Levels and patterns of population genetic diversity in the red seaweed *Chondrus crispus* (Florideophyceae): a direct comparison of single nucleotide polymorphisms and microsatellites

JIM PROVAN*, KEITH GLENDINNING, RUTH KELLY and CHRISTINE A. MAGGS

School of Biological Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK

Single nucleotide polymorphisms (SNPs) are predicted to supersede microsatellites as the marker of choice for population genetic studies in the near future. To date, however, very few studies have directly compared both marker systems in natural populations, particularly in non-model organisms. In the present study, we compared the utility of SNPs and microsatellites for population genetic analysis of the red seaweed *Chondrus crispus* (Florideophyceae). Six SNP loci yielded very different patterns of intrapopulation genetic diversity compared to those obtained using seven moderately (mean 5.2 alleles) polymorphic microsatellite loci, although Bayesian clustering analysis gave largely congruent results between the two marker classes. A weak but significant pattern of isolation-by-distance was observed across scales from a few hundred metres to approximately 200 km using the combined SNP and microsatellite data set of 13 loci. Over larger scales, however, there was little correlation between genetic divergence and geographical distance. Our findings suggest that even a moderate number of SNPs is sufficient to determine patterns of genetic diversity across natural populations, and also highlight the fact that patterns of genetic variation in seaweeds arise through a complex interplay of short- and long-term natural processes, as well as anthropogenic influence. © 2012 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2013, **108**, 251–262.

ADDITIONAL KEYWORDS: algae – Bayesian clustering analysis – gene flow – isolation-by-distance.

INTRODUCTION

In recent years, population genetic studies in seaweeds have seen an increase in the use of microsatellite markers to determine the levels and patterns of genetic diversity and gene flow within and between natural populations. Microsatellites offer many advantages in such studies, including high levels of polymorphism, with multiple alleles observed at many loci, codominant inheritance, robustness, and transferability between studies (Powell, Machray & Provan, 1996). Despite these benefits, there are also some drawbacks in using microsatellites, including homoplasy as a result of the bidirectional mutational processes operating at microsatellite loci, ambiguity

surrounding the exact nature of these mutation models, and a limited number of loci, particularly in seaweeds, where between five and ten microsatellites are usually available for the species in question (but see also Coyer *et al.*, 2008).

Single nucleotide polymorphisms (SNPs) are predicted to become the marker of choice for population genetic studies in the near future (Morin, Luikart & Wayne, 2004; Garvin, Saitoh & Gharrett, 2010; Seeb *et al.*, 2011). They share many of the advantages of microsatellites, such as codominant inheritance and robustness, although they have additional benefits including adherence to a simple infinite sites model of mutation, lack of homoplasy, and potentially thousands of loci available for analysis. Despite these benefits, however, the main drawback of using SNPs in population genetic studies of non-model organisms lies in the lack of data available for marker

*Corresponding author. E-mail: j.provan@qub.ac.uk

ascertainment (Brumfield *et al*., 2003; Seddon *et al*., 2005; Garvin *et al*., 2010). The increasing availability of expressed sequence tag (EST) data for many species means that primers can immediately be developed to amplify and sequence multiple loci to ascertain SNPs. Furthermore, the increased functional constraint on expressed regions of the genome compared to noncoding regions means that, in the absence of EST data for the target species, there is a good chance that primers designed from EST data in a related species will amplify orthologous loci. If EST data are not available for the target species or related taxa, there are alternative ways of generating libraries of random fragments that will yield anonymous sequence data, such as inter-simple sequence repeat cloning (Beatty, Philipp & Provan, 2010; Doonan *et al*., 2012) or the cloning of amplified fragment length polymorphism products (Bensch, Akesson & Irwin, 2002).

In the present study, we compared the utility of EST-derived SNPs and microsatellites for population genetic analysis of the red seaweed *Chondrus crispus* (Florideophyceae). The species is an important component of intertidal and subtidal communities in the northern North Atlantic, as well as being commercially harvested for carrageenans, which are used as thickeners and gelling agents in foodstuffs. Despite this, very little is known about the population genetic structure of *C. crispus*, with only a handful of published studies beng available to date (Cheney & Mathieson, 1979; Donaldson, Chopin & Saunders,

2000; Wang *et al*., 2008; Hu *et al*., 2010; Krueger-Hadfield *et al*., 2011; Provan & Maggs, 2012), the majority of which (excluding Krueger-Hadfield *et al*., 2011) were concerned with large- rather than fine-scale structure. In florideophyte red algae, female gametes remain attached to the female thallus, and there is limited dispersal of male spermatia (Maggs *et al*., 2011). This type of life history would be expected to give rise to high levels of genetic structuring within and between populations (Chen & McLachlan, 1972; Santelices, 1990), although it has also been suggested that movement of detached fertile fronds may facilitate long-distance dispersal in *C. crispus* (Lazo, Greenwell & McLachlan, 1989). The present study aimed: (1) to directly compare the results from SNPs and microsatellites in the same populations because very few studies to date have done so and (2) to assess the levels and patterns of genetic diversity in natural populations of this important species at a variety of spatial scales.

MATERIAL AND METHODS

SAMPLING AND DNA EXTRACTION

Sampling was designed to test intrapopulation and interpopulation genetic variation at a range of spatial scales (Fig. 1, Table 1; see also Supporting information, Table S1). Intrapopulation sampling was carried out to sample individuals within approximately 50 m of each other and with no obvious habitat boundaries

Figure 1. Locations of the sites sampled in the present study.

Table 1. Populations studied

Region	Latitude	Longitude	Population	N
Helen's Bay, Co. Down	54.692°N	5.730°W	HB1	21
			HB2	21
			HB3	22
Bloody Bridge, Co. Down	54.175°N	5.872°W	BB1	20
			BB2	17
			BB3	16
Dublin Bay, Co. Dublin	53.297°N	6.133°W		
Sandy Cove			SC	23
Dun Laoghaire			DL	19
Sea Point			SPT	20
Fanad Head, Co. Donegal	55.276°N	7.634°W	DON	19
Rinville, Co. Galway	53.250°N	8.950°W	GAL	16
Torquay, Devon	50.450°N	3.483°W	TOR	20

N, number of samples analyzed.

separating them. Medium-scale structuring of genetic variation was tested between populations separated by more than 500 m but less than 5 km and was carried out at three locations: Dublin Bay, Co. Dublin, Bloody Bridge, Co. Down, and Helen's Bay, Co. Down. Large-scale structuring of genetic variation was tested between populations separated by more than 50 km from each other and was based on the three locations named above plus single populations from Torquay, Devon, Fanad Head, Co. Donegal, and Rinville, Co. Galway.

Samples were collected by hand, taking care to ensure that only one sample was taken from any individual thallus. Haploid gametophytes were identified using the resorcinol staining test for κ-carrageenan (Garbary & de Wreede, 1988) and total genomic DNA was extracted using the Qiagen Plant DNeasy kit. In total, between 16 and 23 individual gametophytes were analyzed from each population (mean = 19.5; total $N = 234$; Table 1).

SNP ASCERTAINMENT AND ANALYSIS

Primers were designed to amplify ten single copy DNA (scnDNA) regions based on EST data for *C. crispus* from GenBank (Collén *et al.*, 2006) (accession numbers and primers are available on request). These loci were sequenced in an ascertainment set containing a single individual from each population to identify SNPs. To minimize the effects of ascertainment bias on levels of genetic diversity and differentiation (e.g. as a result of selecting the most polymorphic SNPs), the first SNP present in more than one individual was selected for each locus. Where no SNPs were found in more than one individual, the first SNP in the ascertainment panel was used. SNPs were screened in all samples using an allele-specific polymerase chain reaction (AS-PCR). AS-PCR primers were designed *sensu* Provan *et al.* (2008) and are provided in the Supporting information (Table S2). SNP genotyping was carried out *sensu* Doonan *et al.* (2012).

MICROSATELLITE ASCERTAINMENT AND ANALYSIS

Microsatellite markers were developed by searching the *C. crispus* EST sequences for all trinucleotide repeat motifs *sensu* Provan *et al.* (2007). Primer sequences are given in the Supporting information (Table S3). PCR was carried out in a total volume of 10 μL containing 100 ng of genomic DNA, 10 pmol of dye-labelled forward primer (6-FAM or HEX), 1 pmol of tailed forward primer, 10 pmol reverse primer, $1 \times$ PCR reaction buffer, 200 μM each of dNTP, 2.5 mM $MgCl_2$ and 0.25 U of GoTaq Flexi DNA polymerase (Promega). PCR was carried out on a MWG Primus thermal cycler using the parameters: initial denaturation at 94 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 58 °C for 1 min, extension at 72 °C for 2 min and a final extension at 72 °C for 5 min. Genotyping was carried out on an AB3730xl capillary genotyping system (Applied Biosystems). Allele sizes were scored using ROX-500 size standards and were checked by comparison with previously sized control samples.

DATA ANALYSIS

Because both the microsatellite and SNP markers were derived from ESTs, we carried out an analysis to detect whether any of the loci used were potentially under selection, which could skew values of

within-population diversity and between-population differentiation calculated under the assumption of neutrality. We used the BAYESCAN, version 1.0, which implements the Bayesian method described by Foll & Gaggiotti (2008) to identify outlier loci based on population-specific and locus-specific F_{ST}. The software was run for each population-pairwise comparison (66 comparisons) using the default parameters, and loci were classified as potentially under selection where $\log_{10}(BF) > 1$, which corresponds to 'strong' evidence based on the criteria of Jeffreys (1961).

Tests for linkage disequilibrium between pairs of loci in each population were carried out in FSTAT (http://www2.unil.ch/popgen/softwares/fstat.htm). Levels of polymorphism measured as gene diversity (H) averaged over loci were calculated using ARLEQUIN, version 3.5.1.2 (Excoffier & Lischer, 2010). Levels of interpopulation differentiation within regions and differentiation between regions (Table 1) were estimated from allele frequencies using Φ-statistics, which give an analogue of F-statistics (Weir & Cockerham, 1984) calculated within the analysis of molecular variance (AMOVA) framework (Excoffier, Smouse & Quattro, 1992), also using ARLEQUIN. Significance of Φ_{ST} values was estimated using a nonparametric permutation approach with 10 000 permutations. Calculation of diversity statistics and AMOVA was performed for both the SNP and microsatellite data sets separately, as well as for the combined data set. Population-pairwise Φ_{ST} values were also calculated for the SNP and microsatellite data sets, as well as for the combined data set, using ARLEQUIN, and congruence between the matrices based on the SNP and microsatellite data sets was tested using a Mantel test. Finally, a Mantel test based on pairwise F_{ST} values was also carried out to test for isolation-by-distance (IBD; Rousset, 1997) between the Helen's Bay, Bloody Bridge, and Dublin populations based on the combined data set. The ISOLDE test implemented in GENEPOP (Raymond & Rousset, 1995) was used to assess the relationship between genetic distance, measured as $F_{ST}/(1 - F_{ST})$, and geographical distance between population pairs. Distances were measured along coastlines, and 1000 permutations were used for the Mantel test.

To further identify possible spatial patterns of gene flow, BAPS , version 5 (Corander, Waldmann & Sillanpää, 2003) was used to identify clusters of genetically similar populations using a Bayesian approach. Ten replicates were run for all possible values of the maximum number of clusters (K) up to $K = 12$, the number of populations sampled in the study, with a burn-in period of 10 000 iterations followed by 50 000 iterations. Multiple independent runs always gave the same outcome. Again, the analysis was performed for both the SNP and microsatellite data sets sepa-

rately, as well as for the combined data set. A BAPS analysis was also carried out using microsatellite data from the same seven loci for a further 14 populations from across the species entire range previously analyzed by Provan & Maggs (2012). The analysis was run as described above, except that that maximum number of clusters was $K = 26$. Assignment tests were also carried out at the level of the individual in GENECLASS, version 2.0 (Piry *et al*., 2004) using the Bayesian classification method of Rannala & Mountain (1997), coupled with the Monte Carlo resampling method of Paetku *et al*. (2004) to compute assignment probabilities for each individual to each population.

Because of the relatively low number of alleles revealed, particularly for the SNPs, simulations were carried out using POWSIM, version 4.0 (Ryman & Palm, 2006) to determine the power of the markers to detect low levels of population differentiation. Simulations were carried out for an effective population size of $N_e = 1000$ to yield F_{ST} values of 0.01, 0.02, 0.03, 0.04, and 0.05. Although *C. crispus* may have a larger effective population size, this is not relevant to the analysis because N_e only determines the time necessary to reach the target F_{ST}. Thus, the use of larger values of N_e is unjustified because the difference between, for example, $N_e = 1000$ and 10 000 (and higher) is not important at values of F_{ST} as small as those tested in the simulation (N. Ryman, pers. comm.). In all cases, 1000 replicates were run and the power of the analysis was indicated by the proportion of tests that were significant at $P < 0.05$ based on chi-squared tests using the respective allele frequencies at the six SNP and seven microsatellite loci studied.

RESULTS

TESTS FOR LOCI POTENTIALLY UNDER SELECTION

In the 66 population-pairwise tests to identify loci potentially exhibiting the signatures of selection, microsatellite locus CO649963 was identified as possibly being under balancing selection in 17 (approximately 25%) of the tests, with loci CO652249, CO653364, and CO652857 identified as possibly being under divergent selection in separate single tests, and locus CO650794 possibly being under balancing selection in one further test. Thus, we are confident that the markers used in the present study exhibited no consistent effects of selection that could skew values of within-population diversity and between-population differentiation.

LEVELS OF GENETIC DIVERSITY

SNPs were detected in eight of the ten scnDNA loci sequenced. Of these eight, six pairs of AS-PCR

Table 2. Diversity statistics for each population

Population	SNPs (six loci)		Microsatellites (seven loci)		Combined data (13 loci)	
	N_P	H	N_P	H	N_P	H
HB1	4	0.236	5	0.295	9	0.267
HB2	4	0.212	5	0.355	9	0.289
HB3	3	0.124	5	0.348	8	0.245
BB1	3	0.181	5	0.252	8	0.227
BB2	3	0.181	3	0.179	6	0.187
BB3	4	0.241	4	0.213	8	0.220
SC	4	0.342	2	0.153	6	0.246
DL	4	0.287	4	0.258	8	0.273
SPT	4	0.310	3	0.156	7	0.229
DON	6	0.436	3	0.279	9	0.351
TOR	3	0.184	4	0.271	7	0.251
GAL	4	0.262	3	0.234	7	0.237
Mean	3.833	0.250	3.833	0.250	7.667	0.252

N_P, number of polymorphic loci; H, gene diversity; SNP, single nucleotide polymorphism.

primers were successfully developed to assay the SNPs (see Supporting information, Table S2). Analysis of these SNPs using AS-PCR revealed minor allele frequencies of between 0.051 (CO653304) and 0.384 (CO653326). Between two and eleven alleles (mean = 5.142) were detected at the seven microsatellite loci studied (see Supporting information, Table S3). Allele frequencies by locus/population for both marker types are available on request and a list of private alleles (i.e. those found in only a single population) is provided in the Supporting information (Table S4). No evidence of significant linkage disequilibrium was observed between any pairs of the combined set of 13 SNP and microsatellite loci.

Mean diversity values are given in Table 2. Mean values across SNP loci by population ranged from 0.124 (HB3) to 0.436 (DON). The equivalent values for microsatellites ranged from 0.153 (SC) to 0.355 (HB2). Values for the combined microsatellite/SNP data set ranged from 0.187 (BB2) to 0.351 (DON). A Spearman's rank correlation analysis showed no significant correlation between levels of diversity revealed by microsatellites compared to those based on the SNP data ($r_s = -0.371$, $P = 0.236$).

STRUCTURING OF GENETIC DIVERSITY AMONG POPULATIONS

The AMOVA analysis across all populations revealed that 7.77%, 7.01%, and 7.03% of the total variation was partitioned among regions for the SNP, microsatellite, and combined data sets, respectively,
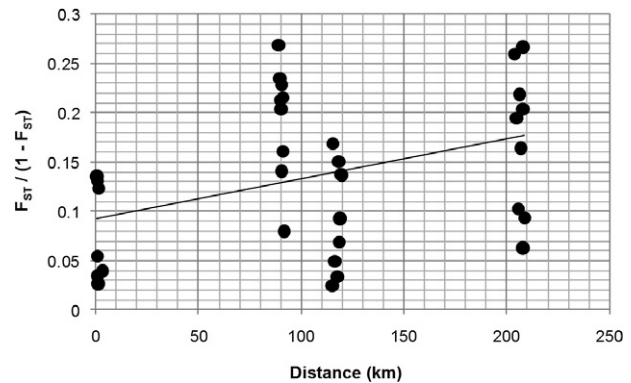

Figure 2. Mantel test for isolation-by-distance (IBD) between the Helen's Bay, Bloody Bridge, and Dublin populations. The significance of the correlation based on Spearman's rank correlation coefficient was $P = 0.046$ based on 1000 permutations).

with 4.89%, 8.06%, and 7.22% partitioned among populations within regions (see Supporting information, Table S5). For individual regions, 7.59% (SNPs), 7.70% (microsatellites), and 7.55% (combined) of the total variation was partitioned among populations at Helen's Bay, with equivalent values of 9.42%, 7.64%, and 8.34% for the Bloody Bridge populations, and 1.20% (NS), 10.68%, and 5.08% for the Dublin populations. Within-region population-pairwise Φ_{ST} values were significantly lower than between-region values for SNP data (Mann–Whitney U-test, $P = 0.044$; see also Supporting information, Table S6) and for the combined data set (Mann–Whitney U-test, $P = 0.004$; see also Supporting information, Table S6) but not for the microsatellite data (Mann–Whitney U-test, $P = 0.056$; see also Supporting information, Table S6). The Mantel test revealed no significant correlation between population-pairwise Φ_{ST} values calculated from the SNP and microsatellite data sets. The Mantel test for IBD using the combined data set indicated a significant correlation between population differentiation and geographical distance (Spearman rank correlation coefficient, $P = 0.046$; Fig. 2).

BAPS analysis based on microsatellite data indicated that $K = 4$ was the most likely number of genetic clusters (Fig. 3). The three Helen's Bay populations formed a single cluster, a second cluster contained the Dun Laoghaire and Torquay populations, the third contained only the Donegal population, and the remaining populations comprised the final cluster. For SNPs, the BAPS analysis indicated $K = 3$, with the only difference from the microsatellite-based clustering being that the Dun Laoghaire and Torquay populations grouped with the majority of the other populations, rather than forming a distinct cluster (Fig. 3). BAPS analysis of the combined data set gave
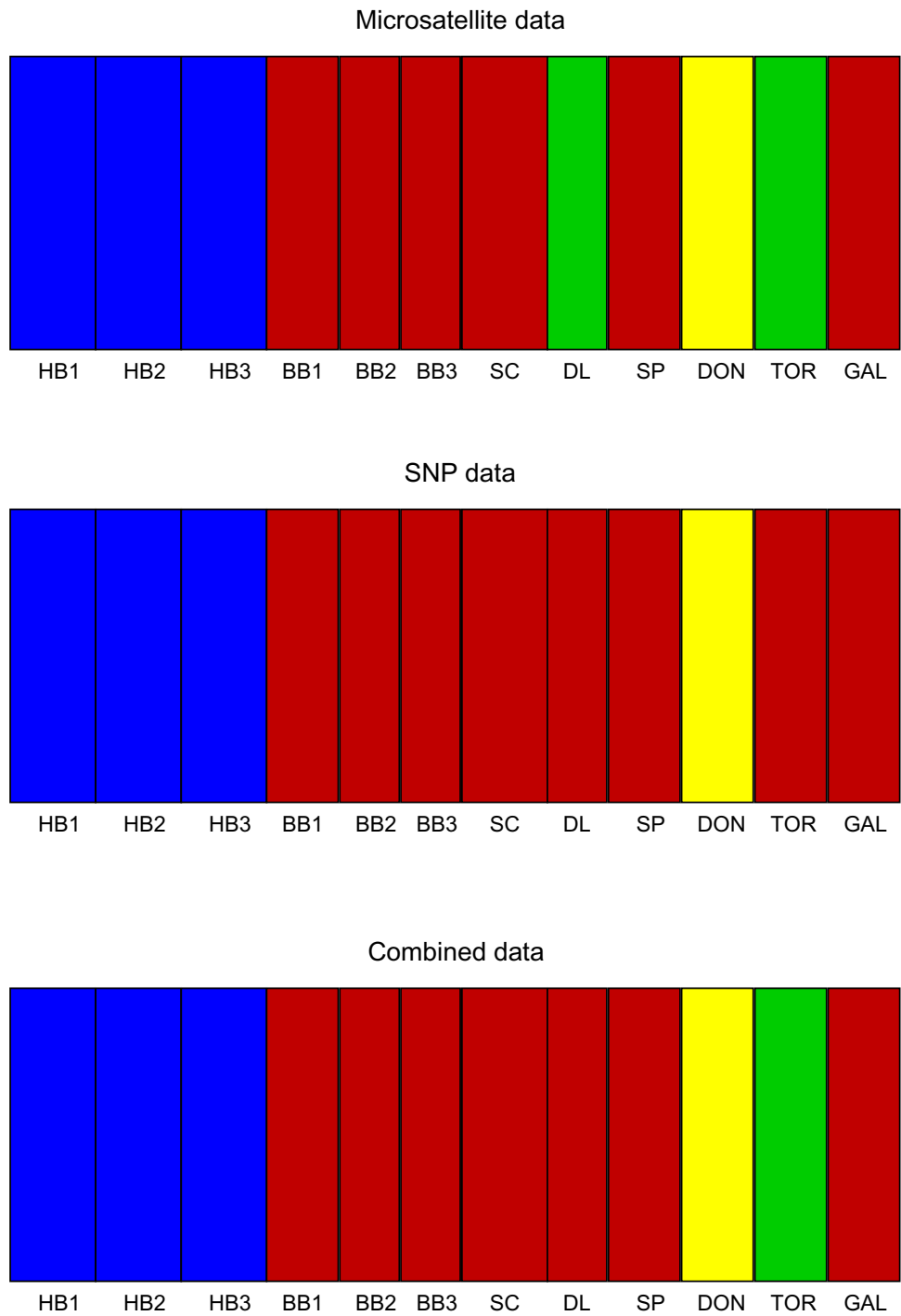
Microsatellite data

HB1   HB2   HB3   BB1   BB2   BB3   SC   DL   SP   DON   TOR   GAL

SNP data

HB1   HB2   HB3   BB1   BB2   BB3   SC   DL   SP   DON   TOR   GAL

Combined data

HB1   HB2   HB3   BB1   BB2   BB3   SC   DL   SP   DON   TOR   GAL

Figure 3. Results of the Bayesian clustering analysis performed using the BAPS software package. Different colours represent the assignment of populations to different genetic clusters; $K = 4$ for the microsatellite and combined data sets; $K = 3$ for the single nucleotide polymorphism (SNP) data set.

a similar result to the microsatellite data ($K = 4$), with the exception of the Dun Laoghaire population, which grouped with the majority of the other populations rather than with the Torquay population (Figs 3, 4).

Finally, the BAPS analysis across the entire range of *C. crispus* based on seven microsatellite loci utilizing the populations from Provan & Maggs (2012) identified the number of genetic clusters as $K = 6$. The
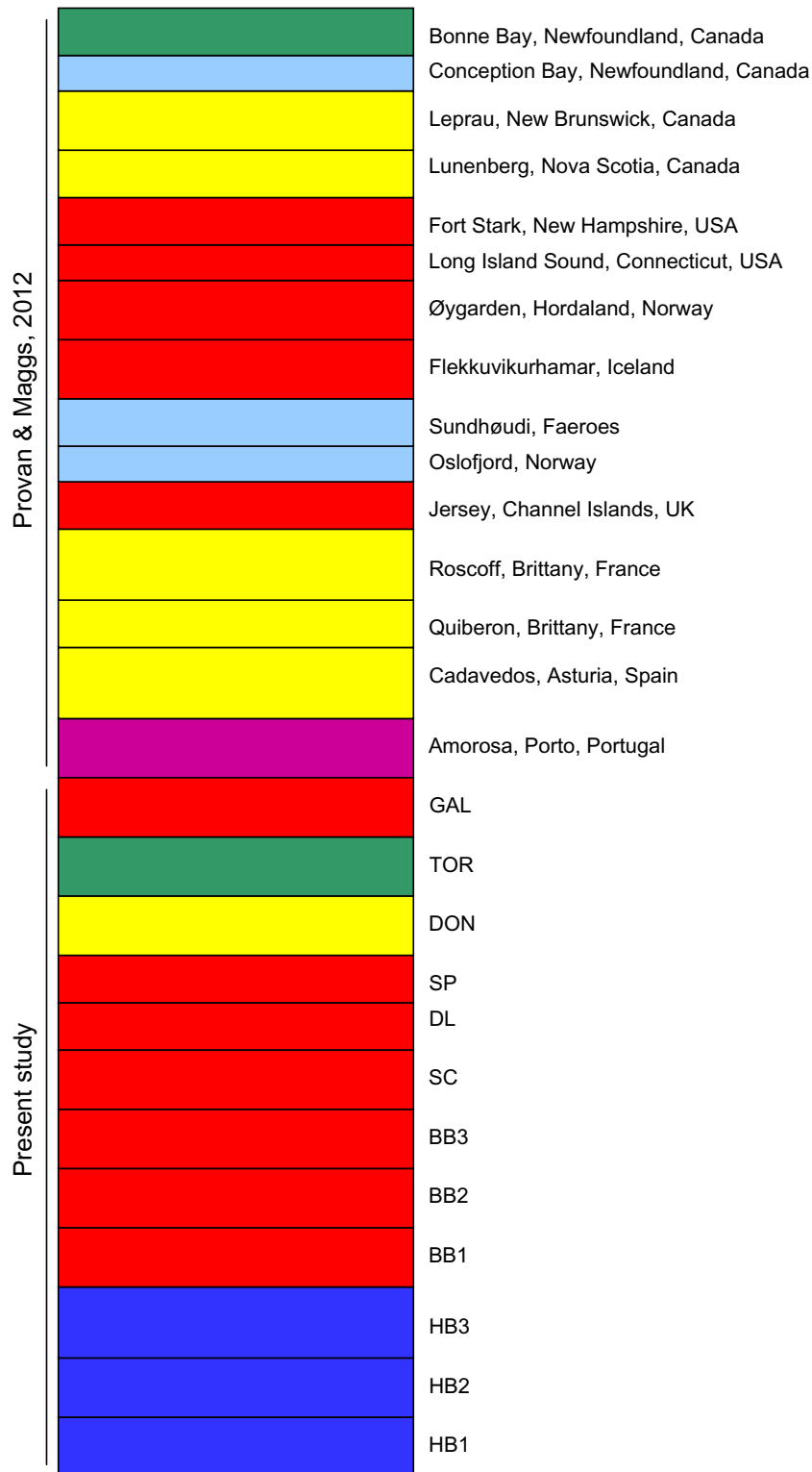
Figure 4. Results of the Bayesian clustering analysis performed using the BAPS software package and including fourteen extra populations from Provan & Maggs (2012). Different colours represent the assignment of populations to one of *K* = 6 different genetic clusters.

Portuguese population was unique, although there was no apparent geographic structuring of the populations in the remaining genetic clusters (Fig. 4). In all the BAPS analyses, the same value for K was recovered across the ten replicate runs. GENECLASS analysis assigned between 0% (BB3) and 84.2% (DON) of individuals to their own population (see Supporting information, Table S7). In total, 37.2% of individuals were assigned to their population of origin.

STATISTICAL POWER TO DETECT LOW LEVELS OF GENETIC DIFFERENTIATION

POWSIM analysis indicated at least a 95% probability (based on the proportion of significant chi-squared tests) of detecting an F_{ST} value of as low as 0.01 using the microsatellite data set (Fig. 5A). For the SNP data set, there was only a 63% probability of detecting $F_{ST} = 0.01$, although a 95% probability of detecting an F_{ST} value of 0.02 (Fig. 5B). For both data sets, the simulations detected higher values of F_{ST} in all cases.

DISCUSSION

COMPARISON OF MICROSATELLITES AND SNPs

Although SNPs have been predicted to supersede microsatellites as the marker of choice for population genetic studies in the near future, only a limited number of studies have been carried out comparing the two types of markers directly in natural populations, particularly for non-model organisms (Seddon *et al*., 2005; Ryynänen *et al*., 2007; Narum *et al*., 2008; Smith & Seeb, 2008; Coates *et al*., 2009; Beacham, McIntosh & Wallace, 2010; Glover *et al*., 2010; Khrustaleva *et al*., 2010; Hauser *et al*., 2011; Hess, Matala & Narum, 2011). Furthermore, only two of these studies (Seddon *et al*., 2005; Coates *et al*., 2009) were carried out on organisms other than commercially important salmonid species. In the present study on populations of the red seaweed *C. crispus*, six SNP loci yielded patterns of intrapopulation genetic diversity very different from those obtained using seven moderately (mean 5.2 alleles) polymorphic microsatellite loci. Conversely, the Bayesian clustering analysis gave largely congruent results between the two marker classes.

Although the mean diversity across populations was identical for both marker classes (0.250 for both SNPs and microsatellites), there were some major discrepancies in diversity values for individual populations. The HB3 population had by far the lowest gene diversity of all populations based on SNPs (0.124) but had the second highest value based on microsatellites (0.348). This was confirmed by the Spearman's rank correlation, which yielded a nega-
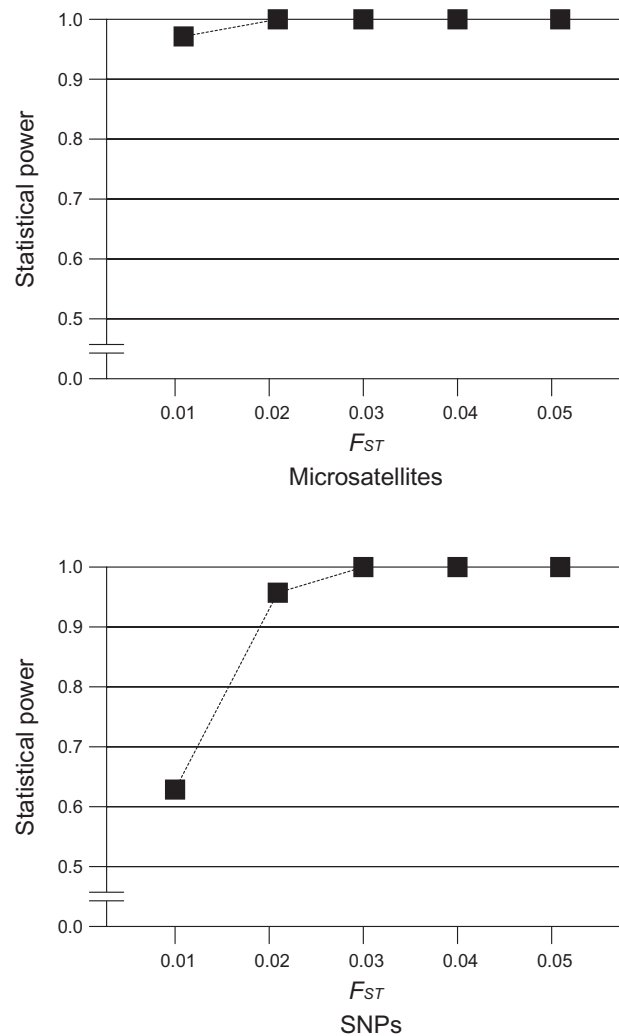

Figure 5. Results of the POWSIM analyses based on microsatellites (top) and single nucleotide polymorphisms (SNPs) (bottom). The *y*-axis represents the power of the markers to successfully recover the value of F_{ST} indicated on the *x*-axis, expressed as the proportion of 1000 simulations (for details, see text).

tive and nonsignificant value for *r*. Simulation studies have suggested that four- to ten-fold more biallelic markers than multiallelic microsatellites are necessary to determine accurately genome-wide levels of diversity (Mariette *et al*., 2002). Thus, it may be that the low number of SNP loci used in the present study was not representative of the true genetic diversity in the populations examined. This may not be the case for all SNP-based studies, however, because a comparison of 7 SNPs and 14 microsatellites in Atlantic salmon revealed a positive correlation between levels of heterozygosity estimated by both classes of markers (Ryynänen *et al*., 2007).

Data from both the six SNPs and the seven microsatellites were sufficient to identify F_{ST} values of as

low as 0.02, based on the results of the simulation studies. This level of resolution for the SNP data set was similar to that obtained using six SNP loci in five populations of the mollusc *Katharina tunicata* (Doonan *et al.*, 2012). By contrast to levels of within-population genetic diversity, patterns of variation between populations were largely congruent between the SNP and microsatellite data sets. This was particularly evident in the Bayesian clustering analysis, where both data sets identified the three Helen's Bay populations and the Donegal population as two distinct genetic entities, separate from the remaining populations (see also below). The AMOVA analyses were also broadly similar, with the exception of Φ_{ST} values of 0.012 and 0.107 for SNPs and microsatellites, respectively, among the Dublin populations. It has been suggested that SNPs should provide more accurate estimates of population divergence than microsatellites (Coates *et al.*, 2009), which will tend to underestimate true levels of differentiation as a result of their high levels of within-population diversity (Hedrick, 1999, 2005; Jost, 2008). This discrepancy was not apparent in the present study, however, most likely as a result of the microsatellite loci employed being only moderately polymorphic (mean = 5.142 alleles per locus).

PATTERNS OF GENETIC DIVERSITY IN *C. CRISPUS*

The limited dispersal potential associated with the various life-history stages of red algae (haploid spermatia and diploid carpospores) has generally resulted in patterns of IBD, where restrictions on gene flow mean that geographically proximal populations are more genetically similar than those separated on greater spatial scales. This has been observed previously in the red algal species *Ahnfeltiopsis pusilla* (Couceiro *et al.*, 2011), *Gelidium canariense* (Bouza *et al.*, 2006), *Grateloupia lanceola* (Maneiro *et al.*, 2011), and *Mazzaella laminarioides* (Faugeron *et al.*, 2001), although not in the invasive *Asparagopsis taxiformis* (Andreakis, Kooistra & Procaccini, 2008). A recent study on *C. crispus* using fine-scale sampling (primarily < 100 m but with replicates across 500 km) also identified potential IBD but highlighted the need to test this further at a range of scales (Krueger-Hadfield *et al.*, 2011). In the present study, analysis of the combined SNP and microsatellite data set also revealed a pattern of IBD in *C. crispus* at intermediate scales of a few hundred metres to approximately 200 km.

Although the range of *C. crispus* also extends well beyond these scales, other factors, rather than simple IBD, tend to shape the structure of genetic variation in seaweeds over such distances. It has been shown that longer-term processes such as postglacial recolo-

nization from refugia are primarily responsible for shaping the range-wide genetic diversity of *C. crispus* and other North Atlantic marine macroalgae across their entire ranges (Provan, Wattier & Maggs, 2005; Hoarau *et al.*, 2007; Olsen *et al.*, 2010; Provan & Maggs, 2012). These patterns may also be modified to a greater or lesser degree by subsequent anthropogenic dispersal. The brown seaweed *Fucus serratus* has been introduced to the Canadian Maritime provinces from western Ireland via rocks used as ballast in shipping (Brawley *et al.*, 2009), and similar processes could explain the genetic clustering of the Donegal population with those from Nova Scotia and New Brunswick, as suggested for other red algae (Maggs *et al.*, 2004). Similarly, although there is a weak pattern of IBD across the populations from the Irish Sea, the Helen's Bay populations were assigned to their own, unique genetic cluster, which could further reflect the importance of nearby Belfast as a major shipping port in the past. The link between populations in the Faeroes, Oslofjord, and Conception Bay, Newfoundland, could have its origins in ballast material in ship voyages associated with the cod-fishing and whaling industries for over three centuries (Ryan, 1986). It is likely that future studies on genetic structuring in seaweeds using multiple, high-resolution, codominant markers such as SNPs and microsatellites, as in the present study, will confirm that patterns of genetic variation arise through a complex interplay of short- and long-term natural processes, as well as anthropogenic influence.

ACKNOWLEDGEMENTS

REFERENCES

Andreakis N, Kooistra WHCF, Procaccini G. 2008. High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Molecular Ecology* **18:** 212–226.

Beacham TD, McIntosh B, Wallace C. 2010. A comparison of stock and individual identification for sockeye salmon (*Oncorhynchus nerka*) in British Columbia provided by microsatellites and single nucleotide polymorphisms. *Canadian Journal of Fisheries and Aquatic Science* **67:** 1274–1290.

Beatty GE, Philipp M, Provan J. 2010. Unidirectional hybridization at a species' range boundary: implications for habitat tracking. *Diversity and Distributions* **16:** 1–9.

Bensch S, Akesson S, Irwin DE. 2002. The use of AFLP to find informative SNP: genetic differentiation across a migratory divide in willow warblers. *Molecular Ecology* 11: 2359–2366.

Bouza N, Caujape-Castells J, Gonzalez-Perez MA, Sosa PA. 2006. Genetic structure of natural populations in the red alga *Gelidium canariense* (Gelidiales, Rhodophyta) investigated by random polymorphic DNA (RAPD) markers. *Journal of Phycology* 42: 304–311.

Brawley SH, Coyer JA, Blakeslee AMH, Hoarau G, Johnson LE, Byers JE, Stam WT, Olsen JL. 2009. Historical invasions of the intertidal zone of Atlantic North America associated with distinctive patterns of trade and emigration. *Proceedings of the National Academy of Sciences of the United States of America* 106: 8239–8244.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution* 18: 249–256.

Chen LCM, McLachlan J. 1972. The life history of *Chondrus crispus* in culture. *Canadian Journal of Botany* 50: 1055–1060.

Cheney DP, Mathieson AC. 1979. Population differentiation in *Chondrus crispus*: preliminary results. *Isozyme Bulletin* 12: 57.

Coates BS, Sumerford DV, Miller NJ, Kim KS, Sappington TW, Siegfried BD, Lewis LC. 2009. Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity* 100: 556–564.

Collén J, Roeder V, Rousvoal S, Collin O, Kloareg B, Boyen C. 2006. An expressed sequence tag analysis of thallus and regenerating protoplasts of *Chondrus crispus* (Gigartinales, Rhodophyceae). *Journal of Phycology* 42: 104–112.

Corander J, Waldmann P, Sillanpää MJ. 2003. Bayesian analysis of genetic differentiation between populations. *Genetics* 163: 367–374.

Couceiro L, Maneiro L, Ruiz JM, Barreiro R. 2011. Multiscale genetic structure of an endangered seaweed *Ahnfeltiopsis pusilla* (Rhodophyta): implications for its conservation. *Journal of Phycology* 47: 259–268.

Coyer JA, Hoarau G, Sjotun K, Olsen JL. 2008. Being abundant is not enough: a decrease in effective population size over eight generations in a Norwegian population of the seaweed *Fucus serratus*. *Biology Letters* 4: 755–757.

Donaldson SL, Chopin T, Saunders GW. 2000. An assessment of the AFLP method for investigating population genetic structure in the red alga *Chondrus crispus* Stackhouse (Gigartinales, Florideophyceae). *Journal of Applied Phycology* 12: 25–35.

Doonan J, Beatty GE, Sigwart JD, Provan J. 2012. Extensive local-scale gene flow and long-term population stability in the intertidal mollusc *Katharina tunicata* (Mollusca: Polyplacophora). *Biological Journal of the Linnean Society* 106: 589–597.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes – application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

Faugeron S, Valero M, Destombe C, Martinez EA, Correa JA. 2001. Hierarchical spatial structure and discriminant analysis of genetic diversity in the red alga *Mazzaella laminarioides* (Gigartinales, Rhodophyta). *Journal of Phycology* 37: 705–716.

Foll M, Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180: 977–995.

Garbary DJ, de Wreede RE. 1988. Life history phases in natural populations of Gigartinaceae (Rhodophyta): quantification using resorcinol. In: Lobban CS, Chapman DJ, Kremer BP, eds. *Experimental phycology: a laboratory manual*. Cambridge: Cambridge University Press, 1988: 174–178.

Garvin MR, Saitoh K, Gharrett AJ. 2010. Application of single nucleotide polymorphisms to non-model species: a technical review. *Molecular Ecology Resources* 10: 915–934.

Glover KA, Hansen MM, Lien S, Als TD, Houheim B, Skaala O. 2010. A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. *BMC Genetics* 11: 2.

Hauser L, Baird M, Hilborn R, Seeb LW, Seeb JE. 2011. An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resources* 11: S150–S161.

Hedrick PW. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53: 313–318.

Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution* 59: 1633–1638.

Hess JE, Matala AP, Narum SR. 2011. Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook salmon in the Columbia River Basin. *Molecular Ecology Resources* 11: S137–S149.

Hoarau G, Coyer JA, Veldsink JH, Stam WT, Olsen JL. 2007. Glacial refugia and recolonization pathways in the brown seaweed *Fucus serratus*. *Molecular Ecology* 16: 3606–3616.

Hu ZM, Guiry MD, Critchley AT, Duan DL. 2010. Phylogeographic patterns indicate transatlantic migration from Europe to North America in the red seaweed *Chondrus crispus* (Gigartinales, Rhodophyta). *Journal of Phycology* 46: 889–900.

Jeffreys H. 1961. *Theory of probability*. Oxford: Clarendon Press.

Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015–4026.

Khrustaleva AM, Volkov AA, Stocklitskaya DS, Mugue NS, Zelenina DA. 2010. Comparative analysis of STR and SNP polymorphism in the populations of sockeye salmon

(*Oncorhynchus nerka*) from Eastern and Western Kamchatka. *Russian Journal of Genetics* **46:** 1362–1372.

Krueger-Hadfield SA, Collén J, Daguin-Thiebaut C, Valero M. 2011. Genetic population structure and mating system in *Chondrus crispus* (Rhodophyta). *Journal of Phycology* **47:** 440–450.

Lazo ML, Greenwell M, McLachlan J. 1989. Population structure of *Chondrus crispus* Stackhouse (Gigartinaceae, Rhodophyta) along the coast of Prince Edward Island, Canada: distribution of gametic and sporophytic fronds. *Journal of Experimental Marine Biology and Ecology* **126:** 45–58.

Maggs CA, Douglas SE, Fenety J, Bird CJ. 2004. A molecular and morphological analysis of the *Gymnogongrus devoniensis* (Rhodophyta) complex in the North Atlantic. *Journal of Phycology* **28:** 214–232.

Maggs CA, Fletcher HL, Fewer D, Loade L, Mineur F, Johnson MP. 2011. Speciation in red algae: members of the Ceramiales as model organisms. *Integrative and Comparative Biology* **51:** 492–504.

Maneiro I, Couceiro L, Barbara I, Cremades J, Ruiz JM, Barreiro R. 2011. Low genetic variation and isolation of northern peripheral populations of a red seaweed (*Grateloupia lanceola*). *Aquatic Conservation – Marine and Freshwater Ecosystems* **21:** 590–600.

Mariette A, Le Corre V, Austerlitz F, Kremer A. 2002. Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology* **11:** 1145–1156.

Morin PA, Luikart G, Wayne RK. 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology and Evoution* **19:** 209–216.

Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, Dekoning J, Elz A, Guthrie CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF, Garza JC. 2008. Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. *Molecular Ecology* **17:** 3464–3477.

Olsen JL, Zechman FW, Hoarau G, Coyer JA, Stam WT, Valero M, Aberg P. 2010. The phylogeographic architecture of the fucoid seaweed *Ascophyllum nodosum*: an intertidal 'marine tree' and survivor of more than one glacial-interglacial cycle. *Journal of Biogeography* **37:** 842–856.

Paetku D, Slade R, Burden M, Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13:** 55–65.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A. 2004. GENECLASS2: a software for genetic assignment and first generation migrant detection. *Journal of Heredity* **95:** 536–539.

Powell W, Machray GC, Provan J. 1996. Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* **1:** 215–222.

Provan J, Beatty GE, Hunter AM, McDonald RA, McLaughlin E, Preston SJ, Wilson S. 2008. Restricted gene flow in fragmented populations of a wind-pollinated tree. *Conservation Genetics* **9:** 1521–1532.

Provan J, Beatty GE, Maggs CA, Savidge G. 2007. Expressed sequence tag-derived microsatellites for the cool-water marine copepod *Calanus finmarchicus*. *Molecular Ecology Notes* **7:** 1369–1371.

Provan J, Maggs CA. 2012. Unique genetic variation at a species' rear edge is under threat from global climate change. *Proceedings of the Royal Society of London Series B, Biological Sciences* **279:** 39–47.

Provan J, Wattier RA, Maggs CA. 2005. Phylogeogaphic analysis of the red seaweed *Palmaria palmata* reveals a Pleistocene marine refugium in the English Channel. *Molecular Ecology* **14:** 793–804.

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 9197–9201.

Raymond M, Rousset F. 1995. GENEPOP (Version 1.2): population genetic software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145:** 1219–1228.

Ryan S. 1986. *Fish out of water: the Newfoundland saltfish trade 1814–1914*. Newfoundland: Breakwater Books, St John's.

Ryman N, Palm S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes* **6:** 600–602.

Ryynänen HJ, Tonteri A, Vasemägi A, Primmer CR. 2007. A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *Journal of Heredity* **98:** 692–704.

Santelices B. 1990. Patterns of reproduction, dispersal and recruitment in seaweeds. *Oceanography and Marine Biology* **28:** 177–276.

Seddon JM, Parker HG, Ostrander EA, Ellegren H. 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology* **14:** 503–511.

Seeb JE, Carvalho G, Hauser L, Naish K, Roberst S, Seeb LW. 2011. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources* **11:** S1–S8.

Smith CT, Seeb LW. 2008. Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single nucleotide polymorphism (SNP) baselines for chum salmon. *Transactions of the American Fisheries Society* **137:** 751–762.

Wang XL, Zhao FJ, Hu ZM, Critchley AT, Morrell SL, Duan DL. 2008. Inter-simple sequence repeat (ISSR) analysis of genetic variation of *Chondrus crispus* populations from North Atlantic. *Aquatic Botany* **88:** 154–159.

Weir BS Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Inter-population distances (km) between Helen's Bay, Bloody Bridge and Dublin Bay populations. For population codes, see Table 1.

Table S2. Allele-specific primers used to genotype six single nucleotide polymorphisms.

Table S3. Microsatellite primers used in the present study. *N*, number of alleles.

Table S4. Private alleles detected.

Table S5a. Analysis of molecular variance (AMOVA) based on six single nucleotide polymorphism loci.

Table S5b. Analysis of molecular variance (AMOVA) based on seven microsatellite loci.

Table S5c. Analysis of molecular variance (AMOVA) based on combined microsatellite and single nucleotide polymorphism data (13 loci).

Table S6a. Matrix of population-pairwise Φ_{ST} values based on six single nucleotide polymorphism loci (below diagonal) and seven microsatellite loci (above diagonal). Values in italics are not significantly different from zero.

Table S6b. Matrix of population-pairwise Φ_{ST} values based on combined single nucleotide polymorphism and microsatellite data (13 loci). Values in italics are not significantly different from zero.

Table S7. Results of assignment tests. Shaded cells represent populations from the same region.