Genetic Diversity and Demographic History of Wild and Cultivated/Naturalised Plant Populations: Evidence from Dalmatian Sage (*Salvia officinalis* L., Lamiaceae)

Ivana Rešetnik[1], Dea Baričevič[2], Diana Batîr Rusu[3], Klaudija Carović-Stanko[4], Paschalina Chatzopoulou[5], Zora Dajić-Stevanović[6], Maria Gonceariuc[7], Martina Grdiša[4], Danijela Greguraš[1], Alban Ibraliu[8], Marija Jug-Dujaković[9], Elez Krasniqi[10], Zlatko Liber[1], Senad Murtić[11], Dragana Pećanac[12], Ivan Radosavljević[1], Gjoshe Stefkov[13], Danijela Stešević[14], Ivan Šoštarić[6], Zlatko Šatović[4]*

1 Faculty of Science, University of Zagreb, Zagreb, Croatia, 2 Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia, 3 Suceava Genebank, Suceava, Romania, 4 Faculty of Agriculture, University of Zagreb, Zagreb, Croatia, 5 Department of Aromatic and Medicinal Plants, Hellenic Agricultural Organisation, Thessaloniki, Greece, 6 Faculty of Agriculture, University of Belgrade, Belgrade, Serbia, 7 Institute of Genetics and Plant Physiology, Academy of Sciences, Chişinău, Moldova, 8 Faculty of Agriculture and Environment, Agricultural University of Tirana, Tirana, Albania, 9 Institute for Adriatic Crops and Karst Reclamation, Split, Croatia, 10 Faculty of Natural Sciences and Mathematics, University of Pristina, Prishtinë, Kosovo, 11 Faculty of Agriculture and Food Science, University of Sarajevo, Sarajevo, Bosnia and Herzegovina, 12 Faculty of Agriculture, University of Banja Luka, Banja Luka, Bosnia and Herzegovina, 13 Faculty of Pharmacy, University of Ss. Cyril and Methodius, Skopje, Macedonia, 14 Faculty of Natural Sciences and Mathematics, University of Montenegro, Podgorica, Montenegro

* zsatovic@agr.hr

Abstract

Dalmatian sage (*Salvia officinalis* L., Lamiaceae) is a well-known aromatic and medicinal Mediterranean plant that is native in coastal regions of the western Balkan and southern Apennine Peninsulas and is commonly cultivated worldwide. It is widely used in the food, pharmaceutical and cosmetic industries. Knowledge of its genetic diversity and spatiotemporal patterns is important for plant breeding programmes and conservation. We used eight microsatellite markers to investigate evolutionary history of indigenous populations as well as genetic diversity and structure within and among indigenous and cultivated/naturalised populations distributed across the Balkan Peninsula. The results showed a clear separation between the indigenous and cultivated/naturalised groups, with the cultivated material originating from one restricted geographical area. Most of the genetic diversity in both groups was attributable to differences among individuals within populations, although spatial genetic analysis of indigenous populations indicated the existence of isolation by distance. Geographical structuring of indigenous populations was found using clustering analysis, with three sub-clusters of indigenous populations. The highest level of gene diversity and the greatest number of private alleles were found in the central part of the eastern Adriatic coast, while decreases in gene diversity and number of private alleles were evident towards the northwestern Adriatic coast and southern and eastern regions of the Balkan Peninsula.

The results of Ecological Niche Modelling during Last Glacial Maximum and Approximate Bayesian Computation suggested two plausible evolutionary trajectories: 1) the species survived in the glacial refugium in southern Adriatic coastal region with subsequent colonization events towards northern, eastern and southern Balkan Peninsula; 2) species survived in several refugia exhibiting concurrent divergence into three genetic groups. The insight into genetic diversity and structure also provide the baseline data for conservation of *S. officinalis* genetic resources valuable for future breeding programmes.

Introduction

For thousands of years, people have gathered plant and animal resources for their needs, resulting in changes to genetic structure of populations over the course of cultivation and domestication. This process is particularly manifested in crop species used for food [1], but is less evident in medicinal and aromatic plants (MAP), which are still harvested primarily from wild populations [2, 3]. Nevertheless, impacts on MAP intra-specific genetic diversity can occur through overharvesting in natural environments [4, 5] or through population genetic bottlenecks caused by collection of seeds from a limited number of wild plants that are subsequently used to found cultivated populations [1, 6]. In either case, the need for comprehensive surveys of genetic diversity in natural and cultivated MAP populations is an imperative for efficient conservation efforts, breeding programmes and agricultural production.

The reductions of gene diversity in domesticated plants vary across species and have usually been examined in crop plants such as soybean [7], maize [8] and wheat [9]. Domestication bottleneck processes reduce neutral genetic diversity across the entire genome [7, 9]; the strength of such a bottleneck is determined by duration and effective population size [10]. One of the key questions relating to the evolutionary processes underlying domestication and cultivation of plant species concerns the identity and geographic origin of populations [11] as well as the tempo and mode of domestication (e.g., single or multiple origin events) [10]. Furthermore, the cultivation of plants in proximity to their natural environment can induce introgressive hybridization between domesticated forms and their wild relatives, thereby impacting the initial loss of genetic diversity [1, 12]. Additionally, similarities in habitat and climate conditions can foster the naturalization of cultivated plants, thus expanding their influence on natural populations and surrounding biodiversity [13–16].

Dalmatian sage (*Salvia officinalis* L.) is an outcrossing, insect-pollinated, perennial sub-shrubby plant of the family Lamiaceae. The genus *Salvia* is one of the largest genera in the family, with nearly 1,000 species distributed worldwide [17, 18]. Recent molecular phylogenetic studies revealed the non-monophyly of the genus [19–21] and the inclusion of the type species *S. officinalis* within the monophyletic clade I (*Salvia* sensu stricto; [20]). *Salvia officinalis* is naturally distributed throughout the coastal region of the western Balkan and central and southern Apennine Peninsulas, where it grows abundantly on dry calcareous rocky soil [22–24]. The species is a well-known aromatic Mediterranean plant and has been widely cultivated since ancient times for medicinal, culinary and ornamental purposes. Extracts of *S. officinalis* have been shown to exhibit antioxidant [25, 26], anti-inflammatory [27, 28], fungicidal and bactericidal [29–31], virucidal [32], antispasmatic [33], antidiabetic [34], gastroprotective [35] and anti-obesity [36] activity. The leaves are broadly used for aromatization in the food industry, and the plant has recently become popular as an ornamental garden plant [37], with several cultivars developed for this purpose.

Despite the medicinal, historical and cultural importance of Dalmatian sage, molecular data describing the population genetics and phylogeography of the species are scarce. The majority of previous studies focused on discovery and characterisation of bioactive compounds [38–41] and assessment of essential oil content and composition in relation to collecting site [42–46], environmental conditions [42, 47, 48], season [42], physiological stage (i.e., time of harvest; [49]), plant parts used for the extraction of essential oil [42, 50], soil mineral fertilization [51], drying procedure [52], and extraction [53] and distillation methods [54].

Random Amplified Polymorphic DNA (RAPD) [55, 56] and Amplified Fragment Length Polymorphism (AFLP) [57] fingerprinting were used to analyse the genetic diversity and structure of natural populations distributed in Croatia and Bosnia and Herzegovina. Both marker types revealed high variability within the populations, while genetic differentiation among populations showed a pattern of isolation by distance. The highest genetic diversity was found in populations from central part of eastern coast of the Adriatic Sea, while the highest frequency down-weighted marker values were found in the northernmost populations and the southernmost inland population. Recently, a plastid DNA phylogeographic study based on eight Balkan populations confirmed the natural origin of four disjunct inland populations and revealed the presence of inland and southern coastal lineages [58]. However, no studies have yet investigated the genetic diversity of wild populations across the whole Balkan area as well as the genetic diversity of cultivated and naturalised populations.

Microsatellites are molecular markers widely used in germplasm conservation, genetic diversity analysis, studies of genetic relationships, genetic mapping, DNA fingerprinting and marker-assisted breeding [59–61]. Due to their abundance, high polymorphism, codominance, stability and suitability for automated analysis, microsatellites provide an accurate outline of the genetic structure of populations and can be used to determine plants origin and phylogeographic history. The isolation and characterization of specific *S. officinalis* microsatellite loci was recently provided by the Molecular Ecology Resources Primer Development Consortium et al. [62], Radosavljević et al. [63] and Radosavljević et al. [64].

The main objectives of this study were to analyse demographic history, genetic diversity and population structure in wild, naturalised and cultivated populations of *S. officinalis* on the Balkan Peninsula using eight microsatellite makers. We assessed the relative levels of genetic diversity of natural populations compared to planted populations in proximity and discuss the extent of genetic diversity reduction that has occurred in naturalised and cultivated populations. In order to reconstruct the species demographic history, and to better understand contemporary genetic structure of wild populations, Maximum Entropy Method along with Approximate Bayesian Computation were implemented. As an endemic as well as economically important plant species, knowledge of population genetics and demographic history of *S. officinalis* is of great importance for the effective conservation and utilization of the wild germplasm.

Materials and Methods
Population sampling

A total of 709 *S. officinalis* specimens from 30 locations across the Balkan Peninsula were sampled during flowering time between May and June in 2009. In order to analyse and compare the genetic diversity and population structure in wild, naturalised and cultivated populations and to assess the putative origin of naturalised and cultivated populations, indigenous populations were collected from 23 locations, while naturalised and cultivated populations were collected from three and four locations, respectively. Leaf tissue was collected from 20 to 25 individuals per location and immediately stored in silica gel. To avoid sampling of closely

related individuals, the minimum distance between samples was set at 20 m. Details of sampling locations and voucher information are given in Figs 1 and 2, Table 1 and S1 Appendix. No permits were required for the field study as only a few leaves of an abundant plant species were collected with no effect significant to individual and ecosystem health. The sampled species is not protected and the study was conducted on public land.

DNA extraction and microsatellites amplification

Genomic DNA samples were extracted from silica gel dried leaves using the GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich®). Eight highly informative microsatellite markers were used for the analysis: SoUZ001, SoUZ002, SoUZ003, SoUZ007, SoUZ011 [62], and SoUZ013, SoUZ014, and SoUZ019 [63]. PCR amplification reactions were performed in a total volume of 20 μL containing, 1 × PCR buffer, 1.5 mM MgCl2, 0.2 mM of each dNTP, 0.5 U of Taq HS polymerase (Takara Bio Inc.), 0.2 μM of fluorescent labelled forward primer (FAM, VIC, NED or PET), 0.2 μM of reverse primer (Applied Biosystems®) and 5 ng of genomic DNA. DNA amplification was performed using a GeneAmp PCR System 9700 (Applied Biosystems®) and a two-step PCR protocol with an initial touchdown cycle. The cycling conditions were as follows: 94°C for 5 min; 5 cycles of 45 s at 94°C, 30 s at 60°C, which was lowered
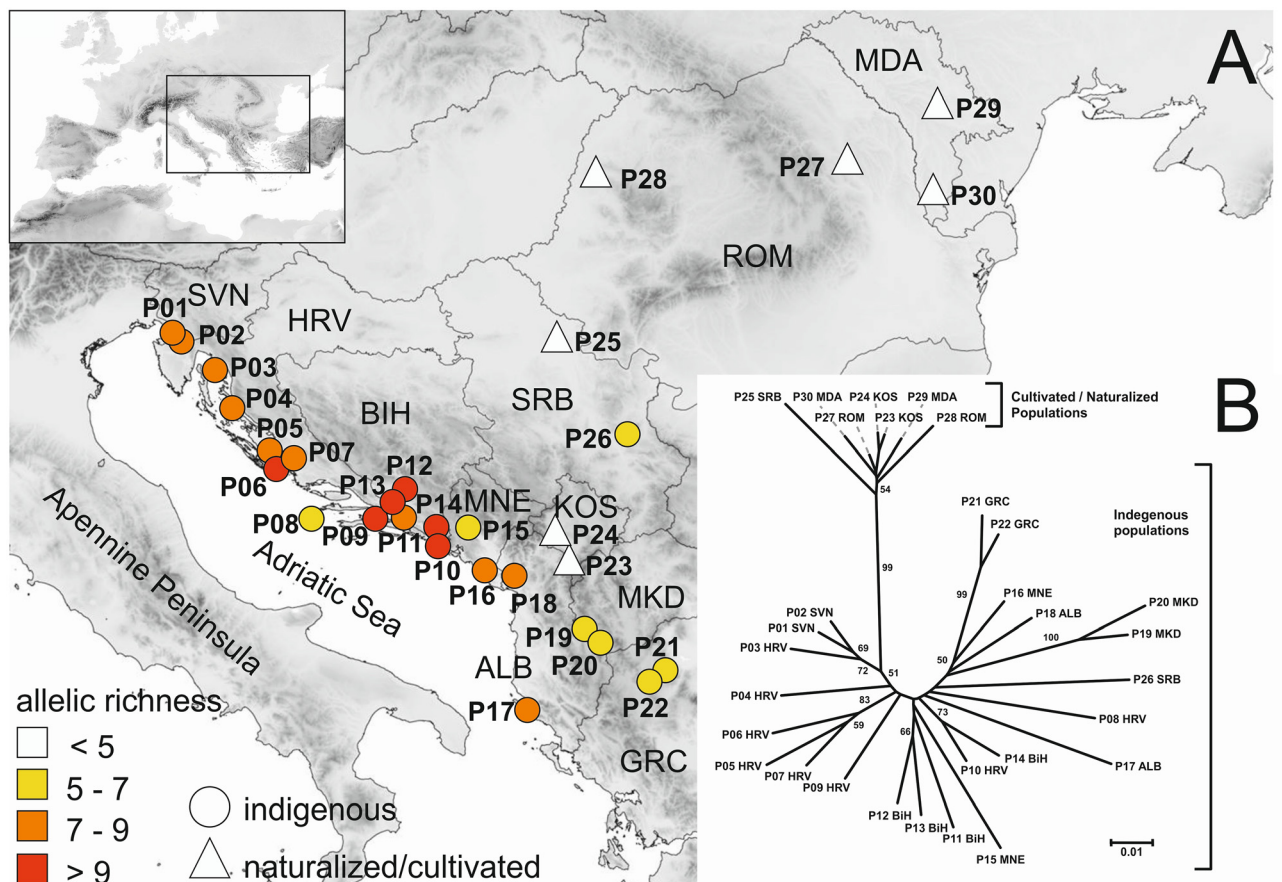
Fig 1. Within-population microsatellite diversity and genetic relationships of 30 Dalmatian sage populations. Populations are numbered as for Table 1. (A) Distribution of sampled populations: circles indicate indigenous populations; triangles indicate cultivated/naturalised populations. Symbol colours correspond to allelic richness (N_{ar}): white < 5, yellow 5–7, orange 7–9 and red >9. (B) Unrooted Fitch-Margoliash tree based on Cavalli-Sforza's chord distance. Bootstrap support values greater than 50% of 1,000 replicates are given near the branches.

doi:10.1371/journal.pone.0159545.g001
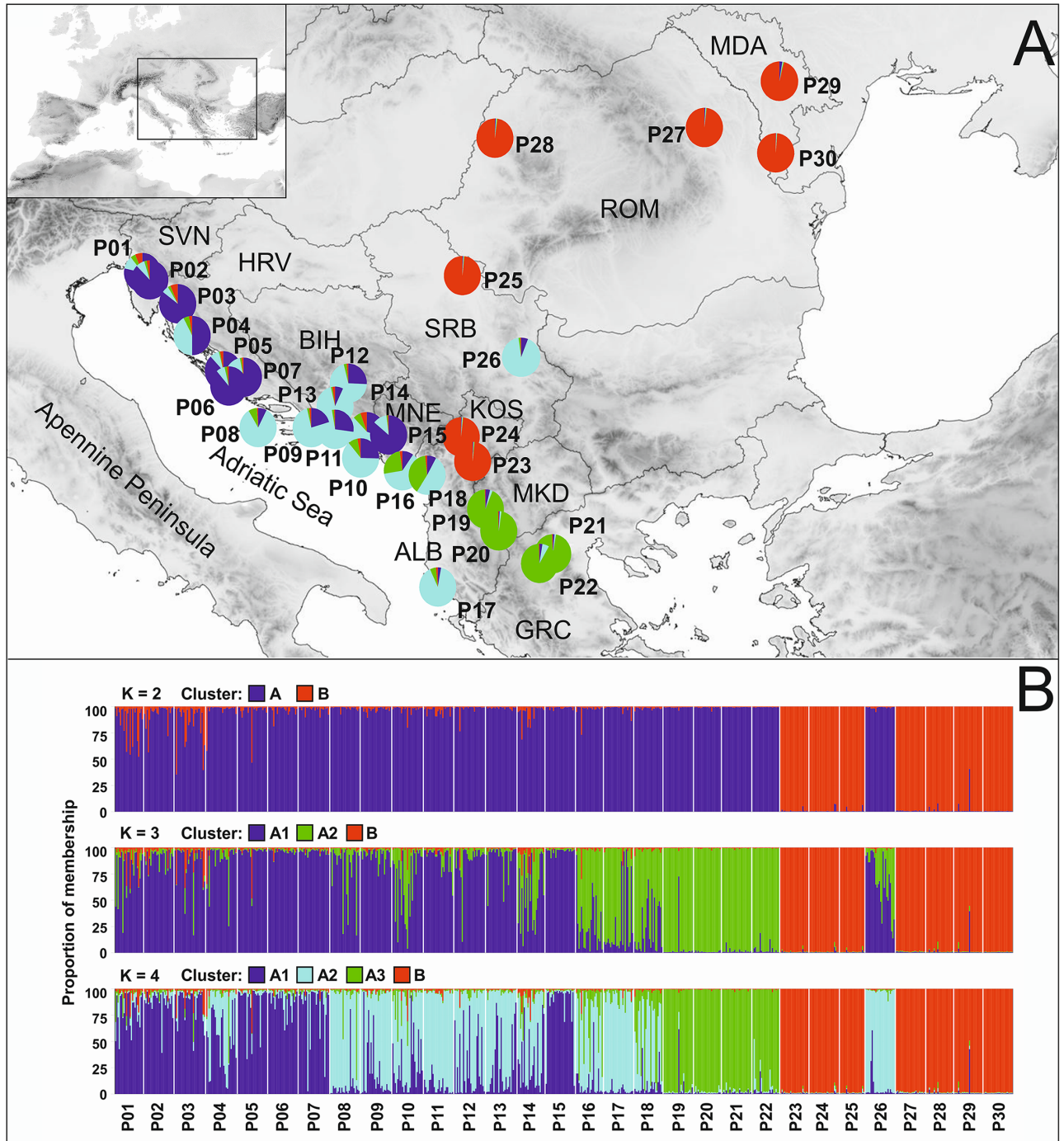
Fig 2. Genetic structure of 30 Dalmatian sage populations as estimated by the software STRUCTURE. The population numbering corresponds to Table 1. (A) Population structure assuming K = 4. (B) Proportions of membership for K = 2 to 4 clusters are given. Each individual plant is represented by a single vertical line divided into colours. Each colour represents one cluster, and the length of the coloured segment shows the individual's estimated proportion of membership in that cluster. White lines separate populations that are labelled below the figure.

doi:10.1371/journal.pone.0159545.g002

Table 1. Sampling sites, genetic diversity and genetic bottlenecks assessed from eight microsatellite markers in 30 Dalmatian sage populations.

No.	Locality	Country	n	Latitude (N)	Longitude (E)	Status	N_{av}	N_{ar}	N_{pr}	H_O	H_E	$H_{E(null)}$	F_{IS}		$P_{Bottleneck}$	M-ratio
P01	Petrinjski Kras, Soligrad	SVN	23	45.57	13.91	Indigenous	8.750	7.930	0	0.745	0.795	0.806	0.063	ns	0.422	0.730
P02	Petrinje	SVN	24	45.57	13.90	Indigenous	9.125	8.258	0	0.854	0.795	0.797	-0.075	ns	0.273	0.785
P03	Krk	HRV	25	45.23	14.57	Indigenous	8.500	7.708	0	0.717	0.734	0.748	0.024	ns	0.973	0.797
P04	Pag	HRV	25	44.43	15.04	Indigenous	8.750	7.948	1	0.754	0.759	0.766	0.008	ns	0.875	0.745
P05	Pirovac	HRV	24	43.83	15.72	Indigenous	8.000	7.107	0	0.707	0.708	0.709	0.002	ns	0.875	**0.676**
P06	Šparadići	HRV	24	43.63	15.96	Indigenous	9.875	9.071	1	0.698	0.743	0.757	0.060	ns	0.980	0.754
P07	Unešić	HRV	25	43.73	16.16	Indigenous	10.000	8.990	1	0.736	0.771	0.787	0.045	ns	0.809	0.756
P08	Vis	HRV	24	43.03	16.14	Indigenous	7.125	6.655	0	0.717	0.712	0.727	-0.007	ns	0.809	**0.673**
P09	Pelješac	HRV	25	42.98	17.27	Indigenous	11.250	9.949	2	0.754	0.768	0.770	0.018	ns	0.770	0.820
P10	Konavle	HRV	25	42.60	18.25	Indigenous	11.375	10.300	0	0.825	0.846	0.849	0.025	ns	0.473	0.847
P11	Hutovo blato	BiH	24	43.04	17.71	Indigenous	8.000	7.372	2	0.817	0.758	0.764	-0.078	ns	0.578	**0.674**
P12	Mostar	BiH	25	43.33	17.75	Indigenous	10.625	9.709	1	0.762	0.796	0.806	0.043	ns	0.963	0.866
P13	Međugorje	BiH	25	43.18	17.69	Indigenous	10.500	9.471	1	0.818	0.815	0.819	-0.004	ns	0.473	0.805
P14	Trebinje	BiH	22	42.71	18.40	Indigenous	10.625	9.847	2	0.777	0.833	0.844	0.067	ns	0.320	0.766
P15	Pješivci, Nikšić	MNE	24	42.36	19.23	Indigenous	7.375	6.721	0	0.693	0.745	0.759	0.070	ns	0.422	0.704
P16	Sutorman, Bar	MNE	22	42.15	19.12	Indigenous	8.375	7.651	1	0.699	0.718	0.725	0.026	ns	0.973	0.800
P17	Llogora	ALB	24	40.20	19.59	Indigenous	9.375	8.633	4	0.744	0.764	0.765	0.027	ns	0.994	0.681
P18	Mt. Rrenci	ALB	23	41.83	19.58	Indigenous	9.625	8.694	0	0.743	0.802	0.813	0.073	ns	0.727	0.724
P19	Mt. Jablanica	MKD	24	41.32	20.58	Indigenous	7.250	6.458	2	0.683	0.699	0.709	0.024	ns	0.680	**0.608**
P20	Mt. Karaormar	MKD	22	41.39	20.62	Indigenous	6.625	6.210	1	0.720	0.725	0.738	0.007	ns	0.273	**0.550**
P21	Lygeri	GRC	24	40.33	21.71	Indigenous	6.375	5.752	0	0.679	0.673	0.682	-0.009	ns	0.527	**0.668**
P22	Skiti	GRC	22	40.32	21.65	Indigenous	7.000	6.600	1	0.672	0.702	0.710	0.042	ns	0.191	0.757
P23	Vermicë	KOS	23	42.17	20.58	Naturalised	4.250	4.089	0	0.598	0.561	0.577	-0.067	ns	0.578	**0.535**
P24	Mirusha	KOS	24	42.52	20.57	Naturalised	4.375	4.076	0	0.505	0.561	0.598	0.100	ns	0.527	**0.552**
P25	Pančevo	SRB	20	44.85	20.72	Cultivated	2.750	2.707	0	0.313	0.377	0.422	0.171	ns	0.527	**0.493**
P26	Gradište	SRB	24	43.33	22.17	Indigenous	5.500	5.134	0	0.537	0.608	0.646	0.118	ns	0.986	0.725
P27	Bacau, Motoc	ROM	24	46.36	27.10	Cultivated	4.250	4.189	0	0.783	0.610	0.623	-0.283	***	**0.037**	**0.542**
P28	Bihor, Avram Iancu	ROM	22	46.67	21.53	Cultivated	3.750	3.639	0	0.793	0.612	0.633	-0.296	***	**0.002**	**0.455**
P29	Chishinau	MDA	23	47.36	28.85	Naturalised	4.125	3.886	0	0.592	0.595	0.610	0.004	ns	0.230	**0.543**
P30	Lopatica, Cahul	MDA	24	45.95	28.41	Cultivated	3.250	3.121	0	0.545	0.497	0.504	-0.096	ns	0.125	**0.493**
	Indigenous							7.920	115	0.733	0.751	0.761	0.025			
	Cultivated/Naturalised							3.672	0	0.590	0.545	0.567	-0.067			
	P^*							0.001		0.001	0.001		0.008			

SVN—Slovenia, HRV—Croatia, BiH—Bosnia and Herzegovina, MNE—Montenegro, ALB—Albania, MKD—Macedonia, GRC—Greece, KOS—Kosovo, SRB—Serbia, ROM—Romania, MDA—Moldova; n—sample size; N_{av}—average number of alleles; N_{ar}—allelic richness; N_{pr}—total number of private alleles; H_O—observed heterozygosity; H_E—expected heterozygosity;; $H_{E(null)}$—expected heterozygosity calculated on allele frequencies corrected for null-alleles; F_{IS}—inbreeding coefficient ([ns]–non-significant value; * significant at $P < 0.05$; ** significant at $P < 0.01$; *** significant at $P < 0.001$); $P_{Bottleneck}$—probability results of a Wilcoxon signed-ranks test used to assess population bottleneck (P-values lower than 0.05 are indicated in bold type-face); M-ratio—Garza—Williamson's M-ratio (values below the critical value, $M_c < 0.68$, are indicated in bold type-face); *P—the significance level of differences in the average values of N_{ar}, H_O, H_E and F_{IS} between groups (indigenous vs. cultivated/naturalised populations)

by 1°C in each cycle, and 90 s at 72°C; 25 cycles of 45 s at 94°C, 30 s at 55°C, and 90 s at 72°C; and an 8 min extension step at 72°C. The products were run on an ABI 3730XL (Applied Biosystems®) analyser using the commercial GeneScan service (Macrogen Inc.). The results were analysed using GeneMapper 4.0 software (Applied Biosystems®).

Data analysis

Within-population diversity. Polymorphism Information Content (*PIC*) [65] was calculated for each microsatellite marker using the PowerMarker ver. 3.23 [66] software. GENEPOP

ver. 4.0 [67] was used to estimate population genetic parameters (i.e., the average number of alleles per locus, N_{av}; the observed heterozygosity, H_O; the expected heterozygosity, H_E; inbreeding coefficient, F_{IS}) and to test population genotypic frequencies for conformation to Hardy-Weinberg (HW) expectations. Hardy-Weinberg tests for each locus in each population as well as global tests across all loci for each population were performed. A sequential Bonferroni adjustment [68, 69] was applied to correct for the effect of multiple tests using SAS Release ver. 9.1 [70].

We used the program MICRO-CHECKER [71] to check for potential problems related to allele dropout and the presence of null alleles. Estimates of null allele frequencies based on the expectation-maximization algorithm [72] were then calculated using FreeNA [73]. The adjusted allele frequencies were used to recalculate the expected heterozygosity values [$H_{E(null)}$] and compare it to the original values using Wilcoxon Two-Sample Test in SAS. Pairwise F_{ST} values were also estimated after correcting for the presence of null alleles using ENA correction method [Excluding Null Alleles; $F_{ST(null)}$] as implemented in FreeNA and compared to original values (as above).

The allelic richness, N_{ar}, was calculated as the number of alleles per locus independent of sample size using FSTAT ver. 2.9.3.2 [74, 75]. FSTAT was also used to test for the significance of differences in average values of N_{ar}, H_O, H_E, F_{IS} between groups (indigenous vs. cultivated/naturalised populations), pairwise F_{ST} (used to measure genetic differentiation between all pairs of populations) and the respective *P*-values (used to determine significant differences from zero).

We used two approaches to test for genetic bottlenecks: the heterozygosity-excess test [76] and the *M*-ratio test [77, 78]. The heterozygosity-excess method as implemented in the program BOTTLENECK ver. 1.2.02 [76, 79] was used to compare the gene diversity observed (H_E) to the gene diversity expected at mutation-drift equilibrium (H_{EQ}), calculated from the observed number of alleles under the two-phase model (TPM) assuming 22% multistep changes and variance of 11.92 as recommended by Peery et al. [80] based on empirical evidence of microsatellite mutations. Based on the number of loci in our dataset, the Wilcoxon sign-rank test [81] was chosen for the statistical analysis of heterozygote excess as recommended by Piry et al. [79]. *M*-ratio was calculated as the mean ratio of the total number of alleles (*k*) and the overall range in allele size (*r*) in each population. The *M*-ratio below the critical value of M_c < 0.68, derived from putatively stable wild populations [77], suggest a reduction in population size.

Genetic differentiation and structure. Pairwise Cavalli-Sforza's chord distances [82] were calculated and the cluster analysis was performed using the Fitch-Margoliash algorithm with 1,000 bootstraps [83] over microsatellite loci as implemented in the SEQBOOT, GENDIST, FITCH, and CONSENSE programs of the PHYLIP ver. 3.6b software package [84].

Analysis of molecular variance (AMOVA) [85] was performed using ARLEQUIN ver. 3.0 [86]. AMOVA was used to partition the total microsatellite diversity among and within populations, as well as among groups (indigenous vs. cultivated/naturalised), among populations within groups and within populations. The variance components were tested statistically by non-parametric randomisation tests using 10,000 permutations.

A model-based clustering method was applied to infer genetic structure and define the number of clusters using the software STRUCTURE ver. 2.3.4 [87]. Given a value for the number of clusters, this method assigns individual genotypes from the entire sample to clusters so that linkage disequilibrium (LD) is maximally explained. Ten runs per cluster (K), with K ranging from 1 to 11, were carried out on the Isabella computer cluster at the University of Zagreb, University Computing Centre (SRCE). Each run consisted of a burn-in period of 200,000 steps followed by 10^6 Monte Carlo Markov Chain (MCMC) replicates, assuming an admixture

model and correlated allele frequencies. The choice of the most likely number of clusters (K) was carried out by comparing the average estimates of the likelihood of the data, $ln[\Pr(X|K)]$, for each value of K [87], as well as by calculating an *ad hoc* statistic ΔK, which was based on the rate of change in the log probability of data between successive K values as described by Evanno et al. [88] and implemented in Structure-sum ver. 2011 [89].

Isolation by distance (IBD) among indigenous Dalmatian sage populations was tested using the method of Rousset [90]. A Mantel test (10^6 permutations of population locations among all locations) on the matrix of pairwise $F_{ST}/(1-F_{ST})$ ratios and that of the natural logarithm of geographical distances (in km) between pairs of populations was performed using NTSYS-pc ver. 2.02 [91].

Demographic history. In order to assess dynamics of suitable habitat distribution, 68 data of *S. officinalis* occurrence evenly distributed throughout the species distribution area on Balkan Peninsula were used. All data were geocoded in WGS84 coordinate system. The potential present and Last Glacial Maximum (LGM) distribution was modelled by MAXENT ver. 3.3.3k, an algorithm used for identifying species' suitable environmental space from incomplete information of occurrence [92]. Present ecological niche model (ENM) was based on WorldClim bioclimatic variables [93] at a resolution of 30 arc-seconds. Correlation among bioclimatic variables was tested using a Pearson's correlation matrix with IBM SPSS Statistics ver. 19.0 (IBM, Armonk, NY, USA). For modelling, only variables with correlation lower than 0.70 were used. In order to produce a model with the best set of bioclimatic variables and the appropriate value of regularization multipliers, model selection in ENMTools [94] was performed. As the best model, the one with the lowest AICc (Akaike information criteria, corrected for small samples) value was chosen, as suggested by Warren and Seifert [95]. Twenty replicates for the model with cross-validation and 10.000 background points were done. For the LGM predictions, the same bioclimatic variables used for present ENM and based on MIROC and CCSM models (http://www.worldclim.org/past; Community Climate System Model from the National Center for Atmospheric Research, Boulder, Colorado, USA) at a resolution of 2.5 arc-minutes were used. Both models were used to create the final average model using raster calculator in ArcGis ver. 10.1. (Esri, Redlands, CA, USA). To obtain binary maps of habitat suitability, "maximum training sensitivity plus specificity" threshold was used [96]. All models were post-processed and visualised in ArcGIS ver. 10.1. (Esri, Redlands, CA, USA).

The evolutionary history of Dalmatian sage on Balkan Peninsula was inferred using Approximate Bayesian Computation (ABC) [97]. We performed coalescent simulations with DIYABC ver. 2.1 [98] software to test alternative historic scenarios of divergence and admixture. We defined three wild populations based on the results of STRUCTURE analysis (clusters A1, A2 and A3) including only individuals displaying a probability of belonging to each of the cluster over 0.90 ($Q > 0.90$). Populations P15 Pješivci, Nikšić and P26 Gradište were omitted from the analysis being out of the region in which the most of the individuals assigned to a given cluster was sampled. Thus, the ABC analysis was carried out based on: Pop1 Northwestern Adriatic cluster (A1; P01-P07; 92 individuals), Pop2 Southern Adriatic cluster (A2; P08-P14; P16-P18; 91 individuals), and Pop3 Macedonian/Greek cluster (A3; P19-P22; 82 individuals). Five simple historic scenarios were specified and explored: (1) Scenario 1—Population Pop1 is derived from population Pop2, itself derived from population Pop3; (2) Scenario 2—Population Pop3 is derived from population Pop2, itself derived from population Pop1; (3) Scenario 3—Both populations Pop1 and Pop3 derived independently from population Pop2; (4) Scenario 4—Population Pop2 was generated by admixture of populations Pop1 and Pop3; and (5) Scenario 5—All three populations diverged at the same time. Detailed description of scenarios is given in the Supporting Information (S2 Appendix). Wide prior parameter distribution were used for all the parameters including current effective population sizes (N_1, N_2, N_3;

uniform; 10–10,000), times of the events counted in generations (t_1, t_2; uniform; 1–100,000), ancestral effective population size (N_A; uniform; 10–10,000), and admixture rate (r; uniform; 0.001–0.999). Additional conditions were imposed for the sequence order of historic events ($t_2 > t_1$) and for the demographic model ($N_A > N_1, N_2, N_3$). The Generalized Stepwise Mutation model of mutation was applied (GSM) [99] with broad prior parameter distribution for mean mutation rate (μ; uniform; 10^{-5}–10^{-3}), mean value of the parameter of the geometric distribution (P; uniform; 0.10–0.30) and mean SNI (single nucleotide insertions/deletions) rate (μ_{SNI}; log-uniform; 10^{-8}–10^{-5}). One million simulations were generated for each explored scenario and the 10,000 simulations closest to observed dataset were used to estimate the posterior probabilities and distributions of parameters. Alternative scenarios were evaluated by comparing posterior probabilities using logistic regression. The reliability of the most likely scenario was evaluated by performing model checking [100] based on 10,000 data sets simulated from the posterior distributions of the following parameters (test quantities; t): mean number of alleles (N_a) and mean expected heterozygosity (H_E) [101] for each populations as well as N_a, H_E, F_{ST}, mean individual assignment likelihoods (classification index) [102] and shared-allele distance (D_{AS}) [103] between pairs of populations. The discrepancies between simulated and observed data were measured by the cumulative distribution function values of each test quantity (t) defined as the probability of $t_{simulated}$ being lower than $t_{observed}$. Principal component analysis (PCA) was performed considering 10,000 data sets simulated with parameter values drawn from prior and the observed data set as well as the 10,000 data sets simulated from the posterior distributions of parameters were added to each plane.

Results

Within-population diversity

The eight microsatellite loci yielded a total of 165 alleles, ranging from 13 (SoUZ013 and SoUZ007) to 30 (SoUZ001) (S3 Appendix). The PIC values ranged from 0.63 to 0.94, with an average of 0.81 (S3 Appendix).

Out of a total of 240 tests, the presence of null alleles was suggested in eleven population/locus combinations (4.58%); the frequencies of null alleles ranged from 0.05 (SoUZ019 in P16) to 0.18 (SoUZ0002 in P07) (S4 Appendix).

The main parameters describing within-population diversity of 30 Dalmatian sage populations are shown in Table 1. The allelic richness varied from 3.12 (P30 from Moldova) to 10.63 (P12 and P14 both from Bosnia and Herzegovina) (Fig 1). Indigenous populations had significantly higher allelic richness than the cultivated/naturalised populations (7.92 vs. 3.67; $P <$ 0.001). Private alleles were observed exclusively in indigenous populations, and in 13 populations a total of 20 private alleles were detected. The highest number of private alleles (4) was observed in population P17 from Albania. Comparing the two main groups, a total of 115 alleles detected in indigenous populations were not present in cultivated/naturalised populations, while cultivated/naturalised populations had no private alleles.

From 240 instances (30 populations, eight loci), 38 significant deviations ($P < 0.05$) from the HWE were observed. Two cultivated populations from Romania (P27 and P28) showed significant deviations from the HWE at more than two loci (four and six, respectively). After applying sequential Bonferroni corrections no significant departures from the HWE were detected at any loci in any population except in case of locus SoUZ007 in P11 (Hutovo blato, BiH). Eight out of 30 populations showed significant deviations from the HWE by applying global tests across all loci in each population before sequential Bonferroni correction, whereas only two remained significant after correction (P27 and P28; Table 1).

Indigenous populations exhibited significantly higher values of both observed (H_O: 0.73 vs. 0.59; P < 0.001) and expected heterozygosity (H_E: 0.75 vs. 0.54; P < 0.001). In the indigenous populations, the values of H_E varied from 0.61 to 0.85 while in the cultivated/naturalised populations, the values were substantially lower, ranging from 0.37 to 0.61. The expected heterozyosity values [$H_{E(null)}$] corrected for the presence of null alleles were only slightly higher than the original values (Tab. 1) and no significant differences were observed between H_E and $H_{E(null)}$ in any population ($P_{Wilcoxon}$ = 0.22–0.50) suggesting that null alleles did not have substantial impact on the results.

The heterozygosity-excess method implemented in BOTTLENECK (Wilcoxon signed rank test assuming a two-phased model) identified significant bottleneck events only in two cultivated populations from Romania (P27 and P28; Table 1). In contrast, the *M*-ratio test detected strong signals of population size reduction in all cultivated/naturalized populations (P23-P25; P27-P30) with *M*-ratios (0.493–0.552) far below the critical value of 0.68. Moreover, the analysis of southernmost wild populations from Macedonia (P19, P20) and Greece (P21) also yielded the *M*-ratios below but close to the critical value (0.550–0.668). Finally, the *M*-ratios slightly lower than the critical value were also observed in three Croatian populations (P05, P03, P11) ranging from 0.673 to 0.676.

Genetic differentiation and structure

All pairwise F_{ST} values between populations were significant ($P < 0.05$) except between two neighbouring indigenous populations from Slovenia (P01/P02), between two neighbouring indigenous populations, one from Croatia (P10) and the other from Bosnia and Herzegovina (P14), and between two neighbouring naturalised populations from Kosovo (P23/P24). The lowest F_{ST} value was observed between two indigenous populations from Slovenia (P01/P02; F_{ST} = 0.003), while the highest value was found between two populations from Serbia (the cultivated population P25 and the indigenous P26; F_{ST} = 0.41). Global F_{ST} values without correction for null alleles vs. F_{ST} values using ENA [$F_{ST(null)}$] were 0.153 (95%-confidence interval; CI = 0.140–0.168) and 0.149 (CI = 0.136–0.163), respectively (S5 Appendix). No significant differences were observed between F_{ST} and $F_{ST(null)}$ ($P_{Wilcoxon}$ = 0.257), suggesting that the presence of putative null alleles did not affect this analysis.

The unrooted Fitch-Margoliash tree based on Cavalli-Sforza's chord distance between 30 Dalmatian sage populations is shown in Fig 1B. Indigenous populations grouped together in accordance with the geographical position of the collection sites, from Slovenia in the northwest to Greece in the south-east region of the study area. The seven cultivated/naturalised populations grouped separately from the rest and formed a well-supported clade (bootstrap support 99%), suggesting a common origin of the cultivated plant material.

The AMOVA showed that most of the total genetic diversity was attributable to differences between individuals within populations (84.73%) (Table 2). However, the significant ϕ-value among populations suggested the existence of population differentiation. Two-way AMOVA revealed that only a minority of variations in the genetic diversity was explained by differences between groups of populations (indigenous vs. cultivated/naturalised; 11.94%; ϕ_{CT} = 0.110; P(ϕ_{CT}) < 0.0001), thus confirming the structuring of populations according to cultivation status. Excluding the seven cultivated/naturalised populations, the among-population component of genetic diversity was 10.80% [ϕ_{ST} = 0.108; P(ϕ_{ST}) < 0.0001], suggesting further geographic structuring of indigenous populations.

The results of the STRUCTURE analysis are shown in Fig 2. The average estimate of the likelihood of the data, conditional on a given number of clusters, ln[Pr(X|K)], increased with increased K, as did the standard deviations among different runs for each K. The highest ΔK

Table 2. Analysis of molecular variance for the partitioning of microsatellite diversity.

Analysis	Source of variation	df	Variance components	% Total variance	ϕ-statistics	$P(\phi)$
A	Among populations	29	0.494	15.25	0.152	< 0.0001
	Within populations	1388	2.744	84.75		
B	Among groups	1	0.419	11.94	0.119	< 0.0001
	Among population within groups	28	0.342	9.77	0.111	< 0.0001
	Within populations	1388	2.744	78.29	0.217	< 0.0001
C	Among populations	22	0.355	10.80	0.108	< 0.0001
	Within populations	1075	2.931	89.20		
D	Among populations	6	0.293	12.25	0.122	< 0.0001
	Within populations	313	2.101	87.75		

(A) Among and within 30 Dalmatian sage populations, (B) between groups (indigenous vs. cultivated/naturalised), among populations within groups and within populations, (C) among and within 23 indigenous populations and (D) among and within seven cultivated/naturalised populations. $P(\phi)$ - ϕ-statistics probability level after 10,000 permutations.

doi:10.1371/journal.pone.0159545.t002

value was observed for K = 2 (3183.82), followed by that for K = 4 (4.10) (S6 Appendix). The average proportion of membership of each individual in each cluster was calculated for K = 2 to K = 4 based on the run with the highest ln[Pr(X|K)]. At K = 2, all the individuals belonging to indigenous populations were assigned to a separate cluster (cluster A) from individuals belonging to cultivated/naturalised populations (cluster B). At K = 4, three subclusters of indigenous populations appeared, thus confirming the geographical structuring of wild populations. Subcluster A1 was predominant in the north-west part of the Adriatic coast and included Slovenian populations (P01 and P02) and five out of eight Croatian populations (P03-P07), as well as a population from Montenegro (P15). Subcluster A2 was found primarily in the southern part of the Adriatic coast and included the most southern Croatian populations (P07-P09), all the populations from Bosnia and Herzegovina (P11-P14), a population from Montenegro (P16) and two Albanian populations (P17 and P18). The indigenous population from Serbia (P26) clearly belongs to the same cluster. The majority of individuals belonging to Macedonian (P19 and P20) and Greek (P21 and P22) populations were assigned to subcluster A3.

The isolation by distance analysis revealed that the correlation between matrices of genetic $[F_{ST}/(1-F_{ST})$ ratios] and geographical $[ln(km)]$ distances was significant ($r = 0.48$; $P_{Mantel} <$ 0.0001), suggesting that 22.50% of the variance could be explained by isolation-by-distance (Fig 3).

Demographic history

The MAXENT models were build using the following bioclimatic variables (the percent of contribution of each variable for the construction of the model is indicated in the brackets): BIO1 (Annual Mean Temperature; 3.69%), BIO3 (Isothermality; 0.12%), BIO7 (Temperature Annual Range; 0.21%), BIO8 (Mean Temperature of Wettest Quarter; 0%), BIO9 (Mean Temperature of Driest Quarter; 23.86%), BIO14 (Precipitation of Driest Month; 20.86%) and BIO19 (Precipitation of Coldest Quarter; 51.26%). Logistic threshold was set according to "maximum training sensitivity plus specificity" at the value of 0.262. Model selection chose the model with regularization multipliers value 5 as the best model. All resulting models are shown in Fig 4. The generated model for the present (Fig 4A) was congruent with *S. officinalis* natural distribution in Balkan Peninsula, as the entire eastern Adriatic coast was characterized by medium to high levels of the environmental suitability. MIROC and CCSM models for the potential species distribution during the LGM differed substantially from each other, especially in prediction of the
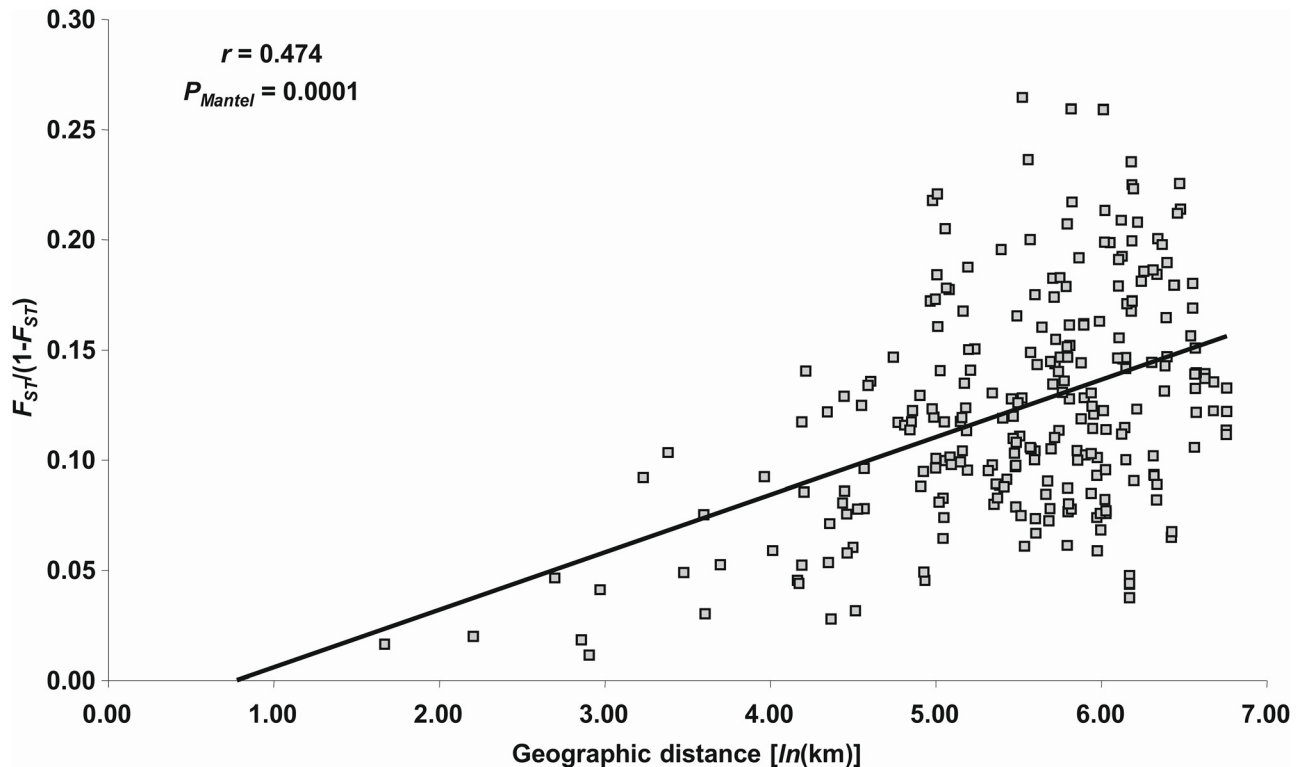
Fig 3. Isolation by distance analysis among 23 indigenous Dalmatian sage populations.

doi:10.1371/journal.pone.0159545.g003

areas of low and medium suitability. MIROC results implied the possibility of existence of multiple refugia throughout the studied region as considerable geographic areas were characterized by suitable environmental conditions for survival of LGM (Fig 4C). On the other hand, CCSM simulation pointed only to parts of southern Adriatic coastal region as the area of medium to high LGM environmental suitability (Fig 4B).

The coalescent-based ABC analyses showed that the median of the posterior probability (PP) estimated for scenario 5 (PP 0.440; 95% CI: 0.365–0.514) was higher than those for scenario 1 (0.166), 2 (0.193), 3 (0.075) and 4 (0.126) with 95% confidence intervals never overlapping those of the other scenarios (S2 Appendix). Scenario 5 included a simple split model in which all three populations diverged from the ancestral population t generation ago. The median values of effective population size of Pop1, Pop2, Pop3, and the ancestral populations were N_1 = 4,330, N_2 = 7,190, N_3 = 2,440, and N_A = 8,550, respectively (S2 Appendix). Estimated median time of divergence was 525 generations ago (95% CI: 157–1,540; 90% CI: 191–1,270). The estimates of model parameters are given for each scenario in S2 Appendix. The model checking of Scenario 5 revealed that none of the 24 test quantities [mean number of alleles (N_a) and mean expected heterozygosity (H_E) for each populations as well as N_a, H_E, F_{ST}, mean individual assignment likelihoods and shared-allele distance (D_{AS}) between pairs of populations] had tail-area probability values (i.e. *P*-values) lower than 0.05 (S7 Appendix) indicating a good fit of the selected model. The results of the Principal component analysis (PCA; S8 Appendix) provided further evidence for the validity of the Scenario 5, since the PCA points of the test quantities generated from the posterior distribution were clearly centred on the target point corresponding to the target point (i.e. observed data set).
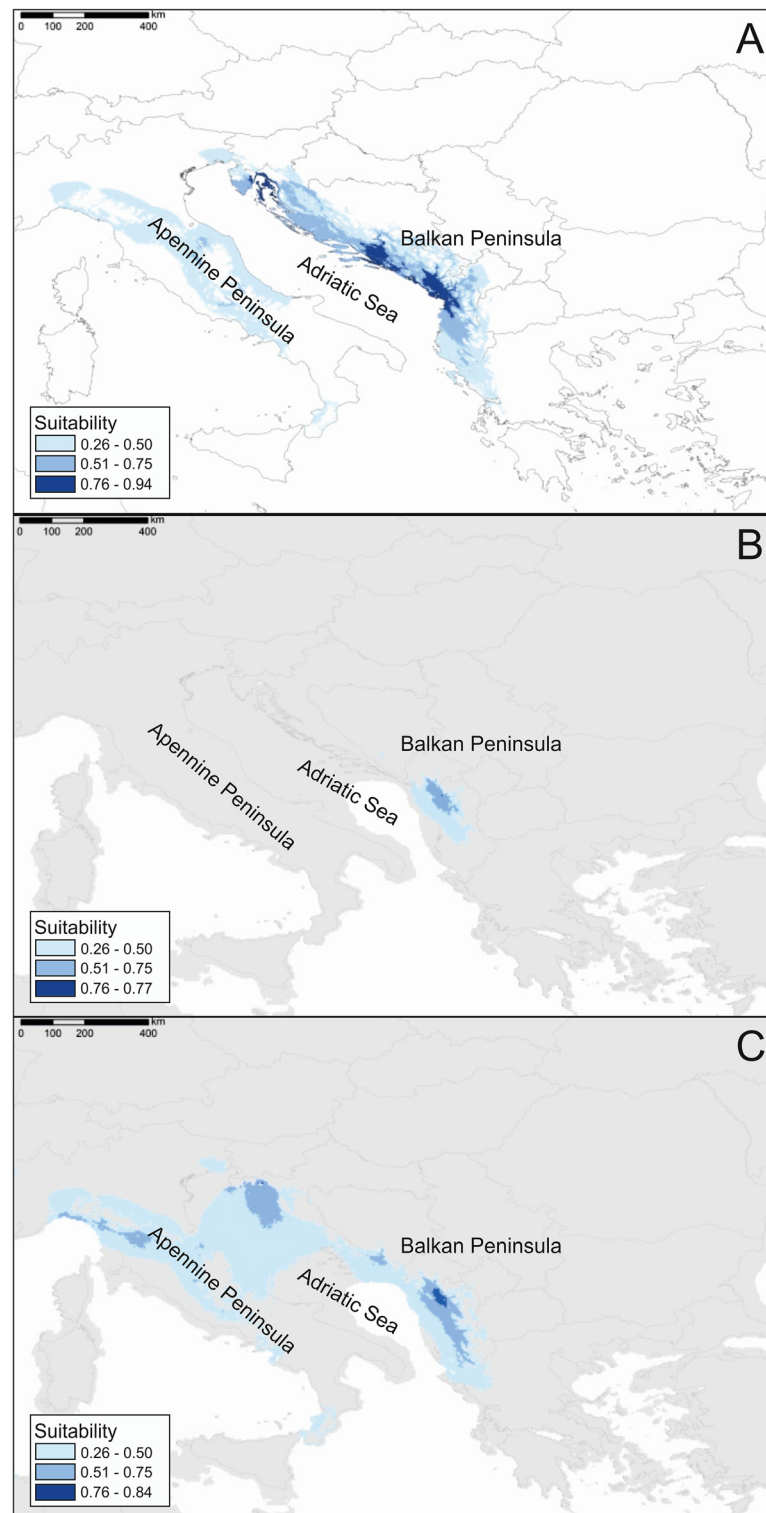
Fig 4. Ecological Niche Modelling of *Salvia officinalis* under (a) present, (b) Last Glacial Maximum CCSM conditions and (c) Last Glacial Maximum MIROC conditions. The colour codes indicate habitat suitability, from light blue (low suitability) to dark blue (high suitability) as calculated by the model.

doi:10.1371/journal.pone.0159545.g004

Discussion

Genetic diversity in wild and cultivated/naturalised populations

The analysis of 30 populations of *S. officinalis* using eight microsatellite markers revealed a high degree of genetic diversity in 23 indigenous populations and significantly lower genetic diversity in seven cultivated/naturalised populations; the two groups on the Balkan Peninsula could clearly be separated according to their origin. The pattern of genetic diversity between wild and cultivated populations of *S. officinalis* observed herein was consistent with other domesticated species [7, 9, 104, 105]. Indigenous populations exhibited significantly higher allelic richness and private alleles were found exclusively in indigenous populations (Table 1, Fig 1). Moreover, cultivated/naturalised populations had significantly lower values of both observed and expected heterozygosity (Table 1). These findings are typical for domesticated species as the cultivation of wild plants always produces genetic bottlenecks, and thus results in loss of genetic diversity because of founder effects and unconscious or conscious artificial selection [106]. However, concerns about the impacts of cultivation bottlenecks are more often expressed when considering economically valuable crops with long histories of domestication, while the genetic impact of cultivation on medicinal plants is usually less pronounced due to shorter cultivation histories, weaker selection pressure and multiple origins [107, 108]. In this study, several lines of evidence suggest that all analysed cultivated/naturalised *S. officinalis* populations originated from the same geographic area, i.e., from the northernmost wild populations of the Balkan Peninsula. Such a uniform geographical origin is suggestive of planned human-mediated movement of seeds and cuttings that probably took place quite recently. *Salvia officinalis* is known to have been cultivated on the Balkan Peninsula since Roman times [109], for such older established cultivated and naturalised populations, multiple geographic origins from adjacent wild populations would have been expected to exist. As cultivated/naturalized populations sampled in our study are located in the central and easternmost part of the Balkan Peninsula we assume that the source of all studied cultivated/naturalised populations is the Institute of Medicinal Plants Research "Dr Josif Pančić", Belgrade, Serbia, which served as a leading MAP breeding institution in the Balkans during the second half of the 20th century. Therefore, to the best of our knowledge, the Institute organised the collection of seeds and/or cuttings not from the nearby southern indigenous populations, but from the northernmost indigenous populations. Despite the probable origin of cultivated material from one restricted geographical area, all cultivated/naturalised populations exhibited similar ϕ_{ST} values relative to indigenous populations (0.122 vs. 0.108, respectively, Table 2) and the majority of overall genetic variation in both groups was explained by differences between individuals within populations (AMOVA, Table 2), which is characteristic for outcrossing and long-lived plants. It is expected that the cultivation of wild plants always leads to reduction in genetic diversity and the severe loss of rare alleles [108, 110]. Our results suggest that a genetic bottleneck has occurred in cultivated/naturalized populations but the results are not unequivocal. While *M*-ratio tests indicate potential bottleneck events in all cultivated/naturalized populations, heterozygosity-excess tests point out that only two cultivated populations from Romania show signatures of population bottlenecks. The ambiguous results suggest that different approaches differ substantially in statistical power depending on timing, duration and severity of the bottleneck event as well as pre-bottleneck genetic diversity [78, 80]. The diverse propagule sources and the outcrossing nature might enable the cultivated populations to regain mutation-drift equilibrium very rapidly hampering the detection of bottlenecks based on heterozygosity-excess tests in contrast to *M*-ratio tests [78]. Moreover, heterozygosity-excess tests are less powerful than *M*-ratio tests when pre-bottleneck genetic diversity is high as shown by the simulation study reported by the simulation study reported by Peery et al. [80].

Biogeographic implications

The results of our study showed evidence of genetic differentiation between indigenous populations (Fig 2) that was most likely because of range shifts imposed by Pleistocene climatic fluctuations. The genetic data presented here support the supposition that central and southern Adriatic coastal region is the centre of diversity for S. *officinalis* and that populations extending to the northwestern Adriatic coast and further south and east on the Balkan Peninsula have lost some of that diversity (Fig 1). S. *officinalis* populations were more diverse in central and southern Adriatic coastal region, as shown by the significantly higher genetic diversity and greater number of private alleles found in that region (Figs 1 and 2). The greater number of private alleles may indicate the existence of glacial refugia [111, 112]; studies of the evolution of other Balkan endemic species have also suggested the presence of such glacial refugia [5, 113–115]. Additionally, both CCSM and MIROC simulations of the LGM climatic conditions also supported the possibility that southern Adriatic coastal region was the refugium area. The lower genetic diversity in the northern and southernmost populations (subclusters A1 and A3, respectively; Fig 2) could be explained by colonization from the core southern populations. The loss of genetic diversity (N_{ar}, N_{pr}, Table 1) in these populations may thus be the result of bottlenecks during migration events [116]. The geographical structuring of indigenous populations is also supported by the IBD analysis showing a significant correlation between geographic distance and genetic variation (Fig 3); this type of pattern is often associated with postglacial colonization [117, 118]. The second most plausible scenario suggests the possibility of three separate refugia during the LGM, corresponding to the three subclusters obtained by STRUCTURE (Fig 2). The concurrent divergence into three genetic groups is also proposed by the ABC analysis (S2 Appendix). Such paleodistribution scenario was only partially supported by the LGM distribution modelling, as only the MIROC simulation gave pattern consistent with ABC analysis results. CCSM and MIROC simulations differ in severe temperature decline modelled by CCSM. Therefore, substantial differences between the two projections are frequently found (e.g. [119, 120, 121]). Acknowledging the fact that microsatellite data are not well suited for estimation of divergence times and absolute dating, it is difficult to temporally scale the spatial divergence of S. *officinalis*. Moreover, there is no data regarding the S. *officinalis* generation time or life span. However, a closely related species, S. *fruticosa* Mill., characterized by similar life traits, is a long lived chamaephyte with individuals living up to 300 years [122]. If we assume similar for S. *officinalis*, the divergence time of detected genetic groups could be placed in Pleistocene indicating that species survived LGM in multiple refugia on Balkan Peninsula. To achieve more solid temporal reconstructions of phylogeographic relationships between genetic groups it is necessary to acquire dating results based on different phylogeographic approaches.

The recent establishment of non-native populations is further supported by the lack of admixture between indigenous and cultivated/naturalised populations. It appears that not enough time has passed since establishment to detect gene flow and genetic homogenization between proximate naturalised and wild populations (e.g., naturalised P23 and P24 and wild P18-20, Fig 2). However, admixture was present among indigenous lineages involving populations P16 and P18 from Montenegro and Albania, respectively, which exhibited characteristics of A2 and A3 gene pools. Similarly, P04 from the island Pag in Croatia and P13 from Bosnia and Herzegovina showed an admixed pattern comprising A1 and A2 gene pools (Fig 2). These results are in accord with Liber et al. [56], who used RAPD markers to show that the central and southern Adriatic populations with the highest diversity and the lowest frequency down-weighted values [123] were at the same time the most admixed populations. Liber et al. [56] suggested that this region is the main contact zone between descendants of

adjacent multiple Pleistocene microrefugia (the 'refugia within refugia model', [124]; confirmed for the Balkans by, e.g., [5, 114, 115, 125]) with recent unrestricted gene flow among large and connected populations. The data presented here supports this view and provides further evidence that the northern Adriatic and the southernmost and eastern populations (not included in the Liber et al. [56] study) represent rear-edge populations characterised by lower genetic diversity, lower allelic richness and fewer private alleles. The disjunct inland P26 population in Serbia exhibited the lowest observed and expected heterozygosities among indigenous populations, a low value of allelic richness (H_O = 0.537, H_E = 0.608, N_{ar} = 5.134, Table 1), and no private alleles. However, the clustering analysis (Fig 2) unequivocally grouped this population into subcluster A2 together with the core southern populations with the highest genetic diversity. These data indicate that the inland P26 population likely resulted from post-glacial recolonization, rather than being a remnant relict and refugial population as postulated by Stojanović et al. [58]. A similar pattern is observed in the most isolated island population P08 on Vis in Croatia, which was included in subcluster A2 by the clustering analysis but has low allelic richness and no private alleles (N_{ar} = 6.655, Table 1); although this population did have relatively high observed and expected heterozygosity (H_O = 0.717, H_E = 0.712, Table 1).

Conservation and breeding efforts

Though it is generally believed that the quality and composition of essential oils in MAP are influenced by various environmental conditions and habitats where the plants are grown and harvested e.g., [42], several studies of *S. officinalis* revealed the dependence of variations in essential oil composition on genetic background [44, 126]. Correspondingly, our result of the STRUCTURE analysis, which separates natural populations in three subclusters (A1, A2 and A3), is mostly in accordance with essential-oil composition (chemotypes D, C and A) as determined by Cvetkovic et al. [46]. By comparison of these two studies, we can conclude that there is a high congruence between groupings based on chemical composition and genetic relatedness. Although the boundaries are somewhat blurred, most of the populations in the subcluster A1 are characterised by high content of camphor and β pinene and low content of both *cis*- and *trans*-thujones, the populations in subcluster A2 by high content of *cis*- and *trans*-thujones and camphor, and the populations in subcluster A3 by *cis*-thujone and camphor and low content of *trans*-thujone. The presented relationship between the genetic profile and chemical composition of the indigenous populations could be an important step in future breeding programs and cultivation, as indigenous populations represent an indispensable source of genetic diversity that is conspecific with the cultivated gene pool. Wild populations can be utilised to improve cultivated populations by introducing genetic diversity through sexual reproduction and consequently increase the efficiency of artificial selection for desired traits. Although the wild *S. officinalis* cannot be treated as an endangered species because populations are numerous and characterised by high levels of genetic diversity, it is essential to conserve the most diverse natural populations for future breeding programmes through efficient management. Special effort should be taken in conservation of populations identified in this study as containing the highest genetic diversity and having the greatest number of private alleles. Rare alleles are often considered a minor element in genetic conservation programmes but they can be very important for the long-term response to selection and the survival of populations and species [110]. Populations from STRUCTURE subcluster A2 have the highest reservoir of genetic diversity. Hence, these populations are the most adequate for future breeding programmes. The

contradiction of protection and commercial gathering should be addressed through controlled reproduction of wild individuals following appropriate guidelines and *in situ* management should take into consideration the results presented here to ensure an efficient way to conserve desirable agronomic traits.

Supporting Information

S1 Appendix. Microsatellite information, voucher information and microsatellite genotyping raw data of eight loci scored in 30 Dalmatian sage populations.
(XLS)

S2 Appendix. Historic scenarios of Dalmatian sage on Balkan Peninsula explored using Approximate Bayesian Computation: description of scenarios, models used in DIYABC, posterior probabilities (PP) after logistic regression on the 10,000 simulations (1% of the total) closest to observed dataset and the estimates of parameters including effective population sizes (N_1, N_2, N_3), times of the events counted in generations (t_1, t_2), and ancestral effective population size (N_A). Median values and 95% confidence intervals are given for each parameter. Pop1 represents the Northwestern cluster, Pop2 the Southern cluster and Pop3 Macedonian/Greek cluster.
(PDF)

S3 Appendix. Allelic diversity of eight microsatellite loci scored in 30 Dalmatian sage populations.
(PDF)

S4 Appendix. Null-allele frequencies, estimated using FreeNA, for population/locus combinations identified as containing a large number of null alleles by MICRO-CHECKER.
(PDF)

S5 Appendix. Pairwise F_{ST} values (lower diagonal) and null-allele corrected F_{ST} values [$F_{ST (null)}$; upper diagonal) among 30 Dalmatian sage populations.
(XLS)

S6 Appendix. The choice of the most likely number of clusters (K) inferred from multilocus microsatellite data using a model-based clustering method of Pritchard et al. (2000): *ln* P (X|K) values for each of the ten independent runs for each K and ΔK values for each K (shown on logarithmic scale) based on the second order rate of change of the likelihood function with respect to K described by Evanno et al. (2005).
(PDF)

S7 Appendix. Historic scenario of Dalmatian sage on Balkan Peninsula explored using Approximate Bayesian Computation; the model checking of Scenario 5: Comparison of 24 test quantities based on the observed data set and 10,000 data sets simulated from the posterior distributions of parameters.
(PDF)

S8 Appendix. The model checking of Scenario 5 explored using Approximate Bayesian Computation: Principal component analysis (PCA) of test quantities calculated from (a) 10,000 data sets simulated with parameter values drawn from prior distributions of parameters (Prior), (b) 10,000 data sets simulated with parameter values drawn from the posterior distributions of parameters (Posterior), and (c) from observed data set (Observed).
(PDF)

Author Contributions

Conceived and designed the experiments: ZS DB ZDS AI MJD ZL GS. Performed the experiments: ZL KCS DG I. Radosavljević I. Rešetnik. Analyzed the data: ZS M. Grdiša. Contributed reagents/materials/analysis tools: DB DBR KCS PC ZDS M. Gonceariuc DG AI MJD EK ZL SM DP I. Radosavljević GS DS IS. Wrote the paper: I. Rešetnik ZL ZS.

References

1. Meyer RS, Purugganan MD. Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet. 2013; 14: 840–852. doi: 10.1038/nrg3605 PMID: 24240513

2. Lange D. Europe's medicinal and aromatic plants: their use, trade and conservation. Cambridge; TRAFFIC International; 1998.

3. Hamilton AC. Medicinal plants, conservation and livelihoods. Biodivers Conserv. 2004; 13: 1477–1517.

4. Cruse-Sanders JM, Hamrick JL. Genetic diversity in harvested and protected populations of wild American ginseng, *Panax quinquefolius* L. (Araliaceae). Am J Bot. 2004; 91: 540–548. doi: 10.3732/ajb.91.4.540 PMID: 21653409

5. Grdiša M, Liber Z., Radosavljević I, Carović-Stanko K, Kolak I, Satovic Z. Genetic Diversity and Structure of Dalmatian Pyrethrum (*Tanacetum cinerariifolium* Trevir. /Sch./ Bip., Asteraceae) within the Balkan Refugium. Plos One. 2014; 9: e105265. doi: 10.1371/journal.pone.0105265 PMID: 25121763

6. Zohary D. Unconscious selection and the evolution of domesticated plants. Econ Bot. 2004; 58: 5–10.

7. Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL, Costa JM, et al. Impacts of genetic bottlenecks on soybean genome diversity. Proc Natl Acad Sci USA. 2006; 103: 16666–16671. PMID: 17068128

8. Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS. Selection versus demography: A multilocus investigation of the domestication process in maize. Mol Biol Evol. 2004; 21: 1214–1225. PMID: 15014173

9. Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C, et al. Grinding up wheat: a massive loss of nucleotide diversity since domestication. Mol Biol Evol. 2007; 24: 1506–1517. PMID: 17443011

10. Olsen KM, Gross BL. Detecting multiple origins of domesticated crops. Proc Natl Acad Sci USA. 2008; 105: 13701–13702. doi: 10.1073/pnas.0807439105 PMID: 18784363

11. Diamond J. Evolution, consequences and future of plant and animal domestication. Nature. 2002; 418: 700–707. PMID: 12167878

12. Small E. Hybridization in the domesticated-weed-wild complex. In: Grant WF, editor. Plant Biosystematics. Toronto: Academic Press; 1984. pp. 195–210.

13. Ellstrand NC. Current knowledge of gene flow in plants: implications for transgene flow. Philos Trans R Soc Lond B Biol Sci. 2003; 358: 1163–1170. PMID: 12831483

14. Ellstrand NC, Prentice HC, Hancock JF. Gene flow and introgression from domesticated plants into their wild relatives. Annu Rev of Ecol Evolu Syst. 1999; 30: 539–563.

15. Haygood R, Ives AR, Andow DA. Consequences of recurrent gene flow from crops to wild relatives. Proc R Soc Lond B Biol Sci. 2003; 270: 1879–1886.

16. Pilson D, Prendeville HR. Ecological effects of transgenic crops and the escape of transgenes into wild populations. Annu Rev of Ecol Evolu Syst. 2004; 35: 149–174.

17. Alziar G. Catalogue synonymique des *Salvia* L. du monde (Lamiaceae). I à VI. Biocosme Mésogéen, Nice. 1988–1993; 5: 87–136; 6: 79–115; 6: 163–204; 7: 59–109; 9: 413–497; 10: 33–117.

18. Harley RM, Atkins S, Budantsey AL, Cantino PD, Conn BJ, Grayer R, et al. Labiatae. In: Kadereit JW, editor. The families and genera of vascular plants 7, Lamiales. Berlin: Springer; 2004. pp. 167–282.

19. Walker JB, Sytsma KJ, Treutlein J, Wink M. *Salvia* (Lamiaceae) is not monophyletic: implications for the systematics, radiation, and ecological specializations of *Salvia* and tribe Mentheae. Am J Bot. 2004; 91: 1115–1125. doi: 10.3732/ajb.91.7.1115 PMID: 21653467

20. Walker JB, Sytsma KJ. Staminal evolution in the genus *Salvia* (Lamiaceae): molecular phylogenetic evidence for multiple origins of the staminal lever. Ann Bot. 2007; 100: 375–391. PMID: 16926227

21. Will M, Claßen-Bockhoff R. Why Africa matters: evolution of Old World *Salvia* (Lamiaceae) in Africa. Ann Bot. 2014; 114: 61–83. doi: 10.1093/aob/mcu081 PMID: 24966353

22. Hedge IC. Salvia L. In: Davis PH, editor. Flora of Turkey and the East Aegean Islands. Vol. 7. Edinburgh: University Press; 1982. pp 400–461.

23. Pignatti S. (ed.) Flora d' Italia, vol. 2. Bologna: Edagricole; 1982.

24. Di Pietro R. New dry grassland associations from Ausoni-Aurunci mountains (central Italy)—syntaxonomical updating and discussion on the higher rank syntaxa. Hacquetia. 2011; 10: 183–231.

25. Cuvelier ME, Richard H, Berset C. Antioxidative activity and phenolic composition of pilot-plant and commercial extracts of sage and rosemary. J Am Oil Chem Soc. 1996; 73: 645–652.

26. Baricevic D, Bartol T. The biological pharmacological activity of the *Salvia* genus V. In: Kintzios SE editor. The Genus *Salvia*. Amsterdam: Harwood Academic Publishers; 2000. pp. 143–184.

27. Amr S, Djordjevic S. The investigation of the quality of sage (*Salvia officinalis* L.) originating from Jordan. Facta Univ Ser Work Living Environ Prot. 2000; 1: 103–108.

28. Baricevic D, Sosa S, Della Loggia R, Tubaro A, Simonovska B, Krasna A, et al. Topical anti-inflammatory activity of *Salvia officinalis* L. leaves: the relevance of ursolic acid. J Ethnopharmacol. 2001; 75: 125–132. PMID: 11297842

29. Delamare APL, Moschen-Pistorello IT, Artico L, Atti-Serafini L, Echeverrigaray S. Antibacterial activity of the essential oils of *Salvia officinalis* L. and *Salvia triloba* L. cultivated in South Brazil. Food Chem. 2007; 100: 603–608.

30. Pinto E, Salgueiro LR, Cavaleiro C, Palmeira A, Gonçalves MJ. *In vitro* susceptibility of some species of yeasts and filamentous fungi to essential oils of *Salvia officinalis*. Ind Crop Prod. 2007; 26: 135–141.

31. Bouaziz M, Yangui T, Sayadi S, Dhouib A. Disinfectant properties of essential oils from *Salvia officinalis* L. cultivated in Tunisia. Food Chem Toxicol. 2009; 47: 2755–2760. doi: 10.1016/j.fct.2009.08.005 PMID: 19682532

32. Tada M, Okuno K, Chiba K, Ohnishi E, Yoshii T. Antiviral diterpenes from *Salvia officinalis*. Phytochemistry. 1994; 35: 539–541.

33. Todorov S, Philianos S, Petkov V, Harvala C, Zamfirova R, Olimpiou H. Experimental pharmacological study of three species from genus *Salvia*. Acta Physiol Pharmacol Bulg. 1984; 10: 13–20.

34. Eidi A, Eidi M. Antidiabetic effect of sage (*Salvia officinalis* L.) leaves in normal and streptozotocin-induced diabetic rats. Diab Metabol Synd: Clin Res & Rev. 2009; 3: 40–44.

35. Mayer B, Baggio KH, Freitas KS, Santos AC, Twardowschy A, Horst H, et al. Gastroprotective constituents of *Salvia officinalis* L. Fitoterapia 2009; 80: 421–426. doi: 10.1016/j.fitote.2009.05.015 PMID: 19481590

36. Ninomiya K, Matsuda H, Shimoda H, Nishida N, Kasajima N, Yoshino T, et al. Carnosic acid, a new class of lipid absorption inhibitor from sage. Bioorg Med Chem Lett. 2004; 14: 1943–1946. PMID: 15050633

37. Armitage AM. Herbaceous perennial plants. Champaign, Illinois: Stipes Publishing; 1997.

38. Corsi G, Bottega S. Glandular hairs of *Salvia officinalis*: new data on morphology, localization and histochemistry in relation to function. Ann Bot. 1999; 84: 657–664.

39. Baričević D, Bernath J, Maggioni L, Lipman E. Report of a working group on medicinal and aromatic plants. Rome: International Plant Genetic Resources Institute; 2002.

40. Böszörményi A, Héthelyi E, Farkas Á, Horváth G, Papp N, Lemberkovics É, et al. Chemical and genetic relationships among sage (*Salvia officinalis* L.) cultivars and Judean sage (*Salvia judaica* Boiss.). J Agric Food Chem. 2009; 57: 4663–4667. doi: 10.1021/jf9005092 PMID: 19449812

41. Lipman E. Report of a working group on medicinal and aromatic plants. Rome: Biodiversity International; 2009.

42. Perry NB, Anderson RA, Brennan NJ, Douglas MH, Hesney AJ, McGimpsey JA. Essential oils from dalmatian sage (*Salvia officinalis* L.): variation among individuals, plant parts, seasons and sites. J Agric Food Chem. 1999; 47: 2048–2054. PMID: 10552494

43. Bernotiené G, Nivinskiené O, Butkiené R, Mockuté D. Essential oil composition variability in sage (*Salvia officinalis* L.). Chemija. 2007; 18: 38–43.

44. Jug-Dujaković M, Ristić M, Pljevljakušić D, Dajić-Stevanović Z, Liber Z, Hančević K, et al. High diversity of indigenous populations of Dalmatian Sage (*Salvia officinalis* L.) in essential oil composition. Chem Biodivers. 2012; 9: 2309–2323. doi: 10.1002/cbdv.201200131 PMID: 23081929

45. Stešević D, Ristić M, Nikolić V, Nedović M, Caković D, Satovic Z. Chemotype Diversity of Indigenous Dalmatian Sage (*Salvia officinalis* L.) Populations in Montenegro. Chem Biodivers. 2014; 11: 101–114. doi: 10.1002/cbdv.201300233 PMID: 24443430

46. Cvetkovic I, Stefkov G, Karapandzova M, Kulevanova S, Satovic Z. Essential Oils and Chemical Diversity of Southeast European Populations of *Salvia officinalis* L. Chem Biodivers. 2015; 12: 1025–1039. doi: 10.1002/cbdv.201400273 PMID: 26172323

47. Kuštrak D, Kuftinec J, Blazevic N. Yields and composition of sage oils from different regions of the Yugoslavian Adriatic Coast. J Nat Prod. 1984; 47: 520–524.

48. Maksimović M, Vidić D, Miloš M, Šolić ME, Abadžić S, Šiljak-Yakovlev S. Effect of the environmental conditions on essential oil profile in two Dinaric *Salvia* species: *S. brachyodon* Vandas and *S. officinalis* L. Biochem. Syst Ecol. 2007; 35: 473–478.

49. Lakušić B, Ristić M, Slavkovska V, Stojanović D, Lakušić D. Variations in essential oil yields and compositions of *Salvia officinalis* (Lamiaceae) at different developmental stages. Bot Serb. 2013; 37: 127–139.

50. Santos-Gomes PC, Fernandes-Ferreira M. Organ and seasoned-pendent variation in the essential oil composition of *Salvia officinalis* L. cultivated in two different sites. J Agric Food Chem. 2001; 49: 2908–2916. PMID: 11409986

51. Piccaglia R, Marotti M, Galletti GC. Effect of Mineral Fertilizers on the Composition of *Salvia officinalis* Oil. J Essent Oil Res. 1989; 1: 73–83.

52. Venskutonis PR. Effect of the drying on the volatile constituents of thyme (*Thymus vulgaris* L.) and sage (*Salvia officinalis* L.). Food Chem. 1997; 59: 219–227.

53. Veličković DT, Randjelović ND, Ristić MS, Veličković AS, Šmelcerović AA. Chemical constituents and antimicrobial activity of the ethanol extract obtained from the flower, leaf and stem of *Salvia officinalis* L. J Serb Chem Soc. 2003; 68: 17–24.

54. Mastelić J. The essential oil co-distillation by superheated vapour of organic solvents from aromatic plants. Flavour Fragrance J. 2001; 16: 370–373.

55. Židovec V. Varijabilnost prirodnih populacija mirisave kadulje (Salvia officinalis L.) (Variability of indigenous populations of Dalmatian sage (Salvia officinalis L.)). PhD Thesis, Faculty of Agriculture, University of Zagreb. 2004.

56. Liber Z, Židovec V, Bogdanović S, Radosavljević I, Pruša M, Filipović M, et al. Genetic Diversity of Dalmatian Sage (*Salvia officinalis* L.) as Assessed by RAPD Markers. Agric Conspec Sci. 2014; 79: 77–84.

57. Jug-Dujaković M. Genetska i biokemijska raznolikost ljekovite kadulje (Salvia officinalis L.) (Genetic and biochemical diversity of Dalmatian sage (Salvia officinalis L.)). PhD Thesis, Faculty of Agriculture, University of Zagreb. 2010.

58. Stojanović D, Aleksić JM, Jančić I, Jančić R. A Mediterranean medicinal plant in the continental Balkans: A plastid DNA-based phylogeographic survey of *Salvia officinalis* (Lamiaceae) and its conservation implications. Willdenowia. 2015; 45:103–118.

59. Jarne P, Lagoda PJL. Microsatellites, from molecules to populations and back. Trends Ecol Evol. 1996; 11: 424–429. PMID: 21237902

60. Sunnucks P. Efficient genetic markers for population biology. Trends Ecol Evol. 2000; 15: 199–203. PMID: 10782134

61. Chistiakov DA, Hellemans B, Volckaert FAM. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. Aquaculture. 2006; 255: 1–29.

62. Molecular Ecology Resources Primer Development Consortium et al. Isolation and characterization of polymorphic microsatellites loci from common sage (*Salvia officinalis* L. (Lamiaceae)). Mol Ecol Resour. 2010; 10: 404–408.

63. Radosavljević I, Jakse J, Javornik B, Šatović Z, Liber Z. New microsatellite markers for *Salvia officinalis* (Lamiaceae) and cross-amplification in closely related species. Am J Bot. 2011; 98: e316–e318. doi: 10.3732/ajb.1000462 PMID: 22003176

64. Radosavljević I, Šatović Z, Jakse J, Javornik B, Greguraš D, Jug-Dujaković M, et al. Development of New Microsatellite Markers for *Salvia officinalis* L. and Its Potential Use in Conservation-Genetic Studies of Narrow Endemic *Salvia brachyodon* Vandas. Int J Mol Sci. 2012; 13: 12082–12093. doi: 10.3390/ijms130912082 PMID: 23109901

65. Botstein D, White RL, Sholnick M, David RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980; 32: 314–331. PMID: 6247908

66. Liu J. Powermarker—A Powerful Software for Marker Data Analysis. Relaigh: North Carolina State University Bioinformatics Research Center. 2002. Available: www.powermarker.net

67. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 1995; 86: 248–249.

68. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979; 6: 65–70.

69. Rice WR. Analyzing tables of statistical tests. Evolution. 1989; 43: 223–225.

70. SAS Institute. SAS/STAT® 9.1 User's Guide. Cary, NC: SAS Institute Inc.; 2004.

71. van Oosterhout C, Weetman D, Hutchinson WF. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Mol Ecol Notes. 2006; 6: 255–256.

72. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc B. 1977; 39: 1–38.

73. Chapuis MP, Estoup A. Microsatellite null alleles and estimation of population differentiation. Mol Biol Evol. 2007; 24: 621–631. PMID: 17150975

74. Goudet J. FSTAT (vers. 1.2): a computer program to calculate F-statistics. J Hered. 1995; 86: 485–486.

75. Goudet J. FSTAT. A program for Windows to estimate and test gene diversities and fixation indices. Version 2.9.3. Available: www.unil.ch/izea/softwares/fstat.html

76. Cornuet JM, Luikart G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics. 1996; 144: 1119–1127.

77. Garza J, Williamson E. Detection of reduction in population size using data from microsatellite loci. Mol Ecol. 2001; 10: 305–318. PMID: 11298947

78. Williamson-Natesan EG. Comparison of methods for detecting bottlenecks from microsatellite loci. Conserv Genet. 2005; 6: 551–562.

79. Piry S, Luikart G, Cornuet JM. BOTTLENECK: A computer programme for detecting recent reductions in the effective population size using allele frequency data. J Hered. 1999; 89: 502–503.

80. Peery MZ, Kirby R, Reid BN, Stoelting R, Doucet-Bëer E, Robinson S, et al. Reliability of genetic bottleneck tests for detecting recent population declines. Mol Ecol. 2012; 21: 3403–3418. doi: 10.1111/j.1365-294X.2012.05635.x PMID: 22646281

81. Luikart G, Allendorf FW, Cornuet JM, Sherwin WB. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered. 1998; 89: 238–247. PMID: 9656466

82. Cavalli-Sforza LL, Edwards AW. Phylogenetic analysis: Models and estimation procedures. Am J Hum Genet. 1967; 19: 233–257. PMID: 6026583

83. Felsenstein J. Confidence limits on phylogenesis: An approach using the bootstrap. Evolution. 1985; 39: 783–791.

84. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.5c. Seattle: Department of Genetics, University of Washington; 1993.

85. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitocondreal DNA restriction sites. Genetics. 1992; 131: 479–491. PMID: 1644282

86. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online. 2005; 1: 47–50.

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

88. 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

89. Ehrich D, Gaudeul M, Assefa A, Koch MA, Mummenhoff K, Nemomissa S, et al. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. Mol Ecol. 2007; 16: 2542–2559. PMID: 17561912

90. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics. 1997; 145: 1219–1228. PMID: 9093870

91. Rohlf FJ. NTSYS-PC 2.02i, Numerical Taxonomy and Multivariate Analysis System. Setauket: Exeter Software; 1997.

92. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modelling of species geographic distributions. Ecol Model. 2006; 190: 231–259.

93. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis. A Very high resolution interpolated climate surfaces for global land areas. Int J Climatol. 2005; 25: 1965–1978.

94. Warren DL, Glor E, Turelli M. ENMTools: A toolbox for comparative studies of environmental niche models. Ecography. 2010; 33: 607–611.

95. Warren DL, Seifert S. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl. 2011; 21: 335–342. PMID: 21563566

96. Liu C, Berry PM, Dawson TP, Pearson RG. Selecting thresholds of occurrence in the prediction of species distributions. Ecography. 2005; 28: 385–393.

97. Beaumont MA, Zhang W-Y, Balding DJ. Approximate Bayesian computation in population genetics. Genetics. 2002; 162: 2025–2035. PMID: 12524368

98. Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, et al. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 2014; 30: 1187.

99. Estoup A, Jarne PJ, Cornuet JM. Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. Mol Ecol. 2002; 11:1591–1604. PMID: 12207711

100. Cornuet JM, Ravigné V, Estoup A. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC Bioinformatics. 2010; 11: 401. doi: 10.1186/1471-2105-11-401 PMID: 20667077

101. Nei M. Molecular evolutionary genetics. New York, NY, USA: Columbia University Press; 1987.

102. Rannala B, Mountain JL. Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci U S A 1997; 94: 9197–9201. PMID: 9256459

103. Chakraborty R, Jin L. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. In: Pena SDJ, Chakraborty R, Epplen JT, Freys AJ, editors. DNA fingerprinting: state of the science. Basel, Germany: Birkhauser; 1993. pp. 153–175.

104. Buckler ES IV, Thornsberry JM, Kresovich S. Molecular diversity, structure and domestication of grasses. Genet Res. 2001; 77: 213–218. PMID: 11486504

105. Bourguiba H, Audergon J-M, Krichen L, Trifi-Farah N, Mamouni A, Trabelsi S, et al. Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. BMC Plant Biol. 2012; 12: 49. doi: 10.1186/1471-2229-12-49 PMID: 22510209

106. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. Cell. 2006; 127: 1309–1321. PMID: 17190597

107. He J, Chen L, Si Y, Huang B, Ban X, Wang Y. Population structure and genetic diversity distribution in wild and cultivated populations of the traditional Chinese medicinal plant *Magnolia officinalis* subsp. *biloba* (Magnoliaceae). Genetica. 2009; 135: 233–243. doi: 10.1007/s10709-008-9272-8 PMID: 18500655

108. Yuan Q, Zhang Z, Hu J, Guo L, Shao A, Huang L. Impacts of recent cultivation on genetic diversity pattern of a medicinal plant, *Scutellaria baicalensis* (Lamiaceae). BMC Genet. 2010; 11: 29. doi: 10.1186/1471-2156-11-29 PMID: 20429879

109. Dweck AC. The folklore and cosmetic use of various *Salvia* species. In: Kintzios SE, editor. Sage. The genus *Salvia*. Amsterdam: Harwood Academic; 2000. pp. 1–25.

110. Allendorf FW, Luikart G. Conservation and the Genetics of Populations. Oxford: Blackwell Publishing; 2007.

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

112. Teixeira H, Rodríguez-Echeverría S, Nabais C. Genetic Diversity and Differentiation of *Juniperus thurifera* in Spain and Morocco as Determined by SSR. PLoS ONE. 2014; 9(2): e88996. d doi: 10.1371/journal.pone.0088996 PMID: 24533164

113. Caković D, Stešević D, Schönswetter P, Frajman B. How many taxa? Spatiotemporal evolution and taxonomy of *Amphoricarpos* (Asteraceae, Carduoideae) on the Balkan Peninsula. Org Divers Evol. 2015; 25: 429–445.

114. Kutnjak D, Schönswetter P, Dullinger S, Kuttner M, Niketić M, Frajman B. Escaping to the summits: phylogeography and predicted range dynamics of *Cerastium dinaricum*, an endangered high mountain plant endemic to the western Balkan Peninsula. Mol Phylogenet Evol. 2014; 78: 365–374. doi: 10.1016/j.ympev.2014.05.015 PMID: 24857887

115. Surina B, Schönswetter P, Schneeweiss GM. Quaternary range dynamics of ecologically divergent species (*Edraianthus serpyllifolius* and *E. tenuifolius*, Campanulaceae) within the Balkan refugium. J Biogeogr. 2011; 38: 1381–1393.

116. Hewitt GM. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc. 1996; 58: 247–276.

117. Ibrahim KM, Nichols RA, Hewitt GM. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity. 1996, 77: 282–291.

118. Pope LC, Domingo-Roura X, Erven K, Burke T. Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. Mol Ecol. 2006; 15: 371–386. PMID: 16448407

119. Garcia-Porta J, Litvinchuk SN, Crochet PA, Romano A, Geniez PH, Lo-Valvo M, et al. Molecular phylogenetics and historical biogeography of the west-palearctic common toads (*Bufo bufo* species complex). Mol Phylogenet Evol. 2012; 63: 113–130. doi: 10.1016/j.ympev.2011.12.019 PMID: 22214922

120. Rebelo H, Froufe E, Brito JC, Russo D, Cistrone L, Ferrand N, et al. Postglacial colonization of Europe by the barbastelle bat: agreement between molecular data and past predictive modelling. Mol Ecol. 2012; 21: 2761–2774. doi: 10.1111/j.1365-294X.2012.05566.x PMID: 22490279

121. Fernández-Mazuecos M, Vargas P. Congruence between distribution modelling and phylogeographical analyses reveals Quaternary survival of a toadflax species (*Linaria elegans*) in oceanic climate areas of a mountain ring range. New Phytol. 2013; 198: 1274–1289. doi: 10.1111/nph.12220 PMID: 23496320

122. Rivera D, Obon C, Cano F. The botany, history and traditional uses of three-lobed sage (*Salvia fruticosa* Miller) (Labiatae). Econ Bot. 1994; 48:190–195.

123. Schönswetter P, Tribsch A. Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). T axon. 2005; 54: 725–732.

124. Gómez A., Lunt DH. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N, editors. Phylogeography of southern European refugia. Berlin: Springer; 2007. pp. 155–188.

125. Surina B, Pfanzelt S, Einzmann HJR, Albach DC. Bridging the Alps and the Middle East: evolution, phylogeny and systematics of the genus *Wulfenia* (Plantaginaceae). Taxon. 2014; 63: 843–858.

126. Putievsky E, Ravid U, Sanderovich D. Morphological observations and essential oils of sage (*Salvia officinalis* L.) under cultivation. J Essent Oil Res. 1992; 4: 291–293.