Contrasting Patterns of Neutral and Adaptive Genetic Variation of Chilean Blue Mussel (*Mytilus chilensis*) due to Local Adaptation and Aquaculture

Cristian Araneda[1], M. Angélica Larraín[2], Benjamin Hecht[3], Shawn Narum[3]

[1]Universidad de Chile, Departamento de Producción Animal. Santiago, Chile.
[2]Universidad de Chile, Departamento de Ciencia de los Alimentos y Tecnología Química. Santiago, Chile.
[3]Columbia River Inter-Tribal Fish Commission. Hagerman, Idaho, USA.

INTRODUCTION

Most marine species with pelagic larvae distributed over wide geographic areas by ocean currents have been thought to be panmictic with limited genetic structure. Chilean blue mussel (*Mytilus chilensis*) larvae have a planktonic stage of about 45 days, thereby they are potentially dispersed over large geographical areas by marine currents or human-mediated activities. This is an important commercial species distributed in Chile from Arauco (latitude 37ºS) to Cape Horn (latitude 55ºS).

Recent studies that have elucidated local adaptation patterns in marine environments, demonstrated the importance of studying adaptive variation (De Wit and Palumbi, 2013; Limborg *et al*. 2012; Sanford and Kelly, 2011). Genetic structure of *M. chilensis* has been explored using RAPD, allozyme and microsatellite markers, in populations from Arauco (37ºS) to Punta Arenas (53ºS), finding limited genetic structure and no evidence of discrete stocks (0.011≤ Global F_{ST} ≤ 0.055), with the possible exception of an austral population from Magallanes strait (53ºS) (Toro *et al*. 2006; Toro *et al*. 2004; Larraín *et al*. 2014).

Recent advances in sequencing technology has made possible to perform population genetic analyses in non-model species using thousands of SNPs markers obtained by genotyping-by-sequencing approach (Miller *et al*. 2007; Narum *et al*. 2013). These large panels of markers allow to test patterns of neutral and adaptive genetic variation related to population structure and local adaptation. High F_{ST} loci that are putative outliers have also been shown to be highly informative to determine the geographic origin of individuals (Ogden 2011).

OBJECTIVE

This study was designed to investigate patterns of neutral and adaptive genetic variation within Chilean blue mussel populations in order to identify a subset of putatively adaptive genetic markers to investigate the population structure and to improve the ability to trace individuals to their geographical origin, especially in the area with strong aquaculture activities (Reloncaví - Zone 1 and Chiloé Island - Zone 2).

METHODS

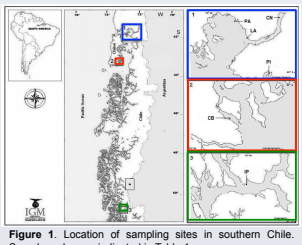

Figure 1. Location of sampling sites in southern Chile. Sample codes are indicated in Table 1.

Table 1. Number of samples, zones and geographic coordinates of sampling sites in southern Chile.

Code	Locations	Zone	N	South latitude West longitude
1-PA	Piedra Azul	1 Reloncaví	30	41º 32' 55,35'' 72º 46' 14,35''
1-LA	Caleta La Arena	1 Reloncaví	39	41º 41' 00,00'' 72º 40' 18,92''
1-CN	Canutillar	1 Reloncaví	33	41º 31' 13,90'' 72º 20' 15,69''
1-PI	Pichicolo	1 Reloncaví	25	42º 02' 23,76'' 72º 35' 27,17''
2-CB	Canal Coldita Piedra Blanca	2 Chiloé Island	31	43º 14' 48,82'' 73º 41' 42,77''
3-IP	Isla Peel	3 Magallanes	32	50º 50' 29,83'' 74º 00' 41,27''

Samples were collected from subtidal zones in 2009. DNA extractions and species identification were performed with standard methods (Larraín *et al*. 2012). Two RAD-tag libraries were prepared with *Sbf*I according to Hess *et al*. (2013) and sequenced on an Illumina® HiSeq1500 genetic analyzer (Illumina Inc., San Diego, CA, USA). 800 million reads were quality filtered (Phred 33) and trimmed to 75 pb. SNP discovery and genotyping was performed with STACKS pipeline 1.08 (Catchen *et al*. 2013). The quality filters applied to remove putative false SNPs were:

1. SNPs with genotyping success lower than 70% and minor allele frequency (MAF) less than 5% per population were removed, obtaining 4,305 SNPs for further steps.
2. SNPs that were not genotyped in all six populations (2,140 SNPs) were removed.
3. Only one bi-allelic SNP per stacks was allowed, so 865 SNPs were removed.
4. SNPs that showed significant deviations from HWE in three or more locations, using BY-FDR corrected critical level of 0.005238 (Narum, 2006) were also removed (60 SNPs).

After applying these filters combination, a total of **1,240** SNPs were genotyped in **190** individuals.

Outlier tests with LOSITAN (Antao et al. 2008) were applied to detect loci under directional selection (loci above upper limit of CI 99,5%) and neutral variation (loci inside CI 80%). Two scenarios were simulated, **Scenario 1** with all six locations and **Scenario 2** excluding the Magallanes location (Zone 3). These scenarios provided three sets of loci that were: Outlier loci (six locations), neutral loci (six locations) and outlier loci (five locations).

Global and pairwise F_{ST} were estimated with Genepop 4.2 (Rousset 2008). Discriminant Analysis of Principal Component (DAPC; Jombart and Ahmed 2011) with outlier loci from both scenarios along with assignment analyses with GeneClass2.0 (Piry et al. 2004) were performed to evaluate the population differentiation of Chilean blue mussels.

RESULTS

Outlier analyses identified 981 neutral loci and 58 loci as candidates for positive selection under scenario 1 (six populations) and 34 loci as candidates for positive selection under scenario 2 (five populations), with 17 outlier SNPs shared for both scenarios.

Global F_{ST} using the neutral panel was 0.0072. Adaptive panels (scenarios 1 and 2) showed global F_{ST} of 0.1139 and 0.0886, respectively. Pairwise F_{ST} are shown in Table 2 and 3.

Table 2. Pairwise F_{ST} using 58 SNPs outlier among six collection sites of *Mytilus chilensis* in southern Chile.

	1-PI	1-LA	1-CN	2-CB	3-IP
1-PA	0.0075	0.0052	0.0074	0.0804	0.2279
1-PI		0.0225	0.0079	0.0602	0.2166
1-LA			0.0174	0.1011	0.2484
1-CN				0.0820	0.2313
2-CB					0.2158

Table 3. Pairwise F_{ST} using 34 SNPs outlier among five collection sites of *Mytilus chilensis* in southern Chile.

	1-PI	1-LA	1-CN	2-CB
1-PA	0.0129	0.0170	0.0126	0.1553
1-PI		0.0344	0.0092	0.1178
1-LA			0.0286	0.1978
1-CN				0.1641

DAPC identified three clusters (k=3) in both scenarios. In Scenario 1, clusters matched to the three geographic zones analyzed: Reloncaví, Chiloé Island and Magallanes (Figure 1, Table 1 and Figure 2). In Scenario 2, one cluster correspond to Chiloé Island (Zone 2) and the other two clusters to Reloncaví area (Zone 1), but without concordance with sampling locations (Figure 3).

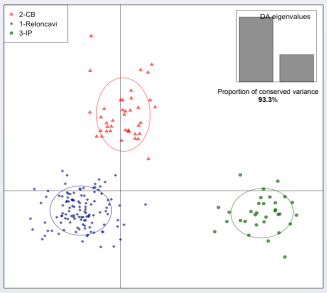

Figure 2. Clusters obtained with DAPC and 58 adaptive SNPs **(Scenario 1)** for *Mytilus chilensis* from southern Chile.

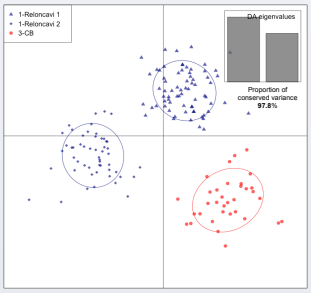

Figure 3. Clusters obtained with DAPC and 34 adaptive SNPs **(Scenario 2)** for *Mytilus chilensis* from southern Chile.

Table 4. Assignment of Chilean blue mussel individuals from southern Chile. obtained with both adaptive SNPs panels

	1-PA	1-PI	1-LA	1-CN	2-CB	3-IP
Scenario 1 Correct Assignment	7 (23%)	8 (32%)	17 (44%)	13 (39%)	26 (84%)	32 (100%)
Scenario 2 Correct Assignment	12 (40%)	9 (36%)	21 (54%)	12 (36%)	27 (87%)	

Assignment values were high for mussels from two of the three areas, with 100% from Magallanes correctly assigned, and 84-87% correct from Chiloé Island (Table 4). Assignment for the three sites within the Reloncaví area were only 42% on average, but higher for Reloncaví as a single reporting group (92%).

CONCLUSIONS

1. Panels of adaptive SNPs were useful to identify genetic structure and geographic origin of Chilean blue mussels in southern Chile.
2. Our analysis support genetic differentiation between the Magallanes population and northern populations (Reloncaví and Chiloé Island) possibly due to local adaptation and isolating effects of the Cape Horn current.
3. Chiloé Island was also highly distinct from the other two regions, which has not been previously detected with other molecular markers.
4. Reloncaví was distinct from the other regions, but the four sites within Reloncaví were not well differentiated. This suggests high gene flow within this aquaculture production region due to exchange of seed stocks among facilities.

REFERENCES

Antao T *et al*. (2008). *BMC Bioinformatics* **9**, 323 .
Catchen J *et al* (2013). *Molecular Ecology* **22**, 3124-3140.
De Wit P, Palumbi SR (2013) *Molecular Ecology* **22**, 2884-2897.
Hess JE *et al* (2013). *Molecular Ecology* **22**, 2898-2916.
Jombart T, Ahmed I (2011). *Bioinformatics* **27**, 3070-3071.
Larrain MA *et al* (2012) *Latin American J. of Aquaculture* **40**, 8.
Larrain MA *et al*. (2014). *Food Research International* **62**, 104-110.
Limborg MT *et al*. (2012) *Molecular Ecology* **21**, 3686-3703.
Miller M *et al*. (2007). *Genome Res* **17**, 240 - 248.
Narum S (2006). *Conservation Genetics* **7**, 783-787.
Piry S *et al*. (2004). *Journal of Heredity* **95**, 536-539
Rousset F (2008). *Molecular Ecology Resources* **8**, 103-106.
Sanford E, Kelly MW (2011) *Ann Rev of Marine Science* **3**, 509-535.
Toro JE *et al*. (2004). *Aquaculture Research* **35**, 1466-1471.
Toro JE *et al*. (2006). *Genetics and Molecular Biology* **29**, 174-179.

ACKNOWLEDGEMENTS

We thank Eugenio Yokota, Julio Bañados, Marcela Cárcamo, Armando Salinas, and José Villarroel for assistance in collecting mussel samples. We also thank Juan Vidal from Instituto Geográfico Militar of Chile, for assistance in the preparation of maps, and Andrea Bravo for DNA extraction and Amanda Matala for preparing RAD-seq libraries and excellent laboratory assistance.

This research was supported by CONICYT, FONDECYT/Regular N°1130302.