



**HAL**  
open science

## Pre-domestication bottlenecks of the cultivated seaweed *gracilaria chilensis*

Oscar R Huanel, Suany Quesada-Calderón, Cristian Ríos, Sarai Morales-González, Pablo Saenz-agudelo, Wendy A Nelson, Natalia Arakaki, Stéphane Mauger, Sylvain Faugeron, Marie-laure Guillemin

► **To cite this version:**

Oscar R Huanel, Suany Quesada-Calderón, Cristian Ríos, Sarai Morales-González, Pablo Saenz-agudelo, et al.. Pre-domestication bottlenecks of the cultivated seaweed *gracilaria chilensis*. *Molecular Ecology*, In press, 10.1111/mec.16672 . hal-03773947

**HAL Id: hal-03773947**

**<https://hal.science/hal-03773947v1>**

Submitted on 9 Sep 2022

**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.

Huanel Oscar R (Orcid ID: 0000-0003-0592-8847)  
Saenz-Agudelo Pablo (Orcid ID: 0000-0001-8197-2861)  
Ríos-Molina Cristian (Orcid ID: 0000-0002-5641-6836)

PRE-DOMESTICATION BOTTLENECKS OF THE CULTIVATED SEAWEED  
*GRACILARIA CHILENSIS*

Oscar R. Huanel

Núcleo Milenio MASH, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Casilla 114-D, Santiago, Chile.

IRL 3614 Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université, Pontificia Universidad Católica de Chile, Universidad Austral de Chile, Station Biologique, Roscoff, France

Suany Quesada-Calderón

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

AUSTRAL-omics, Vicerrectoría de Investigación, Desarrollo y Creación Artística, Universidad Austral de Chile, Valdivia, Chile

Cristian Ríos

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

Sarai Morales-González

Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Casilla 567, Valdivia, 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/mec.16672](https://doi.org/10.1111/mec.16672)

Pablo Saenz-Agudelo

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

IRL 3614 Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université, Pontificia Universidad Católica de Chile, Universidad Austral de Chile, Station Biologique, Roscoff, France.

ANID- Millennium Science Initiative Nucleus (NUTME), Las Cruces, Chile

Wendy A. Nelson

National Institute of Water and Atmospheric Research, Wellington 6021, New Zealand.

School of Biological Sciences, University of Auckland, Auckland 1142, New Zealand.

Natalia Arakaki

Instituto del Mar del Perú, Banco de Germoplasma de Organismos Acuáticos, Esquina Gamarra y General Valle s/n, Chucuito, Callao, Perú.

Stéphane Mauger

IRL 3614 Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université, Pontificia Universidad Católica de Chile, Universidad Austral de Chile, Station Biologique, Roscoff, France.

Sylvain Faugeron

Núcleo Milenio MASH, Facultad de Ciencias Biológicas, Pontificia Universidad Católica

de Chile, Casilla 114-D, Santiago, Chile.

IRL 3614 Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université,  
Pontificia Universidad Católica de Chile, Universidad Austral de Chile, Station Biologique,  
Roscoff, France.

Marie-Laure Guillemin

Núcleo Milenio MASH, Instituto de Ciencias Ambientales y Evolutivas, Facultad de  
Ciencias, Universidad Austral de Chile, Casilla 567, Valdivia, Chile.

IRL 3614 Evolutionary Biology and Ecology of Algae, CNRS, Sorbonne Université,  
Pontificia Universidad Católica de Chile, Universidad Austral de Chile, Station Biologique,  
Roscoff, France.

Co-corresponding authors: Marie-Laure Guillemin: [marielaure.guillemin@gmail.com](mailto:marielaure.guillemin@gmail.com);

Sylvain Faugeron: [sfaugeron@bio.puc.cl](mailto:sfaugeron@bio.puc.cl)

Running title: Population genomics of *Gracilaria chilensis*.

## Abstract

*Gracilaria chilensis* is the main cultivated seaweed in Chile. The low genetic diversity observed in the Chilean population has been associated with the over-exploitation of natural beds and/or the founder effect that occurred during the post-glacial colonization from New Zealand. How these processes have affected its evolutionary trajectory before farming and incipient domestication is poorly understood. In this study, we used 2,232 SNPs to assess how the species evolutionary history in New Zealand (its region of origin), the founder effect linked to transoceanic dispersion and colonization of South America, and the recent over-exploitation of natural populations have influenced the genetic architecture of *G. chilensis* in Chile. The contrasting patterns of genetic diversity and structure observed between the two main islands in New Zealand attest to the important effects of Quaternary glacial cycles on *G. chilensis*. Approximate Bayesian Computation analyses indicated that Chatham Island and South America were colonized independently near the end of the Last Glacial Maximum and emphasized the importance of coastal and oceanic currents during that period. Furthermore, ABC analyses inferred the existence of a recent and strong genetic bottleneck in Chile, matching the period of over-exploitation of the natural beds during the 1970s, followed by rapid demographic expansion linked to active clonal propagation used in farming. Recurrent genetic bottlenecks strongly eroded the genetic diversity of *G. chilensis* prior to its cultivation, raising important challenges for the management of genetic resources in this incipiently domesticated species.

**Keywords:** domestication, genetic resource, phylogeography, genetic bottleneck, SNPs, Rhodophyceae.

## INTRODUCTION

The domestication process can be seen as a sustained relationship between humans and plants or animals during which these organisms are selected from wild populations and progressively evolve adaptations to agroecological niches (Zeder 2015). While domestication of terrestrial species started approximately 10,000 years ago (Zeder 2015), with more recent “pulses of domestication”, including the domestication of most fruit producing trees (i.e., apples, peaches, and almonds) around 4,000 - 3,500 years ago (Cornille et al. 2019, Zohary et al. 2012), most marine species have been cultivated for less than a century (Duarte et al. 2007). Seaweeds are among the most recently cultivated and incipiently domesticated organisms. Large scale seaweed cultivation began about 75 years ago (Valero et al. 2017). Eight species belonging to the genera *Gracilaria*, *Eucheuma*, *Kappaphycus*, *Pyropia*, *Saccharina*, and *Undaria*, represent the major part of commercially grown seaweed crops and are now showing signatures of phenotypic and genetic divergence with wild populations (Valero et al. 2017). Increasing focus on seaweed cultivation stemmed from the rapid increase in demands of food and chemical industries that cannot be met by harvesting biomass from natural populations (Buschmann et al. 2001, Hafting et al. 2015, Valero et al. 2017). Population genetic inferences of the domestication history of terrestrial plants have commonly revealed strong genetic bottlenecks due to selective sweeps and/or drift linked to the use of few/local seeds (i.e., Molina et al. 2011), sometimes with a recovery of genetic diversity after domestication as a result of introgressions with wild relatives (i.e., Hufford et al. 2012). The development of seaweed aquaculture can also lead to genetic bottlenecks and the erosion of genetic diversity (brown algae: Guzinski et al. 2018, Zhang et al. 2017; red algae: Guillemin et al. 2008, Hurtado et al. 2019). This reduction has been particularly severe in the case of algal crops propagated

Accepted Article

vegetatively with many of the farms established outside of the species native range using a very small number of individuals / biomass for seedling (i.e., *Gracilaria* farms in the northern part of Chile: Guillemain et al. 2008; *Eucheuma* and *Kappaphycus* farms in some 30 countries: Conklin and Smith 2005, Zuccarello et al. 2006, Chandrasekaran et al. 2008, Halling et al. 2013, Hayashi et al. 2017, Hurtado et al. 2019). Few innovations and investments have been made in Gracilariales and eucheumatoids to steady the loss of the crops' genetic diversity and to preserve genetic resources present in both natural populations and farming facilities (Brakel et al. 2021, Hurtado et al. 2019). This has been associated with rapid crop failure due to loss of strain vigor and the increase in pest, epiphyte, and disease susceptibility (Hayashi et al. 2017, Hurtado et al. 2019, Vairappan et al. 2008, Valderrama et al. 2015). There is an urgent need to implement long-term sustainable aquaculture practices to counter the setbacks being experienced.

On top of the problems resulting from current cultivation practices, the status of genetic resources prior to the establishment of farming operations is poorly understood. Most species have been harvested for decades, even centuries, before being cultivated (i.e., *Porphyra/Pyropia* spp. have been traditionally harvested for at least 900 years for their consumption as Nori, *Undaria pinnatifida* since 1600 to produce Wakame; Valero et al. 2017). How harvesting has affected population genetic diversity of seaweeds remains largely unknown (Faugeron et al. 2004). On a broader spatiotemporal scale, population dynamics and connectivity have promoted spatial genetic structure that has influenced the success of some breeding strategies. For instance, Asian kelps have experienced population divergence in glacial refugia (Zhang et al. 2015, 2019), that served as a basis for a breeding strategy promoting heterosis, by producing hybrid crosses between divergent populations (Goecke et al. 2020, Hwang et al. 2019, Wang et al. 2020).

*Gracilaria chilensis* is the main cultivated seaweed in Chile (Buschmann et al. 2017). The species has been an important economic resource in the country since the 1950s-60s, when intense harvesting of the natural populations began to supply the phycocolloid industry (Buschmann et al. 2008). Over-exploitation led to the collapse of natural populations in the late 1970s and hundreds of farms were created along the Chilean coast to sustain the production. Today, natural populations are found in a relatively narrow range, from the Bio-Bio region (36°S) to Chiloe Island (42.5°S), while farms are operating from Patagonia (43.5°S) to Northern Chile (23°S). The development of farms was initially limited by the difficulty of completing the sexual life cycle. Indeed, *Gracilaria* species have a triphasic life cycle, with diploid tetrasporophytes that release spores after meiosis, haploid male and female gametophytes that produce gametes, and the diploid carposporophyte that produces diploid spores surrounded by the haploid tissue of the maternal gametophyte. To circumvent this constraint, cultivation of *G. chilensis* is based on a very simple system of clonal propagation of cuttings, collected locally or sometimes exchanged among farmers, that are then grown entwined in ropes or planted in sandy or muddy substrates (Santelices and Doty 1989). This practice has interrupted the sexual life cycle. Currently, farms are dominated by heterozygous diploid individuals with very little spore production (Guillemin et al. 2008, 2014a), which has been interpreted as a signature of incipient domestication (Valero et al. 2017). Low genetic diversity has been reported in both farms and natural populations, and has been related to over-exploitation prior to the start of farming operations (Guillemin et al. 2008). However, another possible explanation for this low genetic diversity comes from phylogenetic and phylogeographic evidence that the species has colonized Chile from New Zealand at the end of the last glacial era (Cohen et al. 2004; Guillemin et al. 2014b; Huanel et al. 2020). Indeed, the species is present all around the two



Accepted Article

main islands of New Zealand and Chatham Island, where most of the genetic diversity is present (Huanel et al. 2020), and at least 11 species of Gracilariales are found (Gurgel et al. 2018; Neill and Nelson 2019; Preuss et al. 2020). Only *G. chilensis* is found in Chile and Peru. Such trans-oceanic dispersal, although likely rare, occurred in several seaweed and invertebrate species (Fraser et al. 2009; 2011; Guillemín et al. 2016). In *G. chilensis*, the founder effect of such trans-oceanic colonization was likely strong, as evidenced by the detection a single ribosomal or mitochondrial haplotype shared between New Zealand and Chile (Guillemín et al. 2014b). Yet, the observation of private alleles in Chilean populations of *G. chilensis* suggests that population expansion after colonization may have restored some level of genetic diversity (Guillemín et al. 2014b). However, the poor polymorphism of the markers used so far (ITS in Guillemín et al. 2014b and Huanel et al. 2020; microsatellites in Guillemín et al. 2008) did not allow precise and statistically powerful assessment of the historical demography of *G. chilensis*.

The relative roles of past founder effects and recent over-harvesting are not clearly understood and the central question addressed in this study is which of such previous events, either natural or human-driven, were most influential for the genetic basis of algal crops. This study aimed at deciphering the pre-domestication history of *G. chilensis* using genomic markers. Three main hypotheses were tested: the low genetic diversity found in Chile is the result of 1) the founder effect during the colonization of Chilean coasts; 2) the overharvesting prior to the onset of aquaculture; or 3) both events have strongly affected the genetic basis of this algal crop. To do so, we first analyzed the spatial structure of the genetic diversity in New Zealand to identify potential source populations. Then we inferred the transoceanic colonization history of Chile and assessed the effect of recent over-harvesting, considering the previous founder effects. Using samples from the whole

distribution range of the species, we analyzed natural populations from the native area, i.e., New Zealand, and natural and cultivated populations from Chile and Peru, using RADseq markers.

## MATERIAL AND METHODS

### *Sampling*

Samples were collected across the entire distributional range of *Gracilaria chilensis* in the Pacific Ocean (Guillemin et al. 2008, Guillemin et al. 2014, Robitzch et al. 2019, Huanel et al. 2020), including North and South Islands of New Zealand, Chatham Island, the south of Peru and Chile (Fig. 1, Table 1, Table S1). Samples were collected from both rocky shores (i.e., attached individuals resulting from spore settlement, and therefore from sexual reproduction) and soft bottoms (i.e., free-floating thalli propagating vegetatively) when available (Table S1). A total of 687 individuals of *G. chilensis* were sampled, including 439 samples from 22 populations in New Zealand and Chatham Island, 8 samples from the only population recorded in Peru and 240 samples from 13 Chilean populations (Table S1). In New Zealand, two species of *Gracilaria* are found in sympatry, *G. chilensis* and *G. transtasmanicum*, that are morphologically indistinguishable in the field (Preuss et al. 2020). The species identity of all samples from New Zealand was corroborated using the molecular diagnostic tool developed by Huanel et al. (2020). The ploidy of each mature individual was determined in the field by direct observation, as well as examination of reproductive structures under a binocular microscope, while amplification of sex molecular markers (Guillemin et al. 2012) was used for the vegetative individuals.

### *DNA extraction and ddRADseq library preparations*

Genomic DNA was extracted using NucleoSpin Plant II (Macherey Nagel, Düren, Germany), quantified using Picogreen assay kit (Molecular Probes, EEUU) and normalized to a standard concentration of 2.5 ng/μl. A double-digested restriction site-associated DNA (ddRAD) library was obtained using the methodology proposed by Peterson et al. (2012) with slight modifications. Briefly, the genomic DNA was digested with *HhaI* and *PstI*. The samples were amplified using Illumina indexed primers following standard protocol for multiplexing of Illumina. The combination of 24 different adapters and 12 primers allowed the multiplexing of 288 individuals per library. All individuals were then mixed to equimolar concentration and the DNA fragments were filtered to 400-800pb with a Pippin Prep (Sage Science). Paired-end sequencing was conducted on Illumina HiSeq2500 using Rapid Run option at Fasteris SA (Geneva, Switzerland).

#### *De novo assembly and filtration*

The libraries were demultiplexed using the *process\_radtags* module in STACKS v2.4 (Catchen et al. 2011). Reads with low quality value (i.e., Phred score of 10) or with ambiguous barcodes were removed. The *de novo* assembly was performed in STACKS v2.4 (Catchen et al. 2011), using the default parameters except for the number of mismatches allowed between stacks within individuals (*-M*), and the number of mismatches allowed between individuals for a given stack (*-n*), which were both set at 4 (parameters chosen following the optimization procedure described in Paris et al. 2017). Using the *population* module, only the first SNP per locus was retained. SNPs characterized by  $H_o > 0.50$  and minor allele frequency  $MAF < 0.02$ , present in less than 60% of the individuals ( $-r = 0.6$ ) and in less than 35 of the 36 populations ( $-p = 35$ ) were discarded. Resulting genotype data were exported as a vcf file and converted into the different file formats

necessary for the following analyses using the program PGDSpider v.2.1.1.5 (Lischer and Excoffier, 2012).

We also verified the ploidy (diploid vs haploid) assigned *a priori* by direct observation and sex markers using the observed heterozygosity ( $H_o$ ); all individuals were ordered from the lowest to the highest values of  $H_o$  in each population sampled and the values of putative haploid and diploid individuals were compared manually with the assigned ploidy ( $H_o = 0$  is expected for all haploid individuals when no paralogs are present). Non-congruent  $H_o$  values were observed for 19 (2.7%) individuals over the 687 sequenced and the ploidy of these individuals was corrected (Fig. S1). The ploidy of 50 vegetative individuals for which sex markers did not amplify was directly deduced from the  $H_o$  values (Fig. S1).

Finally, putative paralogs were identified as the loci for which  $H_o$  was higher than 0 in the corrected set of haploid individuals and were then removed from the data.

In the final data set haploids were artificially diploidized and appear as homozygous for all loci.

#### *Clonal membership assignment*

To remove clonal replicates from the dataset, clone assignment was performed for each population with the software GenoDive v3.0 (Meirmans 2020). GenoDive uses the genetic distance calculated between each pair of samples to define an assignment threshold that indicates the maximum dissimilarity allowed between two individuals to be considered as clones. We explored thresholds of 5, 6, 8, 10 and 11 mismatches between genotypes to assign clones in the complete dataset and selected the highest possible threshold that do not identify clones in individuals attached to hard substratum (i.e., individuals issued from sexual reproduction and spore settlement; Guillemin et al. 2008). Analyses were performed independently for diploids and haploids.

### *Genetic diversity*

The genetic diversity was evaluated for all populations and bioregions using various genetic diversity indices: mean number of alleles ( $N_a$ ), effective number of alleles ( $N_e$ ), percentage of polymorphic loci (%P), expected heterozygosity ( $H_e$ ), observed heterozygosity ( $H_o$ ), number of private alleles ( $A_{priv}$ ), and fixation index ( $F_{IS}$ ). All indices were calculated using GenAEx v6.5 (Peakall and Smouse 2012). The proportion of different genotypes was calculated as  $D = MLG / N$ , where MLG was the number of distinct Multi Locus Genotypes and  $N$  is the total number of studied individuals (Ellstrand and Roose 1987). Student T tests were performed to test for significant differences in genetic diversity ( $N_a$  and %P) between the New Zealand and South American *G. chilensis* populations after data transformation using boxcoxmix package in R Studio v1.2.1335 (R Core Team, 2019).

### *Population structure analysis*

Genetic structure in *G. chilensis* throughout the Pacific Ocean was characterized using four approaches. First, we estimated the pairwise  $F_{ST}$  in ARLEQUIN v3.5 (Excoffier and Lischer 2010), adjusting the  $p$ -values using a false discovery rate (FDR; Benjamini and Hochberg 1995).  $F_{ST}$  values were estimated using only diploid individuals. Second, we estimated the number of genetic clusters ( $K$ ) and the admixture coefficient for each individual, using an analysis of ancestral population assignment available in the package “LEA” in R. The number of ancestral populations tested ranged from  $K=1$  to  $K=36$  (10 repetitions), the best  $K$  was selected based on the cross-entropy criterion. Third, a Principal Component Analysis (PCA) was conducted using the function *pca* of the package “LEA” (Frichot and François 2015). Finally, a phylogenetic network was reconstructed with the algorithm Neighbor-Net, implemented in SplitsTree v4.14.1 (Klopper and Huson 2008). The matrix of genetic distances was calculated using uncorrected distances

(*Uncorrected\_P*) and ignoring the missing data. For the population assignment by LEA, the PCA and Neighbor-Net analyses, haploids and diploids were analyzed jointly.

#### *Inferring trans-Pacific colonization scenarios*

To determine the colonization scenarios most compatible with the observed SNP variation, we first used explicit model comparisons in the program DIYABC-RF (Collin et al. 2021). We examined the most probable origin and route of colonization for Chatham Island and South America from New Zealand. Based on our population structure analyses (PCA), we considered the existence of only five major genetic groups for the DIYABC-RF analyses, namely Northeast coast of North Island ( $N_1$ ; NE and ENI, Table 1), Northern Cook Strait ( $N_2$ ; NCS), South Island ( $N_3$ ; WSI, SCS, ESI, SOUT and FIOR, Table 1) and Chatham Islands ( $N_4$ ; CHT) in New Zealand, and Chile ( $N_5$ ). A subsample of 34 to 37 individuals (including both haploids and diploids characterized by less than 2% of missing data), selected randomly after discarding clones and representing equally all populations assigned to a genetic group, was used as representative of each genetic group. The population of Peru, previously reported as completely clonal (Robitzch et al. 2019), was not included. Likewise, populations from the west coast of North Island, extremely divergent from the rest of the *G. chilensis* sampled (see results), were not included.

To reduce computational efforts, the choice of the best scenario was made in two steps. First, we defined three groups of competing colonization scenarios. The first group of scenarios involved a common New Zealand origin for both Chatham Island and Chile to test different possible origins (Group 1, three scenarios: Sc1, Sc2 and Sc3; Fig. S2). In the second group of scenarios, the colonization followed a stepping-stone process from New Zealand to Chatham Island and then Chile (Group 2, three scenarios: Sc4, Sc5 and Sc6; Fig. S2). The third group of scenarios inferred two independent origins of colonization for

Accepted Article

Chatham Island and Chile (Group 3, six scenarios: Sc7 to Sc12; Fig. S2). Following Guillemin et al. (2014b), the time of colonization from New Zealand to Chatham Island was set as older than the one from New Zealand to Chile and we assumed the existence of a bottleneck associated with each founding event. The prior distributions of historical and demographic parameters are described in Table S2. Training sets included 20,000 and 100,000 simulated datasets per scenario for the scenario choice and parameter estimation steps, respectively, as suggested (Collin et al. 2021). Within each group of colonization scenarios, the most likely scenario was identified using a RF tree-based classification method. The RF classifier was constructed using 1,000 trees per scenario group and the scenario receiving the most votes over all trees was selected. The global and local error rates were estimated by RF regression and used to calculate the posterior probability of the selected scenario (Pudlo et al. 2016). In a second round of scenario testing, the best scenario of each group was kept, and these three remaining scenarios were compared using the same method as described above.

We estimated the posterior distributions of demographic parameters under the best scenario and the precision of each parameter estimation by computing the median as point estimates for each parameter, a 90% credibility interval for each point estimate and global and local accuracy indices using the approach developed by Raynal and collaborators (2019). A total of 10,000 bootstrap samples, randomly chosen from the 100,000 simulated datasets, were used for parameters estimation. These analyses were replicated ten times.

To evaluate the existence, magnitude, and direction of gene flow during the trans-Pacific colonization, the DIYABC-RF analyses were complemented by Diffusion Approximations for Demographic Inference ( $\delta\text{a}\delta\text{i}$ ) analyses (Gutenkunst et al. 2009). Since DIYABC-RF analyses supported a distinct origin for Chatham Island and South America from New

Zealand (see results for more details), we independently evaluated six alternative 2D models to determine the optimal scenario for the colonization of Chatham Island ( $N_4$ ) from the Northeast coast of North Island ( $N_1$ ) and of Chile ( $N_5$ ) from the North Cook Strait ( $N_2$ ). We used the script easySFS.py (available at <https://github.com/isaacovercast/easySFS>) to generate a joint-site frequency spectrum (SFS) for  $N_4$  and  $N_1$  and for  $N_5$  and  $N_2$ . The seven alternative 2D models followed the analysis pipeline of Portik et al. (2017): 1) founder effect with no migration, 2) founder effect with constant symmetric migration, 3) founder effect with constant asymmetric migration (i.e., a higher gene flow from the region of origin to the colonized area), 4) divergence with ancient symmetric migration, 5) divergence with ancient asymmetric migration, 6) allopatric divergence without migration, and 7) divergence with constant asymmetric migration. A visual representation of the demographic models can be found in the Fig. S3. Model optimization routine was performed using the Nelder-Mead method following the pipeline (Optimize\_Functions.py) implemented by Portik et al. (2017). We used four rounds of parameter optimizations with [10, 20, 30, 40] replicates and [5, 20, 125, 200] maximum iterations per round, respectively. The most likely model for each dataset (i.e.,  $N_4$ - $N_1$  and  $N_5$ - $N_2$ ) was chosen based on the lowest AIC score.

In order to estimate the timing of demographic events we used a mutation rate of  $1 \times 10^{-7}$  (mutation rate inferred for *Arabidopsis*; Bashir et al. 2014) in  $\delta a \delta i$ . No mutation rate for SNPs is available in red algae. The *Arabidopsis* mutation rate was chosen considering that red algae, jointly with Chloroplastida (a lineage including green algae and land plants) and glaucophytes, form the supergroup Archaeplastida (Burki 2014). In DIYABC-RF, SNPs do not require mutation model parameterization, and the algorithm of Hudson (2002) was used for all biallelic polymorphic loci as recommended by the authors (Collin et al. 2021). The



Accepted Article

timing - in years - of demographic events was estimated using a generation time of 1.5 years (following results obtained for population dynamics in *G. chilensis*; Vieira et al. 2021).

#### *Testing for scenarios of multiple bottlenecks in Chile*

We tested whether Chilean populations of *G. chilensis* passed through simple or multiple bottlenecks prior to the establishment of seaweed farms, including an old founder event linked to the colonization from New Zealand and a more recent demographic bottleneck associated with over-exploitation. We compared four scenarios using DIYABC-RF: 1) a stable demographic history over time in Chile, 2) the existence of an old bottleneck associated with the founder event, 3) the existence of a recent bottleneck linked to anthropogenic impacts, and 4) the existence of both old and recent bottlenecks in Chile.

The Chilean population was rooted in the New Zealand source population identified from the Trans-Pacific inferences. Parameter prior distributions are shown in Table S3 and a graphic representation of the four scenarios tested is given in Fig. S4. For RF analyses, model choice and estimation of posterior distributions of demographic parameters we followed the protocols described above.

The DIYABC-RF analysis was complemented with a  $\delta a \delta i$  analysis (Gutenkunst et al. 2009) by testing between various 1D models of potential historical demographic changes affecting the Chilean population. We examined the following four alternatives: (1) constant population size model, (2) two-epochs with instantaneous size change, (3) two-epochs with exponential growth, (4) two-epochs with instantaneous size decline followed by exponential growth (i.e., bottle-growth), and (5) three-epochs with instantaneous size change (i.e., variation in population size occurred twice, once in a remote past and once in a more recent past). The most likely model for each dataset was chosen based on the lowest

AIC. As above, for  $\delta a\delta i$  analyses, timing was estimated using a generation time of 1.5 years (Vieira et al. 2021) and a mutation rate of  $1 \times 10^{-7}$  (Bashir et al. 2014).

## RESULTS

An average of 2.5 million reads per individual (ranging from 696,000 to 4,500,000 reads) were obtained for the 687 *G. chilensis* samples. After removing low quality reads, 1.8 billion reads were retained across all samples. On average, RAD loci were of 132 bp in length and the individual depth coverage was of 36.8 reads per locus ( $\pm 16.6SD$ ). The SNP dataset after *de novo* assembly and filtering contained 2,232 biallelic SNPs.

### *Clonal diversity and genetic diversity*

The results obtained in GenoDive, based on thresholds ranging from five to eleven, identified from 24 to 30 repeated genotypes respectively, in our dataset. Whatever the threshold used, the number of clones (i.e., ramets) of a repeated genet detected in each population was highly consistent (Fig. S5) and we therefore set the threshold distance to five for both diploids and haploids for further analyses. Of the 24 repeated genotypes detected, 11 were from New Zealand and 13 from South America, none of which were shared. Over a total of 150 clones detected with the threshold distance of five (Fig. S5), 111 clones (74%) were identified in Peru and Chile and 39 (26%) in New Zealand. The majority of clones were free-floating thalli growing on sand or mud (Peru and Chile: 103 of 111, 93%; New Zealand: 31 of 39, 80%). These results are consistent with the higher number of different genotypes observed in attached sexual populations (New Zealand:  $0.95 < D < 1.0$ ; South America:  $0.75 < D < 1.0$ ) than in floating ones (New Zealand:  $0.38 < D < 1.0$ ; South America:  $0.14 < D < 0.80$ ) (Table S1). The differences between both types of population are more substantial in Chile where asexual reproduction in floating populations was actively promoted by the vegetative propagation by farmers (Table S1). Within each population,

only one sample of each clonal genotype was retained for further analyses of genetic structure (dataset without clones includes 263 haploids and 308 diploids).

The mean number of alleles and percentage of polymorphic alleles were higher in New Zealand than South America ( $N_a = 1.12$  and  $1.07$ ,  $t = 2.35$ ,  $P = 0.03$ ;  $\%P = 13.17$  and  $\%P = 6.60$ ,  $t = 2.71$ , Student T test  $P = 0.01$ , values are given for New Zealand and South America, respectively, Table 1, Table S1). In New Zealand, the highest genetic diversity was observed in the northeast coast of North Island (North Eastern; NE:  $N_a = 1.27$ ;  $\%P = 35.89$ ) followed by the rest of the bioregions of North Island and the South coast of Cook Strait (WNI, ENI, NCS and SCS:  $1.12 < N_a < 1.26$ ;  $22.49 < \%P < 31.59$ ). The diversity of New Zealand bioregions located south of  $41^\circ\text{S}$  on the South Island and Chatham Island are similar to that encountered in South America (WSI, ESI, SOUT and FIOR:  $1.03 < N_a < 1.06$ ;  $2.96 < \%P < 10.66$ ;  $0.01 < H_e < 0.02$ ; CHT:  $N_a = 1.12$ ;  $\%P = 7.44$ ;  $H_e = 0.03$ ). These results are consistent with the number of private alleles detected in our dataset for which the highest values were observed in the West Coast North Island, the North Eastern and the East Coast North Island ( $A_{\text{priv}} = 24$ ; 24 and 25 for WNI, NE and ENI, respectively; Table 1, Table S1).

$F_{IS}$  values were negative in all New Zealand populations except Ohope. In Chile, attached (sexual) populations had values close to 0 ( $-0.12 < F_{IS} < 0.01$ ; Table S1), except for Metri (NMET:  $F_{IS} = -0.50$ ). Contrastingly, floating (clonal) populations had strongly negative  $F_{IS}$ -values when compared with attached (sexual) ones (i.e.  $-0.02 < F_{IS} < -0.21$ ), the most extreme values being observed in Chilean farms ( $-0.21 < F_{IS} < -0.72$ ).

#### *Population differentiation and genetic structure*

Strong and significant differentiation was detected among populations, except for a few pairs of populations located in Chile (Table 2). The highest  $F_{ST}$ -values were observed

Accepted Article

separating populations of Shelly Beach and Raglan sampled in the West Coast North Island (WNI) from the rest of *G. chilensis* ( $0.88 < F_{ST} < 0.96$ ; Table 2). Pairwise  $F_{ST}$  values were smaller when comparing populations from Chatham Island (CHT) with the North Eastern (i.e., NE, Whitianga and Ohope,  $0.57 < F_{ST} < 0.65$ ) than with the rest of the New Zealand samples. The lowest pairwise  $F_{ST}$ -values between Chilean populations and populations from New Zealand were observed for the North Cook Strait (i.e., NCS, Paremata and Pauatahanui,  $0.55 < F_{ST} < 0.74$ ).

The  $\Delta$  cross-entropy method implemented in LEA suggested the existence of 8 clusters as the most likely population structure (Fig. S6). At  $K = 2$ , populations Shelly Beach and Raglan located in the West Coast North Island were clearly separated from the rest of *G. chilensis* (Fig. 1). At  $K = 3$ , South American populations appeared as a distinct genetic cluster and populations from the North Cook Strait (i.e., Paremata and Pauatahanui) were identified as admixed between South America and New Zealand (except the West Coast North Island) (Fig. 1). At  $K = 4$ , samples from Chatham Island, the North Eastern (i.e., Whitianga and Ohope) and the East Coast North Island (i.e., ENI, Ahuriri) were assigned to the same genetic cluster and clearly separated from the rest of the New Zealand populations (Fig. 1). Populations from the North Cook Strait formed their own, non-admixed genetic cluster at  $K = 5$  (Fig. 1). At  $K = 6$ , samples from Whitianga and Ohope were separated from Chatham Island and Ahuriri (Fig. 1). At  $K = 7$ , the populations from the South Island were separated in two genetic clusters with one cluster located at the northwest corner of the South Island, close to the Cook Strait (i.e., Okari River, Whanganui Inlet and Moutere Inlet) and the second cluster grouping all the other populations of the South Island (i.e., from South Brighton to Riverton) (Fig. 1). The two South Island populations located in between these two genetic clusters (i.e., Momorangi Bay in the northeast and Fiordland in

the southwest) showed a mixed ancestry. Finally, at  $K=8$ , Momorangi Bay appeared as a distinct cluster within the Cook Strait (Fig. 1).

The PCA confirmed the results obtained with clustering analysis implemented in LEA. When implemented with the whole dataset, the first principal component (PC) explaining 40.54% of the total variance clearly separated the populations Shelly Beach and Raglan from the rest of *G. chilensis* (Fig. 2A). When excluding Shelly Beach and Raglan (Fig. 2B), PC1 explained 32.05% of the variance and roughly separated South America from New Zealand, with the samples of Paremata and Pauatahanui located in the North Cook Strait showing an intermediate position. When considering only New Zealand samples (Fig. S7), Chatham Island (CHT) clustered with the North Eastern (i.e. NE; Whitianga and Ohope) and appeared close-by to the East Coast North Island (i.e. ENI; Ahuriri).

Two major clusters were distinguished by the Neighbor-Net reconstruction, with a strong split between Shelly Beach and Raglan and the rest of *G. chilensis* (Fig. 3). Twenty genetic groups, mostly corresponding to sampled localities or bioregions, were observed among the remaining populations. All South American samples clustered together and were closely related to samples from Paremata and Pauatahanui (in Fig. 3 PER-INT-MAG and NCS, respectively). The cluster formed by the Chatham Island samples (CHT) showed a close relationship with the East Coast North Island (i.e., ENI; Ahuriri).

#### *Inferring pathways of trans-Pacific colonization*

After the first round of scenario testing in DIYABC-RF a single scenario clearly dominated the others in terms of number of votes and posterior probability within each group of scenarios (Sc3 in Group 1, Sc6 in Group 2, and Sc10 in Group 3; see Table S4 for more details). The second round of scenario testing indicated that Sc10 was the most likely

Accepted Article

of the colonization models (534/1,000 votes; posterior probability: 0.77; Table S5). The lower values of both global and local error rates were obtained for Sc10 (i.e., 0.034 and 0.23, respectively; Table S5), also supporting this model choice. Sc10 consists of two independent founder events linked to the colonization of Chatham Island from the east coast of the North Island and of Chile from the North Cook Strait (Fig. 4, Fig. S2). Parameter estimates for Sc10 had large confidence intervals and absolute errors (especially for the population size during founder events in Chile and Chatham Island, corresponding to parameters  $N_{5b}$  and  $N_{4b}$ , respectively; see Table S6). Assuming one generation every 1.5 years for *G. chilensis* (Vieira et al. 2021), Sc10 suggests that Chatham Island was colonized some 23,000 years ago while the colonization of Chile occurred 19,000 years ago (Chatham Island:  $t_2$  median; 90% CI = [23,588; 11,129]; Chile:  $t_1$  median; 90% CI = [19,871; 11,396]; Table S6). Population sizes during founder events were poorly defined but the analyses suggested that both Chatham Island and Chile experienced a strong bottleneck (Chatham Island:  $N_{4b}$  median, 90% CI = [7,148; 30,278]; Chile:  $N_{5b}$  median, 90% CI = [7,295; 29,671]; Table S6).

Analyses in  $\partial a \partial i$  inferred a model with founder effect and complete absence of gene flow after the initial split between the Chilean and the North Cook Strait populations (best-fit model corresponded to Fig. S3A; AIC = 171.23). The model suggests a population reduction of about 98% during the Chilean founder event. Contrastingly, the best-fit model for the colonization of Chatham Island from the North Island supported the existence of a founder effect and a constant asymmetric gene flow after population divergence (best-fit model corresponds to Fig. S3C; AIC = 1145.46). The model suggests the existence of low gene flow from the North Island to Chatham Island (i.e., approximately 2 effective migrants

per generation). The model also suggests a population reduction of 98% during the founder event in the Chatham Island.

#### *Testing for complex demographic history in Chile*

DIYABC-RF analyses supported two alternative models for the demography of Chilean *G. chilensis*: the existence of a recent bottleneck alone (435 classification votes over 1,000 RF-trees) and the existence of both an old and a recent bottleneck in Chile (485 classification votes over 1,000 RF-trees; Table S7). However, the scenario supporting two bottlenecks was always the best-supported through the ten replicated analyses and it also showed the highest posterior probability ( $p=0.431$ ; Table S7). Scenario 4 suggests the existence of a long genetic bottleneck (i.e., 2,000 generations) 23,000 years ago and a much shorter and recent one (i.e., 49 years ago, and 37 generations). The reduction of the effective population size during the more recent bottleneck is much stronger than the one associated with the founder effect (effective population size during recent bottleneck:  $N_{2r}$  median, 90% CI = [33; 34]; effective population size during founder effect:  $N_{2f}$  median, 90% CI = [18,311; 29,705]; Table S8). A slow and slight increase in population size, up to twofold, was inferred for the founder event (effective population size during founder effect:  $N_{2f}$  median = 18,311; posterior to the founder effect:  $N_{2p}$  median = 33.081). In contrast, a strong and fast demographic expansion was inferred for the recent events (corresponding to 1850-fold increase in 2 to 3 decades; effective population size during recent bottleneck  $N_{2r}$  median = 33; current effective population size:  $N_2$  median = 60,948. The results should, however, be taken with caution since large CI 90% estimates and high global and local error rates were associated with all demographic bottleneck parameter estimates, reflecting poor accuracy (Table S8).

In  $\delta a \delta i$ , the 1D demographic model with the highest AIC was the constant population size scenario (AIC = 294.46). The other four demographic models had a very close fit that did not allow a clear-cut selection of the best-fit model (AIC between 154.30 and 158.22).

## DISCUSSION

Historical processes, especially those associated with Pleistocene climatic variations leading to genetic divergence and bottlenecks of populations isolated in glacial refugia, have shaped the genetic basis for domestication in a large number of terrestrial species (i.e., temperate species as walnuts, peaches and apples: Feng et al. 2018, Pollegioni et al. 2017, Sun et al. 2020, Yu et al. 2018; and tropical species as cacao and chili pepper: Thomas et al. 2012, Scaldaferrero et al. 2018). Here, we discuss three main aspects of the pre-domestication history of *G. chilensis*: 1) its evolutionary history in New Zealand, its region of origin, 2) the process of colonization of Chilean coasts, and 3) the more recent effects of harvesting and onset of cultivation.

### Evolutionary history of *G. chilensis* in its region of origin

The coastal biogeography of New Zealand is characterized by several boundaries observed in macroalgae (i.e., bull kelp species: Collins et al. 2010; Fraser et al. 2010; Laminariales: Zuccarello and Martin 2016; and red algae: Muangmai et al. 2014, 2015, 2016), marine angiosperms (Jones et al. 2008) and benthic marine invertebrates (i.e., Schnabel et al. 2000, Stevens and Hogg 2004, Veale and Lavery 2011). Their location corresponds to major shifts in ocean currents, eddies and upwelling around New Zealand (Apte and Gardner 2002, Ross et al. 2009, Veale and Lavery 2012). Our study revealed the existence of seven genetic clusters in *G. chilensis*, mostly congruent with the described boundaries between bioregions (Ayers and Waters 2005, Ross et al. 2009, Shears et al.



2008). The main boundaries observed in *G. chilensis* were detected around 42°S, on both the west and the east coasts. These areas are characterized by strong and regular upwelling regimes that represent important dispersal barriers due to offshore water movement (Apte and Gardner 2002, Goldstien et al. 2006, Ross et al. 2009). The genetic discontinuity observed at East Cape, on the northeast of the North Island, separating the northernmost populations of Whitianga and Ohope from Ahuriri has been attributed to the permanent anticyclonic East Cape eddy that causes offshore displacement of the East Auckland current and contributes to gene flow restrictions (Shears et al. 2008, Stevens and Hogg 2004, Ross et al. 2009, Zuccarello and Martin 2016). Moreover, the convergence between the warm subtropical East Auckland current and cool northward Wairarapa coastal current may also act as effective dispersal barrier at East Cape (Chiswell 2001, Zuccarello and Martin 2016). The cluster formed by Shelly Beach and Raglan in the Western North Island was highly distinct from all other localities studied, in agreement with previous ITS2 analyses (Huanel et al. 2020). However, while the ITS2 marker only detected a moderate amount of genetic difference between the Western North Island and the rest of New Zealand, our SNPs dataset supports the hypothesis of long-term complete isolation after a possible ancient introgression between populations from the Western North Island and Northern Cook Strait (see clustering analysis implemented in LEA for  $K = 2$ ; Fig. 1) that can be interpreted as ongoing speciation. Huanel and collaborators (2020) suggested that the Western North Island genetic cluster corresponds to a divergence predating the Last Glacial Maximum (LGM). During the Pliocene, diverse tectonic and climatic changes split the North Island in a complex of small islands (Craw 1988, Trewick and Bland 2012), with probably limited connectivity due to sea level decrease, especially during the Pliocene-Pleistocene transition (Stevens and Hogg 2004). These geographic changes have been related to strong population

Accepted Article

divergence both in terrestrial (Craw 1988, Trewick and Wallis 2001) and marine New Zealand taxa (Stevens and Hogg 2004). Allopatric divergence may have begun during the Pliocene-Pleistocene transition and was likely reinforced by the closure of Cook Strait during LGM (Alloway et al. 2007, Fleming 1975), leading to the complete isolation of the Western North Island.

Pleistocene glacial-interglacial dynamics also strongly affected the distribution of the genetic diversity of *G. chilensis* along the southernmost coasts of New Zealand. Indeed, the highest level of genetic diversity was observed in populations located in the North Island (i.e., particularly in the North Eastern bioregion) and in Cook Strait, while the South Island populations were characterized by low level of genetic diversity. All highly divergent genetic clusters were also observed in the North Island and Cook Strait while South Island populations were more genetically homogeneous. This last result is confirmed by the distribution patterns of private alleles along the New Zealand coasts (i.e., higher in the North than in the South Island). The South Island, particularly its west coast, was strongly affected by the advance and retreat of ice cover during the Otira Glaciation (22,000-14,000 years ago), leading to extinction-recolonization dynamics seen in terrestrial and marine taxa (Craw 1988, Lorrey and Bostock 2017, Main 1989, Stevens and Hogg 2004.). Southern postglacial demographic expansion from one or multiple small northern refuges has been hypothesized (Buchanan and Zuccarello 2012, Huanel et al. 2020, Marshall et al. 2009, Wardle 1963). These results are congruent with marine paleo-evidence suggesting cold temperatures at high latitude areas, with at least 4°C less than the current average along the east coast of the South Island of New Zealand (Kohfeld et al. 2013, Lorrey and Bostock 2017). These low temperatures suggest that the southern part of the South Island was not a suitable thermal habitat during glacial periods for *G. chilensis*, a temperate species that can

tolerate temperature as low as 5°C but for which the optimum temperature for growth is 15°C (McLachlan and Bird 1984, Usandizaga et al. 2019). Patterns of genetic diversity and structure observed with our SNPs dataset suggest post-glacial colonization from at least 2 refugia located north of 42°S, one probably located at the eastern opening of Cook Strait and the other one in Tasman/Golden Bay. Indeed, admixture in the Fiordland bioregion, a region with coastal areas highly impacted by ice scour during LGM (Alloway et al. 2007, Craw 1988, Lorrey and Bostock 2017, Trewick and Bland 2012) suggests a secondary contact of two independent recolonization fronts, one descending the east coast of the South Island and the other the west coast.

### **Transpacific colonization, founder effect and genetic bottlenecks**

Passive long-distance dispersal is an important promoter of range expansion and species radiations in marine organisms across the southern hemisphere (Gillespie et al. 2012, Muñoz et al. 2004). Our data show that two independent events of colonization led to the expansion of *G. chilensis* distribution area in the South Pacific at the end of the Pleistocene (Fig. 4). Demographic inferences estimated a colonization of Chatham Island some 23,000 years ago while the colonization of South America occurred 19,000 years ago. This is highly concordant with the estimations based on ITS2 (Guillemin et al. 2014b) and with paleo-reconstructions suggesting a reinforcement of the Antarctic Circumpolar Current (ACC) for several thousands of years, with an intensification of the westerly winds during glacial maxima (Kohfeld et al. 2013, Lorrey et al. 2012, Lorrey and Bostock 2017). Wind and currents favored a higher rate of eastward propagule transport from New Zealand main islands to Chatham Island and South America (Kaiser et al. 2005, Lorrey et al. 2012).

Chatham Island and South American colonization events were characterized by important differences in terms of the area of origin, route of colonization, and level of gene

flow between the area of origin and the newly colonized area. The colonization of Chatham Island from the North Eastern / Eastern North Island bioregions was probably facilitated by the jets and eddies (as the East Cape and Wairarapa eddies, Stevens et al. 2021) connecting to offshore main currents flowing north of the Chatham Rise (Fig. 4). Similarly, genetic data supported a colonization of Chatham Island from North Island for the chiton *Sypharochiton pelliserpentis* (Veale and Lavery 2011), the scallop *Pecten novaezelandiae* (Silva and Gardner 2015), and the clam *Austrovenus stutchburyi* (Ross et al. 2012). On the east coast of New Zealand main islands, both the southward subtropical East Cape Current and the colder Southland Current divert offshore when they encounter the Chatham Rise, flowing eastwards along the rise where they generate a mixture of subtropical and subantarctic waters (i.e., the Subtropical Convergence) (Barnes 1985, Chiswell 2005, Stevens et al. 2021). Despite their geographic isolation (i.e., 700 km), populations from the Chatham Island were remarkably genetically similar to the ones the North Eastern / Eastern North Island (especially to Ahuriri), suggesting that some degree of directional gene flow occurred - and probably still occurs - from that region. Moreover, the observation of a persistent component of the South Cook Strait (i.e., South Island) in Chatham genotypes suggest some gene flow from this region may have occurred as well, leading to this apparent admixture (see Fig. 1). Other studies reported frequent exchanges of migrants between mainland New Zealand and Chatham Island, these being mainly unidirectional from New Zealand to Chatham Island (Silva and Gardner 2015, Veale and Lavery 2011). Currently, with a three to four months estimated dispersal time to Chatham Island from the East Cape (Chiswell and Rickard 2011), a low level of genetic connectivity could be ensured.

Contrastingly, our data suggest a single event of colonization of South America from populations located in the North Cook Strait (Fig. 4), without any subsequent gene flow between the two areas. Distinct water currents could allow an eastward dispersal of marine organisms from the North Cook Strait through the South Pacific. First, the eastward D'Urville Current flows directly through Cook Strait and, second, a component of the northward Southland Current washes along the coasts of Cook Strait between Cloudy Bay and Turakirae Head (Stevens et al. 2021). Both water masses mix with the East Cape current before heading eastward (Barnes 1985). Other studies have proposed a New Zealand origin of Chilean populations for various marine organisms, including brown macroalgae (*Adenocystis utricularis*, Fraser et al. 2013; *Durvillaea antarctica*, Fraser et al. 2009; *Macrocystis pyrifera*, Macaya and Zuccarello 2010), red macroalgae (*Bostrichia intricata*, Fraser et al. 2013; *Capreolia implexa*, Boo et al. 2014), green macroalgae (*Wittrockiella lyallii*, Boedeker et al. 2010) and mollusks (*Ostrea chilensis*, Foighil et al. 1999), among others. However, contrary to *G. chilensis*, most studies on macroalgae have proposed a South Island (Boedeker et al. 2010, Boo et al. 2014) or subantarctic origin of the Chilean populations (Fraser et al. 2013, Macaya and Zuccarello 2010). However, it has to be noted that the limited statistical power of these studies, due to the low variability of the markers, do not clearly pinpoint the area of origin of population expansion. The hypothesis of a single dispersal event is further supported by the reduced speed of the West Wind Drift, which may take up to 3 years for a 14,000 km-long rafting journey (Smith et al. 2018). Even though eastward wind can speed-up the displacement of floating rafts, such long-distance dispersal can be expected to be rare (Cumming et al. 2014, Gillespie et al. 2012, Nathan 2006, Smith et al. 2018). However, it has been recently demonstrated that floating macroalgae can travel for years in the Southern Ocean, carrying a rich flora and

fauna on their thalli over distance of thousands of kilometers (up to 25,000 km; Fraser et al. 2018).

### **Pre-domestication losses of genetic diversity**

Despite the fact that seaweed use traces back to Neolithic times (Dillehay et al. 2008) and seaweed farming is currently rapidly expanding and intensifying, many key aspects of seaweed domestication and resource management are still poorly studied (Brakel et al. 2021, Cottier-Cook et al. 2021, Valero et al. 2017). The availability of genetic diversity is *a priori* critical for selection to operate, and ultimately for domestication to proceed. In the case of *G. chilensis* however, two non-exclusive processes may have driven a genetic erosion associated with the onset of farming operation: first, the collapse of natural populations that occurred in the late 1970s – early 1980s (Avila and Seguel 1993, Norambuena 1996); second, the massive use of clonal propagation. Partially clonal reproduction leads to a rapid dominance of a few clones during population expansion (Rafajlović et al. 2017). In haploid-diploid organisms, this dynamic is further impacted by genetic drift during the haploid phase, which accelerates the loss of genetic diversity (Stoeckel et al. 2021). While the best-supported ABC scenario included both the ancient founder effect and the recent bottleneck, the scenario with only a recent bottleneck was also well supported, suggesting a strong effect of overharvesting on present-day genetic diversity. The estimated timing of the genetic bottleneck matches the onset of this overharvesting period, further confirming its impact on the genetic diversity. ABC inferences also support a rapid increase in effective population size associated with the expansion of farms in Chile; but the speed and magnitude of this increase seems difficult to reconcile with a simple shift in mutation-drift equilibrium. Two lines of hypotheses associated with farming practices may explain the observed results. First, because of the haploid-diploid

life cycle, partially clonal propagation in seaweeds may lead to a rapid increase in effective population size when the population is dominated by diploids (Stoeckel et al. 2021), which is the case in *Gracilaria* farms (Guillemin et al. 2008). Second, farms are dominated by highly heterozygous genotypes, and their allelic richness does not differ from natural populations (Guillemin et al. 2008). Because ABC estimations of present-day effective population size rely essentially on the number of lineages (i.e., alleles) in the populations, the inference of a rapid demographic growth after the bottleneck could be explained by the persistence of a large number of alleles (when compared with natural populations) in these heterozygous genotypes.

The reduced genetic diversity was also associated with a reduced spatial genetic structure of the Chilean populations, when compared with New Zealand where all pairwise  $F_{ST}$  values were statistically significant and up to 7 genetic clusters were detected. Some level of pairwise differentiation was observed only within the Intermediate Area, although not among all populations. One hypothesis was that farming practices, often promoting the exchange of raw material among farmers for seeding clonal fragments, could have homogenized the distribution of genetic diversity within Chile. In this case, we would expect strong genetic homogeneity among farms, but not among natural (fixed) populations, because farmed individuals are asexual and produce only few spores, therefore limiting gene flow from farms to natural populations. Such pattern was not detected as natural populations were also poorly differentiated. Moreover, no systematic differentiation was observed between cultivated and natural populations, suggesting that farming practices did not promote a clear genetic divergence between farmed material and their wild relatives. It is therefore unclear what process led to the current distribution of Chilean genetic diversity.

Our results show that the reduced genetic diversity of *G. chilensis* in Chile results from the combined effects of the founder effect when colonizing South America, the over-harvesting during the 1970s and 1980s, and eventually the first steps of the domestication process after the establishment of farming operations. In this context, we wonder about the status and future of *G. chilensis* domestication. As suggested earlier (Guillemin et al. 2014b), these processes may be driving the species to an extinction vortex in Chile, instead of promoting its sustainable domestication. More specific assessment of neutral versus selective processes under partial and fully clonal reproduction are required to better understand the drivers of current genetic diversity in *Gracilaria* farms, and to enable more effective management of the genetic resources. The introduction of new genetic variants in farms may be necessary to stimulate new strain selection for productive gain or pathogen avoidance. A tempting solution would be selecting these new strains from the region of origin, New Zealand, where most of the genetic diversity is present. Such an approach is being actively promoted for genetically depauperate elite cultivar populations in seaweed agronomy (i.e., Zhang et al. 2017). However, it is imperative to safeguard the eventual adaptations that the species may have evolved in the Chilean coastal ecosystems, including farming habitats, and prevent any negative impact on the quality of the Chilean genetic stock. After all, *Gracilaria chilensis* currently supports an algal industry that positions Chile as one of the principal agarophyte producers worldwide.

## ACKNOWLEDGMENTS

This research was supported by ANID CONICYT (Fondo Nacional de Desarrollo Científico y Tecnológico FONDECYT, Chile) under grant number 1090360 and 1170541 and the project IDEALG (ANR-10-BTBR-04, France). OH acknowledge support from



BECA DE DOCTORADO NACIONAL Grant#21120791 (ANID, Chile), SQ-C from FONDECYT 2021 Postdoctoral grant #3210788 (ANID, Chile), PS-A from FONDECYT 1190710 and Millenium Nucleus NUTME NCN19\_056 (ANID, Chile), and WAN from NIWA SSIF funding (Coasts and Oceans Research Programme 2, Marine Biological Resources, New Zealand). France / Chile cooperation was supported by the International Research Network (IRN) DABMA «Diversity and Biotechnology of Marine Algae» IRN00022. We thank Jessica Beltrán (UC Santiago) for her help in the lab and Joe O'Callaghan (NIWA) for advice on oceanography.

#### DATA ACCESSIBILITY

Data deposited in the Dryad repository: <https://doi.org/10.5061/dryad.v41ns1rzw>

#### BENEFIT-SHARING

No benefits to reports.

#### REFERENCES

- Ainis, A. F., Erlandson, J. M., Gill, K. M., Graham, M. H., & Vellanoweth, R. L. (2019). The potential use of seaweeds and marine plants by native peoples of Alta and Baja California. In Gill, K. M. (Ed.), *An archaeology of abundance: Reevaluating the marginality of California's islands* (pp. 135-170). Gainesville: University Press of Florida.
- Alloway, B. V., Lowe, D. J., Barrell, D. J., Newnham, R. M., Almond, P. C., Augustinus, P. C., ... & NZ-INTIMATE members. (2007). Towards a climate event stratigraphy for New Zealand over the past 30 000 years (NZ-INTIMATE project). *Journal of Quaternary Science*, 22(1), 9-35.
- Apte, S., & Gardner, J. P. A. (2002). Population genetic subdivision in the New Zealand greenshell mussel (*Perna canaliculus*) inferred from single-strand conformation polymorphism analysis of mitochondrial DNA. *Molecular Ecology*, 11(9), 1617-1628.
- Avila, M., & Seguel, M. (1993). An overview of seaweed resources in Chile. *Journal of Applied Phycology*, 5(2), 133-139.
- Ayers, K. L., & Waters, J. M. (2005). Marine biogeographic disjunction in central New Zealand. *Marine Biology*, 147(4), 1045-1052.
- Barnes, E. J. (1985). Eastern Cook Strait region circulation inferred from satellite-derived, sea-surface, temperature data. *New Zealand Journal of Marine and Freshwater Research*, 19(3), 405-411.
- Bashir, T., Sailer, C., Gerber, F., Loganathan, N., Bhoopalan, H., Eichenberger, C., ... & Baskar, R. (2014). Hybridization alters spontaneous mutation rates in a parent-of-origin-dependent fashion in *Arabidopsis*. *Plant Physiology*, 165(1), 424-437.
- Beers, L., Rowland, L. J., & Drummond, F. (2019). Genetic diversity of lowbush blueberry throughout the United States in managed and non-managed populations. *Agriculture*,

9(6), 113.

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289-300.
- Boedeker, C., Ramírez, M. E., & Nelson, W. A. (2010). *Cladophoropsis brachyartra* from southern South America is a synonym of *Wittrockiella lyallii* (Cladophorophyceae, Chlorophyta), previously regarded as endemic to New Zealand. *Phycologia*, 49(6), 525-536.
- Boo, G. H., Mansilla, A., Nelson, W., Bellgrove, A., & Boo, S. M. (2014). Genetic connectivity between trans-oceanic populations of *Capreolia implexa* (Gelidiales, Rhodophyta) in cool temperate waters of Australasia and Chile. *Aquatic Botany*, 119, 73-79.
- Brakel, J., Sibonga, R. C., Dumilag, R. V., Montalescot, V., Campbell, I., Cottier-Cook, E. J., ... & Gachon, C. M. M. (2021). Exploring, harnessing and conserving marine genetic resources towards a sustainable seaweed aquaculture. *Plants, People, Planet*, 3, 337-349.
- Buchanan, J., & Zuccarello, G. C. (2012). Decoupling of short- and long-distance dispersal pathways in the endemic New Zealand seaweed *Carpophyllum maschalocarpum* (Phaeophyceae, Fucales). *Journal of Phycology*, 48(3), 518-529.
- Burki, F. (2014). The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harbor Perspective Biology*, 6, a016147.
- Buschmann, A. H., Camus, C., Infante, J., Neori, A., Israel, Á., Hernández-González, M. C., ... & Critchley, A. T. (2017). Seaweed production: Overview of the global state of exploitation, farming and emerging research activity. *European Journal of Phycology*, 52(4), 391-406.
- Buschmann, A. H., Correa, J. A., Westermeier, R., Hernández-González, M. D. C. & Norambuena, R. (2001). Red algal farming in Chile: A review. *Aquaculture*, 194(3-4), 203-220.
- Buschmann, A. H., Hernández-González, M. C. & Varela, D. (2008). Seaweed future cultivation in Chile: Perspectives and challenges. *International Journal of Environment and Pollution*, 33(4), 432-456.
- Buschmann, A. H., Westermeier, R., & Retamales, C. A. (1995). Cultivation of *Gracilaria* on the sea-bottom in southern Chile: A review. *Journal of Applied Phycology*, 7(3), 291-301.
- Camus, P. A. (2001). Marine biogeography of continental Chile. *Revista Chilena de Historia Natural*, 74, 587-617.
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Gene/Genomes/Genetics*, 1(3), 171-182.
- Chandrasekaran, S., Nagendran, N. A., Pandiaraja, D., Krishnankutty, N., & Kamalakannan, B. (2008). Bioinvasion of *Kappaphycus alvarezii* on corals in the Gulf of Mannar, India. *Current Science*, 1167-1172.
- Chiswell, S. M. (2001). Eddy energetics in the subtropical front over the Chatham Rise, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 35(1), 1-15.
- Chiswell, S.M. (2005). Mean and variability in the Wairarapa and Hikurangi eddies. *New Zealand Journal of Marine and Freshwater Research*, 39, 121e134.

- Chiswell, S. M., Bostock, H. C., Sutton, P. J., & Williams, M. J. (2015). Physical oceanography of the deep seas around New Zealand: A review. *New Zealand Journal of Marine and Freshwater Research*, 49(2), 286-317.
- Chiswell, S. M., & Rickard, G. J. (2011). Larval connectivity of harbours via ocean currents: A New Zealand study. *Continental Shelf Research*, 31(10), 1057-1074.
- Cohen, S., Faugeron, S., Martínez, E. A., Correa, J. A., Viard, F., Destombe, C. & Valero, M. (2004). Molecular identification of two sibling species under the name *Gracilaria chilensis* (Rhodophyta, Gracilariales). *Journal of Phycology*, 40(4), 742-747.
- Collin, F. D., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R., Marin, J. -M., & Estoup, A. (2021). Extending Approximate Bayesian Computation with Supervised Machine Learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Molecular Ecology Resources*, 21, 2598-2613.
- Collins, C., Fraser, C., Ashcroft, A., & Waters, J. (2010). Asymmetric dispersal of southern bull-kelp (*Durvillaea antarctica*) adults in coastal New Zealand: Testing an oceanographic hypothesis. *Molecular Ecology*, 19(20), 4572-4580
- Conklin, E. J., & Smith, J. E. (2005). Abundance and spread of the invasive red algae, *Kappaphycus* spp., in Kane'ohe Bay, Hawai'i and an experimental assessment of management options. *Biological Invasions*, 7(6), 1029-1039.
- Cornille, A., Antolín, F., Garcia, E., Vernesi, C., Fietta, A., Brinkkemper, O., ... & Roldán-Ruiz, I. (2019). A multifaceted overview of apple tree domestication. *Trends in Plant Science*, 24(8), 770-782.
- Cottier-Cook, E., Nagabhatla, N., Asri, A., Beveridge, M., Bianchi, P., Bolton, J., ... & Yarish, C. (2021). Ensuring the sustainable future of the rapidly expanding global seaweed aquaculture industry-A vision.
- Craw, R. (1988). Continuing the synthesis between panbiogeography, phylogenetic systematics and geology as illustrated by empirical studies on the biogeography of New Zealand and the Chatham Islands. *Systematic Biology*, 37(3), 291-310.
- Cumming, R. A., Nikula, R., Spencer, H. G., & Waters, J. M. (2014). Transoceanic genetic similarities of kelp-associated sea slug populations: Long-distance dispersal via rafting? *Journal of Biogeography*, 41(12), 2357-2370.
- Davies, B. J., Darvill, C. M., Lovell, H., Bendle, J. M., Dowdeswell, J. A., Fabel, D., ... & Thorndycraft, V. R. (2020). The evolution of the Patagonian Ice Sheet from 35 ka to the present day (PATICE). *Earth-Science Reviews*, 204, 103152.
- Dillehay, T. D., Ramírez, C., Pino, M., Collins, M. B., & Rossen, J. (2008). Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science*, 320(5877), 784-786.
- Doebly, J. F., Gaut, B. S. & Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell*, 127(7), 1309-1321.
- Duarte, C. M., Marba, N. & Holmer, M. (2007). Rapid domestication of marine species. *Science*, 316(5823), 382-3.
- Ellstrand, N. C., & Roose, M. L. (1987). Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, 74(1), 123-131.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564-567.
- Faugeron, S., Martínez, E. A., Correa, J. A., Cárdenas, L., Destombe, C. & Valero, M. (2004). Reduced genetic diversity and increased population differentiation in

- peripheral and overharvested populations of *Gigartina skottsbergii* (Rhodophyta, Gigartinales) in Southern Chile. *Journal of Phycology*, 40(3), 454-462.
- Feng, X., Zhou H., Zulfiqar S., Luo X., Hu Y., Feng L., Malvolti, M. E., Woeste, K., & Zhao P. (2018). The Phylogeographic history of common walnut in China. *Frontiers in Plant Science*, 9, 1399.
- Fleming, C. A. (1975). The geological history of New Zealand and its biota. In G. Kuschel (Ed.), *Biogeography and ecology in New Zealand* (pp. 1-86). Springer, Dordrecht.
- Foighil, D. O., Marshall, B. A., Hilbish, T. J., & Pino, M. A. (1999). Trans-Pacific range extension by rafting is inferred for the flat oyster *Ostrea chilensis*. *The Biological Bulletin*, 196(2), 122-126.
- Fraser, C. I., Hay, C. H., Spencer, H. G. & Waters, J. M. (2009). Genetic and morphological analyses of the southern bull kelp *Durvillaea antarctica* (Phaeophyceae: Durvillaeales) in New Zealand reveal cryptic species. *Journal of Phycology*, 45(2), 436-443.
- Fraser, C. I., Morrison, A. K., Hogg, A. M., Macaya, E. C., van Seville, E., Ryan, P. G., Padovan, A., Jack, C., Valdivia, N., & Waters, J. M. (2018). Antarctica's ecological isolation will be broken by storm-driven dispersal and warming. *Nature Climate Change*, 8(8), 704-708.
- Fraser, C., Winter, D., Spencer, H., & Waters, J. (2010). Multigene phylogeny of the southern bull-kelp genus *Durvillaea* (Phaeophyceae: Fucales). *Molecular Phylogenetics and Evolution*, 57(3), 1301-1311.
- Fraser, C. I., Zuccarello, G. C., Spencer, H. G., Salvatore, L. C., Garcia, G. R. & Waters, J. M. (2013). Genetic affinities between trans-oceanic populations of non-buoyant macroalgae in the high latitudes of the Southern Hemisphere. *PLoS One*, 8(7), e69138.
- Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925-929.
- Gillespie, R. G., Baldwin, B. G., Waters, J. M., Fraser, C. I., Nikula, R., & Roderick, G. K. (2012). Long-distance dispersal: A framework for hypothesis testing. *Trends in Ecology & Evolution*, 27(1), 47-56.
- Goecke, F., Klemetsdal, G., & Ergon, Å. (2020) Cultivar development of kelps for commercial cultivation-Past lessons and future prospects. *Frontiers in Marine Science*, 7, 110.
- Goldstien, S. J., Schiel, D. R., & Gemmill, N. J. (2006). Comparative phylogeography of coastal limpets across a marine disjunction in New Zealand. *Molecular Ecology*, 15(11), 3259-3268.
- Graf, L., Shin, Y., Yang, J. H., Choi, J. W., Hwang, I. K., Nelson, W., Bhattacharya, D., Viard, F., & Hwang, H. S. (2021). A genome-wide investigation of the effect of farming and human-mediated introduction on the ubiquitous seaweed *Undaria pinnatifida*. *Nature Ecology & Evolution*, 5(3), 360-368.
- Guillemin, M. -L., Faugeron, S., Destombe, C., Viard, F., Correa, J. A. & Valero, M. (2008). Genetic variation in wild and cultivated populations of the haploid-diploid red alga *Gracilaria chilensis*: How farming practices favor asexual reproduction and heterozygosity. *Evolution*, 62(6), 1500-1519.
- Guillemin, M. -L., Huanel, O. R., & Martínez, E. A. (2012). Characterization of genetic markers linked to sex determination in the haploid-diploid red alga *Gracilaria chilensis*. *Journal of Phycology*, 48(2), 365-372.
- Guillemin, M. -L., Valenzuela, P., Gaitán-Espitia, J. D., & Destombe, C. (2014a). Evidence

- of reproductive cost in the triphasic life history of the red alga *Gracilaria chilensis* (Gracilariales, Rhodophyta). *Journal of Applied Phycology*, 26(1), 569-575.
- Guillemin, M. -L., Valero, M., Faugeron, S., Nelson, W. & Destombe, C. (2014b). Tracing the trans-Pacific evolutionary history of a domesticated seaweed (*Gracilaria chilensis*) with archaeological and genetic data. *PLoS One*, 9(12), e114039.
- Guzinski, J., Ballenghien, M., Daguin-Thiébaud, C., Lévêque, L., & Viard, F. (2018). Population genomics of the introduced and cultivated Pacific kelp *Undaria pinnatifida*: Marinas—not farms—drive regional connectivity and establishment in natural rocky reefs. *Evolutionary Applications*, 11(9), 1582-1597.
- Gutenkunst, R. N., Hernandez, R.D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetic*, 5(10), e1000695.
- Hafting, J. T., Craigie, J. S., Stengel, D. B., Loureiro, R. R., Buschmann, A. H., Yarish, C., Edwards, M. D., & Critchley, A. T. (2015). Prospects and challenges for industrial production of seaweed bioactives. *Journal of Phycology*, 51(5), 821-837.
- Halling, C., Wikström, S. A., Lilliesköld-Sjöo, G., Mörk, E., Lundsør, E., & Zuccarello, G. C. (2013). Introduction of Asian strains and low genetic variation in farmed seaweeds: Indications for new management practices. *Journal of Applied Phycology*, 25(1), 89-95.
- Hayashi, L., Reis, R. P., dos Santos, A. A., Castelar, B., Robledo, D., Batista de Vega, G. B., Msuya, F. E., Eswaran, K., Yasir, S. M., Ali, M. K. M. & Hurtado, A. Q. (2017). The cultivation of *Kappaphycus* and *Euclidean* in Tropical and sub-Tropical waters. In Hurtado, A. Q., Critchley, A. T., & Neish, I. C. (Eds.). *Tropical Seaweed Farming Trends, Problems and Opportunities* (pp. 55-90). Springer, Cham.
- Huanel, O. R., Nelson, W. A., Robitzsch, V., Mauger, S., Faugeron, S., Preuss, M., Zuccarello, G. C., & Guillemin, M. -L. (2020). Comparative phylogeography of two *Agarophyton* species in the New Zealand archipelago. *Journal of Phycology*, 56(6), 1575-1590.
- Hudson, R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, 18: 337-338.
- Hufford, M. B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J. M., Cartwright, R. A., ... & Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44(7), 808-811.
- Hurtado, A. Q., Neish, I. C., & Critchley, A. T. (2019) Phyconomy: The extensive cultivation of seaweeds, their sustainability and economic value, with particular reference to important lessons to be learned and transferred from the practice of euclidean farming. *Phycologia*, 58(5), 472-483.
- Hwang, E. K., Yotsukura, N., Pang, S. J., Su, L. & Shan, T. F. (2019). Seaweed breeding programs and progress in eastern Asian countries. *Phycologia*, 58(5), 484-495.
- Jones, T. C., Gemmill, C. E., & Pilditch, C. A. (2008). Genetic variability of New Zealand seagrass (*Zostera muelleri*) assessed at multiple spatial scales. *Aquatic Botany*, 88(1), 39-46.
- Kaiser, J., Lamy, F., & Hebbeln, D. (2005). A 70-kyr sea surface temperature record off southern Chile (Ocean Drilling Program Site 1233). *Paleoceanography*, 20(4).
- Klopper, T. H., & Huson, D. H. (2008). Drawing explicit phylogenetic networks and their integration into SplitsTree. *BMC Evolutionary Biology*, 8(1), 1-7.
- Kohfeld, K. E., Graham, R. M., De Boer, A. M., Sime, L. C., Wolff, E. W., Le Quéré, C.,

- & Bopp, L. (2013). Southern Hemisphere westerly wind changes during the Last Glacial Maximum: Paleo-data synthesis. *Quaternary Science Reviews*, 68, 76-95.
- Li, Y., Colleoni, C., Zhang, J., Liang, Q., Hu, Y., Ruess, H., ... & Huang, B. (2018). Genomic analyses yield markers for identifying agronomically important genes in potato. *Molecular Plant*, 11(3), 473-484.
- Lischer, H. E., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298-299.
- Liu, Q., Zhou, Y., Morrell, P. L., & Gaut, B. S. (2017). Deleterious variants in Asian rice and the potential cost of domestication. *Molecular Biology and Evolution*, 34(4), 908-924.
- Lorrey, A. M., & Bostock, H. (2017). The climate of New Zealand through the Quaternary. In Landscape and quaternary environmental change in New Zealand. In Shulmeister J. (Ed.) *Landscape and Quaternary Environmental Change in New Zealand*, Atlantis Advances in Quaternary Science 3 (pp. 67-139). Paris: Atlantis Press.
- Lorrey, A. M., Vandergoes, M., Almond, P., Renwick, J., Stephens, T., Bostock, H., ... & Fowler, A. M. (2012). Palaeocirculation across New Zealand during the Last Glacial Maximum at ~ 21 ka. *Quaternary Science Reviews*, 36, 189-213.
- Macaya, E. C. & Zuccarello, G. C. (2010). Genetic structure of the giant kelp *Macrocystis pyrifera* along the southeastern pacific. *Marine Ecology Progress Series*, 420, 103-112.
- Main, M. R. (1989). Distribution and post-glacial dispersal of freshwater fishes in South Westland, New Zealand, *Journal of the Royal Society of New Zealand*, 19(2), 161-169.
- Marshall, D. C., Hill, K. B., Fontaine, K. M., Buckley, T. R., & Simon, C. (2009). Glacial refugia in a maritime temperate climate: Cicada (*Kikihia subalpina*) mtDNA phylogeography in New Zealand. *Molecular Ecology*, 18(9), 1995-2009.
- McKey, D., Elias, M., Pujol, B., & Duputié, A. (2010). The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, 186(2), 318-332.
- McLachlan, J., & Bird, C. J. (1984). Geographical and experimental assessment of the distribution of *Gracilaria* species (Rhodophyta: Gigartinales) in relation to temperature. *Helgoländer Meeresuntersuchungen*, 38(3-4), 319-334.
- Meirmans, P. G. (2020). Genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, 20(4), 1126-1131.
- Molina, J., Sikora, M., Garud, N., Flowers, J. M., Rubinstein, S., Reynolds, A., ... & Purugganan, M. D. (2011). Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences*, 108(20), 8351-8356.
- Muangmai, N., Fraser, C. I., & Zuccarello, G. C. (2015). Contrasting patterns of population structure and demographic history in cryptic species of *Bostrychia intricata* (Rhodomelaceae, Rhodophyta) from New Zealand. *Journal of Phycology*, 51(3), 574-585.
- Muangmai, M., von Ammon, U., & Zuccarello, G. C. (2016). Cryptic species in sympatry: Nonrandom small-scale distribution patterns in *Bostrychia intricata* (Ceramiaceae, Rhodophyta), *Phycologia*, 55(4), 424-430.
- Muangmai, N., West, J. A., & Zuccarello, G. C. (2014). Evolution of four Southern

Hemisphere *Bostrychia* (Rhodomelaceae, Rhodophyta) species: Phylogeny, species delimitation and divergence times. *Phycologia*, 53(6), 593-601.

- Muñoz, J., Felicísimo, A. M., Cabezas, F., Burgaz, A. R., & Martínez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, 304(5674), 1144-1147.
- Nathan, R. (2006). Long-distance dispersal of plants. *Science*, 313(5788), 786-788.
- Norambuena, R. (1996). Recent trends of seaweed production in Chile. *Hydrobiologia*, 326/327, 371-379.
- Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, 8(10), 1360-1373.
- Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288-295.
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135.
- Pollegioni, P., Woeste, K., Chiocchini, F., Del Lungo, S., Ciolfi, M., Olimpieri, I., ... & Malvolti, M. E. (2017). Rethinking the history of common walnut (*Juglans regia* L.) in Europe: Its origins and human interactions. *PLoS One*, 12(3), e0172541.
- Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., Rodel, M. -O., Blackburn, D. C., & Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian Tropical Forest frog using demographic model selection. *Molecular Ecology*, 26(19), 5245-5263.
- Preuss, M., Muangmai, N., Nelson, W. A., Guillemin, M. L., West, J. A., & Zuccarello, G. C. (2020). *Agarophyton transtasmanicum* sp. nov. from Australia and New Zealand. *Phycologia*, 59(3), 238-245.
- Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859-866.
- R Core Team. (2019). R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.Rproject.org/>
- Rafajlović, M., Kleinhan, D., Gulliksson, C., Fries, J., Johansson, D., Ardehed, A., Sundqvist, L., Pereyra, R. T., Mehlig, B., Jonsson, P. R., & Johannesson, K. (2017). Neutral processes forming large clones during colonization of new areas. *Journal of Evolutionary Biology*, 30(8), 1544-1560.
- Ramu, P., Esuma, W., Kawuki, R., Rabbi, I. Y., Egesi, C., Bredeson, J. V., Bart, R. S., Verma, J., Buckler, E. S. & Lu, F. (2017). Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nature Genetics*, 49(6), 959-963.
- Raynal, L., Marin, J. M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2019). ABC random forests for Bayesian parameter inference. *Bioinformatics*, 35(10), 1720-1728.
- Robitcz, V., Arakaki, N., Mauger, S., Rojas, J. C. Z., & Guillemin, M. -L. (2019). Stranded alone: The first reported Peruvian population of *Agarophyton chilensis* is a single-male's clone. *Algal Research*, 41, 101527.
- Ross, P. M., Hogg, I. D., Pilditch, C. A. & Lundquist, C. J. (2009). Phylogeography of New Zealand's coastal benthos. *New Zealand Journal of Marine and Freshwater Research*, 43(5), 1009-1027.
- Ross, P. M., Hogg, I. D., Pilditch, C. A., Lundquist, C. J., & Wilkins, R. J. (2012). Population genetic structure of the New Zealand estuarine clam *Austrovenus*

- stutchburyi* (Bivalvia: Veneridae) reveals population subdivision and partial congruence with biogeographic boundaries. *Estuaries and Coasts*, 35(1), 143-154.
- Santelices, B. & Doty, M. S. (1989). A review of *Gracilaria* farming. *Aquaculture*, 78(2), 95-133.
- Scaladaferro, M. A., Barboza, G. E., & Acosta, M. C. (2018). Evolutionary history of the chili pepper *Capsicum baccatum* L. (Solanaceae): Domestication in South America and natural diversification in the Seasonally Dry Tropical Forests. *Biological Journal of the Linnean Society*, 124(3), 466-478.
- Schnabel, K. E., Hogg, I. D., & Chapman, M. A. (2000). Population genetic structures of two New Zealand corophiid amphipods and the presence of morphologically cryptic species: Implications for the conservation of diversity. *New Zealand Journal of Marine and Freshwater Research*, 34(4), 637-644.
- Stevens, C. L., O'Callaghan, J. M., Chiswell, S. M., & Hadfield, M. G. (2021). Physical oceanography of New Zealand/Aotearoa shelf seas-a review. *New Zealand Journal of Marine and Freshwater Research*, 55(1), 6-45.
- Shears, N. T., Smith, F., Babcock, R. C., Duffy, C. A., & Villouta, E. (2008). Evaluation of biogeographic classification schemes for conservation planning: Application to New Zealand's coastal marine environment. *Conservation Biology*, 22(2), 467-481.
- Silva, C. N., & Gardner, J. P. (2015). Emerging patterns of genetic variation in the New Zealand endemic scallop *Pecten novaezelandiae*. *Molecular Ecology*, 24(21), 5379-5393.
- Smith, T. M., York, P. H., Broitman, B. R., Thiel, M., Hays, G. C., van Sebille, E., Putman, N. F., Macreadie, P. I. & Sherman, C. D. H. (2018) Rare long-distance dispersal of a marine angiosperm across the Pacific Ocean. *Global Ecology and Biogeography*, 27(4), 487-496.
- Stevens, M. I., & Hogg, I. D. (2004). Population genetic structure of New Zealand's endemic corophiid amphipods: Evidence for allopatric speciation. *Biological Journal of the Linnean Society*, 81(1), 119-133.
- Stoeckel, S., Arnaud-Haond, S., & Krueger-Hadfield, S. A. (2021). The combined effect of haplodiplonty and partial clonality on genotypic and genetic diversity in a finite mutating population. *Journal of Heredity*, 112(1), 78-91.
- Sun, X., Jiao, C., Schwaninger, H., Chao, C. T., Ma, Y., Duan, N., Khan, A., Ban, S., Xu, K., Cheng, L., Zhong, G. -Y., & Fei, Z. (2020). Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nature Genetics*, 52(12), 1423-1432.
- Thomas, E., van Zonneveld, M., Loo, J., Hodgkin, T., Galluzzi, G., & van Etten, J. (2012). Present spatial diversity patterns of *Theobroma cacao* L. in the Neotropics reflect genetic differentiation in Pleistocene refugia followed by human-influenced dispersal. *PLoS One*, 7, e47676.
- Tomczak, M., & Godfrey, J. S. (1994). *Regional Oceanography: An Introduction*. Pergamon, New York.
- Trewick, S. A., & Bland, K. J. (2012) Fire and slice: Palaeogeography for biogeography at New Zealand's North Island/South Island juncture. *Journal of the Royal Society of New Zealand*, 42(3), 153-183.
- Trewick, S. A., & Wallis, G. P. (2001). Bridging the "beech-gap": New Zealand invertebrate phylogeography implicates Pleistocene glaciation and Pliocene isolation. *Evolution*, 55(11), 2170-2180.



- Usandizaga, S., Camus, C., Kappes, J. L., Guillemin, M. -L., & Buschmann, A. H. (2019). Effect of temperature variation in *Agarophyton chilensis*: Contrasting the response of natural and farmed populations. *Journal of Applied Phycology*, 31(4), 2709-2717.
- Vairappan, C. S., Chung, C. S., Hurtado, A. Q., Soya, F. E., Lhonneur, G. B., & Critchley, A. (2008). Distribution and symptoms of epiphyte infection in major carrageenophyte-producing farms. *Journal of Applied Phycology*, 20(5), 477-483.
- Valderrama, D., Cai, J., Hishamunda, N., Ridler, N., Neish, I. C., Hurtado, A. Q., Msuya, F. E., Krishnan, M., Narayanakumar, R., Kronen, M., Robledo, D., Gasca-Leyva, E., & Fraga, J. (2015). The economics of *Kappaphycus* seaweed cultivation in developing countries: A comparative analysis of farming systems. *Aquaculture Economics & Management*, 19(2), 251-277.
- Valero, M., Guillemin, M. -L., Destombe, C., Jacquemin, B., Gachon, C. M. M., Badis, Y., Badis, Y., Buschmann, A. H., Camus, C., & Faugeton, S. (2017). Perspectives on domestication research for sustainable seaweed aquaculture. *Perspectives in Phycology*, 4(1), 33-46.
- Veale, A. J., & Lavery, S. D. (2011). Phylogeography of the snakeskin chiton *Sypharochiton pelleris* (Mollusca: Polyplacophora) around New Zealand: Are seasonal near-shore upwelling events a dynamic barrier to gene flow? *Biological Journal of the Linnean Society*, 104(3), 552-563.
- Veale, A. J., & Lavery, S. D. (2012) The population genetic structure of the waratah anemone (*Actinia tenebrosa*) around New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 46(4), 523-536.
- Vieira, V. M., Engelen, A. H., Huanel, O. R., & Guillemin, M. -L. (2021). Differential frond growth in the isomorphic haploid-diploid red seaweed *Agarophyton chilense* by long-term in situ monitoring. *Journal of Phycology*, 57(2), 592-605.
- Wang, X., Yao, J., Zhang, J., & Duan, D. (2020). Status of genetic studies and breeding of *Saccharina japonica* in China. *Journal of Oceanology and Limnology*, 38(4), 1064-1079.
- Wardle, P. (1963). Evolution and distribution of the New Zealand flora, as affected by quaternary climates, *New Zealand Journal of Botany*, 1(1), 3-17.
- Yu, Y., Fu, J., Xu, Y., Zhang, J., Ren, F., Zhao, H., ... & Xie, H. (2018). Genome re-sequencing reveals the evolutionary history of peach fruit edibility. *Nature Communications*, 9(1), 1-13.
- Zeder, M. A. (2015). Core questions in domestication research. *Proceedings of the National Academy of Sciences USA*, 112(11), 3191-3198.
- Zhang, J., Wang, X., Yao, J., Li, Q., Liu, F., Yotsukura, N., Krupnova, T. N., & Duan, D. (2017). Effect of domestication on the genetic diversity and structure of *Saccharina japonica* populations in China. *Scientific Reports*, 7(1), 1-11.
- Zhang, J., Yao, J., Hu, Z. M., Jueterbock, A., Yotsukura, N., Krupnova, T. N., Nagasato, C., & Duan, D. (2019). Phylogeographic diversification and postglacial range dynamics shed light on the conservation of the kelp *Saccharina japonica*. *Evolutionary Applications*, 12(4), 791-803.
- Zhang, J., Yao, J. T., Sun, Z. M., Fu, G., Galanin, D. A., Nagasato, C., Motomura, T., Hu, Z. -M., & Duan, D. L. (2015). Phylogeographic data revealed shallow genetic structure in the kelp *Saccharina japonica* (Laminariales, Phaeophyta). *BMC Evolutionary Biology*, 15(1), 1-12.
- Zhou, Y., Massonnet, M., Sanjak, J. S., Cantu, D., & Gaut, B. S. (2017). Evolutionary

genomics of grape (*Vitis vinifera* ssp. *vinifera*) domestication. *Proceedings of the National Academy of Sciences USA*, 114(44), 11715-11720.

Zohary, D., Hopf, M., & Weiss., E. (2012). *Domestication of Plants in the Old World*. Oxford University Press, Oxford.

Zuccarello, G. C., Critchley, A. T., Smith, J., Sieber, V., Lhonneur, G. B., & West, J. A. (2006). Systematics and genetic variation in commercial shape *Kappaphycus* and shape *Eucheuma* (Solieriaceae, Rhodophyta). *Journal of Applied Phycology*, 18(3), 643-651.

Zuccarello, G. C. & Martin, P. (2016). Phylogeography of the *Lessonia variegata* species complex (Phaeophyceae, Laminariales) in New Zealand. *Algae*, 31(2),91-103.

Tables and Figures

Table 1: Summary of genetic diversity obtained in each bioregions included in the study using 2,232 SNPs. Bioregions following the New Zealand Coastal Biogeographic Regions Classification map (available at <https://www.doc.govt.nz/documents/conservation/marine-and-coastal/marine-protected-areas/coastal-marine-habitats-marine-protected-areas.pdf>, the Peru-Chile biogeographic classification of Camus (2001). For details about genetic diversity estimated in each locality sampled in the South Pacific Ocean, see Supplementary Table 1.

Country	Bioregion (code)	N	Na <sup>a</sup>	Ne <sup>a</sup>	%P	He <sup>a</sup>	D	Apriv
<b>New Zealand</b>		<b>t=439</b>	<b>1.12</b> ( <b>&lt;0.01</b> )	<b>1.05</b> ( <b>0.04</b> )	<b>13.17</b>	<b>0.03</b> ( <b>0.02</b> )	<b>0.94</b>	<b>97</b>
	West Coast North Island (WNI)	t=43	1.20 (0.05)	1.07 ( <b>&lt;0.01</b> )	30.51	0.05 ( <b>&lt;0.01</b> )	1	24
	North Eastern (NE)	t=47	1.27 (0.03)	1.13 (0.01)	35.89	0.08 ( <b>&lt;0.01</b> )	1	24
	East Coast North Island (ENI)	t=20	1.20 ( <b>&lt;0.01</b> )	1.02 ( <b>&lt;0.01</b> )	22.49	0.03 ( <b>&lt;0.01</b> )	0.95	25
	North Cook Strait (NCS)	t=40	1.26 (0.02)	1.10 ( <b>&lt;0.01</b> )	31.09	0.07 ( <b>&lt;0.01</b> )	0.95	0
	West Coast South Island (WSI)	t=22	1.03 ( <b>&lt;0.01</b> )	1.01 ( <b>&lt;0.01</b> )	2.96	0.01 ( <b>&lt;0.01</b> )	1	7
	South Cook Strait (SCS)	t=63	1.12 (0.05)	1.04 (0.04)	31.59	0.02 (0.02)	1	10
	East Coast South Island (ESI)	t=38	1.06 (0.02)	1.03 (0.01)	7.53	0.02 ( <b>&lt;0.01</b> )	0.71	0
	Southern (SOUT)	t=92	1.05 (0.02)	1.02 (0.01)	10.66	0.02 ( <b>&lt;0.01</b> )	0.90	7
	Fiordland (FIOR)	t=40	1.05 (0.03)	1.02 (0.02)	8.96	0.02 ( <b>&lt;0.01</b> )	0.98	0
	Chatham Islands (CHT)	t=34	1.12 ( <b>&lt;0.01</b> )	1.05 ( <b>&lt;0.01</b> )	7.44	0.03 ( <b>&lt;0.01</b> )	0.95	0
<b>South America</b>		<b>t=248</b>	<b>1.07</b> ( <b>0.03</b> )	<b>1.04</b> ( <b>0.02</b> )	<b>6.60</b>	<b>0.02</b> ( <b>0.01</b> )	<b>0.60</b>	<b>0</b>
	Peruvian bioregion (PER)	t=34	1.05 ( <b>&lt;0.01</b> )	1.03 ( <b>&lt;0.01</b> )	9.05	0.02 ( <b>&lt;0.01</b> )	0.26	0
	Intermediate Area (INT)	t=145	1.06 (0.2)	1.03 (0.02)	10.08	0.02 ( <b>&lt;0.01</b> )	0.68	0
	Magellanic Bioregion (MAG)	t=69	1.08 (0.02)	1.06 ( <b>&lt;0.01</b> )	9.50	0.03 ( <b>&lt;0.01</b> )	0.61	0
<b>Total</b>		<b>t=687</b>	<b>1.23</b> ( <b>&lt;0.01</b> )	<b>1.08</b> ( <b>&lt;0.01</b> )	<b>10.62</b>	<b>0.06</b> ( <b>&lt;0.01</b> )	<b>0.82</b>	<b>97</b>

Na: mean number of alleles, Ne: effective number of alleles, %P: percentage of polymorphic loci, He: expected heterozygosity, *D*: proportion of different genotypes, Apriv: number of private alleles. <sup>a</sup>: Standard deviation noted between brackets



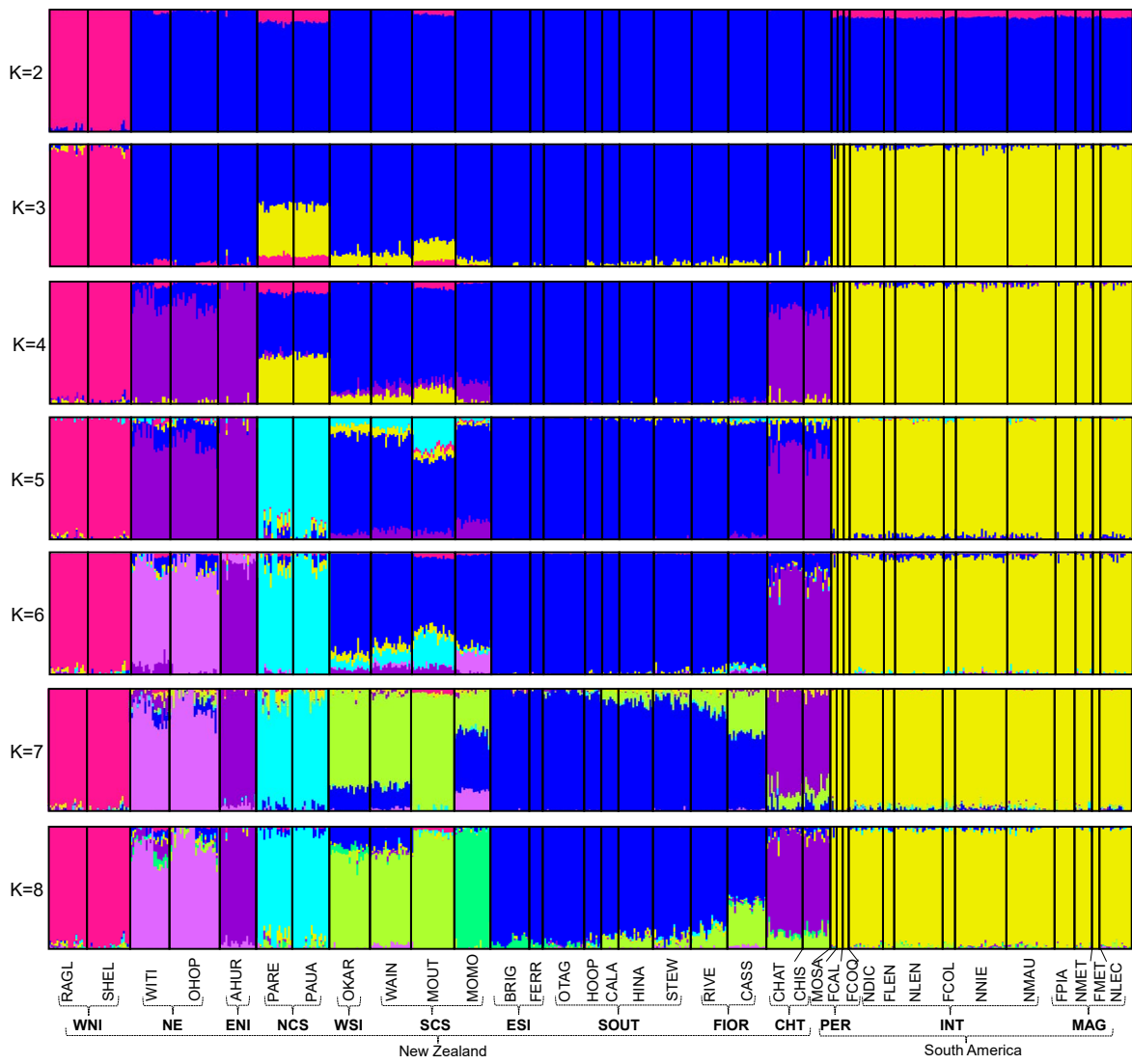


Figure 1: Bart plot representing *Gracilaria chilensis* individual genetic cluster assignment from the R package “LEA” results from K = 2 to K = 8. Clones were not considered in the analysis. Codes for localities and bioregions are as in Table 1 and S1.

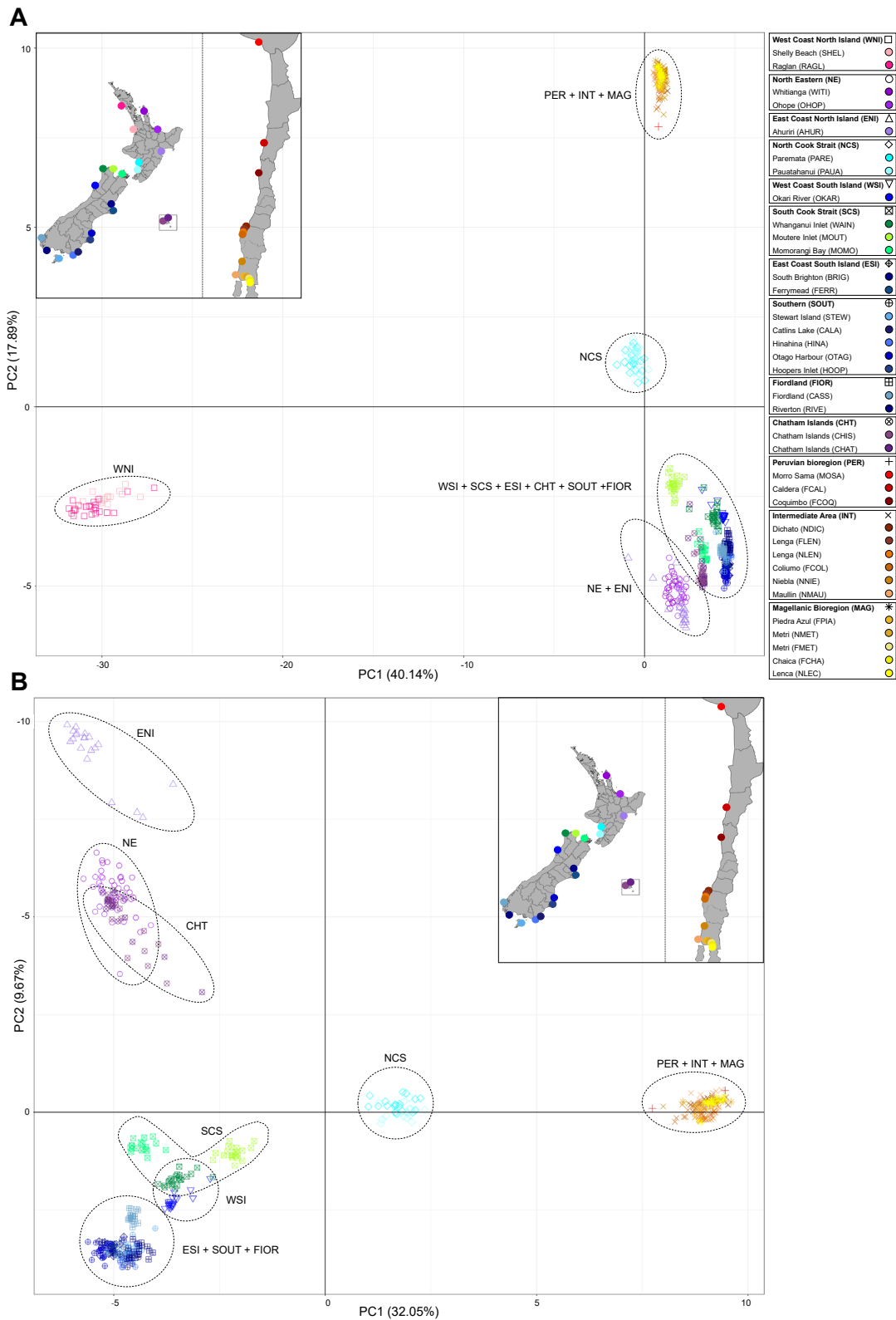


Figure 2: Principal components analysis (PCA) showing the clustering of *Gracilaria chilensis* samples according to the sampled locality and bioregion and implemented (A) with the whole dataset, (B) with all samples except the ones from the localities of SHEL

and RAGL located in the Western North Island (WNI). Whatever the dataset, clones were excluded from the analysis. Codes for localities and bioregions are as in Table 1 and S1.



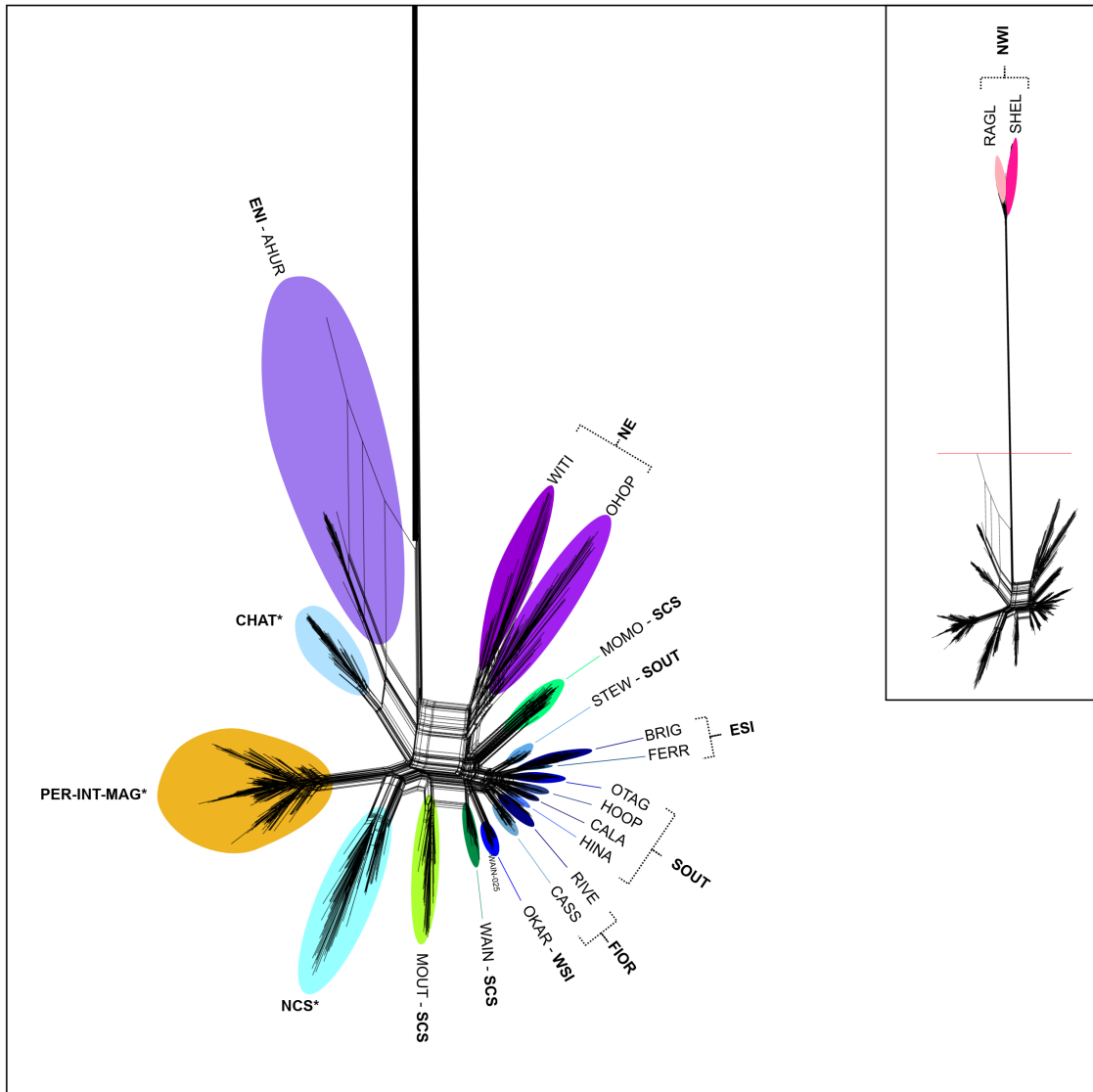


Figure 3: Neighbor-Net network based on uncorrected distances for the 571 *Gracilaria chilensis* samples, clones excluded. Codes for localities (plain text) and bioregions (in bold) are as in Table 1 and S1. Bioregions for which no clear sub-clustering by localities was observed are indicated by an asterisk.

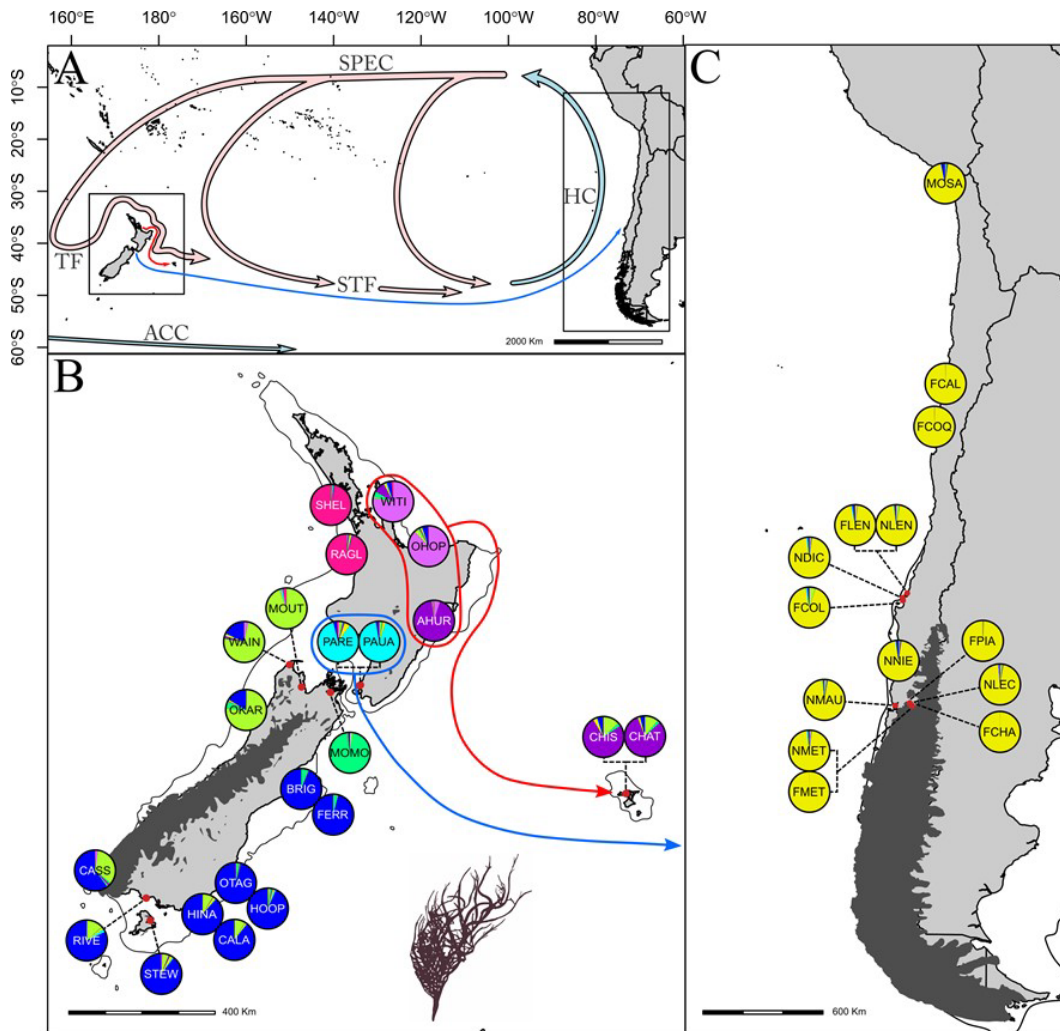


Figure 4: Sampling areas across the current distribution and geographical distribution of the seven genetically distinct groups of *Gracilaria chilensis*. Pie colors correspond to the population assignment results implemented using the R package LEA, assuming  $K = 8$ , and 2,232 SNPs. The black rectangle delimits the two main sampling areas: New Zealand and South America. Codes for localities are as in Table S1. The contour in fine black lines indicates the extent of the exposed shelves and dark grey shading the extent of ice during the Last Glacial Maximum (Alloway et al. 2007, Davies et al. 2020). Prevailing currents are modified from Tomczak and Godfrey (1994) and Chiswell et al. (2015): South Pacific Equatorial Current (SPEC), Tasman Current (TC), Sub Tropical Front (STF) Antarctic Circumpolar Current (ACC), and Humboldt Current (HC). Reconstruction of the most likely colonization routes followed by *G. chilensis* are shown as colored arrows (route of colonization from New Zealand to Chatham Island in red and to Chile in blue).