

Supplemental Information for:

**Evolutionary rate and genetic load in an emblematic
Mediterranean tree following an ancient and prolonged
population collapse**

Juan P. Jaramillo-Correa¹, Francesca Bagnoli², Delphine Grivet³, Bruno Fady⁴, Filippos A. Aravanopoulos⁵, Giovanni G. Vendramin², Santiago C. González-Martínez^{6,*}

¹Department of Evolutionary Ecology, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico

²Institute of Biosciences and Bioresources, Division of Florence, National Research Council, Sesto Fiorentino (FI), Italy

³Department of Forest Ecology and Genetics, Forest Research Centre, INIA-CIFOR, Madrid, Spain

⁴INRAE, Unité de Recherche Écologie des Forêts Méditerranéennes (URFM), Avignon, France

⁵Laboratory of Forest Genetics and Tree Breeding, Department of Forestry and Natural Environment, Aristotle University of Thessaloniki, Thessaloniki, Greece

⁶INRAE, Université de Bordeaux, UMR BIOGECO, Cestas, France

*Corresponding author: santiago.gonzalez-martinez@inrae.fr

TABLE S1 Geographic coordinates, average sample size (N), mean number of alleles (A), mean effective number of alleles (A_E), number of private alleles (A_{PT}), and mean observed (H_O) and expected heterozygosities (H_E) estimated with 13 nuclear microsatellites (nuSSRs) for 33 stone pine (*P. pinea*) populations distributed across its current natural range. See Figure 1 for population location.

Ref	Name	Country	Latitude °N	Longitude °E	N	A	A_E	A_{PT}	H_O	H_E
1	Quinta do Lago	Portugal	37.20	-8.1	21.23	1.46	1.26	1	0.069	0.150
2	Garrovillas	Spain	39.41	-6.35	23.85	1.46	1.09	1	0.036	0.064
3	Bogarra	Spain	38.32	-2.12	23.92	1.31	1.17	0	0.046	0.095
4	Palafrugell	Spain	41.57	-3.6	23.15	1.46	1.25	0	0.109	0.154
5	Doñana	Spain	36.55	-6.25	23.08	1.46	1.21	0	0.087	0.119
6	Biar	Spain	38.38	0.45	22.23	1.54	1.29	0	0.052	0.167
7	San Martín de Valdeiglesias	Spain	40.20	-4.25	28.85	1.62	1.41	1	0.176	0.221
8	Tarazona	Spain	39.10	-1.50	23.31	1.69	1.36	2	0.162	0.203
9	Cap Spartel	Morocco	35.47	-5.55	23.00	1.46	1.22	0	0.119	0.130
10	Villeneuve/Herault	France	43.37	3.24	23.39	1.54	1.22	1	0.090	0.137
11	Saintes Maries	France	43.27	4.26	20.39	1.39	1.30	0	0.092	0.164
12	Saint Aygulf	France	43.26	6.41	17.39	1.39	1.27	0	0.113	0.155
13	Cap Bon	Tunisia	36.57	9.59	23.15	1.62	1.25	1	0.175	0.159
14	Rimel	Tunisia	37.14	9.56	21.92	1.54	1.31	0	0.141	0.188

15	Migliarino	Italy	47.47	10.17	22.85	1.46	1.21	0	0.110	0.129
16	Feniglia	Italy	42.25	11.17	21.85	1.39	1.27	0	0.144	0.155
17	Sardegna/Pula	Italy	39.01	9.00	15.00	1.39	1.19	0	0.102	0.113
18	Bkassine	Lebanon	33.57	35.60	21.77	1.39	1.22	0	0.115	0.125
19	Beit Mounzer	Lebanon	34.29	35.91	23.54	1.39	1.22	0	0.074	0.126
20	Kornael	Lebanon	33.84	35.77	22.00	1.23	1.18	0	0.087	0.099
21	Qsaibeh	Lebanon	33.88	35.68	23.77	1.46	1.19	1	0.106	0.116
22	Masade/Golden Heights	Israel	35.44	33.13	23.77	1.31	1.22	0	0.124	0.126
23	Quirit/Anavin	Israel	35.07	31.48	21.54	1.31	1.18	0	0.100	0.104
24	Lemessos	Cyprus	34.48	32.56	18.85	1.39	1.21	0	0.020	0.126
25	Pafos	Cyprus	34.48	32.24	17.08	1.31	1.10	0	0.042	0.070
26	Chalkidiki Agios	Greece	40.14	23.34	23.62	1.08	1.02	0	0.014	0.013
27	Parthenonas	Greece	40.26	23.27	15.00	1.00	1.00	0	0.043	0.000
28	Kaiafas	Greece	37.31	21.36	14.54	1.15	1.09	0	0.026	0.058
29	Strofilia	Greece	38.49	23.24	16.54	1.15	1.11	0	0.052	0.062
30	Skopelos	Greece	39.7	23.43	22.46	1.23	1.08	0	0.027	0.051
31	Skiathos	Greece	39.11	23.28	23.15	1.23	1.03	0	0.025	0.026
32	Bursa Yalova	Turkey	40.32	29.22	23.46	1.39	1.06	1	0.014	0.045
33	Yatağan Katrancı	Turkey	37.22	28.03	23.62	1.23	1.06	1	0.040	0.036
	<i>Mean</i>				<i>20.96</i>	<i>1.38</i>	<i>1.19</i>		<i>0.083</i>	<i>0.112</i>

TABLE S2 Mean DNA sequence diversity and divergence from loblolly pine (*P. taeda*), and estimates of adaptive evolution for the two genetic clusters detected in stone pine (*P. pinea*). See Figures 2 and S2 for cluster assignment and geographic location.

Parameter*	Western genetic cluster	Eastern genetic cluster
θ_w (<i>SD</i>)	0.00016 (0.00055)	0.00015 (0.00066)
π (<i>SD</i>)	0.00022 (0.00127)	0.00018 (0.00131)
π_a (<i>SD</i>)	0.00019 (0.00126)	0.00017 (0.00128)
π_s (<i>SD</i>)	0.00043 (0.00355)	0.00042 (0.00357)
π_a/π_s (<i>SD</i>)	0.4488 (0.5241)	0.4492 (0.5244)
K_a (<i>SD</i>)**	0.0099 (0.0114)	0.0099 (0.0115)
K_s (<i>SD</i>)**	0.0381 (0.0289)	0.0381 (0.0288)
K_a/K_s (<i>SD</i>)**	0.4597 (1.0668)	0.4598 (1.0670)
α (95% <i>CI</i>)	0.171 (-0.025 – 0.321)	0.170 (-0.022 – 0.314)
ω_a (95% <i>CI</i>)	0.061 (-0.002 – 0.121)	0.059 (-0.003 – 0.142)

* Abbreviations: *SD*, standard deviation; 95%*CI*, 95% confidence intervals; θ_w , average Watterson's nucleotide diversity per site; π , average Tajima's nucleotide diversity per site; π_a , average π for non-synonymous sites; π_s , average π for synonymous sites; π_a/π_s , mean per gene ratio of non-synonymous to synonymous nucleotide diversity; K_a , average number of non-synonymous substitutions per non-synonymous site; K_s , average number of synonymous substitutions per synonymous site; K_a/K_s , mean per gene ratio of non-synonymous to synonymous substitutions; α , proportion of (non-shared) non-synonymous substitutions fixed by adaptive evolution; and ω_a , relative rate of (non-shared) adaptive substitutions.

** Estimates determined with Jukes-Cantor correction.

TABLE S3a Description of gene-loci with more exclusive replacement substitutions in *P. pinea* than the average gene (>5). Presence of potentially deleterious SNPs according to Provean (Ref. 2) is highlighted in first column (Effect). Annotation is based on Blast similarity searches on pine data. 'Candidate gene' denotes if loci have been reported as candidate genes for adaptation in other pines.

Effect	Name	Annotation	No. fixed mutations	Coding	Non synonymous	Non coding	Candidate gene	Expression in <i>P. pinea</i>)	Refs
	0_4042_01	GST; glutathione S-transferase	5	5	5	0	<i>P. pinaster</i>		6
	0_4229_01	protein JINGUBANG-like	5	3	1	2			
	0_4285_01	amino acid permease 3-like	9	2	0	7			
	0_4756_01	F-box family protein	5	5	1	0			
	0_4787_01	Histone H2A	7	1	0	6			
	0_7183_01	Phospholipase A2-like	6	2	1	4	<i>P. halepensis</i>		1
	0_8463_01	Putative KH domain-containing protein	5	2	0	3			
	0_9383_01	Ubiquitin carboxyl-terminal hydrolase-like protein	6	6	4	0			
	0_10162_01	Carbonic anhydrase	12	4	0	8			
	0_11649_01	TUBB; tubulin beta	9	5	0	4	<i>P. pinaster</i>	Upexpressed (stem)	7, 8
	0_12077_01	Putative squamosa promoter-binding-like protein 14	6	5	1	1			
	0_12535_01	hydrolase, alpha/beta fold family protein	5	0	0	5	<i>P. halepensis</i>		1
	0_13484_01	Putative LOB domain-containing protein	6	5	3	0			
	0_13929_02	Basic pentacysteine 5	5	1	1	4	<i>P. lambertiana</i>	Upexpressed (stem and needle)	5, 9
	0_14221_01	Serine-tRNA ligase	7	2	0	5			
	0_15741_01	Respiratory burst oxidase homologue D	7	0	0	7	<i>P. cembra</i>		8
	0_16860_01	AP3D; AP-3 complex subunit delta	7	0	0	6			
	0_17127_01	putative U1 small nuclear ribonucleoprotein 70 kDa	6	2	0	4			
	0_17758_01	GST; glutathione S-transferase	7	5	4	2			
	0_17758_01	GST; glutathione S-transferase	9	6	4	2			

0_17790_01	Beta-glucuronidase	6	2	2	4			
0_17971_01	putative signal recognition particle receptor	7	1	0	6			
2_1014_01	Heat shock factor protein HSF24	6	4	4	0	<i>P. mugo and P. pinaster</i>	Upexpressed (stem and roots)	7, 8, 9
2_2936_01	NA	9	1	1	8			
2_2960_02	HAT3 transcription factor	6	5	4	1	<i>P. mugo</i>		8
2_3867_02	profilin-like protein	6	3	0	3			
2_3919_01	Cation/H(+) antiporter	14	na	na	na	<i>P. pinaster</i>		7
2_3941_01	B3 domain-containing protein	7	2	0	5	<i>P. halepensis</i>		4
2_5668_01	alpha-L-fucosidase 3	6	1	0	5			
2_7725_01	Beta-galactosidase 8	7	2	2	4	<i>P. mugo</i>		8
CL516Contig1_07	ATPase/hydrogen-translocating pyrophosphatase (AVP1)	5	5	1	0	<i>P. mugo and P. cembra</i>		8
CL708Contig1_02	dxr; 1-deoxy-D-xylulose-5-phosphate reductoisomerase	5	2	0	3			
CL1004Contig1_08	T-complex protein 1 subunit epsilon-like	9	0	0	9			
CL1029Contig1_01	Imbibition protein	5	2	0	3	<i>P. mugo</i>		8
CL1430Contig1_06	Pyrophosphate—fructose-6-phosphate 1-phosphotransferase	5	0	0	5	<i>P. lambertiana</i>		5
CL3037Contig1_06	60S ribosomal protein L13a	7	1	0	6			
CL3539Contig1_01	VHS domain-containing protein	5	1	0	4	<i>P. taeda</i>		3
CL4147Contig1_01	Hexokinase (HXK1)	5	1	0	4	<i>P. mugo</i>		8
UMN_1911_01	NA	13	na	na	13			
UMN_6426_02	Cellulose synthase/transferase	6	3	1	3	<i>P. mugo and P. cembra</i>		8

Effect:



Only synonymous changes

All non-synonymous changes marked as neutral by Provean

At least one non-synonymous change marked as deleterious by Provean

No data available (i.e. unannotated loci, no coding region in loci, or protein not found in Provean database)

References:

1. Budde, K.B., et al. (2017) *Annals of Botany*, 119, 1061–1072.
2. Choi, Y., Chan, A.P. (2015) *Bioinformatics*, btv195.
3. Chhatre, V.E. et al. (2013) *Tree Genetics and Genomes*, 9, 1161–1178.
4. Daniels, R.R. et al. (2018) *Molecular Ecology*, 27, 3331-3345.
5. Eckert, A.J. et al. (2015) *Tree Genetics and Genomes*, 11, 42.
6. Grivet, D., et al. (2017) *Molecular Ecology*, 26, 6857–6870.
7. Jaramillo-Correa, J.P., et al. (2015) *Genetics*, 199, 793-807.
8. Mosca, E. et al. (2016) *Tree Genetics and Genomes*, 12, 9.
9. Perdiguero, P., et al. (2013) *Plant Physiology and Biochemistry*, 67, 199-208.

TABLE S3b Description of gene-loci with exclusive polymorphisms in *P. pinea*. Presence of potentially deleterious SNPs according to Provean (Ref. 2) is highlighted in first column (Effect). Annotation is based on Blast similarity searches on pine data. ‘Candidate gene’ denotes if loci have been reported as candidate genes for adaptation in other pines. See Table S3a for references (Refs).

Effect	Name	Annotation	No. new mutations	Fixed	Variable	Coding	Non synonymous	Non coding	Candidate gene	Refs
	0_9448_01	putative xylanase Xyn2	4	2	2	4	1	0		
	0_10384_02	NA	4	3	1	3	2	1		
	0_11351_01	putative pentatricopeptide repeat-containing protein	5	3	2	5	5	0		
	0_12156_02	ATP binding protein	2	1	1	2	2	0	<i>P. sylvestris</i>	6
	0_12190_01	NA	1	0	1	1	1	0		
	0_13484_01	Putative LOB domain-containing protein	6	5	1	6	4	0		
	0_16009_02	ELF2C; eukaryotic translation initiation factor 2C	4	3	1	4	1	0		
	0_16068_01	Sterol glucosyltransferase-like protein	1	0	1	1	1	0	<i>P. lambertiana</i>	5
	0_16860_01	AP3D; AP-3 complex subunit delta	7	6	1	0	0	7		
	0_17758_01	GST; glutathione S-transferase	9	8	1	7	5	2		
	2_5996_01	aspartyl protease-like protein	2	1	1	2	1	0		
	2_7725_01	Beta-galactosidase 8	7	6	1	2	2	5	<i>P. mugo</i>	8
	2_8011_02	Putative LRR receptor-like serine/threonine protein kinase MRH1	5	3	2	5	1	0		
	CL1489Contig1_01	ruBisCO large subunit-binding protein subunit alpha-like	4	1	3	4	0	0		
	CL1634CONTIG1_01	Putative transmembrane 9 superfamily member	3	2	1	1	1	2		

Effect:

- Only synonymous changes
- All non-synonymous changes marked as neutral by Provean
- At least one non-synonymous change marked as deleterous by Provean
- No data available (i.e. unannotated loci, no coding region in loci, or protein not found in Provean database)

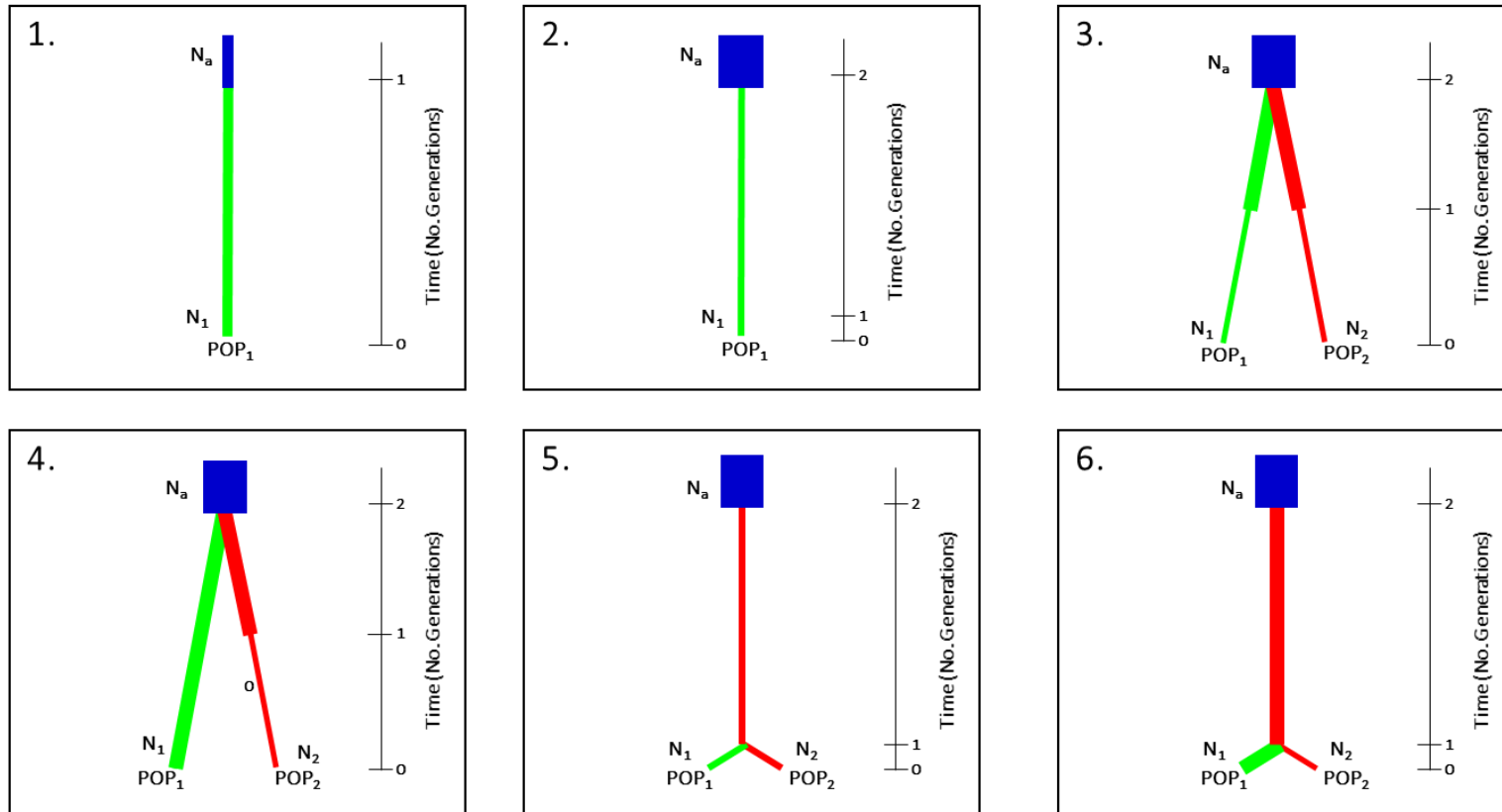


FIGURE S1 Schematic representation of the demographic scenarios tested for stone pine using DiyABC (based on the variation of 13 nuSSRs), and population parameters estimated for each scenario.

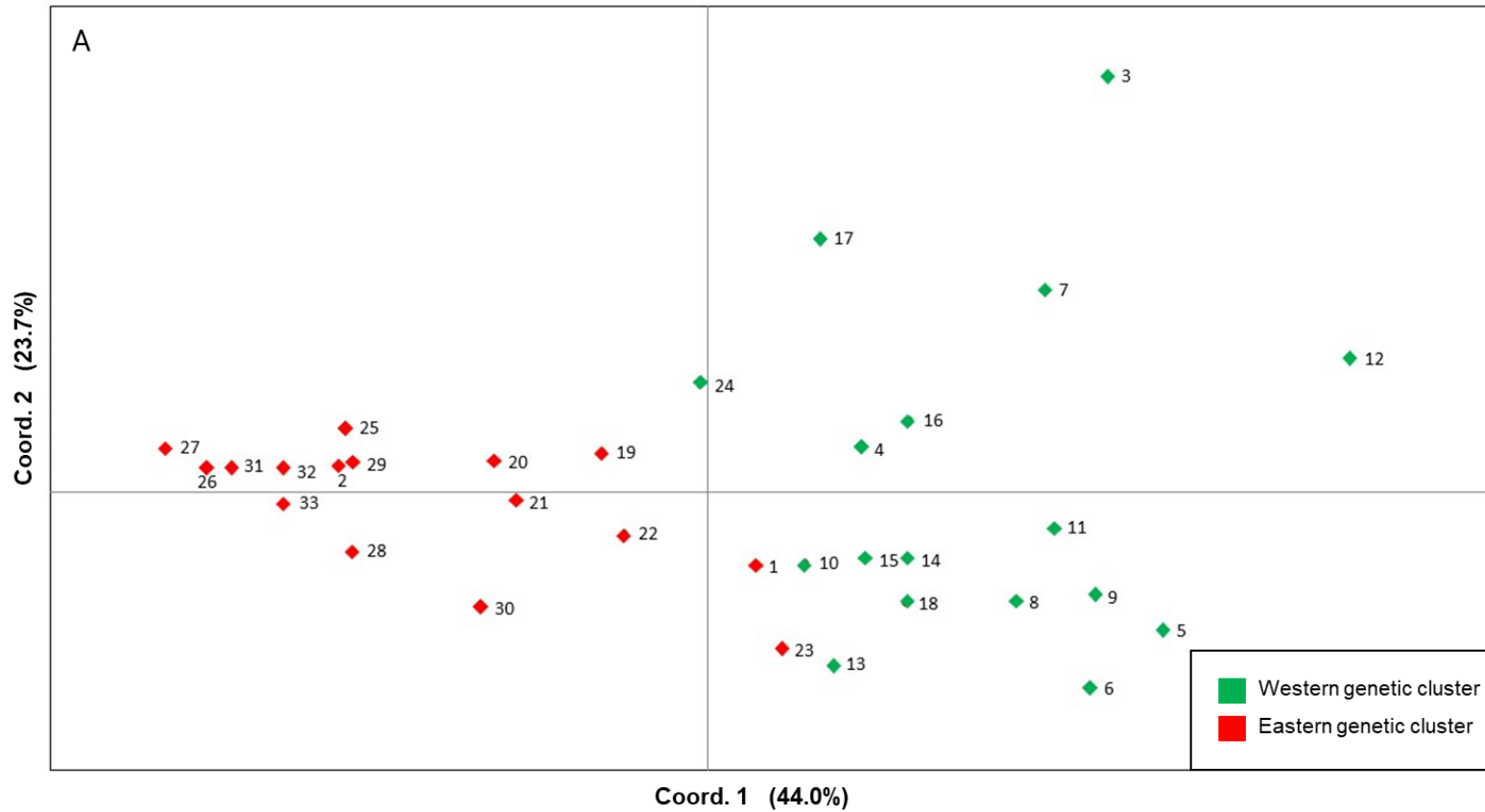


FIGURE S2a Scatterplot for the first two coordinates of a Principal Coordinates Analysis performed on stone pine nuSSR variation. Populations are color-coded according to assignment based on STRUCTURE results for $K=2$ (Q -value threshold of 50%; see Fig. S2b). See Table S1 for population codes and Fig. 1 for map location.

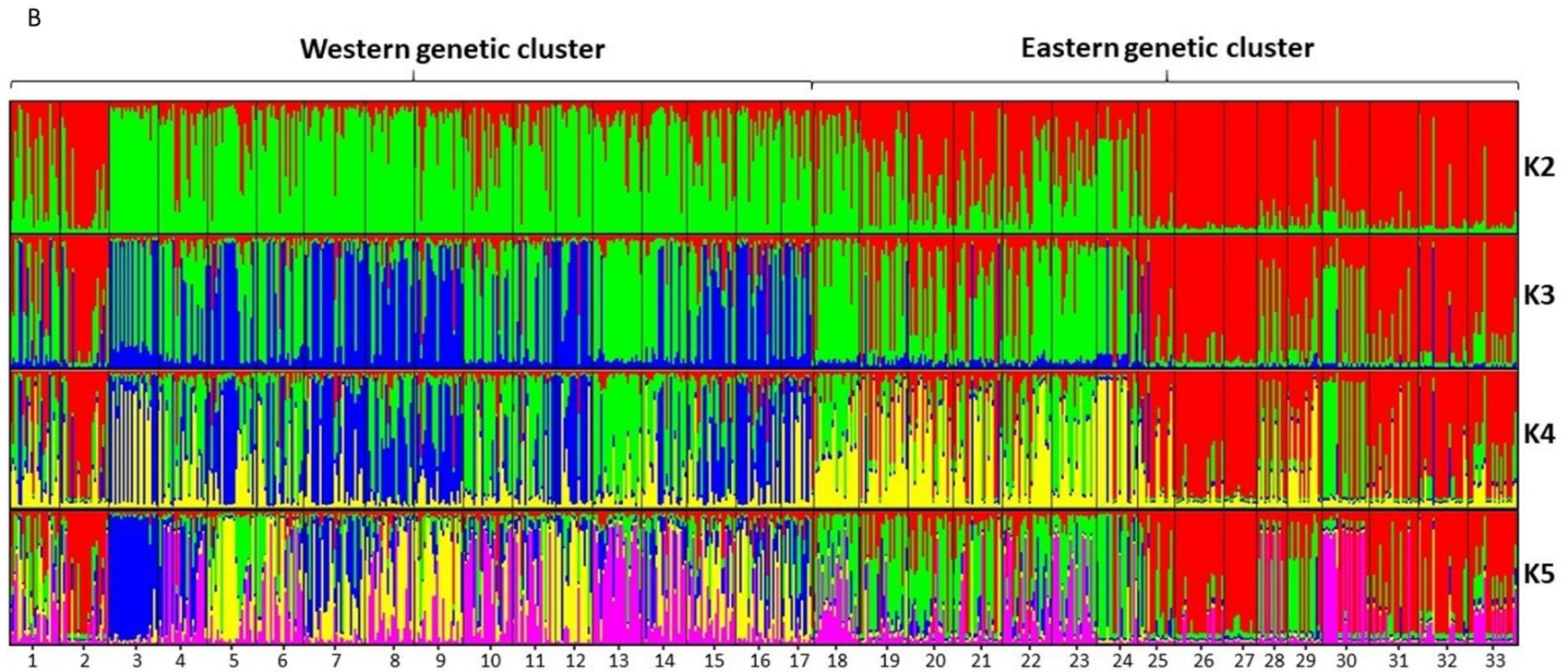


FIGURE S2b Bar-plot of genetic-cluster membership coefficients obtained with STRUCTURE from stone pine nuSSR variation; two to five genetic clusters ($K = 2 - 5$) were assumed. See Table S1 for population codes and Figs. 1-2 for map location.

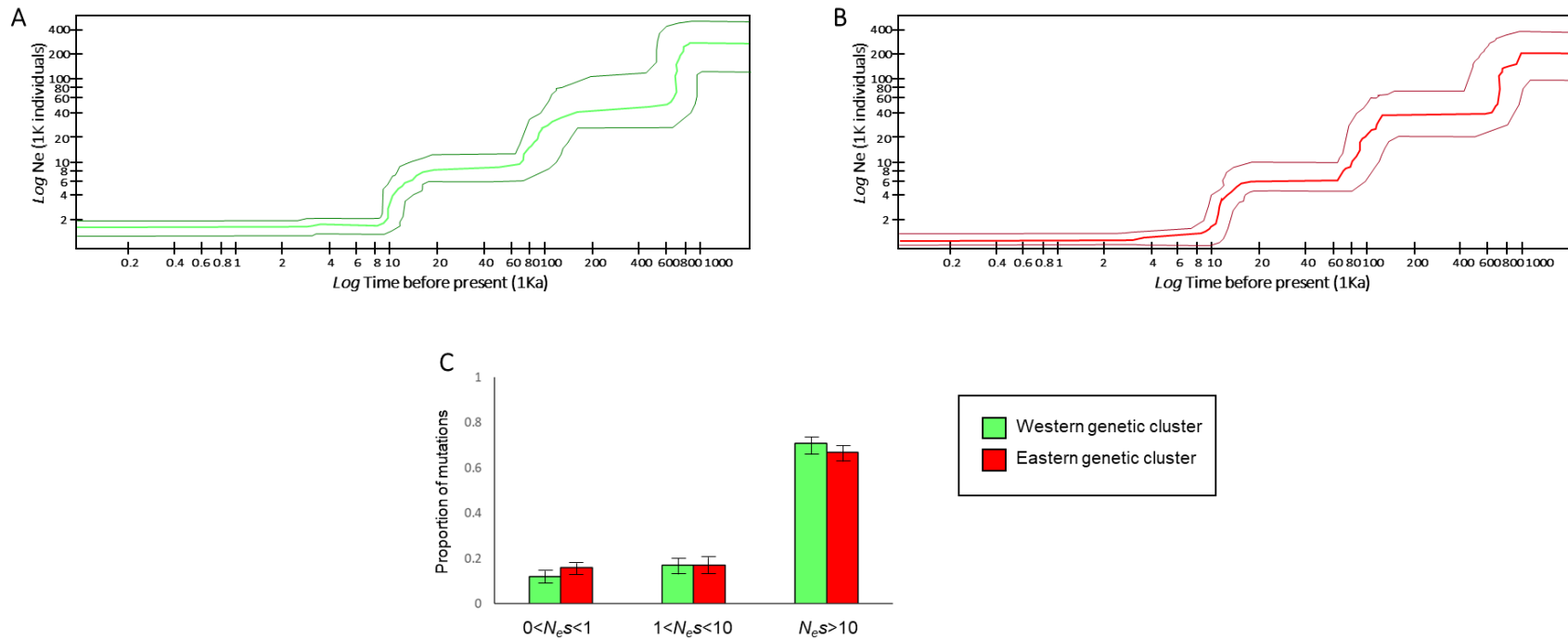


FIGURE S3 Changes of population size over time for the western (A) and eastern (B) genetic clusters of stone pine inferred with a stairway-plot derived from the site frequency spectrum (SFS) of 55,833 DNA-sequence sites. Dark and light lines represent the median and 95% CI, respectively. Estimates based on the assumption of a mutation rate of 1.64×10^{-9} per site per year, and a generation time of 25 years. (C) Distribution of fitness effects (DFE) of exclusive mutations for the same genetic clusters.

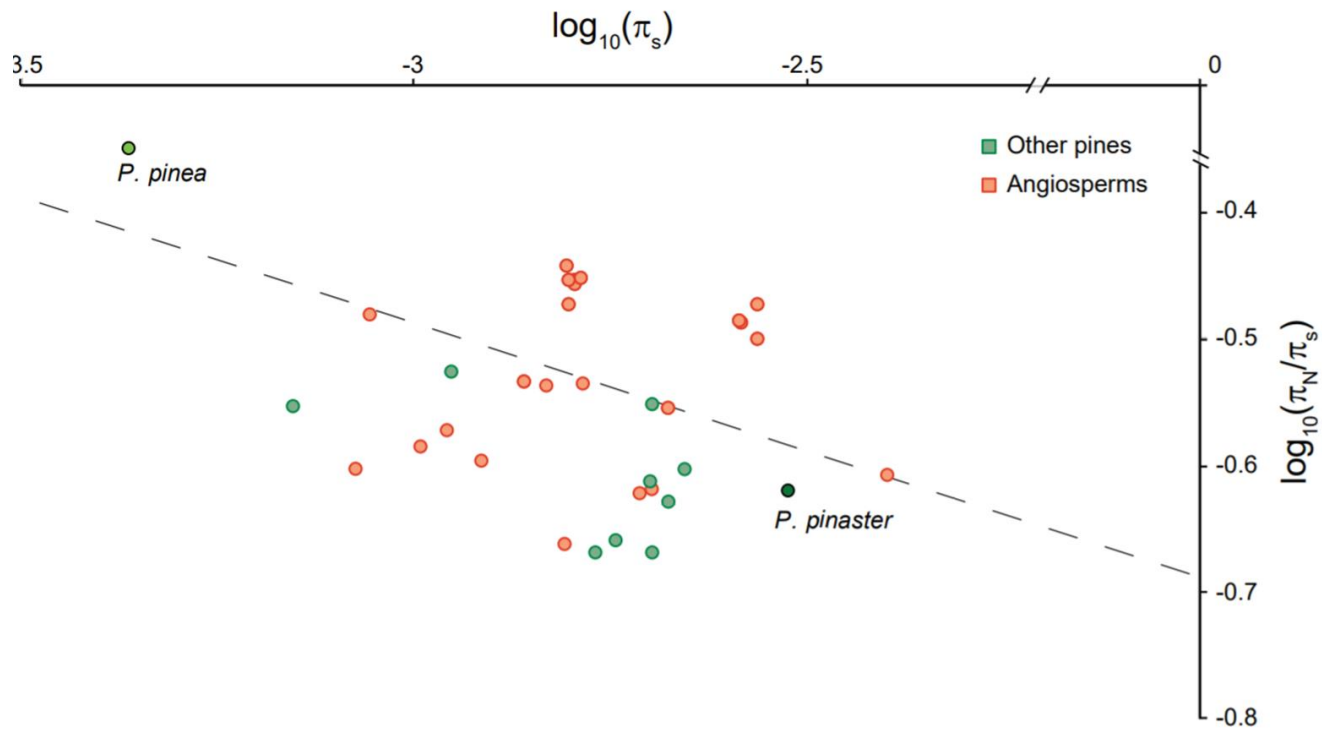


FIGURE S4 Correlation between π_s and π_N/π_s for woody/shrubby plant species. Values for stone and maritime pine are highlighted. Data for other species and regression line taken from Chen et al. (2017).