**MATERIALS AND METHODS**

*Sampling collection and molecular protocols*

*Patella aspera* individuals were collected from six locations in the islands Madeira, Desertas, and Porto Santo from Madeira Archipelago in the North Atlantic Ocean (Table 1; Fig. 1). Twelve individuals of *P. aspera* were collected within each location to develop microsatellite markers. Microsatellite markers can help to detect recent changes in the genetic composition of populations due to its high mutation rate (105), and thus they could reflect signals of recent connectivity and increase or decrease of the genetic diversity due to the conservation of Marine Protected Areas (MPA) and overexploitation respectively. In order to develop microsatellite markers for *P. aspera*, DNA was extracted using an E.Z.N.A Mollusc DNA kit (Omega Bio-Tek) following the manufacturer's instructions. PCRs were performed in final reaction volume of 12.5 µL, containing 1 µL of DNA, 6.25µL of the Type-it Microsatellite PCR Kit (Qiagen), 4 µL of PCR-grade water, and 1.25µL of the primer mix. The optimal PCR protocol consisted in an initial denaturation step at 95° C for 5 min, followed by 30 cycles of 95° C for 30 s, 57° C for 90 s, 72° C for 30 s; 8 cycles of 95° C for 30 s, 53° C for 90 s, 72° C for 30 s; and final extension step at 68° C for 30 min. Alleles were called using Geneious 10.2.3 (Biomatters Ltd) to observe the allelic peaks obtained for each individual. All DNA extractions and microsatellite markers development presented here were performed by Allgenetics (https://www.allgenetics.eu).

*Descriptive analyses*

Hardy-Weinberg Equilibrium (HWE) and Linkage Disequilibrium (LD) tests were estimated in Genepop on the web (http://genepop.curtin.edu.au/) to assess any markers or population differing from the assumptions made to estimate gene flow among populations. The False Discovery Rate (FDR) approach was used to account for multiple tests in both the HWE and LD using FDR. Number of alleles (allelic richness (*A*), and observed heterozygosity (*Ho*) were calculated in Genalex 6.5 (Peakall &Smouse 2012) to compare the genetic diversity between collection sites using different parameters. While the first two parameters account for recent processes affecting the genetic composition of populations, the third parameter accounts for historical processes. Thus both set of parameter could reflect processes acting at different periods of time. Since three sampled locations (Rocha do Navio, Garajau, and Desertas) are Marine Protected Areas (MPA) and one of them (Paúl do Mar) has been overexploited, it is expected that *Na* and *A* should reflect any impact of MPA in contemporary genetic diversity of populations.

A *FST* distance matrix created in GENALEX 6.5 (Peakall &Smouse 2012) was constructed to observe the genetic distance among *Patella aspera* individuals from the 6 sampled locations. Genetic distance should reflect similarity and differences in the genetic composition of individuals from different locations and thus could help to infer the relationship between individuals within locations.

*Population differentiation and gene flow*

The number of populations and the degree of admixture among them was assessed using a Discriminant Analysis of Principal Components (DAPC; Jombart 2010). DAPC is a non-model-based method that search for linear combinations of alleles showing differences between groups as best as possible while minimizing variation within clusters without relying on assumptions required in other software such as LD and HWE. The analysis was run in ADEGENET server (Jombart 2008) to construct a DAPC plot, the proportion of successful reassignment of individuals to their original clusters (*assign.per.pop*), and a membership probability plot. Cross-validation analysis was made to determine whether the number of PCs retained in the analysis was correct. For this purpose, the analysis was executed with a maximum number of PCA = 64, training set = 90%, and 30 replicates. In order to investigate any population differentiation among Madeira Island individuals, allele frequencies obtained in GENALEX 6.5 were plotted to identify loci explaining this differentiation. Population structure was also investigated using the software STRUCTURE (Pritchard 2000) using sampling location (locprior) and 10 replicates with 100000 burnin and 1000000 iterations. Runs were summarized with STRUCTURE HARVESTER (Earl et al. 2012) and CLUMPP (Jakkobsson & Rosenber 2007) and visualized with DISTRUCT (Rosenberg 2004).

Recent migration rates were calculated in BAYESASS 3.0 (Wilson & Rannala 2003), identifying the gene flow signature from the current generation up to two generations in the past. Four replicates with 10000000 iterations were used, sampling every 1000 samples with burn-in equal to 1000000, deltaA = 0.30, and deltaF = 0.50. Trace results were observed in Tracer (Rambaut et al. 2018) to identify that the probability of all parameters reached a plateau and that all parameters presented a smooth distribution

**RESULTS**

*Microsatellite development*

A total of 17 new polymorphic markers were developed, amplified and sequenced. Six markers were trinucleotide repeat motifs and eleven were dinucleotide repeat motifs (Table S1). Despite the laboratory efforts, 3 markers did not amplify for all individuals, producing missing values above 40% within location/markers. Therefore they were not included in all analyses, giving a total of 14 markers useful for all populations.

*Descriptive analyses*

Although 12 individuals were collected within each location, not all individuals were successfully amplified for the 14 polymorphic microsatellite markers developed for *P. aspera*. One marker ASP2F presented a 20% of missing values among all 64 individuals amplified and although it was included for descriptive analyses (n = 14), was not considered in gene flow and population differentiation analyses (n = 12). HWE tests after applying the FDR showed that locus ASP3F deviated from the assumption of HWE, therefore this locus was removed from posterior gene flow analyses in order to accomplish with all assumptions of BAYESASS analysis. Deviations from HWE occurred in 7 other cases but they were not population nor marker specific, therefore no other marker was removed from the analyses. No evidence of Linkage Disequilibrium was found among samples.

Number of alleles (*Na*) and allelic richness (*Ar*), both allelic-based methods displaying the genetic diversity caused by recent processes, were lower in Paúl do Mar (*Na* = 5.929, *Ar* = 4.189) and higher in Garajau and Desertas (*Na* = 7.500 and 7.500, *Ar* = 4.967 and 4.921; Table 1). This patterns was observed when including all 14 microsatellite markers as well as with 12 markers, after removing the locus ASP3F that deviated from HWE. Observed heterozygosities were lower in Porto Moniz (*Ho* = 0.398) and higher in Porto Santo (*Ho* = 0.628), a pattern maintained when removing locus ASP3F (Table 1).

*FST* -based genetic distances (Table 2) ranged from 0.025 to 0.046, revealed that the lowest distances were found between Desertas (*FST* = 0.025), and Garahau Rocha do navio and Desertas (*FST* = 0.026), Rocha do navio and Garajau (*FST* = 0.028). The largest genetic distances were found between Porto Santo and Paúl do Mar (*FST* = 0.046), Porto Santo and Garajau (*FST* = 0.047) as well as between Paúl do Mar and Porto Maniz (*FST* = 0.046).

*Population differentiation and gene flow*

Population differentiation observed in the DAPC analysis using 20 PCs (RMSE = 0.775) showed a high relationship between individuals from Porto Moniz, Rocha do navio, Garajau, and Desertas and a slight differentiation of individuals from the distant Porto Santo and the less diverse Paúl do Mar (Fig. 2). Cluster analysis of group membership probability (Fig. 3) exhibited, in general, an important amount of admixture in individuals within locations and less admixture among individuals from Porto Santo, supporting the higher differentiation of individuals from this location observed in DAPC. Assignment per population ranged from 0.75 in Desertas to 0.91 in Porto Santo, confirming the higher differentiation in individuals from this location. Allele frequency plots (Fig. S1) revealed that individuals from Paúl de Mar in Madeira Island presented a pattern in the allele frequencies of loci ASP21F, ASP34F, ASP38F, and ASP17F with a Gaussian distribution while individuals from the other three locations in the island presented two or more peaks in their allele frequency plots (Fig. S1), suggesting that the loss of alleles in individuals from Paúl do Mar. STRUCTURE results (Fig. 4) revealed a very subtle differentiation among individuals from Porto Santo and all other five locations, supporting the differentiation of Porto Santo population.

Contemporary gene flow calculated in BAYESASS (Table 3, Fig. 1) evidenced a low to nil connectivity among all populations except from individuals from the MPA located in Garajau, suggesting than this population could act as source of genetic variability among individuals from this location into all other locations.

**DISCUSSION**

A recent study by Faria et al. (2018) attempted to reveal the main genetic features of *P. aspera* in the Macaronesia, finding that populations from Madeira and Gran Canaria were well differentiated from all other islands. Collected individuals of *P. aspera* from that study corresponded to a single location and therefore did not intend to identify the genetic patterns of *P. aspera* within Madeira Archipelago. Here, by studying *P. aspera* individuals from six locations within the Madeira Archipelago, using a novel set of microsatellite markers, found subtle signals of differentiation and higher genetic diversity in Marine Protected Areas (MPA) and lower genetic diversity in overexploited areas.

Figure S1 shows the effects of genetic drift and loss of alleles in loci ASP21F, ASP34F, ASP38F, and ASP17F in the overexploited population of *P. aspera* in Paúl do Mar. Allelic diversity can be lost as consequence of population size declining due to overexploitation (Allendorf et al. 2008), reducing the possibility of populations to evolve in the future (Crow & Kimura 1970, Ryman et al. 1995). Although individuals from Paúl do Mar did not have the lowest *Ho* (*Ho* = 0.398), it is still low when compared among the six sampled locations and only compared to *Ho* in Porto Moniz (*Ho* = 0.396). It is known that although heterozygosity is the preferred measure of loss of genetic diversity it might not reflect the consequences of bottlenecks reducing populations size neither the consequences of recent processes (Nei et al. 1976, Maruyama & Fuerst 1985, El Mousadik & Petit 1996). Therefore it is expected to observe the consequences of overexploitation (population size reduction and genetic drift) in *P. aspera* in allele based methods rather than in *Ho*.

Gene flow estimates displayed a high connectivity from the MPA Garajau into all other locations, homogenizing allele frequencies and reducing population differentiation (Slatkin 1985). Thus, the subtle population differentiation signal observed with DAPC and STRUCTURE could be explained by distance among Madeira and Porto Santo Island. Although high gene tends to homogenize populations it can also help to increase the genetic diversity of overexploited populations, since these populations can act as sink populations, receiving a high amount of alleles that can help to increase or recover genetic diversity (Aitken et al. 2013, Carlson et al. 2014). Therefore, the high gene flow into Paúl do Mar can help to recover the genetic diversity and be proposed as genetic rescue measure (Frankham 2015, Whiteley et al. 2015).

Table 1. Location, coordinates, number of samples (N), Number of alleles (*Na*), Allelic richness (*Ar*), and Observed heterozygosity (*Ho*) for individuals of *Patella aspera* from 6 locations in the Madeira Archipelago. Subscript 12 and 14 specify the number of microsatellite markers used to calculate each parameter.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Location** | **Coordinates** | **N** |  | ***Na*14** | ***Ar*14** | ***Ho*14** | ***Na*12** | ***Ar*12** | ***Ho*12** |
| Rocha do Navio | 32º48'26''N | 10 | **Mean** | 6,429 | 4,886 | 0,493 | 6.615 | 4.960 | 0.485 |
|  | 16º51'35''W |  | **SE** | 0,803 | 0,697 | 0,066 | 0.844 | 0.749 | 0.070 |
|  |  |  |  |  |  |  |  |  |  |
| Porto Moniz | 32º51'49''N | 10 | **Mean** | 6,214 | 4,403 | 0,398 | 6.308 | 4.467 | 0.413 |
|  | 17º09'51''W |  | **SE** | 0,743 | 0,579 | 0,068 | 0.796 | 0.621 | 0.072 |
|  |  |  |  |  |  |  |  |  |  |
| Paúl do Mar | 32º45'49''N | 9 | **Mean** | 5,929 | 4,189 | 0,396 | 6.000 | 4.183 | 0.427 |
|  | 17º14'05''W |  | **SE** | 0,569 | 0,519 | 0,075 | 0.610 | 0.561 | 0.074 |
|  |  |  |  |  |  |  |  |  |  |
| Garajau | 32º38'45''N | 12 | **Mean** | 7,500 | 4,967 | 0,585 | 7.692 | 5.129 | 0.623 |
|  | 16º53'15''W |  | **SE** | 0,959 | 0,759 | 0,062 | 1.015 | 0.801 | 0.052 |
|  |  |  |  |  |  |  |  |  |  |
| Desertas | 32º30'22''N | 12 | **Mean** | 7,500 | 4,921 | 0,461 | 7.615 | 4.936 | 0.490 |
|  | 16º30'33'' |  | **SE** | 0,669 | 0,594 | 0,061 | 0.712 | 0.641 | 0.057 |
|  |  |  |  |  |  |  |  |  |  |
| Porto Santo | 33º06'16''N | 11 | **Mean** | 6,857 | 4,269 | 0,628 | 6.846 | 4.133 | 0.663 |
|  | 16º19'56''W |  | **SE** | 0,533 | 0,419 | 0,079 | 0.576 | 0.428 | 0.022 |

Table 2. *FST*-based genetic distance among *Patella aspera* individuals from 6 populations located in the Madeira Archipelago. Values range from no genetic distance (*FST* = 0) to high genetic distance (*FST* = 1).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Rocha do Navio | Porto Moniz | Paúl do Mar | Garajau | Desertas | Porto Santo |
| Rocha do Navio | 0,000 |  |  |  |  |  |
| Porto Moniz | 0,031 | 0,000 |  |  |  |  |
| Paúl do Mar | 0,042 | 0,046 | 0,000 |  |  |  |
| Garajau | 0,028 | 0,044 | 0,036 | 0,000 |  |  |
| Desertas | 0,026 | 0,035 | 0,037 | 0,025 | 0,000 |  |
| Porto Santo | 0,036 | 0,040 | 0,046 | 0,047 | 0,040 | 0,000 |

Table 3. Gene flow estimates calculated with BAYESASS among individuals of *Patella aspera* from 6 locations in the Madeira Archipelago. Source populations are in the left and receiving populations are in the top.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Rocha do Navio | Porto Moniz | Paúl do Mar | Garajau | Desertas | Porto Santo |
| Rocha do Navio | 0.6876(0.0195) | 0.0203(0.0182) | 0.0217(0.0202) | 0.0206(0.0189) | 0.0201(0.0183) | 0.0194(0.0176) |
| Porto Moniz | 0.0209(0.0205) | 0.6878(0.0199) | 0.0235(0.0224) | 0.0224(0.0200) | 0.0207(0.0201) | 0.0195(0.0184) |
| Paúl do Mar | 0.0225(0.0203) | 0.0203(0.0185) | 0.6893(0.0213) | 0.0217(0.0208) | 0.0214(0.0192) | 0.0192(0.0187) |
| Garajau | 0.2265(0.0384) | 0.2287(0.0377) | 0.2206(0.0398) | 0.8890(0.0366) | 0.2268(0.0362) | 0.2347(0.0362) |
| Desertas | 0.0212(0.0192) | 0.0209(0.0190) | 0.0219(0.0205) | 0.0217(0.0204) | 0.6876(0.0199) | 0.0204(0.0193) |
| Porto Santo | 0.0212(0.0187) | 0.0220(0.0206) | 0.0229(0.0215) | 0.0247(0.0226) | 0.0234(0.0216) | 0.6868(0.0184) |

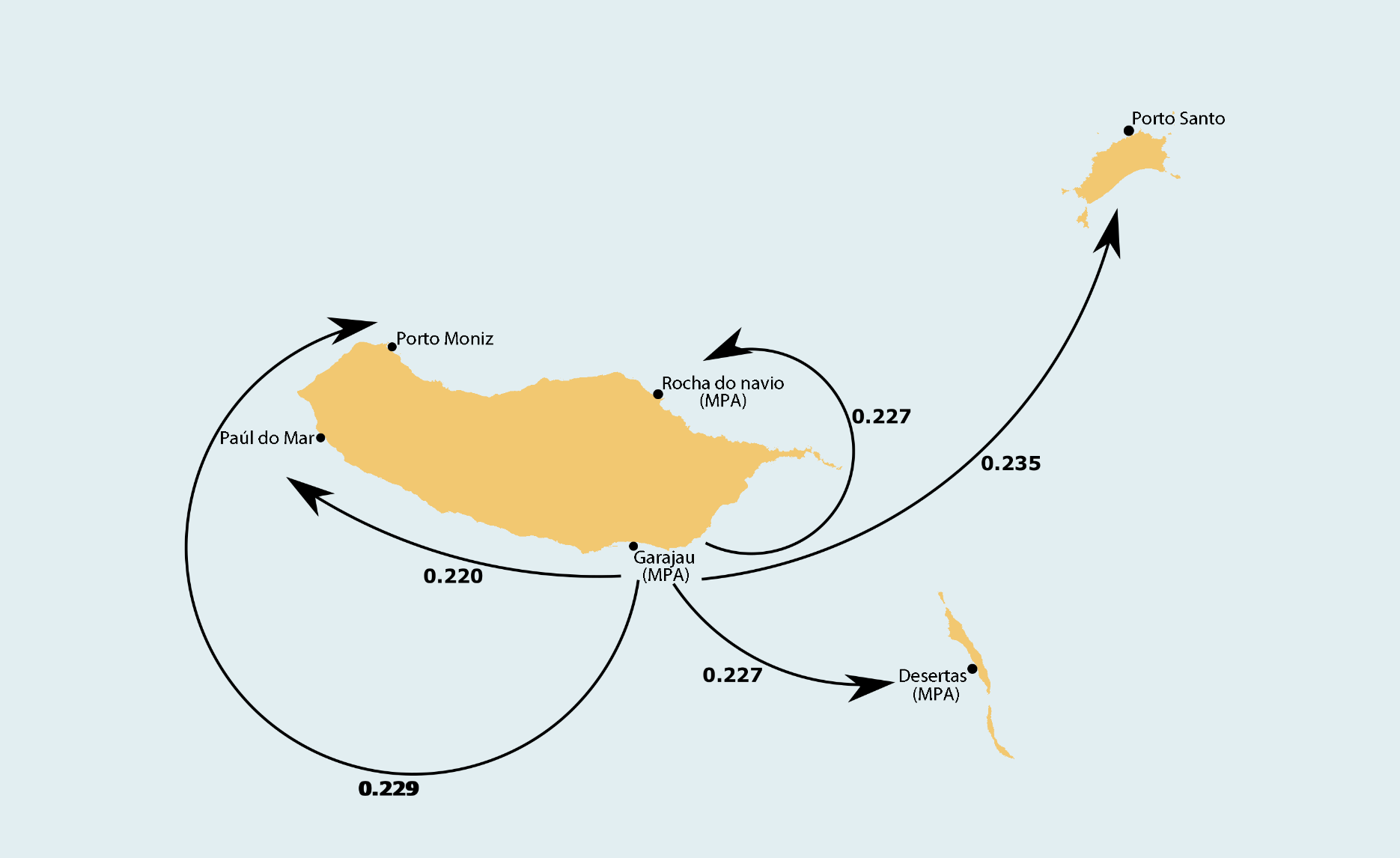


Fig. 1 Sampled locations for *Patella aspera* individuals in the Madeira Archipelago and gene flow estimates different of zero calculated with BAYESASS.

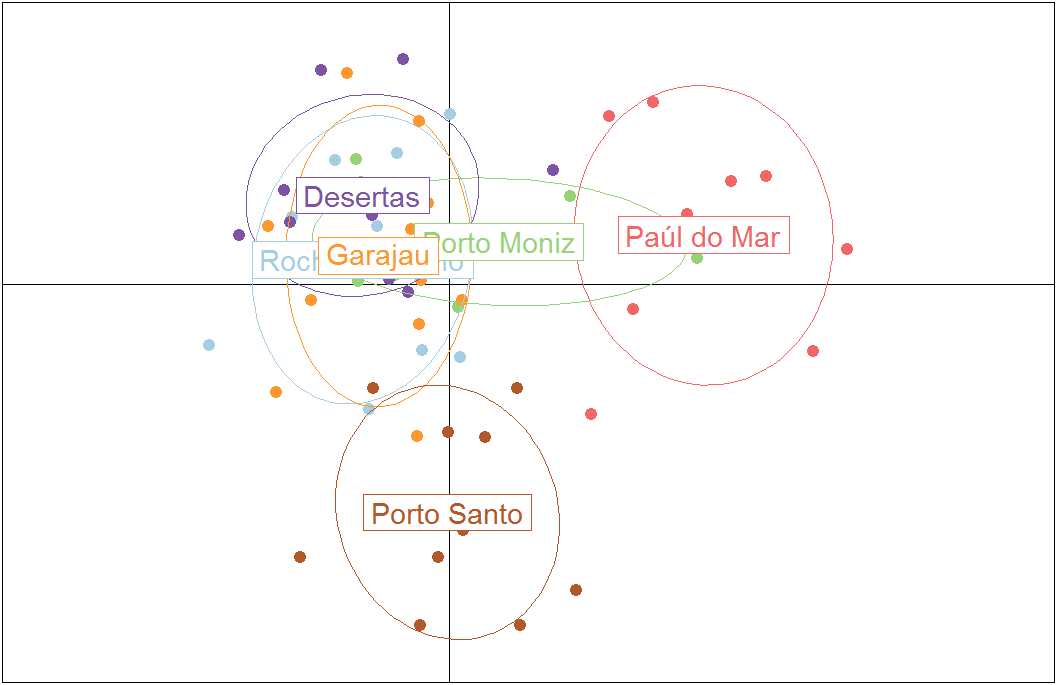


Fig. 2. DAPC analysis of individuals of *Patella aspera* from 6 locations in the Madeira Archipelago.

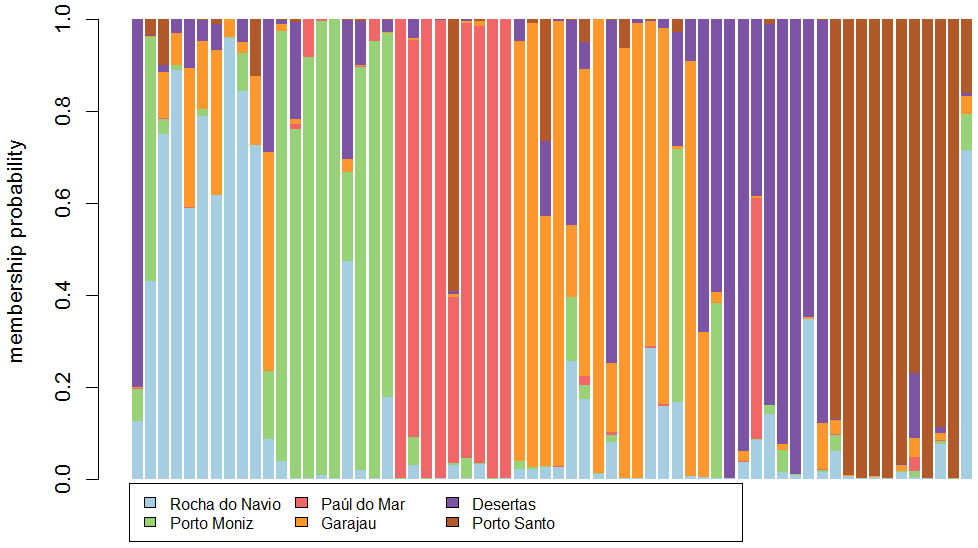


Fig. 3. Membership probability for each *Patella aspera* sampled in the Madeira Archipelago.

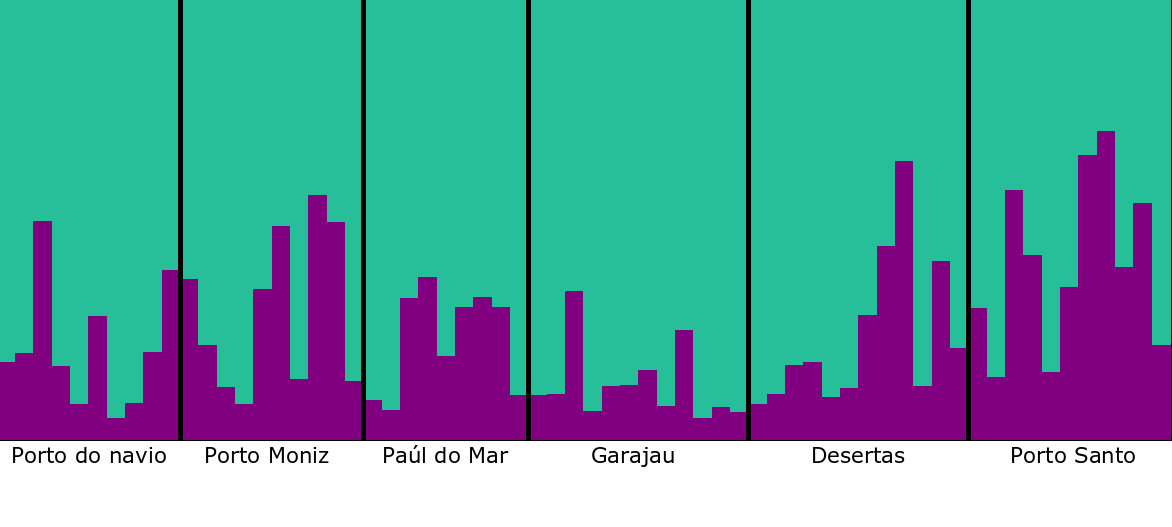


Fig. 4 STRUCTURE results showing k = 2 (k = number of populations) running 10 replicates.

Table S1. Characteristics of 17 microsatellite loci developed for *Patella aspera* developed for this study. F: forward primer sequence; R: reverse primer sequence.

|  |  |  |  |
| --- | --- | --- | --- |
| Locus | Primer sequences (5'-3') | Amplicon size | Repeat motif |
| ASP2F | F: CGTACTTCAATTGGCGAAGC | 110-146 | (AAG)n |
|  | R: ACGATCTAGCAACCCCTGACT |  |  |
| ASP3F | F: TTTCTGTCATGTGACTTTCTATCCTC | 110-134 | (TCT)n |
|  | R: TCTATGATGCGACACCTCCTT |  |  |
| ASP21F | F: GTGCATAAAATTTGGTTGCG | 168-188 | (AG)n |
|  | R: GAACCTGCAATAAGCAGATTTACA |  |  |
| ASP23F | F: TATGAACCGCCCCTTTAAGA | 165-216 | (TCA)n |
|  | R: AGCTGAGGGTGATGGTGATGT |  |  |
| ASP33F | F: GGTTCTACGTAAGACCGTCTGG | 235-275 | (AG)n |
|  | R: TCTTCAAATGAGGACTAAGCAGTT |  |  |
| ASP34F | F: AGATGTCCCATCTGAGGTGC | 245-275 | (GA)n |
|  | R: TACCCACCCGAACGACTATG |  |  |
| ASP15F | F: ACAGTCACGAACAGGGTATGTT | 156-222 | (TTC)n |
|  | R: ACCTCCCACAAGAAAGACTCCT |  |  |
| ASP24F | F: ATGTTCAAGCGATTGGAAGG | 160-214 | (GA)n |
|  | R: GCAGAACATTGTGACCCAAA |  |  |
| ASP29F | F: TCTTCAAATGAGGACTAAGCAGTT | 211-248 | (CT)n |
|  | R: GCCTGTCTGTTTTACAGGGG |  |  |
| ASP36F | F: ACCCTTTTGTGTGATGAGGG | 288-344 | (GA)n |
|  | R: ATTGTTTGTAGTGGATGTTGAAGC |  |  |
| ASP38F | F: GAAGTTTATATCACTCAGGGCCTA | 302-352 | (AG)n |
|  | R: AGTCTAGAGTGCCGCGCTT |  |  |
| ASP39F | F: TGTTGGATATAGAGCGTGTTTCA | 315-342 | (ACC)n |
|  | R: TTCACCTAGGGGAGGGATAGA |  |  |
| ASP7F | F: CTGTCTTTCTCGTCTCACTCTCA | 133-148 | (ATC)n |
|  | R: AGGTTGTGGACGTTGAACTGAT |  |  |
| ASP17F | F: ATAAATAAATGTACAACCATTGACACA | 148-168 | (AG)n |
|  | R: TACGGTTGTACGTGACAAGGA |  |  |
| ASP26F | F: ATTGGTGGACACCCACAATTA | 194-232 | (AG) |
|  | R: ATTGAGTCACCGGCGTAGTT |  |  |
| ASP27F | F: TTTTCTCAGGGTACTCCGGTT | 195-225 | (TC) |
|  | R: GGCATAATGGCAGGGTGAAT |  |  |
| ASP40F | F: CAATTTCATTGACGCAAAGC | 311-353 | (CA) |

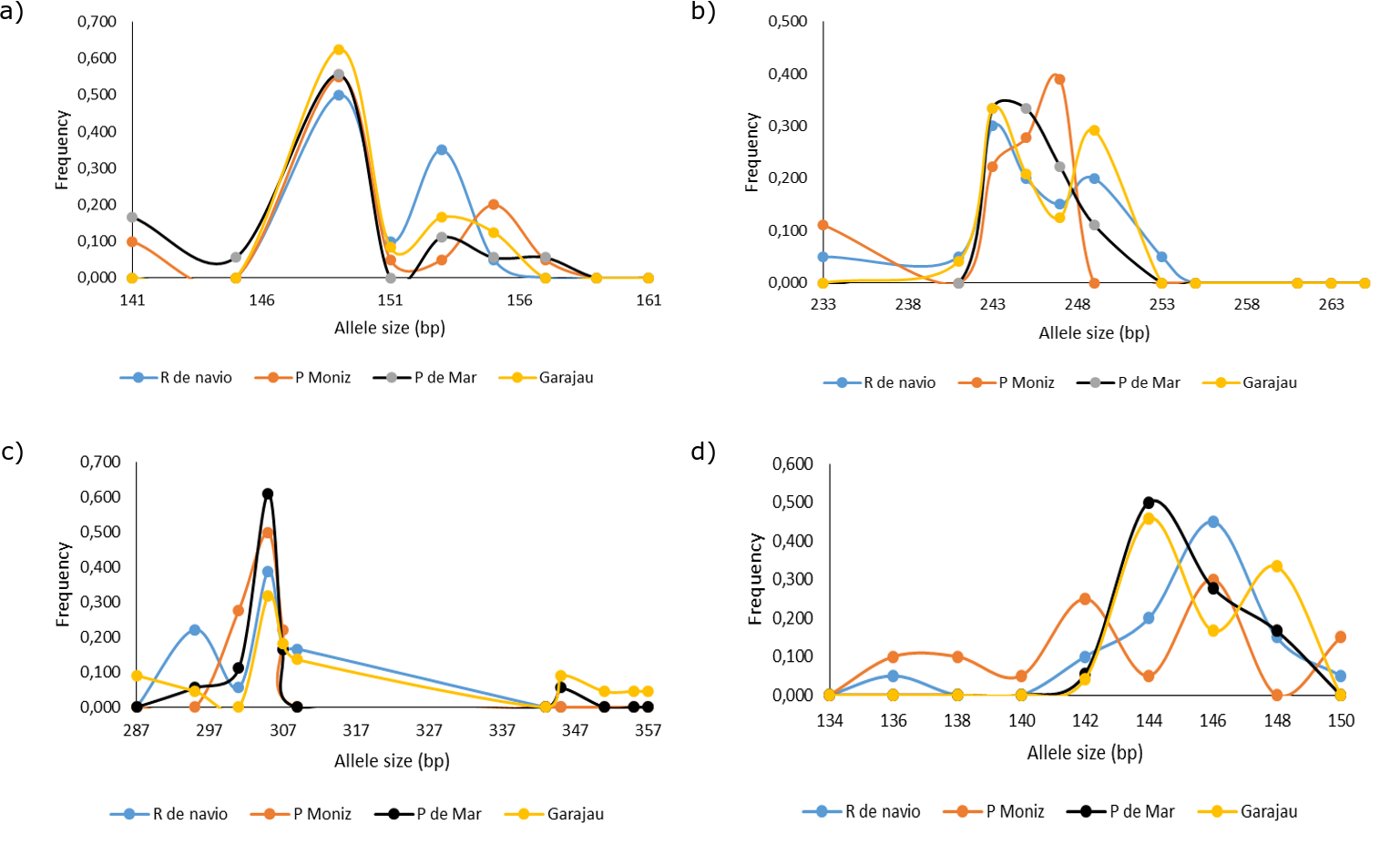


Figure S1. Allele frequencies in loci ASP21F, ASP34F, ASP38F, and ASP17F among four populations in Madeira Island, showing the low number of alleles in Porto Moniz. R. de navio = Rocha do navio, P. Moniz = Porto Moniz, P. de Mar = Paúl do mar.