to global questions of taxonomy and hybridisation for mussels of the genus *Mytilus* will greatly enhance our knowledge of how many species exist, where they are found, and how they interact in terms of their interbreeding and introgression.

Beside *M. chilensis* and *M. galloprovincialis* no other *Mytilus* taxa were detected at Chilean sites in this study. Elsewhere, Oyarzún et al. (2016) reported *M. edulis* individuals in the Strait of Magellan (based on genotyping of the *Me15-16* locus), consistent with the western-most natural occurrence of *M. edulis*-like mussels or *M. edulis platensis* on the Atlantic coast of South America (e.g., McDonald et al., 1991; Hilbish et al., 2000; Borsa et al. 2012). In addition, *M. trossulus* alleles have been described in Chile in very low frequencies using the *Me15-16* marker (Larrain et al., 2012; Oyarzún et al., 2016). We found no evidence of either *M. edulis* or *M. trossulus* in our samples from Chile using the 49 SNP panel. In order to address the question of the taxonomic status of native blue mussels from the Atlantic Ocean coast of South America (*M. edulis, M. platensis* or another species) additional studies analysing mussels from this area (i.e., Argentina, Uruguay, Falkland Islands) will be needed.

4.2 The role of taxonomy in protecting and exploiting blue mussels

Systematics and taxonomy are important to delimit species, the fundamental units of biodiversity. The newly revealed SNPs-based difference between *M. chilensis* and all other studied taxa within the genus highlights the key role that molecular approaches to support taxonomy may play in disciplines such as conservation, aquaculture and food security.

From an ecological point of view, our results highlight the need for policies to mitigate the impact of or to prevent further expansion of non-indigenous species (NIS) along the Chilean coast. Specifically, the SNPs panel has confirmed several previous reports (e.g., Daguin and Borsa, 2000; Westfall and Gardner, 2010; Tarifeño et al., 2012; Astorga et al., 2015) of the presence of the highly invasive *M. galloprovincialis* in the Gulf of Arauco, but shows that the COCL sample is genetically more similar to the ORIT (Italy, Mediterranean Sea) *M. galloprovincialis* sample than to the CAES (Spain, Atlantic Ocean) *M. galloprovincialis* individuals. This strongly suggests a possible recent introduction from the Mediterranean Sea via human-mediated activities.

Introductions of alien species are one of the most important environmental issues today (Ojaveer et al., 2015). Considering the differences between *M. chilensis* and the other commercial blue mussel species revealed in this study, the potential biosecurity risk posed by the anthropogenic introduction and spread of *M. galloprovincialis* needs further attention. Given that the Mediterranean mussel is listed amongst the hundred worst invasive species, and taking into account the dispersal capacity of this invasive mussel (McQuaid and Phillips, 2000; Branch and Steffani, 2004), for the Chilean aquaculture sector the spread of *M. galloprovincialis* to other mussel production area poses a threat to the native Chilean blue mussel. This has led local producers to express concerns about a negative effect on production of the native *M. chilensis*. In Chile, protection and control measures to avoid the introduction of marine species that constitute pests, to isolate their presence on occurrence,

and to prevent their spread and promote their eradication, are regulated by the general law of fisheries and aquaculture (Chilean law Nº 18892, www.leychile.cl/Navegar?idNorma=30265). Although *M. galloprovincialis* is not considered to be a pest under Chilean law, this species cannot be imported live into the country and can be cultivated only at an experimental scale in the Bío-Bío region (regulation Nº 96 2015, Ministry of Economy, Development and Tourism, www.leychile.cl/Navegar?idNorma=256174). As with all such cases worldwide, once established in Chile it is highly unlikely that an invasive mussel such as *M. galloprovincialis* can be eradicated. The only future option appears to be monitoring its distribution with molecular tools, such as those used here, and to take action directed to containing it within limited areas.

Hybridisation between the NIS and native taxa may also occur, incorporating exotic DNA/genes into local populations and thereby modifying localised gene pools by introgression with unpredictable consequences over time, which may vary from cryptic to substantial changes in ecosystem structure and function. Only one individual (0.53 %) from the Mch-mix group was assigned to *M. galloprovincialis*, indicating that at the time of sample collection (2009) only a limited dispersal of this species had occurred from the Gulf of Arauco to the *M. chilensis* aquaculture zone in the south. Detailed investigation and ongoing monitoring is required to check for further spread and interaction between the two taxa. From an economic point of view, in addition to the unpredictable consequences that introgression can have on productivity of mussel farms (Michalek et al., 2016), the recognition of the name *M. chilensis* to designate the native Chilean blue mussel will contribute to traceability, authenticity and compliance with seafood labelling regulations, promoting transparency in the seafood trade. This recognition also provides an opportunity to the local mussel industry to apply for a geographically protected origin indication for the Chilean blue mussel.

4.3 Highly informative loci

The best ranking criterion to select the most informative loci for species identification was minor allele frequency (MAF) > 0.2. This selection criterion includes 19 loci that showed a power comparable to the total panel of 49 SNPs for species identification. In practical terms, these 19 loci constitute a reduced multi-locus SNP panel with high performance that permits identification of commercial *Mytilus* taxa without the inconsistencies associated with the use of a single or a few molecular markers. The performance of the reduced SNP panels selected with the MAF and BELS criteria, to identify mussel geographical origin region, was always lower than for species identification, and shows a small increment only when F_{ST} outlier loci were included (Table 4). This finding is probably explained by the fact that loci used to perform species identification must ideally be variable amongst species but fixed amongst populations within species. On the other hand, loci for identification of geographical origin must be more variable amongst populations within species, because they can reflect adaptation to local conditions, as demonstrated recently by Araneda et al. (2016).

5 CONCLUSIONS

A new multi-locus SNP marker panel reveals a clear separation between *M. chilensis* and all other *Mytilus* taxa, a degree of separation that is greater than that between *M. edulis* and *M. galloprovincialis*. Because *M. edulis* and *M. galloprovincialis* are already considered different species, this finding provides conclusive evidence of the status of smooth-shelled blue mussels native to the Chilean coast as a species within the *Mytilus* genus. Taxonomic priority indicates that these mussels should be recognised as *Mytilus chilensis* Hupé, 1854. The SNP markers also corroborate the previously reported presence of the highly invasive *M. galloprovincialis* in the Gulf of Arauco, revealing its probable recent introduction from the Mediterranean Sea. Recognition of this invasion poses challenges in order to avoid

environmental and economic damage in Chile. The 49 SNP panel was able to assign 99.7% of individuals correctly to species, whilst a comparable (99.4%) assignment success was obtained with a reduced panel of the 19 most informative SNP loci. These SNPs are a particularly valuable tool in terms of increasing our understanding of (1) *Mytilus* phylogeography, population genetics and connectivity, (2) elucidating evolutionary processes such as natural hybridisation and introgression, (3) helping to enforce aquaculture and conservation policies, and (4) increasing transparency in seafood labelling, authenticity and traceability field.

ACKNOWLEDGEMENTS

This research was funded in part by the 2011/01/B/NZ9/04352 NCN project to R.W. and statutory topic IV in the IOPAS, and CONICYT grants FONDECYT 1130302 and FONDEF IDeA ID16I10013 to M.A.L. The authors thank Eugenio Yokota, Julio Bañados, Marcela Cárcamo, Armando Salinas and José Villarroel for in collecting Chilean mussel samples and Dr Natalia Lam for assistance in a figure preparation. The authors declare no conflict of interest.

DATA ARCHIVING STATEMENT

The sequences of SNP originated mRNA fragments are available from GenBank (accession numbers KJ871032 - 74, KT713368 – 88). Allele frequencies of SNP loci and clusters identified for samples of *Mytilus* are presented in Supporting information.

REFERENCES

Ab Rahim, E. S, Nguyen, T. T. T, Ingram, B., Riginos, C., Weston, K. J., & Sherman, C. D. H. (2016). Species composition and hybridisation of mussel species (Bivalvia: Mytilidae) in Australia. *Marine and Freshwater Research*, 67, 1955-1963.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. 2008. LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, 9, 323.

Araneda, C., Larraín, M. A., Hecht, B., & Shawn, N. (2016). Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecology and Evolution*, 6, 3632-3644.

Astorga, M. P. (2014). Genetic considerations for mollusc production in aquaculture: current state of knowledge. *Frontiers in Genetics*, 5 (435), 1-6.

Astorga, M. P., Cardenas, L., & Vargas, J. (2015). Phylogenetic approaches to delimit genetic lineages of the *Mytilus* complex of South America: How Many Species are there? *Journal of Shellfish Research*, 34, 919-930.

Avise, J. C. (2000). *Phylogeography: the history and formation of species*: Harvard University Press.

Benjamini, Y., & Yekutieli, D. (2011). The control of the False Discovery Rate in multiple testing under dependency. *The Annals of Statistics*, 29, 1165-1188.

Borsa, P., Rolland, V., & Daguin-Thiébaut, C. (2012). Genetics and taxonomy of Chilean smooth-shelled mussels, *Mytilus* spp. (Bivalvia: Mytilidae). *Comptes Rendus Biologies*, 335, 51-61.

Branch, G. M., & Steffani, C. N. (2004). Can we predict the effects of alien species? A case-history of the invasion of South Africa by *Mytilus galloprovincialis* (Lamarck). *Journal of Experimental Marine Biology and Ecology*, 300, 189-215.

Bromaghin, J. F. (2008). BELS: backward elimination locus selection for studies of mixture composition or individual assignment. *Molecular Ecology Resources*, 8, 568-571.

Cárcamo, C., Comesaña, Á. S., Winkler, F. M., & Sanjuan, A. (2005). Allozyme identification of mussels (Bivalvia: *Mytilus*) on the pacific coast of South America. *Journal of Shellfish Research*, 24, 1101-1115.

Chichvarkhin, A. Yu., Kartavtsev, Yu. F., & Kafanov, A. I. 2000. Genetic relationships among some species of Mytilidae (Mollusca: Bivalvia) from the northern Pacific Ocean. *Russian Journal of Genetics, 36, 1003–1016. Translated from Genetika*, 36, 1206–1220.

Crego-Prieto, V., Ardura, A., Juanes, F., Roca, A., Taylor, J. S., & Garcia-Vazquez, E. (2015). Aquaculture and the spread of introduced mussel genes in British Columbia. *Biological Invasions*, 17, 2011-2026.

Daguin C., Borsa P. (2000) Genetic relationships of *Mytilus galloprovincialis* Lmk. populations worldwide: evidence from nuclear- DNA markers, Geol. Soc. Lond. Spec. Publ. 177, 389-397.

Dias, P. J., Fotedar, S., & Snow, M. (2014). Characterisation of mussel (*Mytilus* sp.) populations in Western Australia and evaluation of potential genetic impacts of mussel spat translocation from interstate. *Marine and Freshwater Research*, 65, 486-496.

E.U. (2013). Regulation (E.U.) No 1379/2013 of the European Parliament and of the Council on the Common Organisation of the Markets in fishery and aquaculture products, amending Council Regulations (EC) No 1184/2006 and (EC) No 1224/2009 and repealing Council Regulation (EC) No 104/2000. In *Official Journal of the European Union*, edited by E. Parliament: Official Journal of the European Union.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359-361.
Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620.
FAO. (2016). Global Statistical Collections.
FDA. (2011). US Food Drug Administration. Food safety modernization act (FSMA). In *Public Law*.
Fernandez-Tajes, J., Longa, A., Garcia-Gil, J., Chiu, Y. W., Huang, Y. S., Mendez, J. & Lee, R. S. (2011). Alternative PCR-RFLP methods for mussel *Mytilus* species identification. *European Food Research and Technology*, 233, 791-796.
Ferreira, J. G., & Bricker, S. B. (2016). Goods and services of extensive aquaculture: shellfish culture and nutrient trading. *Aquaculture International*, 24, 803-825.
Filipowicz, M., Burzyński, A., Śmietanka, B., & Wenne, R. (2008). Recombination in mitochondrial DNA of European mussels *Mytilus*. *Journal of Molecular Evolution*, 67, 377-388.
GAA. 2013. Best aquaculture practices certification. *Mussel farms. BAP Standards, Guidelines. USA: Global Aquaculture Alliance*.
Gabriel, S., Ziaugra, L., & Tabbaa, D. (2009). SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics*, 60, 2.12.1−12.12.18
Gaitán-Espitia, J. D., Quintero-Galvis, J. F., Mesas, A., & D'Elía, G. (2016). Mitogenomics of southern hemisphere blue mussels (Bivalvia: Pteriomorphia): Insights into the evolutionary characteristics of the *Mytilus edulis* complex. *Scientific Reports* 6, 26853.
Gao, X. & Starmer, J. (2008). AWclust: point-and-click software for non-parametric population structure analysis. *BMC Bioinformatics*, 9, 77.
Gardner, J. P. A., Zbawicka, M., Westfall, K. M., & Wenne, R. (2016). Invasive blue mussels threaten regional scale genetic diversity in mainland and remote offshore locations: the need for baseline data and enhanced protection in the Southern Ocean. *Global Change Biology*, 22, 3182-3195.
Gardner, J. P. A. (1997). Hybridization in the sea. *Advances in Marine Biology*, 31, 1-78.
Gérard, K., Bierne, N., Borsa, P., Chenuil, A., & Féral, J. P. (2008). Pleistocene separation of mitochondrial lineages of *Mytilus* spp. mussels from Northern and Southern Hemispheres and strong genetic differentiation among southern populations. *Molecular Phylogenetics and Evolution*, 49, 84-91.
Hess, J. E., Matala, A. P., & Narum, S. R. (2011). Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River Basin. *Molecular Ecology Resources*, 11, 137-149.
Hilbish, T. J., Mullinax, A., Dolven, S. I., Meyer, A., Koehn, R. K., & Rawson, P. D. (2000). Origin of the antitropical distribution pattern in marine mussels (*Mytilus* spp.): routes and timing of transequatorial migration. *Marine Biology*, 136, 69-77.
Horton, T., Kroh, A., Bailly, N., Boury-Esnault, N., Brandão, S. N., Costello, M. J., … Zeidler, W. (2017). World Register of Marine Species (WoRMS). WoRMS Editorial Board. Retrieved from http://www.marinespecies.org
Jilberto, F., Araneda, C., & Larraín, M. A. (2017). High resolution melting analysis for identification of commercially-important *Mytilus* species. *Food Chemistry*, 229, 716-720.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24, 1403-1405.

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

Kafanov, A.I. (1987). The subfamily Mytilinae Rafinesque 1815 (Bivalvia, Mytilidae) in the Caenozoic of the northern Pacific. In N.M. Baranovskaya, T.S. Beniaminson, A.I. Kafanov, V.G. Chavtur, (Eds.), *Fauna and distribution of molluscs: northern Pacific and the Polar Basin*, (pp. 65-103). Acad. Sci. USSR, Far-East Science Centre, Vladivostok (In Russian).

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, 4, 782-788.

Kijewski, T., Śmietanka, B., Zbawicka, M., Gosling, E., Hummel, H., & Wenne, R. (2011). Distribution of *Mytilus* taxa in European coastal areas as inferred from molecular markers. *Journal of Sea Research*, 65, 224-234.

Kijewski, T., Wijsman, J. W. M., Hummel, H., & Wenne, R. (2009). Genetic composition of cultured and wild mussels *Mytilus* from The Netherlands and transfers from Ireland and Great Britain. *Aquaculture*, 287, 292-296.

Koehn, R. K. (1991). The genetics and taxonomy of species in the genus *Mytilus*. *Aquaculture*, 94, 125-145.

Larraín, M.A., Diaz, N. F., Lamas, C., Vargas, C. & Araneda, C. (2012). Genetic composition of *Mytilus* species in mussel populations from southern Chile. *Latin American Journal of Aquatic Research*, 40, 1077-1084.

Larraín, M. A., Díaz, N. F., Lamas, C., Uribe, C., Jilberto, F., & Araneda, C. (2015). Heterologous microsatellite-based genetic diversity in blue mussel (*Mytilus chilensis*) and differentiation among localities in southern Chile. *Latin American Journal of Aquatic Research*, 43, 998-1010.

Larraín, M. A., Díaz, N. F., Lamas, C., Uribe, C., & Araneda, C. (2014). Traceability of mussel (*Mytilus chilensis*) in southern Chile using microsatellite molecular markers and assignment algorithms. Exploratory survey. *Food Research International*, 62, 104-110.

Lee, J. E. & Chown, S.L. (2007). *Mytilus* on the move: transport of an invasive bivalve to the Antarctic. *Marine Ecology Progress Series*, 339, 307-310.

Mace, G. M. (2004). The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359, 711-719.

Mathiesen, S. S., Thyrring, J., Hemmer-Hansen, J., Berge, J., Sukhotin, A., Leopold, P., Bekaert, M., Sejr, M. K., & Nielsen, E. E. (2017). Genetic diversity and connectivity within *Mytilus* spp. in the subarctic and Arctic. *Evolutionary Applications*, 10, 39-55.

McDonald, J. H., Seed, R., & Koehn, R. K. (1991). Allozymes and morphometric characters of three species of *Mytilus* in the Northern and Southern Hemispheres. *Marine Biology*, 111, 323-333.

McQuaid, C. D. & Phillips, T. E. (2000). Limited wind-driven dispersal of intertidal mussel larvae: in situ evidence from the plankton and the spread of the invasive species *Mytilus galloprovincialis* in South Africa. *Marine Ecology Progress Series*, 201, 211-220.

Michalek, K., Ventura, A., & Sanders, T. (2016). *Mytilus* hybridisation and impact on aquaculture: A minireview. *Marine Genomics*, 27, 3-7.

Molinet, C., Díaz, M., Arriagada, C., Cares, L., Marín, S., Astorga, M., & Niklitschek, E. (2015). Spatial distribution pattern of *Mytilus chilensis* beds in the Reloncaví fjord: hypothesis on associated processes. *Revista chilena de historia natural*, 88, 11.

Morin, P. A., Luikart, G., & Wayne R. K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19, 208-216.

Ojaveer, H., Galil, B. S., Campbell, M. L., Carlton, J. T., Canning-Clode, J., Cook, E. J., Davidson, A. D., Hewitt, C. L., Jelmert, A. & Marchini A. (2015). Classification of non-indigenous species based on their impacts: considerations for application in marine management. *PLoS Biol*, 13, e1002130.

Ouagajjou, Y., Presa, P., Astorga, M, & Pérez, M. (2011). Microsatellites of *Mytilus chilensis*: A genomic print of its taxonomic status within *Mytilus* sp. *Journal of Shellfish Research*, 30, 325-330.

Oyarzún, P. A., Toro, J. E., Cañete, J. I., & Gardner, J.P.A. (2016). Bioinvasion threatens the genetic integrity of native diversity and a natural hybrid zone: smooth-shelled blue mussels (*Mytilus spp*.) in the Strait of Magellan. *Biological Journal of the Linnean Society*, 117, 574-585.

Paetkau, D., Calvert, W., Stirling, I., & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 4, 347-354.

Pereira, F., Carneiro, J., & Amorim A. (2008). Identification of species with DNA-based technology: current progress and challenges. *Recent patents on DNA & gene sequences*, 2, 187-200.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., & Estoup A. (2004). GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95, 536-539.

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

Rawson, P.D., Agrawal, V., & Hilbish, T.J. (1999). Hybridization between the blue mussels *Mytilus galloprovincialis* and *M. trossulus* along the Pacific coast of North America: evidence for limited introgression. *Marine Biology*, 134, 201-211.

Riginos, C., & Cunningham, C. W. (2005). Local adaptation and species segregation in two mussel (*Mytilus edulis* × *Mytilus trossulus*) hybrid zones. *Molecular Ecology*, 14, 381-400.

Saarman, N. P. & Pogson, G. H. (2015). Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology*, 24, 4723-4738.

Santaclara, F. J., Espineira, M., Cabado, G., Aldasoro, A., Gonzalez-Lavin, N., & Vieites, J. M. (2006). Development of a method for the genetic identification of mussel species belonging to *Mytilus*, *Perna*, *Aulacomya*, and other genera. *Journal of Agricultural and Food Chemistry*, 54, 8461-8470.

Seddon, J.M., Parker, H.G., Ostrander, E.A., & Ellegren, H. (2005). SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology*, 14, 503-511.

Shaw, J. D., Terauds, A., Riddle, M. J., Possingham, H. P., & Chown, S. L. (2014). Antarctica's protected areas are inadequate, unrepresentative, and at risk. *PLoS Biol*, 12, e1001888.

Śmietanka, B., & Burzyński A. (2017). Complete female mitochondrial genome of *Mytilus chilensis*. *Mitochondrial DNA Part B*, 2, 101-102.

Storer, C. G, Pascal, C. E., Roberts, S. B., Templin, W. D., Seeb, L., W., & Seeb, J. E. (2012). Rank and order: Evaluating the performance of SNPs for individual assignment in a non-model organism. *Plos One*, 7, e49018.

Takezaki, N., Nei, M., & Tamura, K. (2014). POPTREEW: Web Version of POPTREE for Constructing Population Trees from Allele Frequency Data and Computing Some Other Quantities. *Molecular Biology and Evolution*, 31, 1622-1624.

Tarifeño, E., Galleguillos, R., Llanos-Rivera, A., Arriagada D., Ferrada, S., Canales-Aguirre, C. B., & Seguel, M., (2012). Erroneous identification of the mussel, *Mytilus galloprovincialis* (Lamarck 1819) as the species, *Mytilus chilensis* (Hupe 1854) in the Bay of Concepción, Chile. *Gayana (Concepción)*, 76, 167-172.

Toro, Jorge E. (1998). PCR-based nuclear and mtDNA markers and shell morphology as an approach to study the taxonomic status of the Chilean blue mussel, *Mytilus chilensis* (Bivalvia). *Aquatic Living Resources*, 11, 347-353.

Toro, J. E., Ojeda, J. A., Vergara, A. M., Castro, G. C., & Alcapán, A. C. (2005). Molecular characterization of the chilean blue mussel (*Mytilus chilensis* Hupe 1854) demonstrates evidence for the occurrence of *Mytilus galloprovincialis* in southern Chile. *Journal of Shellfish Research*, 24, 1117-1121.

Vermeij, G. J. 1992. Trans-equatorial connections between biotas in the temperate eastern Atlantic. *Marine Biology*, 112, 343-348.

Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population-structure. *Evolution*, 38, 1358-1370.

Wenne, R., Bach L., Zbawicka M., Strand J., & McDonald, J.H. (2016). A first report on coexistence and hybridization of *Mytilus trossulus* and *M. edulis* mussels in Greenland. *Polar Biology*, 39, 343-355.

Wennerstrom, L., Laikre, L., Ryman, N., Utter, F.M., Ab Ghani, N.I., Andre, C., DeFaveri, J., Johansson, D., Kautsky, L., Merila, J., Mikhailova, N., Pereyra, R., Sandstrom, A., Teacher, A.G.F., Wenne, R., Vasemagi, A., Zbawicka, M., Johannesson, K., Primmer, C.R. (2013). Genetic biodiversity in the Baltic Sea: species-specific patterns challenge management. *Biodiversity and Conservation*, 22, 3045-3065.

Westfall, K. M., & Gardner, J. P. A. (2010). Genetic diversity of Southern hemisphere blue mussels (Bivalvia: Mytilidae) and the identification of non-indigenous taxa. *Biological Journal of the Linnean Society*, 101, 898-909.

Westfall, K. M., & Gardner, J. P. A. (2013). Interlineage *Mytilus galloprovincialis* Lmk. 1819 hybridization yields inconsistent genetic outcomes in the Southern hemisphere. *Biological Invasions*. 15, 1493-1506.

Zbawicka, M., Drywa, A. Śmietanka, B., & Wenne, R. (2012). Identification and validation of novel SNP markers in European populations of marine *Mytilus* mussels. *Marine Biology*, 159, 1347-1362.

Zbawicka, M., Sanko, T., Strand, J., & Wenne, R. (2014). New SNP markers reveal largely concordant clinal variation across the hybrid zone between *Mytilus* spp. in the Baltic Sea. *Aquatic Biology*, 21, 25-36.

Zbawicka, M., Wenne R., & Burzyński, A. (2014). Mitogenomics of recombinant mitochondrial genomes of Baltic Sea Mytilus mussels. *Molecular Genetics and Genomics*, 289, 1275-1287.

Zbawicka, M., Wenne, R., & Skibinski, D. (2003). Mitochondrial DNA variation in populations of the mussel *Mytilus trossulus* from the Southern Baltic. *Hydrobiologia*, 499, 1-12.

SUPPORTING INFORMATION

Additional Supporting Information for this article is presented online in the tab.

Figure titles:

FIGURE 1 Locations and codes of the 14 *Mytilus* sampling sites. Four letter code of locations can be found in Table 1. Colour indicates species according to the PCR-RFLP Me15-16, *Aci*I assay. (•) *Mytilus chilensis,* (•) *Mytilus galloprovincialis,* (•) *Mytilus edulis,* (•) *Mytilus trossulus.* Background topographic map from GeoMapApp (http:// www.geomapapp.org).

FIGURE 2 Neighbour-joining tree of *Mytilus* populations based on F_{ST} distance matrix.

FIGURE 3 Clusters obtained by Discriminant Analysis of Principal Components of 338 *Mytilus* individuals from 14 locations (K=14) along with Discriminant Analysis eigenvalues retained (left) and PCA eigenvalues retained (right). Different shapes represent species according to their genotype assays using Me15-16 or SNP BM151A (• Southern hemisphere mussels, including *M. chilensis* (Chile) and Southern Hemisphere *M. galloprovincialis* WENZ (New Zealand) samples, ▲ Northern Hemisphere *M. galloprovincialis*, □ *M. edulis,* ■ *M. trossulus*). Each symbol in the graph represents an individual. Colours represent the sampling location: in red - Chile (CLMch-mix), in black - Pacific coast of Canada (VACA), in blue - Northern Ireland (LFGB), in yellow - Italy (Mediterranean - ORIT), in gold - Spain (Atlantic - CAES), in green - New Zealand (WENZ). As clusters are abstract objects that are not necessarily coincident with sampling sites, the colour of all individuals included in the cluster was given by the location that contributes the highest number of individuals. Detailed composition of individuals in each cluster is shown in Table 2.

FIGURE 4 Proportional membership (Q) of *Mytilus* individuals to each of the: **a)** three clusters (K=3) and **b)** seven clusters (K=7) inferred by STRUCTURE.

Table 1 Sample location, genotype designation based on Me 15-16 and SNP BM151A assays, number of individuals, sampling date and stage of development.

Code	Sampling location - Country Latitude / Longitude	Genotype Me 15-16	Genotype SNP BM151A	Nº of individuals	Sampling date	Stage of development
Chilean samples						
COCL	Cocholgue - Chile S: 35° 7'38.62" / W: 73°11'25.56"	126 / 126	GG	20	2007	Adult
QICL	Quillaipe - Chile S: 41º 32' 55.35" / W: 72º 46' 14.35"	126 / 126	TT	20	2009	Seed
LACL	Caleta La Arena - Chile S: 41º 41' 00.00"' / W: 72º 40' 18,92"	126 / 126	TT	27	2009	Seed
PICL	Pichicolo - Chile S: 42º 02' 23.76" / W: 72º 35' 27.17"	126 / 126	TT	30	2009	Seed
ABCL	Abtao - Chile S: 42°24'0.54"/ W: 74°10'48.49"	126 / 126	TT	17	2013	Adult
CBCL	Canal Coldita - Piedra blanca - Chile S: 43º 14' 48.82" / W: 73º 41' 42.77"'	126 / 126	TT	29	2009	Seed
GTCL	Golfo Trinidad - Chile S: 49°57'59.69"/ W: 75°11'12.99"	126 / 126	TT	5	2002	Adult
IPCL	Isla Peel - Chile S: 50º 50' 29.83" / W: 74º 00' 41.27'	126 / 126	TT	28	2009	Adult
PUCL	Punta Arenas - Chile S: 53° 9'16.12"/ W: 70°54'59.31"	126 / 126	TT	33	2012	Adult
Reference samples						
ORIT#	Oristano - Italy N: 39°47'59.88" /	126 / 126	GG	29	2004	Adult

	E: 8°31'9.72"					
CAES	Punta Camarinal - Spain N: 36° 4'48.01" / W: 5°47'58.00"	126 / 126	GG	29	2004	Adult
WENZ	Wellington - New Zealand S: 41°22'12.19" / E: 174°47'39.89"	126 / 126	TT	27	2005	Adult
LFGB	Lough Foyle – Northern Ireland N: 55° 5'35.50" / W: 7° 4'48.92"	180 / 180	GG	27	2006	Adult
VACA#	Vancouver - Canada N: 49°18'33.75"/ W: 123°49'15.41"	168 / 168	GG	17	2006	Adult
Total				338		

These samples were also used by Gardner et al. (2016).

Table 2. Clusters obtained by Discriminant Analysis of Principal Components of SNPs variation amongst 338 *Mytilus* individuals from 14 locations.

	•	•	•	•	•	•	•	□	■	△	△	△	△	•	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	
Location	CL-mix	CL-mix	CL-mix	CL-mix	CL-mix	CL-mix	CL-mix	LFGB	VACA	ORIT-COCL-others	ORIT-COCL	CAES-others	CAES	WENZ	**Total**
COCL	0	0	0	0	0	0	0	0	0	**13**	**6**	1	0	0	20
QICL	0	1	**3**	**5**	**5**	**1**	**4**	0	0	0	1	0	0	0	20
LACL	**3**	**4**	**7**	**3**	**2**	**4**	**4**	0	0	0	0	0	0	0	27
PICL	**6**	**2**	**4**	**1**	**6**	**9**	**2**	0	0	0	0	0	0	0	30
ABCL	**2**	**3**	**2**	**3**	**3**	**3**	**1**	0	0	0	0	0	0	0	17
CBCL	**2**	**4**	**8**	**7**	**2**	**5**	**1**	0	0	0	0	0	0	0	29
GTCL	**2**	**0**	**0**	**1**	**0**	**2**	**0**	0	0	0	0	0	0	0	5
IPCL	**6**	**1**	**1**	**4**	**0**	**7**	**9**	0	0	0	0	0	0	0	28
PUCL	**9**	**8**	**3**	**5**	**3**	**2**	**3**	0	0	0	0	0	0	0	33
LFGB	0	0	0	0	0	0	0	**25**	0	0	0	2	0	0	27
VACA	0	0	0	0	0	0	0	0	**17**	0	0	0	0	0	17
ORIT	0	0	0	0	0	0	0	0	0	**14**	**13**	2	0	0	29
CAES	0	0	0	0	0	0	0	0	0	4	0	**10**	**15**	0	29
WENZ	0	0	0	0	0	0	0	0	0	0	0	0	0	**27**	27
	30	23	28	29	21	33	24	25	17	31	20	15	15	27	**338**

Different shape symbols represent species according to their genotype assays using Me15-16 or SNP BM151A (• *M. chilensis* and WENZ sample (Southern Hemisphere M. galloprovincialis) ▲ Northern Hemisphere *M. galloprovincialis*, □ *M. edulis*, ■ *M. trossulus*). Each symbol in the graph represents an individual. Colours represent the sampling location: in red: Chile (CLMch-mix), in black: Canada (VACA), in blue: Northern Ireland (LFGB), in yellow: Italy (Mediterranean - ORIT), in gold: Spain (Atlantic - CAES), in green: New Zealand (WENZ). As clusters are abstract objects that are not necessarily coincident with sampling sites, the color of all individuals included in the cluster was given by the location that contribute with the high number of individuals.

Table 3. Number of individuals (%) correctly reassigned to A) species and B) region of origin by the leave-one-out procedure implemented in GeneClass2.0 (Piry et al. 2004) using the frequency criteria of Paetkau et al. (1995).

A)	**Assigned to species**			
Location	***M. chilensis***	***M. galloprovincialis***	***M. edulis***	***M. trossulus***
COCL	-	20 (100.0)	-	-
QICL	19 (95.0)	1 (5.0)	-	-
LACL	27 (100.0)	-	-	-
PICL	30 (100.0)	-	-	-
ABCL	17 (100.0)	-	-	-
CBCL	29 (100.0)	-	-	-
GTCL	5 (100.0)	-	-	-
IPCL	28 (100.0)	-	-	-
PUCL	33 (100.0)	-	-	-
ORIT	-	29 (100.0)	-	-
CAES	-	29 (100.0)	-	-
WENZ	-	27 (100.0)	-	-
LFGB	-	-	27 (100.0)	-
VACA	-	-	-	17 (100.0)

B)	**Assigned to region of origin**					
Location	**CLMch-mix**	**ORIT**	**CAES**	**WENZ**	**LFGB**	**VACA**
COCL	-	18 (90.0)	2 (10.0)	-	-	-
QICL	19 (95.0)	1 (5.0)	-	-	-	-
LACL	27 (100.0)	-	-	-	-	-
PICL	30 (100.0)	-	-	-	-	-
ABCL	17 (100.0)	-	-	-	-	-
CBCL	29 (100.0)	-	-	-	-	-
GTCL	5 (100.0)	-	-	-	-	-
IPCL	28 (100.0)	-	-	-	-	-
PUCL	33 (100.0)	-	-	-	-	-
ORIT	-	27 (93.1)	2 (6.9)	-	-	-
CAES	-	4 (13.8)	25 (86.2)	-	-	-
WENZ	-	-	-	27 (100.0)	-	-
LFGB	-	-	2 (7.4)	-	25 (92.6)	-
VACA	-	-	-	-	-	17 (100.0)

Table 4. Number of individuals (%) correctly reassigned to species and region of origin by panels composed of highly informative SNP loci selected by different ranking criteria.

Ranking criterion	**All**	**F_{ST} outlier loci** (1)	**Minor allele frequency (MAF)**				**BELS** (2)
			> 0.1	**> 0.2**	**> 0.3**	**> 0.4**	
Nº of selected loci	49	11	25	19	13	6	19
To species	337 (99.7)	262 (77.5)	336 (99.4)	336 (99.4)	335 (99.1)	331 (97.9)	317 (93.8)
To region of origin	309 (91.4)	267 (79.0)	307 (90.8)	307 (90.8)	297 (87.9)	269 (79.6)	300 (88.8)

(1) F_{ST} outlier loci (CI > 0.95)
(2) Loci whose removal caused the assignment performance measure to drop down 0.55 in the BELS software program (Bromaghin, 2008).