



# Microsatellite markers and cytoplasmic sequences reveal contrasting pattern of spatial genetic structure in the red algae species complex *Mazzaella laminarioides*

Marie-Laure Guillemin, Myriam Valero, Kennia Morales Collio, Ramona Pinochet Sanchez, Miguel Henríquez Espinosa, Andrea X. Silva

## ► To cite this version:

Marie-Laure Guillemin, Myriam Valero, Kennia Morales Collio, Ramona Pinochet Sanchez, Miguel Henríquez Espinosa, et al.. Microsatellite markers and cytoplasmic sequences reveal contrasting pattern of spatial genetic structure in the red algae species complex *Mazzaella laminarioides*. *Journal of Phycology*, 2016, 10.1111/jpy.12440 . hal-01333426

**HAL Id: hal-01333426**

**<https://hal.science/hal-01333426>**

Submitted on 17 Jun 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1   MICROSATELLITE MARKERS AND CYTOPLASMIC SEQUENCES REVEAL  
2   CONTRASTING PATTERN OF SPATIAL GENETIC STRUCTURE IN THE RED  
3   ALGAE SPECIES COMPLEX *MAZZAELLA LAMINARIOIDES*

4

5   Marie-Laure Guillemin

6   Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad  
7   Austral de Chile, Casilla 567, Valdivia, Chile.

8   CNRS, Sorbonne Universités, UPMC University Paris VI, UMI 3614, Evolutionary  
9   Biology and Ecology of Algae, Station Biologique de Roscoff, CS 90074, Place G.  
10   Tessier, 296888 Roscoff, France.

11   Email: marielaure.guillemin@gmail.com

12   Phone: +56 063 222 1703

13   Myriam Valero

14   CNRS, Sorbonne Universités, UPMC University Paris VI, UMI 3614, Evolutionary  
15   Biology and Ecology of Algae, Station Biologique de Roscoff, CS 90074, Place G.  
16   Tessier, 296888 Roscoff, France.

17   Kennia Morales Collio

18   Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad  
19   Austral de Chile, Casilla 567, Valdivia, Chile.

20   Ramona Pinochet Sanchez

21   Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad  
22   Austral de Chile, Casilla 567, Valdivia, Chile.

23   Miguel Henríquez Espinosa

24   Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad  
25   Austral de Chile, Casilla 567, Valdivia, Chile.

26   Andrea X. Silva

27   AUSTRAL-omics, Facultad de Ciencias, Universidad Austral de Chile, Casilla 567,  
28   Valdivia, Chile.

29

30   Running Title: Genetic discordance in *Mazzaella laminarioides*

Abstract: *Mazzaella laminarioides* (Bory) is a common haploid-diploid red alga that forms dense beds. This alga has a wide distribution range, covering 3,500km of the Chilean coast, but is restricted to high rocky intertidal zones. Recently, the existence of three highly divergent genetic lineages was demonstrated for this taxon, and two cytoplasmic markers were used to determine that these lineages are distributed in strict parapatry. Here, using 454 next-generation sequencing, we developed polymorphic microsatellite loci that cross amplify in all three cytoplasmic lineages. Six sites (i.e. two sites within each lineage) were analyzed using nine microsatellite loci. Our work shows that, although substantial cytoplasmic differentiation occurs within *M. laminarioides*, the microsatellite loci did not retrieved three nuclear genetic clusters as expected. Indeed, while the northernmost and southernmost cytoplasmic lineages form two strongly divergent nuclear groups characterized by diagnostic alleles, the third cytoplasmic lineage did not form a third nuclear independent group. It is possible that inter-lineage gene exchange has occurred, particularly at sites along the contact zone between the different cytoplasmic lineages. This nuclear-cytoplasmic incongruence in *M. laminarioides* could be explained by incomplete lineage sorting of the nuclear genes or asymmetric introgressive hybridization between the lineages. Finally, highly significant heterozygote deficiencies (suggesting occurrence of intergametophytic selfing) were observed in the three small northernmost sites while the large southernmost sites generally approached panmixia.

Keywords: Chile, gene flow, intergametophytic selfing, microsatellites, Rhodophyta, species complex

55

56 List of abbreviations:

57 COI, cytochrome c oxidase sub-unit1

58 cpDNA, chloroplast DNA

59 ENSO, El Niño Southern Oscillation

60 HRMA, high resolution melting analysis

61 LD, linkage disequilibrium

62 mtDNA, mitochondrial DNA

63 PCA, principal component analysis

64 RAPD, random amplified polymorphic DNA

65 rbcL, large subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase enzyme

66

67 Introduction

68 The Chilean coast is mostly linear from north to south and characterized by very dynamic

69 and regionalized tectonic, oceanographic and climatic processes (Thiel et al. 2007,

70 Guillemin et al. 2015). These coastal and oceanic features have led to contrasted pattern

71 of marine biodiversity distributed over three main biogeographic regions (Camus 2001,

72 Thiel et al. 2007). The particularities of the Chilean coast have stimulated a strong

interest in deciphering the phylogeographic patterns of the marine realm in this region and studies have accumulated rapidly during the last five years (see for review in invertebrates: Haye et al. 2014 and in seaweeds: Guillemain et al. 2015). More specifically, these two reviews report the use of molecular markers in species that cross the biogeographic boundaries to compare concordance among phylogeographic and biogeographic breaks. In seaweeds, based on the occurrence of divergent mitochondrial lineages, several putative cryptic species were uncovered along the Chilean coast: in *Lessonia* (Tellier et al. 2009), *Durvillaea* (Fraser et al. 2009), *Adenocystis* (Fraser et al. 2013), *Mazzaella* (Montecinos et al. 2012) and *Nothogenia* (Lindstrom et al. 2015). Most of these studies were only based on the analysis of the cytoplasmic genomic compartment, except for *Lessonia* in which comparison of divergence at nuclear and cytoplasmic markers supported the same pattern (Tellier et al. 2009). Furthermore, in this last case, nuclear microsatellites markers were used to demonstrate that the two cryptic species did not share any alleles and were, thus, reproductively isolated (Tellier et al. 2011).

However, conflicting geographic patterns between mitochondrial and nuclear genetic markers have been observed when demographic asymmetries produce dissimilar movement in the two marker types or when different selective pressures affect the mitochondrial and the nuclear genome (Toews and Brelsford 2012). For example, in the barnacle *Notochthamalus scabrosus*, Zakas et al. (2014) reported that the nuclear genome homogeneity throughout the central and northern regions of Chile contrasted with the strong mitochondrial divergence pattern described previously (Zakas et al. 2009). They concluded that there is little reason to treat the two mitochondrial groups as distinct species. In contrast with invertebrates, dispersal is generally limited to less than a few

kilometers in seaweed and such pattern of nuclear homogeneity along the Chilean coast is not expected in those organisms (Kinlan and Gaines, 2003). It is thus interesting to test for algae cryptic species, which were defined on the basis of divergent cytoplasmic lineages, if their pattern of nuclear genetic structure is congruent with their cytoplasmic divergence.

*Mazzaella laminarioides* is a haploid-diploid rocky shore species that forms dense beds in high intertidal zones. This carrageenophyte is an economically important resource in Chile and is harvested from natural populations by small fishing communities (Buschmann et al. 2001). *M. laminarioides* is non-buoyant and is considered to be a poor disperser (Faugeron et al. 2001). The species distribution range encompasses a high variety of environmental conditions as it covers 3,500km (28-56°S) of Chilean coastline (Thiel et al. 2007). Using two cytoplasmic genes (COI, mitochondrial and *rbcL*, chloroplast), Montecinos et al. (2012) revealed strong genetic structure within *M. laminarioides* with the existence of three divergent genetic lineages distributed along the Chilean coast. They reported the presence of a northern lineage from 28°S to 32°S, a central lineage from 34°S to 37°S, and a southern lineage from 39°S to 56°S. Guillemín et al. (2015) confirmed that the three lineages are distributed in strict parapatry with sharp phylogeographic breaks of a few kilometers in width. However, the presence of reproductive barriers has not been tested between these three lineages.

In *M. laminarioides*, the three cytoplasmic lineages were separated for the *rbcL* by 0.6 to 1.0% divergence and for COI by 2.6 to 7.8% divergence and no cytoplasmic incongruence was observed (Montecinos et al. 2012). In red algae, where both plastid and mitochondria DNA maternal inheritance have been observed (Zuccarello and West 2011),

cytoplasmic incongruence have generally been related to events of interspecific hybridization (Destombe et al. 2010). These results suggest that strong reproductive barriers probably evolved between the three lineages of *M. laminarioides* limiting hybridization even in the contact zones. Indeed, in the Rhodophyta, laboratory crosses between phylogenetic species have generally revealed complete reproductive incompatibility that correlates with cytoplasmic genetic distances (Brodie and Zuccarello 2006, Zuccarello and West 2011 and reference therein). In the three species complex thoroughly studied (i.e. *Spyridia*, *Bostrychia*, and *Mastocarpus*), the experiments showed that strains sharing the same chloroplastic haplotypes (rubisco spacer) were always fully compatible while strains differing by only 0.6 to 2.1% were not able to be crossed (Zuccarello and West 2002, Zuccarello and West 2003, Zuccarello et al. 2005).

We can predict that, in agreement with plastid and mitochondrial information, the nuclear genome should present strong genetic discontinuities with no or very limited gene flow among the three cytoplasmic lineages of *M. laminarioides*. In this context, we developed nine microsatellite markers for the red alga *M. laminarioides* in order to confirm the existence of nuclear genetic structure and to test for potential hybridization between the three previously described parapatric cytoplasmic lineages.

## Material and Methods

*Development of microsatellite markers* - In order to construct the 454 libraries, a single haploid (i.e. a female gametophyte) specimen was used as the source of DNA for each three cytoplasmic lineages of *M. laminarioides*. Samples from Fray Jorge (30°40'S/71°42'W), Constitución (35°19'S/72°26'W) and Chiloe (41°52'S/71°01'W) were

used for the northern, central and southern lineage respectively (Montecinos et al. 2012). DNA was extracted following the protocol described by Saunders (1993); slight modifications were made according to Faugeron et al. (2001). PicoGreen<sup>TM</sup> fluorescence enhancement (Ahn et al. 1996) was used to test DNA quality and quantity. DNA sequencing was performed using a 454 GS Junior Titanium Series (Roche) at the AUSTRAL-omics Core-Facilities. Briefly, each DNA sample was tagged using different multiplex identifiers (MIDs). DNA library fragments were captured onto beads and clonally amplified within individual emulsion droplets. Amplified fragments from all three lineages were evenly mixed and sequenced on 3 PicoTiterPlates. Library preparation, amplification, and sequencing were carried out following the manufacturer protocols (Roche Diagnostics Corporation, Branford, Connecticut USA). The assembly of the reads was performed using the MIRA (Chevreux et al. 1999) and CAP3 (Huang and Madan 1999) software programs. In order to determine which contigs correspond to nuclear sequences, nucleotide BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) searches were performed using the complete mitochondrial genome of *Chondrus crispus* (25.836 bp NC\_001677) and using the complete plastid genome of *C. crispus* (180.086 bp, HF562234). Moreover, in order to identified loci sequenced in more than one of our three 454 libraries, we performed nucleotide BLAST searches between the three cytoplasmic lineages contig files.

Nucleotide repeats from di- to hexa-nucleotides were identified using MSATCOMMANDER (Faircloth 2008). Generally, the BLAST searches between the three cytoplasmic lineages contig files show that the same microsatellite locus have been sequenced in more than one of our three 454 libraries (see Table S1). For microsatellite

loci present in more than one of our three 454 libraries, alignment between the different cytoplasmic lineages were performed in GENEIOUS R6 for each locus (Biomatters Ltd.). For 30 loci with a high number of repetitions (at least 7 repetitions) and long flanking regions located in the nuclear contigs, primers pairs were designed using GENEIOUS R6 (Biomatters Ltd.). When the same microsatellite locus was encountered in more than one 454 library (see Table S1) the primer pairs were designed within the more conserved part of the flanking region. All loci for which alignment between the different cytoplasmic lineages were performed showed clear homologies in their flanking regions.

Following primer design, PCR cross amplification of all three lineages was performed; for this, three individuals from Fray Jorge, three individuals from Constitución, and three individuals from Chiloe were used. For the 20 loci that amplified successfully in all three lineages, high resolution melting analysis (HRMA, Mackay et al. 2008) was used to make a preliminary assessment of variability. For each lineage, three sites were tested: Fray Jorge, Puerto Oscuro, and Maitencillo for the northern lineage; Topocalma, Constitución, and Concepción for the central lineage; and Pucatrihue, Chiloe, and Punta Arenas for the southern lineage. Fifteen haploid individuals from each site were used. The COI regions of all of the haploid samples were previously sequenced in the study of Montecinos et al. (2012). The HRMA analyses showed that twelve loci were polymorphic for *M. laminarioides*.

*Genotyping* – The twelve polymorphic loci were used to genotype 96 diploid individuals sampled from Caleta Sauce (SAU, N=16), Mina Talca (MIT, N=18), Montemar (MTM, N=16), Lebu (LEB, N=16), Pilolcura (PIL, N=16) and Melinka (MLK, N=14) (see Figure 1A). The samples from SAU and MIT represent the northern cytoplasmic lineage, the

samples from MTM and LEB represent the central cytoplasmic lineage, and the samples from PIL and MLK represent the southern cytoplasmic lineage (Figure 1A). Cytoplasmic markers sequences are not available for all the samples genotyped with microsatellites. However, since the geographic distribution of the three different cytoplasmic clades (Montecinos et al., 2012) was shown to be strictly parapatric (i.e. each region corresponding to a single cytoplasmic clade with no mixture and no overlapping of distribution of haplotypes between regions), geographical origin of the individuals was considered as a good proxy for cytoplasmic clades' assignation. Direct observations of the reproductive organs (tetrads) in the field were used to select only diploid tetrasporophytes in each localities sampled.

PCR reactions were performed in 12.5  $\mu$ L reactions containing 50 ng of template DNA, 1X PCR buffer, 0.2 mM dNTPs, 0.5  $\mu$ M of each primer and 1 U Top Taq DNA polymerase (Qiagen, Valencia, CA, USA). The concentrations of BSA and  $MgCl_2$  are shown in Table S1. The PCR program consisted of an initial denaturation of 3 min at 94°C, 35 cycles each with 94°C for 40 s, 40 s at optimized Ta (Table S1), and 30 s at 72 °C, and with a final extension of 7 min at 72 °C. PCR products were sent to the Servicio de Secuenciacion Depto. Ecologia, Pontificia Universidad Catolica de Chile, Chile, for fragment analyses. Fragments were separated using an Applied Biosystems 3130XL Genetic Analyzer (Life Technologies Corporation).

*Statistical analyses* – Allele size was determined with the GENEMARKER software (SoftGenetics, State Collage, PA, USA). Prior to analyses, the frequency of null alleles was estimated for each locus using MICRO-CHECKER (Van Oosterhout et al. 2004; Brookfield equation 2, Brookfield 1996). Linkage disequilibrium at all locus pairs was

211 assessed using FSTAT version 2.9.3.2 (Goudet 2001). Statistical significance of LD was  
212 calculated based on 1000 permutations using Bonferonni correction for  $\alpha = 0.05$ . Single  
213 and multilocus estimates of genetic diversity were calculated as the mean number of  
214 alleles per locus ( $N_a$ ), expected heterozygosity ( $H_e$ , sensus Nei 1978) and observed  
215 heterozygosity ( $H_o$ ) using GENETIX 4.05 (Belkhir et al. 1996-2004). Allele frequencies  
216 were calculated for each study site at each locus and plotted using the R-package  
217 StandArich (available at <http://www.ualg.pt/ccmar/maree/software.php>, F. Alberto,  
218 University of Algarve, Faro, Portugal). Single and multilocus estimates of deviation from  
219 random mating ( $F_{IS}$ ) were calculated according to Weir and Cockerham (1984),  
220 significance was determined by running 1000 permutations of alleles among individuals  
221 within sites using GENETIX 4.05 (Belkhir et al. 1996-2004). Genetic differentiation  
222 among sites was analyzed by estimating  $F_{ST}$  ( $\theta$ ) (Weir and Cockerham 1984);  
223 significance of  $F_{ST}$  values was assessed by running 1000 permutations using GENETIX  
224 4.05 (Belkhir et al. 1996-2004). Principal component analysis (PCA) was conducted  
225 using PCAGEN (Goudet 1999) to visualize pairwise differentiation among sites ( $F_{ST}$ );  
226 1,000 randomizations of genotypes were used to determine axis significance. Bayesian  
227 inference was implemented in STRUCTURE v. 2.2 (Pritchard et al. 2000) in order to  
228 detect potential signs of hybridization and/or introgression between lineages. We used the  
229 admixture model and allowed for correlated allele frequencies between sites. A range of  
230 clusters ( $K$ ), from 1 to 10 were tested. Each run, replicated 20 times, consisted of 400000  
231 iterations after a “burn-in” of 200000. To infer which  $K$  best fit the data, we applied the  
232 ad hoc  $\Delta K$  statistic developed by Evanno et al. (2005). The results from STRUCTURE  
233 were then compared to those of INSTRUCT (Gao et al. 2007), which relaxes

STRUCTURE's assumption of Hardy-Weinberg equilibrium as it was designed for species with significant self-fertilization. INSTRUCT was run with the parameters “-K 3 -v 2 -x 0 -w 0 -j 20000 -e 0 -f 0 -L 9 -N 96 -p 2 -u 600000 -b 200000 -t 10 -c 20 -sl 0.95 -g 1 -r 20000 -ik 1 -kv 1 10 -df 1 -af 0 -mm 2.0e9” with K ranging from 1 to 10 (i.e. -kv). For both STRUCTURE and INSTRUCT analyses, combined results of the independent runs were obtained using the greedy algorithm with 100000 random input orders in CLUMMP (Jakobsson and Rosenberg 2007) before exporting the results to DISTRUCT (Rosenberg 2004) for viewing.

## Results

We obtained 172872 reads for the northern lineage, 238901 reads for the central lineage and 111945 reads for the southern lineage (average length: 436 bp). In total, 1091 microsatellite inserts were observed for the northern lineage, 1805 for the central lineage and 667 for the southern lineage (see Table S2 for more details). Most of the microsatellites recovered were di- (87.3%) and trinucleotides (10.8%) while larger repeated motifs were rare (0.9% tetranucleotides and 0.9% pentanucleotides) (Table S2).

Of the 12 polymorphic loci developed during this study (Table S1, Table S3), three loci (MI\_39C37, MI\_106C1748 and MI\_106C203) were not retained for population genetic analyses. The locus MI\_39C37 was not retained because it presented strong and significant linkage disequilibrium with locus MI\_106C32. The locus MI\_106C1748 failed to amplify in 33 of the 96 DNA samples (NI/n = 0.6, Table S3) and the locus MI\_106C203 presented a very low level of polymorphism. Indeed, MI\_106C203 was

fixed for all sites from the northern and southern lineage and only two alleles were observed over the whole dataset (Table S3).

For the nine loci selected (i.e. shaded in grey in Table S3), observed heterozygosities ranged from 0.11 to 0.57 while expected heterozygosities ranged from 0.59 to 0.87 for the whole data set (Table S3). The number of alleles encountered in each site varied from 1 to 10 while the average number of alleles per locus was 3.11 (SE = 1.69) (Table S3). Three loci were moderately polymorphic (4–6 alleles) and six were highly variable (9–26 alleles). The frequency of null alleles was significant for five loci in some of the sites studied (MI\_106C75 in four sites; MI\_106C32 in two sites; and MI\_39C1451, MI\_39C5118 and MI\_39C4313 in only one site) (Table S3). The estimated frequency of null alleles ranged from 0.12 to 0.36 (Table S3).

$F_{IS}$  values were highly variable among loci and sites. The number of significant  $F_{IS}$  values per locus was only slightly higher in the central lineage (8/16) than in the northern (6/15) and southern lineages (6/17) (Table S3). However, the  $F_{IS}$  multilocus estimates (Table 1) show that only LEB and PIL were close to panmixia ( $F_{IS} = -0.03$  and  $-0.07$ , respectively). Alternatively, a negative and significant  $F_{IS}$  was encountered for MLK ( $F_{IS} = -0.17$ ), and positive and significant  $F_{IS}$  values were encountered in the three northern most sites of SAU, MIT and MTM ( $F_{IS} = 0.16, 0.14$  and  $0.45$ , respectively, Table 1). Similar results were obtained when corrections for null alleles were made (Table 1).

Allele size distributions generally overlapped between cytoplasmic lineages when all six sites were taken into account (Figure S1). However, for three of the 12 loci (MI\_106\_C462, MI\_39\_C69 and MI\_106\_C32, Figure S1), allele distributions were

totally disjoint between the southernmost and northernmost cytoplasmic clades (i.e. these three loci are diagnostic for each of the two clades). In addition, a clear gradient of variation in allele size with respect to latitude was observed for those three loci (Figure S1). For the loci MI\_39C69 and MI\_106C462, shorter alleles were encountered in sites from the northern lineage while for the locus MI\_106C32 shorter alleles were observed in sites from the southern lineage (Figure S1). Sites belonging to the central lineage had alleles of intermediate size; the MTM site presented allele sizes that were more similar to the northern lineage while the LEB site presented allele sizes was more similar to the southern lineage (Figure S1). For the MI\_106\_C462, MI\_39\_C69 and MI\_106\_C32 loci, the LEB site presented 27.3% of private alleles, 54.5% of alleles that were shared with the two sites representing the southern cytoplasmic clade (PIL and MLK) and 9.1% of alleles that were shared with the two sites representing the northern cytoplasmic clade (SAU and MIT) (Figure S1). For the same three loci, the MTM site presented 76.9% of private alleles, 7.7% of alleles that were shared with the two sites representing the southern cytoplasmic clade and 23.1% of alleles that were shared with the two sites representing the northern cytoplasmic clade (Figure S1).

The PCA summarized the information given by the nine loci (Figure 1B) and showed results that were congruent with the allele size distribution results. The first two axes explained 39.4% (PC1) and 26.0% (PC2) of the total genetic differentiation ( $F_{ST}$ ) (Figure 1B). The first axis separated sites of the northern lineage (SAU and MIT) from all other lineages (MTM, LEB, PIL and MLK) while the second axis discriminated sites of the central lineage from those of the southern lineage (Figure 1B). However, the MTM site holds a unique position in the PCA and was located in-between the group composed

of the three southernmost sites (LEB, PIL and MLK) and the two sites from the northern lineage along the first axis and was separated from the other five study sites along the second axis (Figure 1B). All pairwise  $F_{ST}$  multilocus values were significantly different from zero and ranged from 0.253 (between PIL and MLK, both from the southern lineage) to 0.581 (between PIL and SAU, northern vs southern lineage differentiation, Table S4). Similar results were found when using the dataset after corrections for null alleles were made (pairwise  $F_{ST}$  ranges from 0.249 to 0.572, Table S4).

Results from the clustering analysis performed with both STRUCTURE and INSTRUCT, revealed that the posterior probability of the data increased steadily from  $K = 1$  to  $K = 6$  while the curves dropped after  $K = 7$  (Figure S2 and S3). The  $\Delta K$  method of Evanno et al. (Evanno et al. 2005) clearly shows that both  $K=2$  and  $K=6$  clusters are the optimal numbers of clusters in our study (Figure S2). Contrary to our expectations, the number of nuclear genetic clusters retrieved was two or even six but definitely not three as predicted with organelle sequences. For  $K=6$ , both STRUCTURE and INSTRUCT identified the same 6 clusters each corresponding to a different sampling site (STRUCTURE: Figure 2; INSTRUCT: Figure S3). Almost all individuals from the same site had similar membership coefficients and, overall, a very low level of admixture was observed (Figure 2,  $K=6$  and Figure S3). This result indicated that differentiation occurs between all of the sampling sites. The Bayesian clustering assignment in STRUCTURE for  $K = 2$  shows that all individuals from the northern lineage (i.e. SAU and MIT) grouped together while the three southernmost sites belonging to the central and southern lineages (i.e. LEB, PIL and MLK) formed another cluster. However, clustering of the MTM individuals to one or the other cluster did vary among the different STRUCTURE

runs. Eleven of the 20 runs (Figure 2, K2 upper graph) assigned all MTM individuals to the northern genetic cluster while individuals from MTM were assigned to the southern genetic cluster for the other 9 runs (Figure 2, K2 lower graph). Only central lineage cytoplasmic COI sequences were observed (M-L. Guillemin, unpublished data) in this site located only 13 kilometres from the contact zone with the northern lineage (Figure 1A).

## Discussion

In this study, nine variable microsatellite markers have been developed for *M. laminarioides* using 454 next-generation sequencing. Recently, this method has been used successfully to develop microsatellite markers for various eukaryotic species (see the review by Megléc et al. 2012) including some red algae (Couceiro et al. 2011a, Pardo et al. 2014, Ayres-Ostrock et al. 2015). Repetitive sequences were encountered in less than 0.7% of the *M. laminarioides* reads. In other red algae (Couceiro et al. 2011a, Pardo et al. 2014, Ayres-Ostrock et al. 2015) and in corals (Ruiz-Ramos and Baums 2014), the percentage of repetitive sequences is generally much higher and ranges from 1.3% in the candelabrum coral *Eunicea flexuosa* to 11.6% in the coralline alga *Phymatolithon calcareum*. This scarcity of repetitive sequences clearly reduced our capacity to develop microsatellites markers for *M. laminarioides* where less than 0.03% of the reads yielded microsatellites with sufficient motif repetitions (i.e. >7). Considering that the main aim of this study was to investigate the pattern of gene flow across the 33°S and the 38°S transition zones, only loci that amplified in all three lineages were selected. A reported drawback of cross-amplifying loci is the high probability of generating an

uneven amplification of the target loci, which could result from defective primers or deviations from optimal amplification conditions, between different taxa (Selkoe and Toonen 2006). In our study, five loci revealed a moderate to high frequency of null alleles in some sites. In *M. laminarioides*, no clear differences were observed between the  $F_{IS}$  and the  $F_{ST}$  estimates calculated using the uncorrected or corrected datasets. As such, the  $F_{IS}$  and  $F_{ST}$  values obtained in our study, even if null alleles are potentially present, can be used to infer biological processes in this species.

The results found here only partially supported the previously defined *M. laminarioides* lineages based on cytoplasmic markers (Montecinos et al. 2012). Indeed, contrary to our expectations, the number of main nuclear genetic clusters retrieved by STRUCTURE was two and not three as predicted with organelle sequences. Based on the nuclear marker data, the two northernmost sites (SAU and MIT) that correspond to the northern cytoplasmic lineage and the two southernmost sites (PIL and MLK) that correspond to the southern cytoplasmic lineage formed two differentiated groups. The presence of several unique, high frequency diagnostic alleles supports the existence of a long-term divergence between these two lineages. On the other hand and contrary to the previous results found using cytoplasmic markers (Montecinos et al. 2012), the sites from the central cytoplasmic lineage (MTM and LEB) do not represent a third well-separated group according to nuclear microsatellites. While LEB clearly grouped with sites of the southern cytoplasmic lineage, the position of MTM stays unclear. Indeed, MTM appeared highly differentiated from the other five sites sampled in the PCA. However, shared alleles with sites that correspond to both the northern and southern cytoplasmic lineages were observed in MTM and the site was not assigned to a third main nuclear genetic

371 cluster in the STRUCTURE analyses but actually shifted between the northern and the  
372 southern genetic cluster, depending on the STRUCTURE run for K2. The discordance  
373 observed between the present study, based on nuclear markers, and the previously  
374 published work based on cytoplasmic markers could simply be explained by incomplete  
375 lineage sorting of the nuclear genes. Due to the differences in effective population size  
376 between the maternally and the bi-parentally inherited genomes (nuclear effective  
377 population size is approximately four times greater than that of the mitochondria and  
378 chloroplast) lineage sorting is expected to be faster for organelle genomes than for the  
379 nucleus. It is thus expected to observe greater genetic differentiation using mtDNA and  
380 cpDNA genes than when using nuclear genes (see Zink and Barrowclough 2008 for a  
381 review). Discordance between genomes can also arise if there are differences in the way  
382 selection acts on the non-recombinant cytoplasmic genomes as compared to the nuclear  
383 genome; or when past and / or present gene flow took / takes place between lineages  
384 (Toews and Brelsford 2012). In their study, Montecinos et al. (2012) recognized that  
385 cytoplasmic markers showed shallower divergence between the central and southern  
386 lineages than between the northern lineage and the central/southern clade. Accordingly,  
387 complete reproductive barriers may not have yet evolved between the more recently  
388 diverging central and southern lineages; this suggests that contemporary gene flow might  
389 still connect the three southernmost sites of LEB, PIL, and MLK. Moreover, asymmetric  
390 introgressive hybridization in the past could also have lead to the incongruence pattern  
391 observed in LEB. In algae, such nuclear / cytoplasmic incongruences were observed in  
392 various species and interpreted as footprints of hybridization (Destombe et al. 2010,  
393 Hoarau et al. 2015) or past introgression (Neiva et al. 2010, Hoarau et al. 2015). In

contrast, the pattern observed for MTM, the other central cytoplasmic lineage site, does not fit this hypothesis of asymmetric introgression. MTM was highly differentiated from both the northern sites (SAU and MIT, ~ 230 km away,  $F_{ST} > 0.5$ ) and the central and southern sites (LEB, PIL and MLK, ~ 550 km away,  $F_{ST} > 0.4$ ). Individuals from this site exhibited a high number of private alleles with intermediate sizes between the sizes of the diagnostic alleles of the two nuclear groups. The MTM site is located only 13 kilometers away from the contact zone with the northern lineage. However, only one individual of MTM was determined to be a possible hybrid. This individual correspond to a genotype assigned mostly to the MTM cluster but for which the secondary cluster of MIT represents more than 5% of the genome in both the STRUCTURE and the INSTRUCT clustering results for K6 (see Figure 2 and S3). The genetic composition of this individual was more consistent with an old hybridization event followed by consecutive backcrosses than with a recent hybridization event. This result reinforce the idea that gene leakage is reduced even at the 33°S transition zone margin.

It is hypothesized that progressive, multiple genome-wide changes accumulated by drift or selection between two gene pools will lead to the build-up of reproductive isolation over time (Nosil and Feder 2012). Hybridization and introgression between two diverging taxa should then diminish with time of divergence. At first sight, this supports the idea that, while reaching a determined threshold of genetic distance, all taxa in a taxonomic group will reach complete reproductive isolation and could then be recognized as a biological species. The existence of such a correlation is the key argument provided to support the creation of a universal threshold to define species using DNA barcoding methods (the 1% threshold for the COI mitochondrial marker as mentioned in BOLD,

Ratnasingham and Hebert 2007). However, the timeframe of speciation is strongly influenced by factors such as drift, opportunities for ecological adaptation, and intrinsic rates of mutations (Sobel et al., 2010), and these factors have been shown to vary depending on the taxa being studied (Dufresnes et al. 2014 and 2015). For COI in red algae, genetic distances between 4.5 to 13.6% were reported between species of the same genus (Saunders 2005, Clarkston and Saunders 2010) but sister species of *Mazzaella* that diverged by less than 1.2% were also reported (Saunders 2005). In *M. laminarioides*, the genetic distance between the three COI-defined cytoplasmic lineages were well within the red algae threshold to delimit species in this genus (2.6 to 7.8%, Montecinos et al. 2012). Our work shows that, although substantial mtDNA and cpDNA differentiation occurs within *M. laminarioides* (Montecinos et al. 2012), this does not necessarily correlate with the existence of complete reproductive isolation. In this study, as in others focused on recently diverging lineages distributed in parapatry (Colliard et al. 2010, Zarza 2011, Dufresnes et al. 2014 and 2015), cytoplasmic markers and microsatellites did not reveal the same pattern of genetic subdivision. However, due to the low number of sites sampled in our study, the true number of *M. laminarioides* nuclear genetic groups will remain unresolved until a more detailed sampling is conducted.

The very high  $F_{ST}$  estimates (all  $> 0.24$ ) obtained in our study suggest that gene flow between the sampled sites is minimal even when sites were previously classified as part of the same nuclear genetic group and cytoplasmic lineage. The retrieval of six clusters, each corresponding to a different sampling site and showing a very low level of admixture, by both STRUCTURE and INSTRUCT, confirm the high differentiation between our six sampling sites. Red algae, that lack long-lived motile gametes, spores,

and floating structures, are ranked among the poorest dispersers when compared to other marine taxa (Kinlan and Gaines 2003). Genetic differentiation has been reported at the scale of kilometres or even meters in these organisms (Faugeron et al. 2001, Zuccarello et al. 2001, Engel et al. 2004, Couceiro et al. 2011b, Krueger-Hadfield et al. 2013). As reported for *M. laminarioides* by Faugeron et al. (2001) using random amplified polymorphic DNA (RAPD) markers, high and significant genetic structure exists between sites located less than 40km apart ( $F_{ST} = 0.39$  between SAU and MIT).  $F_{ST}$  estimates between the three southernmost sites (LEB, PIL and MLK), spanning more than 730km of coast, were significant but lower ( $0.24 < F_{ST} < 0.35$ ) than estimates calculated for the northern sites. Interestingly, the same pattern was reported for the cytoplasmic markers (Montecinos et al. 2012) where genetic structure was much more pronounced in the northern lineage than in the southern lineage. The authors associated the attributed differences between the northern and southern groups to historical demographic differences. They concluded that southern populations of *M. laminarioides* were strongly affected by Quaternary glaciations that would have caused recent demographic expansion while northern populations would have experienced repeated local extinctions and range fragmentations due to the El Niño Southern Oscillation (ENSO) (Montecinos et al. 2012).

These regional differences between the north and the south could also explain the marked differences between the  $F_{IS}$  estimates reported in our study. Highly significant heterozygote deficiencies were observed in the three northernmost site of SAU, MIT and MTM while the sites of LEB and PIL had observed heterozygosities close to random mating. Departure from random mating in the north could be explained by the recurrence

of massive mortality due to ENSO events observed in this region; this would have resulted in patches of small fragmented populations (as reported in algae, Martínez et al. 2003). In the site of MTM, located at the edge of the contact zone with the highly divergent northern lineage (Montecinos et al. 2012), selfing (intergametophytic selfing, see Krueger-Hadfield et al. 2015) could limit gene flow and protect parental genotypes from the formation of hybrid progeny (Antonovics 1968). The prominent role of selfing in the emergence of reproductive barriers has been observed in plants (Martin and Willis 2007, Matallana et al. 2010), fungi (Gibson et al. 2011) and brown algae (Engel et al. 2005, Billard et al. 2010, Hoarau et al. 2015). Small scale sampling within the 33°S contact zone is needed to both determine the exact role of selfing as a reinforcement mechanism in *M. laminarioides* and to better characterize the possible mechanisms responsible of the onset and progress of divergence between the northern and central/south lineages.

In conclusion, this study revealed discordant pattern of geographic variation between nuclear microsatellite markers and cytoplasmic sequences in the red alga *Mazzaella laminarioides*. The presence of several diagnostic alleles did support the existence of a long-term divergence between the northernmost and southernmost cytoplasmic lineages. However, the central cytoplasmic lineage did not form a third nuclear independent group and a high number of shared alleles are observed, principally with the southern cytoplasmic lineage. This shared polymorphism may be explained either by retention of ancestral polymorphism or by hybridization between cytoplasmic lineages. The likelihood of shared ancestral polymorphism vs. recurrent gene flow in these closely related parapatric lineages is a question difficult to test but that we cannot

resolve. Indeed, our study reveals intriguing results for the two sites of the central lineage and jointly leads us to ask if MTM (located at the northern limit of the central lineage) could represent a third nuclear group while the pattern observed in LEB (located in the southern part of the central lineage distribution) is mostly due to ancient or recent gene flow with the neighboring southern lineage. A much denser sampling, especially between the 33°S and the 38°S where the central lineage is located, is needed to better characterize the complex genetic pattern observed in this region. The southern part of the 33-38°S coast forms a mosaic of sandy beaches and rocky coast characterized by rapid changes due to high tectonic activity. Processes, linked to extinction and recolonization, have probably shaped the intertidal algae population of this region leading to complex pattern of genetic differentiation. Nuclear microsatellites clearly show that *M. laminarioides* does not correspond to a single species widely distributed over 3,500km of Chilean coast. These markers rather suggest that one or several speciation processes have probably shaped the present-day patterns of genetic diversity in this species complex. Unraveling the relations between the *M. laminarioides* taxa should be the goal of future works.

## Acknowledgements

This research was supported by FONDECYT N°1130797 (CONICYT, Comisión Nacional de Investigación Científica y Tecnológica, Gobierno de Chile). The authors thank S. Faugeron for the access to *Mazzaella laminarioides* DNA for the 454-library construction. Additional support came from the International Research Network

“Diversity, Evolution and Biotechnology of Marine Algae” (GDRI N°0803). The authors have declared that no competing interests exist.

## Bibliography

Ahn, S. J., Costa, J. & Emanuel, J. R. 1996. PicoGreen quantitation of DNA: effective evaluation of samples Pre-or post-PCR. *Nucleic acids res.* 24(13):2623–2625.

Ayres-Ostrock, L. M., Mauger, S., Plastino, E. M., Oliveira, M. C., Valero, M. & Destombe, C. 2015. Development and characterization of microsatellite markers in two agarophyte species, *Gracilaria birdiae* and *Gracilaria caudata* (Gracilariaceae, Rhodophyta), using next-generation sequencing. *J. Appl. Phycol.* 1–10.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. 1996-2004. Genetix 4.02, Logiciel Sous Windows TM Pour La Génétique Des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR. 5000, Université de Montpellier II, Montpellier.

Billard, E., Serrão, E., Pearson, G., Destombe, C. & Valero, M. 2010. *Fucus vesiculosus* and *spiralis* species complex: a nested model of local adaptation at the shore level. *Mar. Ecol. Prog. Ser.* 405:163–174.

Brodie, J. & Zuccarello, G. C. 2006. Systematics of the species rich algae: red algal classification, phylogeny and speciation. In Hodkinson, T. R. & Parnell, J. A. N. [Eds.] *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*. The Systematics Associations, Boca Raton, Florida, pp.317–30.

Buschmann, A. H., Correa, J. A., Westernmeier, E., Hernandez- Gonzalez, M. D.

531 C. & Normabuena, R. 2001. Red algal farming in Chile: a review. *Aquaculture* 194:203–  
532 220.

533 Brookfield, J. F. Y. 1996. A simple new method for estimating null allele  
534 frequency from heterozygote deficiency. *Mol. Ecol.* 5:453–455.

535 Camus, P. A. 2001. Biogeografía marina de Chile continental. *Rev. Chil. Hist. Nat.*  
536 74(3):587–617.

537 Chevreux, B., Wetter, T. & Suhai, S. 1999. Genome sequence assembly using  
538 trace signals and additional sequence information. *Compt. Sci. Biol. Proc. Ger. Conf.*  
539 *Bioinformatics* 99:45–56.

540 Clarkston, B. E. & Saunders, G. W. 2010. A comparison of two DNA barcode  
541 markers for species discrimination in the red algal family Kallymeniaceae (Gigartinales,  
542 Florideophyceae), with a description of *Euthora timburtonii* sp. nov. *Botany* 88(2):119–  
543 131.

544 Colliard, C., Sicilia, A., Turrise, G. F., Arculeo, M., Perrin, N. & Stöck, M. 2010.  
545 Strong reproductive barriers in a narrow hybrid zone of West-Mediterranean green toads  
546 (*Bufo viridis* subgroup) with Plio-Pleistocene divergence. *BMC Evol. Biol.* 10(1):232.

547 Couceiro, L., Maneiro, I., Mauger, S., Valero, M., Ruiz, J. M. & Barreiro, R.  
548 2011a. Microsatellite development in Rhodophyta using high - throughput sequence data.  
549 *J. Phycol.* 47(6):1258–1265.

550 Couceiro, L., Maneiro, I., Ruiz, J. M. & Barreiro, R. 2011b. Multiscale genetic  
551 structure of an endangered seaweed *Ahnfeltiopsis pusilla* (Rhodophyta): implications for  
552 its conservation. *J. Phycol.* 47(2):259–268.

553 Destombe, C., Valero, M. & Guillemin, M-L. 2010. Delineation of two sibling red

554 algal species, *Gracilaria gracilis* and *Gracilaria dura* (Gracilariales, Rhodophyta), using  
 555 multiple DNA markers: resurrection of the species *G. dura* previously described in the  
 556 northern Atlantic 200 years ago. *J. Phycol.* 46:720–7.

557 Dufresnes, C., Bonato, L., Novarini, N., Betto-Colliard, C., Perrin, N., & Stöck,  
 558 M. 2014. Inferring the degree of incipient speciation in secondary contact zones of  
 559 closely related lineages of Palearctic green toads (*Bufo viridis* subgroup). *Heredity*  
 560 113(1):9–20.

561 Dufresnes, C., Brelsford, A., Crnobrnja-Isailović, J., Tzankov, N., Lymberakis, P.  
 562 & Perrin, N. 2015. Timeframe of speciation inferred from secondary contact zones in the  
 563 European tree frog radiation (*Hyla arborea* group). *BMC Evol. Biol.* 15(1):155.

564 Engel, C. R., Daguin, C. & Serrão, E. A. 2005. Genetic entities and mating system  
 565 in hermaphroditic *Fucus spiralis* and its close dioecious relative *F. vesiculosus* (Fucaceae,  
 566 Phaeophyceae). *Mol. Ecol.* 14:2033–2046.

567 Engel, C. R., Destombe, C. & Valero, M. 2004. Mating system and gene flow in  
 568 the red seaweed *Gracilaria gracilis*: effect of haploid-diploid life history and intertidal  
 569 rocky shore landscape on finescale genetic structure. *Heredity* 92:289–298.

570 Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of  
 571 individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14:2611–  
 572 2620.

573 Faircloth, B. C. 2008. MSATCOMMANDER: detection of microsatellite repeat  
 574 arrays and automated, locus-specific primer design. *Mol. Ecol. Res.* 8:92–94.

575 Faugeron S., Valero M., Destombe C., Martínez E. A. & Correa, J. A. 2001.  
 576 Hierarchical spatial structure and discriminant analysis of genetic diversity in the red alga

577 *Mazzaella laminarioides* (Gigartinales, Rhodophyta). *J. Phycol.* 37:705–716.

578 Fraser, C. I., Nikula, R., Spencer, H. G. & Waters, J. M. 2009. Kelp genes reveal  
579 effects of subantarctic sea ice during the Last Glacial Maximum. *Proc. Natl. Acad. Sci.*  
580 *USA* 106:3249–3253.

581 Fraser, C. I., Zuccarello, G. C., Spencer, H. G., Salvatore, L. C., Garcia, G. R., &  
582 Waters, J. M. 2013. Genetic affinities between trans-oceanic populations of non-buoyant  
583 macroalgae in the high latitudes of the Southern Hemisphere. *PLoS ONE* 8(7):e69138.

584 Gao, H. S., Williamson, S. & Bustamante, C. D. 2007. A Markov chain Monte  
585 Carlo approach for joint inference of population structure and inbreeding rates from  
586 multilocus genotype data. *Genetics* 176:1635–1651.

587 Gibson, A. K., Hood, M. E. & Giraud, T. 2012. Sibling competition arena: selfing  
588 and a competition arena can combine to constitute a barrier to gene flow in sympatry.  
589 *Evolution* 66(6):1917–1930.

590 Goudet, J. 1999. [<http://www2.unil.ch/popgen/softwares/pcagen.htm>].

591 Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and  
592 fixation indices (version 2.9. 3).

593 Guillemin, M-L., Valero, M., Tellier, F., Macaya, E. C., Destombe, C. &  
594 Faugeron, S. 2015. Phylogeography of seaweeds in the South East Pacific: complex  
595 evolutionary processes along a latitudinal gradient. *In* Hu Z-M. & Fraser, C. [Eds.]  
596 *Seaweed phylogeography*. Springer Verlag, Berlin.

597 Hoarau, G., Coyer, J. A., Giesbers, M. C. W. G., Jueterbock, A., & Olsen, J. L.  
598 2015. Pre-zygotic isolation in the macroalgal genus *Fucus* from four contact zones  
599 spanning 100–10 000 years: a tale of reinforcement? *Royal Society Open Science*

600 2(2):140538.

601 Huang, X. Q. & Madan, A. 1999. CAP3: a DNA sequence assembly program.

602 *Genome Res.* 9:868–877.

603 Jakobsson, M. & Rosenberg, N. A. 2007. CLUMPP: a cluster matching and

604 permutation program for dealing with label switching and multimodality in analysis of

605 population structure. *Bioinformatics* 23(14):1801–1806.

606 Kinlan, B. P. & Gaines, S. D. 2003. Propagule dispersal in marine and terrestrial

607 environments: a community perspective. *Ecology* 84:2007–20.

608 Krueger-Hadfield, S. A., Roze, D., Correa, J. A., Destombe, C. & Valero, M.

609 2015. O father where art thou? Paternity analyses in a natural population of the haploid-

610 diploid seaweed *Chondrus crispus*. *Heredity* 114:185–194.

611 Krueger-Hadfield, S. A., Roze, D., Mauger, S. & Valero, M. 2013.

612 Intergametophytic selfing and microgeographic genetic structure shape populations of the

613 intertidal red seaweed *Chondrus crispus*. *Mol. Ecol.* 22(12):3242–3260.

614 Lindstrom, S. C., Gabrielson, P. W., Hughey, J. R., Macaya, E. C., & Nelson, W.

615 A. 2015. Sequencing of historic and modern specimens reveals cryptic diversity in

616 *Nothogenia* (Scinaiaaceae, Rhodophyta). *Phycologia* 54 (2):97–108.

617 Mackay, J. F., Wright, C. D. & Bonfiglioli, R. G. 2008. A new approach to

618 varietal identification in plants by microsatellite high resolution melting analysis:

619 application to the verification of grapevine and olive cultivars. *Plant Methods* 4(1):8.

620 Matallana, G., Godinho, M. A. S., Guilherme, F. A. G., Belisario, M., Coser, T. S.

621 & Wendt, T. 2010. Breeding systems of Bromeliaceae species: evolution of selfing in the

622 context of sympatric occurrence. *Plant. Syst. Evol.* 289:57–65.

623           Martin, N. H., & Willis, J. H. 2007. Ecological divergence associated with mating  
624   system causes nearly complete reproductive isolation between sympatric *Mimulus species*.  
625   *Evolution* 61:68–82.

626           Martínez, E. A., Cárdenas, L. & Pinto, R. 2003. Recovery and genetic diversity of  
627   the intertidal kelp *Lessonia nigrescens* (phaeophyceae) 20 years after el Niño 1982/83. *J.*  
628   *Phycol.* 39(3):504–508.

629           Megléczy, E., Nève, G., Biffin, E. & Gardner, M. G. 2012. Breakdown of  
630   phylogenetic signal: a survey of microsatellite densities in 454 shotgun sequences from  
631   154 non model eukaryote species. *PLoS One* 7(7):e40861.

632           Montecinos, A., Broitman, B. R., Faugeron, S., Haye, P. A., Tellier, F. &  
633   Guillemin, M. L. 2012. Species replacement along a linear coastal habitat:  
634   phylogeography and speciation in the red alga *Mazzaella laminarioides* along the south  
635   east pacific. *BMC Evol. Biol.* 12(1):97.

636           Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a  
637   small number of individuals. *Genetics* 89(3):583–590.

638           Neiva, J., Pearson, G. A., Valero, M. & Serrao, E. A. 2010. Surfing the wave on a  
639   borrowed board: range expansion and spread of introgressed organellar genomes in the  
640   seaweed *Fucus ceranoides* L. *Mol. Ecol.* 19(21):4812–4822.

641           Nosil, P., & Feder, J. L. 2012. Genomic divergence during speciation: causes and  
642   consequences. *Philos. T. Roy. Soc. B* 367(1587):332–342.

643           Pardo, C., Peña, V., Bárbara, I., Valero, M. & Barreiro, R. 2014. Development  
644   and multiplexing of the first microsatellite markers in a coralline red alga (*Phymatolithon*  
645   *calcareum*, Rhodophyta). *Phycologia* 53(5):474–479.

646 Pritchard, J. K., Stephens, M. & Donnelly, P. 2000. Inference of population  
647 structure using multilocus genotype data. *Genetics* 155:945–959.

648 Ratnasingham, S. & Hebert, P. D. 2007. BOLD: The Barcode of Life Data System  
649 (<http://www.barcodinglife.org>). *Mol. Ecol. Notes* 7(3):355–364.

650 Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of  
651 population structure. *Mol. Ecol. Notes* 4(1):137–138.

652 Ruiz-Ramos, D. V. & Baums, I. B. 2014. Microsatellite abundance across the  
653 Anthozoa and Hydrozoa in the phylum Cnidaria. *BMC genomics* 15(1):939.

654 Saunders, G. W. 2005. Applying DNA barcoding to red macroalgae: a  
655 preliminary appraisal holds promise for future applications. *Philos. T. Roy. Soc. B*  
656 360(1462):1879–1888.

657 Selkoe, K. A. & Toonen, R. J. 2006. Microsatellites for ecologists: a practical  
658 guide to using and evaluating microsatellite markers. *Ecol. Lett.* 9(5):615–629.

659 Sobel, J. M., Chen, G. F., Watt, L. R. & Schemske, D. W. 2010. The biology of  
660 speciation. *Evolution* 64(2):295–315.

661 Tellier, F., Meynard, A. P., Correa, J. A., Faugeron, S., & Valero, M. 2009.  
662 Phylogeographic analyses of the 30°S south-east Pacific biogeographic transition zone  
663 establish the occurrence of a sharp genetic discontinuity in the kelp *Lessonia nigrescens*:  
664 Vicariance or parapatry? *Mol. Phyl. Evol.* 53(3):679–693.

665 Tellier, F., Tapia, J., Faugeron, S., Destombe, C. & Valero, M. 2011. The  
666 *Lessonia nigrescens* species complex (Laminariales, Phaeophyceae) shows strict  
667 parapatry and complete reproductive isolation in a secondary contact zone. *J. Phycol.*  
668 47:894–903.

669 Thiel, M., Macaya, E. C., Acuña, E., Arntz, W. E., Bastias, H., Brokordt, K.,  
 670 Camus, P. A., Castilla, J. C., Castro, L. R., Cortés, M. et al. 2007. The Humboldt Current  
 671 System of northern and central Chile. *Oceanogr. Mar. Biol.* 45:195–344.  
 672 Toews, D. P. & Brelsford, A. 2012. The biogeography of mitochondrial and  
 673 nuclear discordance in animals. *Mol. Ecol.* 21(16):3907–3930.  
 674 Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. & Shipley, P. 2004.  
 675 MICRO - CHECKER: software for identifying and correcting genotyping errors in  
 676 microsatellite data. *Mol. Ecol. Notes* 4(3):535–538.  
 677 Weir, B. S. & Cockerham, C. C. 1984. Estimating F-statistics for the analysis of  
 678 population structure. *Evolution* 38:1358–1370.  
 679 Zakas, C., Binford, J., Navarrete, S. A., & Wares, J. P. 2009. Restricted gene flow  
 680 in Chilean barnacles reflects an oceanographic and biogeographic transition zone. *Mar.*  
 681 *Ecol. Prog. Ser.* 394:165–177.  
 682 Zakas, C., Jones, K., & Wares, J. P. 2014. Homogeneous nuclear background for  
 683 mitochondrial cline in northern range of *Notochthamalus scabrosus*. *G3: Genes Genomes*  
 684 *Genetics* 4(2):225–230.  
 685 Zarza, E., Reynoso, V. H. & Emerson, B. C. 2011. Discordant patterns of  
 686 geographic variation between mitochondrial and microsatellite markers in the Mexican  
 687 black iguana (*Ctenosaura pectinata*) in a contact zone. *J. Biogeo.* 38(7):1394–1405.  
 688 Zink, R. M. & Barrowclough, G. F. 2008. Mitochondrial DNA under siege in  
 689 avian phylogeography. *Mol. Ecol.* 17(9):2107–2121.  
 690 Zuccarello, G. C., Schidlo, N., Mc Ivor, L. & Guiry, M. D. 2005. A molecular re-  
 691 examination of speciation in the intertidal red alga *Mastocarpus stellatus* (Gigartinales,

692 Rhodophyta) in Europe. *Eur. J. Phycol.* 40:337–44.

693         Zuccarello, G. C. & West, J. A. 2002. Phylogeography of the *Bostrychia*  
694 *calliptera*-*B. pinnata* complex (Rhodomelaceae, Rhodophyta) and divergence rates based  
695 on nuclear, mitochondrial and plastid DNA markers. *Phycologia* 41(1):49–60.

696         Zuccarello, G. C. & West, J. A. 2003. Multiple cryptic species: molecular  
697 diversity and reproductive isolation in the *Bostrychia radicans*/*B. moritziana* complex  
698 (Rhodomelaceae, Rhodophyta) with focus on North American isolates. *J. Phycol.*  
699 39:948–959.

700         Zuccarello, G. C. & West, J. A. 2011. Insights into evolution and speciation in the  
701 red alga *Bostrychia*: 15 years of research. *Algae* 26(1):21–32.

702         Zuccarello, G. C., Yeates, P. H., Wright, J. T. & Bartlett, J. 2001. Population  
703 structure and physiological differentiation of haplotypes of *Caloglossa leprieurii*  
704 (Rhodophyta) in a mangrove intertidal zone. *J. Phycol.* 37:235–44.

705

706

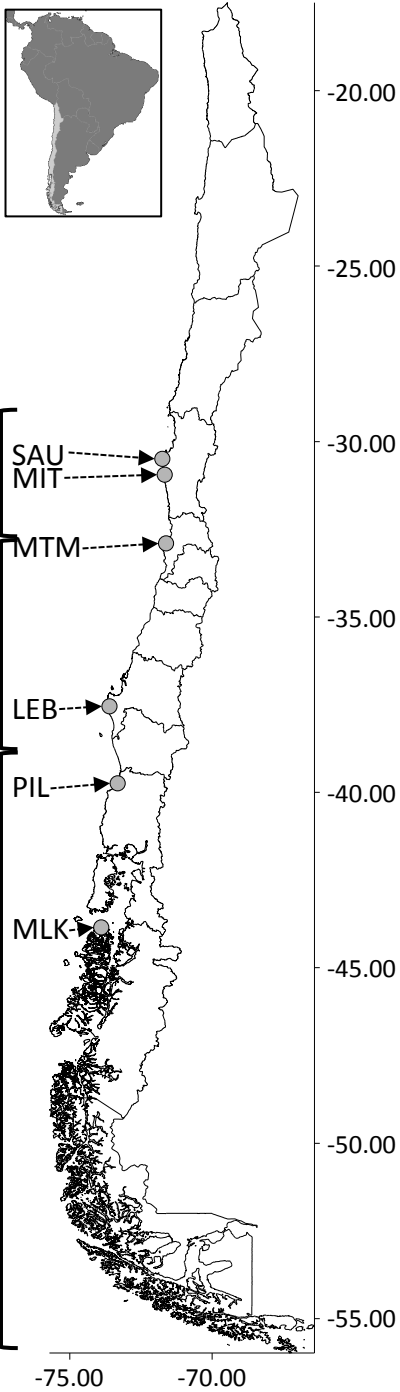
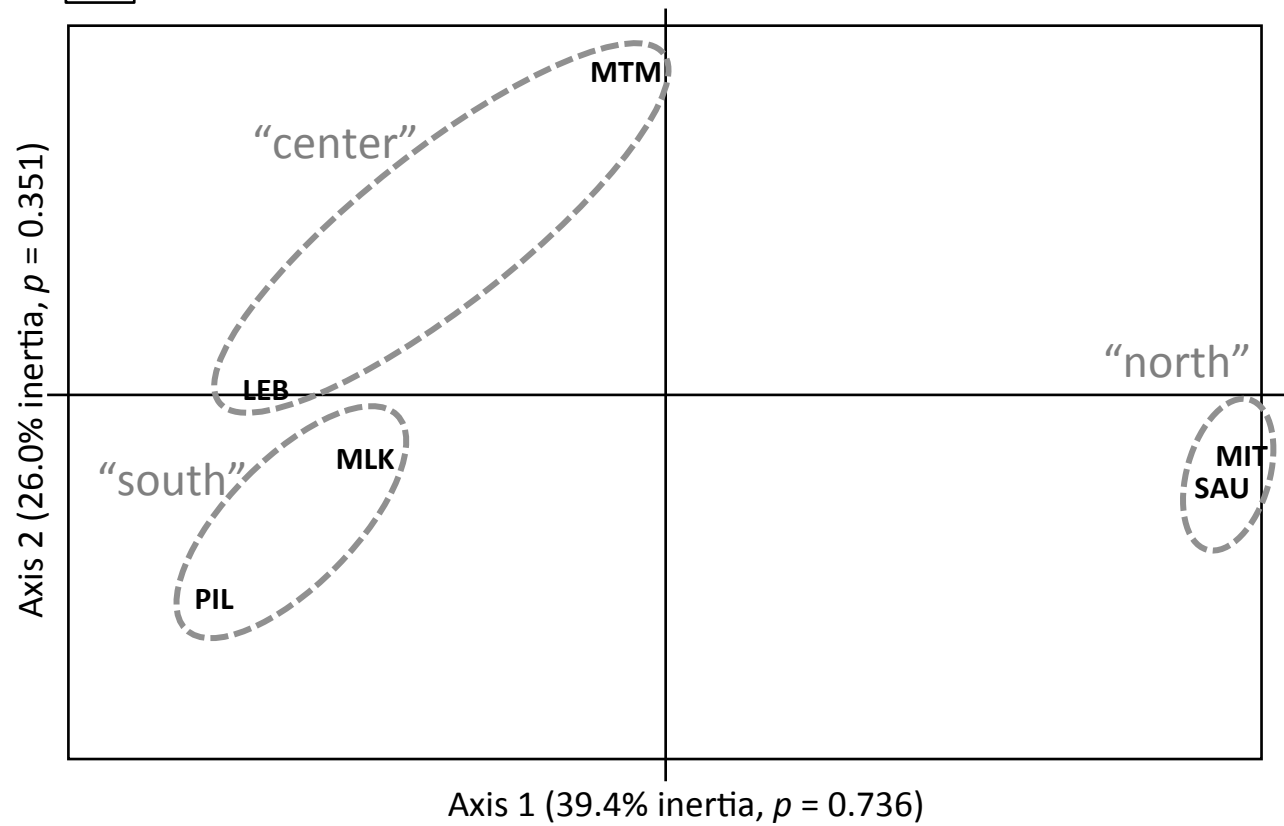
707

Figure 1: Geographic location of the study sites and cytoplasmic clades along the Chilean coast (A) and principal component analysis (PCA) based on mean pairwise  $F_{ST}$  averaged over the nine selected loci (B). In the map, brackets show the cytoplasmic lineages of *M. laminarioides*: “north”, “center” and “south”. Ellipsoids delineating the three cytoplasmic lineages of *Mazzaella laminarioides* in the PCA were drawn for better visualization, but they do not have any statistical support.

Figure 2: Bayesian analysis using STRUCTURE for the six study sites of *Mazzaella laminarioides*. Each vertical bar represents a different individual. Each shade represents the proportion of individual genome assigned to each genetic group and \* indicates individuals for which a secondary cluster represents more than 5% of the genome (i.e. admixed individuals). Individuals are ordered from north to south. Results are shown for K2 and K6. For K2, since the 20 independent runs give contrasting results depending on the run, two graphs are presented: one that corresponds to 11 of the 20 runs (upper graph) and the other that corresponds to the remaining 9 of the 20 runs (lower graph).

Table 1: Multilocus estimates of the number of alleles per locus ( $N_a$ ), expected heterozygosity ( $H_e$ ), observed heterozygosity ( $H_o$ ) and  $F_{IS}$  calculated for the nine selected microsatellite loci. Significant departure from panmixia was tested by running 1,000 permutations of alleles among individuals within sites using GENETIX 4.05 (Belkhir et al. 1996-2004).  $F_{IS}$  values significantly different from zero are shown in bold. Both an uncorrected dataset and a dataset corrected for null alleles (Oosterhout et al. 2004) were used for the calculations. For  $N_a$ ,  $H_e$  and  $H_o$ , mean and standard deviation computed over the nine loci are noted.

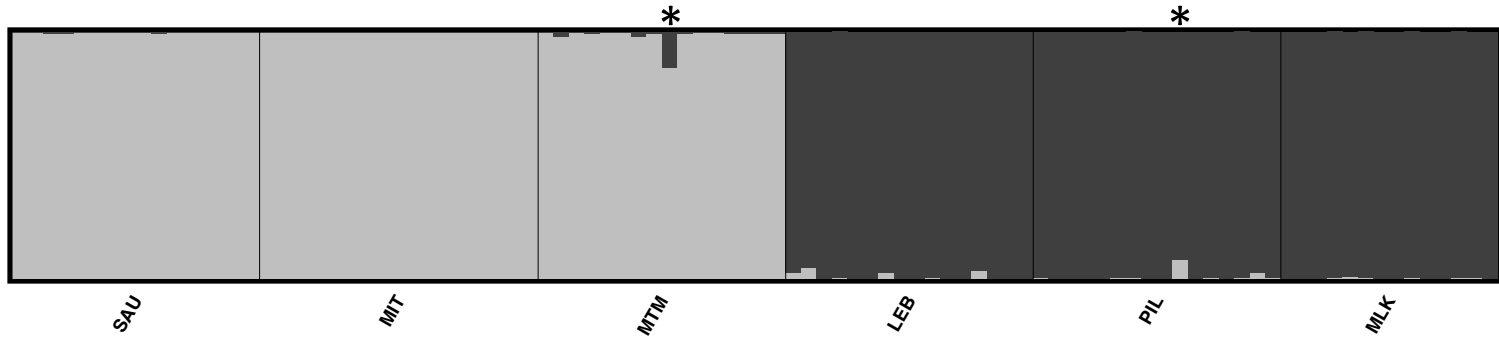
Geographic origin of sampled sites	Site	$N_a$	$H_e$	$H_o$	$F_{IS}$	$p$
<b>Uncorrected data set</b>						
Northern cytoplasmic lineage	SAU	3.11 (1.69)	0.30 (0.23)	0.26 (0.26)	<b>0.16</b>	<b>0.0180</b>
	MIT	3.67 (2.69)	0.46 (0.28)	0.41 (0.30)	<b>0.14</b>	<b>0.0360</b>
Central cytoplasmic lineage	MTM	3.56 (2.35)	0.37 (0.28)	0.21 (0.23)	<b>0.45</b>	<b>0.0000</b>
	LEB	3.89 (1.83)	0.48 (0.14)	0.51 (0.32)	-0.03	0.3580
Southern cytoplasmic lineage	PIL	3.11 (1.54)	0.35 (0.22)	0.38 (0.27)	-0.07	0.1230
	MLK	3.22 (1.09)	0.55 (0.07)	0.66 (0.22)	<b>-0.17</b>	<b>0.0110</b>
<b>Data set corrected for the presence of null alleles</b>						
Northern cytoplasmic lineage	SAU	3.11 (1.69)	0.30 (0.24)	0.27 (0.26)	<b>0.12</b>	<b>0.0450</b>
	MIT	3.67 (2.69)	0.47 (0.29)	0.42 (0.31)	<b>0.13</b>	<b>0.0180</b>
Central cytoplasmic lineage	MTM	3.56 (2.35)	0.40 (0.31)	0.24 (0.24)	<b>0.45</b>	<b>0.0000</b>
	LEB	3.89 (1.83)	0.49 (0.14)	0.54 (0.29)	-0.06	0.2100
Southern cytoplasmic lineage	PIL	3.11 (1.54)	0.35 (0.22)	0.39 (0.27)	-0.08	0.1350
	MLK	3.22 (1.09)	0.55 (0.08)	0.66 (0.22)	<b>-0.17</b>	<b>0.0180</b>

**A****B**

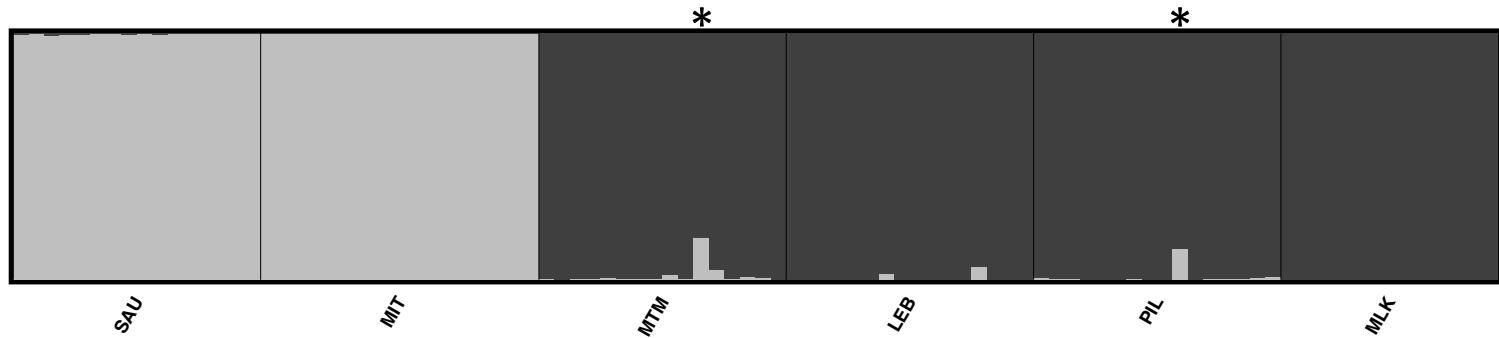
# Cytoplasmic lineages

North	Center	South
-------	--------	-------

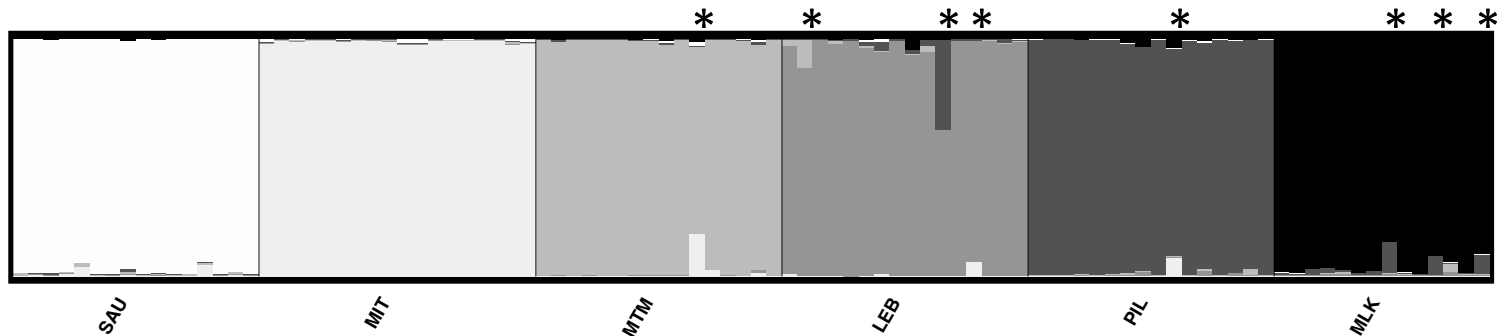
11 of the  
20 runs



9 of the  
20 runs



All the  
20 runs



K2

K6

Table S1: Characteristics of 12 polymorphic microsatellite loci for the three lineages of *Mazzaella laminarioides*. Repeat motifs are given for every lineage for which the microsatellite sequence was encountered in the 454 nuclear contigs (N = northern lineage, C = central lineage and S = southern lineage). PCR conditions are given for each lineage.

Locus	Primer sequence (5'-3')	Repeat motif <sup>t</sup>	BSA ( $\mu\text{g}/\mu\text{L}$ )	Mg (mM)	Ta (°C)
MI_106_C462	F: GCAAGGGATGACCATGAC R: CTTGTGCCACCTCTATTCTAAGC	N: _	0.2	1	62
		C: _	0-0.2*	1-1.5*	62
		S: (TGA) <sub>12</sub>	0	1	62
MI_106_C10	F: GCGTGTAGCACAGTACTTCTAC R: GAAAGCACCGAAACACAGCC	N: (CA) <sub>10</sub> AATA(CA) <sub>4</sub>	0	1	65-67*
		C: CATA(CA) <sub>9</sub>	0	1	67
		S: (CA) <sub>13</sub> AATA(CA) <sub>4</sub>	0	1	65-67*
MI_39_C69	F: GCTGTCGAGTGTACGTTTCG R: GCCTCTGTGAAGCAAGCA	N: _	0.1	1.2	60
		C: (GA) <sub>9</sub>	0.1	1.2	60
		S: (GA) <sub>8</sub>	0.1	1.2	60
MI_106_C75	F: CGCAATCGGGAGCCATCG R: CCCTATCGTGTGTTGCCACCG	N: (GGCTC) <sub>12</sub>	0	0.8	67
		C: (GGCTC) <sub>7</sub>	0	0.8	67
		S: (GGCTC) <sub>5</sub>	0	0.8	65-67*
MI_106_C32	F: CTGGTACAGTACCGAAGATGTC R: GTTGGGTAATCGAAGAAGATGA	N: (AC) <sub>12</sub> (ATACACACAC) <sub>6</sub>	0.15	1.2	63-64*
		C: (AC) <sub>23</sub> ATACTC(AC) <sub>2</sub>	0.15	1.2	63
		S: (AC) <sub>13</sub> ATACTC(AC) <sub>2</sub>	0.15	1.2	64
MI_106_C1748	F: CGTGATGTGTGTCATTTTATCCC  R: CTGTTCTGTTGTTCTGCG	N: (CCTC) <sub>4</sub> N <sub>44</sub> (TG) <sub>2</sub> TA(TG) <sub>2</sub> (TG TGTA) <sub>3</sub>	0.2	1	58
		C: (CCTC) <sub>4</sub> N <sub>44</sub> (TG) <sub>2</sub> TA(TG) <sub>2</sub> GA (TG) <sub>11</sub>	0.05	1	61
		S: (CCTC) <sub>4</sub> N <sub>44</sub> (TG) <sub>2</sub> TA(TG) <sub>11</sub>	0.05	1	61

MI_39_C1451	F: GGGAAAAAGCGAGCAATTTG R: CTGTGCCGTCTGACATTG	N: _	0	-	58
		C: (TG) <sub>8</sub> CG(TG) <sub>4</sub>	0.2	1.5	58
		S: (TG) <sub>6</sub> CG(TG) <sub>4</sub>	0.2	1.5	58
MI_39_C5118	F: GACTTCGCTGTCCATCCA R: CAACGAATCATCCAGACGAC	N: _	0	-	64
		C: (TG) <sub>7</sub> TC(TG) <sub>12</sub>	0	-	64
		S: _	0	-	64
MI_106_C203	F: CGGAGGCGACGGAGGAAG R: CCTCAATTCCCTCGGTTGCTG	N: (TG) <sub>5</sub> TA(TG) <sub>2</sub> CG(TG) <sub>4</sub>	0	1	65
		C: (TG) <sub>8</sub> CG(TG) <sub>4</sub>	0	1	66
		S: ((TG) <sub>6</sub> CG(TG) <sub>4</sub>	0	1	65
MI_39_C4313	F: ATCGTTTCAGGGCAATCACTG R: TCATCCCTCCGTACCTGC	N: (AG) <sub>6</sub> (TG) <sub>5</sub>	0	-	56-58*
		C: (AG) <sub>6</sub> (TG) <sub>6</sub>	0	-	62
		S: (AG) <sub>6</sub> (TG) <sub>5</sub>	0	-	62
MI_39_C37	F: CGGCGATGATCGACTGAGATAGAA R: GATGTCCCACCAACGATTGATGG	N: (TGTGTGTGTA) <sub>5</sub> (TG) <sub>12</sub>	0.05	1	65
		C: (TG) <sub>24</sub>	0.05	1	65
		S: (TG) <sub>14</sub>	0.05	1	65
MI_39_C3942	F: CAATGATCGTATTTACTTCGTAGCG R: GCTAACCACAATACTGGCT	N: (CGT) <sub>4</sub> N <sub>12</sub> (AT) <sub>5</sub> (CT) <sub>2</sub>	0.2	1	59-61*
		C: (CGT) <sub>4</sub> N <sub>10</sub> (AT) <sub>9</sub> C(AT) <sub>2</sub> (CT) <sub>3</sub>	0.2	1	61
		S: _	0.2	1	59-61*

Ta annealing temperature for the PCR program.

\* When PCR conditions were different between sites from the same lineage, the range of Ta and BSA and Mg2 concentration used are given.

<sup>t</sup> A single haploid (i.e. a female gametophyte) specimen was used as the source of DNA for each three cytoplasmic lineages of *M. laminarioides* to construct the 454 libraries. Samples from Fray Jorge (30°40'S/71°42'W), Constitución (35°19'S/72°26'W) and Chiloe (41°52'S/71°01'W) were used for the northern, central and southern lineage respectively (Montecinos et al. 2012). BLAST searches between the three cytoplasmic lineages contig files were performed to identify microsatellite loci sequenced in more than one of our three 454 libraries (see Material and Methods).

Table S2: Counts of loci for each combination of microsatellite category (di-, tri-, tetra-, and pentanucleotides) and number of perfect tandem repeat units in each *Mazzaella laminarioides* cytoplasmic lineage (i.e. “north”, “center” and “south”).

<i>M. laminarioides</i> "north"		N° Rep.	N		N° Rep.	N		N° Rep.	N		N° Rep.	N
	Di	4	791	Tri	4	99	Tetra	4	9	Penta	4	10
	Di	5	113	Tri	5	15	Tetra	5	0	Penta	5	0
	Di	6	22	Tri	6	6	Tetra	6	1	Penta	6	0
	Di	7	10	Tri	7	2	Tetra	7	0	Penta	7	0
	Di	8	3	Tri	8	0	Tetra	8	0	Penta	8	0
	Di	9	0	Tri	9	1	Tetra	9	0	Penta	9	0
	Di	10	2	Tri	10	0	Tetra	10	0	Penta	10	0
	Di	>10	7	Tri	>10	0	Tetra	>10	0	Penta	>10	0
<i>M. laminarioides</i> "center"		N° Rep.	N		N° Rep.	N		N° Rep.	N		N° Rep.	N
	Di	4	1317	Tri	4	129	Tetra	4	10	Penta	4	11
	Di	5	175	Tri	5	35	Tetra	5	3	Penta	5	1
	Di	6	41	Tri	6	11	Tetra	6	0	Penta	6	0
	Di	7	22	Tri	7	1	Tetra	7	1	Penta	7	1
	Di	8	11	Tri	8	4	Tetra	8	0	Penta	8	0
	Di	9	5	Tri	9	8	Tetra	9	0	Penta	9	0
	Di	10	3	Tri	10	0	Tetra	10	0	Penta	10	0
	Di	>10	13	Tri	>10	3	Tetra	>10	0	Penta	>10	0

<i>M. laminarioides</i> "south"		N° Rep.	N		N° Rep.	N		N° Rep.	N		N° Rep.	N
	Di	4	471	Tri	4	49	Tetra	4	7	Penta	4	8
	Di	5	68	Tri	5	12	Tetra	5	1	Penta	5	0
	Di	6	17	Tri	6	5	Tetra	6	1	Penta	6	1
	Di	7	7	Tri	7	2	Tetra	7	1	Penta	7	0
	Di	8	4	Tri	8	0	Tetra	8	0	Penta	8	0
	Di	9	2	Tri	9	0	Tetra	9	0	Penta	9	0
	Di	10	1	Tri	10	2	Tetra	10	0	Penta	10	0
	Di	>10	6	Tri	>10	2	Tetra	>10	0	Penta	>10	0

Table S3: Genetic variability within sites of *Mazzaella laminarioides* for the 12 microsatellite loci. Number of individuals analyzed ( $n$ ); ratio of individuals successfully genotyped ( $NI/n$ ), with  $NI$  = number of individuals successfully genotyped and  $n$  = number of individuals analyzed; number of alleles per locus ( $Na$ ); expected heterozygosity ( $He$ ); observed heterozygosity ( $Ho$ ). For each single locus  $F_{IS}$  estimates, significant departure from panmixia was tested by running 1,000 permutations of alleles among individuals within sites using GENETIX 4.05 software (Belkhir et al. 1996-2004).  $F_{IS}$  values significantly different from zero are shown in bold. Null allele frequency ( $Nf$ ) obtained with MICRO-CHECKER (Oosterhout et al. 2004; Brookfield equation 2, Brookfield 1996), locus showing a significant frequency of null alleles are shown in bold.

Geographic origin of sampled sites		Northern cytoplasmic lineage		Central cytoplasmic lineage		Southern cytoplasmic lineage		
Loci	Site	SAU	MIT	MTM	LEB	PIL	MLK	All
	$n$	16	18	16	16	16	14	96
Ml_106C462	$NI/n$	1	1	0.94	1	1	1	0.99
	$Na$	1	1	1	4	4	4	9
	$He$	0	0	0	0.23	0.23	0.56	0.76
	$Ho$	0	0	0	0.13	0.13	0.43	0.11
	$F_{IS}$	—	—	—	<b>0.48</b>	<b>0.48</b>	0.27	
	$p$	—	—	—	<b>0.03</b>	<b>0.04</b>	0.172	
	$Nf$	—	—	—	0.08	0.08	0.09	
Ml_106C10	$NI/n$	1	1	1	1	1	1	1
	$Na$	5	4	3	3	4	3	12
	$He$	0.33	0.7	0.23	0.53	0.48	0.46	0.81

	<i>Ho</i>	0.31	0.72	0.25	1	0.5	0.64	0.57
	<i>F<sub>IS</sub></i>	0.08	0	-0.08	<b>-0.88</b>	-0.02	-0.38	
	<i>p</i>	0.389	0.588	0.802	<b>0</b>	0.616	0.142	
	<i>N<sub>f</sub></i>	0.01	0	0	0	0	0	
Ml_39C69	<i>NI/n</i>	1	1	1	1	0.94	1	0.99
	<i>Na</i>	2	2	3	3	4	5	11
	<i>He</i>	0.22	0.5	0.28	0.54	0.55	0.69	0.83
	<i>Ho</i>	0.25	0.51	0.25	0.38	0.87	0.71	0.57
	<i>F<sub>IS</sub></i>	-0.11	<b>-0.89</b>	0.12	0.33	<b>-0.54</b>	0.01	
	<i>p</i>	0.814	<b>0.001</b>	0.316	0.109	<b>0.014</b>	0.579	
	<i>N<sub>f</sub></i>	0	0	0.02	0.11	0.08	0	
Ml_106C75	<i>NI/n</i>	1	1	1	1	1	1	1
	<i>Na</i>	4	5	4	5	6	4	10
	<i>He</i>	0.54	0.73	0.59	0.54	0.71	0.64	0.81
	<i>Ho</i>	0.25	0.56	0.31	0.31	0.5	0.43	0.4
	<i>F<sub>IS</sub></i>	<b>0.56</b>	0.26	<b>0.5</b>	<b>0.45</b>	<b>0.33</b>	<b>0.37</b>	
	<i>p</i>	<b>0.006</b>	0.056	<b>0.007</b>	<b>0.015</b>	<b>0.018</b>	<b>0.036</b>	
	<i>N<sub>f</sub></i>	<b>0.19</b>	0.1	<b>0.18</b>	<b>0.15</b>	<b>0.12</b>	0.13	
Ml_106C32	<i>NI/n</i>	1	1	0.88	1	1	1	0.98
	<i>Na</i>	5	10	9	4	2	4	26
	<i>He</i>	0.56	0.8	0.8	0.5	0.22	0.56	0.87
	<i>Ho</i>	0.63	0.5	0.71	0.19	0.25	0.57	0.47
	<i>F<sub>IS</sub></i>	-0.08	<b>0.4</b>	0.15	<b>0.65</b>	-0.11	0.01	
	<i>p</i>	0.472	<b>0</b>	0.183	<b>0</b>	0.831	0.576	
	<i>N<sub>f</sub></i>	0	<b>0.17</b>	0.26	<b>0.21</b>	0	0	
Ml_106C1748	<i>NI/n</i>	0.5	0.72	0.38	0.94	0.5	0.93	0.65
	<i>Na</i>	5	7	6	2	8	8	16
	<i>He</i>	0.61	0.83	0.79	0.28	0.83	0.8	0.88

	<i>Ho</i>	0.38	0.69	0.5	0.33	0.75	0.77	0.57
	<i>F<sub>IS</sub></i>	0.44	0.21	<b>0.44</b>	-0.17	0.16	0.07	
	<i>p</i>	0.055	0.083	<b>0.03</b>	0.653	0.226	0.437	
	<i>Nf</i>	0.72	0.45	0.82	0.19	0.66	0.16	
Ml_39C1451	<i>NI/n</i>	0.94	1	1	1	1	1	0.99
	<i>Na</i>	3	2	4	2	2	3	4
	<i>He</i>	0.18	0.28	0.56	0.31	0.43	0.52	0.66
	<i>Ho</i>	0.07	0.22	0.57	0.38	0.63	0.64	0.32
	<i>F<sub>IS</sub></i>	<b>0.66</b>	0.23	<b>1</b>	-0.2	-0.43	-0.21	
	<i>p</i>	<b>0.033</b>	0.396	<b>0</b>	0.582	0.152	0.28	
	<i>Nf</i>	0.28	0.04	<b>0.36</b>	0	0	0	
Ml_39C5118	<i>NI/n</i>	1	1	1	1	1	1	1
	<i>Na</i>	1	3	4	8	3	2	10
	<i>He</i>	0	0.45	0.59	0.64	0.23	0.5	0.72
	<i>Ho</i>	0	0.47	0.06	0.63	0.25	0.5	0.3
	<i>F<sub>IS</sub></i>	—	0.17	<b>0.9</b>	0.05	-0.08	0.03	
	<i>p</i>	—	0.304	<b>0</b>	0.488	0.8	0.625	
	<i>Nf</i>	—	0.04	<b>0.33</b>	0.01	0	0	
Ml_106C203	<i>NI/n</i>	1	1	1	1	1	1	1
	<i>Na</i>	1	1	2	2	1	1	2
	<i>He</i>	0	0	0.06	0.26	0	0	0.47
	<i>Ho</i>	0	0	0.06	0.31	0	0	0.06
	<i>F<sub>IS</sub></i>	—	—	0	-0.15	—	—	
	<i>p</i>	—	—	0.8	0.682	—	—	
	<i>Nf</i>	—	—	0	0	—	—	
Ml_39C4313	<i>NI/n</i>	1	1	0.94	1	1	1	0.99
	<i>Na</i>	5	4	3	2	1	2	6
	<i>He</i>	0.6	0.6	0.29	0.45	0	0.5	0.59

	<i>Ho</i>	0.69	0.28	0.33	0.69	0	1	0.48
	<i>F<sub>IS</sub></i>	-0.11	<b>0.56</b>	-0.13	-0.5	—	<b>-1</b>	
	<i>p</i>	0.396	<b>0</b>	0.685	0.075	—	<b>0</b>	
	<i>Nf</i>	0	<b>0.2</b>	0.19	0	—	0	
Ml_39C37	<i>NI/n</i>	1	1	0.88	1	1	0.93	0.97
	<i>Na</i>	4	9	10	4	3	4	23
	<i>He</i>	0.56	0.8	0.84	0.53	0.27	0.7	0.86
	<i>Ho</i>	0.44	0.44	0.64	0.25	0.31	0.31	0.4
	<i>F<sub>IS</sub></i>	0.25	<b>0.47</b>	<b>0.27</b>	<b>0.55</b>	-0.12	<b>0.58</b>	
	<i>p</i>	0.186	<b>0</b>	<b>0.012</b>	<b>0.05</b>	0.675	<b>0</b>	
	<i>Nf</i>	0.08	<b>0.2</b>	<b>0.3</b>	<b>0.18</b>	0	<b>0.34</b>	
Ml_106C3942	<i>NI/n</i>	1	1	1	1	0.94	1	0.99
	<i>Na</i>	2	2	1	4	2	2	5
	<i>He</i>	0.22	0.11	0	0.61	0.28	0.5	0.63
	<i>Ho</i>	0.13	0.11	0	0.94	0.33	1	0.4
	<i>F<sub>IS</sub></i>	<b>0.46</b>	-0.03	—	<b>-0.52</b>	-0.17	<b>-1</b>	
	<i>p</i>	<b>0.009</b>	0.978	—	<b>0.009</b>	0.679	<b>0</b>	
	<i>Nf</i>	0.08	0	—	0	0.19	0	

Table S4: Pairwise multilocus estimates of  $F_{ST}$  ( $\theta$ ) (Weir and Cockerham 1984). Values above the diagonal correspond to estimates calculated using the uncorrected 9 loci dataset while the values below the diagonal correspond to estimates calculated using the 9 loci dataset corrected for null alleles using Brookfield equation 2 (Brookfield 1996) implemented in MICRO-CHECKER (Oosterhout et al. 2004). Regardless of the dataset tested, all estimates were associated with significant values of  $p$ , as assessed by running 1000 permutations using GENETIX 4.05 (Belkhir et al. 1996-2004).

	SAU	MIT	MTM	LEB	PIL	MLK
SAU	–	0.398	0.563	0.549	0.581	0.482
MIT	0.389	–	0.505	0.469	0.537	0.431
MTM	0.562	0.494	–	0.434	0.574	0.402
LEB	0.542	0.457	0.420	–	0.358	0.318
PIL	0.571	0.535	0.571	0.350	–	0.253
MLK	0.479	0.431	0.389	0.312	0.249	–

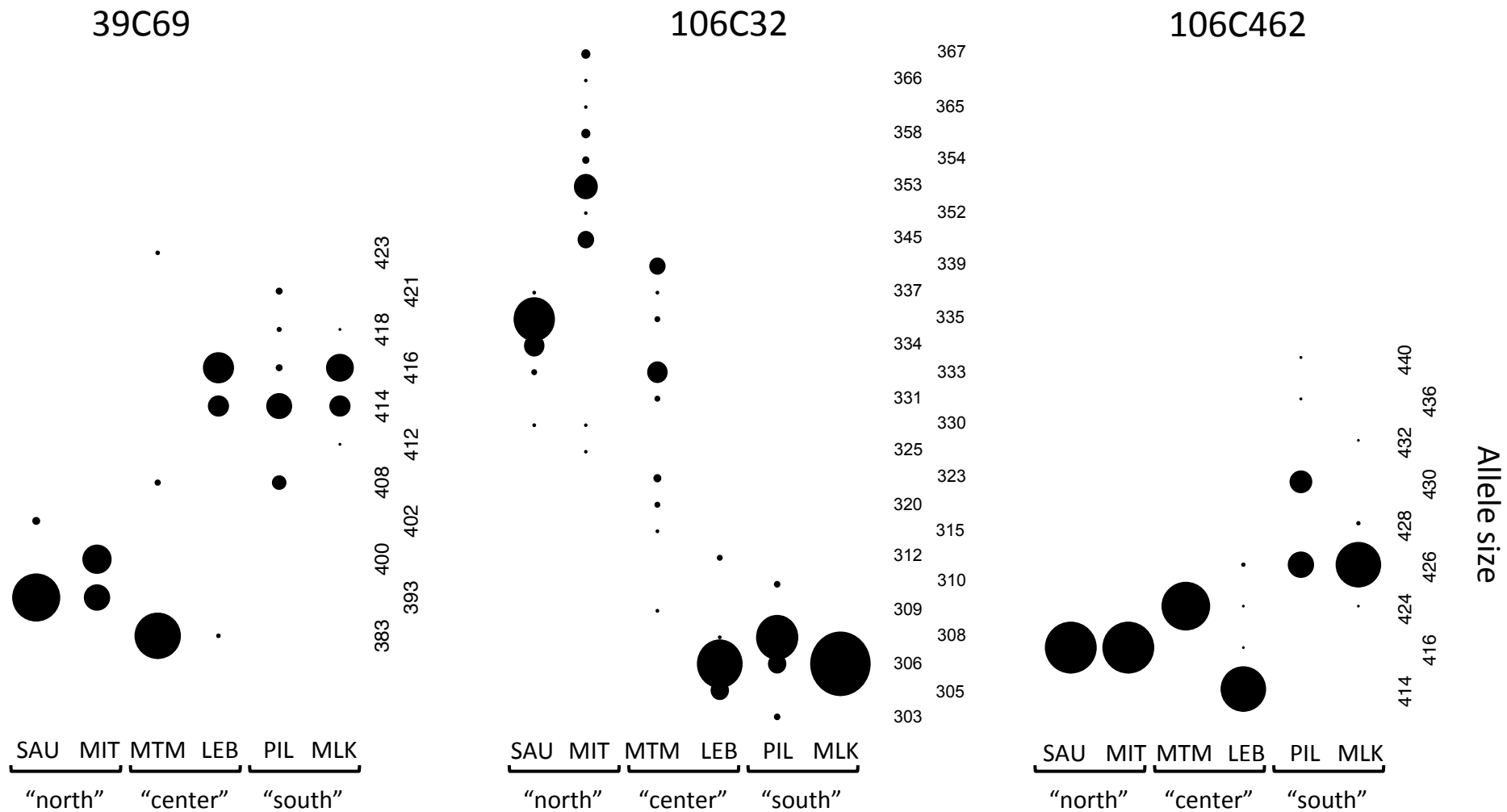


Figure S1: Allele frequency distributions for the nine microsatellite loci observed in each of the six study sites. Sites on the x-axis are ranged from north to south. Numbers on the y-axis are allele sizes in base pairs for each locus. Each circle indicates presence of corresponding allele; diameter of circle represents frequency of that allele in the site. Cytoplasmic lineages (*M. laminarioides* "north", "center" and "south") are noted below the site codes.

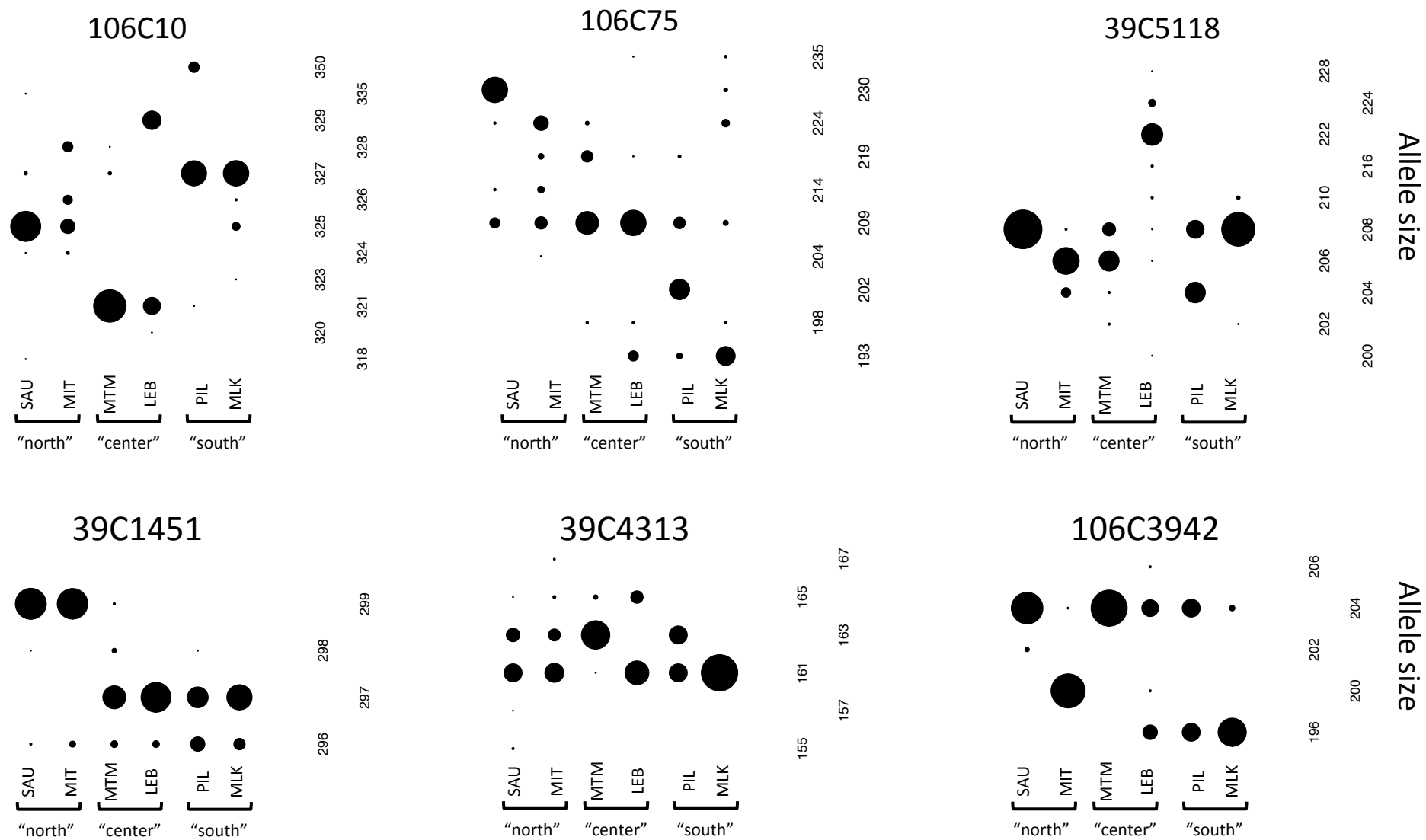


Figure S1: Continued from previous page

Figure S2: The posterior probability of the data given K (noted  $P(X|K)$ , dark circle) and  $\Delta K$  (white circle) results are given for  $K = 1$  to 7 using the results of the Bayesian analysis obtained with STRUCTURE.

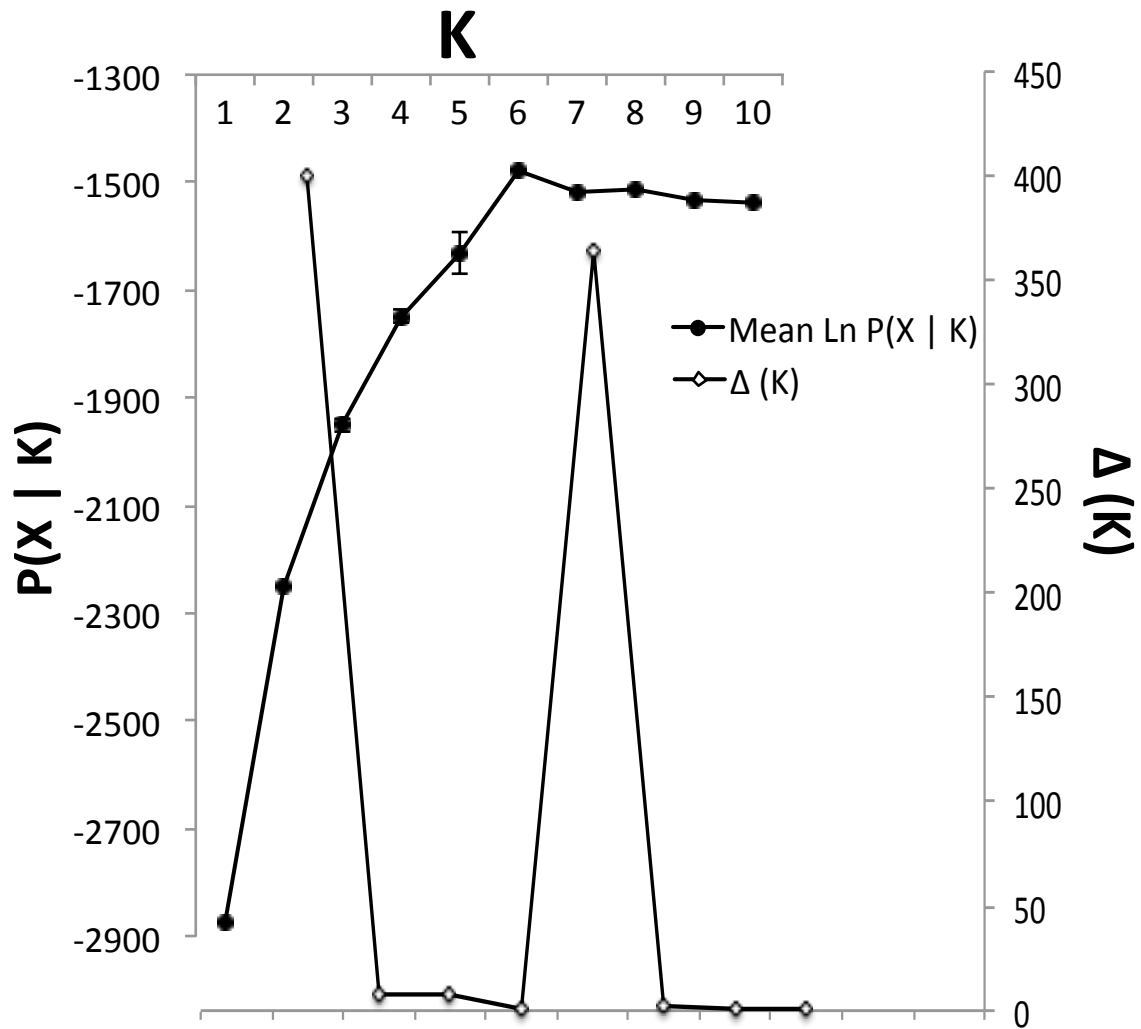


Figure S3: Bayesian analysis using INSTRUCT for the six study sites of *Mazzaella laminarioides*. A) Estimated Deviance Information Criterion (DIC) for values of K = 1 to 7. B) Results of the genetic clustering of the 96 *Mazzaella laminarioides* samples assuming K6. Each vertical bar represents a different individual. Each color represents the proportion of individual genome assigned to each genetic group and \* indicates individuals for which a secondary cluster represents more than 5% of the genome (i.e. admixed individuals). Individuals are ordered from north to south.

