Genomic signatures of clonality in the deep water kelp *Laminaria rodriguezii*

Lauric REYNES[1], Thierry Thibaut[1], Stephane Mauger[2], Aurélie Blanfuné[1], Florian Holon[3], Corinne Cruaud[4], Arnaud Couloux[4], Myriam Valero[5], and Didier Aurelle[1]

[1]Mediterranean Institute of Oceanography
[2]Station Biologique de Roscoff
[3]Andromède Oceanologie
[4]CEA Genoscope centre national de séquençage
[5]Centre National de la Recherche Scientifique

August 4, 2020

Abstract

Understanding the evolution of sexual vs asexual reproduction and their consequences in population genetics is a central tenet of evolutionary biology. Nevertheless, until now, it has proved unachievable to assess partially clonal reproduction when the rate of clonality is less than 95%, especially without the use of genome-wide data or temporal samples. Here, we investigate the genomic signatures of partial clonality in the deep water kelp *Laminaria rodriguezii*, known to reproduce by both sexual and asexual means. The results of these investigations have been interpreted by comparison with the sexually reproducing congeneric species *Laminaria digitata*. Genome-wide variation was assessed by dd-RAD sequencing using 4077 SNPs in *L. rodriguezii* and 7364 SNPs in *L. digitata*. As predicted for partially clonal populations, the distribution of FIS within populations of L. rodriguezii was centered in negative values, with heterozygote excess at most loci along the genomes. This finding is the opposite of what we reported within sexual populations of *L. digitata*, characterized by a generalized deficit in heterozygotes. Furthermore, two distinct distribution patterns of FIS were reported among populations of *L. rodriguezii*, consistent with the results predicted by the theoretical model for different levels of clonality. These findings highlight that the investment in clonal growth could differ among populations of *L. rodriguezii*, confirming that the full distribution of FIS is a promising feature to take into account for the study of asexuality in natural populations. We discuss the implications of these results for the conservation of the rare deep water kelp *L. rodriguezii*.

Partial asexuality, mating system, heterozygote excess, population genomics, kelp forest

Understanding the evolution of sexual vs asexual reproduction and their consequences in population genetics is a central tenet of evolutionary biology. Nevertheless, until now, it has proved unachievable to assess partially clonal reproduction when the rate of clonality is less than 95%, especially without the use of genome-wide data or temporal samples. Here, we investigate the genomic signatures of partial clonality in the deep water kelp *Laminaria rodriguezii* , known to reproduce by both sexual and asexual means. The results of these investigations have been interpreted by comparison with the sexually reproducing congeneric species *Laminaria digitata* . Genome-wide variation was assessed by dd-RAD sequencing using 4077 SNPs in *L. rodriguezii* and 7364 SNPs in *L. digitata* . As predicted for partially clonal populations, the distribution of F_{IS}within populations of *L. rodriguezii* was centered in negative values, with heterozygote excess at most loci along the genomes. This finding is the opposite of what we reported within sexual populations of*L. digitata* , characterized by a generalized deficit in heterozygotes. Furthermore, two distinct distribution patterns of F_{IS} were reported among populations of *L. rodriguezii* , consistent with the results predicted by

the theoretical model for different levels of clonality. These findings highlight that the investment in clonal growth could differ among populations of*L. rodriguezii* , confirming that the full distribution of F_{IS} is a promising feature to take into account for the study of asexuality in natural populations. We discuss the implications of these results for the conservation of the rare deep water kelp*L. rodriguezii* .

INTRODUCTION

Partial clonality, defined as the capacity of species to reproduce both sexually and asexually, occurs in a wide variety of organisms across the tree of life, and is particularly widespread in fungi, plants, rotifers, cladocerans, and insects (Halkett et al., 2005). This mating strategy has attracted considerable attention among evolutionary biologists due to the potential effects of asexuality on the evolutionary trajectory of species, since reproductive modes have major impacts on the level of genome-wide genetic diversity and its distribution within and among populations (Halkett et al., 2005; Duminil et al., 2007). Moreover, reproductive modes profoundly affect other biological traits that play important roles in colonization processes and the spread of populations, such as population growth and dispersal (Kettenring & Mock, 2012; González de León et al., 2016), and in buffering and resilience after catastrophic events (Becheler et al., 2020).

Mathematical models predict that higher rates of asexual events will increase heterozygosity (Marshall & Weir, 1979; Balloux et al., 2003; De Meeus et al., 2006), which may ultimately lead to higher levels of observed heterozygosity (Ho) than those expected (He) under panmixia. This deviation has been observed empirically through negative values of the inbreeding coefficient F_{IS}, reported in various partially clonal populations (in animals Adjeroud et al., 2014; Halkett et al., 2005; angiosperms Alberto et al., 2002; Alberto et al., 2005; Arnaud-Haond et al., 2007, and red and brown macroalgae Guillemin et al., 2008; Ardehed et al., 2015; Coleman & Wernberg,2018, Pardo et al., 2019). Without sex, alleles do not segregate independently, which implies that heterozygosity can be preserved rather than reduced along generations (Judson & Normark, 1996). In addition, if asexuality is stable over generations, the accumulation of mutations in different alleles will increase sequence differences among alleles, which corresponds to the so-called "Meselson effect" (Birky 1996; Judson & Normark, 1996; Balloux et al., 2003). This effect has been empirically demonstrated in the bdelloid rotifers, that have gone without sex for more than 80–100 million years (Welch & Meselson, 2000; Butlin, 2002), and more recently in ancient asexual *Timema* stick-insects (Schwander et al., 2011). However, some taxa (e.g. *Daphnia* Omilia et al., 2006; *Oribatida* Schaefer et al., 2006) fail to exhibit high levels of allelic divergence despite the absence of recombination over a long generation time, implying that other homogenizing mechanisms (e.g. strong selection, automixis) may shape the genetic composition of asexual organisms. More generally, apart from the rate of asexual reproduction, the genetic diversity of partially clonal populations is evidently shaped by the interaction of reproductive mode with genetic drift, gene flow, and mutation (Stoeckel & Masson, 2014; Reichel et al., 2016; Stoeckel et al., 2019). Indeed, mathematical models have predicted that such evolutionary forces can lead to specific patterns of distribution of F_{IS} over the genome, characterized by a high occurrence of extreme values, mainly negative (Stoeckel & Masson, 2014; Reichel et al., 2016). These studies reported, among other points, that increasing the rates of asexuality increased the variance of the distribution of F_{IS} compared to that obtained from fully sexual populations (Stoeckel & Masson, 2014).

Nevertheless, until now, most empirical studies of the genetic diversity and composition of partially clonal populations have relied on a low number of genetic markers such as from 5 to 8 microsatellites (Arnaud-Haond et al., 2007; Guillemin et al., 2008; Adjeroud et al., 2014; Coleman & Wernberg, 2018; Pardo et al., 2019). With the development of thousands of SNP loci, it becomes possible to compare empirical data with the aforementioned theoretical expectations on the distribution of F_{IS}, and to investigate how heterozygosity and departures from panmixia are distributed along the genome for different inferred levels of clonality. Apart from an interest with regard to inferences on clonality rates with high-throughput molecular markers, such results are important to better understand the impact of clonality on the intra-individual genomic diversity. The development of reduced representation sequencing methods (see for review Davey et al., 2011) now allows such investigations for non-model species. Among these methods, Restriction Site Associated DNA sequencing (RAD sequencing) seems to be promising to more precisely delineate Multi-locus lineages

(MLLs), where microsatellite loci have failed to detect any genetic variation, as in the case of the selfing species *Sargassum muticum* in the introduction range (Le Cam et al., 2019). This study illustrates the fact that genotypic richness (R) in partially clonal populations could be underestimated if clonal assessment is performed with a reduced number of molecular markers (Arnaud-Haond et al., 2005).

Here, we used a double-digest Restriction site Associated DNA (dd-RAD) sequencing methodology (Peterson et al., 2012) to study the genome-wide variation of the kelp *Laminaria rodriguezii* . *Laminaria rodriguezii* is a deep water kelp endemic to the Mediterranean Sea, mainly reported from depths >70 m, rarely found in shallower waters on seamounts or in upwelling systems (Ballesteros, 2006) with a maximum depth recorded at 260 m in the Adriatic Sea (Ercegović, 1960). Because of its depth, little is known regarding its biology and ecology, but sporophytes longer than 2m form canopies that support complex food webs on coralligenous assemblages (Ballesteros, 2006). The rare populations are exposed to various levels of disturbance, such as direct destruction by fishing gear and a decline in water transparency, that has led to *L. rodriguezii* being classified as endangered in the Mediterranean (Zuljevic et al., 2016). Besides sexual reproduction involving the alternation of microscopic dioecious haploid gametophytes, with large diploid sporophytes, only two*Laminaria* species have been described as reproducing asexually by stoloniferous or rhizoidal growth: *L. rodriguezii* (Huvé, 1955) and the northeast Pacific *L. sinclairii* (Demes & Graham, 2011).*Laminaria rodriguezii* is particularly interesting from an evolutionary perspective because, in addition to the effect of partial clonality, these deep water populations are scarce and highly fragmented (Araújo et al., 2016). We therefore hypothesized that partially clonal reproduction, but also genetic drift, had strongly influenced the current patterns of genetic variation in *L. rodriguezii* . To put these results into perspective, and to better understand the consequences of clonal reproduction, we compared the genome-wide diversity of *L. rodriguezii* with fully sexual populations from a congeneric Atlantic species, *Laminaria digitata* . The aims of the study were (i) to test the extent of clonal reproduction in populations of *L. rodriguezii* by delimiting Multi-Locus Lineages (MLLs), (ii) to analyze the distribution of F_{IS} among loci in the genome of *L. rodriguezii* , in the light of theoretical expectations and results obtained with the same methods on the congeneric species without clonal reproduction. The implications of the inferred genetic diversity and genetic structure for the management and conservation of this rare deep water kelp *L. rodriguezii* are discussed.

MATERIAL AND METHODS

Sampling, DNA extraction

Individuals of *Laminaria rodriguezii* were sampled from four geographical Mediterranean localities, three in Eastern Provence (Banc Magaud_1, Banc Magaud_2, Cap Camarat) and one in the Southern Corsica (Bonifacio), between 65 to 76 meters depth (Figure S1). Sampling was done by scuba divers between June and August 2018. Sampling for DNA collection was performed by collecting a small piece of tissue from the blade of sporophytes. More precisely, DNA samples were collected from sporophytes not connected to each other by a stolon, and with a minimum distance of two meters between sampled blades. The aim of this strategy was to promote the sampling of distinct genets. For one of the two Banc du Magaud sites, namely Banc Magaud_2, we used a distinct sampling design: the spatial distance between sampled sporophytes was recorded along a regular transect of 30 meters to study the distribution of distance between ramets for a given genet. Regarding *L. digitata* , we used samples from five northeastern Atlantic populations of *L. digitata* collected in 2018 (Table S1, Supporting Information). Genomic DNA was extracted using the Nucleospin(r) 96 plant kit (Macherey-Nagel, Duren, Germany), according to the manufacturer's protocol.

ddRAD-sequencing

Two double-digest RAD-sequencing libraries (ddRAD-seq) with 49 individuals of *L. rodriguezii* and 130 individuals of *L. digitata* were prepared according to Peterson et al., (2012). To minimize bias in library preparation, all DNA samples were randomized across libraries, and two and five technical replicates of *L. rodriguezii* and *L. digitata* , respectively, were included. DNA concentrations were estimated with PicoGreen (Invitrogen, Carlsbad, CA, USA), and 100 ng of genomic DNA was double digested using PstI and HhaI (NEB) for 12 hours at 37degC. The purification of digested DNA was conducted by the NucleoMag NGS

clean-up and Size Select kit (Macherey-Nagel, Duren, Germany) according to the manufacturer's protocol. When these steps had been completed, the ligation of sequencing adapters to restriction digested DNA was performed as follows: P1 adapters (overhang PstI), including unique barcodes (6-12 bp) and P2 universal adaptor (overhang Hhal) were ligated to DNA using 0.5 µl (400 U/µL) of T4 DNA ligase (NEB), 6 µl of 10X T4 ligase buffer (NEB) and incubated for 12 hours at room temperature. Before PCR amplification, an additional purification step was done with the NucleoMag NGS clean-up and Size Select kit. Amplification by PCR was restricted to 15 cycles and performed with the Q5 hot Start High-Fidelity DNA polymerase kit (NEB). Then, samples were multiplexed, and automated size selected for a range between 400 and 900 bp using a Pippin Prep (Sage Science, Beverly, MA, USA). Finally, the libraries were sequenced with the paired-end method (2×150 bp) on an Illumina Hiseq 4000 platform (Génome Québec Innovation Centre, McGill Univ., Montreal, Canada).

Genotyping and SNP filtering

Raw quality reads were checked using FastQC v.0.11.7 (Andrews, 2010) and trimmed to 137 bp after the removal of adapters sequencing by Trimmomatic (Bolger et al., 2014). Then, paired-reads were demultiplexed by both index and barcode using Stacks's process_radtags (Catchen et al., 2011). Thanks to the phylogenetic proximity between *L. rodriguezii* and*L. digitata*(Žuljević et al., 2016), we were able to use the draft genome of *L. digitata* (unpublished data, Mark Cock) to align our reads with the BWA-mem algorithm in BWA 0.7.17 (Li & Durbin, 2009). Aligned reads were then assembled into loci and finally, we called single nucleotide polymorphisms (SNPs) with the Stacks v2.4 pipeline (Rochette et al., 2019). Aligned reads of mapping quality below 30 were excluded to minimize errors resulting from the erroneous assembly of paralogous loci as a single locus. To avoid a strong linkage between SNPs, we kept only the first SNP of each RAD locus. Further filtering was performed on the basis of minor allele frequency, missing data rates for both individuals and locus, using VCFtools v0.1.17 (Danecek et al., 2011) (Table 1). Then, we tested for the presence of selected loci with pcadapt v3.5 (Luu et al., 2017) to conserve putatively neutral loci. Pcadapt starts to identify population structure by PCA, then it estimates to what extent a SNP is related to the first K principal components to identify outlier loci (Luu et al., 2017). Finally, the bias due to potential genotyping errors was estimated from genotypes of sequencing replicate (Mastretta-Yanes et al., 2015). These estimations were performed twice for each species dataset: before and after filtering in VCFtools (Table S2, Supporting Information).

Clonal assessment in *Laminaria rodriguezii*

individuals belonging to the same clonal lineage were expected to share the same MLG (Multi Locus Genotype), which means that their genotypes appear almost identical over dozens of loci. As expected in large SNP datasets, such as those generated by RAD-sequencing, all samples corresponded to distinct MLGs. Because of somatic mutations and sequencing errors, we do not expect to get two MLGs strictly identical over thousands of loci even if they belong to the same MLL (Multi Locus Lineage). Consequently, a threshold distance has been defined beyond which two MLGs could be considered as distinct MLLs. Considering the maximum genetic distance observed between technical replicates (d = 0.022 for *L. rodriguezii* , see Results), we chose this distance as the threshold to separate distinct MLLs. For this purpose, we used the UPGMA clustering, implemented in Poppr v2.8.2 R packages (Kamvar et al., 2014; Kamvar et al., 2015). This threshold distance was also compared to the distribution of pairwise genetic distances among individuals (Figure 1A). The same analysis was repeated on *Laminaria digitata* to determine whether individual genetic distances can be below the threshold of technical replicates (greatest distance d = 0.020 for*L. digitata*) over sexual reproduction (Figure 1B). The genetic distance matrix was additionally used for the construction of networks with the neighbor-net method implemented in SplitsTree, which was used to visually identify samples with high proximity (Huson, 1998). Furthermore, the distinction of MLLs was done with an independent likelihood clustering (LC) method implemented in COLONY (Wang, 2004). The LC method is based on the computation of a threshold of mismatches (T_M), which considers the number of loci, mistyping and missing data rates, and the allele frequencies of each locus (Wang, 2016). Taking account of these parameters helps to give an unbiased estimate of the average number of mismatches between true duplicated MLGs.

COLONY was parametrized for dioecious species with polygamous mating systems, including inbreeding and clonality. We computed the genotypic richness (R) index within populations (Dorken & Eckert, 2001), as follows: R = (G - 1)/(N – 1), where G is the number of distinct MLLs detected in the population and N the total number of individuals sampled. Two additional diversity statistics were computed in Poppr v2.8.2 R packages to appreciate clonal evenness within populations. The first, relating to the Simpson evenness index (V), while the second the Pareto (β) describes the scaling of the partitioning of MLGs among MLL size classes (Arnaud-Haond et al., 2007). Finally, the average values of genetic diversity and population differentiation (see below) of *L. rodriguezii* were estimated in two complementary ways: first, including all genotyped individuals, second by keeping only one individual per MLL.

Genetic diversity

The within-population genetic diversity was assessed with the datasets of *L. rodriguezii* and *L. digitata* , after the removal of potentially selected loci (Table 1). Given that the percentage of polymorphic loci (%P) may be affected by the variations in sample sizes, a random sampling approach implemented in a custom bash script (available at*https://github.com/rebecca-cj/revegetation*) was employed. The percentage of polymorphic loci was calculated from 100 random resamples (with replacement between samples) for each sample size at population level (Figure S2). In addition, average estimates of observed heterozygosity (Ho), expected heterozygosity (He) and F_{IS} (Weir & Cockerham, 1984) were calculated with the GENEPOP v1.0.5 R package (Rousset, 2008). The same data sets were also analyzed with the Hierfstat R package (Goudet, 2005) to compute the distribution of F_{IS} values across loci. Then, a Statistical Wilcoxon signed-rank test was applied to assess pairwise differences in F_{IS} distributions among samples, and the skewness and Kurtosis were computed to describe the shape of distributions. Ten discrete classes of F_{IS}values were defined and represented as histograms showing the frequency of each class in populations. Considering the observation of generalized negative F_{IS} for *L. rodriguezii* , and positive for*L. digitata* , we tested departure from panmixia with the alternative hypothesis of heterozygosity excess and heterozygosity deficit, respectively. Global Hardy-Weinberg tests [Score (U) test] (Rousset & Raymond, 1995) implemented in GENEPOP were performed among loci. Unbiased estimates of the associated P-values were calculated using a Markov chain (MCMC) algorithm with the following settings; dememorization: 10 000, batch: 100, iterations per batch: 5 000. Then, multiple P values (e.g. one value per locus) were combined by the Fishers method. Finally, pairwise linkage disequilibrium (LD) was analyzed with the standardized index of multilocus association rd (Agapow & Burt, 2001) in Poppr v2.8.2. Genotypes were permuted 1000 times for each population to test for linkage disequilibrium.

Population differentiation in *Laminaria rodriguezii*

Several methods were used to compute population differentiation. First, global and pairwise F_{ST} (Weir & Cockerham, 1984) values were estimated with HierFstat v0.04-22 R package (Goudet, 2005), and second, exact tests of genotypic differentiation, implemented in Genepop v4.2.2, were computed for each locus, testing the null hypothesis that genotypes are drawn from the same distribution in all populations (Rousset, 2008). Unbiased estimates of the associated P-values were calculated using a Markov chain (MCMC) algorithm with default parameters. Fisher's exact test was then performed to compute P-values for each population pair across all loci. Furthermore, population structure was assessed using principal component analysis (PCA) implemented in the adegenet R package (Jombart & Ahmed, 2011). For PCA, the dataset was centered, and missing values were replaced with population mean allele frequencies. We also analyzed the genetic structure with a sparse non-negative matrix factorization (sNMF) analysis (Frichot et al., 2014) implemented in LEA v2.4.0 R package (Frichot & François, 2015). The sNMF algorithm does not assume panmixia, which is interesting for species with mixed reproductive systems. The sNMF was run with one hundred replicates for k values between one and ten, with default parameters.

RESULTS

Sequencing and SNPs filtering

Approximately 1.1 billion (1 154 828 457) raw reads were produced from 49 individuals of *Laminaria rodriguezii* , 130 individuals of*Laminaria digitata* , and 397 samples of *Cystoseira* spp. as part of a distinct project. After quality filtering and read trimming, an average of 2.7 million (± 1.5 million SE) and 2.3 million (± 2.1 million SE) high-quality reads per individual were generated from*L. rodriguezii* and *L. digitata* , respectively. The mapping rate, defined for each individual as the percentage of reads mapped to the *L. digitata* genome, ranged from an average of 85.4% (min: 53.9%; max: 92.2%) for *L. rodriguezii* , to 94.2% (min: 67.9%; max: 97.6%) for *L. digitata* . The initial filtering of loci, which retained those shared by a high proportion of individuals (Table 1), has resulted in catalogs of 44 716 loci and 60 942 loci for *L. rodriguezii* and *L. digitata* , respectively. At this stage, the number of SNPs was more than eight times higher for *L. digitata*than for *L. rodriguezii* (Table 1). Finally, additional quality filtering steps followed by outlier exclusion resulted in the selection of 4077 putatively neutral SNPs per locus among 43 individuals of*L. rodriguezii* (not including two technical replicates) and 7364 putatively neutral SNPs per locus among 116 individuals of *L. digitata* (not including five technical replicates) (Table 1). The rate of missing data was slightly higher in the dataset of *L. digitata*(i.e. 3.36%) than in *L. rodriguezii* (i.e. 2.24%). Furthermore, the SNP error rate estimated from technical replicates ranged from an average of 1.06% for *L. digitata* to 3.89% for *L. rodriguezii* (Table S2, Supporting Information).

Clonal structure of *Laminaria rodriguezii*

The clustering methods implemented in Poppr (i.e distance threshold d = 0.022, defined using pairs of technical replicates) and COLONY resulted in 32 multilocus lineages (MLLs) among the 43 individuals genotyped. Besides, the clustering of individuals within MLLs was identical whether with Poppr or COLONY. In addition, the distribution of genetic distance revealed the occurrence of a gap between the threshold of technical replicate (d = 0.022) and the lowest pairwise distance above this threshold (d = 0.042) (Figure 1A). None of the individual genetic distances of *Laminaria digitata* were below the threshold of technical replicates (Figure 1B). Furthermore, the peak in the distribution (around d = 0.035, Figure 1B) corresponds to intra-population differences within Helgoland. In *L. rodriguezii* , the 32 MLLs included 28 unique and four distinct repeated MLLs. These repeated MLLs, corresponded to the occurrence of multiple ramets belonging to the same genet (3, 3, 7 and 2 for the four repeated MLLs: MLL-A, MML-C, MML-D and MML-E, respectively, Figure 2. Eastern Provence displayed the greatest levels of genotypic richness (R) and evenness (V, β) with average values of R = 0.90, V = 0.88 and β = 2.66 (Table S3, Supporting Information). Conversely, Southern Corsica had the lowest values for these indexes (i.e. R = 0.20, V = 0.51 and β = 0.56). The linear transect at Banc de Magaud_2 allowed measurement of the spatial distance among samples of the single repeated MLLs detected at this site (MLL-C Figure 2): these distances ranged from 0.2 to 0.9 meters. For the other sites, the sampling protocol (with a minimum spacing of two meters), did not allow a similar investigation of the distance among repeated MLLs

Genetic diversity

The partial clonal kelp *L. rodriguezii* showed high values of observed heterozygosity (average within populations 0.16). By contrast, the sexual kelp *L. digitata* exhibited low to moderate levels of Ho (average within populations 0.11; Table 2). The average percentage of polymorphic loci within populations (at n = 10 from 100 random resamples) was 54.42% in*L. rodriguezii* and 39.39% in *L. digitata* (see Table 2, Figure S2). A noticeable result was that observed heterozygosity (Ho) was higher than expected in panmixia (He) within populations for*L. rodriguezii* and conversely strictly lower within populations of *L. digitata* , (see Table 3 for *L. rodriguezii* and Table S1, Supporting Information for*L. digitata*). The heterozygotes excess was statistically significant (Score [U] test, P < 0.001) within populations for *L. rodriguezii* , except for Banc de Magaud_1. The heterozygotes deficit was uniformly significant (Score [U] test, P < 0.001) within populations of *L. digitata* . Considering the distribution of F_{IS} among loci, two main types of distribution were characterized within populations of *L. rodriguezii.* One is reported in Eastern Provence where the distributions of F_{IS} were homogeneous among Banc Magaud_1, Banc Magaud_2, and Cap Camarat and differed from that reported at Bonifacio (pairwise Wilcoxon test, Bonferroni adjusted: p < 0.001). Considering the variance of the F_{IS}among loci, the highest and lowest values were reported at Bonifacio (var = 0.13)

and Eastern Provence, respectively (var = 0.08, 0.07 and 0.06 for Banc Magaud_1, Banc Magaud_2, and Cap Camarat, respectively). Based on the skewness and kurtosis values (skewness = 0.69; kurtosis = 4.07) the distribution in Eastern Provence was moderately skewed to negative values with a long tail to positive values, which contrast with the distribution observed at Bonifacio, which was highly skewed toward negative values (skewness = -0.95; kurtosis = 3.14) (Figure 3A). Furthermore, the F_{IS} distribution at Bonifacio was massed on negative values with a longer tail to negative values below -0.5 and also marked by the absence of F_{IS} values greater than 0.1, except at seven loci. An opposite trend was observed within the populations of *L. digitata* despite the fact that all between-population comparisons reach statistical significance (pairwise Wilcoxon test, Bonferroni adjusted: $p < 0.001$). These distributions were uniformly massed on positives values and seem to be skewed (skewness = 0.98; kurtosis = 3.32) toward positive values greater than 0.5 (Figure 3B). Furthermore, the F_{IS} distributions in *L. digitata* were characterized by a very low occurrence of loci with negative F_{IS} below -0.3. The mean values of F_{IS} within populations ranged from 0.12 to 0.43 for*L. digitata* (Table S1, Supporting Information) whereas mean values were uniformly negative within populations in *L. rodriguezii* , with a range from -0.47 in Bonifacio to -0.02 at Banc de Magaud_1 (Table 3A). After applying the clone correction for *L. rodriguezii* , the average values of F_{IS} increased in a range from F_{IS} = -0.26 at Bonifacio to F_{IS} = -0.01 at Banc de Magaud_2 (Table 3B) in comparison with average values previously observed from the 43 individuals. After MLL correction the heterozygosity excess only remained significant at Bonifacio (Score [U] test, $P < 0.001$). Finally, the linkage disequilibrium, estimated with the r_d index, was statistically significant ($P < 0.001$) within populations of *L. rodriguezii* , before and after the clone correction (Table 3), with the highest values reported before the clone correction at Cap Camarat (r_d = 0.11) and Bonifacio (r_d = 0.21).

Population differentiation in *Laminaria rodriguezii*

L. rodriguezii exhibited high levels of population structure (global F_{ST} = 0.28), including all genotyped individuals (n = 43). However, after applying the clone correction (i.e. to keep one individual per MLL; for details see Table 3B), genetic differentiation decreased among localities (global F_{ST} = 0.18). The highest pairwise F_{ST} was observed between Bonifacio and Cap Camarat (F_{ST} = 0.48 and 0.44 for the non-corrected and corrected datasets, respectively) and the lowest pairwise F_{ST} was between the two sampling sites of Banc de Magaud, with F_{ST} = 0.02 and F_{ST} = 0.01 for the non-corrected and the corrected datasets respectively (Table S4, Supporting Information). Accordingly, all pairwise tests of genotypic differentiation were significant (exact G tests, $p < 0.001$), except between Banc de Magaud_1 and Banc de Magaud_2 (Table S4, Supporting Information). Following the minimum cross-entropy criterion in sNMF, genetic variation was partitioned among all genotyped individuals into six distinct clusters (K = 6). At this K value, all localities appeared separated from one another, except for Banc de Magaud_1 and Banc de Magaud_2 that belonged to the same genetic cluster (Figure S3B). The additional clusters corresponded to repeated MLLs identified with COLONY and the distance network within localities (see Figure 2; MLL-C, MLL-A, and MLL-D). However, on the clustering plot of sNMF, three individuals belonging to different regions, and to distinct MLLs, appeared to be more closely related to each other than individuals of native populations (Figure S3B, the grey cluster). These three outlier individuals shared a common characteristic: the highest proportion of heterozygote genotypes among loci per individual, with an average of 33% against 16% in other individuals. Finally, the PCA analysis (Figure S3A) agreed with the population structure inferred by pairwise F_{ST} values and sNMF. Precisely, the first principal components, accounting for 28 % of the genetic variance, separated Southern Corsica from Eastern Provence. The second principal component with 11.9% of the genetic variance discriminated Cap Camarat from both Banc de Magaud sites. The two Banc de Magaud sites were not separated by PCA analysis. The third axis of the PCA, accounting for 5.9 % of the genetic variance (Figure S4), was relevant to identify outlier individuals observed with sNMF and part of the clusters corresponding to repeated MLLs (i.e. MLL-A and MLL-C).

DISCUSSION

Here, we have shown that high throughput technologies, such as RAD-sequencing, are powerful tools to study the reproductive system in a partially clonal species. The method used here enabled us to not only

to assess clonal membership in *Laminaria rodriguezii* but also to consider how clonality affects the intra-individual genomic diversity. Our empirical results can be confronted with theoretical expectations under different levels of clonality.

RAD Seq as a tool to study partially clonal species

Over the last few years, RAD-sequencing has been used to study partially clonal species, such as the moss *Syntrichia caninervis* (Baughman et al., 2017), the tussock species (*Kobresia tibetica* and*Blysmus sinocompressus* , Ning et al., 2018), the Cardamine (*Kobresia tibetica* and *Blysmus sinocompressus* , Tsujimoto et al., 2020) and scleractinian corals (*Montipora capitate* ,*Porites compressa* in Locatelli & Drew, 2019; *Leptopsammia pruvoti* in Boscari et al., 2019). These studies have enabled the detection of multilocus lineages (MLLs) for determining the spatial architecture of clonal lineages. However, while RAD-sequencing data provide a better resolution than other genotyping methods to assess multilocus lineages (MLLs) (see Le Cam et al., 2019), genotyping errors and missing data introduce additional noise than microsatellite data. It follows that investigating clonality with these data cannot rely on the detection of strict Multilocus Genotypes (MLGs). Rather, inferences on clonality with a high number of SNPs require establishing a distance threshold for which multiple ramets of the same genet would be collapsed in MLLs. Nevertheless, a key question remains: how to find an adequate threshold? Notably, the threshold was defined from the distribution of pairwise genetic distance in admitting that the first peak of small distance was derived from somatic mutations or genotyping errors within genets, and the second peak of long-distance represented variation between genets (Meirmans & Tienderen, 2004). Sometimes, a fixed threshold of similarity was applied to collapse individuals in MLLs following the logic that clonal individuals should be nearly 100% identical (Locatelli & Drew, 2019). However, this is not a universal trend, especially when the background of genotyping errors is high enough to confuse the identification of closely related genotypes, or if independent somatic mutations are numerous enough to accentuate the differences beyond the detection threshold. This also depends on the genetic diversity of the populations: MLLs are more easily identified in populations where the diversity among sexually produced individuals is high. In this study, we have proposed defining a threshold on the basis of the technical of replicates. Although this procedure has the benefit of taking into account the effect of distance induced by genotyping errors, and secondarily by somatic mutations, it has rarely been used for delimiting distinct MLLs, except in *Leptopsammia pruvoti*(Boscari et al., 2019). The value of the distance threshold between ramets of the same genets estimated in our study (d=0.022) is lower than that defined in *Leptopsammia pruvoti* (d = 0.043) from replicates with 2b-RAD (Boscari et al., 2019). Although none of the pairwise-distances in the sexual species *Laminaria digitata* were below the maximum distance from replicates of both species (i.e.*L. rodriguezii* , d = 0.022, *L. digitata* , d = 0.020), the distribution has a peak around d = 0.035 (Figure 1B). Yet these low values all relating to intra-population genetic differences in Helgoland were indicative of a high level of inbreeding (F_{IS} = 0.43) in the population.

The genomic patterns of clonality

The identification of repeated MLGs in addition to mean negative values of F_{IS}, and significant linkage disequilibrium across loci observed in *L. rodriguezii* are features that all agree with signatures of clonality in natural populations (Balloux et al., 2003; Halkett et al., 2005). Nevertheless, we must be aware that inferring the rates of clonality by taking into account both the genotypic richness (R) and the average F_{IS} values remains unachievable, except for high levels of clonality, greater than 95% (Balloux et al., 2003; Stoeckel et al., 2019), or requires the availability of temporal samples (Ali et al., 2016; Becheler et al., 2017). Furthermore, the estimates of R are strongly dependent on sampling strategy and density, basically decreasing with increasing sampling effort (Arnaud-Haond et al., 2007; Becheler et al., 2017; Gorospe et al., 2015). In that context, recent studies reported that R had poor relevance for measuring the relative importance of sexual versus clonal reproduction (Arnaud-Haond et al., 2019), even using relatively large sample sizes (from 100 to 500 individuals) (Stoeckel et al., 2019). Therefore, we suspect that the genotypic richness (R) of *L. rodriguezii* was probably overestimated in our study, due to the complexity of sampling at these depths (i.e. between 65 to 76 meters depth). However, by looking at the F_{IS} distribution in the genome, we propose a new analysis which seems useful to detect partial clonal reproduction. F_{IS} rather than

R is considerably less sensitive to sampling density (Arnaud-Haond et al., 2019; Stoeckel et al., 2019), thus its distribution could be used for detecting genomic signatures of clonal reproduction. Furthermore, inter-locus variance of F_{IS} (rather than an average value) was predicted to increase as clonal rates increase (Balloux et al., 2003; Halkett et al., 2005; Stoeckel & Masson 2014; Reichel et al., 2016; Stoeckel et al., 2019). Further investigations showed that the distribution of F_{IS} along genomes is expected to be concentrated between -0.3 and 0.1 under intermediate levels of clonality, while strongly shifted to negative values for high levels of clonality (Stoeckel & Masson, 2014). Here, we reported (1) two significant different F_{IS} distributions among populations of *L. rodriguezii* suggesting different levels of clonality, and (2) these distributions are shifted towards negative F_{IS} values in contrast with those shifted towards positive F_{IS} values in the fully sexual populations of*L. digitata* (Figure 3). For *L. rodriguezii* , the distributions of F_{IS} in Eastern Provence are in line with the range of values predicted under intermediate levels of clonality, while that of Bonifacio would be expected under high levels of clonality. In accordance with these predictions, we reported higher inter-locus variance in F_{IS} at Bonifacio (var = 0.13) than in Eastern Provence (var = 0.06 to 0.08). Although R should be interpreted carefully, levels of genotype richness were considerably lower at Bonifacio (R = 0.20) than in Eastern Provence (R = 0.80 to 1). Furthermore, recent advances suggested that the Pareto (β) is a better estimate of clonality rates than R, since β decreases further as the rate of clonality increase regardless of the sample size (Stoeckel et al., 2019). For instance, $\beta = 2$, as scored at Banc de Magaud_2 (Eastern Provence) was predicted to reach clonal rates of approximately 0.8 to 0.9 whatever population sizes (Stoeckel et al., 2019), while at Bonifacio, the lowest value ($\beta = 0.56$) agrees with the hypothesis that the rate of clonality could be higher than 0.9 and even close to 1. In comparison with previous studies on other partially clonal species, the Pareto (β) reached the minimum value of $\beta = 0.06$ for a population of *Posidonia oceanica* dominated by a very large genet surrounded by several marginally represented MLGs (see for review Arnaud-Haond et al., 2007). This finding appeared quite consistent with the clonal evenness reported at Bonifacio characterized by the highly skewed distribution of MLGs ($V = 0.51$, $\beta = 0.56$), with one predominant genet (MLL-D, Figure 2) shared among seven MLGs. Finally, genotypic richness and evenness corroborate with the distributions of F_{IS} and their variances, confirming that the rate of clonality is probably considerably higher at Bonifacio than in populations of Eastern Provence.

Confidence in the methodology and data

We reported an intriguing signal of genome-wide heterozygosity in three of the 43 individuals. These outlier individuals (Figure 2, Figure S3B, Figure S4) manifested an exceptionally high proportion of heterozygous genotypes compared to other individuals. High heterozygosity at particular loci can correspond to paralogs assembled in the same RAD locus depending on assembly parameters (Verdu et al., 2016; Mckinney et al., 2017; Ravindran et al., 2018). However, for the following reasons, we would argue that polymorphisms detected within these outlier individuals do not correspond to originating in artefactual loci: firstly, sequencing reads were mapped to the reference genome of a closely related species, followed by stringent filter criteria, including on maximum heterozygosity (Table 1). Secondly, paired-end sequencing protocols provide additional information on correct read placement since RAD loci are retained when both reads are properly paired (i.e. according to the insert length between forward and reverse reads). On the other hand, recent studies have revealed that both cross-contamination among samples and environmental contamination, especially during library preparation and/or sequencing, can lead to artefactual results, with the sequencing of loci of different origins in the same sample (Laurin-Lemay et al., 2012; Ballenghien et al., 2017). Here, the three outlier individuals were also those that had the highest proportion of private alleles among all individuals (average frequency of private alleles (Pr) = 0.203 for outlier individuals and Pr = 0.010 among other 40 genotyped individuals, see Figure S5). Such private alleles can not be explained by cross-contamination among samples of the same species. Rather they could be explained by the accumulation of mutations along with clonal reproduction, thus increasing the divergence among loci in the same individuals, the so-called Meselson effect (Birky, 1996; Welsh & Meselson, 2000). These outlier individuals could then belong to relatively ancient asexual lineage.

Implications for conservation

We reported high genetic differentiation among populations of *L. rodriguezii* , suggesting restricted gene flow between populations, as previously reported for other species of kelp (Billot et al., 2003; Coleman et al., 2009; Coleman et al., 2011; Durrant et al., 2018). These findings were expected given the low dispersal ability of the genus*Laminaria* via haploid spores (Santelices, 1990) and the patchy distribution of *L. rodriguezii* , with few populations. Such a high genetic structure is quite unusual for marine species, and there are still relatively few similar data for comparisons in the same geographical area. As a comparison, strong genetic differentiation was also observed among Corsican and continental populations in the red coral *Corallium rubrum* with RAD sequencing, albeit lower than that observed in *L. rodriguezii* ($F_{ST} = 0.24$ in*Corallium rubrum* , Pratlong et al., 2018). This confirms the expected differentiation among island and continental populations for these low-dispersal species. At a more local scale, a strong genetic differentiation was also observed with microsatellites in other brown algae, *Cystoseira amentacea* , and *C. rubrum* along French Mediterranean coasts (Ledoux et al., 2010; Thibaut et al., 2016). The identification of such a local genetic structure has important implications for management and conservation. In the case of *L. rodriguezii* , but also for the other previously cited species, this points to a structure with very local and poorly connected demes. Considering this structure, we expect that the combined effects of genetic drift and restricted gene flow should have dramatically reduced the within-population genetic variation of *L. rodriguezii*compared to species with wide geographical distribution, exhibiting semi-continuous and dense populations, such as *L. digitata* (see Hamrick & Godt, 1990; Frankham, 1996; Levy et al., 2016). Furthermore, the range of the Mediterranean kelp has been drastically reduced during the last half-century, especially in the Adriatic Sea, where the species has suffered losses of 85% from its historical range (Žuljević et al., 2016). However, when we considered the proportion of polymorphic loci (%P) and the observed heterozygosity (Ho), such genomic patterns are not consistent with the hypothesis that *L. rodriguezii* had low levels of genetic variation. It is remarkable to show that the highest and the lowest values of both indexes (%P, Ho) were reported within populations of *L. rodriguezii* and *L. digitata* , respectively (Table 2). Apart from effects inherent to demographic variations of these two species, such differences in genetic diversity are probably conditioned by their different mating systems. As previously argued with regard to the rare insular plant *Ruta microcarpa* (Meloni et al., 2013), it is permissible that clonal reproduction could retain higher intra-individual polymorphism compared to sexual isolated populations subject to the deleterious effects of genetic drift. This is mainly because through asexual reproduction, the absence of recombination preserves heterozygous genotypes (Judson & Normark, 1996; Balloux et al., 2003) and genetic drift acts on genotype frequencies rather than allele frequencies (Stoeckel & Masson, 2014; Reichel et al., 2016). However, although the genomic signatures of clonality were particularly marked in *L. rodriguezii* , it is more difficult to address if current patterns of genetic diversity bear the molecular signature of past demographic events. Further studies will be needed to clarify whether recent changes in population sizes (see Beichman et al., 2019) have occurred, especially if the contemporary genetic diversity in *L. rodriguezii* is the relic of a historically large population that has since become fragmented. Thus the current distribution of these populations is restricted to the coralligenous rocky bottom which is the last refuge of *L. rodriguezii* in the Mediterranean Sea (Ballesteros, 2006). Through extensive cartography and monitoring by different governmental French programs, all the biocenoses down to 100 m have been explored and monitored for two decades. Thus, we can reasonably claim that we have sampled all the populations of *L. rodriguezii* along the French Mediterranean coasts. Conservation of these last isolated populations should become a high priority for stakeholders or that species could face regional extinction.

The present study is a pioneering work confronting the genomic signatures of partially clonal populations with theoretical predictions. We show the usefulness of RAD sequencing to analyze clonality at the population and genome levels. Of particular interest is the observation of individuals with high genome-wide heterozygosity, a result which could not have been obtained with microsatellites, for example. Further genomic analyzes would be necessary to study the distribution of heterozygosity along the genome and to test, for example, whether such excess is concentrated in particular in low recombining regions. It would also be interesting to test for selective effects along the genome and its potential interaction with clonality and (lack of) recombination. In addition, our results confirm the interest of comparative genomic approaches in conservation. The comparison with*L. digitata* enabled us to highlight differences in genomic diversity among species and populations. This paves the way for demographic inferences to better understand the evolving trends in the scattered and

isolated populations of *L. rodriguezii* . Nevertheless, our results have already provided a sufficient basis to confirm the importance of the conservation of these local populations, which are of high ecological value.

ACKNOWLEDGEMENTS

Funding was provided by the Région PACA, MARFOR Biodiversa/0004/2015, GDR Génomique Environnementale, Agence Nationale de la Recherche, Grant/ Award Number: ANR-18-CE32-0001 (Clonix2D). The project leading to this publication has received funding from the European FEDER Fund under project 1166-39417. This work was also funded by a Ph.D. Fellowship from Région PACA and the Calanques National Park. The authors would like to thank Julie Deter, in charge of the scientific coordination during sampling, the team of research divers of Andromède Océanologie for the prowess in sampling *Laminaria rodriguezii* at 76 meters depth, Claire Daguin-Thiébaut & Quentin Langevin for useful discussions regarding the optimization processes in purpose-extracted high-quality DNA in brown macroalgae, Marck Cock providing the draft genome of*Laminaria digitata* , Thierry Gosselin for very useful discussions on the quality control of RADseq data, and finally Solenn Stoeckel for his comments on the last version of the manuscript.

REFERENCES

[dataset] Reynes, L., Thibaut, T., Mauger, S., Blanfuné, A., Holon, F., Cruaud, C., ... Aurelle, D. (2020). Individual high-quality reads (dd-RADsequencing) "Genomic signatures of clonality in the deep water kelp *Laminaria rodriguezii* ". Dryad. doi:10.5061/dryad.hmgqnk9dq

Adjeroud, M., Guérécheau, A., Vidal-Dupiol, J., Flot, J.-F., Arnaud-Haond, S., & Bonhomme, F. (2014). Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis* : a multi-scale analysis in an insular, fragmented reef system. *Marine Biology* , *161* (3), 531–541. doi: 10.1007/s00227-013-2355-9

Agapow, P.-M., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* , *1* (1–2), 101–102. doi: 10.1046/j.1471-8278.2000.00014.x

Alberto, F., Gouveia, L., Arnaud-Haond, S., Perez-Llorens, J. L., Duarte, C. M., & Serrao, E. A. (2005). Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa* . *Molecular Ecology* ,*14* (9), 2669–2681. doi: 10.1111/j.1365-294X.2005.02640.x

Ali, S., Soubeyrand, S., Gladieux, P., Giraud, T., Leconte, M., Gautier, A., ... Enjalbert, J. (2016). Cloncase: Estimation of sex frequency and effective population size by clonemate resampling in partially clonal organisms. *Molecular Ecology Resources* , *16* (4), 845–861. doi: 10.1111/1755-0998.12511

Andrew, S. (2010). FastQC: a quality control tool for high throughput sequence data. (Version 0.11.7). Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Araújo, R. M., Assis, J., Aguillar, R., Airoldi, L., Bárbara, I., Bartsch, I., ... Sousa-Pinto, I. (2016). Status, trends and drivers of kelp forests in Europe: an expert assessment. *Biodiversity and Conservation* , *25* (7), 1319–1348. doi: 10.1007/s10531-016-1141-7

Ardehed, A., Johansson, D., Schagerström, E., Kautsky, L., Johannesson, K., & Pereyra, R. T. (2015). Complex spatial clonal structure in the macroalgae *Fucus radicans* with both sexual and asexual recruitment. *Ecology and Evolution* , *5* (19), 4233–4245. doi: 10.1002/ece3.1629

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrão, E. A., & Duarte, C. M. (2005). Assessing Genetic Diversity in Clonal Organisms: Low Diversity or Low Resolution? Combining Power and Cost Efficiency in Selecting Markers. *Journal of Heredity* ,*96* (4), 434–440. doi: 10.1093/jhered/esi043

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van De Vliet, M. S., Alberto, F., ... Serrao, E. A. (2007). Vicariance patterns in the Mediterranean Sea: east–west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica* . *Journal of Biogeography* , *34* (6), 963–976. doi: 10.1111/j.1365-2699.2006.01671.x

Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrao, E. A. (2007). Standardizing methods to address clonality in population studies.*Molecular ecology* , 16(24), 5115-5139. doi: 10.1111/j.1365-294X.2007.03535.x

Arnaud-Haond, S., Stoeckel, S., & Bailleul, D. (2020). New insights into the population genetics of partially clonal organisms: when seagrass data meet theoretical expectations. *Molecular Ecology* , mec.15532. doi: 10.1111/mec.15532

Ballenghien, M., Faivre, N., & Galtier, N. (2017). Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. *BMC Biology* ,*15* (1), 25. doi: 10.1186/s12915-017-0366-6

Ballesteros, E. (2006). Mediterranean coralligenous assemblages: a synthesis of present knowledge. *Oceanography and Marine Biology: An Annual Review* , *44* , 123–195.

Balloux, F., Lehmann, L., & de Meeus, T. (2003). The Population Genetics of Clonal and Partially Clonal Diploids. *Genetics Society of America* , 164(4), 1635-1644.

Baughman, J. T., Payton, A. C., Paasch, A. E., Fisher, K. M., & McDaniel, S. F. (2017). Multiple factors influence population sex ratios in the Mojave Desert moss *Syntrichia caninervis* . *American Journal of Botany* , *104* (5), 733–742. doi: 10.3732/ajb.1700045

Becheler, R., Guillemin, M.-L., Stoeckel, S., Mauger, S., Saunier, A., Brante, A., ... Valero, M. (2020). After a catastrophe, a little bit of sex is better than nothing: Genetic consequences of a major earthquake on asexual and sexual populations. *Evolutionary Applications* . doi: 10.1111/eva.12967

Beichman, A. C., Koepfli, K.-P., Li, G., Murphy, W., Dobrynin, P., Kliver, S., ... Wayne, R. K. (2019). Aquatic Adaptation and Depleted Diversity: A Deep Dive into the Genomes of the Sea Otter and Giant Otter. *Molecular Biology and Evolution* , *36* (12), 2631–2655. doi: 10.1093/molbev/msz101

Billot, C., Engel-Gautier, C., 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. *Marine Ecology Progress Series* , *253* , 111–121. doi: 10.3354/meps253111

Birky, C. W. (1996). Heterozygosity, heteromorphy ,and phylogenetic trees in asexual eukaryotes. *Genetics* ,144(1), 427-437.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* ,*30* (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Boscari, E., Abbiati, M., Badalamenti, F., Bavestrello, G., Benedetti-Cecchi, L., Cannas, R., ... Congiu, L. (2019). A population genomics insight by 2b-RAD reveals populations' uniqueness along the Italian coastline in *Leptopsammia pruvoti*(Scleractinia, Dendrophylliidae). *Diversity and Distributions* ,*25* (7), 1101–1117. doi: 10.1111/ddi.12918

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3: Genes, Genomes, Genetics* , *1* (3), 171–182. doi: 10.1534/g3.111.000240

Coleman, M. A., Gillanders, B. M., & Connell, S. D. (2009). Dispersal and gene flow in the habitat-forming kelp, *Ecklonia radiata* : relative degrees of isolation across an east–west coastline.*Marine and Freshwater Research* , *60* (8), 802–809. doi: 10.1071/MF08268

Coleman, M., & Wernberg, T. (2018). Genetic and morphological diversity in sympatric kelps with contrasting reproductive strategies.*Aquatic Biology* , *27* , 65–73. doi: 10.3354/ab00698

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics* , *27* (15), 2156–2158. doi: 10.1093/bioinformatics/btr330

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* , *12* (7), 499–510. doi: 10.1038/nrg3012

De Meeus, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids: A quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution* , *6* (2), 163–170. doi: 10.1016/j.meegid.2005.02.004

Demes, K. W., & Graham, M. H. (2011). Abiotic Regulation of Investment in Sexual Versus Vegetative Reproduction in the Clonal Kelp*Laminaria Sinclairii* (laminariales, Phaeophyceae)1. *Journal of Phycology* , *47* (3), 463–470. doi: 10.1111/j.1529-8817.2011.00981.x

Dorken, M. E., & Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations of a clonal plant, Decodon verticillatus (Lythraceae). *Journal of Ecology* , *89* (3), 339–350. doi: 10.1046/j.1365-2745.2001.00558.x

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can Population Genetic Structure Be Predicted from Life-History Traits? *The American Naturalist* , *169* (5), 662–672. doi: 10.1086/513490

Durrant, H. M. S., Barrett, N. S., Edgar, G. J., Coleman, M. A., & Burridge, C. P. (2018). Seascape habitat patchiness and hydrodynamics explain genetic structuring of kelp populations. *Marine Ecology Progress Series* , *587* , 81–92. doi: 10.3354/meps12447

Ercegović, A. (1960). Quelques traits caracteristiques de la végétation des algues de l'Adriatique. *Acta Botanica Croatica* ,*18–19* .

Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology* , *10* (6), 1500–1508. doi: 10.1046/j.1523-1739.1996.10061500.x

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* ,*6* (8), 925–929. doi: 10.1111/2041-210X.12382

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and Efficient Estimation of Individual Ancestry Coefficients. *Genetics* , *196* (4), 973–983. doi: 10.1534/genetics.113.160572

González de León, S., Herrera, I., & Guevara, R. (2016). Mating system, population growth, and management scenario for *Kalanchoe pinnata*in an invaded seasonally dry tropical forest. *Ecology and Evolution* , *6* (13), 4541–4550. doi: 10.1002/ece3.2219

Gorospe, K. D., Donahue, M. J., & Karl, S. A. (2015). The importance of sampling design: spatial patterns and clonality in estimating the genetic diversity of coral reefs. *Marine Biology* , *162* (5), 917–928. doi: 10.1007/s00227-015-2634-8

Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* , *5* (1), 184–186. doi: 10.1111/j.1471-8286.2004.00828.x

Guillemin, M.-L., Faugeron, S., Destombe, C., Viard, F., Correa, J. A., & Valero, M. (2008). Genetic variation in wild and cultivated populations of the haploid– diploid red alga *Gracilaria chilensis* : how farming practices favor asexual reproduction and heterozygosity. *Evolution* , *62* (6), 1500–1519. doi: 10.1111/j.1558-5646.2008.00373.x

Halkett, F., Plantegenest, M., Prunier-Leterme, N., Mieuzet, L., Delmotte, F., & Simon, J. C. (2005). Admixed sexual and facultatively asexual aphid lineages at mating sites: admixed sexual and asexual lineages in aphids. *Molecular Ecology* , *14* (1), 325–336. doi: 10.1111/j.1365-294X.2004.02358.x

Halkett, F., Simon, J.-C., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms.*Trends in Ecology & Evolution* , *20* (4), 194–201. doi: 10.1016/j.tree.2005.01.001

Hamrick, J. L., & Godt, M. J. W. (1990). Allozyme diversity in plant species. *Plant Population Genetics, Breeding, and Genetic Resources.* , 43–63.

Huson, D. H. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics . 14* (1), 68–73. doi: 10.1093/bioinformatics/14.1.68

Huvé, H. (1955). Présence de *Laminaria rodriguezii* Bornet sur les cotes françaises de la Méditerranée. *Recueil Des Travaux de La Station Marine d'Endoume , 15 .*

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics , 27* (21), 3070–3071. doi: 10.1093/bioinformatics/btr521

Judson, O. P., & Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution , 11* (2), 41–46. doi: 10.1016/0169-5347(96)81040-8

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics , 6 .* doi: 10.3389/fgene.2015.00208

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ , 2 ,* e281. doi: 10.7717/peerj.281

Kettenring, K. M., & Mock, K. E. (2012). Genetic diversity, reproductive mode, and dispersal differ between the cryptic invader, *Phragmites australis* , and its native conspecific. *Biological Invasions , 14* (12), 2489–2504. doi: 10.1007/s10530-012-0246-5

Laurin-Lemay, S., Brinkmann, H., & Philippe, H. (2012). Origin of land plants revisited in the light of sequence contamination and missing data. *Current Biology , 22* (15), R593–R594. doi: 10.1016/j.cub.2012.06.013

Le Cam, S., Daguin-Thiebaut, C., Bouchemousse, S., Engelen, A. H., Mieszkowska, N., & Viard, F. (2019). A genome-wide investigation of the worldwide invader *Sargassum muticum* shows high success albeit (almost) no genetic diversity. *Evolutionary Applications* , 13(3), 500-514. doi: 10.1111/eva.12837

Ledoux, J.-B., Mokhtar-Jamai, K., Roby, C., Feral, J.-P., Garrabou, J., & Aurelle, D. (2010). Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation. *Molecular Ecology , 19* (4), 675–690. doi: 10.1111/j.1365-294X.2009.04516.x

Levy, E., Byrne, M., Coates, D. J., Macdonald, B. M., McArthur, S., & van Leeuwen, S. (2016). Contrasting Influences of Geographic Range and Distribution of Populations on Patterns of Genetic Diversity in Two Sympatric Pilbara Acacias. *PLoS ONE , 11* (10). doi: 10.1371/journal.pone.0163995

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics , 25* (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Locatelli, N., & Drew, J. (2019). Population structure and clonal prevalence of scleractinian corals (*Montipora capitata and Porites compressa*) in Kaneohe Bay, Oahu. BioRxiv doi: 10.1101/2019.12.11.860585

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources , 17* (1), 67–77. doi: 10.1111/1755-0998.12592

Marshall, D. R., & Weir, B. S. (1979). Maintenance of genetic variation in apomictic plant populations. *Heredity* , 42(2), 159-172. doi: 10.1038/hdy.1979.20

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Pinero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources , 15* (1), 28–41. doi: 10.1111/1755-0998.12291

McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations.*Molecular Ecology Resources* , *17* (4), 656–669. doi: 10.1111/1755-0998.12613

Meirmans, P. G., & Tienderen, P. H. V. (2004). Genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms.*Molecular Ecology Notes* , *4* (4), 792–794. doi: 10.1111/j.1471-8286.2004.00770.x

Meloni, M., Reid, A., Caujape-Castells, J., Marrero, A., Fernandez-Palacios, J. M., Mesa-Coelo, R. A., & Conti, E. (2013). Effects of clonality on the genetic variability of rare, insular species: the case of *Ruta microcarpa* from the Canary Islands.*Ecology and Evolution* , *3* (6), 1569–1579. doi: 10.1002/ece3.571

Nadukkalam Ravindran, P., Bentzen, P., Bradbury, I. R., & Beiko, R. G. (2018). PMERGE: Computational filtering of paralogous sequences from RAD-seq data. *Ecology and Evolution* , *8* (14), 7002–7013. doi: 10.1002/ece3.4219

Ning, Y., Wu, G.-J., Ma, H., Guo, J.-L., Zhang, M.-Y., Li, W., ... Duoerji, S.-L. (2018). Contrasting fine-scale genetic structure of two sympatric clonal plants in an alpine swampy meadow featured by tussocks.*PLOS ONE* , *13* (12), e0209572. doi: 10.1371/journal.pone.0209572

Omilian, A. R., Cristescu, M. E. A., Dudycha, J. L., & Lynch, M. (2006). Ameiotic recombination in asexual lineages of Daphnia.*Proceedings of the National Academy of Sciences* , *103* (49), 18638–18643. doi: 10.1073/pnas.0606435103

Pardo, C., Guillemin, M.-L., Pena, V., Barbara, I., Valero, M., & Barreiro, R. (2019). Local Coastal Configuration Rather Than Latitudinal Gradient Shape Clonal Diversity and Genetic Structure of*Phymatolithon calcareum* Maerl Beds in North European Atlantic.*Frontiers in Marine Science* , *6* , 149. doi: 10.3389/fmars.2019.00149

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE* , *7* (5). doi: 10.1371/journal.pone.0037135

Pratlong, M., Haguenauer, A., Brener, K., Mitta, G., Toulza, E., Garrabou, J., ... Aurelle, D. (2018). Separate the wheat from the chaff: genomic scan for local adaptation in the red coral*Corallium rubrum* . *BioRxiv* 306456, ver. 3 peer-reviewed and recommended by Peer Community In Evolutionary Biology. doi: 10.1101/306456

Becheler, R., Masson, JP., Arnaud-Haond, S., Halkett, F., Mariette, S., Guillemin, ML., ...

Stoeckel, S. (2017). ClonEstiMate, a Bayesian method for quantifying rates of clonality of populations genotyped at two-time steps.*Molecular Ecology Resources* , *17* (6), e251–e267. doi: 10.1111/1755-0998.12698

Reichel, K., Masson, J.-P., Malrieu, F., Arnaud-Haond, S., & Stoeckel, S. (2016). Rare sex or out of reach equilibrium? The dynamics of FIS in partially clonal organisms. *BMC Genetics* , *17* (1), 76. doi: 10.1186/s12863-016-0388-z

Rochette, N. C., Rivera-Colon, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology* , *28* (21), 4737–4754. doi: 10.1111/mec.15253

Rousset, F., & Raymond, M. (1995). Testing heterozygote excess and deficiency. *Genetics* , *140* (4), 1413–1419.

Rousset, Francois. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* , *8* (1), 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

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

Schaefer, I., Domes, K., Heethoff, M., Schneider, K., Schon, I., Norton, R. A., ... Maraun, M. (2006). No evidence for the "Meselson effect" in parthenogenetic oribatid mites (*Oribatida, Acari*).*Journal of Evolutionary Biology* , *19* (1), 184–193. doi: 10.1111/j.1420-9101.2005.00975.x

Schwander, T., Henry, L., & Crespi, B. J. (2011). Molecular Evidence for Ancient Asexuality in *Timema* Stick Insects. *Current Biology* , *21* (13), 1129–1134. doi: 10.1016/j.cub.2011.05.026

Stoeckel, S., & Masson, J.-P. (2014). The Exact Distributions of FIS under Partial Asexuality in Small Finite Populations with Mutation.*PLoS ONE* , *9* (1). doi: 10.1371/journal.pone.0085228

Stoeckel, S., Porro, B., & Arnaud-Haond, S. (2019). The discernible and hidden effects of clonality on the genotypic and genetic states of populations: improving our estimation of clonal rates. *arXiv* , 1902.09365, ver. 4 peer-reviewed and recommended by Peer Community in Evolutionary Biology. https://arxiv.org/abs/1902.09365v4 submitted to Molecular Eciology Resources

Thibaut, T., Bottin, L., Aurelle, D., Boudouresque, C.-F., Blanfune, A., Verlaque, M., ... Millet, B. (2016). Connectivity of Populations of the Seaweed *Cystoseira amentacea* within the Bay of Marseille (Mediterranean Sea): Genetic Structure and Hydrodynamic Connections.*Cryptogamie, Algologie* , *37* (4), 233–255. doi: 10.7872/crya/v37.iss4.2016.233

Tsujimoto, M., Araki, K. S., Honjo, M. N., Yasugi, M., Nagano, A. J., Akama, S., ... Kudoh, H. (2020). Genet assignment and population structure analysis in a clonal forest-floor herb, *Cardamine leucantha* , using RAD-seq. *AoB PLANTS* , *12* (1), plz080. doi: 10.1093/aobpla/plz080

Verdu, C. F., Guichoux, E., Quevauvillers, S., De Thier, O., Laizet, Y., Delcamp, A., ... Mariette, S. (2016). Dealing with paralogy in RADseq data: in silico detection and single nucleotide polymorphism validation in *Robinia pseudoacacia* L. *Ecology and Evolution* , *6* (20), 7323–7333. doi: 10.1002/ece3.2466

Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics* , *166* (4), 1963–1979. doi: 10.1534/genetics.166.4.1963

Wang, J. (2016). Individual identification from genetic marker data: developments and accuracy comparisons of methods. *Molecular Ecology Resources* , *16* (1), 163–175. doi: 10.1111/1755-0998.12452

Weir, B. S., & Cockerham, C. C. (1984). Estimating f-statistics for the analysis of population structure. *Evolution* , *38* (6), 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x

Welch, D. M., & Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange.*Science* , *288* (5469), 1211–1215. doi: 10.1126/science.288.5469.1211

Žuljević, A., Peters, A. F., Nikolić, V., Antolić, B., Despalatović, M., Cvitković, I., ... Küpper, F. C. (2016). The Mediterranean deep-water kelp *Laminaria rodriguezii* is an endangered species in the Adriatic Sea. *Marine Biology* , *163* (4), 69. doi: 10.1007/s00227-016-2821-2

DATA ACCESSIBILITY

Individual high-quality reads (dd-RADseq) that support the findings of this study are openly available in a Dryad Digital Repository (doi:10.5061/dryad.hmgqnk9dq). The home-made scripts used to generate the distribution of F_{IS} values into classes, the histograms of pairwise genetic distances, and the graphs representing the percentage of polymorphic loci are available in public repositories with GitHub.*https://github.com/LauricReynes*

AUTHOR CONTRIBUTION

Thierry Thibaut devised the project and was responsible for the main conceptual ideas. Florian Holon performed the sampling of*Laminaria rodriguezii* . Aurélie Blanfuné worked out almost all of the technical details and contributed to sample preparation. Stephane Mauger and Lauric Reynes performed the RAD-sequencing experiment and analyzed sequencing reads. Corinne Cruaud and Arnaud Couloux provided the draft genome of *Laminaria digitata* . Lauric Reynes performed research, analyzed data, and wrote the manuscript in close

consultation with Didier Aurelle and Myriam Valero. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

TABLES AND FIGURES

Tables captions

Table 1 : Number of SNPs at each filtering step of the bioinformatic processing from datasets of *Laminaria rodriguezii*and *Laminaria digitata* .

Table 2 : Average (and range) of the intra-population genetic variation in *Laminaria rodriguezii* (n = 43) and *Laminaria digitata* (n = 116) computed across 4 077 and 7364 putatively neutral SNPs, respectively. %P, average percentage of polymorphic loci within populations at n = 10 (from 100 random resamples, see Figure S2), Ho, observed heterozygosity, He, expected heterozygosity, F_{IS}, inbreeding coefficient.

Table 3 : Intra-population genetic variation of *Laminaria rodriguezii* with (A) all genotyped individuals (n = 43) and (B) one individual per Multilocus Lineage (MLLs) (n = 32). N, number of individuals, Ho, observed heterozygosity, He, expected heterozygosity, F_{IS}, average inbreeding coefficient, highly significant heterozygote excess (P-val < 0.001) highlighted in bold, r_d, coefficient of Multilocus Linkage Disequilibrium (LD), highly significant linkage disequilibrium (P-val < 0.001) highlighted in bold.

Table 1

Filtering SNP candidates	*L. rodriguezii* **SNP counts**	*L. digitata* **SNP counts**
Stacks catalogue	283 864	1 194 460
Population module		
>75% of the samples (-p)	70 026	609 699
>75% of the populations (-r)		
Hobs <0.8		
Single SNP per locus	27 769	58 830
Minor allele frequency (MAF)		
MAF > 0.02	16 865	21 090
Maximum amount of missingness		
Per individual < 0.30	13 235	15 150
Per site < 0.20		
Average depth of coverage per site		?¿?
15X and <100X	4 366	7 538
Outliers detection		
Putatively neutral	4 077	7 364

Table 2

		%P		**Ho**	**Ho**	**Ho**	**He**	**He**	**He**	**F_{IS}**	**F_{IS}**	**F_{IS}**
L. rodriguezii												
Average value	**54.42**	**54.42**	**54.42**	**0.16**	**0.16**	**0.16**	**0.14**	**0.14**	**0.14**	**-0.14**	**-0.14**	**-0.14**
Range of values	40.96	**-**	72.14	0.11	**-**	0.19	0.11	**-**	0.18	-0.47	**-**	-0.02
L. digitata												
Average value	**39.39**	**39.39**	**39.39**	**0.11**	**0.11**	**0.11**	**0.15**	**0.15**	**0.15**	**0.26**	**0.26**	**0.26**
Range of values	21.44	**-**	54.89	0.09	**-**	0.14	0.12	**-**	0.19	0.12	**-**	0.43

Table 3	**Table 3**					
	A		**A**	**A**	**A**	**A**	**A**

Table 3	**Table 3**						
Population	**Population**	**N**	**Ho**		**He**	**FIS**	**rd**
Banc Magaud_1	Banc Magaud_1	10	0.18		0.17	-0.02	**0.01**
Banc Magaud _2	Banc Magaud _2	11	0.19		0.18	**-0.04**	**0.08**
Cap Camarat	Cap Camarat	11	0.11		0.11	**-0.05**	**0.11**
Bonifacio	Bonifacio	11	0.16		0.11	**-0.47**	**0.21**
	B	**B**	**B**	**B**	**B**	**B**	
Population	**Population**	**N**	**Ho**		**He**	**FIS**	**rd**
Banc Magaud_1	Banc Magaud_1	10	0.18		0.17	-0.02	**0.01**
Banc Magaud _2	Banc Magaud _2	9	0.19		0.19	-0.01	**0.05**
Cap Camarat	Cap Camarat	10	0.11		0.11	-0.02	**0.10**
Bonifacio	Bonifacio	3	0.20		0.16	**-0.26**	**0.21**

Figures captions

Figure 1 : Frequency of pairwise genetic distances among individuals of (A) *Laminaria rodriguezii* (n = 43), (B)*Laminaria digitata* (n = 116), and (C) technical replicates of each species (n = 7). The distributions were computed across 4077 and 7364 putatively neutral SNPs *for L. rodriguezii* and *L. digitata* , respectively. The vertical dotted line shows the distance threshold (d = 0.022) as reported as the maximum value between technical replicates. in Figure (A), all pairwise comparisons below the threshold were compatible with clonal reproduction, while in Figure (B), the peak in the distribution (around d = 0.035) related to intra-population differences within Helgoland in the sexual species.

Figure 2 : Phylogenetic networks of *Laminaria rodriguezii*(n = 43) represented with the neighbor-net method. Unique and repeated Multilocus Lineage (MLLs) identified among individuals (see Results) as related to the color in the legend. Outlier individuals (n = 3) characterized by a high proportion of heterozygote genotypes were indicated.

Figure 3: Distributions of the inbreeding coefficient F_{IS} along genomes within populations of (A)*Laminaria rodriguezii* (n = 43) and (B) *Laminaria digitata*(n = 116) computed across 4 077 and 7364 putatively neutral SNPs, respectively. The F_{IS} values were distributed into 10 discrete classes and the number of occurrences for each class was indicated on top of histogram bars.