

SCIENTIFIC REPORTS



OPEN

Morphological, genetic and epigenetic aspects of homoploid hybridization between *Salvia officinalis* L. and *Salvia fruticosa* Mill.

Ivan Radosavljević^{1,2}, Sandro Bogdanović^{2,3}, Ferhat Celep⁴, Maja Filipović¹, Zlatko Satovic^{2,5}, Boštjan Surina^{6,7} & Zlatko Liber^{1,2}

The inheritance of phenotypic, genetic and epigenetic traits in hybridization events is difficult to predict, as numerous evolutionary, ecological, and genetic factors can play a crucial role in the process of hybridization. In the middle Adriatic island of Vis, we investigated hybridization between *Salvia officinalis* and *S. fruticosa* at morphological, genetic and epigenetic levels. SSR results revealed that hybrid individuals were characterized by diploid set of chromosomes suggesting homoploid hybridization. A well-defined group that mostly comprised of F₁ generation individuals was detected. For the majority of analysed morphological characteristics, hybrids were placed in-between parental taxa, while at the same time, values of different genetic parameters were mostly higher in hybrids than in parental species. The results revealed a high contrast in the levels of phenotypic variability and epigenetic excitation between parental taxa. Environmental niche modelling confirmed that in the studied location *S. officinalis* experiences optimal climatological conditions, while *S. fruticosa* struggles with unsuitable conditions. Very low levels of gene flow between the parental species were detected. In addition, contrasting levels of epigenetic excitation in the studied groups clearly demonstrated the importance of an epigenetic response to an altered environment and confirmed the trans-generational nature of the epigenetic changes.

Introduction of plant species into new areas occurs naturally, but the frequency of this phenomenon has increased dramatically because of human activity during the last few millennia. Consequently, hybridization between congeneric species occurs more often, causing significant genetic changes in both the native and alien species¹. Hybridization is one of the fundamental forces of evolution, and this process can result in a diverse range of outcomes, including hybrid divergence and sympatric speciation, introgression, hybrid zone formation, erosion of species boundaries and possible species extinction through assimilation^{2,3}. New lineages of hybrid origin are characterized either by allopolyploidy, meaning that the complete genomes of the parental species have joined, or homoploidy. As defined by Nieto Feliner *et al.*⁴, homoploid hybrid speciation “is the formation of a new-hybrid-species, independent from its parents, via hybridization with no whole-genome duplication and thus no increase in ploidy”. Although this definition seems to be straightforward, key aspects of hybridization process are under debate^{4,5}. Schumer *et al.*⁵ proposed following criteria that must be met to confirm inter-species homoploid hybridization: reproductive isolation of hybrids must be recognized, there must be evidences that this isolation was triggered by hybridization itself and hybridization must be confirmed at a genomic level. However,

¹University of Zagreb, Faculty of Science, Department of Biology, Division of Botany, Marulićev trg 9A, HR, 10000, Zagreb, Croatia. ²Centre of Excellence for Biodiversity and Molecular Plant Breeding (CroP-BioDiv), Svetošimunska cesta 25, HR, 10000, Zagreb, Croatia. ³University of Zagreb, Faculty of Agriculture, Department of Agricultural Botany, Svetošimunska cesta 25, HR, 10000, Zagreb, Croatia. ⁴Department of Biology, Faculty of Arts and Sciences, Kirikkale University, Kirikkale, Turkey. ⁵University of Zagreb, Faculty of Agriculture, Department of Seed Science and Technology, Svetošimunska cesta 25, HR, 10000, Zagreb, Croatia. ⁶Natural History Museum Rijeka, Lorenzov prolaz 1, HR, 51000, Rijeka, Croatia. ⁷University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technologies, Glagoljaška 8, SI, 6000, Koper, Slovenia. Correspondence and requests for materials should be addressed to Z.L. (email: zlatko.liber@biol.pmf.hr)

in a recently published paper, Nieto Feliner *et al.*⁴ argues these criteria, especially the importance of reproductive isolation as a crucial criterion as an evidence of occurrence of a homoploid hybrid speciation.

Once an organism finds itself in an altered environment, its genotype can produce numerous phenotypes, of which some will presumably be better adapted to the new ecological conditions. This phenomenon, known as phenotypic plasticity, has an integral role in an organism's attempt to cope with environmental variation and enables plants to respond to such variations within their lifetime^{6,7}. Such expression of different phenotypes in changing environments may be caused by alteration of gene expression without changes in the primary sequence of the DNA, i.e., epigenetics^{8,9}. Molecular mechanisms that are considered responsible for modifications in gene expression include histone and chromatin modifications and methylation, all of which are directed by small RNAs¹⁰. Thanks to transgenerational inheritance of epigenetic variants, phenotypic responses to an altered environment remain memorized and at disposal for generations to come¹¹. Interspecies hybridization triggers not only genetic but also epigenetic changes through modification of DNA methylation patterns^{12,13}. Consequently, new phenotypic variants found in hybrid offspring more often originate from epigenetic than genetic recombination, and thanks to their transgenerational heritability, they can play a very important role in the process of speciation^{11,14,15}.

With approximately 1000 recorded species, the genus *Salvia* is the largest genus of the Lamiaceae family¹⁶. Thirty-six taxa are present in Europe, and these taxa are divided into seven sections¹⁷. The *Salvia* section includes 13 species, with the common or Dalmatian sage (*Salvia officinalis* L.) as its type species. Another easily recognizable member of this section is the Greek sage (*S. fruticosa* Mill.). Both species are well known due to their high essential oil content and commercial value^{18–20}. The natural distribution range of common sage includes the eastern Adriatic coast from the Gulf of Trieste in the north to northwestern Greece in the south^{21,22}, as well as parts of the central and southern Apennines^{23–25}. Greek sage extends from Cyrenaica, Sicily, and southern Italy through the southern part of the Balkan Peninsula to West Syria^{21,26}. Because both species have been highly prized for their numerous beneficial healing properties, Phoenicians, Greeks and Romans have introduced these species into many Mediterranean countries throughout history²⁶. Ancient Mediterranean cultures were prone to transfer and grow useful plants near their settlements and colonies scattered across the Mediterranean basin. Consequently, sub-spontaneous populations and dubiously native populations of certain species can today be found in different locations where the ancient colonies once settled^{27–29}.

In the central Adriatic region, a small and isolated population of Greek sage can be found on the island of Vis (Croatia)³⁰, where the plant grows sympatrically with common sage. Because of the well-organized Greek colony Issa that was founded on the island of Vis in the 4th century BC^{31–33}, it can be assumed that this disjunct population of Greek sage located approximately 500 km from the nearest population of its type is of sub-spontaneous origin rather than being a member of an indigenous flora.

Hybridization within the section *Salvia* has been previously recorded (*S. lavandulifolia* Vahl. subsp. *vellerea* × *S. officinalis*³⁴, *S. officinalis* × *S. lavandulifolia* subsp. *lavandulifolia*³⁵, *S. tomentosa* Mill. × *S. officinalis*¹⁸); thus, it appears that erosion of species boundaries between closely related species of the section is present. Hybridization between common sage and Greek sage has never been documented to occur spontaneously in nature but was artificially performed for breeding purposes in Israel^{18,19,36}. An obtained commercial sage hybrid (*Salvia officinalis* × *Salvia fruticosa* cv. *Newe Ya'ar* No. 4) was well suited for intensive agricultural conditions in Israel, but in terms of essential oil quality, it was more similar to common sage than Greek sage¹⁹.

Although the distributions of both species overlap in parts of southern Albania and northern Greece²⁶, hybridization between these species in the region seems unlikely because of the different flowering periods³⁷. However, in the western coastal area of the Island of Vis, common and Greek sage have overlapping flowering periods from late April to early May, and hybridization between the two species is frequent, as a full spectrum of intermediate individuals sharing morphological traits of both parental taxa can be observed. Greek sage and hybrids grow in dense and admixed stands, whereas common sage is present sporadically. However, in the surrounding area, common sage grows abundantly as a typical representative of the indigenous flora. In addition, it is important to note that both species share the same number of chromosomes ($2n = 14$)²⁶ and rely on the same pollinator assemblage^{38,39}.

Because of their high variability, codominant inheritance and high reproducibility and reliability, for the last two decades microsatellites have been one of the most popular molecular markers used in many different areas, from genome mapping^{40,41} to population and conservation genetics⁴². However, unlike universal molecular markers (e.g. AFLP⁴³) that require no prior knowledge regarding target species genome, microsatellites are species-specific and require *de novo* isolation for any species being studied for the first time⁴⁴. Also, in situations of weak population differentiation, an analysis based on limited number of microsatellite loci lacks the statistical power for individual population assignment⁴⁵. Since this limitation can be overcome only by significant increase of the number of used loci⁴⁶ which may be technically challenging and time consuming⁴⁵, the utilization of AFLP technique (Amplified fragment length polymorphism) in such situations represents a good alternative⁴⁵. However, AFLPs are also characterized by some major drawbacks when compared to microsatellites. The most important ones are homoplasy⁴⁷ and their dominant nature⁴⁵. While the homoplasy (i.e. either non-homologous fragments share the same or very similar size or independent fragment loss occur) decreases the resolution and reliability of the analysis⁴⁷, AFLPs dominance makes it impossible to assess the levels of departure from Hardy-Weinberg equilibrium⁴⁵. In summary, microsatellites and AFLPs both have certain advantages against each other that should be weighted when choosing the most appropriate genotyping technique for a specific study. The MSAP technique⁴⁸ (Methylation sensitive amplification polymorphism) as the most widely used tool in epigenetic studies of natural populations⁴⁹, is a modification of the AFLP method. In the majority of MSAP based studies, the rare cutter *EcoRI* restriction enzyme was used as in the AFLP analysis, while the frequent cutter *MseI* was replaced by *HpaII* and *MspI* methylation-sensitive restriction enzymes⁵⁰. Since these two methods are technically very similar and their results can be easily compared to each other, they are methods of choice in studies of both genetic and epigenetic diversity of the same target species.

The main goal of this study was to perform a comprehensive analysis of hybridization between common sage and Greek sage at morphological, genetic and epigenetic levels. For this purpose, in sampled individuals, we analysed 39 morphometric traits and performed DNA analysis using microsatellite and AFLP nuclear genetic markers as well as MSAP epigenetic markers. By exploring the different aspects of the hybridization event, we aimed to (1) obtain insight into the extent of hybridization between common sage and Greek sage on a phenotypic level; (2) determine the prevalence of hybridization and assess the nature and levels of inter-species gene flow (if present); (3) assess the levels of methylation variability in hybrids and parental groups; and (4) discuss the complexity of the interactions among the environment, phenotype, genotype and epigenotype through the prism of core vs. marginal population.

Results

Morphological diversity. The mean Shannon's diversity index based on 23 qualitative traits was significantly different among taxa ($P < 0.05$), with means (calculated from all traits) of 0.155, 0.417 and 0.684 for common sage, hybrids and Greek sage, respectively (Supplementary Table S1). The number of monomorphic traits was the highest in common sage (17), the lowest in Greek sage (one) and intermediate in the group of hybrids (nine). Fourteen of the 23 qualitative traits were the most variable in Greek sage and six in the group of hybrids, whereas only two traits were the most variable in common sage. A single trait (Trichome type: Eglanular, Subadpressed) that was monomorphic for all analysed individuals, was excluded from further analysis. For common sage, 14 out of 15 individuals had the same specific trichome pattern, whereas only two pairs out of 34 individuals had the same pattern for Greek sage. Each of the remaining 32 individuals exhibited a unique trichome pattern. Among 26 hybrid individuals, 11 specific trichome patterns were observed, eight of which could not be found in either parental species. For common sage, only three different trichome types were identified; for the other two groups, six different types were identified (Supplementary Table S2). Analysis of variance (ANOVA) indicated significant differences in all analysed quantitative traits. The mean values of the most traits in the hybrid specimens were intermediate between the mean values of the two parental taxa (Supplementary Table S3). The correlation between Gower's distance matrices based on qualitative and quantitative traits was highly significant ($r = 0.41$; $P < 0.0001$).

The mean intraspecies Gower's distance was the highest in Greek sage and lowest in common sage, while an intermediate value was detected in the hybrids (0.274, 0.097 and 0.239, respectively). All pairwise differences in mean intraspecies Gower's distances were significant at $P < 0.05$.

A neighbour-net diagram based on Gower's distance matrices of combined qualitative and quantitative traits unambiguously separated the parental taxa from each other and from the hybrids (Fig. 1). The hybrids were positioned between the parental taxa.

SSR diversity. All seven microsatellite loci were successfully amplified in all samples. For each individual, analysis revealed either one (i.e. homozygous state) or two (i.e. heterozygous state) allele per locus. Overall, 64 alleles were detected, out of which 48 were present in common sage, 24 in Greek sage and 52 in hybrids, with a considerable number of alleles shared between taxa (Supplementary Fig. S1). Private alleles were detected in common sage (10) and in the hybrids (five), but not in the Greek sage. Consequently, common sage had the highest level of private allelic richness, the hybrids had an intermediate level, and Greek sage had the lowest level (2.00, 0.52 and 0.13, respectively).

Further comparison across the studied taxa indicated significant differences in patterns of genetic variation. The levels of observed (H_O) and expected heterozygosity (H_E) were significantly higher in common sage (0.610 and 0.749, respectively) and in the hybrids (0.778 and 0.678, respectively) than those in Greek sage (0.289 and 0.333, respectively). Significant and positive F_{IS} values were detected in both parental species (0.187 and 0.132 in common sage and Greek sage, respectively), whereas hybrids exhibited a negative but not significant F_{IS} value of -0.147 (Supplementary Table S4).

The neighbour-net analysis (Fig. 1) clearly separated the parental taxa from the hybrids. The highest levels of within-taxon genetic distances were observed in common sage, followed by the hybrids and then by Greek sage (0.684, 0.522 and 0.289, respectively). Expectedly, the highest average genetic distance was detected between the parental species (0.876) (Supplementary Table S5).

Although the AMOVA indicated that 77% of the genetic variation was observed within the groups (Supplementary Table S6), the highly significant ϕ_{ST} values confirmed the genetic differentiation among the taxa (Supplementary Table S5).

As expected, the highest pairwise ϕ_{ST} was noted between the parental species (0.431), while the value was lowest between the hybrids and Greek sage (0.115).

The results from the Bayesian analysis implemented in STRUCTURE strongly supported the presence of two clusters representing the parental species (Fig. 2). In individuals *a priori* identified as hybrids, wide range of admixture proportions were observed. The average estimates of the likelihood of the data, conditional on a given number of clusters, $\ln[\text{Pr}(X|K)]$, increased up to $K = 3$, with a standard deviation substantially higher at $K = 3$ compared to that at $K = 2$. The highest ΔK was obtained at $K = 2$ (177.54), and the second highest ΔK was obtained at $K = 3$ (4.41).

The assignment of hybrids using NEWHYBRIDS was congruent with the results of STRUCTURE; all individuals characterized by admixed proportions were confirmed to be of a hybrid origin (Fig. 2). The majority of the hybrid individuals (22 out of 26) were classified as F_1 hybrids with posterior probabilities (PP) ranging from 0.608 to 0.990. A single hybrid individual was classified as an F_2 (PP = 0.715), and three individuals were classified as back-crosses to Greek sage (BC_2) (PP = 0.510, 0.544 and 0.745). All common sage and Greek sage samples were classified as pure parental (P_1 and P_2 , respectively) with posterior probabilities ranging from 0.994 to 0.999 and

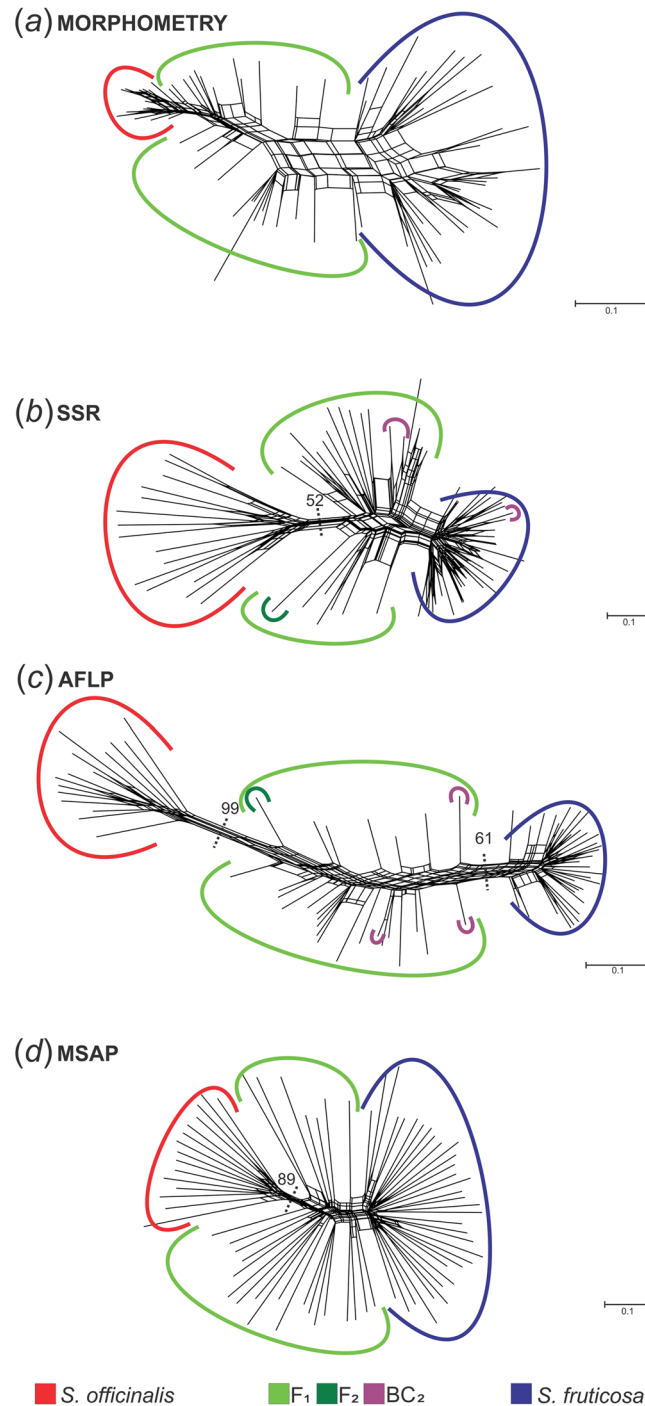


Figure 1. Neighbour-net diagrams showing relationships among studied individuals based on results from different analysis methods. Diagrams are based on (a) Gower's distances of 22 qualitative and 16 quantitative morphological characters for a morphometry analysis, (b) a proportion-of-shared-alleles distance matrix for SSR analysis, (c) a Dice distance matrix for AFLP analysis and (d) a Dice distance matrix for MSAP analysis. In both neighbour-net diagrams based on SSR and AFLP results, affiliation of individuals to different classes was based on the NewHybrids results from a microsatellite dataset. Bootstrap support values were derived from Neighbour Joining analysis, and those greater than 50% are indicated.

0.906 to 0.998, respectively. Assignment of individuals into classes (parental: *S. officinalis* and *S. fruticosa*, hybrid: *F*₁, *F*₂ and *BC*₂) carried out using NEWHYBRIDS was also indicated in the neighbour-net diagrams (Fig. 1).

AFLP diversity. The estimated error rate per primer combination in the AFLP analysis ranged from 0.68% to 2.9%. The mean error rate was 1.94%, which is similar to the error rate found in other AFLP studies⁵¹. The highest proportion of detected polymorphic loci was observed in the group of hybrids, followed by the Greek sage and

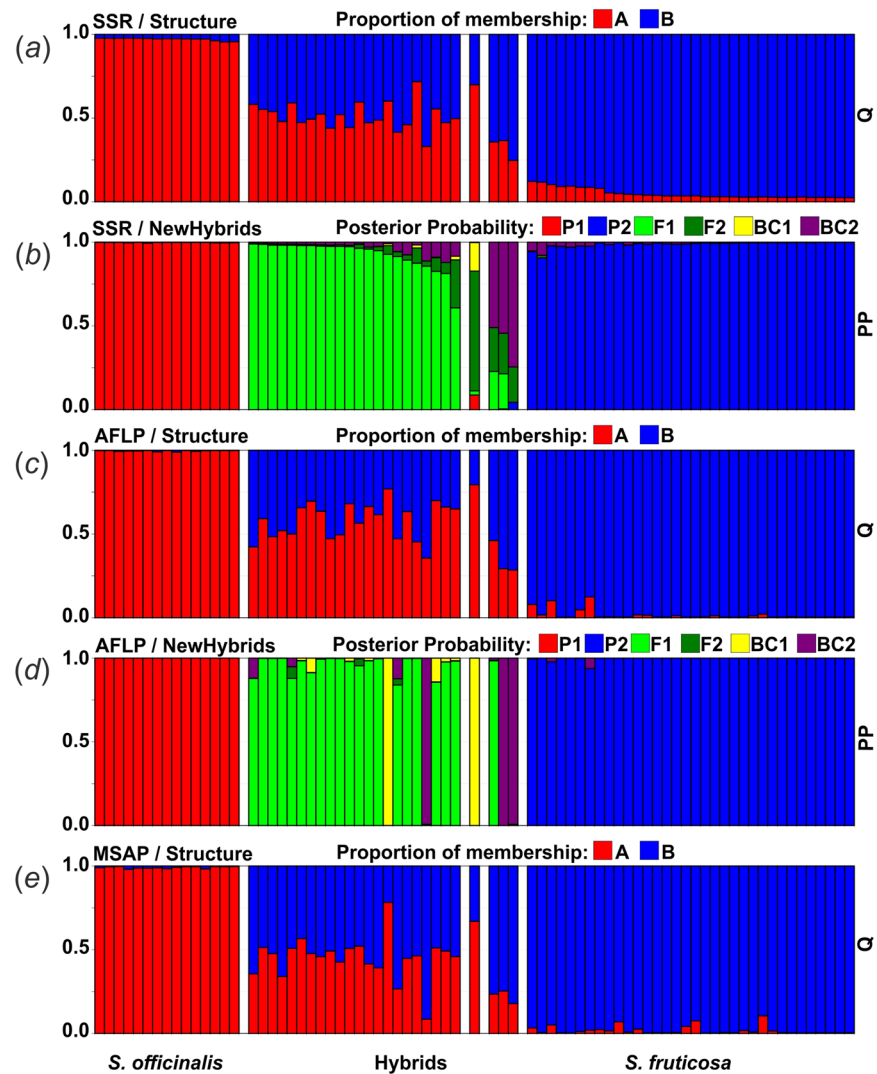


Figure 2. Genetic structure and assignment of individuals into classes as assessed by the computer programs STRUCTURE and NEWHYBRIDS. Each individual plant is represented by a single vertical line. For STRUCTURE (a,c,e), each colour (red and blue) represents a cluster (*S. officinalis* and *S. fruticosa*, respectively), and the length of the coloured segments indicates the individual's estimated proportions of membership in those clusters. For NEWHYBRIDS (b,d), each colour (red, blue, light green, dark green, yellow and violet) represents a genealogical class (purebred *S. officinalis*, purebred *S. fruticosa*, F₁ generation, F₂ generation, back-cross to *S. officinalis* and back-cross to *S. fruticosa*, respectively). The length of the coloured segments indicates the individual's estimated posterior probability of assignment to each class.

common sage (0.739, 0.582 and 0.455, respectively). According to Shannon's information index, a significantly ($P < 0.05$) higher level of diversity was detected in the group of hybrids as compared to that of the parental groups, while no significant difference in levels of genetic diversity was observed between the parental groups (Supplementary Table S7).

Average Dice's genetic distance at the taxon level was significantly higher ($P < 0.05$) in both common sage and the hybrids than that in the Greek sage (0.297, 0.281 and 0.178, respectively), while at the inter-taxa level the greatest genetic distance was expectedly observed between the parental taxa (0.836), while the least was between the Greek sage and the hybrids (0.421) (Supplementary Table S5). Neighbour-net analysis supported the obtained results (Fig. 1). Analysis of molecular variance showed that the majority (57%) of total genetic variability can be attributed to variability detected among groups (Supplementary Table S6), while highly significant pairwise ϕ_{ST} values between the involved taxa supported the existence of three differentiated groups of individuals (Supplementary Table S5). Similar to the results of microsatellite analysis, the highest pairwise ϕ_{ST} value was detected between the parental taxa (0.733).

We have chosen a NEWHYBRIDS result based on SSR analysis as a referent result in terms of assigning individuals to different classes (i.e., parental species, F₁ generation, F₂ generation and back-crosses to either of the parental species) in the AFLP and MSAP analyses. The results from STRUCTURE and NEWHYBRIDS based on the AFLP dataset strongly supported the results obtained by microsatellite analysis (Fig. 2).

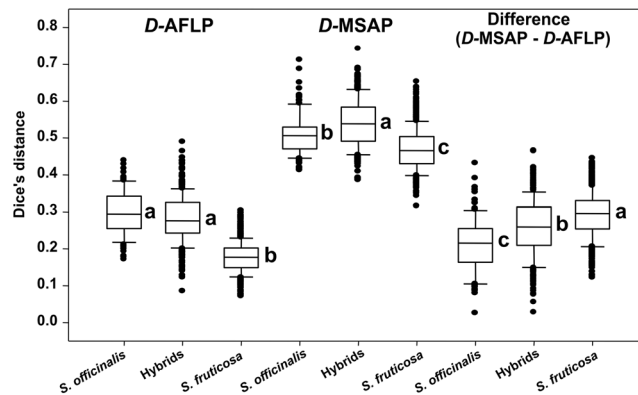


Figure 3. Box plots of intraspecies distances based on AFLP and MSAP data by species.

The average estimates of the likelihood of the data, conditional on a given number of clusters, $\ln[\Pr(X|K)]$, increased up to $K = 3$, with the standard deviation substantially higher at $K = 3$ compared to $K = 2$. The highest ΔK was obtained at $K = 2$ (3204.10), and the second highest ΔK was obtained at $K = 3$ (609.19).

In the STRUCTURE results, two parental clusters were well recognized, while in the group of hybrids, a wide range of admixture proportions can be observed. The results obtained by NEWHYBRIDS were for the most part congruent with the NEWHYBRIDS results for the microsatellites, with the exception of three individuals within the group of hybrids. Two individuals recognized as F_1 hybrids in the SSRs result were now recognized as a back-cross to common sage (BC_1) and a back-cross to Greek sage (BC_2). The individual classified as an F_2 hybrid in the SSRs result was now recognized as a back-cross to common sage (BC_1).

MSAP diversity. Analysis of MSAP data revealed that the highest proportion of polymorphic loci could be observed in the group of hybrids, followed by Greek sage and common sage (0.731, 0.715 and 0.545 respectively), as was also the case with the AFLP results (Supplementary Table S7). Additionally, Shannon's information index suggested the same ratio of diversity levels among studied groups as that seen in the AFLP results: parental groups were characterized by significantly lower levels of epigenetic diversity (0.325 for common sage and 0.328 for Greek sage) than was the group of hybrids (0.367), while there was no significant difference between the diversity levels of the parental groups (Supplementary Table S7). Dice distance at an intra-taxa level was significantly lower ($P < 0.05$) in Greek sage (0.470) than in common sage (0.508), while in the group of hybrids (0.540) it was significantly higher than that seen in both the parental species (Supplementary Table S5). However, when comparing values of Dice distances from the MSAP and AFLP results for each individual group, the detected difference was significantly higher ($P < 0.05$) in the Greek sage (0.292) than it was in the group of hybrids (0.258), while the latter was significantly higher than that of common sage (0.212) (Fig. 3).

The results for Dice distance between researched groups in the MSAP analysis were fully congruent with genetic distance results from the AFLP data analysis, as the greatest distance was detected between the parental taxa (0.717) and the least distance between Greek sage and the hybrids (0.565) (Supplementary Table S5). These results were also well supported by neighbour-net analysis (Fig. 1).

In contrast to the AFLP results, AMOVA analysis based on the MSAP dataset revealed that the majority (81%) of the total genetic variability was attributed to variability detected within groups, rather than among them (Supplementary Table S6). Pairwise ϕ_{ST} values between the involved taxa were again highly significant but of smaller amplitude than those in the AFLP results. (Supplementary Table S5).

The average estimates of the likelihood of the data, conditional on a given number of clusters, $\ln[\Pr(X|K)]$, increased up to $K = 3$, with a standard deviation substantially higher at $K = 3$ compared to $K = 2$. The highest ΔK was obtained at $K = 2$ (1778.52), and the second highest ΔK was obtained at $K = 3$ (664.64).

Results obtained by STRUCTURE for the MSAP dataset were consistent with STRUCTURE results for microsatellite and AFLP analysis as the parental species were well recognized. All individuals belonging to different generations of hybrids as defined by NEWHYBRIDS analysis based on microsatellites were again characterized by a range of admixture proportions (Fig. 2).

Environmental-niche modelling. Based on the multicollinearity test, 12 bioclimatic variables for *S. officinalis* (i.e. BIO1, BIO2, BIO5, BIO6, BIO7, BIO11, BIO12, BIO15, BIO16, BIO17, BIO18 and BIO19) and 13 for *S. fruticosa* (i.e. BIO1, BIO2, BIO5, BIO6, BIO7, BIO10, BIO11, BIO13, BIO14, BIO16, BIO17, BIO18 and BIO19) were excluded from further analysis. For both species, models with the smallest AICc values were selected. All bioclimatic variables and relative contributions to the Maxent model of the selected variables are listed in Supplementary Table S8. Yielded models are well congruent with the common sage distribution area in the western parts of the Balkan Peninsula and the Greek sage distribution area in the central and eastern Mediterranean (Fig. 4). Moran's I correlograms revealed a significant autocorrelation of ENM residuals for distances smaller than 150 km (for *S. officinalis*) or 250 km (for *S. fruticosa*). However, since both species are characterized by relatively small distribution areas that extends in narrow coastal areas, such result was inevitable, because reduction of number of occurrences to avoid ENM residual autocorrelation would lead to insufficiently large data set for this analysis. Besides, this result have not any influence on comparison of two ENMs which illustrates strong spatial

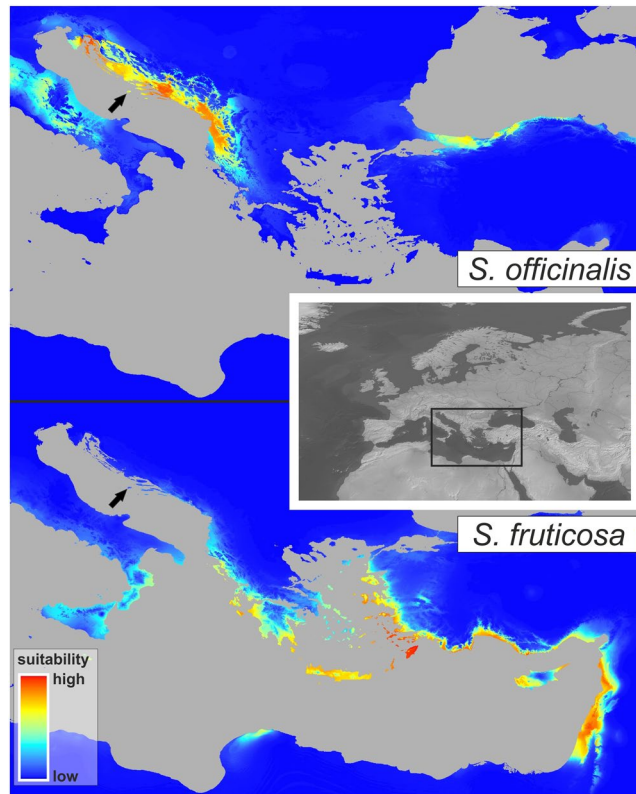


Figure 4. Environmental suitability map for *S. officinalis* and *S. fruticosa* based on ecological-niche modelling. The satellite imagery was obtained from Natural Earth public domain map dataset (<https://www.naturalearthdata.com/downloads/10m-raster-data/10m-gray-earth/>). Environmental suitability maps were generated from the climate layers obtained from <http://www.worldclim.org> and modified using ArcGIS ver. 10.1. (Esri, Redlands, CA, USA). The final figure was constructed using CorelDraw Graphics Suite X7 (Corel Corp., Ottawa, Canada).

segregation of *S. officinalis* and *S. fruticosa* environmental niches and positioning of hybrid population examined in this paper 1) far away from the environmental niche of *S. fruticosa* and 2) within the environmental niche of *S. officinalis*.

Discussion

Morphological aspects of inter-species hybridization. Analyses of morphological variation resulted in a clear delimitation of studied groups. Common sage individuals were characterized by nearly complete morphological uniformity, while in Greek sage, significantly higher levels of phenotypic variation were found (Fig. 1, Supplementary Table S3). Such contrasting results could likely be explained by the fact that the population of common sage is native to the Island of Vis and is positioned very close to the central part of the species' distribution area. Additionally, the Island of Vis was a part of a larger refugium area where the species likely survived Quaternary climatic oscillations^{52,53}. The long persistence of the species in this region led to its adaptation to local ecological conditions and resulted in it reaching a phenotypic optimum. On the other hand, the population of Greek sage is far away from the Eastern Mediterranean where it grows abundantly²⁶. According to MaxEnt analysis (Fig. 4), the studied Greek sage population is faced with marginal environmental conditions that likely caused the high levels of phenotypic variation known as phenotypic plasticity^{7,54,55}. It is known that when a population is faced with altered environmental conditions, a full spectrum of phenotypes is produced⁵⁶. Over time, only those phenotypes with the highest fitness and best-fitted traits will be chosen by natural selection⁵⁷.

As expected, hybrid plants were mostly characterized by admixed and/or intermediate morphological traits, as well as by unique traits that could not be identified in either of the parental species were found i.e., bract length (BL) and calyx lobe length (CLMI) were significantly longer in the hybrids than in either of the parental taxa (Supplementary Table S3). The occurrence of hybrid-specific characters was also confirmed by other studies⁵⁸. The emergence of such novel and unpredictable traits may be related to numerous mechanisms, such as the combination and complementary action of parental alleles⁵⁸, epistasis⁵⁹, phenotypic plasticity⁵⁸ and combinations of these phenomena.

Multilocus genetic characterization of hybridization. Analysis of seven microsatellite loci revealed that the common sage group was characterized by significantly higher levels of genetic diversity than those of the Greek sage group (Supplementary Table S4). Such a result might be explained by specific evolutionary histories

and past demographic fluctuations of these populations. Although today the Island of Vis is located 45 km from the mainland, during the Last Glacial Maximum (ca. 20,000 years ago), when sea level was 120 metres lower than it is today^{60,61}, the island was part of the mainland. It seems that the last several thousand years of isolation was not enough for genetic depauperation of the large common sage population on the Island of Vis. In high contrast to the common sage results, the obtained microsatellite results for Greek sage strongly support assumption of its non-native origin in this area⁶². If we assume that the population originated from very few individuals (i.e., the founder effect) ca. 2500 years ago when Ancient Greeks started their colony on the island^{31–33} and consider that even today the population is comprised of a limited number of individuals, the detected levels of genetic variability are not surprising.

In the group of hybrids, the highest levels of genetic diversity and allelic richness were observed (Supplementary Table S4). Common sage is characterized by greater genetic variability and participates in a hybridization event with more alleles than that of Greek sage, resulting in a high similarity of allelic richness levels between the hybrids and common sage. As already mentioned, only three out of the 26 hybrids were identified as backcrosses to Greek sage. Although such a result implicates that the gene flow is oriented towards the Greek sage as a sink population, only three detected BC₂ individuals make such a conclusion speculative. However, such orientation of the gene flow would be in accordance with a neutral model which suggests that in a hybridization between the native and non-native species, the gene flow is expected to be biased towards the non-native species⁶³. In addition, the model proposes that the asymmetry of the inter-species gene flow arises because of “a demographic imbalance between the two species at the wave front, where the invading species is at lower densities⁶³”; thus supporting our assumptions regarding the non-native origin of the Greek sage in the studied location. Similar patterns of introgression were also detected by some other researches^{64,65}. Based on obtained results it is possible to discuss the nature of homoploid hybridization in accordance to Schumer *et al.*⁵. The case study here presented obviously meet only one criterion without any doubts, as the genome level hybridization was confirmed. The criterion implying the necessity of reproduction isolation of hybrids from parental species was not met, although such a small number of back-crosses leaves some doubts. Finally, even if the back-crosses were not detected, based on implemented research strategies it would be impossible to address the issue of the origin of the reproduction isolation.

Epigenetic analysis of the hybridization event: AFLP vs. MSAP. For an additional reliable assessment of genome-wide genetic and epigenetic variability in the hybridization event, technically very similar AFLP and MSAP approaches were used. In AFLP analysis there was no significant difference in Shannon's information indexes between the parental taxa, which was not the case with results obtained by microsatellites. As already suggested, it is possible that such results were obtained as a consequence of cross-amplification of microsatellites in the target species (i.e., Greek sage). Based on both the AFLP and MSAP results, three groups of individuals clearly differentiated from each other. The results from STRUCTURE for both AFLP and MSAP and from NEWHYBRIDS for AFLP (Fig. 2) are fully congruent with NEWHYBRIDS results for SSRs. The only differences were in the positioning of the three individuals when comparing the SSR and AFLP results and one individual when comparing the SSR and MSAP results. Because these individuals were all recognized as hybrids but of different generation, such small differences in results are likely not of major significance and can be attributed to the noise of estimation.

In both the AFLP and MSAP results, the proportions of polymorphic alleles and Shannon's information index have significantly ($P < 0.05$) higher values for the group of hybrids than those of the parental species, and such results are not surprising, as similar results were observed in other research⁶⁶. When comparing differences between genetic and epigenetic distances (Fig. 3) within individual taxa, contrasting results occurred. The smallest difference was observed for common sage, intermediate for the hybrids and most pronounced for Greek sage. Such a result supports the assumption that the common sage population is well adapted to local environmental conditions and is consequently exposed to low levels of stress. At the same time, a very pronounced difference between the levels of genetic and epigenetic diversity in the Greek sage group most likely occurs because of the stressful environmental conditions encountered outside of the species' optimal ecological niche (Fig. 4). Such results clearly confirm the strong influence of unfavourable environmental conditions on epigenetic excitation in the plants^{9,67,68}.

Being characterized by intermediate values of difference between genetic and epigenetic diversity of parental species, the results for the group of hybrids support our knowledge of heritability of epigenetic traits^{11,69–71}. The obtained results are well supported and visualized by the Neighbour-Net diagram (Fig. 1).

Analysis of molecular variance of the partitioning of AFLP diversity between and within studied groups revealed that 57.2% of the overall genetic diversity can be attributed to diversity among groups. Bearing in mind that the analysed hybridization event was at an inter-species level, such a result is expected. However, for epigenetic diversity, a highly contrasting result was obtained, as only 18.7% of the overall genetic diversity can be attributed to diversity among groups. Such results strongly support the idea that the majority of epigenetic changes are specific for each individual and in addition, that the epigenetic variability is not tightly dependent on an individual's genetic diversity and structure^{66,72,73}.

Conclusions

In our research, different morphometric and molecular methods were used to analyse hybridization between two closely related yet mostly allopatric species. By doing this, we gained information on the different aspects of inheritance of phenotypic, genetic and epigenetic traits in contrast to different levels of adaptation of the parental populations to local environmental conditions. The obtained results revealed a high contrast between the levels of genetic and phenotypic variability in the parental taxa, as common sage was characterized by high levels of genetic and extremely low levels of phenotypic variability. For Greek sage, morphometric analysis revealed exceptionally high levels of variability that can likely be attributed to phenotypic plasticity, while results for genetic

variability obtained using two methods were not entirely congruent with each other. Such results for a parental species can likely be explained by their different demographic histories in the studied location. A well defined group of hybrids was detected at morphological, genetic and epigenetic levels and were mostly comprised of F_1 generation individuals, suggesting that the hybrids were predominantly sterile. For most analysed morphological characteristics, hybrids were placed in-between the parental taxa, while at the same time, their values of different genetic parameters were mostly higher than those of the parental species. In addition, epigenetic analysis of the involved taxa improved our understanding of the importance of epigenetic variation in natural populations in contrast to altered environmental conditions and confirmed the trans-generational nature of epigenetic changes. Although the obtained results provided substantial insight into the hybridization between closely related native and non-native plant species, additional knowledge would be gained by more comprehensive sampling of higher resolution accompanied by geocoding of each sample and genetic analysis of chloroplast DNA. By performing the spatial analysis of both inter-species gene flow and fertilization direction, further unravelling of the complexity of the hybridization event would be achieved.

Methods

Sampling of plant material. Leaf tissue and branches with inflorescences from 75 *Salvia* individuals were collected in the western part of the Island of Vis (Croatia) in late April. The morphological characteristic used in the field for identification of parental species and hybrids was the calyx shape (Supplementary Fig. S2).

The common sage calyx is larger and strictly bilabiate, whereas the calyx of Greek sage is smaller, tubular and actinomorphic^{17,26,36}. Individuals characterized by a slightly bilabiate calyx were considered to be of hybrid origin. In total, 15 samples were identified as common sage, 34 as Greek sage and 26 as their putative hybrid.

In defining the sampling area, the main criteria were that all samples must be (1) within a pollinator's range and (2) strictly from the area of sympatry. Because the expected foraging distances for honey bees is greater than 1 km^{74–76}, up to few hundred metres for solitary bees^{77,78} and at least 250 m for bumble bees^{79,80}, all sampled specimens were obtained within a single pollinator's range (i.e., the greatest distance between two sampled individuals was no greater than 1 km). To meet the second criterion, three groups (i.e., parental species and hybrids) were sampled in unequal proportions and in limited numbers. Since studied species do not naturally reproduce clonally and are characterized by distinctive bush-like habit, there was no risk of sampling genetically identical individuals. All of the samples were collected from neglected and abandoned vineyards and olive groves surrounding a local settlement, on a site of S-SW exposition, at the altitude ranging from 30 to 110 m. The majority of observed individuals was grouped in a large stand and a typical distance between sampled individuals ranged from 5 to 10 meters. The sampling was performed in a way that the entire stand was equally sampled.

Laboratory methods. Based on a previous taxonomic revision of the section *Salvia*³⁶, 23 qualitative and 16 quantitative traits were selected and measured from digitized photographs using a Dino-Lite pro Digital Microscope and DinoCapture 2.0 software (AnMo Electronics Corp, New Taipei City, Taiwan).

Total genomic DNA was extracted from silica-gel-dried leaf tissue using a GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, St. Louis, Missouri, USA). The DNA concentrations were measured using a P330 Nanophotometer (IMPLEN, Munich, Germany).

For fine-scale analysis of the population genetic structure of the sampled groups, seven microsatellite loci (SoUZ003, SoUZ006, SoUZ007, SoUZ009, SoUZ013, SoUZ014 and SoUZ020) that were previously characterized in common sage and successfully cross-amplified in Greek sage^{81–83} were amplified. Details on the polymerase chain reaction (PCR) conditions and cycle profiles may be found in Radosavljević *et al.*⁶². Polymerase chain reaction products were analysed by capillary electrophoresis on a ABI 3730xL DNA analyser (Applied Biosystems, Foster City, CA, USA) provided by Macrogen DNA service (Seoul, Korea), and alleles were scored using GeneMapper 4.0 (Applied Biosystems).

For a genome-wide assessment of the genetic diversity, the AFLP technique was used based on the protocol of Carović-Stanko *et al.*⁸⁴ with some modifications. For selective amplification, four primer combinations were used as follows: VIC-EcoRI-ACG + TruII-CGA, NED-EcoRI-AGA + TruII-CAC, FAM-EcoRI-ACA + TruII-CAC and PET-EcoRI-ACC + TruII-CGA. Amplified PCR products were analysed using an ABI 3730xL analyser (Applied Biosystems), and the GeneMapper 4.0 software (Applied Biosystems) was used for final allele scoring. Reactions of genomic DNA restriction, adapter ligation, and preselective and selective amplification were performed using a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA). The selection of markers for further statistical analysis and the estimation of the error rate per primer combination were performed as suggested by Herrmann *et al.*⁸⁵.

To assess the level of epigenetic variability of the studied groups, MSAP analysis was applied in accordance with Nuskern *et al.*⁸⁶. Analysis and scoring of the amplified PCR products were the same as in the AFLP analysis.

Morphometric data analysis. Shannon's information index as a measure of qualitative trait diversity was calculated for each trait according to Lewontin⁸⁷ (1972) as well as separately for each species. Repeated measure analysis of variance was performed using PROC GLM in SAS⁸⁸ with species membership as categorical independent variable and Shannon's information index of each qualitative trait as response variable. *Post hoc* Bonferroni's adjustments were used to compare the mean values of Shannon's information index between species at a significance level of $P < 0.05$. Univariate analysis of variance was performed using PROC GLM in SAS⁸⁸ to test for differences in mean values between species (categorical independent variable) for 19 quantitative morphological traits (response variables). *Post hoc* comparisons of the means were performed using Tukey's Studentized Range test at $P < 0.05$.

To combine the information from the qualitative and quantitative traits into a single distance measure, Gower's distance⁸⁹ was calculated between all pairs of individuals with the algorithm implemented in PAST ver.

2.01⁹⁰. Pairwise Gower's distance matrices among individuals were calculated separately based on qualitative and quantitative data and compared by calculating the correlation coefficients and performing Mantel's test⁹¹. The randomization procedure as implemented in NTSYS-pc⁹² included 1,000 permutations. To test the significance of the differences in mean intraspecies Gower's distances among the species, Kruskal-Wallis (among all species) and Wilcoxon rank-sum (between all pairs of species) non-parametric tests were performed using SAS Release 9.2⁸⁸. A neighbour-net diagram, which is well suited to depict reticulate relationships⁹³, was constructed from Gower's distances using SplitsTree 4⁹⁴.

SSR data analysis. GENEPop 4.0⁹⁵ was used to estimate population genetic parameters (the average number of alleles per locus, N_{ar} ; the observed heterozygosity, H_O ; the expected heterozygosity, H_E ; and the inbreeding coefficient, F_{IS}) and to test genotypic frequencies across all loci for conformance to Hardy-Weinberg (HW) expectations. As a measure of the number of alleles per locus is independent of sample size, allelic richness (N_{ar}) was calculated using the FSTAT v.2.9.3.2 package^{96,97}, and the number of private alleles (N_{pr}) per species was assessed using MICROSAT⁹⁸. The estimates of N_{ar} , H_O and H_E in each species were compared by repeated measure analysis of variance followed by *post hoc* Bonferroni's adjustments using PROC GLM in SAS⁸⁸. The species membership was used as categorical independent variable and the estimates of N_{ar} , H_O and H_E as response variables. The private allelic richness (N_{par}) for each species was calculated by averaging the number of private alleles across possible subsets of sampled chromosomes while adjusting for differences in sample size across species. The calculations were performed using the rarefaction procedure as implemented in ADZE 1.0⁹⁹.

Genetic distances between pairs of samples were calculated using the proportion-of-shared-alleles distances (D_{psa})¹⁰⁰ as implemented in MICROSAT⁹⁸. A neighbour-net diagram was produced as described previously. Bootstrap support was obtained using 1,000 replicates generated by MICROSAT and subsequently used in the cluster analysis performed by the neighbor-joining method as implemented in the NEIGHBOR and CONSENSE programs in the PHYLIP ver. 3.6b software package¹⁰¹. The significance of the differences in mean intraspecies distances among species was assessed as previously described.

Analysis of molecular variance¹⁰² (AMOVA) was used to partition the total genetic variance into within and among species in Arlequin ver. 2.000¹⁰³. Pairwise comparisons examined with analysis of molecular variance resulted in values of ϕ_{ST} that are equivalent to the proportion of the total variance that is partitioned between the two species.

A model-based clustering method was applied to infer the genetic structure and define the number of clusters using STRUCTURE ver. 2.3.3¹⁰⁴. Given a value for the number of clusters, this method assigns individual genotypes from the entire sample to clusters so that linkage disequilibrium (LD) was maximally explained. Thirty runs for each cluster (K) ranging from one to 11 were performed on the Isabella computer cluster at the University of Zagreb (Croatia), University Computing Centre (SRCE). Each run consisted of a burn-in period of 200,000 steps, followed by 10⁶ 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 performed by comparing the average estimates of the likelihood of the data, $\ln[Pr(X|K)]$, for each value of K , as well as by calculating an ad hoc statistic ΔK based on the rate of change in the log probability of data between successive K values, as described by Evanno *et al.*¹⁰⁵ and implemented in STRUCTURE HARVESTER v0.6.92¹⁰⁶. Runs were clustered and averaged using CLUMPAK¹⁰⁷.

The Bayesian method implemented by NewHybrids 1.1¹⁰⁸ was used to assign individuals to one of six classes: two pure (parental *S. officinalis* and *S. fruticosa*) and four putative (F_1 , F_2 , and backcrosses with the parental populations) hybrids. The program ran without any prior information regarding the putative hybrid status of collected individuals and with the uninformative Joffrey's prior option for both mixing proportions and allele frequencies. The results were based on the average of five independent runs consisting of a burn-in phase of 100,000 steps and 500,000 MCMC sweeps. Following the suggestions of Anderson and Thompson¹⁰⁸, individual genetic assignment to classes was based on a minimum posterior probability threshold (Tq) of 0.50.

AFLP data analysis. AFLP-amplified fragments were scored for the presence (1) or absence (0) of homologous bands to create binary matrices. The AFLP diversity within species was estimated by determining the percentage of polymorphic bands (%P) and Shannon's information index. The mean values of Shannon's information index among species were tested as previously described.

Pairwise genetic distances were calculated using Dice's coefficient¹⁰⁹ (Dice, 1945), and a neighbour-net diagram was constructed as previously described. Bootstrap support was obtained using 1,000 samples in the cluster analysis performed by the neighbour-joining method as implemented in TREECON for Windows version 1.3 b¹¹⁰. The significance of the differences in mean intraspecies distances among species was assessed as previously described. Analysis of molecular variance and the analyses in Structure and NewHybrids were performed using the same parameters as previously described.

MSAP data analysis. In the MSAP analyses, after the *EcoRI/HpaII* and *EcoRI/MspI* reactions, four distinct profiles were obtained as follows: (I) fragments present in both profiles (1/1), unmethylated state; (II) fragments present only in the *EcoRI/MspI* profiles (0/1), hemi- or full-methylation of internal cytosine (^{HMe}CG or ^{Me}CG-sites); (III) fragments present only in *EcoRI/HpaII* profiles (1/0), hemi-methylation of external cytosine (^{HMe}CCG-sites); and (IV) fragments absent in both profiles (0/0), uninformative state. A 'Mixed-Scoring 2' approach as implemented in R script MSAP_calc.r¹¹¹ was used to generate a matrix of binary data for each of the three classes of markers corresponding to three subepiloci: *u*-subepilocus (coding of unmethylated sites), *m*-subepilocus (coding of ^{HMe}CG or ^{Me}CG-sites), and *h*-subepilocus (coding of ^{HMe}CCG-sites). Further analyses were performed in the same manner as in the case of the AFLP data.

Environmental-niche modelling. To test the presumption that the central Adriatic region is well suited for survival of common sage, but not Greek sage, environmental-niche modelling for both species was independently performed using MAXENT ver. 3.3.3k¹¹². The analysis was based on 90 occurrence records for common sage and 70 occurrence records for Greek sage. From the WorldClim dataset, 19 bioclimatic variables for the present time and at a resolution of 30 arc-seconds were downloaded¹¹³. To test the multicollinearity among the bioclimatic variables for each species separately, “usdm” R package¹¹⁴ was used to calculate variance inflation factors (VIF) and only the bioclimatic variables with VIF < 10 were used for further analysis. For a selection of an appropriate value of Maxent’s regularization multipliers for obtained sets of bioclimatic variables, ENMTools was used¹¹⁵. After the environmental suitability model was constructed, ENM residuals (calculated for each occurrence record as 1 – Y, where Y was ENM result) were used to construct Moran’s I correlogram (i.e. to calculate the spatial auto-correlation factors (Moran’s I) for multiple distance classes) as implemented in SAM v4.0 software¹¹⁶. Obtained environmental suitability models were visualized in ArcGIS ver. 10.1. (Esri, Redlands, CA, USA).

Data Availability

The datasets used and/or analysed during the current study are available from the corresponding author upon request.

References

1. Strauss, S. Y., Lau, J. A. & Carroll, S. P. Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecol Lett.* **9**, 357–374 (2006).
2. Rieseberg, L. H. & Wendel, J. F. Introgression and its consequences in plants in *Hybrid zones and the evolutionary process* (ed. Harrison, R. G.) 70–109 (Oxford University Press, 1993).
3. Rhymer, J. M. & Simberloff, D. Extinction by hybridization and introgression. *Annu Rev Ecol Syst.* **27**, 83–109 (1996).
4. Nieto Feliner, G. *et al.* Is homoploid hybrid speciation that rare? An empiricist’s view. *Heredity.* **118**, 513–516 (2017).
5. Schumer, M., Rosenthal, G. G. & Andolfatto, P. How common is homoploid hybrid speciation? *Evolution.* **68**, 1553–1560 (2014).
6. West-Eberhard, M. J. Developmental plasticity and the origin of species differences. *PNAS.* **102**, 6543–6549 (2005).
7. Ghalambor, C. K., McKay, J. K., Carroll, S. P. & Reznick, D. N. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct Ecol.* **21**, 394–407 (2007).
8. Bossdorf, O., Richards, C. L. & Pigliucci, M. Epigenetics for ecologists. *Ecol Lett.* **11**, 106–115 (2008).
9. Zhang, Y.-Y., Fischer, M., Colot, V. & Bossdorf, O. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytol.* **197**, 314–322 (2013).
10. Simon, S. A. & Meyers, B. C. Small RNA-mediated epigenetic modifications in plants. *Curr Opin Plant Biol.* **14**, 148–155 (2011).
11. Jablonka, E. & Raz, G. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol.* **84**, 131–176 (2009).
12. Salmon, A., Ainouche, M. L. & Wendel, J. F. Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Mol Ecol.* **14**, 1163–1175 (2005).
13. Marfil, C. F., Masuelli, R. W., Davison, J. & Comai, L. Genomic instability in *Solanum tuberosum* x *Solanum kurtzianum* interspecific hybrids. *Genome* **49**, 104–113 (2006).
14. Kalisz, S. & Purugganan, M. D. Epialleles via DNA methylation: consequences for plant evolution. *Trends Ecol Evol.* **19**, 309–314 (2004).
15. Rapp, R. A. & Wendel, J. F. Epigenetics and plant evolution. *New Phytol.* **168**, 81–91 (2005).
16. Walker, J. B., Sytsma, K. J., 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.* **91**, 1115–1125 (2004).
17. Hedge, I. C. *Salvia* in *Flora Europaea*, Vol. 3. (eds Tutin, T. G., Heywood, V. H., Burges, D. M. M., Valentine, D. H., Walters, S. M. & Webb, D. A.) 188–192 (Cambridge University Press, 1972).
18. Putievsky, E., Ravid, U., Diwan-Rinzler, N. & Zohary, D. Genetic affinities and essential oil composition of *Salvia officinalis* L., *S. fruticosa* Mill., *S. tomentosa* Mill., and their hybrids. *Flavour Frag J.* **5**, 121–123 (1990).
19. Dudai, N. *et al.* Dynamics of yield components and essential oil production in a commercial hybrid sage. *J Agr Food Chem.* **47**, 4341–4345 (1999).
20. Baser, C. & Husnu, K. Production of *Salvia* oil in Mediterranean countries in *SAGE: The Genus Salvia* (ed. Kintzios, S. E.) 263–268 (Harwood Academic Publishers, 2000).
21. Hedge, I. C. *Salvia* L. In *Flora of Turkey and the East Aegean*, Vol 7. (ed. Davis, P. H.) 400–461 (Edinburgh University Press, 1982).
22. Pignatti, S. *Flora d’Italia*, Vol. 2. (Edagricole, 1982).
23. Lucchese, F., Persia, G. & Pignatti, S. I prati a *Bromus erectus* Hudson dell’Appennino Laziale. *Fitosociologia* **30**, 145–180 (1995).
24. Cutini, M., Cancellieri, L., Ceschin, S., Lucchese, F. & Caneva, G. Analisi cenologica e sintassonomia delle garighe a *Salvia officinalis* L. lucane nel quadro dei salvieti peninsulari (Basilicata, Appennino meridionale). *Webbia* **62**, 225–244 (2007).
25. Di Pietro, R. New dry grassland associations from Ausoni-Aurunci mountains (central Italy) - syntaxonomical updating and discussion on the higher rank syntaxa. *Hacquetia* **10**, 183–231 (2011).
26. Karousou, R., Hanlidou, E. & Kokkini, S. The sage plants of Greece: distribution and infraspecific variation in *SAGE: The Genus Salvia* (ed. Kintzios, S. E.) 27–46 (Harwood Academic Publishers, 2000).
27. Ruberto, G., Biondi, D., Meli, R. & Piattelli, M. Volatile flavour components of Sicilian *Origanum onites* L. *Flavour Frag J.* **8**, 197–200 (1993).
28. Rivera, D., Obon, C. & Cano, F. The botany, history and traditional uses of three-lobed sage (*Salvia fruticosa* Miller) (Labiatae). *Econ Bot.* **48**, 190–195 (1994).
29. Gali-Muhtasib, H., Hilan, C. & Khater, C. Traditional uses of *Salvia libanotica* (East Mediterranean sage) and the effects of its essential oils. *J Ethnopharmacol.* **71**, 513–520 (2000).
30. Flora Croatica Database. Faculty of Science, University of Zagreb, <https://hirc.botanic.hr/fcd/> (2016).
31. Gabričević, B. Antički spomenici otoka Visa (in Croatian). *Contributions to the History of Art in Dalmatia.* **17**, 5–60 (1968).
32. Zimi, E. Illirians in *Encyclopedia of Ancient Greece* (ed. Wilson, N.) 377 (Taylor & Francis group, 2006).
33. Kirigin, B. Issa in *The Encyclopedia of Ancient History* (eds Bangall, R.S., Brodersen, K., Champion, C. B., Erskine, A. & Huebner, S. R.) 3527 (Wiley-Blackwell, 2012).
34. Sánchez-Gómez, P. & Morales, R. A new hybrid of *Salvia* (Labiatae). *An Jardín Bot Madrid.* **57**, 421–422 (1999).
35. Herraiz-Peñalver, D., Ortiz De Elguea-Culebras, G. O., Sánchez-Vioque, R. & Santana Méridas, O. Identification of a hybrid species of sage (*Salvia officinalis* L. x *S. lavandulifolia* subsp. *lavandulifolia*) through the study of the essential oil. *J of Essent Oil Res.* **27**, 363–372 (2015).
36. Reales, A., Rivera, D., Palazon, J. A. & Obon, C. Numerical taxonomy study of *Salvia* sect. *Salvia* L. (Labiatae). *Bot J Linn Soc.* **145**, 353–371 (2004).

37. Schmiderer, C., Torres-Londoño, P. & Novak, J. Proof of geographical origin of Albanian sage by essential oil analysis. *Biochem Syst Ecol.* **51**, 70–77 (2013).
38. Corsi, G. & Bottega, S. Glandular hairs of *Salvia officinalis*: New data on morphology, localization and histochemistry in relation to function. *Ann Bot.* **84**, 657–664 (1999).
39. Neeman, G. & Dafni, A. Fire, bees and seed production in a Mediterranean key species *Salvia fruticosa* (Lamiaceae). *Isr J Plant Sci.* **47**, 157–163 (1999).
40. Schuler, G. D. *et al.* A gene map of the human genome. *Science.* **274**, 540–546 (1996).
41. Knapik, E. W. *et al.* A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nat Genet.* **18**, 338–343 (1998).
42. Jarne, P. & Lagoda, P. J. Microsatellites, from molecules to populations and back. *Trends Ecol Evol.* **11**, 424–429 (1996).
43. Vos, P. *et al.* AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**, 4407–4414 (1995).
44. Zane, L., Bargelloni, L. & Patarnello, T. Strategies for microsatellite isolation: a review. *Mol Ecol.* **11**, 1–16 (2002).
45. Campbell, D., Duchesne, P. & Bernatchez, L. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Mol Ecol.* **12**, 1979–1991 (2003).
46. Bernatchez, L. & Duchesne, P. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can J Fish Aquat Sci.* **57**, 1–12 (2000).
47. Meudt, H. M. & Clarke, A. C. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends Plant Sci.* **12**, 106–117 (2007).
48. Reyna-López, G. E., Simpson, J. & Ruiz-Herrera, J. Differences in DNA methylation pattern are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Mol Gen Genet.* **253**, 703–710 (1997).
49. Schrey, A. W. *et al.* Ecological epigenetics: beyond MS-AFLP. *Integr comp biol.* **53**, 340–350 (2013).
50. Alonso, C., Pérez, R., Bazaga, P., Medrano, M. & Herrera, C. M. MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species. *Mol Ecol Resour.* **16**, 80–90 (2016).
51. Bonin, A. *et al.* How to track and assess genotyping errors in population genetics studies. *Mol Ecol.* **13**, 3261–3273 (2004).
52. Medail, F. & Diadema, K. Glacial refugia influence plant diversity patterns in the Mediterranean basin. *J Biogeogr.* **36**, 1333–1345 (2009).
53. Rešetnik, I. *et al.* Genetic diversity and demographic history of wild and cultivated/naturalised plant populations: evidence from dalmatian sage (*Salvia officinalis* L., Lamiaceae). *PLoS ONE* **11**(7), e0159545 (2016).
54. Schlichting, C. D. The evolution of phenotypic plasticity in plants. *Annu Rev Ecol Syst.* **17**, 667–693 (1986).
55. Agrawal, A. A. Phenotypic plasticity in the interactions and evolution of species. *Science* **294**, 321–326 (2001).
56. Pigliucci, M. *Phenotypic Plasticity: Beyond Nature and Nurture.* (Johns Hopkins University Press, 2001).
57. Wund, M. A. Assessing the impacts of phenotypic plasticity on evolution. *Integr Comp Biol.* **52**, 5–15 (2012).
58. Rieseberg, L. H. & Ellstrand, N. C. What can molecular and morphological markers tell us about plant hybridization? *CRC Crit Rev Plant Sci.* **12**, 213–241 (1993).
59. Orr, H. A. The genetics of species differences. *Trends Ecol Evol.* **16**, 343–350 (2001).
60. Rohling, E. J. *et al.* Magnitudes of sea-level lowstands of the past 500,000 years. *Nature* **394**, 162–165 (1998).
61. Clark, P. U. *et al.* The last glacial maximum. *Science* **325**, 710–714 (2009).
62. Radosavljević, I., Satovic, Z. & Liber, Z. Causes and consequences of contrasting genetic structure in sympatrically growing and closely related species. *AoB PLANTS.* **7**, plv106 (2015).
63. Currat, M., Ruedi, M., Petit, R. J. & Excoffier, L. The hidden side of invasions: massive introgression by local genes. *Evolution.* **62**, 1908–1920 (2008).
64. Rieseberg, L. H. *et al.* Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica.* **129**, 149–165 (2007).
65. Du, F. K. *et al.* Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. *New Phytol.* **192**, 1024–1033 (2011).
66. Herrera, C. M. & Bazaga, P. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytol.* **187**, 867–876 (2010).
67. Chinnusamy, V. & Zhu, J. K. Epigenetic regulation of stress responses in plants. *Curr Opin Plant Biol.* **12**, 133–139 (2009).
68. Mirouze, M. & Paszkowski, J. Epigenetic contribution to stress adaptation in plants. *Curr Opin Plant Biol.* **14**, 267–274 (2011).
69. Verhoeven, K. J., Jansen, J. J., van Dijk, P. J. & Biere, A. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* **185**, 1108–1118 (2010).
70. Becker, C. & Weigel, D. Epigenetic variation: origin and transgenerational inheritance. *Curr Opin Plant Biol.* **15**, 562–567 (2012).
71. Iwasaki, M. & Paszkowski, J. Epigenetic memory in plants. *EMBO J.* **33**, 1987–1998 (2014).
72. Keyte, A. L., Percifield, R., Liu, B. & Wendel, J. F. Intraspecific DNA methylation polymorphism in cotton (*Gossypium hirsutum* L.). *J Hered.* **97**, 444–450 (2006).
73. Vaughn, M. W. *et al.* Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol.* **5**, e174 (2007).
74. Visscher, P. K. & Seeley, T. D. Foraging strategy of honeybee colonies in a temperate deciduous forest. *Ecology* **63**, 1790–1801 (1982).
75. Beekman, M. & Ratnieks, F. L. W. Long-range foraging by the honey-bee, *Apis mellifera* L. *Funct Ecol.* **14**, 490–496 (2000).
76. Steffan-Dewenter, I. & Kuhn, A. Honeybee foraging in differentially structured landscapes. *Proc R Soc Lond B Biol Sci.* **270**, 569–575 (2003).
77. Gathmann, A. & Tschardtke, T. Foraging ranges of solitary bees. *J Anim Ecol.* **71**, 757–764 (2002).
78. Zurbuchen, A. *et al.* Maximum foraging ranges in solitary bees: only few individuals have the capability to cover long foraging distances. *Biol Conserv.* **143**, 669–676 (2010).
79. Osborne, J. L. *et al.* A landscape-scale study of bumble bee foraging range and constancy, using harmonic radar. *J Appl Ecol.* **36**, 519–533 (1999).
80. Wolf, S. & Moritz, R. F. A. Foraging distance in *Bombus terrestris* L. (Hymenoptera: Apidae). *Apidologie* **39**, 419–427 (2008).
81. Radosavljević, I., Jakse, J., Javornik, B., Satovic, Z. & Liber, Z. New microsatellite markers for *Salvia officinalis* (Lamiaceae) and cross-amplification in closely related species. *Am J Bot.* **98**, e316–e318 (2011).
82. Molecular Ecology Resources Primer Development Consortium, An, J. *et al.* Permanent genetic resources added to molecular ecology resources database 1 October 2009 – 30 November 2009. *Mol Ecol Resour.* **10**, 404–408 (2010).
83. Radosavljević, I. *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.* **13**, 12082–12093 (2012).
84. Carović-Stanko, K. *et al.* Molecular and chemical characterization of the most widespread *Ocimum* species. *Plant Syst Evol.* **294**, 253–262 (2011).
85. Herrmann, D. *et al.* Selection criteria for scoring amplified fragment length polymorphisms (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome* **53**, 302–310 (2010).
86. Nuskern, L., Ježić, M., Liber, Z., Mlinarec, J. & urković-Perica, M. Cryphonectria hypovirus 1-induced epigenetic changes in infected phytopathogenic fungus *Cryphonectria parasitica*. *Microb Ecol.* **75**, 790–798 (2017).
87. Lewontin, R. C. Testing the theory of natural selection. *Nature* **236**, 181–182 (1972).
88. SAS Institute Inc. SAS/STAT 9.1 User's Guide (2004).
89. Gower, J. C. A general coefficient of similarity and some of its properties. *Biometrics* **27**, 857–874 (1971).
90. Hammer, Ø., Harper, D. A. T. & Ryan, P. D. Past: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol Electron.* **4**, 1–9 (2001).
91. Mantel, N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220 (1967).

92. Rohlf, F. J. NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.2. (2005).
93. Bryant, D. & Moulton, V. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol.* **21**, 255–265 (2004).
94. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* **23**, 254–267 (2006).
95. Raymond, M. & Rousset, F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered.* **86**, 248–249 (1995).
96. Goudet, J. FSTAT (Vers. 1.2): a computer program to calculate F-statistics. *J Hered.* **86**, 485–486 (1995).
97. Goudet, J. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2), <http://www2.unil.ch/popgen/softwares/fstat.htm> (2002).
98. Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M. & Cavalli-Sforza, L. L. Microsat: a computer program for calculating various statistics on microsatellite allele data (version 1.5d), <http://hplg.stanford.edu/projects/microsat> (1998).
99. Szpiech, Z. A., Jakobsson, M. & Rosenberg, N. A. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498–2504 (2008).
100. Bowcock, A. M. *et al.* High resolution human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455–457 (1994).
101. Felsenstein, J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5**, 164–166 (1989).
102. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction sites. *Genetics* **131**, 479–491 (1992).
103. Schneider, S., Roessli, D. & Excoffier, L. Arlequin version 2.001: A software for population genetics data analysis, <http://cmpg.unibe.ch/software/arlequin/software/2.001/doc/whatsnew/whatsnew.html> (2002).
104. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
105. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* **14**, 2611–2620 (2005).
106. Earl, D. A. & von Holdt, B. M. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conserv Genet Resour.* **4**, 359–361 (2012).
107. Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour.* **15**, 1179–1191 (2015).
108. Anderson, E. C. & Thompson, E. A. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217–1229 (2002).
109. Dice, L. R. Measures of the Amount of Ecologic Association Between Species. *Ecology* **26**, 297–302 (1945).
110. Van de Peer, Y. & De Wachter, R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci.* **10**, 569–570 (1994).
111. Schulz, B., Eckstein, R. L. & Durka, W. Scoring and analysis of methylation-sensitive amplification polymorphisms for epigenetic population studies. *Mol Ecol Resour.* **13**, 642–653 (2013).
112. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol Model.* **190**, 231–259 (2006).
113. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* **25**, 1965–1978 (2005).
114. Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K. & Toxopeus, A. G. Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**, 191–203 (2014).
115. Warren, D. L., Glor, R. E. & Turelli, M. ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611 (2010).
116. Rangel, T. F., Diniz-Filho, J. A. F. & Bini, L. M. SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography* **33**, 46–50 (2010).

Acknowledgements

This research has been funded by the project KK.01.1.1.01.0005 Biodiversity and Molecular Plant Breeding, Centre of Excellence for Biodiversity and Molecular Plant Breeding (CoE CroP-BioDiv), Zagreb, Croatia. The authors would like to thank Oleg Antonić and Martina Temunović for helpful discussion on environmental niche modelling.

Author Contributions

I.R. performed the genotypization, wrote and edited the manuscript and created all the figures. S.B. and M.F. performed the morphometric analysis. F.C. assisted in writing the manuscript. Z.S. conducted statistical analyses and revised the manuscript. B.S. participated in data interpretation and assisted in writing the manuscript. Z.L. conceived the idea for the research and extensively edited the manuscript. All authors have read and approved the final version of this manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-40080-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019