Characterization of newly developed expressed sequence tag-derived microsatellite markers revealed low genetic diversity within and low connectivity between European Saccharina latissima populations

Jaromir Guzinski, Stéphane Mauger, J. Mark Cock, Myriam Valero

Characterization of newly developed expressed sequence tag-derived microsatellite markers revealed low genetic diversity within and low connectivity between European *Saccharina latissima* populations

Jaromir Guzinski [1,3], Stéphane Mauger[1], J. Mark Cock[2], Myriam Valero[1]

[1] CNRS, UPMC, Sorbonne Universités, PUCCH, UACH, UMI 3614, Evolutionary Biology and Evolution of Algae, Station Biologique de Roscoff, Place Georges Teissier, CS 90074 Roscoff, France

[2] CNRS, UPMC, Sorbonne Universités, UMR8227, Laboratoire Végétaux Marins et Biomolécules, Équipe Génétique des Algues, Station Biologique de Roscoff, Place Georges Teissier, CS 90074 Roscoff, France

[3] CNRS, UPMC, Sorbonne Universités, UMR7144, Adaptation et diversité en Milieu Marin, Équipe Diversité et Connectivité dans le Paysage Marin Côtier, Station Biologique de Roscoff, Place Georges Teissier, CS 90074 Roscoff, France

Abstract

The kelp *Saccharina latissima* is a species of high ecological and economic importance. We developed a novel set of *S. latissima*-specific genetic markers that will find applications in conservation biology, biodiversity assessment, and commercial exploitation of this macroalga. Thirty two expressed sequence tag (EST)-derived microsatellite markers (SSRs) were developed and characterized in this study using publically available EST sequences. Twenty seven percent of the 7,064 analysed ESTs contained repeat motifs, and polymerase chain reaction (PCR) amplification primers were designed for 96 selected loci. Fifty one (53 %) of the primer pairs amplified their target loci, of which 32 (33 %) were polymorphic within a sample of 96 *S. latissima* sporophytes collected from six localities distributed along the European Atlantic coast from Southern Brittany (France) to Spitzbergen (Norway). The 32 loci harboured moderate levels of polymorphism with 2-13 alleles per locus (mean

5.4). The 25 loci that were retained for population genetic analyses revealed substantial genetic differentiation among the European populations (pairwise F_{ST} values ranging from 0.077 to 0.562) that did not follow any pattern of isolation by distance. In addition, within-population genetic diversity was generally low (*Hs* < 0.323). Two non-mutually exclusive hypotheses were proposed to explain this low diversity pattern: (1) lower variability of the EST-derived microsatellites compared to the random distribution of SSRs developed from genomic DNA since the former are frequently located in coding regions, which are generally less variable, or (2) reduced effective population size of *S*. *latissima*. The particularly high genetic differentiation between the French and Scandinavian *S*. *latissima* populations is in agreement with the reported ecotypic differentiation, which may reflect an important resource for genetic improvement. The pattern of genetic diversity revealed in this study thus suggests that care should be taken to avoid the transfer of strains between different geographic regions.

Keywords: *Saccharina latissima,* Phaeophyceae, NE Atlantic, EST, population genetic structure, simple sequence repeat

Introduction

Saccharina latissima (Linnaeus), known by the common names of sea belt or sugar kelp, is a perennial brown macroalga (class Phaeophyceae), of the family Laminariaceae. This species exhibits circumpolar distribution in the northern hemisphere, occurring in polar and temperate coastal waters from intertidal to a lower depth limit of about 30 meters (Bolton et al. 1983; Gerard 1988). *S. latissima* forms continuous dense stands ("kelp forests") on shallow rocky coasts, occurring in clear and turbid waters, preferentially in areas sheltered from strong wave action (Bekby and Moy 2011). In Europe, its range extends from Spitsbergen, Norway as far south as Portugal. *S*. *latissima* requires rocky substrate as an attachment point and does not occur where this surface is replaced by long stretches of sandy coast (Bartsch et al. 2008). Both the wide latitudinal and vertical distribution of

this species suggest the occurrence of ecotypic differentiation between populations with respect to light and temperature (Gerard and Du Bois 1988; Müller et al. 2008). However, knowledge about genetic diversity and connectivity is still lacking and this information will be important not only for elucidating population structure across the range of *S. latissima* but also for improved exploration of genetic resources in the wild.

In general, the dispersal capacities of macroalgae are limited, on the order of few meters up to several kilometres, compared with distances of hundreds of meters to several hundred kilometres for other marine organisms such as invertebrates or fish characterized by a pelagic larval dispersal phase (Kinlan and Gaines 2003). In kelps, dispersal is primarily mediated by short lived propagules, and the extent of dispersal depends on factors such as the morphology of the blade (i.e. location of the sporophylls on the thallus and whether they possess swimming aids such as air bladders, see Valero et al. 2011 for a review), as well as on the orientation of currents and habitat continuity (Billot et al. 2003; Alberto et al. 2010; Coleman et al. 2011). In a recent meta-analysis of genetic differentiation in macroalgae, Durrant et al. (2014) found that most species exhibited a clear pattern of isolation by distance with high level genetic differentiation at distances greater than 50-100 km. However, occasional long distance dispersal, even on the scale of tens to hundreds of kilometres, can occur due to pieces of fertile thalli breaking off and being carried away from the point of origin by currents (Thiel and Gutow 2005; Alberto et al. 2011; Coleman et al. 2011). Additionally, Cie and Edwards (2011) have uncovered a mechanism facilitating vertical transport of kelp zoospores into the shallower portions of the water column, which has been proposed to enable long distance dispersal of zoospores via current action. Indeed, it has recently been suggested that spore dispersal is not likely to be a limiting factor in maintaining connectivity between Norwegian *S. latissima* populations (Andersen 2013). Therefore, it remains an open question as to whether *S. latissima* populations are genetically distinct (limited dispersal - spores recruiting within their population of origin), or whether there is some degree of genetic homogenization and thus lack of pronounced spatial or geographic structure (long distance dispersal, even if only occasional, resulting in a certain number of spores

79 recruiting to distant populations). Until now this question has not been investigated, largely because
80 the necessary genetic tools have been lacking.
81 As a primary producer, *S. latissima* plays a crucial role in the functioning of the subtidal
82 ecosystem and dense *S. latissima* stands provide a habitat for a plethora of marine organisms (Bekby
83 and Moy 2011). Additionally, this is a species of high economic importance, with applications in the
84 food, bioenergy, cosmetic and pharmaceutical industries. *S. latissima* is considered to be a species of
85 choice for the production of brown seaweed biomass in Europe (Peteiro 2013) because it has a
86 shorter life cycle than those of *L. digitata* and *Laminaria hyperborea*, the other two most abundant
87 kelps along the European coast. Furthermore, *S. latissima* is closely related to *Saccharina japonica*, a
88 domesticated species that occurs along the north-western coasts of the Pacific Ocean. Indeed, *S.*
89 *japonica* is one of the most economically important types of seaweed in the seaweed farming
90 industry with a long history of breeding programs and artificial seedling-rearing techniques (Tseng
91 2001). The complete genome of *S. japonica* has been recently published providing a vital resource for
92 algal genome studies and improvement of traits of economic importance (Ye et al. 2015). For these
93 reasons *S. latissima* is a good model species for studying algal domestication in Europe, and, as a
94 likely commercially important species, it is about to become a subject of studies aimed at identifying
95 strains exhibiting traits of interest for cultivation.
96 Given the significant ecological and commercial importance of *S. latissima*, it is of particular
97 concern that over the past five to 10 years this species has experienced large scale population
98 declines at sites in Norway, where 50 % of Europe's *S. latissima* forests are found, as well as in
99 Sweden and Germany (Bekby and Moy 2011). It is therefore urgent to develop a set of species-
100 specific genetic markers that could be employed to elucidate and monitor *S. latissima* population
101 structure and hence provide guidance on how to best implement conservation and management
102 strategies.
103 Microsatellites, also known as simple sequence repeats (SSRs), are among the most
104 commonly used marker types in evolutionary and ecological studies. SSRs are appropriate for the

105 determination of genetic diversity because they are abundant in the genome, are highly
106 polymorphic, are highly reliable, and usually have a codominant mode of inheritance. The
107 development of microsatellites has traditionally involved a large input of time and labour for the
108 construction, enrichment and sequencing of genomic libraries (Edwards et al. 1996). In contrast,
109 identification of SSRs from expressed sequences (EST-SSRs) is a fast, efficient and relatively low cost
110 alternative (Bouck and Vision 2007). EST-SSRs facilitate physical mapping and tend to be widely
111 transferable between species and even between genera (Liu et al. 2012). Here we describe the
112 development of a suite of EST-based (Heinrich et al. 2012), *S. latissima*-specific microsatellite
113 markers. Additionally, we test the ability of these newly developed SSRs to differentiate between
114 individuals from six natural populations of *S. latissima* distributed from Brittany to Norway and
115 investigate the importance of inter-population connectivity. We also discuss the potential uses of
116 these markers for applied research (conservation, strain selection) on this kelp species.
117
118 **Material and Methods**
119 *Sample collection and DNA extraction*
120 Sixteen mature *S. latissima* sporophytes were collected from each of six sampling locations. The six
121 sites, SB - Mousterlin, France, NB - Roscoff, France, PDC - Audresselles, France, DEN - Ebeltoft,
122 Denmark, SWE - Tjarno, Sweden, and NOR - Ny-Alesund, Spitsbergen, Norway, followed a latitudinal
123 gradient (Figure 1). Upon collection, a disk of tissue (diameter 2 cm) was cut out from a spot that was
124 free of epiphytes and stored with silica gel. If possible, the sample was taken from close to the
125 junction of the stipe and the blade as the aim was to sample the youngest tissue.
126 Total genomic DNA was extracted from five to 10 mg of dry tissue using the Nucleospin® 96
127 plant kit (Macherey-Nagel, Düren, Germany). The extractions were performed according to the
128 manufacturer's instructions except that samples were left in the lysis buffer at room temperature for
129 one hour rather than being heated to 65 °C for 30 minutes. The extracted DNA was eluted in 200 µL
130 of the supplied elution buffer.

EST-SSR screening and primer design

A total of 400,503 *S. latissima* ESTs obtained from the Sequence Read Archive (SRA) database (accession number: SRR305166 (Heinrich et al. 2012)) were screened for di-, tri-, tetra-, and pentanucleotide motif SSRs using the Sputnik program (http://espressosoftware.com/sputnik/). Seven thousand and sixty four ESTs were found to contain at least one SSR, with 438 (6 %) containing a dinucleotide repeat microsatellite, 2,330 (33 %) containing a trinucleotide repeat microsatellite, 1,765 (25 %) containing a tetranucleotide repeat microsatellite, and 2,531 (36 %) containing a pentanucleotide repeat microsatellite (Table 1). For loci where the flanking sequence on either side of the repeat was of sufficient length, PCR amplification primers were designed in Primer3 (Rozen and Skaletsky 2000). In-house scripts were used to specify the primer selection conditions, which used default values for all parameters except the optimum melting temperature, which was set to 60 °C (range 57 - 62 °C) and the expected product size, which was set at 350 bp.

This resulted in the design of 1,878 primer pairs corresponding to 270 (14.4 %) di-, 849 (45.2 %) tri-, 428 (22.8 %) tetra-, and 331 (17.6 %) pentanucleotide repeat loci (Table 1). In order to cost-effectively test these primers for amplification, and then check the amplified loci for levels of polymorphism within the sampled individuals, we reduced the number of loci used in the trials to 96 (i.e. capacity of a microtiter plate), employing the following criteria (see Table 1 for details). First, we eliminated all loci which contained an unknown base ("N") within the repeat region. Next, for the di- and trinuclotide motif SSRs only those loci which comprised at least eight repeat units were retained, and for the tetra- and pentanucleotide motif SSRs only those loci which comprised at least five repeat units were retained (184 loci remaining - Table 1). Finally, the number of loci was reduced to 96 such that loci with the highest number of repeats were selected for each of the four motif categories.

Amplification trials using the 96 primer pairs

Ninety six primer pairs were tested for amplification using DNA from six randomly chosen *S. latissima* individuals, one individual from each locality. Amplifications were carried out in 10 µL reaction volumes with each reaction comprising: 2 µl of DNA template diluted 1:50, 1x GoTaq® Flexi buffer (Promega Corporation), 2 mM $MgCl_2$, 150 µM of each dNTP (Thermo Fisher Scientific Inc., Waltham, MA, USA), 30 pmol forward primer, 30 pmol reverse primer, 30 pmol fluorescent-labelled M13(-21) primer (Eurofins MWG Operon, Ebersberg, Germany) and 0.35 U GoTaq® Polymerase (Promega Corporation). The 5' end of each forward primer included an M13 (-21) universal sequence tag (5'-TGTAAAACGACGGCCAGT-3') to enable the incorporation of the universal fluorescent-labelled M13 (-21) primer for detection of the amplified products on an ABI3130 XL DNA Analyser (Applied Biosystems, Foster City, CA, USA). The 96 loci were divided into four randomly selected groups (24 loci per group), and each of the four groups of loci was labelled with a different color. Amplifications were carried out in a BioRad DNA Engine Peltier Thermal Cycler under the following conditions: initial denaturation at 95 °C for 5 min, 12 cycles of denaturation at 95 °C for 30 s, annealing at 65 °C for 30 s (-1 °C/cycle - touchdown until 53 °C), and extension at 72 °C for 30 s, and 30 cycles of denaturation at 95 °C for 30 s, annealing at 53 °C (optimum M13 (-21) annealing temperature) for 30 s, extension at 72 °C for 30 s, followed by a final extension of 72 °C for 10 min.

The PCR products labelled with the four different colours were pooled (24 pools in total) and diluted 1:10 with water. Next, two µL of the diluted PCR product pool were added to 10 µL of loading buffer made up of 0.5 µL of the SM594 size standard (Mauger et al. 2012) and 9.5 µL of Hi-Di formamide, denatured at 95 °C for three minutes, and run in an ABI 3130 XL capillary sequencer (Applied Biosystems, Foster City, CA, USA). Genotypes were scored manually in Genemapper version 4.0 (Applied Biosystems, Foster City, CA, USA).

Forty five of 96 primer pairs did not amplify a product from any of the six individuals despite several re-amplifications and were not further analysed.

Investigating levels of polymorphism for 51 EST-SSR loci

The 51 loci retained after the amplification tests were subsequently checked for levels of polymorphism within the sample of 96 *S*. *latissima* individuals. Amplifications were carried out in same manner as described for the amplification trials.

Within the sample of 96 *S*. *latissima* individuals, 36 of the 51 loci were polymorphic. However, two of the polymorphic loci failed to amplify in many of the individuals originating from NB, PDC, DEN, and SWE (Sacl-70 amplified in 76 out of 96 samples, and Sacl-77 in 72 samples). As re-amplifications did not resolve this issue, these two loci were not retained for further analyses (see Supplementary Table 1 for the primer sequences and repeat motif details for these two loci, and Supplementary Table 2 for information on the 15 monomorphic loci). Two pairs of linked loci were detected among the remaining 34 loci (see *Statistical analyses* section) and, after removal of one locus from each pair, the remaining 32 loci were used for population genetics analyses.

Multiplex development

In order to optimize the usage of the 32 loci in future population genetics studies, we used Multiplex Manager v 1.2 (Holleley and Geerts 2009) with default parameters to organize the amplification of the loci in as few multiplex reactions as possible. Three multiplexes were created: multiplexes 1 and 2 comprised 10 loci each, and multiplex 3 comprised 12 loci (Table 2). For each multiplex, two disparate sets of loci were amplified separately to minimize primer dimer formation, but the two sets were pooled before genotyping. For multiplex 1 these were set M1i and set M1ii, whereas for multiplex 2 these were set M2i and set M2ii, and for multiplex 3 these were set M3i and set M3ii (Table 2). Each of the six multiplex sets was amplified in the same manner as described for the amplification trials, except that we added 40 pmol of each forward fluorescently-labelled primer (Eurofins MWG Operon, Ebersberg, Germany), 40 pmol of each reverse primer, and 0.5 U GoTaq® Polymerase. Amplifications were carried out in a BioRad DNA Engine Peltier Thermal Cycler under the following conditions: initial denaturation at 95 °C for 5 min, 10 cycles of denaturation at 95 °C for 30 s, annealing at 65 °C for 30 s (-1 °C/cycle – touchdown to 55 °C), and extension at 72 °C for 30 s, and

30 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s, followed by a final extension of 72 °C for 10 min.

Statistical analyses

We used Genepop v 4.3 (Rousset 2008) to test for linkage disequilibrium between all pairs of the 34 loci within each of the six sampling sites as well as for all localities combined (global test employing Fisher's method). This analysis used the following Markov chain parameters: dememorisation: 100,000; batches: 1,000; iterations per batch: 50,000. All significance levels were adjusted from an alpha value of 0.01 using sequential Bonferroni correction to allow for multiple tests on the same dataset (Rice 1989). We omitted locus Sacl-14 from further analyses since it exhibited significant linkage disequilibrium with locus Sacl-66 in three localities (NB, SB, DEN) and across all populations. We also omitted locus Sacl-72 since its allele patterns were identical to those of locus Sacl-08 (i.e. Sacl-72 alleles were always one base larger than Sacl-08 alleles) for all individuals with the exception of a genotype for a single NOR individual.

For each of the remaining 32 loci we calculated the following statistics using the combined sample of all 96 amplified individuals with the indicated R (R Development Core Team 2015) package. Number of alleles (N_A) and the polymorphic information content (PIC) were estimated in *PopGenKit* (Paquette 2012), whereas *hierfstat* (Goudet 2014) was used to obtain the observed heterozygosity (H_O) and the overall gene diversity (H_T) (Nei 1987). All significance levels were adjusted using sequential Bonferroni corrections (Rice 1989). Next, we computed for each of the 32 loci within each locality, allelic richness (Ar), H_O, and within-population gene diversity (H_S) (Nei 1987) using *hierfstat*, with the per locus "All Pops" Ar and H_S obtained by averaging these two parameters over the six localities. F_{IS} values for each locus within each locality were calculated in *Demerelate* (Kraemer and Gerlach 2013). Significance for each estimate was obtained by generating 10,000 bootstraps. The "All Pops" F_{IS} and F_{ST} parameters were calculated in *hierfstat*.

The frequencies of null alleles of the 32 loci were estimated using the Brookfield 1 equation (Brookfield 1996). The equation was run for each of the genotyped loci, separately for each of the six sampling locations, using Micro-Checker 2.2.3 (Van Oosterhout et al. 2004). The confidence interval for the Monte Carlo simulations of the homozygote frequencies was set to 95 %.

Subsequent analyses were performed using a reduced set of 25 loci (after removing loci with null alleles and/or exhibiting odd allelic patterns, see *Results* and Supplementary Material 1). For each of the six localities, we estimated the mean Ar and mean H_S by averaging over the 25 locus specific values presented in Supplementary Table 3 (for underlined loci only). The H_S means were compared between the six populations using the Kruskal-Wallis test. F_{IS} for each of the six populations was estimated as the multilocus estimate of Wright's inbreeding coefficient using *diveRsity* (Keenan et al. 2013). Significance was tested with 10,000 bootstrap iterations, and the values were considered to be significant if the bias-corrected 95 % confidence intervals did not overlap zero. The global (i.e. over all loci and all populations) values for H_S, F_{IS}, and F_{ST} were estimated by calculating the average of the over-population means (obtained from *hierfstat*) for each of these parameters over the 25 loci.

To estimate the level of genetic differentiation between the six sampling localities, pair-wise F_{ST} (Weir and Cockerham 1984) values were computed in *hierfstat*. The significance of the comparisons was estimated by performing 10,000 bootstraps over loci, with the comparisons judged significant if the bootstrap generated confidence intervals that did not overlap zero.

To gain further insights into *S. latissima* population subdivisions across our sampling range, and in particular to determine if the six clusters defined based on the sampling localities reflect the biological reality, we used the Discriminant Analysis of Principal Components (DAPC) method implemented in the *adegenet* R package (Jombart 2008; Jombart et al. 2010). This multivariate statistical approach seeks to provide an efficient description of genetic substructure through the use of a limited number of synthetic variables called the discriminant functions. DAPC infers the number of clusters of genetically related individuals (i.e. the number of distinct populations) by partitioning

the original variables (allele frequencies) into within-group and between-group components such that between-group variance is maximized. We used the six predefined clusters as the input for DAPC as the F_{ST} analysis showed all pairwise comparisons to be significant (see *Results*).

Finally, we tested for isolation by distance (IBD) between the six *S. latissima* populations. The geographic distance was the "as crow flies" direct distance in kilometres between each of the localities, whereas the genetic distance was Rousset's $F_{ST}/(1- F_{ST})$ estimate for each pair of populations (Rousset 1997). The significance of the IBD relationship was tested statistically with a Mantel test in R package *vegan* (Oksanen et al. 2015) with 999 permutations.

Results

Of the 96 primer pairs selected for the amplification trials, amplification was successful for 51 loci (53 %) (Supplementary Figure 1). Of these, two loci (Sacl-70 and Sacl-77) were removed due to inconsistent amplification in a larger sample of individuals, and two were dropped because they were found to be in linkage with other loci (Sacl-08 and Sacl-14). Of the 47 remaining EST-SSR loci, 32 loci (68 %) were polymorphic within the analysed sample of 96 *S. latissima* sporophytes originating from six disparate geographic locations (Figure 1). These 32 loci displayed variable levels of polymorphism, with PIC ranging from 0.010 to 0.823 with a mean of 0.350, and 11 (34 %) of the loci displaying values above 0.450. The number of alleles per locus ranged from two to 13 while the average over 32 loci was 5.41 alleles per locus (Table 2). Observed heterozygosity H_O was the highest at 0.875 for locus Sacl-49, whereas locus Sacl-27 was homozygous within all the scored individuals (Table 2). Mean H_O was 0.253 and the average overall gene diversity (H_T) was 0.385. H_O was lower than H_T at 24 of the 32 loci, probably because of a Wahlund effect due to pooling of individuals from multiple genetic clusters, but other contributing causes (inbreeding or the presence of null alleles) might also be involved. To minimize the possibility of the Wahlund effect, the per locus statistics were estimated within each of the six geographic localities (Supplementary Table 3).

285 Based on these data, seven loci were considered unusable for population genetics analyses
286 within the context of our dataset, and were thus removed from all subsequent analyses (details are
287 given in Supplementary Material 1 and in Supplementary Table 3). Briefly, these seven loci (Sacl-06,
288 Sacl-09, Sacl-27, Sacl-47, Sacl-49, Sacl-88, Sacl-90) were removed because they exhibited significant
289 null allele frequency at multiple geographic localities, displayed low levels of polymorphism, or they
290 exhibited odd allelic patterns with for instance fixed heterozygosities in some populations (highly
291 negative F_{IS}, Supplementary Table 3). When significant null allele frequency was detected for highly
292 polymorphic loci, but only for a single population, such loci were retained for subsequent analyses
293 (i.e. Sacl-24, Sacl-54, or Sacl-65, Supplementary Table 3). Note that, although the seven loci that were
294 removed were deemed unsuitable for our analyses, they could be of use for other types of studies or
295 studies on other populations of *S. latissima*, thus they are also reported in this manuscript and were
296 optimized as part of the multiplexes specified in Table 2.
297 Of the 25 loci retained, 18 were monomorphic within at least one of the six localities
298 (Supplementary Table 3). Thus the within-population observed allelic richness averaged over the 25
299 loci was low, ranging from 2.080 in SWE to 2.840 in SB (Figure 2, Supplementary Table 3). The highest
300 observed heterozygosity value was 0.938 for locus Sacl-38 within the NOR samples, while the lowest
301 value (0.063) was observed within NB for five loci and within DEN for three loci (Supplementary Table
302 3). Similarly, the locus-specific, within-population gene diversity (H_S) values were generally low. The
303 multilocus within-population H_S means ranged from 0.204 for PDC to 0.323 for NOR (Figure 2,
304 Supplementary Table 3). The H_S means were not significantly different between the six geographic
305 localities (Kruskal-Wallis test $X^2(5) = 4.549$, p-value = 0.473), and thus did not vary with latitude. This
306 was also the case for the Ar means (Figure 2). The within-population multilocus F_{IS} values were not
307 significantly different from zero for any of the six geographic localities, which indicates that random
308 mating commonly occurs within the studied populations.
309 Despite the generally low level of within-population polymorphisms displayed by the 25 loci
310 (grand H_S mean of 0.245 - Supplementary Table 3), there were 47 private alleles at 23 of 25 loci. The

311 pattern of genetic differentiation among populations was highly significant with an overall F_{ST} mean
312 of 0.259 and 10 loci exhibiting F_{ST} values above 0.300 (Supplementary Table 3).
313 All of the pairwise multilocus F_{ST} comparisons were significant. The least differentiated
314 populations were SB and NB, with F_{ST} of 0.077 (Table 3). These populations exhibited approximately
315 equal levels of divergence from PDC, with the geographically more distant SB being also slightly more
316 distant genetically. The F_{ST} between DEN and SWE was the second lowest at 0.120, with both these
317 populations being highly genetically differentiated from NOR (F_{ST} above 0.250 in both comparisons)
318 (Table 3). The highest levels of differentiation were between the three north European populations
319 and the three French populations, with F_{ST} ranging from 0.358 to 0.562 (Table 3). The IBD model was
320 rejected (Mantel statistic r = 0.275, significance = 0.112) (Figure 3), since the more geographically
321 distant NOR population was less differentiated from the three French localities than the less distant
322 DEN and SWE populations.
323 The DAPC results correlate well with the F_{ST} estimates (Figure 4). Axis 1 (which explains
324 45.12% of the variance) separates the three French sampling localities from the three Scandinavian
325 sampling sites (Figures 4a and 4b), while axis 2 (which explains 10.83 % of the variance) separates
326 NOR from the other five sampling localities (Figure 4a) and axis 3 (which explains 6.03 % of the
327 variance) partitions the French sampling site into a PDC and an SB/NB cluster (Figure 4b). Indeed, the
328 two pairs of localities displaying the lowest pairwise F_{ST} values (SB and NB, and DEN and SWE) were
329 the least well defined by DAPC. There was no clear separation of these two pairs of sites within any
330 of the first three principal components of the DAPC, as there was always some overlap between
331 individuals from these two localities (Figures 4a and 4b). The admixture plot (Figure 4c) however
332 revealed that most of the SB/NB individuals were assigned to their population of origin, with only
333 nine out of 32 individuals originating from the SB/NB localities exhibiting some level of admixture.
334 The level of admixture between PDC and the SB/NB localities was even lower (Figure 4c). The
335 admixture situation within the Scandinavia sampling localities mirrored that within the French sites,
336 although the DEN/SWE admixture was less substantial in comparison to the SB/NB admixture (Figure

337 4c). Finally, there was no evidence of admixture between the two sampling regions (i.e. France and
338 Scandinavia, Figure 4c).
339
340 **Discussion**
341 Of the 96 primer pairs that were tested for amplification with six *S. latissima* sporophytes, 51
342 (53 %) were deemed successful (though subsequently two of these loci were dropped due to
343 inconsistency of amplification within a larger sample of individuals) (Supplementary Figure 1). Our
344 markers therefore exhibited a higher positive amplification rate than was reported by Wang et al.
345 (2011) (36.5 %) for *Laminaria*-specific EST-SSRs. For the loci presented herein, the success of
346 amplification was poorest for loci comprising tetranucleotide motifs (31.3 %) and best for those
347 comprising dinucleotide motifs (81.3 %), whereas it was just under 60 % for loci comprising the
348 trinucleotide and pentanucleotide motif types (Supplementary Figure 1).
349 Utilization of expressed sequence tags to identify so called genic SSRs is a relatively recent
350 approach, but one that has been successfully employed in a vast array of taxa (Ellis and Burke 2007).
351 The main advantage over the traditional method of isolation of microsatellite markers (in this case
352 termed genomic SSRs) is that obtaining SSRs from ESTs is considerably less labour intensive and also
353 cheaper (Zane et al. 2002; Ellis and Burke 2007). In this study, there was a high attrition rate amongst
354 the newly developed markers, with only around 50 % of the amplification trials being successful, 33
355 % of markers being polymorphic, and just over a quarter of the markers (25 out of 96 – 26 %) being
356 retained for the population genetics analyses. However, the time and effort spent on acquiring these
357 markers was not excessive, and a reasonably high number of markers were obtained. The traditional
358 method of isolating SSRs, partly due to the above mentioned difficulties, rarely results in more than
359 20 markers being produced. Indeed, 16 genomic SSRs have recently been developed for *S. latissima*
360 (Møller Nielsen et al., submitted). Note that, as 184 loci were retained after the reduction step
361 “Repeat no. > 8 (di & tri) & > 5 (tetra & penta)” (Table 1) and only 96 of these have been tested for

362 amplification and polymorphism so far, there is potential to develop additional *S*. *latissima* specific
363 EST-SSRs from this analysis if needed.
364 A potential disadvantage of EST-SSRs as genetic markers is that they are commonly
365 considered to harbour lower levels of polymorphism than genomic SSRs because they are derived
366 from expressed, and thus conserved, genes (Gupta et al. 2003; Ellis and Burke 2007; Guichoux et al.
367 2011). However, various studies have demonstrated that the EST-derived SSRs were sufficiently
368 diverse to be useful markers for population genetics analyses (Dong et al. 2009; Cubry et al. 2014;
369 Teshome et al. 2015), and that conclusions regarding population structure and connectivity were
370 highly congruent whether elucidated with genomic or EST-derived SSRs (Woodhead et al. 2005; Kim
371 et al. 2008; Hu et al. 2011).
372 The number of alleles per locus we detected in our study (5.41 alleles per locus averaged
373 over 32 loci) was higher than or comparable to the genomic or EST-SSR sets described in the
374 literature. For example, in a closely related congener species S. *japonica* the average number of
375 alleles over 18 genomic SSR loci was 4.7 alleles per locus (Shi et al. 2007), whereas on average 5.7
376 alleles per locus were obtained for nine EST-SSR loci (Liu et al. 2010).
377 To demonstrate the applicability of this newly developed set of *S*. *latissima* EST-SSR markers
378 for effective management of this ecologically and economically important species, we investigated
379 the patterns of within-population genetic diversity and connectivity between six localities spanning a
380 latitudinal gradient from South Brittany, France to Spitsbergen, Norway using the reduced panel of
381 25 loci.
382 Our data indicate that the six studied populations were significantly genetically differentiated
383 (significant F_{ST}, Table 3), and formed genetically well-defined clusters even if there was some degree
384 of connectivity between the populations that were closest geographically (Figure 4). The clustering of
385 the two Brittany populations (SB and NB) into two distinct but not completely isolated groups fits
386 well with what was reported for two other kelps, *L*. *digitata* and *L*. *hyperborea* (Couceiro et al. 2013;
387 Robuchon et al. 2014) in the same region. The Ushant tidal front, which constitutes a seasonal

388 dispersal barrier, is probably one of the major factors that influence the observed pattern of kelp
389 population structure in this region as discussed in Couceiro et al. (2013).
390 Further north along the French coast, the PDC sampled sporophytes exhibited substantially
391 lower connectivity to the two Brittany localities than was detected within Brittany (i.e. between SB
392 and NB) (Table 3, Figure 4). The pattern of isolation between the Breton sites and PDC is likely
393 explained by habitat discontinuity as the shore line either side of the PDC sampling site is
394 characterized by long stretches of sandy substrate. Lack of suitable (i.e. rocky) substrate has been
395 proposed to be a major dispersal barrier for several other kelps (Fraser et al. 2010; Alberto et al.
396 2011; Couceiro et al. 2013, Robuchon et al. 2014). However, the status of PDC as a small, fragmented
397 and/or isolated population cannot be fully established from our data given that the expected
398 decrease in the within-population gene diversity (H_S) and allelic richness (Ar) compared to the
399 supposedly larger and better connected populations (i.e. SB and NB) was not significant (Figure 2).
400 Consequently, the genetic isolation of PDC from the Breton samples could simply be due to a
401 combination of the distance between the two sites and the limited dispersal capabilities of *S*.
402 *latissima*, and not necessarily be due to habitat discontinuity. On the other hand, as the H_S and Ar
403 estimates for the two Brittany sampling sites were particularly low, i.e. at least twice as low as those
404 described for *L*. *digitata* (Billot et al. 2003; Couceiro et al. 2013; Robuchon et al. 2014) and *L*.
405 *hyperborea* (Robuchon et al. 2014) sampled from nearby points along the Brittany coast, this could
406 suggest that *S*. *latissima* populations are characterized by smaller population size in comparison to
407 those of other kelps. Indeed, along the Brittany coast, *S*. *latissima* populations generally exhibited
408 more patchy distribution when compared to the dense and continuous forests of the other two kelp
409 species (Valero, personal observation). Genotyping of samples collected from additional localities
410 should provide more information about *S*. *latissima* population structure along the coast of France,
411 and in particular of Brittany.
412 Within Scandinavia, our data revealed moderate levels of connectivity between the *S*.
413 *latissima* populations at the SWE and DEN sampling localities, and clear delimitation of the NOR

414 population from those of the two former sampling sites. The observed admixture was less prevalent
415 than between SB and NB (Figure 4). Overall, this intermediate degree of connectivity between the
416 DEN and SWE populations suggests here also a role for currents in modulating *S*. *latissima* population
417 structure in this section of the North Sea (the Skagerrak strait). There was almost no admixture
418 between SWE/DEN and NOR (Figure 4c). This near absence of connectivity between the NOR and the
419 DEN/SWE sampling localities is in accordance with expectations, given that these sites are separated
420 by more than 2,000 km. Overall, these analyses validated the applicability of the 25 microsatellite loci
421 in elucidating *S*. *latissima* population structure.
422 Phylogeographic studies have demonstrated the effects of glacial-interglacial cycles during
423 the Quaternary (~2.6 Myr to present-times) on the patterns of contemporary genetic diversity,
424 revealing several genetically rich refugia zones and a general trend of decrease in diversity with
425 latitude (Hewitt 2004). In the northern Atlantic, several of these putative long-term climatic refugia
426 have been identified for several marine species including seaweed. These refugia comprise
427 Brittany/South Ireland (corresponding to the palaeo-Celtic Sea/Channel), northwestern Iberia, and
428 for a smaller set of species, central-south Iberia and northwest Africa (Provan 2013; Neiva et al. 2016
429 for a recent review of North Atlantic Fucaceae). However, in the present study we did not detect a
430 latitudinal gradient in genetic diversity as none of the localities displayed significantly higher genetic
431 diversity than others (Figure 2) even if populations from Brittany, located in a putative refugia, were
432 included in our data set. Erosion of genetic diversity may have occurred in this region due to low
433 effective population size because Brittany was at the northern range limit of the species distribution
434 during the colder periods of the Last Glacial Maximum. Reduced genetic diversity was observed for
435 the kelp *Saccorhiza polyschides* in that region and was explained by models predicting marginal,
436 although persistent, populations in the Brittany/Charentes regions (Assis et al. 2016). In addition,
437 note that the *S*. *latissima* range extends as far south as Portugal (Bartsch et al. 2008), another
438 predicted refugia of the northern Atlantic, and thus more genetically rich refugial populations might
439 be found on the Iberian Peninsula. The three Scandinavian sampling localities all exhibited very low

within-population gene diversity and allelic richness, similar to those of the French localities. This was a somewhat surprising result as the lowest genetic diversity might have been expected for the NOR population since it was sampled at the northern limit of the *S. latissima* distribution (Bartsch et al. 2008). This population is expected to show the same pattern of reduced diversity as has been reported for most North Atlantic Fucaceae (Neiva et al. 2016), which is the result of gene surfing at the leading edge of expansion during recolonization after the last glaciation period.

Relatively low within-population genetic diversity (measured in terms of Ar and H_S) was detected within each of the six *S. latissima* sampling localities as these values were two-fold lower than the 4.24 alleles per locus and expected heterozygosity of 0.550 averaged over 16 genomic SSRs analysed in *S. latissima* populations originating from the North Sea-Baltic Sea transition zone (Møller Nielsen et al., submitted), and two-fold lower than the values observed within the Brittany populations of *L. digitata* and *L. hyperborea* (Billot et al. 2003; Couceiro et al. 2013; Robuchon et al. 2014). As discussed above, the relatively low within-population genetic diversity observed along the French coast is most probably explained by the small effective population size since populations are small and patchily distributed in this region. In contrast, the low H_S estimates detected within the DEN and SWE sampling localities are counterintuitive as these sites are expected to form part of a large interconnected population(s) as they are located within the Skagerrak strait and the Norwegian side of the strait is known to harbour some of the largest kelp forests anywhere in Europe (Bekkby and Moy 2011). However, it has been observed that over the last five to 10 years extensive areas of the *S. latissima* forests have all but disappeared from along the coast of Norway (including in the Skagerrak strait), as well as from the coastlines of Germany, Denmark, and Sweden (Pehlke and Bartsch 2008; Andersen et al. 2011; Bekkby and Moy 2011; Andersen 2013). Thus it is conceivable that the fragmentation of populations, driven by anthropogenically induced kelp forest demise (Wahl et al. 2015), has occurred at such speed and to such an extent as to have resulted in a number of small, isolated patches that have retained only some fraction of their genetic diversity. However, we cannot exclude the hypothesis that the low within-population variability of the EST-derived

466 microsatellites that was detected in our study is due to the fact that a proportion of the EST-SSR
467 markers are located in coding regions and are thus generally less variable compared to randomly
468 distributed genomic SSRs (see below).

469 The multilocus F_{IS} was not significant within any of the six localities, ranging from -0.066 in SB
470 to 0.095 in NB (Supplementary Table 3). Therefore, it is not possible to conclude that inbreeding was
471 occurring within any of the sampled localities. Moreover, we did not detect any identical multilocus
472 genotypes indicating that vegetative propagation (clonality) was not predominant within the studied
473 localities.

474 Comparison of the *S*. *latissima* genetic diversity at the European scale revealed very high
475 levels of differentiation and a total absence of admixture between the three French and the three
476 Scandinavian localities (Table 3, Figure 4). The pairwise F_{ST} estimates are particularly high, resembling
477 the between rather than within-species estimates, and they are likely explained by the large
478 geographic scale concerned in addition to the low dispersal capabilities of this seaweed species.
479 However, there could also be other underlying factors, as discussed below.

480 Although SSRs are generally considered neutral genetic markers, EST-derived microsatellites
481 are more likely to be involved in regulatory processes and to influence the differential expression of
482 genes than microsatellites from other parts of the genome (Li et al. 2004). Thus the substantial
483 genetic differentiation within *S*. *latissima* that was detected at the European level may not reflect
484 only neutral genetic diversity, but also adaptive genetic diversity (Holderegger et al. 2006). In other
485 words, if some of the newly developed loci are important for the regulation of genes vital for
486 adaptation to different environmental regimes, then these loci would be under diversifying selection
487 and hence become more differentiated between the disparate sampling localities than neutral SSR
488 loci. Several studies have provided evidence of ecotypic differentiation between *S*. *latissima*
489 populations from temperate (North Sea, Helgoland) or Arctic (Spitsbergen, Svalbard) regimes (Müller
490 et al. 2008; Olischläger et al. 2014) with strong indications that differential expression of genes plays
491 a major role in the adaption of the distinct *S*. *latissima* ecotypes to the local conditions (Heinrich et

492 al. 2012). Taking together 1) the fact that our samples originated from across a latitudinal gradient
493 and from populations that have been exposed to different environmental conditions 2) the evidence
494 for local adaptation within *S*. *latissima* populations, and 3) the potential role of EST-derived SSRs in
495 regulating gene expression, it is conceivable that the high level of inter-population genetic
496 differentiation detected in this study could be influenced by these factors rather than simply reflect
497 the degree of inter-population connectivity. This scenario could also explain in part the low overall
498 (i.e. over all loci) within-population heterozygosity, as it is conceivable that selection has acted to
499 conserve the number of repeats for at least some of the loci within the localities. Indeed, a recent
500 study that used 13 *S*. *latissima*-specific genomic SSRs to genotype 225 *S*. *latissima* sporophytes
501 collected from one Swedish, one Norwegian, and six Danish sites, uncovered within-population gene
502 diversity to range from 0.451 to 0.613 (Møller Nielsen et al. submitted), thus two to three times
503 higher than what we have detected with the EST-SSRs. Therefore, the fact that the SSRs presented
504 herein are EST-derived should be taken into account when interpreting our findings. A comparative
505 study of both sets of markers on a larger sample size would allow outlier loci to be detected. In
506 addition, the genomic location of our EST-derived SSRs could be investigated in order to find out
507 which genes they are associated with and thus evaluate whether they potentially have a role in the
508 adaptivity of the disparate *S*. *latissima* populations to the local conditions.
509 The six *S*. *latissima* sampling localities did not follow the IBD pattern, with no significant
510 positive correlation detected between the genetic and geographic distance measurements (Figure 3).
511 This is a surprising result in light of a recent outcome of a meta-analysis, which found IBD for all of
512 the macroalgae species analysed (Durrant et al. 2014). The macroalgae species included in the study
513 by Durrant et al. (2014) originated from multiple taxonomic groups, exhibited various life histories,
514 and inhabited varied environments. In the present study, the IBD signal may be obscured due to the
515 DEN/SWE samples being more genetically differentiated from the Brittany localities than the levels of
516 genetic differentiation between NOR and the Brittany localities (Table 3), even though the
517 geographic distance between DEN/SWE and Brittany is much shorter than the geographic distance

between the NOR locality and Brittany (Figure 1). Moreover, for all the Scandinavian versus French comparisons, the furthest locality (SB) was the least differentiated genetically whereas the closest locality (PDC) was the most diverse (Table 3). We have already noted that habitat discontinuity probably isolates the PDC population from other localities, resulting in increased pairwise F_{ST} values that are greater than expected given the geographic distance. However, the reasons underlying the patterns involving the two Breton and three Scandinavian populations are open to speculation. Historical events could also have shaped the current population substructure, in particular vicariance due to the environmental and climatic changes that occurred during the last glacial maximum. As the spatial arrangement of the habitat patches will influence the IBD patterns (van Strien et al. 2015), another reason for the lack of detection of the IBD signal could be that the relative distances between the sampled populations are poorly replicated (for instance small distance between the two Brittany populations) and thus there is an uneven spread of the data points. Moreover, we cannot exclude the hypothesis that recent human impact may have modified *S*. *latissima* population structure, in particular in the Scandinavian region (see refs above), or that the effect of selection on EST-derived microsatellites prevents the detection of IBD signal (see above).

Based on the results we present here, we are confident that this newly developed set of *S*. *latissima*-specific SSRs will prove to be a very useful genetic tool for this commercially important kelp species. These markers will find applicability in elucidating the population structure and dispersal patterns of this species, information that will be crucial for making informed and effective conservation and population management decisions (i.e. avoiding risk of transplantation by taking care to avoid the transfer of strains between different geographic regions). Moreover, this set of markers could be used to identify strains/genotypes of interest (for instance strains displaying commercially important traits), to measure the environmental impact of cultivation and harvesting of wild *S*. *latissima* populations on biodiversity (Klinger 2015), and to construct genetic linkage maps for this species.

Acknowledgments This work benefited from the support of the French Government through the National Research Agency with regards to an investment expenditure program IDEALG, which reference is stated as ANR-10-BTBR-04, and was carried out within the context of the international research network "Diversity, Evolution and Biotechnology of Marine Algae" (GDRI N° 0803). The authors are deeply indebted to the Service Mer & Observation (SMO) of Roscoff, G. Cervin, and J. Coyer for sampling. The authors would also like to thank Gildas Le Corguille of the ABiMS Platform, Roscoff for his help with bioinformatics analyses. We are most grateful to the Biogenouest Genomics core facility for its technical support. We also would like to thank the two anonymous reviewers for their suggestions and comments.

References

Alberto F, Raimondi PT, Reed DC, Coelho NC, Leblois R, Whitmer A, Serrão EA (2010) Habitat continuity and geographic distance predict population genetic differentiation in giant kelp. Ecology 91 (1):49-56

Alberto F, Raimondi PT, Reed DC, Watson JR, Siegel DA, Mitarai S, Coelho N, Serrão EA (2011) Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. Mol Ecol 20 (12):2543-2554

Andersen SG, Steen H, Christie H, Fredriksen S, Moy FE (2011) Seasonal patterns of sporophyte growth, fertility, fouling, and mortality of *Saccharina latissima* in Skagerrak, Norway: Implications for forest recovery. J Mar Biol 2011:1-8

Andersen SG (2013) Patterns of *Saccharina latissima* recruitment. PLOS One 8 (12):e81092

Assis J, Coelho NC, Lamy T, Valero M, Alberto F, Serrão EA (2016) Deep reefs are climatic refugia for genetic diversity of marine forests. J Biogeogr

Bartsch I, Wiencke C, Bischof K, Buchholz CM, Buck BH, Eggert A, Feuerpfeil P, Hanelt D, Jacobsen S, Karez R, Karsten U, Molis M, Roleda MY, Schubert H, Schumann R, Valentin K, Weinberger F, Wiese J (2008) The genus *Laminaria sensu lato*: recent insights and developments. European J Phycol 43 (1):1-86

Bekkby T, Moy FE (2011) Developing spatial models of sugar kelp (*Saccharina latissima*) potential distribution under natural conditions and areas of its disappearance in Skagerrak. Estuar Coast Shelf S 95 (4):477-483

Billot C, Engel CR, Rousvoal S, Kloareg B, Valero M (2003) Current patterns, habitat discontinuities and population genetic structure: the case of the kelp *Laminaria digitata* in the English Channel. Mar Ecol-Prog Ser 253 (111):21

Bolton JJ, Germann I, Luning K (1983) Hybridization between Atlantic and Pacific representatives of the Simplices section of *Laminaria* (Phaeophyta). Phycologia 22 (2):133-140

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Mol Ecol 16 (5):907-924

Brookfield J (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5 (3):453-455

Cie DK, Edwards MS (2011) Vertical distribution of kelp zoospores. Phycologia 50 (4):340-350

584 Coleman MA, Roughan M, Macdonald HS, Connell SD, Gillanders BM, Kelaher BP, Steinberg PD (2011)
585 Variation in the strength of continental boundary currents determines continent-wide
586 connectivity in kelp. J Ecol 99 (4):1026-1032
587 Couceiro L, Robuchon M, Destombe C, Valero M (2013) Management and conservation of the kelp
588 species *Laminaria digitata*: using genetic tools to explore the potential exporting role of the
589 MPA "Parc naturel marin d'Iroise". Aquat Living Resour 26 (2):197-205
590 Cubry P, Pujade-Renaud V, Garcia D, Espeout S, Le Guen V, Granet F, Seguin M (2014) Development
591 and characterization of a new set of 164 polymorphic EST-SSR markers for diversity and
592 breeding studies in rubber tree (*Hevea brasiliensis* Müll. Arg.). Plant Breeding 133 (3):419-
593 426
594 Dong P, Wei YM, Chen GY, Li W, Wang JR, Nevo E, Zheng YL (2009) EST-SSR diversity correlated with
595 ecological and genetic factors of wild emmer wheat in Israel. Hereditas 146 (1):1-10
596 Durrant HM, Burridge CP, Kelaher BP, Barrett NS, Edgar GJ, Coleman MA (2014) Implications of
597 macroalgal isolation by distance for networks of marine protected areas. Conserv Biol 28
598 (2):438-445
599 Edwards K, Barker J, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several
600 microsatellite sequences in plants. Biotechniques 20 (5):758
601 Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99 (2):125-
602 132
603 Fraser CI, Thiel M, Spencer HG, Waters JM (2010) Contemporary habitat discontinuity and historic
604 glacial ice drive genetic divergence in Chilean kelp. BMC Evol Biol 10:203
605 Gerard VA, Du Bois K (1988) Temperature ecotypes near the southern boundary of the kelp
606 *Laminaria saccharina*. Mar Biol 97 (4):575-580
607 Gerard VA (1988) Ecotypic differentiation in light-related traits of the kelp *Laminaria saccharina*. Mar
608 Biol 97 (1):25-36
609 Goudet J (2014) hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-14.
610 Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E,
611 Salin F (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11 (4):591-611
612 Gupta P, Rustgi S, Sharma S, Singh R, Kumar N, Balyan H (2003) Transferable EST-SSR markers for the
613 study of polymorphism and genetic diversity in bread wheat. Mol Genet Genomics 270
614 (4):315-323
615 Heinrich S, Valentin K, Frickenhaus S, John U, Wiencke C (2012) Transcriptomic analysis of acclimation
616 to temperature and light stress in *Saccharina latissima* (Phaeophyceae). PLOS One 7
617 (8):e44342
618 Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Phil Trans R Soc
619 Lond B 359 (1442):183-195
620 Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for
621 landscape genetics. Landscape Ecol 21 (6):797-807
622 Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans
623 and optimizes multiplex PCR. Biotechniques 46 (7):511-517
624 Hu J, Wang L, Li J (2011) Comparison of genomic SSR and EST-SSR markers for estimating genetic
625 diversity in cucumber. Biol Plantarum 55 (3):577-580
626 Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers.
627 Bioinformatics 24 (11):1403-1405
628 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method
629 for the analysis of genetically structured populations. BMC Genet 11:94
630 Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: an R package for the
631 estimation and exploration of population genetics parameters and their associated errors.
632 Method Ecol Evol 4 (8):782-788
633 Kim KS, Ratcliffe ST, French BW, Liu L, Sappington TW (2008) Utility of EST-derived SSRs as population
634 genetics markers in a beetle. J Hered 99 (2):112-124

635 Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community
636 perspective. Ecology 84 (8):2007-2020
637 Klinger T (2015) The role of seaweeds in the modern ocean. Pers Phycol:31-39
638 Kraemer P, Gerlach G (2013) Demerelate: Functions to calculate relatedness on diploid genetic data.
639 R package version 0.8-1
640 Li Y-C, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and
641 evolution. Mol Biol Evol 21 (6):991-1007
642 Liu F, Wang X, Yao J, Fu W, Duan D (2010) Development of expressed sequence tag-derived
643 microsatellite markers for *Saccharina* (*Laminaria*) *japonica*. J Appl Phycol 22 (1):109-111
644 Liu F, Wang F, Duan D (2012) EST-SSR markers derived from *Laminaria digitata* and its transferable
645 application in *Saccharina japonica*. J Appl Phycol 24 (3):501-505
646 Mauger S, Couceiro L, Valero M (2012) A simple and cost-effective method to synthesize an internal
647 size standard amenable to use with a 5-dye system. Prime Research on Biotechnology 2:40-
648 46
649 Møller Nielsen M, Paulino C, Neiva J, Krause-Jensen D, Bruhn A, Serrão EA (Submitted) Genetic
650 variability of Saccharina latissima along a salinity boundary in the North Sea-Baltic transition
651 zone. J Phycol
652 Müller R, Wiencke C, Bischof K (2008) Interactive effects of UV radiation and temperature on
653 microstages of Laminariales (Phaeophyceae) from the Arctic and North Sea. Clim Res 37 (2-
654 3):203-213
655 Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, New York
656 Neiva J, Serrão EA, Assis J, Pearson GA, Coyer JA, Olsen JL, Hoarau G, Valero M (2016) Climate
657 oscillations, range shifts and phylogeographic patterns of North Atlantic Fucaceae. In: Hu Z-H,
658 Fraser CI (eds) Seaweed phylogeography. Springer
659 Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens
660 MHH, Wagner H (2015) vegan: Community Ecology Package. R package version 2.3-0
661 Olischläger M, Iniguez C, Gordillo FJ, Wiencke C (2014) Biochemical composition of temperate and
662 Arctic populations of *Saccharina latissima* after exposure to increased pCO_2 and temperature
663 reveals ecotypic variation. Planta 240 (6):1213-1224
664 Paquette SR (2012) PopGenKit: Useful functions for (batch) file conversion and data resampling in
665 microsatellite datasets. R package version 1.0
666 Pehlke C, Bartsch I (2008) Changes in depth distribution and biomass of sublittoral seaweeds at
667 Helgoland (North Sea) between 1970 and 2005. Clim Res 37 (2-3):135-147
668 Peteiro C, Freire Ó (2013) Biomass yield and morphological features of the seaweed *Saccharina*
669 *latissima* cultivated at two different sites in a coastal bay in the Atlantic coast of Spain. J Appl
670 Phycol 25 (1):205-213
671 Provan J (2013) The effects of past, present and future climate change on range-wide genetic
672 diversity in northern North Atlantic marine species. Fron Biogeogr 5 (1):60-66
673 R Core Team (2015) R: A language and environment for statistical computing. R Foundation for
674 Statistical Computing, Vienna, Austria
675 Rice WR (1989) Analyzing tables of statistical tests. Evolution:223-225
676 Robuchon M, Le Gall L, Mauger S, Valero M (2014) Contrasting genetic diversity patterns in two sister
677 kelp species co-distributed along the coast of Brittany, France. Mol Ecol 23 (11):2669-2685
678 Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation
679 by distance. Genetics 145 (4):1219-1228
680 Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows
681 and Linux. Mol Ecol Resour 8 (1):103-106
682 Rozen S, Skaletsky H (1999) Primer3 on the WWW for General Users and for Biologist Programmers.
683 In: Misener S, Krawetz S (eds) Bioinformatics Methods and Protocols, vol 132. Methods in
684 Molecular Biology™. Humana Press, pp 365-386
685 Shi Y, Yang G, Liu Y, Liao M, Li X, Cong Y (2007) Development of 18 polymorphic microsatellite DNA
686 markers of *Laminaria japonica* (Phaeophyceae). Mol Ecol Notes 7 (4):620-622

van Strien MJ, Holderegger R, Van Heck HJ (2015) Isolation-by-distance in landscapes: considerations for landscape genetics. Heredity 114 (1):27-37
Teshome A, Bryngelsson T, Dagne K, Geleta M (2015) Assessment of genetic diversity in Ethiopian field pea (*Pisum sativum* L.) accessions with newly developed EST-SSR markers. BMC Genet 16 (1):102
Thiel M, Gutow L (2005) The ecology of rafting in the marine environment. I. The floating substrata. Oceanography and Marine Biology: An Annual Review 42:181-264
Tseng C (2001) Algal biotechnology industries and research activities in China. J Appl Phycol 13 (4):375-380
Valero M, Destombe C, Mauger S, Ribout C, Engel CR, Daguin-Thiebaut C, Tellier F (2011) Using genetic tools for sustainable management of kelps: a literature review and the example of *Laminaria digitata*. CBM-Cahiers de Biologie Marine 52 (4):467
Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4 (3):535-538
Wahl M, Molis M, Hobday AJ, Dudgeon S, Neumann R, Steinberg P, Campbell AH, Marzinelli E, Connell S (2015) The responses of brown macroalgae to environmental change from local to global scales: direct versus ecologically mediated effects. Pers Phycol
Wang G, Tan X, Shen J, Li J, Zhang L, Sun J, Wang B, Weng M, Liu T (2011) Development of EST-SSR primers and their practicability test for *Laminaria*. Acta Oceanol Sin 30 (3):112-117
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38 (6):1358-1370
Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Mol Ecol 14 (6):1681-1695
Ye N, Zhang X, Miao M, Fan X, Zheng Y, Xu D, Wang J, Zhou L, Wang D, Gao Y (2015) *Saccharina* genomes provide novel insight into kelp biology. Nature Communications 6
Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11 (1):1-16

TABLE 1. Steps involved in the selection of 96 *Saccharina latissima* SSR loci from 7064 initial candidates obtained by analysis of cDNA library sequence data. The percentage of loci with specific motif types is also presented for each step of the selection process.

Reduction step	Di	Tri	Tetra	Penta	Number of retained loci
Sputnik output	6%	33%	25%	36%	7064
Primer3 output	14%	45%	23%	18%	1878
Loci with "N" base removed	14%	45%	23%	18%	1871
Repeat no. > 8 (di & tri) & > 5 (tetra & penta)	20%	30%	32%	19%	184
Final selection	17%	33%	33%	17%	96

Di refers to SSR loci with dinucleotide repeat motifs; Tri to SSR loci with trinucleotide repeat motifs; Tetra to SSR loci with tetranucleotide repeat motifs; Penta to SSR loci with pentanucleotide repeat motifs.

TABLE 2. Technical information (primer sequence, dye, and multiplex), characteristics (repeat motif, expected product size, and observed size range), and polymorphism estimates (N_A, H_O, H_T, *PIC*) for 32 EST-SSR loci based on amplification from genomic DNA of 96 *Saccharina latissima* sporophytes.

Locus	Primer sequence (5'-3')	Repeat motif in the EST	Dye	Contig no. on GenBank BioProject PRJNA80101	Expected product size (bp)	Observed size range (bp)	N_A	H_O	H_T	*PIC*	Multiplex
Sacl-03	F: GTTGTTACGGTTGGCGTTG R: TTCAATAATCGCAGGAAGCAC	$(CA)_9$	PET	10167	212	204-214	6	0.479	0.679	0.625	M1i
Sacl-06	F: TTTGTAGATAGTGCCGCTGGT R: GTAGCCTGCCGAAGAAATAAA	$(AG)_{15}$	PET	13119	158	141-165	9	0.271	0.565	0.519	M1ii
Sacl-08	F: CAGACCTCACCAAGGCAAAG R: TCCACGCACATACAGCAAA	$(TA)_{11}$	PET	14056	130	118-144	4	0.073	0.071	0.069	M3ii
Sacl-09	F: ACCGTTCCCAATAACTACCC R: ATGGACCTATCGCCCACA	$(AGC)_{13}$	PET	7028	332	320-332	5	0.531	0.425	0.385	M1i
Sacl-11	F: AAGAGGTGGTTGCTCGGTT R: CACTGATGGATGCCCTGTC	$(CTG)_{12}$	FAM	8341	302	294-306	5	0.281	0.356	0.322	M1i
Sacl-13	F: CTCAGGTGTTCGGTGCCCT R: GCTCGGTAGATGGTTAGTC	$(GCA)_9$	FAM	13741	268	235-244	3	0.083	0.147	0.139	M3i
Sacl-19	F: GCGTATTCCTAAACACCTCCC R: CGATGACTGCCACCACAG	$(ACC)_9$	PET	2896	175	163-184	5	0.156	0.533	0.428	M2i
Sacl-21	F: TCTCAACTCCAAACAGAGCGT R: AGCAGCAGAAGCAGGCGA	$(GCT)_9$	FAM	17486	184	151-166	5	0.177	0.203	0.194	M3i
Sacl-24	F: GCCAACCTATCATTCAAAGCA R: TGTGGGAGCAACATCCTCA	$(TGA)_{11}$	PET	16267	103	87-132	13	0.375	0.515	0.474	M2i
Sacl-25	F: TTCGTTTCAGTTGGTGGGTT R: TGCTGTAGTAGTGATTCTTCGCC	$(TTGA)_{23}$	NED	7352	339	323-343	4	0.156	0.472	0.373	M1i
Sacl-27	F: GCGTTTGGTTAGCAGTGTGA R: GTGCCCTCCCTACTCCCGT	$(AT)_8$	NED	17361	314	276-312	5	0.000	0.082	0.081	M3i
Sacl-32	F: CTCGCTTGTCTGCTTGCTC R: CTTTCGCCGTCACCTACTACA	$(GCTT)_5$	PET	2655	218	192-200	3	0.031	0.031	0.031	M3i
Sacl-33	F: TTTCCGCTCTGTCTCTCTCCT R: AAATCACAACACAAGGCTGCT	$(CTTT)_6$	FAM	20594	204	182-186	2	0.042	0.118	0.110	M3ii
Sacl-37	F: AACCGCTCTTTGTGTTTGATG R: CTCCTTTCCCTCCCTCTCC	$(GGGA)_6$	NED	12415	149	146-154	3	0.292	0.311	0.271	M1i
Sacl-38	F: TACGATTGCGTGCGTTGT R: ACGAAGAGATTGCGACCAAA	$(TA)_7$	NED	15061	136	122-152	8	0.635	0.846	0.823	M2i
Sacl-41	F: GTGGCGTTAGATGCTGTATGG R: CGTGGACAAAGTAGGAAAGGG	$(GCTTA)_{11}$	VIC	136	340	262-352	6	0.104	0.139	0.134	M3ii
Sacl-47	F: CGACACAATCGCAGTCAG R: GCCCGACACACACTCAAGAC	$(CTTGA)_6$	FAM	11611	165	154-174	3	0.365	0.411	0.353	M1i
Sacl-49	F: GCTCTCCACCTCGCACTAA R: TACCGTCAGCCTCCCAGA	$(AT)_{13}$	NED	7113	333	315-335	11	0.875	0.796	0.765	M2ii
Sacl-54	F: CGACGCTGACCTGACACA R: CCTTCCAATCTCTTCCTTCCA	$(GA)_8$	NED	16526	225	221-227	4	0.323	0.625	0.552	M1ii
Sacl-56	F: GGAGAGCGTCGGATAGACC	$(AT)_{10}$	VIC	1453	132	112-118	2	0.021	0.041	0.040	M3i

	R: GAGAATAGCACAGCAGCGAAC										
SacI-60	F: GTCCAGGTGCGTGCGTTAG	$(TGC)_8$	VIC	7664	200	179-182	2	0.010	0.010	0.010	M3ii
	R: GCGAGCAGTTGAAAGGTGG										
SacI-65	F: ATCTCCCAAACCACACACAAG	$(CCA)_{10}$	PET	6095	344	339-369	6	0.115	0.643	0.580	M2i
	R: CATCATCGTCAAGAACTCGG										
SacI-66	F: TATGTATGTCGGGAGACGGG	$(TTG)_9$	PET	20551	264	202-238	12	0.521	0.835	0.813	M2ii
	R: GGGATTAGCAACTGAAACCA										
SacI-68	F: GGTGGGATTCTTTGGACGA	$(GGT)_8$	FAM	19139	162	155-158	2	0.073	0.501	0.375	M2i
	R: AAATGTGCTTGGGTCGGG										
SacI-74	F: CCTAAAGTTCTACCTGGGCAA	$(CCTT)_6$	PET	19201	279	256-284	7	0.260	0.291	0.279	M1i
	R: TCACAAGGACCACATTCCAAC										
SacI-75	F: CTCGTGTCGTCCCTTCATC	$(TC)_5$	NED	13067	256	236-238	2	0.010	0.010	0.010	M3ii
	R: CTGTCTCCCAGAACTCGCC										
SacI-78	F: GTTGGTCGTGCTTCTAATCGG	$(GCTT)_5$	FAM	17312	291	265-289	3	0.042	0.041	0.041	M3i
	R: GTCCATTTCTTGCTGTCGTG										
SacI-81	F: ACTTTGGCTCGGTCTGCTT	$(CG)_6$	FAM	9492	329	321-337	6	0.469	0.694	0.642	M2ii
	R: CCTCCTCCCTTACCTACCTCC										
SacI-88	F: GAAACGGTGACTGACTGATGAC	$(AC)_5$	PET	1052	189	171-175	3	0.021	0.062	0.059	M3ii
	R: ACAAGGACGAACAGAGAACGA										
SacI-90	F: ATTGTGTTGCTGGATGAGGAC	$(GCAA)_9$	PET	4328	292	274-314	8	0.240	0.577	0.532	M2ii
	R: ACCTTCCGCTCTCTCGCT										
SacI-94	F: TGCCAAATAACACATTCCAGAG	$(CGGTA)_5$	VIC	11651	256	243-298	12	0.573	0.779	0.758	M1ii
	R: TGGTGCGAAGTCACGAAATAG										
SacI-95	F: GGGAAGGAGGAAGAAGGTG	$(GTTGC)_6$	VIC	19139	201	188-203	4	0.479	0.505	0.400	M2i
	R: TAGCGGGAAAGAACGGGTAGT										

Primer sequences exclude the M13(-21) universal sequence tag; F, forward primer (labelled with florescent dye); R, reverse primer; Expected product size, size in base pairs (bp) of the expressed sequence tag (EST) that possessed the microsatellite locus; N_A, number of alleles; H_O, observed heterozygosity; H_T, overall gene diversity; *PIC*, polymorphic information content. 100% amplification was achieved for all loci.

TABLE 3. Pairwise F_{ST} estimates for the six sampling localities.

	SB	NB	PDC	DEN	SWE	NOR
SB	0.000					
NB	0.077	0.000				
PDC	0.179	0.157	0.000			
DEN	0.481	0.517	0.562	0.000		
SWE	0.472	0.509	0.559	0.120	0.000	
NOR	0.358	0.390	0.431	0.291	0.268	0.000

All comparisons are significant

SUPPLEMENTARY TABLE 1. Technical information (primer sequence, and dye), characteristics (repeat motif, expected product size, and observed size range), and polymorphism estimates (N_A, H_O , H_T, *PIC*) for two EST-SSR loci based on amplification from genomic DNA of 76 *Saccharina latissima* sporophytes for Sacl-70 and 72 *S. latissima* sporophytes for Sacl-77.

Locus	Primer sequence (5'-3')	Repeat motif in the EST	Dye	Contig no. on GenBank BioProject PRJNA80101	Expected product size (bp)	Observed size range (bp)	N_A	H_O	H_T	*PIC*
Sacl-70	F: GTTAGGGCAAGGTGTGTGGT R: AGAACCCAGAAGAAGCAGCAA	$(TAC)_8$	M13-VIC	1667	258	258-264	2	0.211	0.317	0.267
Sacl-77	F: GTCACTCAGGCAGCGGTT R: ACCACAAGAGCGTCAACAGAG	$(CT)_6$	M13-VIC	18198	340	236-338	7	0.375	0.730	0.686

Primer sequences exclude the M13(-21) universal sequence tag; F, forward primer (labelled with florescent dye); R, reverse primer; Expected product size, size in base pairs (bp) of the expressed sequence tag (EST) that possessed the microsatellite locus; N_A, number of alleles; H_O, observed heterozygosity; H_T, overall gene diversity; *PIC*, polymorphic information content.

SUPPLEMENTARY TABLE 2. Technical information (primer sequence, and dye), and characteristics (repeat motif, expected product size, and observed product size) for 15 EST-SSR loci that were monomorphic within a sample of 96 *S. latissima* sporophytes.

Locus	Primer sequence (5'-3')	Repeat motif in the EST	Dye	Contig no. on GenBank BioProject PRJNA80101	Expected product size (bp)	Observed product size (bp)
Sacl-16	F: GGAGGAGGAGGAGGTTTAGGT R: CTTATGTCGCCTTGGATGGA	$(CAA)_{10}$	M13-PET	10750	223	223
Sacl-20	F: CGGAAACGGGAGTCTATCTG R: GCTGCTGGAGTTGCTGTTG	$(CAG)_{10}$	M13-PET	2190	187	186
Sacl-22	F: CTCACAGCCAACATCACACAC R: AACGAGAAGAGCAAGAGCCAC	$(CAC)_{9}$	M13-VIC	1442	183	180
Sacl-29	F: GCACTTTCTTCGTCTCGTGTG R: ACCACCACCACCACTTTGTT	$(TTTA)_{5}$	M13-NED	11922	264	264
Sacl-44	F: AGCAAAGAACCGAACAAAGAA R: AGACCGAAGAAGAAATGACGG	$(ACCGA)_{6}$	M13-PET	9226	230	230
Sacl-45	F: AGTAGCGGCAGGGAGGAAC R: GACGCAGAGAGACACAGAACC	$(TGGTC)_{7}$	M13-FAM	17892	203	203
Sacl-46	F: TTTCGGTATCATCGTGTTC R: TCCATCCATCCATCAATCTAA	$(GGGTT)_{8}$	M13-VIC	1566	187	187
Sacl-48	F: AAAGGGACTGGATTTGTCT R: CCTACTTCTCGGCTTGTTGA	$(TATGC)_{5}$	M13-PET	3016	136	136
Sacl-50	F: AGCACCCGTAGGTAGAAGGAA R: GTGGGAGGAGGAAACCATAAC	$(CT)_{8}$	M13-VIC	292	247	243
Sacl-59	F: GGAGGTGTGTGGTGGACTTG R: ACAATGGCAGATGAGGAACT	$(TGA)_{8}$	M13-NED	8272	250	247
Sacl-63	F: TTGTTGTTGGTGTTGCTGTGT R: CTGGGATAGATGGAGATGGG	$(AGC)_{8}$	M13-NED	15110	230	229
Sacl-71	F: CGAAGGAAGTGAGAGTGACGA R: TTCTGAGGAGCAACGACAAG	$(GAG)_{8}$	M13-NED	11027	215	215
Sacl-80	F: AGCCTGCGTCTCAACACC R: CCGTGCTACAGTTTCCTATCC	$(GCAT)_{6}$	M13-PET	3817	203	203
Sacl-86	F: ATGTTGCCGTTGTTTAGGTCA R: TTGATTGGTCTGGTTTGGTTT	$(CATC)_{6}$	M13-VIC	2253	258	254
Sacl-91	F: TGCTTGGATTTATTAGAGTTT R: TCTGCCTTGCCTTGCTTG	$(AGGCC)_{5}$	M13-NED	19844	248	243

Primer sequences exclude the M13(-21) universal sequence tag; F, forward primer (labelled with florescent dye); R, reverse primer; Expected product size, size in base pairs (bp) of the expressed sequence tag (EST) that possessed the microsatellite locus. 100% amplification was achieved for all loci.

SUPPLEMENTARY TABLE 3. Characteristics of 32 polymorphic *S. latissima* microsatellite loci in three French populations: SB (Mousterlin), NB (Roscoff), and PDC (Audresselles).

		SB					NB					PDC				
Locus	*n*	Ar	H_O	H_S	F_{IS}	Null Allele Freq	Ar	H_O	H_S	F_{IS}	Null Allele Freq	Ar	H_O	H_S	F_{IS}	Null Allele Freq
Sacl-03	16	5	0.750	0.754	0.006	-0.011	5	0.813	0.748	-0.086	-0.050	5	0.563	0.600	0.063	0.011
Sacl-06	16	5	0.563	0.569	0.011	-0.008	4	0.563	0.592	0.049	0.006	7	0.250	0.533	0.531*	0.171*
Sacl-08	16	3	0.250	0.231	-0.081	-0.021	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-09	16	4	0.500	0.427	-0.171	-0.059	2	0.188	0.175	-0.071	-0.015	2	0.063	0.063	N/A	-0.002
Sacl-11	16	3	0.250	0.231	-0.081	-0.021	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-13	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-19	16	2	0.250	0.225	-0.111	-0.026	1	N/A	N/A	N/A	N/A	5	0.375	0.550	0.318	0.100
Sacl-21	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-24	16	3	0.125	0.123	-0.017	-0.005	5	0.250	0.238	-0.053	-0.016	8	0.500	0.440	-0.137	-0.051
Sacl-25	16	1	N/A	N/A	N/A	N/A	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-27	16	5	N/A	0.450	1.000*	0.297*	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-32	16	2	0.063	0.063	N/A	-0.002	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-33	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	2	0.250	0.492	0.492	0.149
Sacl-37	16	2	0.188	0.175	-0.071	-0.015	3	0.375	0.331	-0.132	-0.040	1	N/A	N/A	N/A	N/A
Sacl-38	16	6	0.750	0.775	0.032	N/A	7	0.688	0.819	0.160	0.057	4	0.563	0.488	-0.154	-0.060
Sacl-41	16	5	0.188	0.294	0.362	0.073	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-47	16	3	0.125	0.444	0.718*	0.208*	2	0.563	0.413	-0.364	-0.113	2	0.625	0.438	-0.429	-0.137
Sacl-49	16	8	1.000	0.838	-0.194	-0.101	7	0.938	0.769	-0.220	-0.107	3	0.813	0.665	-0.223	-0.100
Sacl-54	16	3	0.688	0.554	-0.241	-0.095	3	0.188	0.454	0.587*	0.171*	3	0.250	0.406	0.385	0.100
Sacl-56	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-60	16	1	N/A	N/A	N/A	N/A	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-65	16	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
Sacl-66	16	5	0.813	0.696	-0.168	-0.080	6	0.688	0.790	0.129	0.042	3	0.063	0.183	0.659	0.095
Sacl-68	16	1	N/A	N/A	N/A	N/A	2	0.125	0.229	0.455	0.077	2	0.063	0.063	N/A	-0.002
Sacl-74	16	4	0.375	0.335	-0.118	-0.037	4	0.313	0.452	0.309	0.085	4	0.250	0.235	-0.062	-0.018
Sacl-75	16	1	N/A	N/A	N/A	N/A	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-78	16	1	N/A	N/A	N/A	N/A	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-81	16	5	0.563	0.602	0.066	0.012	4	0.250	0.342	0.268	0.059	3	0.563	0.533	-0.055	-0.030
Sacl-88	16	2	0.063	0.063	N/A	-0.002	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A
Sacl-90	16	6	0.563	0.515	-0.093	-0.042	4	0.250	0.235	-0.062	-0.018	2	0.250	0.225	-0.111	-0.026
Sacl-94	16	10	0.813	0.854	0.049	0.008	9	0.813	0.835	0.027	-0.002	3	0.313	0.567	0.449	0.148
Sacl-95	16	2	0.563	0.496	-0.135	-0.054	3	0.500	0.535	0.066	0.012	3	0.750	0.540	-0.390	-0.144
Mean (SE)		2.8 (0.4)		0.259 (0.058)	-0.066		2.8 (0.4)		0.243 (0.059)	0.095		2.3 (0.4)		0.204 (0.049)	0.085	

n, number of amplified individuals (*n* is same for three populations); Ar, allelic richness; H_O, observed heterozygosity; H_S, within-population gene diversity; F_{IS} for each locus within each locality estimated using the Nei & Chesser 1983 equation; Null allele frequency estimated using Brookfield I equation (Brookfield 1986); SE, standard error. * indicates significant F_{IS} (p-value below 0.01), or loci showing evidence for null alleles. Underlined loci are the 25 retained for the population genetics analyses. Within population Ar and H_S averages (bottom row) calculated for the 25 underlined loci only. F_{IS} for each population is the multilocus estimate of Wright's inbreeding coefficient (calculated for the 25 underlined loci only).

SUPPLEMENTARY TABLE 3 continued. Characteristics of 32 polymorphic *S. latissima* microsatellite loci in three Scandinavian populations: DEN (Ebeltoft, Denmark), SWE (Tjarno, Sweden), and NOR (Ny-Alesund, Norway), and in all six populations (French and Scandinavian) combined (ALL POPS).

		DEN					SWE					NOR				ALL POPS			
Locus	*n*	Ar	H_O	H_S	F_{IS}	Null Allele Freq	Ar	H_O	H_S	F_{IS}	Null Allele Freq	Ar	H_O	H_S	F_{IS}	Null Allele Freq	Ar (SE)	H_S (SE)	F_{IS}	F_{ST}
Sacl-03	16	3	0.250	0.231	-0.081	-0.021	3	0.125	0.123	-0.017	-0.005	2	0.375	0.388	0.032	N/A	3.833 (0.543)	0.474 (0.109)	-0.011	0.302
Sacl-06	16	3	0.125	0.338	0.630	0.148*	6	0.125	0.531	0.765*	0.251*	1	N/A	N/A	N/A	N/A	4.333 (0.882)	0.427 (0.093)	0.366	0.244
Sacl-08	16	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A	2	0.125	0.121	-0.035	-0.007	1.667 (0.333)	0.069 (0.038)	-0.055	0.032
Sacl-09	16	2	0.938	0.500	-0.875*	-0.293	2	0.813	0.488	-0.667*	-0.223	3	0.688	0.517	-0.331	-0.121	2.500 (0.342)	0.361 (0.079)	-0.470	0.149
Sacl-11	16	3	0.688	0.596	-0.154	-0.068	2	0.500	0.517	0.032	N/A	4	0.250	0.235	-0.062	-0.018	2.333 (0.494)	0.263 (0.103)	-0.069	0.260
Sacl-13	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	3	0.500	0.590	0.152	0.044	1.333 (0.333)	0.098 (0.098)	0.152	0.330
Sacl-19	16	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A	2	0.250	0.225	-0.111	-0.026	2.167 (0.601)	0.177 (0.086)	0.118	0.668
Sacl-21	16	3	0.188	0.179	-0.047	-0.012	2	0.438	0.350	-0.250	-0.071	4	0.438	0.533	0.180	0.050	2.000 (0.516)	0.177 (0.091)	N/A	0.128
Sacl-24	16	2	0.500	0.442	-0.132	-0.049	4	0.625	0.573	-0.091	-0.044	4	0.250	0.552	0.547*	0.181*	4.333 (0.843)	0.394 (0.073)	0.049	0.234
Sacl-25	16	3	0.188	0.179	-0.047	-0.012	2	0.375	0.313	-0.200	-0.054	2	0.313	0.471	0.336	0.096	1.833 (0.307)	0.171 (0.078)	0.085	0.638
Sacl-27	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.667 (0.667)	0.075 (0.075)	1.000	0.085
Sacl-32	16	2	0.063	0.063	N/A	-0.002	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.500 (0.224)	0.031 (0.014)	N/A	-0.007
Sacl-33	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.167 (0.167)	0.082 (0.082)	0.492	0.305
Sacl-37	16	2	0.438	0.350	-0.250	-0.071	2	0.563	0.513	-0.098	-0.043	2	0.188	0.175	-0.071	-0.015	2.000 (0.258)	0.257 (0.073)	-0.134	0.173
Sacl-38	16	3	0.500	0.540	0.073	0.014	2	0.375	0.446	0.159	0.038	5	0.938	0.742	-0.264	-0.123	4.500 (0.764)	0.635 (0.066)	-0.001	0.249
Sacl-41	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	2	0.438	0.417	-0.050	-0.024	1.833 (0.654)	0.118 (0.077)	0.120	0.147
Sacl-47	16	3	0.188	0.179	-0.047	-0.012	3	0.500	0.640	0.218	0.071	2	0.188	0.175	-0.071	-0.015	2.500 (0.224)	0.381 (0.073)	0.044	0.072
Sacl-49	16	4	1.000	0.585	-0.708*	-0.266	3	0.875	0.523	-0.673*	-0.236	4	0.625	0.627	0.003	-0.011	4.833 (0.872)	0.668 (0.048)	-0.311	0.161
Sacl-54	16	2	0.250	0.225	-0.111	-0.026	1	N/A	N/A	N/A	N/A	3	0.563	0.450	-0.250	-0.086	2.500 (0.342)	0.348 (0.082)	0.073	0.443
Sacl-56	16	2	0.125	0.229	0.455	0.077	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.167 (0.167)	0.038 (0.038)	0.455	0.070
Sacl-60	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.167 (0.167)	0.010 (0.010)	N/A	N/A
Sacl-65	16	1	N/A	N/A	N/A	N/A	3	0.375	0.402	0.067	0.010	3	0.250	0.558	0.552*	0.184*	1.833 (0.401)	0.170 (0.100)	0.328	0.735
Sacl-66	16	5	0.625	0.640	0.023	-0.004	3	0.375	0.323	-0.161	-0.046	3	0.563	0.600	0.063	0.011	4.167	0.539	0.033	0.355

																(0.543)	(0.096)		
Sacl-68	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	2	0.250	0.225	-0.111	-0.026	1.500 (0.224)	0.086 (0.046)	0.153	0.828
Sacl-74	16	1	N/A	N/A	N/A	N/A	2	0.125	0.121	-0.035	-0.007	2	0.500	0.442	-0.132	-0.049	2.833 (0.543)	0.264 (0.074)	0.015	0.093
Sacl-75	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.167 (0.167)	0.010 (0.010)	N/A	N/A
Sacl-78	16	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	2	0.188	0.175	-0.071	-0.015	1.333 (0.211)	0.040 (0.029)	-0.053	0.038
Sacl-81	16	3	0.500	0.460	-0.086	-0.036	4	0.563	0.567	0.007	-0.009	2	0.375	0.313	-0.200	-0.054	3.500 (0.428)	0.469 (0.049)	0.002	0.324
Sacl-88	16	2	N/A	0.233	1.000	0.180*	1	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A	1.500 (0.224)	0.060 (0.037)	0.651	0.029
Sacl-90	16	3	0.063	0.596	0.895*	0.319*	5	0.188	0.629	0.702*	0.256*	3	0.125	0.181	0.310	0.042	3.833 (0.601)	0.397 (0.084)	0.396	0.312
Sacl-94	16	8	0.688	0.725	0.052	0.008	8	0.313	0.748	0.582*	0.233*	6	0.500	0.629	0.205	0.066	7.333 (1.022)	0.726 (0.046)	0.211	0.068
Sacl-95	16	3	0.438	0.488	0.103	0.023	3	0.563	0.521	-0.080	-0.038	3	0.063	0.238	0.737	0.132*	2.833 (0.167)	0.469 (0.047)	-0.021	0.070
Mean (SE)		2.3 (0.3)		0.219 (0.048)	-0.049		2.1 (0.3)		0.221 (0.050)	0.004		2.5 (0.3)		0.323 (0.046)	0.063			0.245 (0.042)	0.078 (0.031)	0.259 (0.048)

n, number of amplified individuals (*n* is same for three populations); Ar, allelic richness; H_O, observed heterozygosity; H_S, within-population gene diversity; F_{IS} for each locus within each locality estimated using the Nei & Chesser 1983 equation; Null allele frequency estimated using Brookfield I equation (Brookfield 1986); SE, standard error. * indicates significant F_{IS} (p-value below 0.01), or loci showing evidence for null alleles. For each locus the ALL POPS Ar and H_S are the means of these two parameters over the six sampling localities (SB, NB, PDC, DEN, SWE, NOR), F_{IS} was calculated using equation $F_{IS} = 1 - H_O/ H_S$, and F_{ST} was calculated using equation $F_{ST} = 1 - H_O/ H_T$, where H_O and H_S are averages of these values across the six populations, and H_T is the overall gene diversity (shown in Table 1). Underlined loci are the 25 retained for the population genetics analyses. Within population Ar and H_S averages and the ALL POPS H_S, F_{IS}, and F_{ST} averages (bottom row) calculated for the 25 underlined loci only. Within population F_{IS} is the multilocus estimate of Wright's inbreeding coefficient (calculated for the 25 underlined loci only).

FIGURE 1. Map of the six sampled locations: SB - Mousterlin, France (47°50’27.4”N, 4°02'18.1”W), NB - Roscoff, France (48°43’41.5”N, 4°00’20.6”W), PDC - Audresselles, France (50°49'27.8”N, 1°35'40.9”E), DEN - Ebeltoft, Denmark (56°10'04.5“N, 10°43’51.2”E), SWE - Tjarno, Sweden (58°52'31.9”N, 11°06'12.6”E), NOR - Ny-Alesund, Spitsbergen, Norway (78°55'30“N, 11°55’20”E).

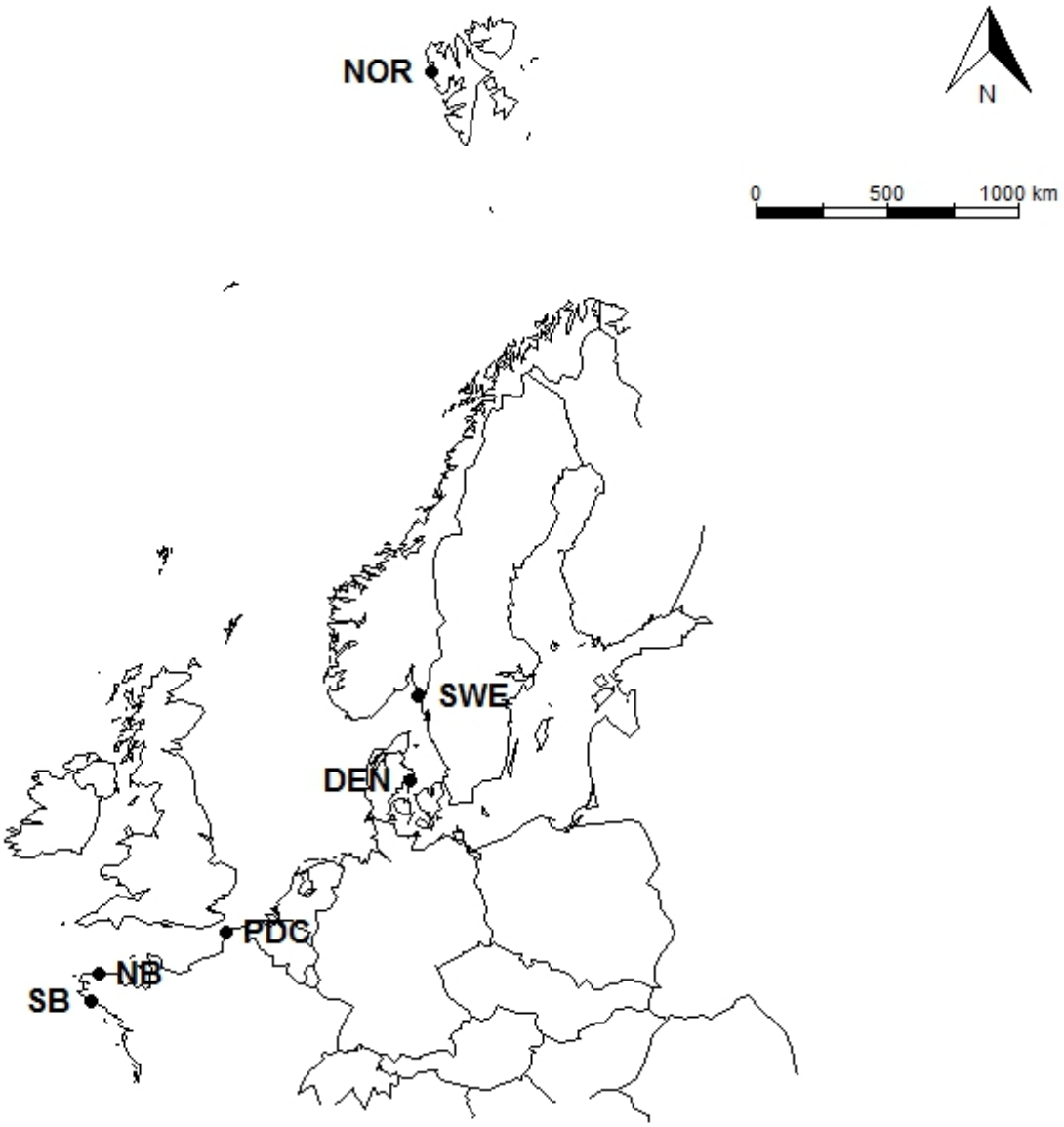


FIGURE 2. Allelic richness (Ar) (filled symbol – left hand side y-axis) and within-population gene diversity (H_S) (open symbol – right hand side y-axis) averaged over 25 EST-SSR loci for each of the six sampling localities (16 individuals per population) ordered by latitude. Error bars indicate standard error.

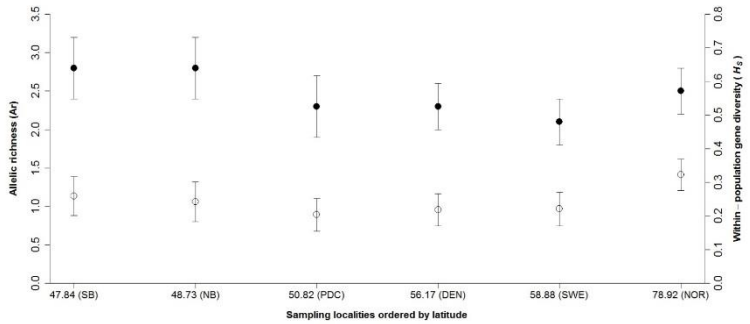


FIGURE 3. Genetic estimates of pairwise differentiation (F_{ST}/(1- F_{ST})) plotted against geographic distance (as the crow flies) for the six sampling localities.

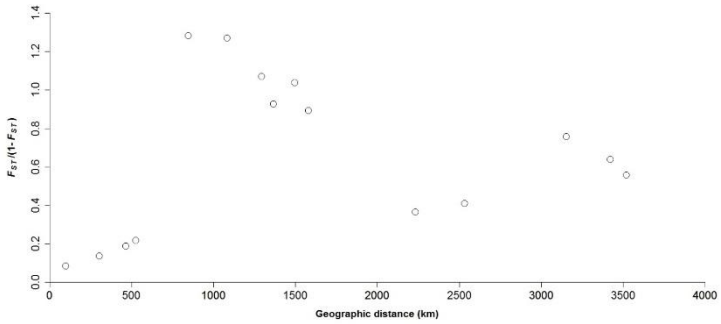
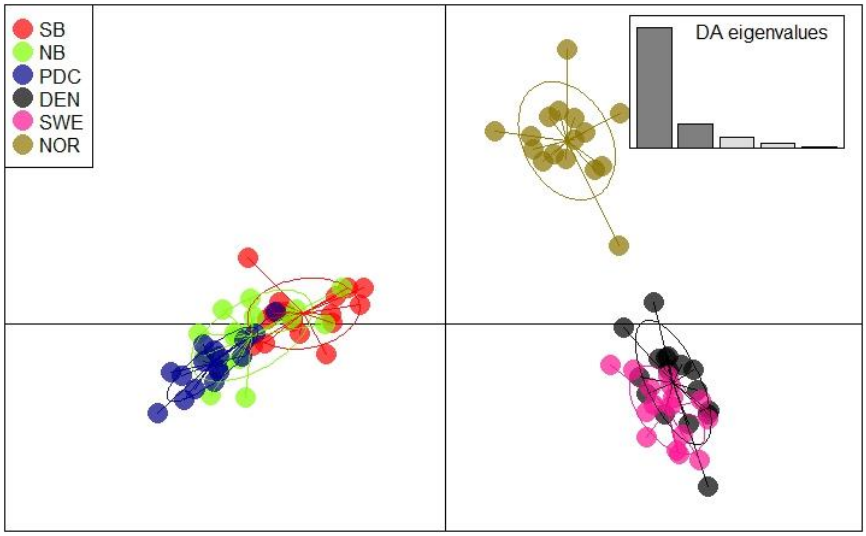
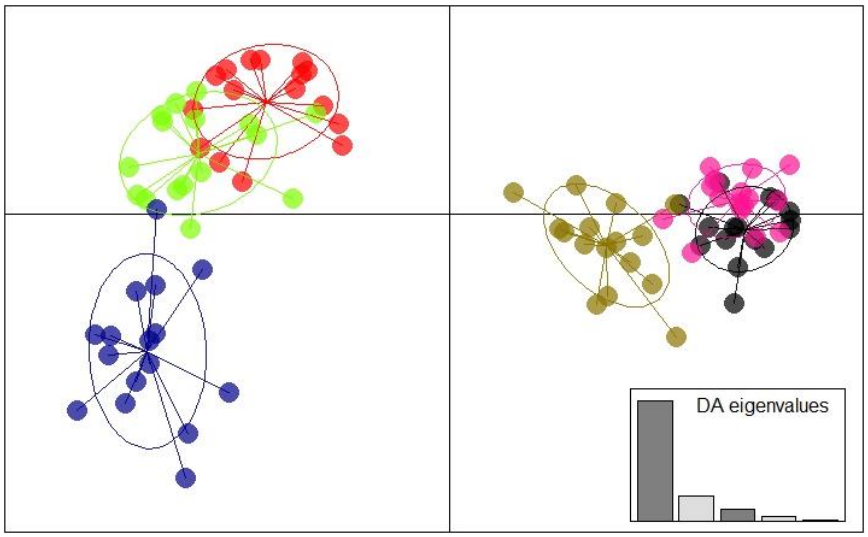


FIGURE 4. Scatterplot of the **A)** first two components (axes 1 and 2), and **B)** of the first and third components (axes 1 and 3), of the discriminant analysis of principal components (DAPC) performed on 96 *S. latissima* sporophytes amplified at 25 microsatellite loci, using sampling location as prior clusters. Sampling groups are shown by different colours and inertia ellipses, and each point corresponds to a single individual. **C)** Membership probability of each *S. latissima* individual to each of the six sampling localities (differentiated by colour). Individuals are represented by vertical bars. Each individual's sampling locality (i.e. population of origin) is specified along the x-axis.

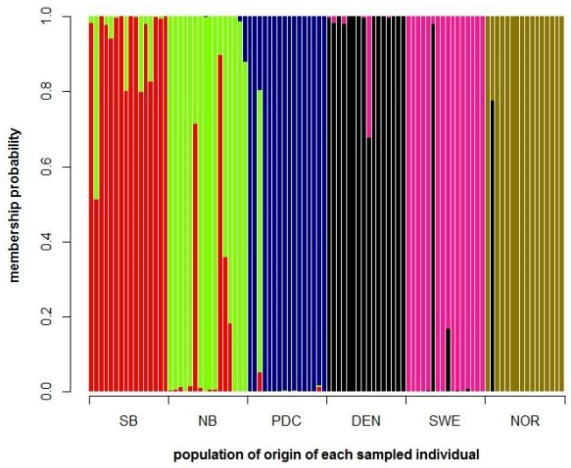
A)



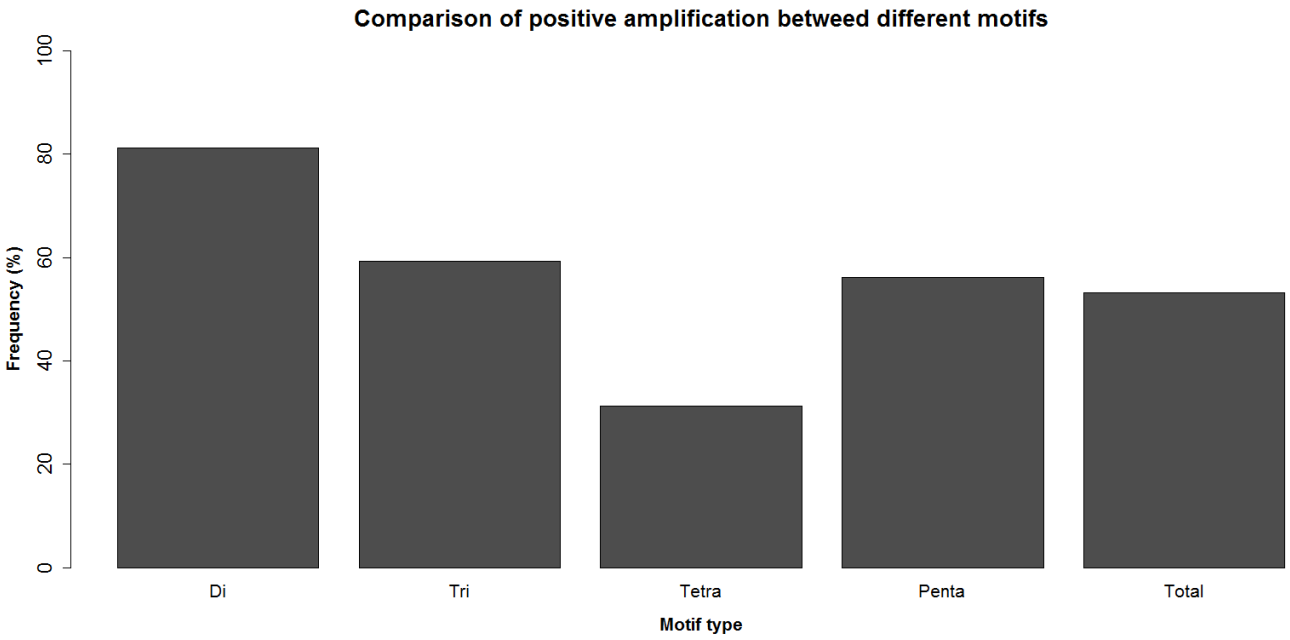
B)



C)



SUPPLEMENTARY FIGURE 1. Comparison of the success of amplification of the EST-SSR loci with different repeat types in the test set of 96 loci.



SUPPLEMENTARY MATERIAL 1. Justification for removal of seven EST-SSR loci from the panel of 32 polymorphic loci listed in Table 2.

Locus Sacl-06 was removed because it exhibited significant null allele frequency within PDC, DEN, and SWE, and significantly positive F_{IS} (i.e. heterozygote deficit) in two of these three populations. Locus Sacl-09 was removed because it exhibited significantly negative F_{IS} (i.e. heterozygote excess) in DEN and SWE and highly negative "All Pops" F_{IS} (in fact the most negative "All Pops" F_{IS} of any the 32 loci – Supplementary Table 3). Additionally, the allelic patterns for this locus in these two populations were odd due to the fact that 15 of the 16 DEN samples and 13 of the 16 SWE samples were heterozygous for the same two alleles. Sacl-27 also exhibited a strange allelic pattern because, despite having five alleles in SB, all 16 SB individuals were homozygous (thus significantly positive F_{IS} and significant null allele frequency – Supplementary Table 3). In the remaining five populations this locus was monomorphic for the same allele. Locus Sacl-47 was removed because within SB it exhibited significantly positive F_{IS} and significant null allele frequency, and within the other five populations it displayed low levels of polymorphism (Supplementary Table 3). For locus Sacl-49, as with Sacl-09, the "All Pops" F_{IS} was highly negative, and its F_{IS} was significantly negative within DEN and SWE. For Sacl-88, there was significant null allele frequency within DEN (no heterozygotes despite this locus possessing two alleles within that population), and this locus was monomorphic or displayed low levels of polymorphism within the other five localities. Lastly, locus Sacl-90 displayed significantly positive F_{IS} and significant null allele frequency within two of the six localities (DEN and SWE), and was thus removed. Significant null allele frequency and/or significantly positive F_{IS} was also detected for several other loci, such as Sacl-24, Sacl-54, or Sacl-65 (Supplementary Table 3). However, as in each of those cases these problems were only observed for a single population, and the loci in question were highly polymorphic and/or showed high levels of F_{ST}, we nonetheless retained those loci for the population genetics analyses. Note that, although the seven loci that were removed were deemed unsuitable for our analyses, they could be of use for other types of studies or

studies on other populations of *S. latissima*, thus they are also reported in this manuscript and were optimized as part of the multiplexes specified in Table 2