# Biogeography and Genetic Structure in Populations of a Widespread Lichen (*Parmelina tiliacea*, *Parmeliaceae*, *Ascomycota*)

Jano Núñez-Zapata[¤][☯], Paloma Cubas*[☯], David L. Hawksworth, Ana Crespo

Departamento de Biología Vegetal II, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid, Spain

☯ These authors contributed equally to this work.
¤ Current address: Museo de Zoología, Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México, Mexico D. F., Mexico
* cubas@ucm.es

## Abstract

The genetic diversity and population structure of the foliose lichenized fungus *Parmelina tiliacea* has been analyzed through its geographical range, including samples from Macaronesia (Canary Islands), the Mediterranean, and Eurosiberia. DNA sequences from the nuclear ribosomal internal transcribed spacer, the mitochondrial large subunit ribosomal RNA gene, and the translation elongation factor 1-α were used as molecular markers. The haplotypes of the three markers and the molecular variance analyses of multilocus haplotypes showed the highest diversity in the Canary Islands, while restricted haplotypes occurred at high frequencies in Mediterranean coastal samples. The multilocus haplotypes formed three unevenly distributed clusters (clusters 1-3). In the Canary Islands all the haplotypes were present in a similar proportion, while the coastal Mediterranean sites had almost exclusively haplotypes of cluster 3; cluster 2 predominated in inland Mediterranean sites; and cluster 1 was more abundant in central and northern Europe (Eurosiberian area). The distribution of clusters is partially explained by climatic factors, and its interaction with local spatial structure, but much of the variation remains unexplained. The high frequency of individuals in the Canary Islands with haplotypes shared with other areas suggests that could be a refugium of genetic diversity, and the high frequency of individuals of the Mediterranean coastal sites with restricted haplotypes indicates that gene flow to contiguous areas may be restricted. This is significant for the selection of areas for conservation purposes, as those with most genetic variation may reflect historical factors and biological properties of the species.

## Introduction

Many lichenized fungi have long been considered to have exceptionally wide areas of distribution. However, detailed molecular studies are increasingly showing that, in some cases, different lineages are hidden under one species name, allowing recognition of more than one species [1]. Further, in other cases, a single lineage distributed over a wide area may display major genetic variation and geographical differentiation in the populations [2–3]. Large-scale differences in population structure may be explained by different post-glacial histories of coastal and inland populations [4], persistence in glacial refugia and subsequent genetic drift [5] or southern refugia during the last glaciations and subsequent recolonization into central and northern Europe [6]. In other cases, the genetic structure of the populations has been related to the interaction of climatic and geographical factors, with locally adapted algal partners and isolation by distance of the fungal ones [2, 7–8]. Additionally, for lichens growing mostly on trees, changes in the vegetation may affect the populations, for example by fragmentation of the distributional area of the host tree, either naturally or by human action [9–12].

*Parmelina tiliacea* is a foliose lichen occurring throughout Europe, North Africa, and the Canary Islands, and extending eastward to Asia. We decided to investigate the genetic structure within this species to ascertain if it showed any genetic variation. We hypothesized that this could be the case as the lichen occurs from sea level to 1500 m (and even to 3200 m in India), and under different ecological conditions. It is found on nutrient-rich bark of broad-leaved deciduous or coniferous trees, acid or siliceous rocks, as well as roofing tiles and memorials in well-lit situations. In Europe the species is particularly common around the Mediterranean, especially in central Iberia, extending through central into northern Europe, becoming rare in northern and western Scotland and with scattered localities in Ireland and southern Scandinavia. This species has asexual propagules (isidia, with both the algal and fungal partner), and also reproduces sexually by ascospores, although the spore-producing apothecia are often scarce or absent.

We used samples from a wide geographic range and selected three molecular markers of the fungal partner to build haplotype networks, analyze the population structure through clustering models and molecular variance analyses, and explore the geographic and climatic distribution of the global genetic variation of this widespread species. We did not use material previously recognized as a separate cryptic species, *P. cryptotiliacea* [13], in this study.

The study aimed to: (1) establish whether *P. tiliacea* consists of a single widely distributed lineage or shows any population subdivision; (2) determine if any genetic differentiation found corresponds with different geographical areas or other discontinuities that might be barriers to dispersal; and (3) ascertain if the distribution of any group of panmictic individuals correlates with, for example, macroclimatic variables or spatial trends. The analyses detected differences in haplotype composition between the regions, and allowed recognition of three clusters of populations in *P. tiliacea*. These clusters were partly explained by geography, local spatial structure, and climate variables. We also discuss the possible influence of past climate and vegetation.

## Material and Methods

### Taxon sampling

We studied 364 samples of *Parmelina tiliacea* from 62 localities representing different altitudes and habitats. No specific permissions were required for collecting at these locations because this species is very common and neither endangered nor protected. Sampling sites were from three broad geographical areas: Macaronesian (MA), Mediterranean *s.l.* (sensu [14]) and Eurosiberian (EU). EU included the Alpine, Atlantic, Boreal and Continental biogeographical

regions as defined by [15]. Sites from the most extensively sampled area, the Mediterranean, were separated into two subgroups, one comprising inland (MI: Mediterranean-Inland) and the other coastal localities (MC: Mediterranean-Coastal). Details of localities, voucher specimens, and GenBank accession numbers are given in S1 Table. Nineteen bioclimatic variables of the 62 sampling sites were extracted from the 30 arc seconds resolution (~1 km) layers in the WorldClim repository (http://www.worldclim.org/) with Diva-GIS 7.5 [16].

## Molecular methods

DNA extraction and amplification of the nuclear ribosomal internal transcribed spacer (ITS) and partial large subunit mitochondrial ribosomal rRNA gene (mtLSU) followed the procedures described in [13] with some modifications: primers ITS1F [17] and ITS4A [18] were also used for amplification of ITS; and for the mtLSU region a second protocol was implemented, decreasing the annealing temperature by 1°C per 6 cycles from 63°C to 58°C, followed by 34 cycles of annealing at 40°C for 30 sec. A partial sequence of the translation elongation factor 1-$\alpha$ (*EF1-$\alpha$*) was amplified using primers EFA-F and EFA-R [19], with similar parameters as for ITS but with an annealing temperature of 52 °C. PCR products were cleaned using ExoSAP-IT (USB, Cleveland, Ohio, USA), following manufacturers' instructions. Sequencing was performed using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Life Technologies) with the PCR primers. Sequencing products were electrophoresed on a 3730 DNA analyzer (Applied Biosystems, Life Technologies) at the Unidad de Genómica (Parque Científico de Madrid). The sequence fragments obtained were assembled and manually adjusted in BioEdit 7.0.9.0 [20]. Three separate matrices (one for each locus) were built. Sequences were aligned using the ClustalW Multiple alignment program [21], and manually inspected.

## Individual loci, haplotype networks and multilocus analyses

Measures of genetic diversity were calculated for each sampling site and for the four geographic areas: the number of polymorphic sites (s), number of haplotypes (h), haplotype diversity (Hd) and nucleotide diversity ($\pi$) were estimated with DnaSP v5 [22] for each gene, and Tajima's D tests were performed with the same program to test for selective neutrality [23]. To show the relatedness of haplotypes, we constructed median-joining haplotype networks for the three DNA loci using Hamming distance with the *haplonet* function of the *pegas* package [24], available through the Comprehensive R Archive Network (CRAN) [25]. The distribution of haplotypes with respect to geographical areas was displayed graphically on a network of each locus.

To assess whether the ITS, mtLSU and *EF1-$\alpha$* loci could be combined in a joint analysis, we used the congruence among distance matrices (CADM) test [26] based on Kimura 2-parameter distance matrices [27] with function *CADM.global* implemented in the R library *ape* [28] which calculates the Kendall's coefficient of concordance W [29–30]. The null hypothesis of complete incongruence among DNA distance matrices was tested with 999 permutations in R. Tests to detect recombination events (RDP, GENECONV, CHIMAERA, MAXCHI) were performed for each locus with the options implemented in RDP3 [31], and the Perl script IMgc [32].

A matrix with the multilocus haplotypes (MLHs) of the three combined loci was built for each sampling site and AMOVA analyses were run with Arlequin 3.5.1.3 [33] to estimate the apportionment of genetic variation within and between sites and areas. The analyses run included: (a) all samples; and (b) non-redundant MLHs within sites to correct for any clonal signal. We investigated isolation by distance (IBD) using a Mantel test between MLHs genetic distances ($D_R$, Rogers distance) and geographic distances of the sampling sites with the *mantel*.

*randtest* function of the *adegenet* 1.4–1 package [34–35]. Correlations between these distances, however, can be due to IBD, which results in either continuous clines of genetic differentiation, or to the existence of distant and differentiated populations (distant patches). We plot local densities of distances to disentangle both processes, which were measured using a 2-dimensional kernel density estimation (function *kde2d)* and the results displayed with a customized color palette using *image* in MASS package [36].

## Population structure analyses

Clustering analyses were performed with Structure v2.3 [37] using the matrix of MLHs for inference of the number of clusters (k) and assignation of individual thalli to genepools. We used MLHs instead of sequences to avoid data being dominated by one or a few non- or low-recombining regions [38]. The analyses were run with an admixture model to allow individuals to have ancestry from multiple populations, a uniform prior of individual admixture from all genepools, and independent panmictic genepool allele frequencies [39]. We performed three replicate runs for each value of k (between 1 and 10), with $10^5$ iterations after a burn-in period of $10^4$ iterations. Analyses were run with a data set that included all 364 individuals, and with a collapsed data set excluding redundant haplotypes within sites. The best k value was selected based on the rate of change in the log likelihood of the data between successive k values runs (ΔK) using the methods in [40], through the Structure Harvester portal (http://taylor0.biology.ucla.edu/structureHarvester/). We used Distruct V1.1 [41] to graphically represent the individual assignment output of Structure. Individuals with over 85% probability of ancestry in a given gene pool were regarded as 'assigned' to that gene pool, whereas all other individuals were classified as 'unassigned'. We performed $\chi^2$ tests of heterogeneity to assess whether the observed frequencies of clusters differed significantly from the null expectation of equal frequency across areas.

## Genetic structure, climate and geography

We used redundancy analysis (RDA)-based variation partitioning [42–43] to determine the fraction of variation revealed in the clustered genetic structure of *P. tiliacea* explained by the macroclimate and geography. As descriptors of spatial relationships we used principal coordinates of neighbor matrices (PCNM) obtained by principal coordinate analysis of a truncated matrix of Euclidean (geographic) distances among the sampling sites using the PCNM R package [44]. The PCNMs with positive spatial correlation were retained for redundancy analysis as suggested by [45]. For the RDA analysis we used the proportion of assignment to the clusters of the sites as the response variables, and climatic variables and spatial PCNMs as explanatory variables. Prior to the RDA, the cluster data were Hellinger-transformed [46] using the *decostand* function in the R package *vegan* [47], and climatic and PCNM variables were selected with function *forward.sel* of the *packfor* R package, retaining only those that were significant following the double stopping criterion [48]. The significance of the fractions of interest was estimated with functions *varpart* and *rda* coupled with a permutation test (*anova*) implemented in *vegan* [47].

## Results

We generated 1086 new sequences (348 of ITS, 348 of mtLS, and 364 of *EF1-α*) and downloaded the rest (up to 364 sequences per locus) from GenBank. The alignments included 462 bp from ITS, 716 bp from mtLSU and 546 bp from *EF1-α*. No recombination events were detected in the single matrices with either RPD3 or Imgc programs. Genetic diversity indices obtained for each marker globally and by geographical area are detailed in Table 1. The three loci

**Table 1. Summary of genetic diversity statistics of ITS, mtLSU, and *EF1-α* sequences of *Parmelina tiliacea* in the four geographical areas considered.**

| | | N | s | h | Hd | π | D |
|---|---|---|---|---|---|---|---|
| ITS | Overall | 364 | 26 | 25 | 0.868 | 0.00682 | -0.72235 |
| | MA | 68 | 9 | 7 | 0.801 | 0.00725 | 2.07518* |
| | MI | 111 | 19 | 15 | 0.774 | 0.00518 | -1.14458 |
| | MC | 82 | 14 | 12 | 0.705 | 0.00356 | -1.29077 |
| | EU | 103 | 14 | 10 | 0.783 | 0.00681 | 0.25464 |
| mtLSU | Overall | 364 | 17 | 8 | 0.679 | 0.00657 | 1.93858 |
| | MA | 68 | 16 | 7 | 0.722 | 0.00690 | 1.41370 |
| | MI | 111 | 11 | 3 | 0.548 | 0.00624 | 2.94599** |
| | MC | 82 | 1 | 2 | 0.178 | 0.00025 | -0.13463 |
| | EU | 103 | 12 | 5 | 0.592 | 0.00730 | 3.34958** |
| *EF1-α* | Overall | 364 | 44 | 24 | 0.838 | 0.01846 | 0.99951 |
| | MA | 68 | 35 | 10 | 0.765 | 0.01695 | 0.52846 |
| | MI | 111 | 36 | 14 | 0.739 | 0.01531 | 0.39475 |
| | MC | 82 | 25 | 11 | 0.748 | 0.00842 | -0.57509 |
| | EU | 103 | 30 | 11 | 0.724 | 0.01683 | 1.23510 |

Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas. N: number of individuals, s: number of polymorphic sites, h: number of haplotypes, Hd: Haplotype (gene) diversity, π: nucleotide diversity (average over loci), D: Tajima's D (significance *: $p < 0.05$; **: $p < 0.10$).

doi:10.1371/journal.pone.0126981.t001

were polymorphic, yielding 25 ITS, 8 mtLSU, and 24 *EF1-α* haplotypes. ITS and *EF1-α* had similar values of haplotype diversity (Hd) in all samples and in the different areas, while mtLSU values were lower and heterogeneous in the different areas. The lowest values of nucleotide diversity for the three loci were in the MC area. Globally, the three loci appeared to be selectively neutral according to Tajima's D values. The significant positive D value for the ITS from the MA samples and the mtLSU from the MI and EU samples might suggest sudden population contraction or balancing selection at these loci. The frequency of the haplotypes in the four regions, a summary of the number of restricted haplotypes (i.e. haplotypes found only in a single area) for each marker, and the percentage of samples with those haplotypes, are shown in Table 2 and S2 Table.

In the ITS network (Fig 1) the four most frequent haplotypes (2, 3, 4, and 6) were connected by short branches, one to two changes in length. Many rare or restricted haplotypes were connected peripherally to the predominant ones. One ITS haplotype (15) was restricted to, and very frequent in, the MC area. In the mtLSU network, two very common haplotypes (2 and 3)

**Table 2. Summary of the distribution of haplotypes restricted to a single area (restricted haplotypes).**

| | Percentage of restricted haplotypes/number of haplotypes in the area. | | | | Percentage of individuals with restricted haplotypes/individuals in the area | | | |
|---|---|---|---|---|---|---|---|---|
| | ITS | mtLSU | *EF1-α* | MLHs | ITS | mtLSU | *EF1-α* | MLHs |
| MA | 14.3 | 42.8 | 30 | 67.6 | 1.5 | 5.9 | 4.4 | 5.9 |
| MI | 40 | 0 | 28.6 | 57.1 | 5.4 | 0 | 9.9 | 21.6 |
| MC | 50 | 0 | 36.6 | 55 | 48.8 | 0 | 13.1 | 51.2 |
| EU | 30 | 20.0 | 18.12 | 55 | 14.6 | 0.9 | 18.2 | 13.6 |

Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas.
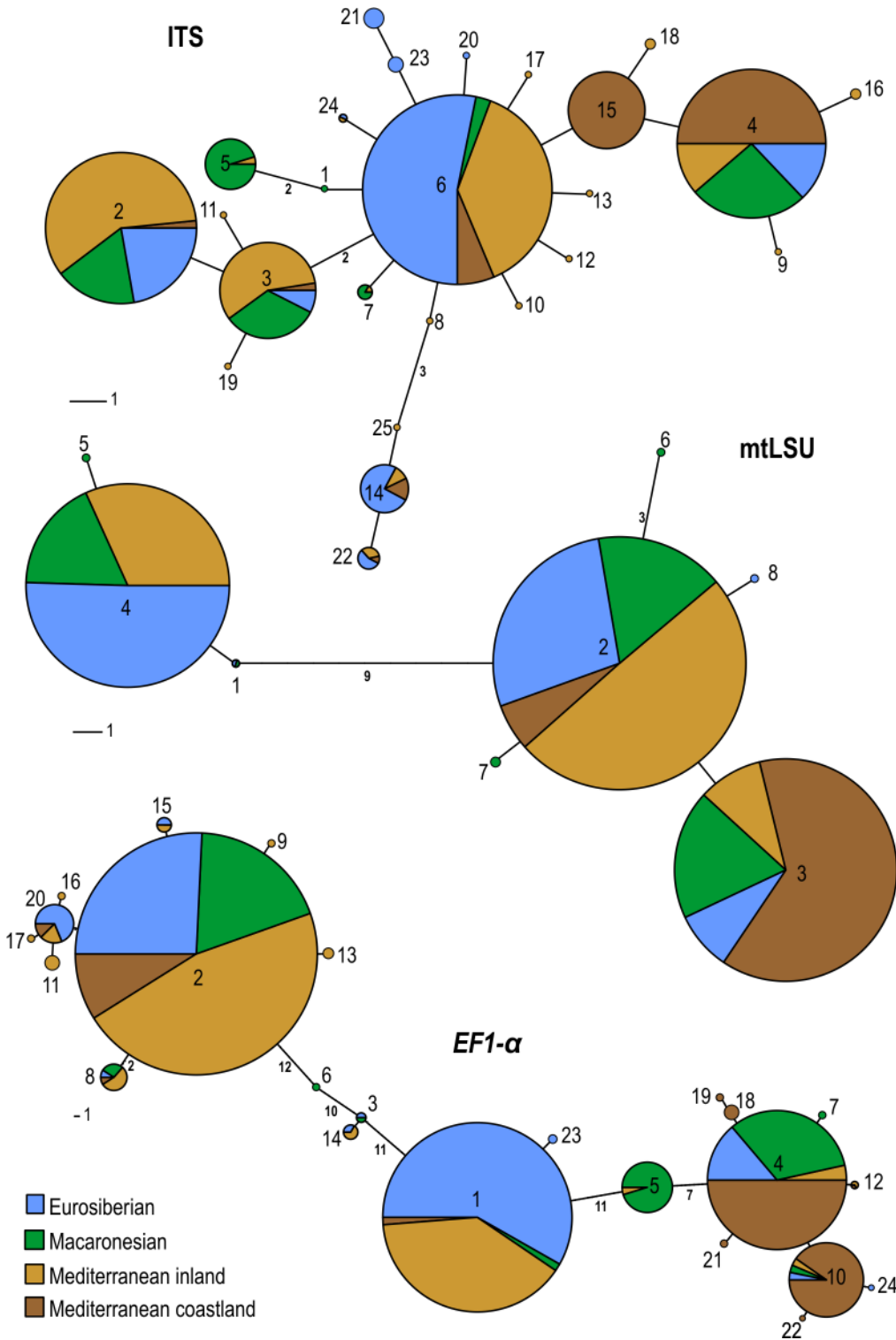
doi:10.1371/journal.pone.0126981.t002

**Fig 1. Median joining networks of three fungal markers.** Pie charts numbered with haplotype codes as in S2 Table, proportional in size to the number of samples containing this haplotype, and colored according to the geographical area. Small numbers on lines indicate branch lengths.

doi:10.1371/journal.pone.0126981.g001

**Table 3. Analysis of molecular variance of multilocus haplotypes (MLHs) in *Parmelina tiliacea*.**

| | Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation | Fixation indexes | *p* |
|---|---|---|---|---|---|---|---|
| Four regions, all samples | Between regions | 3 | 69.754 | 0.211 | 16.84 | 0.168 | <0.001 |
| | Among sites within regions | 58 | 181.093 | 0.444 | 35.39 | 0.426 | <0.001 |
| | Within sites | 302 | 181.093 | 0.600 | 47.78 | 0.522 | <0.001 |
| | Total | 363 | 432.827 | 1.255 | | | |
| Four regions, redundant haplotypes excluded | Between regions | 3 | 30.318 | 0.194 | 15.27 | 0.153 | <0.001 |
| | Among sites within regions | 58 | 72.157 | 0.088 | 6.91 | 0.019 | <0.001 |
| | Within sites | 124 | 122.433 | 0.987 | 77.82 | 0.222 | <0.001 |
| | Total | 184 | 224.909 | 1.269 | | | |
| Mediterranean regions (MI vs MC), all samples | Between regions | 1 | 43.253 | 0.423 | 30.51 | 0.305 | <0.001 |
| | Among sites within areas | 37 | 84.891 | 0.350 | 25.24 | 0.363 | <0.001 |
| | Within sites | 154 | 94.510 | 0.614 | 44.25 | 0.558 | <0.001 |
| | Total | 192 | 222.653 | 1.387 | | | |
| Mediterranean regions (MI vs MC), redundant haplotypes excluded | Between regions | 1 | 17.909 | 0.318 | 23.64 | 0.236 | <0.001 |
| | Among sites within areas | 37 | 42.176 | 0.065 | 4.83 | 0.063 | 0.047 |
| | Within sites | 69 | 66.433 | 0.963 | 71.53 | 0.285 | <0.001 |
| | Total | 107 | 126.519 | 1.346 | | | |

Genetic variation was partitioned among sites and regions, including all the samples or excluding redundant haplotypes within sites.

doi:10.1371/journal.pone.0126981.t003

were closely connected, while another frequent one (4) was separated by a long branch of nine changes. Haplotype 4 was not found in MC area, but was frequent in the others, and haplotype 3 predominated in MC. In the *EF1-α* network the dominant haplotypes were separated by very long branches with rarer haplotypes intermingled between them. Haplotype 10 was almost exclusive to MC. Overall, 4 ITS haplotypes, 2 mtLSU, and 5 *EF1-α* occurred in all the geographical areas, and 16 ITS, 4 mtLSU, and 13 *EF1-α* haplotypes were found in only a single area (S2 Table). In MC there was a high proportion of restricted ITS and *EF1-α* haplotypes, and no restricted mtLSU ones. In contrast, the MA area had a high proportion of mtLSU restricted haplotypes.

Kendall's coefficient of concordance (W) rejected the incongruence among the three matrices, globally (W = 0.5462059, *p* = 0.03) and by pairs, so we combined the loci into a multilocus haplotype (MLHs) matrix. Of the 78 MLHs found (S2 Table), only three (2, 3, and 4) were present in all the areas, and most were restricted to one. Two MLHs were very frequent and restricted: (1) MLH 5 was mostly restricted to MA; and (2) MLH 51 (19 samples) was confined to MC. The AMOVA of the MLHs matrix (Table 3) indicated little structure between the regions ($\Phi_{RT}$ = 0.17) when considering all four areas, but a very substantial structure ($\Phi_{RT}$ = 0.31) when MI sites were compared with MC ones. MC also had the highest values of pairwise differences (FSTs) against the MA, MI, and EU (0.20, 0.32 and 0.32, respectively; Table 4). Similar results were obtained when redundant haplotypes were excluded (0.12, 0.24 and 0.27, respectively).

A weak pattern indicating isolation by distance between populations was detected in Mantel random tests with all samples (r = 0.111, p = 0.003) and without redundant haplotypes (r = 0.09987512, p = 0.019). However, the scatterplots of local densities of distances (Fig 2)

**Table 4. Pairwise differences (FSTs) among areas.**

|  | Area | MA | MI | MC |
|---|---|---|---|---|
| All samples | MI | 0.10965 |  |  |
|  | MC | 0.20201 | 0.32323 |  |
|  | EU | 0.14049 | 0.05786 | 0.31819 |
| Redundant haplotypes excluded | MI | 0.04685 | - |  |
|  | MC | 0.12280 | 0.23890 |  |
|  | EU | 0.11596 | 0.06612 | 0.27459 |

Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas.

doi:10.1371/journal.pone.0126981.t004

detected some degree of discontinuity. This indicated that the genetic differentiation of the populations showed an incipient patched pattern and did not form a continuous cline.

## Structure of populations

Structure analyses based on the distribution of the MLHs of the different sites showed that the best estimate for the number of clusters was k = 3. Similar groups were obtained when the analysis was run excluding redundant MLHs from each site. The proportion of the three clusters at the different sites is shown in Fig 3 and S3 Table. The distribution of the clusters across the geographical areas was heterogeneous (Table 5 and Fig 4). In the Canary Islands (MA), the three clusters were in similar proportions; at MI sites, especially in Iberia, the clusters were spatially intermingled, although cluster 2 had a higher frequency than the two other clusters; in EU, cluster 1 predominated but was intermingled with cluster 2; and most MC sites had cluster 3, except for some eastern ones where cluster 2 was also found. This last result must, however, be accepted with caution as we had few eastern Mediterranean samples. Of the 364 individuals, 357 were unequivocally assigned to a single cluster with a probability $\geq$ 85%, while seven showed admixed ancestry (Fig 5). Similar results were found when redundant haplotypes were excluded from the analysis (Table 5). In both analyses the hypothesis of independence between areas and clusters was rejected (p$\leq$ 0.0001).

## Relative contribution of climate and geography to the genetic variation

A preliminary redundancy analysis (RDA) of the cluster data showed that the X and Y geographical coordinates were not significant, whereas three PCNM variables (accounting for fine scale patterns) had a positive Moran's I index of spatial autocorrelation at low distances (Fig 6). Five bioclimatic variables were also found to be significant and selected for the analysis: BIO3 (isothermality), BIO6 (minimum temperature of the coldest month), BIO9 (mean temperature of the driest quarter), BIO13 (precipitation of the wettest month), and BIO16 (precipitation of the wettest quarter). The relative contribution of climate and fine spatial distribution to the variance of the clusters is shown in Table 6. Climate accounted for 19% of the variation of the cluster structure, and the interaction of fine scale patterns and climate for another 17%. A similar analysis using the genetic cluster structure without redundant haplotypes at the different sites (Table 7) was undertaken to exclude the clonal signal. In this analysis four climatic variables (BIO3, BIO6, BIO9 and BIO13) and five PCNM were used. The proportion of variation related to climate was slightly lower (15%) and the interaction of climate with geography slightly higher (19%). Both analyses indicated that climate and the interaction between climate and local spatial patterns explained about 35% of the genetic variation.
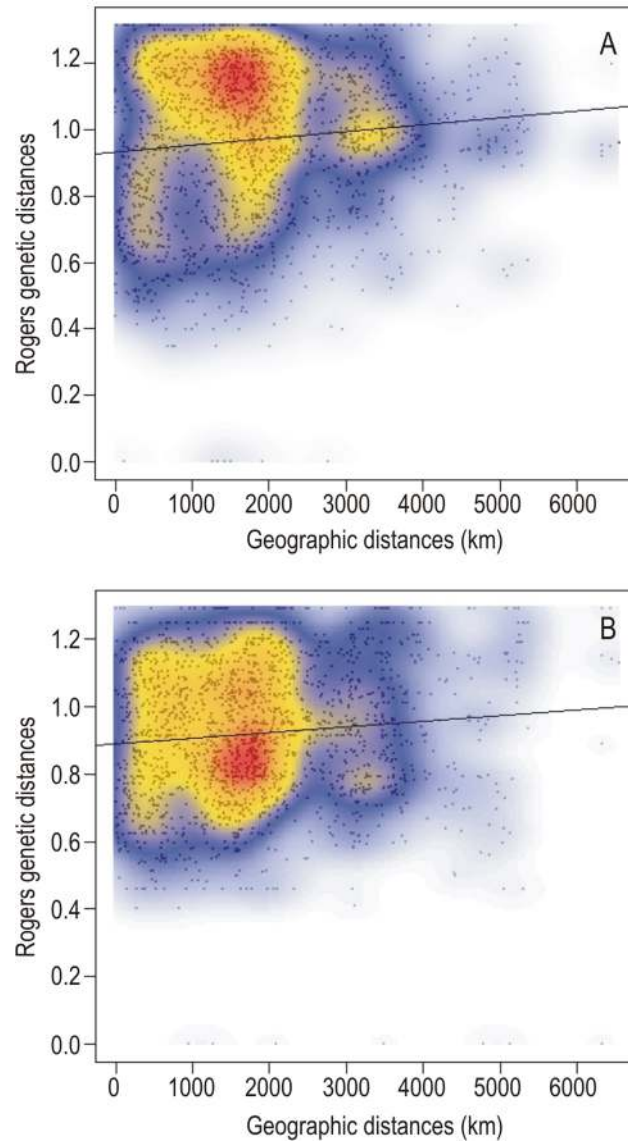
**Fig 2. Isolation by distance plot.** Scatterplots of local densities of Rogers genetic distances of MLHs against pairwise geographical distances of the sites. A) Including all samples per site, adjusted $R^2$ = 0.01193, $p < 0.001$; and B) excluding redundant haplotypes per site, adjusted $R^2$ = 0.00945, $p < 0.001$.

doi:10.1371/journal.pone.0126981.g002

## Discussion

The three molecular markers of the fungal partner we used (ITS, mtLSU, and *EF1-α*) demonstrated genetic variability within *Parmelina tiliacea*. The ITS region is frequently used for phylogenetic studies in fungi and it has been adopted as a universal DNA barcode marker for species [49], although it does not discriminate between all taxa [50–51]. In our case, the ITS marker detected intraspecific genetic differences even at the same locality. Some ITS haplotypes were frequently found in well separated areas. However, in MC up to 50% of the haplotypes were restricted to this zone, and one of those was very frequent. Conversely, in MA most of the haplotypes were shared with other regions. The mtLSU showed lower variability than the ITS, as has been observed in some other lichen-forming fungi [2, 5]. In *P. tiliacea* this marker
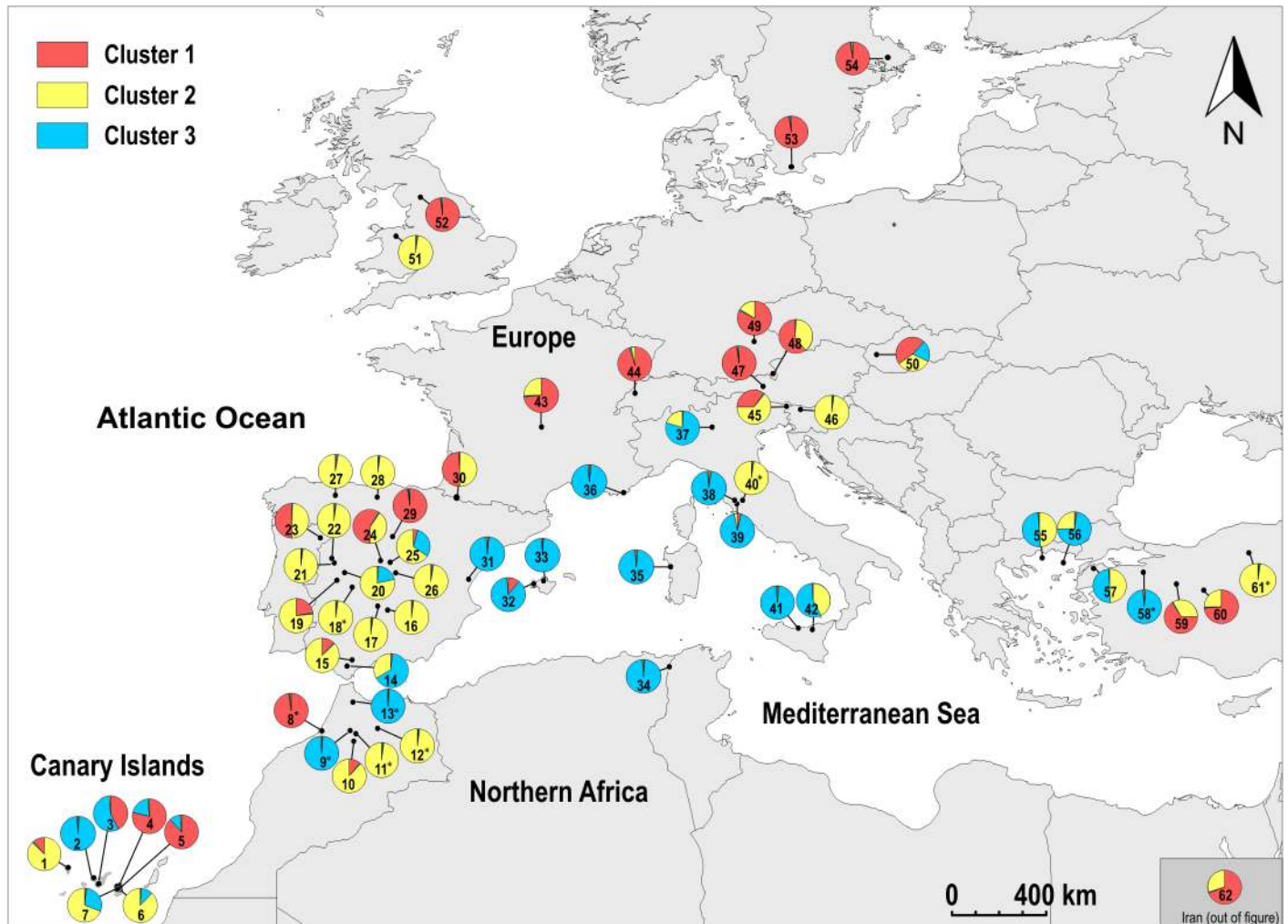
**Fig 3. Distribution of the clusters at each site.** Proportion of membership of each site in the three clusters inferred by Structure. Collection sites numbered as in S1 Table. Localities where only one individual was studied are marked with an asterisk.

doi:10.1371/journal.pone.0126981.g003

detected only three main haplotypes, one absent and another predominant in the MC area, as well as many restricted haplotypes at low frequencies in MA. The *EF1-α* marker, which has not often been used in studies of the population genetics of lichenized fungi [52–53], showed a

**Table 5. Number of individuals assigned to each genetic cluster in each area.**

| | All individuals | | | | Redundant haplotypes excluded | | | |
|---|---|---|---|---|---|---|---|---|
| | Cluster 1 | Cluster 2 | Cluster 3 | Total | Cluster 1 | Cluster 2 | Cluster 3 | Total |
| MA | 22 | 24 | 22 | 68 | 5 | 11 | 9 | 25 |
| MI | 34 | 66 | 9 | 109 | 14 | 41 | 6 | 61 |
| MC | 0 | 8 | 72 | 80 | 0 | 7 | 36 | 43 |
| EU | 55 | 36 | 9 | 100 | 27 | 20 | 2 | 49 |
| Total | 111 | 134 | 112 | 357 | 46 | 79 | 53 | 178 |

Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas.

doi:10.1371/journal.pone.0126981.t005
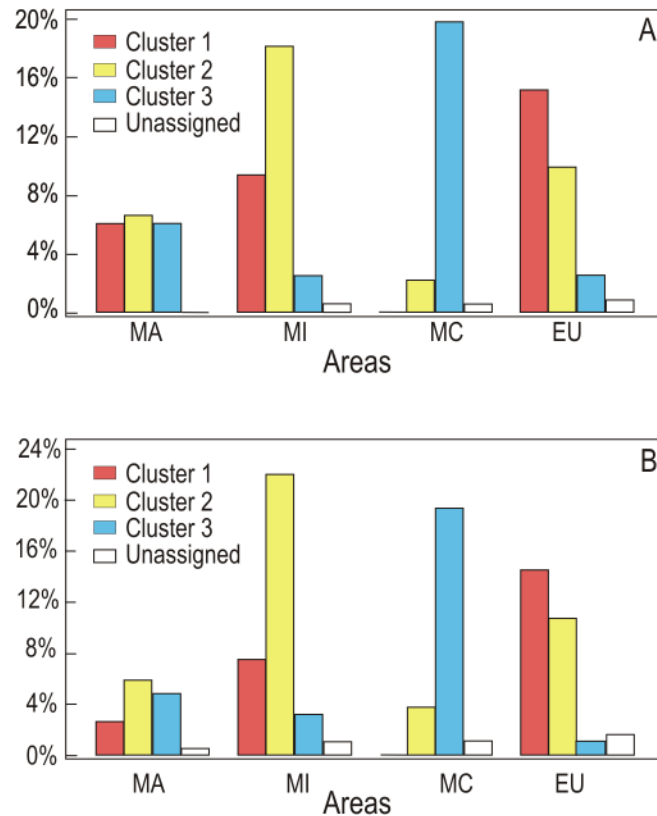
Fig 4. Percentage of the 357 individuals assigned to the different clusters and areas Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas. A) Including all samples per site, and B) excluding redundant haplotypes per site.

high level of variability. Here again, the MC area showed a higher proportion of restricted haplotypes. The three networks had haplotypes present in all regions, and in addition others which were either restricted to, or unevenly distributed amongst, the regions. Further, MC had a high proportion of restricted ITS and *EF1-α* haplotypes, but none of mtLSU. In contrast, MA had a high proportion of mtLSU restricted haplotypes.

The combined MLHs also indicated that the Canary Islands (MA) had the highest number of restricted and rare MLHs, but most of the individuals had haplotypes widespread in other areas. On the contrary, the high number of exclusive MLHs found in the coastal Mediterranean sites (MC) were present in many individuals. Although the relative contribution of asexual and sexual processes in *P. tiliacea* remains unknown, the three markers detected genetic variability between and within populations. This would be expected in an organism with at least occasional sexual reproduction. The asexual propagules (isidia) can be expected to be dispersed within short distances from the parents as in other lichens [54] forming clonal populations. This is indicated by the number of redundant haplotypes found at some sites.

*Parmelina tiliacea* were subdivided into three clusters with heterogeneous distributions between regions, the extreme cases being the Canary Islands (MA) and the coastal sites (MC) around the Mediterranean Basin (as opposed to more continental ones). In the Canary Islands (MA area), the three genetic clusters occurred in similar proportions, and there were also many restricted MLHs at very low frequencies, most individuals sharing MLHs with other areas. Our data indicate that these islands retain the greatest amount of diversity in this species and
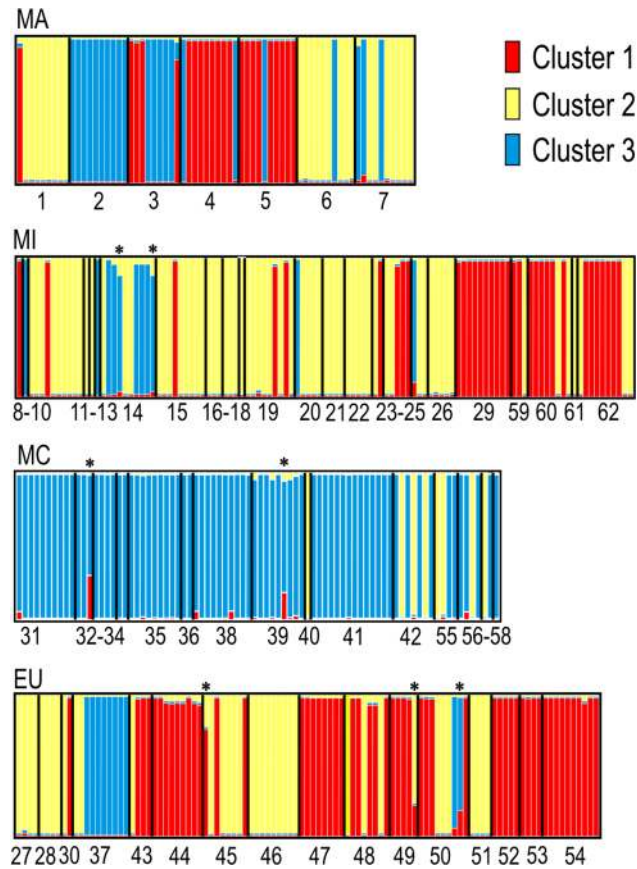
**Fig 5. Inferred ancestry of individuals of *Parmelina tiliacea* from each site grouped by geographical areas.** Collection sites numbered as in S1 Table. Each individual is represented by a single vertical line, broken into colored segments with lengths proportional to each of the three inferred clusters. Admixed individuals marked with an asterisk.

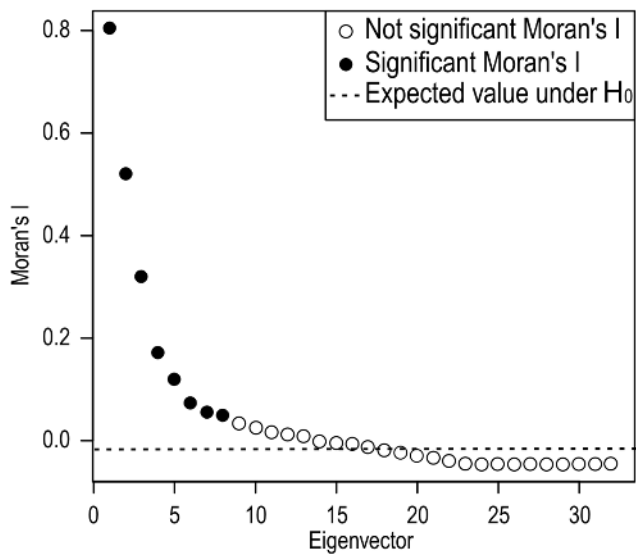doi:10.1371/journal.pone.0126981.g005



**Fig 6. Moran's I index of the cluster structure of *Parmelina tiliacea*.**

doi:10.1371/journal.pone.0126981.g006

**Table 6. Summary of variance partitioning analysis of cluster structure including all samples.**

| | | d.f. | $R^2$ | Adj $R^2$ | Pr(>F) |
|---|---|---|---|---|---|
| Interactions | Climate+Interaction | 5 | 0.41679 | 0.36472 | — |
| | Geographyl+Interaction | 3 | 0.23370 | 0.19406 | — |
| | Climate+Geography+Interaction | 8 | 0.46652 | 0.38600 | |
| Individual fractions | Climate | 5 | — | 0.19194 | 0.005[**] |
| | Interaction | 0 | — | 0.17278 | — |
| | Geography | 3 | — | 0.02128 | 0.115 |
| | Residuals | | | 0.61400 | — |

Two explanatory matrices were used with five climatic variables (BIO3, BIO6, BIO9, BIO13, and BIO16) and three spatial (PCNM) variables. Total variation (SS): 24.327. Variance: 0.3988. Significance codes: **$\leq$ 0.01, $\cdot$ $\leq$0.1

doi:10.1371/journal.pone.0126981.t006

suggest that they may have served as a refugium for it. The Canary Islands, as other oceanic islands at low latitudes, are generally considered to have been buffered from the pronounced climate changes of the last 2 Ma [55], and to have served as a refugium and a source area for the postglacial recolonization of Europe by spore-producing plants (e.g. [56]).

In the coastal Mediterranean sites (MC), however, in all those around the Mediterranean Basin, cluster 3 was dominant and a very high proportion of the individuals had MLHs restricted to this area. This may indicate a geographic barrier for dispersal, as seems to occur with the Pyrenees and the Alps in the case of some plants and insects [57–58]. These differences could, however, also be a response to competition with other species, and the persistence (or colonization) of only the best adapted haplotypes to this region. It is conceivable that the eastern Mediterranean area could have been another refugium, as in plants [59–60], but more extensive sampling within that region would be needed to test that hypothesis.

The geographical genetic differentiation found cannot be explained by isolation by distance (IBD) alone, although a weak pattern of spatial autocorrelation at low distances was detected. Part of the variation of the cluster distribution can be explained by the macroclimatic variables alone, independently of any spatial structure. Another fraction reflects the interaction between climate and the local spatial structure generated by other natural processes, factors that are often mixed together. Both variables have been found to explain population structure in other lichens [2, 7]. Nevertheless, there is still a large percentage of the variation discovered in *P. tiliacea* that remains unexplained, and which could perhaps be related to intrinsic biological properties or ecological requirements of the species. Factors related to the historical biogeography

**Table 7. Summary of variance partitioning analysis of cluster structure excluding redundant haplotypes.**

| | | d.f. | $R^2$ | Adj $R^2$ | Pr(>F) |
|---|---|---|---|---|---|
| Interactions | Climate+Interaction | 4 | 0.38377 | 0.34053 | — |
| | Geographyl+Interaction | 5 | 0.29843 | 0.23579 | — |
| | Climate+Geography+Interaction | 9 | 0.47427 | 0.38327 | |
| Individual fractions | Climate | 4 | — | 0.14748 | 0.005[**] |
| | Interaction | 0 | — | 0.19305 | — |
| | Geography | 5 | — | 0.04275 | 0.077 [.] |
| | Residuals | | | 0.61673 | |

Two explanatory matrices were used with four climatic variables (BIO3, BIO6, BIO9, BIO13) and five spatial variables (PCNM). Total variation (SS): 22.878. Variance: 0.37505. Significance codes: **$\leq$ 0.01, $\cdot$ $\leq$0.1

doi:10.1371/journal.pone.0126981.t007

of the species (particularly bottlenecks, founder effect, or glacial refugial areas) could also have been involved in the initiation of differentiation in this widespread species.

The macroclimatic variables we used are only rough approximations for the ecological requirements of the individuals, as they do not take into account either local microclimatic conditions that may be important for the establishment of lichens [61], or past climatic events that permitted the colonization of an area. Considering the estimated age (8.45 Ma) for the origin of the genus *Parmelina* [62], the diversification of *P. tiliacea* must be a relatively young process. Thus, the present structure of the continental populations (i.e. the non Macaronesian sites) could well have been influenced by Miocene and Pleistocene events, as suggested for some other organisms [57] and also lichens [63].

Our study highlights the special position of the *Parmelina tiliacea* populations in the Canary Islands and in the Mediterranean coastal sites in relation to the differentiation of lineages within species and the conservation of its genetic diversity. These two areas serve as significant reservoirs of its genetic diversity and so merit preservation. More discriminating markers and more detailed information on microclimatic variables are needed to establish the contribution of clonal *vs* sexual reproduction to the fine-scale structure of the populations. The contribution of the algae component to the establishment of a particular fungal genotype in a location has been demonstrated in other species [8], but this was beyond the scope of the present paper. It is, nevertheless, another factor which merits exploration in a future study.

## Supporting Information

**S1 Table. Locality information with herbarium codes and GenBank accessions of the samples used in the present study.**
(PDF)

**S2 Table. Frequency of the haplotyes of ITS, mtLSU, *EF1-α*, and MLHs. Macaronesian (MA), Mediterranean inland (MI) and coastal (MC), and Eurosiberian (EU) areas.** Haplotypes exclusive of a single area in bold. PH: private haplotypes/number of haplotypes in the area. SPH: samples with private haplotype/number of samples in the area.
(PDF)

**S3 Table. Proportion of membership of each locality in clusters 1 to 3 inferred by Structure.** The analyses were run with an admixture model and uncorrelated loci. Locality codes as in S1 Table and Fig 1; n = number of individuals.
(PDF)

## Acknowledgments

## Author Contributions

Conceived and designed the experiments: JNZ PC AC. Performed the experiments: JNZ. Analyzed the data: JNZ PC. Contributed reagents/materials/analysis tools: JNZ PC DLH AC. Wrote the paper: JNZ PC AC DLH.

## References

1. Crespo A, Lumbsch HT. Cryptic species in lichen-forming fungi. IMA Fungus 2010; 1: 167–170 PMID: 22679576

2. Fernández-Mendoza F, Domaschke S, García MA, Jordan P, Martín MP, Printzen C. Population structure of mycobionts and photobionts of the widespread lichen *Cetraria aculeata*. Mol Ecol 2011; 20: 1208–1232. doi: 10.1111/j.1365-294X.2010.04993.x PMID: 21324011

3. Widmer I, Dal Grande F, Excoffier L, Holderegger R, Keller C, Mikryukov VS, et al. European phylogeography of the epiphytic lichen fungus *Lobaria pulmonaria* and its green algal symbiont. Mol Ecol 2012; 21: 5827–5844. doi: 10.1111/mec.12051 PMID: 23094600

4. Walser JC, Holderegger R, Gugerli F, Hoebee SE, Scheidegger C. Microsatellites reveal regional population differentiation and isolation in *Lobaria pulmonaria*, an epiphytic lichen. Mol Ecol 2005; 14: 457–467. PMID: 15660937

5. Printzen C, Ekman S, Tønsberg T. Phylogeography of *Cavernularia hultenii*: evidence of slow genetic drift in a widely disjunct lichen. Mol Ecol 2003; 12: 1473–1486. PMID: 12755876

6. Arnerup J, Högberg N, Thor G. Phylogenetic analysis of multiple loci reveal the population structure within *Letharia* in the Caucasus and Morocco. Mycol Res 2004; 108: 311–316. PMID: 15185981

7. Sork VL, Werth S. Phylogeography of *Ramalina menziesii*, a widely distributed lichen-forming fungus in western North America. Mol Ecol 2014; 23: 2326–2339. doi: 10.1111/mec.12735 PMID: 24673777

8. Werth S, Sork VL. Ecological specialization in *Trebouxia* (Trebouxiophyceae) photobionts of *Ramalina menziesii* (Ramalinaceae) across six range-covering ecoregions of Western North America. Am J Bot 2014; 101: 1127–1140. PMID: 25016008

9. Printzen C, Lumbsch HT, Schmitt I, Feige GB. A study on the genetic variability of *Biatora helvola* Körb. ex Hellb. using RAPD markers. Lichenologist 1999; 31: 491–499.

10. Werth S, Wagner HH, Holderegger R, Kalwij JM, Scheidegger C. Effect of disturbances on the genetic diversity of an old-forest associated lichen. Mol Ecol 2006; 15: 911–921. PMID: 16599956

11. Werth S, Gugerli F, Holderegger R, Wagner HH., Csencsics D, Scheidegger C. Landscape-level gene flow in *Lobaria pulmonaria*, an epiphytic lichen. Mol Ecol 2007; 16: 2807–2815. PMID: 17594449

12. Scheidegger C, Bilovitz PO, Werth S, Widmer I, Mayrhofer H. Hitchhiking with forests: population genetics of the epiphytic lichen *Lobaria pulmonaria*. Ecol Evol 2012; 2: 2223–2240. doi: 10.1002/ece3.341 PMID: 23139881

13. Núñez-Zapata J, Divakar PK, Del-Prado R, Cubas P, Crespo A. Conundrums in species concepts: the discovery of a new cryptic species segregated from *Parmelina tiliacea* (Ascomycota: Parmeliaceae). Lichenologist 2011; 43: 603–616.

14. Quézel P, Barbero M. Definition and characterization of Mediterranean-type ecosystems. Ecol Mediterr 1982; 8: 15–29.

15. European Environment Agency. Europe 2011—Rev. 1—The biogeographical regions dataset contains the official delineations used in the Habitats Directive (92/43/EEC) and for the EMERALD Network set up under the Convention on the Conservation of European Wildlife and Natural Habitats (Bern Convention)". Available: http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-1.

16. Hijmans RJ, Guarino L, Mathur P. DIVA-GIS Version 7.5, A geographic information system for the analysis of species distribution data. 2012. Manual available at http://www.diva-gis.org.

17. Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes—application to the identification of mycorrhizae and rusts. Mol Ecol 1993; 2: 113–118. PMID: 8180733

18. White TJ, Bruns T, Lee S, Taylor JW. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR Protocols: A Guide to Methods and Applications. New York: Academic Press; 1990. pp. 315–322,

19. Johannesson HS, Johannesson KHP, Stenlid J. Development of primer sets to amplify fragments of conserved genes for use in population studies of the fungus *Daldinia loculata*. Mol Ecol 2000; 9: 375–377. PMID: 10736039

20. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999; 41: 95–98.

21. Thompson JD, Higgins DG, Gibson TJ. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weigh matrix choice. Nucleic Acids Res 1994; 22: 4673–4680. PMID: 7984417

22. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 2009; 25:1451–1452. doi: 10.1093/bioinformatics/btp187 PMID: 19346325

23. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 1989; 123:585–595. PMID: 2513255

24. Paradis E. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 2010; 26:419–420 doi: 10.1093/bioinformatics/btp696 PMID: 20080509

25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014. Available: http://www.R-project.org/.

26. Campbell V, Legendre P, Lapointe F-J. The performance of the Congruence Among Distance Matrices (CADM) test in phylogenetic analysis. BMC Evol Biol 2011; 11: 64. doi: 10.1186/1471-2148-11-64 PMID: 21388552

27. Kimura M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 1980; 16: 111–120. PMID: 7463489

28. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 2004; 20: 289–290. PMID: 14734327

29. Kendall MG, Babington Smith B. The problem of m rankings. Ann Math Statist 1939; 10: 275–287.

30. Legendre P. Coefficient of concordance. In: Salkin NJ, editor. Encyclopedia of Research Design. Los Angeles; SAGE Publications. 2010; 1: 164–169.

31. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 2010; 26: 2462–2463. doi: 10.1093/bioinformatics/btq467 PMID: 20798170

32. Woerner AE, Cox MP, Hammer MF. Recombination-filtered genomic datasets by information maximization. Bioinformatics 2007; 23: 1851–1853. http://hammerlab.biosci.arizona.edu/imgc_online.html. PMID: 17519249

33. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010; 10: 564–567. doi: 10.1111/j.1755-0998.2010.02847.x PMID: 21565059

34. Jombart T. *adegenet*: a R package for the multivariate analysis of genetic markers. Bioinformatics 2008; 24: 1403–1405. doi: 10.1093/bioinformatics/btn129 PMID: 18397895

35. Jombart T, Ahmed I. adegenet 1.3–1: new tools for the analysis of genome-wide SNP data. Bioinformatics 2011; 27: 3070–3071. doi: 10.1093/bioinformatics/btr521 PMID: 21926124

36. Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002.

37. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000; 155: 945–959. PMID: 10835412

38. Pritchard JK, Wena X, Falush D. Documentation for Structure software: Version 2.3. 2010. Available: http://pritch.bsd.uchicago.edu/structure.html.

39. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003; 164: 1567–1587. PMID: 12930761

40. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 2005; 14: 2611–2620. PMID: 15969739

41. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 2004; 4: 137–138.

42. Borcard D, Legendre P, Drapeau P. Partialling out the spatial component of ecological variation. Ecology 1992; 73: 1045–55.

43. Peres-Neto PR, Legendre P, Dray S, Borcard D. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 2006; 87: 2614–2625. PMID: 17089669

44. Borcard D, Legendre P. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecol Modell 2002; 153: 51–68.

45. Borcard D, Legendre P, Avois-Jacquett C, Tuonisto H. Dissecting the spatial structure of ecological data at multiple scales. Ecology 2004; 85: 1826–1832.

46. Legendre P, Gallagher ED. Ecologically meaningful transformations for ordination of species data. Oecologia 2001; 129: 271–280.

47. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. R package version 2.0–5. 2012. Available: http://CRAN.R-project.org/package = vegan.

48. Blanchet F., Legendre P, Borcard D. Forward selection of explanatory variables. Ecology 2008; 89: 2623–2632. PMID: 18831183

49. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, André Levesque C, et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci U S A 2012; 109: 6241–6246. doi: 10.1073/pnas.1117018109 PMID: 22454494

50. Samson RA, Visagie CM, Houbraken J, Hong S-B, Hubka V, Klaassen CHW, et al. Phylogeny, identification and nomenclature of the genus *Aspergillus*. Stud Mycol 2014; 78: 141–173. doi: 10.1016/j.simyco.2014.07.004 PMID: 25492982

51. Visagie CM, Houbraken J, Frisvad JC, Hong S- B, Klaassen CHW, Perrone G, et al. Identification and nomenclature of the genus *Penicillium*. Stud Mycol 2014; 78: 343–371. doi: 10.1016/j.simyco.2014.09.001 PMID: 25505353

52. Yahr R, Vilgalys R, DePriest PT. Geographic variation in algal partners of *Cladonia subtenuis* (Cladoniaceae) highlights the dynamic nature of a lichen symbiosis. New Phytol 2006; 171, 847–860. PMID: 16918555

53. Werth S, Sork VL. Local genetic structure in a North American epiphytic lichen, *Ramalina menziesii* (Ramalinaceae). Am J Bot 2008; 95: 568–576. doi: 10.3732/ajb.2007024 PMID: 21632383

54. Walser JC. Molecular evidence for limited dispersal of vegetative propagules in the epiphytic lichen *Lobaria pulmonaria*. Am J Bot 2004; 91: 1273–1276. doi: 10.3732/ajb.91.8.1273 PMID: 21653485

55. Axelrod DI. Evolution and biogeography of Madrean–Tethyan sclerophyll vegetation. Ann Mo Bot Gard 1975; 62: 280–334.

56. Hutsemékers V, Szövényi P, Shaw AJ, González-Mancebo J-M, Muñoz J, Vanderpoorten A. Oceanic islands are not sinks of biodiversity in spore-producing plants. Proc Natl Acad Sci U S A 2011; 108: 18989–18994 doi: 10.1073/pnas.1109119108 PMID: 22084108

57. Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. Comparative phygeographic and postglacial colonization in Europe. Mol Ecol 1998; 7: 453–464. PMID: 9628000

58. Mende M, Biström O, Meichssner E, Kölsch G. The aquatic leaf beetle *Macroplea mutica* (Coleoptera: Chrysomelidae) in Europe: Population structure, postglacial colonization and the signature of passive dispersal. Eur J Entomol 2010; 107: 101–113.

59. Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, et al. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For Ecol Manage 2002; 156: 49–74.

60. Médail F, Diadema K. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. J Biogeogr 2009; 36: 1333–1345.

61. Pearson LC. Influence of temperature and humidity on distribution of lichens in a Minnesota bog. Ecology 1969; 50: 740–746.

62. Amo de Paz G, Cubas P, Divakar PK, Lumbsch HT, Crespo A. Origin and diversification of major clades in parmelioid lichens (Parmeliaceae, Ascomycota) during the Paleogene inferred by Bayesian analysis. PLoS One 2011; 6(12): e28161. doi: 10.1371/journal.pone.0028161 PMID: 22174775

63. Printzen C. Uncharted terrain: the phylogeography of arctic and boreal lichens. Plant Ecol Divers 2008; 1: 265–271.