SUPPORTING INFORMATION

Genome-wide assessment of putative endemism and phylogeography of *Cladonia sandstedei* (Ascomycota: Cladoniaceae) in the Caribbean

Joel A. Mercado-Díaz, Felix Grewe, Thorsten Lumbsch

TABLE OF CONTENTS

Appendix S1. Barcoding sequencing.

Exploratory analysis of genetic divergence between Puerto Rican populations of *C. sandstedei* was carried out by generating single-locus data for three samples from Maricao and three samples from Vega Baja. This work entailed obtaining sequences for the Translation Elongation Factor 1-Alfa (EF1) and the RNA polymerase I subunit II (RPB2). Primers and PCR conditions used in this study are described in Table S1. PCR amplification and sequencing followed protocols described in Mercado-Díaz et al., (2020). Reference sequences were downloaded from GenBank or obtained from Rebecca Yahr (Table S2).

Table S1. Primers and PCR conditions used for single-locus sequencing.

Locus	Primer	Primer sequence 5'-3'	PCR protocol	Reference
Translation elongation factor 1-alpha (~ 1,000 bp) Program: EF1TD	EF1-526f EF1-1567R	GTC GTY GTY ATY GGH CAY GT ACH GTR CCR ATA CCA CCR ATC TT	94°C for 4 mins;10 cycles: 94 °C for 30 s, 66 °C for 30 s (decreasing 1 °C per cycle), 72 °C for 90 s; 30 cycles: 94 °C for 30 s, 56 °C for 30 s, 72 °C for 90 s; 72 °C for 7 mins	(Rehner 2001)
RNA polymerase II subunit 2 (RPB2) (~ 800 bp) Program: IGS52_2	RPB2-5f RPB2-7cR	GAY GAY MGW GAT CAY TTY GG CCC ATR GCT TGY TTR CCC AT	94°C for 3 min; 34 cycles: 94°C for 45 s, 50°C for 60 s, 72°C for 90 s; 72°C for 7 min	(Liu, Whelen, and Hall 1999)

Table S2. Samples and GenBank accession numbers used for barcoding sequencing. Asterisks denote pending GenBank accession numbers. Exclamation marks show sequences obtained from R. Yahr that are not available in GenBank.

ID	Species	Area	EF1	RPB2
LK46	*Cladonia confusa*	Brazil		KP941559
Burgaz 96193	*Cladonia rangiformis*	Spain	JN811444	JF288838
DNA15497	*Cladonia sandstedei*	Maricao, PR	*	*
DNA15498	*Cladonia sandstedei*	Maricao, PR	*	*
DNA15499	*Cladonia sandstedei*	Maricao, PR	*	*
DNA15500	*Cladonia sandstedei*	Vega Baja, PR	*	*
DNA15501	*Cladonia sandstedei*	Vega Baja, PR	*	*
DNA15502	*Cladonia sandstedei*	Vega Baja, PR	*	*
RY1004	*Cladonia subtenuis*	Florida, USA	DQ490098	DQ522287

RY1123	*Cladonia subtenuis*	North Carolina, USA	DQ490096	
RY1128	*Cladonia subtenuis*	North Carolina, USA	DQ490101	
RY1129	*Cladonia subtenuis*	North Carolina, USA	DQ490093	
RY1151	*Cladonia subtenuis*	North Carolina, USA	DQ490095	
RY1189	*Cladonia subtenuis*	North Carolina, USA	DQ490105	
RY1190	*Cladonia subtenuis*	North Carolina, USA	DQ490104	DQ522286
RY1208	*Cladonia subtenuis*	Georgia, USA	!	DQ522282
RY1210	*Cladonia subtenuis*	Georgia, USA	!	DQ522283
RY1213	*Cladonia subtenuis*	Georgia, USA	!	DQ522284
RY1215	*Cladonia subtenuis*	Georgia, USA	DQ490102	
RY1216	*Cladonia subtenuis*	Georgia, USA	DQ490100	
RY1224	*Cladonia subtenuis*	Pennsylvania, USA	!	DQ522289
RY909	*Cladonia subtenuis*	Florida, USA	DQ490103	
RY910	*Cladonia subtenuis*	Florida, USA	DQ490091	
RY911	*Cladonia subtenuis*	Florida, USA	DQ490097	
RY913	*Cladonia subtenuis*	Florida, USA	DQ490092	
RY941	*Cladonia subtenuis*	Florida, USA	DQ490094	
RY942	*Cladonia subtenuis*	Florida, USA	!	DQ522285
RY943	*Cladonia subtenuis*	Florida, USA	DQ490099	
RY999	*Cladonia subtenuis*	Florida, USA	!	DQ522288

Appendix S2. *process_radtags* **command-line usage**

1- Single-end sequences were already demultiplexed (ipyrad), thus *process_radtags* was only used for quality control. No barcode file required:

process_radtags -p /home/FM/jmercado/CladRad/C_sandstedei_demultiplexed_files/ -o /home/FM/jmercado/CladRad/stacks/samples_original –inline_null -e apeKI -t 55 -r -c -q

2- Paired-end sequencing quality control and demultiplexing reads (two plates):

process_radtags -P -p /home/FM/jmercado/CladRad/201113_AHLTJKDSXY/Plate1 -o /home/FM/jmercado/CladRad/stacks/Plate1 -b /home/FM/jmercado/CladRad/201113_AHLTJKDSXY/Plate1/GBS-ApeKI-1-96_barcodes2_stacks.txt –inline_null -e apeKI -t 55 -r -c -q

process_radtags -P -p /home/FM/jmercado/CladRad/201113_AHLTJKDSXY/Plate2 -o /home/FM/jmercado/CladRad/stacks/Plate2 -b /home/FM/jmercado/CladRad/201113_AHLTJKDSXY/Plate2/GBS-ApeKI-2-68_barcodes2_stacks.txt –inline_null -e apeKI -t 55 -r -c -q

Appendix S4. RADseq data processing statistics

Raw fungal reads vs. within-sample clusters
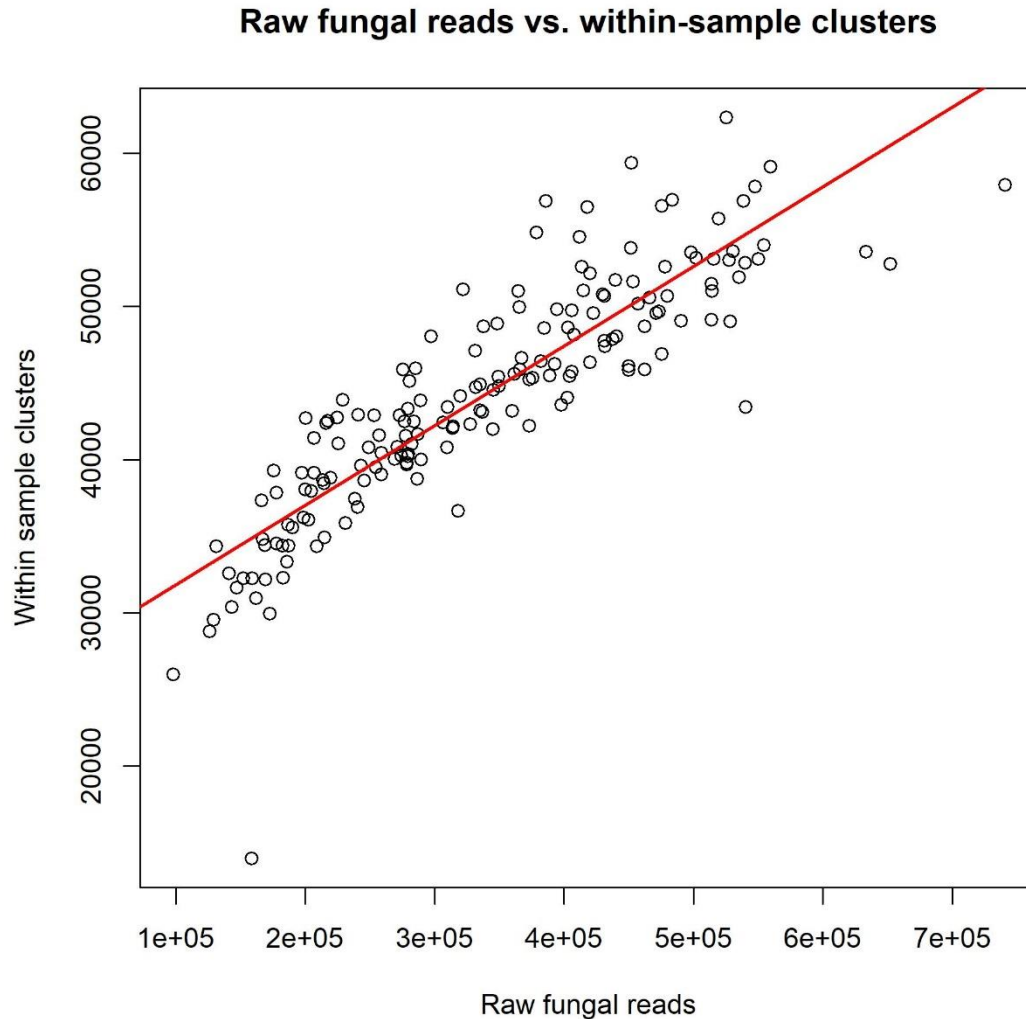
Call:
lm(formula = stats$Within.sample.clusters ~ stats$Reads.mapped.to.reference)
Residuals:
  Min   1Q Median   3Q   Max
-20900 -1916   -296   2056 10228

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.665e+04 8.076e+02   33.01   <2e-16 ***
stats$Reads.mapped.to.reference 5.193e-02 2.239e-03   23.19   <2e-16 ***
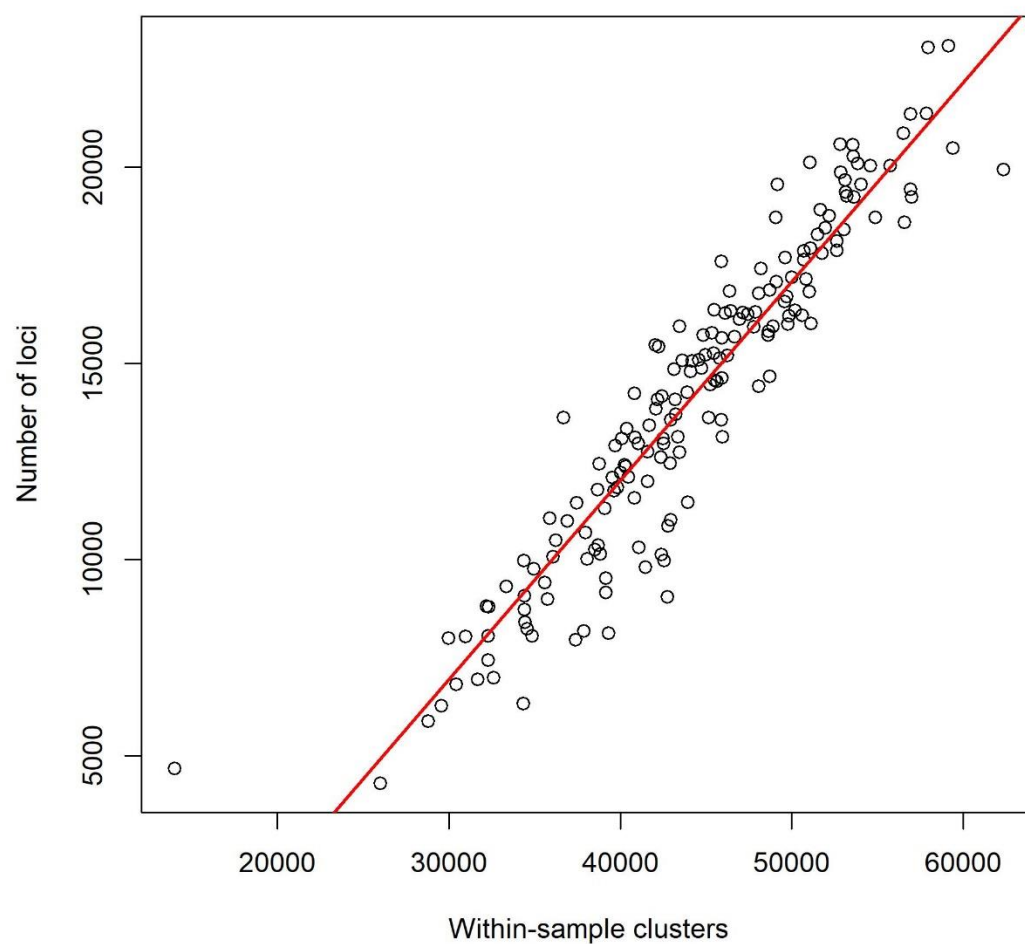---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3715 on 170 degrees of freedom
Multiple R-squared: 0.7599,   Adjusted R-squared: **0.7584**
F-statistic: 537.9 on 1 and 170 DF, p-value: < **2.2e-16**

Within-sample clusters vs. Final number of loci

Call:
lm(formula = stats$Number.of.loci ~ stats$Within.sample.clusters)

Residuals:
    Min   1Q Median   3Q   Max
-4347.2 -774.6   203.5 744.1 5859.6

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -8.276e+03 6.288e+02 -13.16   <2e-16 ***
stats$Within.sample.clusters 5.071e-01 1.403e-02   36.15   <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1387 on 170 degrees of freedom
Multiple R-squared: 0.8849,   Adjusted R-squared: **0.8842**
F-statistic: 1307 on 1 and 170 DF, p-value: < **2.2e-16**

Appendix S5. DAPC results for a-priori and de-novo clustering at K=3 and K=4
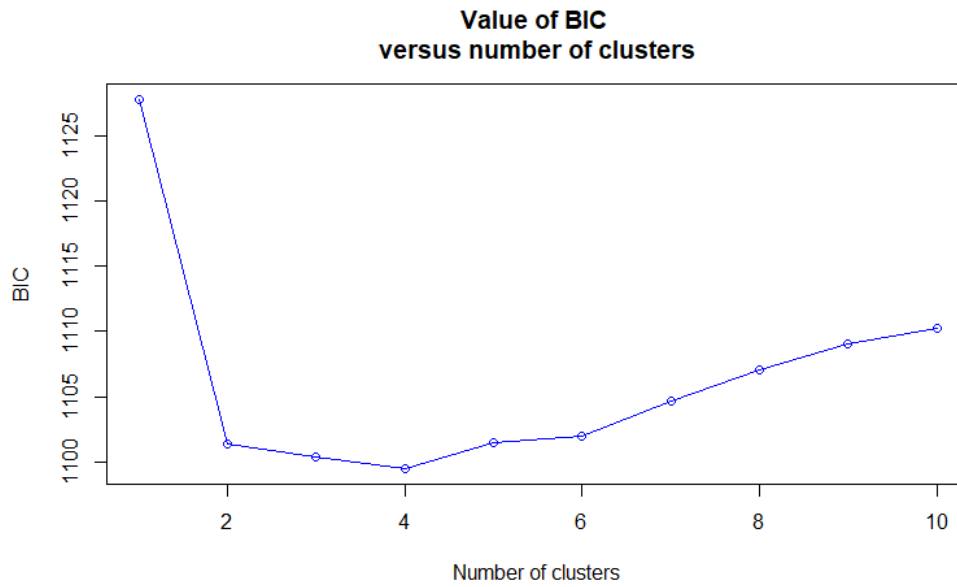

Figure S5.1. Plot for selecting the "best" number of populations (K) based on a Bayesian information criterion. Part of the function *find.clusters* (R package "adegenet").
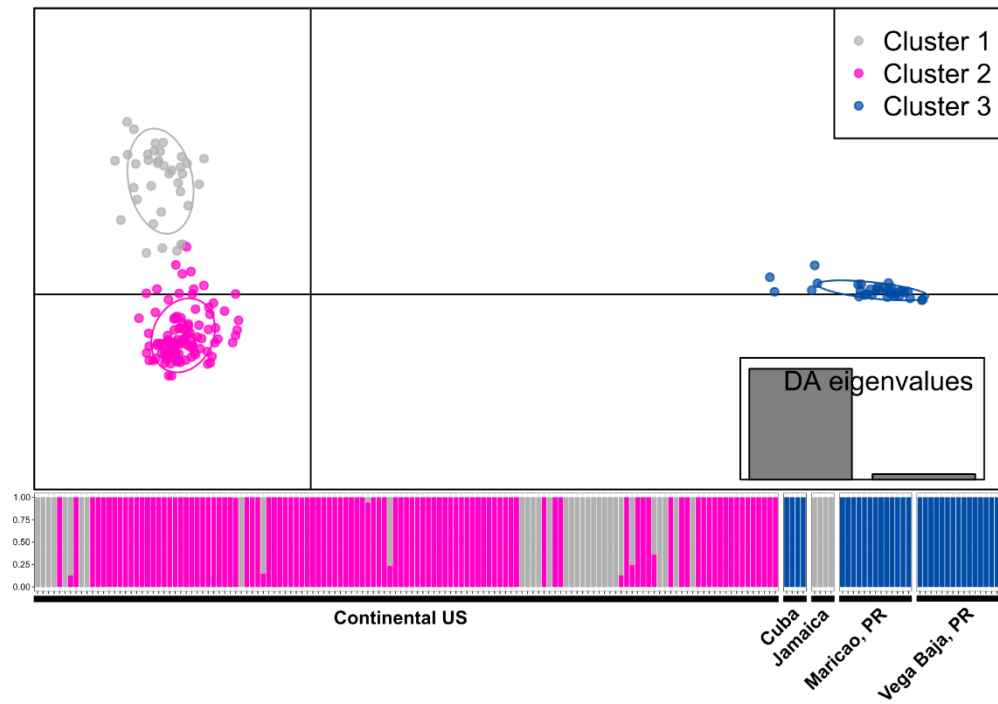

Figure S5.2. Results from de-novo clustering with DAPC at K = 3. Upper part shows scatterplot for discriminant functions whereas the lower part show barplot with assigned membership probabilities. Each dot and bar represent an individual.
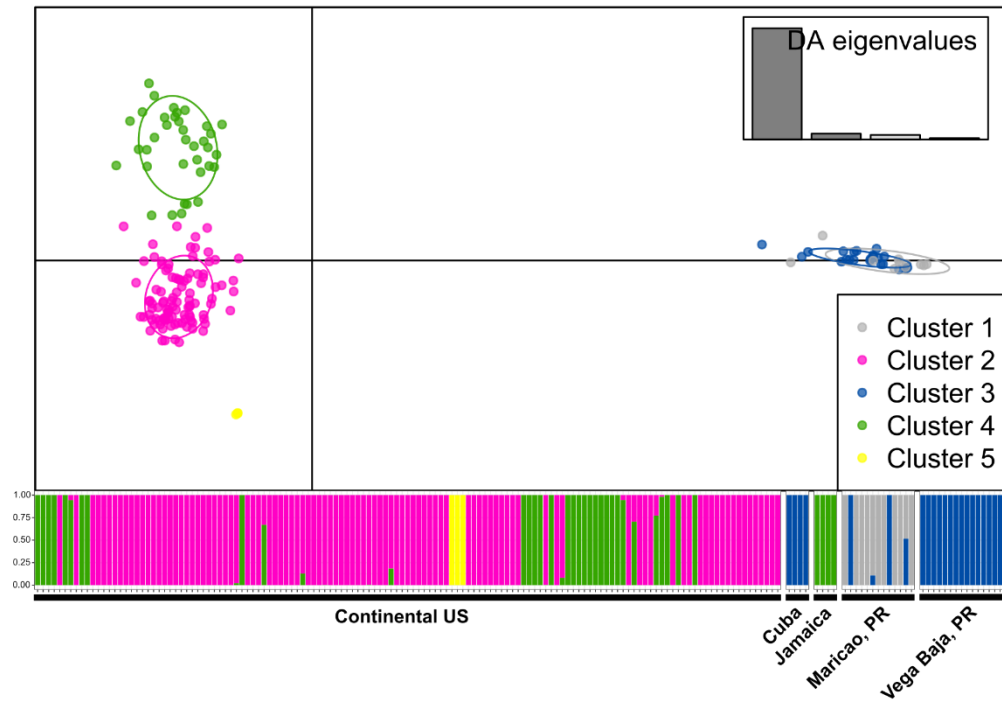
Figure S5.3. Results from de-novo clustering with DAPC at K = 5. Upper part shows scatterplot for discriminant functions whereas the lower part show barplot with assigned membership probabilities. Each dot and bar represent an individual.
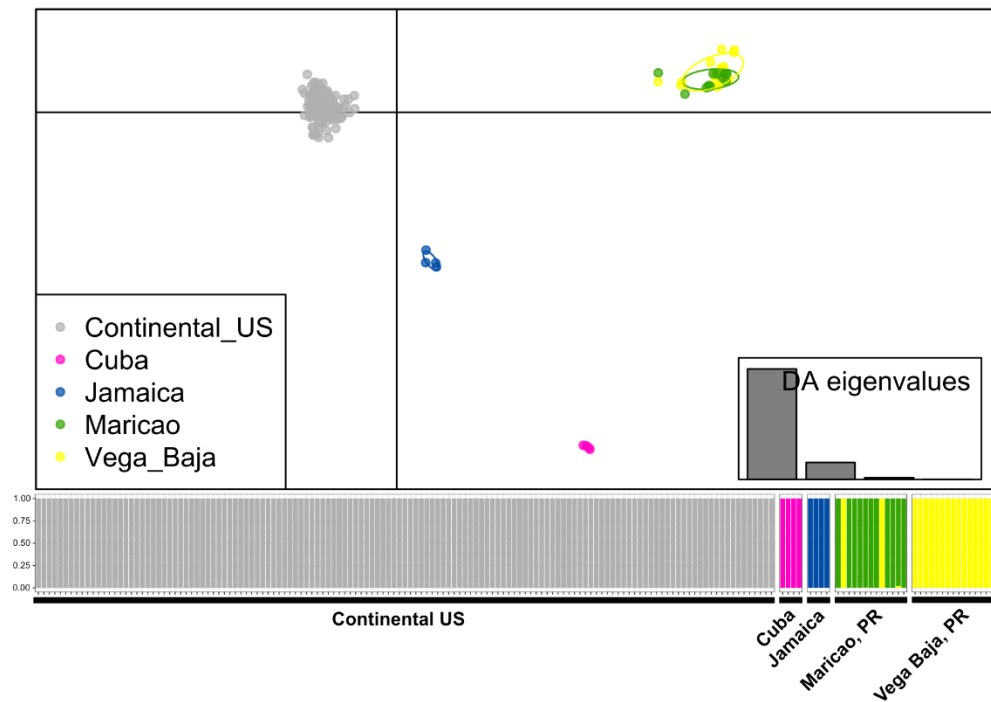

Figure S5.4. Results from a-priori clustering with DAPC. Upper part shows scatterplot for discriminant functions whereas the lower part show barplot with assigned membership probabilities. Each dot and bar represent an individual.

Appendix S6. Genetic dissimilarity vs. geographic distance correlation analysis

       Genetic dissimilarity and geographic distances between continental individuals were computed to assess if isolation by distance could partly explain population partitioning found using de-novo clustering with DAPC. We used the function *diss.dist* from the R package "poppr" (Kamvar, Tabima, and Grïunwald 2014) to calculate pairwise allelic distances between individuals. For geographic distances, sample coordinates were tabulated and the function *distm* (fun = distGeo) from the R package "geosphere" (Hijmans 2019) was used to calculate physical distance between these samples. Matrices were converted to distance objects with the R function *as.dist*. Statistical significance of correlation was assessed using a mantel test (function *mantel* in R package "vegan"(Oksanen et al. 2019)).

       We found weak correlation between genetic dissimilarity and geographic distance between continental individuals (r=0.11). However, the association between these variables was found to be significant (p=0.001) (Fig S6).
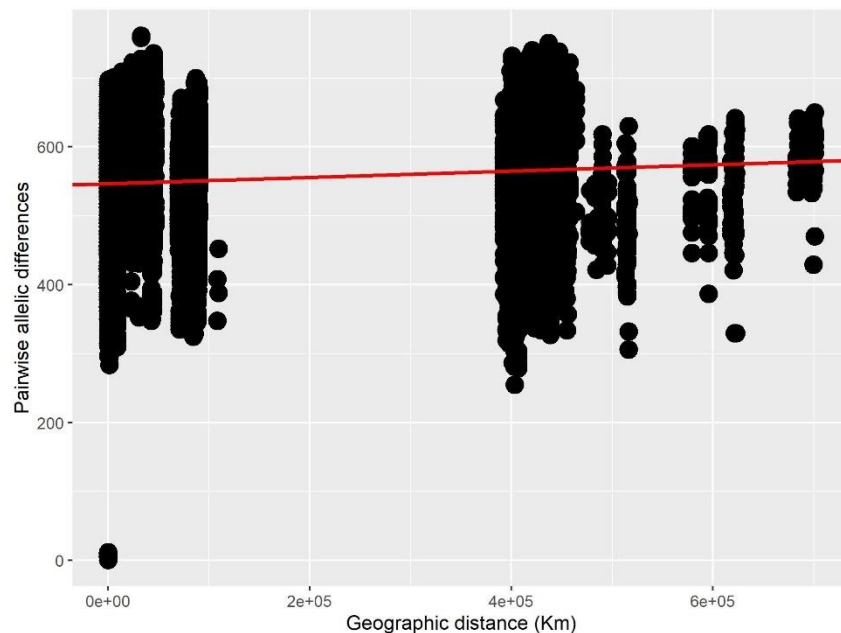

Fig S6. Correlation between geographic distance and genetic dissimilarity (as pairwise allelic differences) between continental individuals. The red line denotes the linear regression function.

References

Hijmans, RJ. 2019. "Geosphere: Spherical Trigonometry. R Package Version 1.5-10." https://cran.r-project.org/package=geosphere.

Kamvar, Zhian N., Javier F. Tabima, and Niklaus J. Grünwald. 2014. "Poppr: An R Package for Genetic Analysis of Populations with Clonal, Partially Clonal, and/or Sexual Reproduction." *PeerJ* 2014 (1): 1–14. https://doi.org/10.7717/peerj.281.

Liu, Yajuan J., Sally Whelen, and Benjamin D. Hall. 1999. "Phylogenetic Relationships among Ascomycetes: Evidence from an RNA Polymerse II Subunit." *Molecular Biology and Evolution* 16 (12): 1799–1808. https://doi.org/10.1093/oxfordjournals.molbev.a026092.

Mercado-Díaz, Joel A., Robert Lücking, Bibiana Moncada, Todd J. Widhelm, and H. Thorsten Lumbsch. 2020. "Elucidating Species Richness in Lichen Fungi: The Genus Sticta (Ascomycota: Peltigeraceae) in Puerto Rico." *Taxon* 69 (5): 851–91. https://doi.org/10.1002/tax.12320.

Oksanen, Author Jari, F. Guillaume Blanchet, Micheal Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, R. B. Minchin, et al. 2019. "Vegan: Community Ecology Package." https://doi.org/ISBN 0-387-95457-0.

Rehner, Stephen A. 2001. "Primers for Elongation Factor 1-Alpha (EF1-Alpha)." http://www.aftol.org/pdfs/EF1primer.pdf.