Preslia 95: 447–474, 2023
doi: 10.23855/preslia.2023.447
Preslia

Niche breadth and overlap of pseudogamous apomictic *Crataegus* hybrids and their progenitors in north-western Romania

Thomas Kuhn* & Eszter Ruprecht

Hungarian Department of Biology and Ecology, Babeş-Bolyai University, RO-400317 Cluj-Napoca, Romania
*corresponding author: kuhnthomas89@yahoo.com

Abstract: Hybridization followed by introgression is well documented within the genus *Crataegus* and although many hybrids are known and widespread, factors influencing gene flow are still poorly understood. In this study, the niche breadth and overlap between hybrids and their progenitors are compared assuming that a relatively wide niche is advantageous in the spread of hybrids outside of the hybrid zone and the ecological factors potentially influencing hybrid and progenitor distribution and frequency are investigated. For this, specimens of *Crataegus* were sampled in four subregions in north-western Romania that differed in terms of elevation, forest cover, fragmentation and the extent of anthropogenic effect on the landscape. We used dynamic range boxes to estimate niche breadth and overlap based on climatic, landscape and microenvironmental variables. CCA models were used to explore the effects of environmental factors on the distribution of taxa of *Crataegus*. Cluster analysis was used to explore the phytosociological affiliation of the taxa studied. In this field investigation, 10 *Crataegus* taxa were identified: five species and five hybrids. Most *Crataegus* hybrids were relatively rare compared to their progenitors, except *Crataegus ×subsphaerica* (*C. monogyna* × *C. rhipidophylla*), which was the most common taxon in the area studied, being present even outside of the hybrid zone, suggesting range expansion. Based on the results, it seems that the distribution of the studied *Crataegus* taxa is determined primarily by climatic factors and light availability, and they appear to be affiliated with major habitat categories, such as low altitude broad-leaved forest or grasslands and scrublands. The results also revealed that most hybrids had parent-like niche breadths, and rare hybrids had surprisingly wide niches, suggesting that hybrid frequency is influenced more by phenological overlap between progenitors and effective long-distance dispersal of apomictic seeds, rather than environmental selection against hybrids. In addition, the presumed intermediate position in ecological preferences of the hybrid *Crataegus* relative to its progenitors was not as evident as expected, probably due to introgression towards one of the progenitors. Considering the frequent occurrence of hybrids outside the hybrid zone and the mosaic distribution pattern of the habitat of sympatric species, a combination of the mosaic and evolutionary novelty model best describe hybrid zone of the studied species.

Keywords: climate, *Crataegus*, hybridization, hybrid zone, landscape, light availability, macroenvironment, microenvironment, niche breadth, niche overlap, progenitors, Romania

Received: 14 Dec 2022; Revised: 10 Oct 2023; Accepted: 14 Oct 2023; Published: 21 Dec 2023

Introduction

Interspecific hybridization occurs frequently between angiosperms and it is estimated interspecific hybrids occur in around 40% of angiosperm families, with an overall frequency of 0.09 hybrids per nonhybrid species (Arnold 1997, Whitney et al. 2010). As hybridization has great evolutionary consequences, the comparative ecology of plant hybrids and their parental taxa has been of great interest to botanists and plant ecologists for decades (Arnold & Hodges 1995, Arnold 1997, Coughlan et al. 2017, Dickinson et al. 2021). Considering the mechanism acting upon the fitness of hybrids, Arnold (1997) proposed four main models describing the structure and composition of the hybrid zone: (a) the 'tension zone model' assumes fitness inferiority of hybrids due to genetic incompatibilities between parental genomes. The hybrid zone is shaped by an equilibrium between the continuous dispersal of parental individuals into the contact zone, followed by negative endogenous selection against hybrids, preventing introgression (Arnold 1997); (b) the 'bounded hybrid superiority model' assumes that hybrids exhibit superior fitness in intermediate environments to those of their parental species, but lower fitness in the parental environments. Furthermore, hybrids often present transgressive trait expression, and parental species and hybrids usually show niche differentiation (Arnold 1997, Wang et al. 1999, Favre & Karrenberg 2011, Gallego-Tévar et al. 2018); (c) the 'mosaic model', where hybrid fitness may vary considerably, and hybrids can have lower, higher or similar fitness than their progenitors. This model proposes that the hybridizing species are sympatric, adapted to different environments and hybridization occurs in the ecotone between progenitor environments, resulting in a mosaic distribution of genotypes (Arnold 1997, Abbott 2017, Gramlich & Hörandl 2016). The mosaic structure prevents the total extinction of either of the purebred species while allowing for introgression; (d) the 'evolutionary novelty model' assumes higher fitness of certain hybrid lineages combined with transgressive trait expression, allowing hybrids to occupy novel habitats far from the hybrid zone or outcompete one particular parental species. Nevertheless, hybrid zones often do not perfectly fit any of these models (Arnold 1997, Curry 2015). This is especially true in the case of the tension zone model, where reconsideration of old data suggests that beside endogenous selection that acts against hybrids, exogenous selection also shapes the genetic patterns in the hybrid zone. While the models presented above constitute a basic framework for hybrid ecology, assumptions or predictions about niche breadth and niche overlap of hybrids and their progenitor species are scarce or missing. Such assumptions can be made based on case studies invoking different hybrid systems and considering niche sizes and overlaps between progenitors and hybrids (Lopez-Alvarez et al. 2015, Blaine Marchant et al. 2016). According to Blaine Marchant et al. (2016), during establishment, hybrids can experience niche contraction (the hybrid has a narrower niche but overlaps considerably that of its progenitors), niche expansion (the hybrid has a wider niche, which greatly overlaps that of its progenitors), niche intermediacy (hybrid niche breadth intermediate and greatly overlaps that of its progenitors) and niche novelty (breadth of hybrid niche narrower, intermediate or wider, but there is very little niche overlap between the hybrid and its progenitors). Based on the two frameworks discussed, regarding the structure of the hybrid zone and niche breadths of hybrids relative to progenitors, in the tension zone model, hybrids should manifest niche contraction, whereas in the bounded hybrid superiority model, niche expansion (Arnold 1997, Blaine Marchant et al. 2016). Furthermore, in the case of the evolutionary novel model, hybrid niche can be narrower, intermediate or wider than that of its progenitors; however, overlap should be small between hybrid and progenitor niches (Arnold 1997, Blaine Marchant et al. 2016). In the case of the mosaic model, a more complex interpretation of niche breadth and overlap is required. Based on macroenvironmental niche constraints, hybrids and parental taxa should have similar niche breadths and overlap considerably, while in the case of microenvironmental niche constraints, hybrids and parental taxa should be different in niche breadth and overlap less. This assumption is based on the sympatric nature of progenitor distributions, with progenitors differentiating in their niche along small scale environmental clines, while overlapping in their range on a broader geographical scale and macroenvironmental parameters, e.g. type of climate (Arnold 1997, Li et al. 2018, Cai et al. 2021).

The genus *Crataegus* includes a few- to several hundred shrub and small tree species native to the temperate and subtropical regions in the Northern Hemisphere, with many species being widespread and common across a wide range of habitats, of great economic and ecological importance (Christensen 1992, Kuhn et al. 2021). However, confusion and controversy still surround many aspects of *Crataegus* taxonomy, genetics and ecology, due to frequent hybridization, followed by introgression towards one or both parental species, resulting in blurred species boundaries (Byatt 1975, 1976, Christensen 1992, Depypere et al. 2006, Kuhn et al. 2021). In the case of European species of *Crataegus*, hybridization occurs mostly between sympatric species, and the resulting hybrids are usually alloploid pseudogamous apomicts (Talent & Dickinson 2005, Vašková & Kolarčik 2019). Existing studies indicate that such asexually reproducing alloploid taxa may have larger geographical distributions and thus wider niches relative to their sexually reproducing diploid parental species due to niche expansion and geographic parthenogenesis (Blaine Marchant et al. 2016, Coughlan et al. 2017, Dickinson et al. 2021). Dickinson et al. (2021) present evidence of geographic parthenogenesis in North-American species of *Crataegus*. Their study confirmed that wider geographic ranges in the case of alloploid *Crataegus* taxa of hybrid origin are the result of self-compatibility, thus pseudogamous agamospermy. In contrast, narrower ranges of diploid parental species can be associated with self-incompatibility and sexual reproduction. Furthermore, they suggest that sexual reproduction favours adaptation and investment in seedling-establishment, while geographic parthenogenesis has less to do with adaptation than reproductive insurance (Dickinson et al. 2021). While a large distribution can be associated with a wider macroenvironmental niche (e.g. climatic niche) in the case of alloploid apomictic hybrids (Blaine Marchant et al. 2016, Li et al. 2018), little is known about the smaller-scale ecological niche components that may affect hybrid distribution. Furthermore, there is little data on the effect of anthropogenic habitat alteration on hybrid formation and distribution. Ecological studies across Europe suggest that some hybrid taxa of *Crataegus* can be regionally common and may increase in abundance under anthropogenic pressure related to landscape fragmentation, while pure populations of the forest specialist species of *Crataegus* are restricted to more intact, forested landscapes (Byatt 1975, 1976, Gosler 1990, Christensen 1992, Oklejewicz et al. 2013). However, in contrast to the North-American alloploid taxa of *Crataegus* of hybrid origin, no studies or data on temperate European *Crataegus* taxa demonstrate that hybrids, despite being abundant and widespread, have a greater distribution than their progenitors, thus manifesting significant niche expansion or even niche novelty (Christensen 1992, Coughlan et al. 2017, Dickinson et al. 2021). Nevertheless, most studies investigating niche shifts in allopoid hybrids only consider macroenvironmental parameters when comparing niche breadth and overlap between hybrids and parents, e.g. climate and distribution (Lopez-Alvarez et al. 2015, Coughlan et al. 2017, Dickinson et al. 2021). However, there has not been an attempt to compare the estimated niche breadth and niche overlap between species of *Crataegus* and their hybrids, based on the disassembly of niche space into macro- and microenvironmental parameters.

This study had two main objectives: (i) to investigate the habitat requirements, frequency and distribution of species of *Crataegu*s and their hybrids in north-western Romania, where many sympatric species of *Crataegus* with diverging habitat requirements co-occur, and (ii), relying on these results, to compare the niche breadth and niche overlap of hybrids and their parental taxa using multiple niche parameter groups defined based on macro- and microenvironmental variables.

To address these objectives, the following hypotheses were proposed: (i) Hybrids with progenitors with divergent environmental requirements will have wider niches than their progenitors, since they combine parental environmental requirements. (ii) Niche breadth of the hybrid taxa and progenitor species will also correlate with their range size and abundance. (iii) Hybrids will have wider microenvironmental niche breadth, but narrower macroenvironmental niche breadth than their progenitors, due to the more restricted occurrence of hybrids in the ‘hybrid zone’, where both progenitor species meet. (iv) The niche overlap between hybrids and their parental species is expected to be greater, relative to the overlap between the two hybridizing progenitors, assuming a broader spectrum of environmental tolerance/adaptation of hybrids compared to their parents. (v) Hybrids are expected to occur in a wider range of plant associations relative to their progenitors.

Methods

Study area and site selection

Our study area was located in the north-western part of Romania and included plains and hills located in the eastern and western foothills of the Apuseni Mountains (Fig. 1). Regions at elevations higher than 800 m were not sampled, as previous field studies revealed that the number of species of hawthorn dramatically decline in abundance and frequency with altitude (Kuhn et al. 2021, Thomas et al. 2021). The climate in this area (except the higher elevations upon ~1000 m), based on the Köppen-Geiger climate classification is humid continental, characterized by four distinct seasons and large seasonal differences in temperature, with warm to hot (and often humid) summers and freezing cold winters (Beck et al. 2018). The western part of the study area is characterized by a milder climate with oceanic influences, while the eastern part has a more continental climate (Barbu et al. 2014).

Within the area studied, four regions of approximately 30 km × 30 km located in the Timiş Plain, the Zărand Mountains, the Călata Hills region and the Transylvanian Plain (Câmpia Transilvaniei) (Supplementary Fig. S1) were selected. The selection of these regions was based on their including a wide range of climatic, topographic and landscape compositions. In the case of landscape composition, two major types of ecosystem were

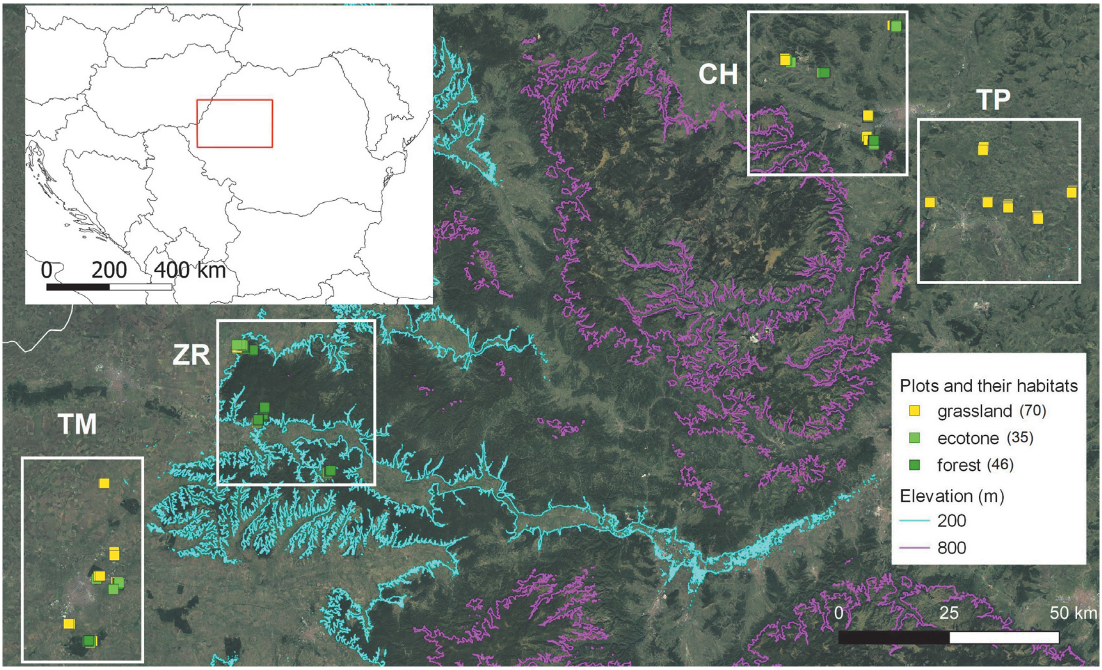


Fig. 1. Location of the study area (transparent red square) within central-eastern Europe (country borders are represented by black lines); and the distribution of the plots sampled in the four study regions in north-western Romania (abbreviations in legend: CH – Călata-Hills (continental fragmented), CT – Transylvanian Plain (continental open), TM – Timiş Plain (suboceanic fragmented), ZR – Zărandului Mountains (suboceanic forested); for each habitat category the number of plots is indicated in brackets).

selected, which are hence forth referred to as habitats, were included in the study: forest and grassland habitats. The grassland habitats are all natural and semi-natural open habitats, where tree cover was below 20%. Thus, grasslands with various degrees of shrub cover, scrublands and wooded pastures with isolated trees were included. Furthermore, forests with a canopy cover exceeding 50% were included. Forest ecosystems with more open canopies were not characteristic of the regions studied, and forest and grassland ecosystems are usually clearly distinguishable. The Timiş Plain study region is located around the city of Timişoara at an elevation around 100 m. It has suboceanic climatic influences, and is characterized by a mixed, highly fragmented landscape, consisting of large agricultural fields, while the natural vegetation is a mosaic of grasslands and forests (this region is referred to as a 'suboceanic fragmented region', Supplementary Fig. S1). The Zărand Mountains region is characterized by hills of between 200–550 m, with suboceanic climatic influences from the west and a landscape dominated by intact forests, with a small but significant amount of natural and semi-natural grassland (this region will be referred to as the 'suboceanic forested region', Supplementary Fig. S1). The Călata Hills region lies between 350–600 m elevation, has a more continental climate, natural and semi-natural vegetation predominates with agricultural land restricted mostly to the valleys, while the landscape is moderately fragmented and characterized by a mosaic of forest and patches of grassland (this region will be referred to as the ‘continental fragmented region’, Supplementary Fig. S1). The Transylvanian Plain, lies between 300–500 m elevation, and is dominated by open habitats with large areas of agricultural land with small patches of grassland between them. Forests are almost completely absent and usually consist of black-pine plantations (this region will be referred to as the ‘continental open region’, Supplementary Fig. S1).

In each of the four regions studied, six sites of approximately 3 km^2 were selected. Based on the habitat composition of the landscape, this resulted in three types of sites: sites with large areas of grassland with various degrees of shrub cover and no forest within 1 km, forested sites with no grassland within 1 km, and mixed sites with both grassland and forests present usually in a 1:1 ratio. The field survey was carried out in late summer and early autumn in 2017 and 2018. At the study sites, three 10 m × 10 m plots were established for each habitat category (grassland or forest): for sites with only one major habitat, 3–5 plots were established, while for sites with both forest and grassland, three plots were established in both grassland and forest, and three plots in the ecotone (Fig. 1). Region studied and site selection was based on aerial photographs and previous field surveys that included landscape structure, habitat fragmentation and shrub cover in grasslands, and survey plots were established during the field sampling in shrubby places where specimens of *Crataegus* were observed from a distance. Sites at which shrubs had been cleared were avoided. Plots were placed to include a maximum number of mature, fruit-bearing specimens of *Crataegus*, as young, sterile individuals are not identifiable (Christensen 1992). In total, 151 plots were sampled: 70 plots in grasslands and scrublands, 46 in forests and 35 in forest-grassland ecotones. Herbarium material from each fertile specimen of *Crataegus* growing inside the plots was collected for identification following the method described in detail by Kuhn et al. (2021). *Crataegus* taxa (both species and nothospecies or hybrids) were identified according to the system of Christensen (1992), with a few minor nomenclatural modifications: *C. laevigata* subsp. *palmstruchii* was recognized as a separate taxon, the name *C.* ×*subsphaerica* was used for the hybrid of *C. monogyna* and *C. rhipidophylla*, and *C.* ×*kyrtostyla* for the hybrid between *C. monogyna* and *C. lindmanii* (Christensen 1992, Kuhn et al. 2021). The cover (%) of each species of vascular plant within the plots was estimated separately for each vegetation layer (herbaceous, shrub and tree layer). Additionally, we recorded elevation, geographic coordinates, cover (%) of each vegetation layer, slope inclination and orientation were recorded.

Environmental variables for niche models

Environmental variables were selected and grouped for niche modelling according to the hierarchical scale at which they have been measured in terms of macroenvironmental and microenvironmental variables (Li et al. 2018). Macroenvironmental niche components were placed in two categories: climatic and landscape-metric niche components. Variables related to soil chemistry and topography were also considered; however, due to their strong correlation with climatic factors and negligible variation within the regions studied they were not included in the analysis. Five climatic variables were derived from monthly temperature and precipitation data at 30' resolution from the WorldClim database for the period 1970–2000 (https://www.worldclim.org/data/bioclim.html, Fick &

Hijmans 2017): the minimum temperature in the coldest month, temperature seasonality, precipitation seasonality, the climatic water balance for April and the climatic water balance for August (Table 1). Climatic water balance was defined as the difference between the sum of the monthly precipitation and monthly potential evapotranspiration, obtained from the mean monthly temperature using the Thorthwaite equation (Senf et al. 2020). Furthermore, considering the evidence for a slight biogeographical differentiation between some taxa of *Crataegus* (Christensen 1992), climatic factors related to potential environmental stress, such as drought and frost, that might be involved in selection against certain taxa, were recorded (Table 1). Thus, all five climatic variables were selected based on their differentiation on an east to west gradient, as being diagnostic of either continental or oceanic climates (Fick & Hijmans 2017, Beck et al. 2018). Hence forth, this niche category will be referred to as the 'climatic niche'.

As a second group of macroenvironmental components, three landscape metrics were selected, characterizing the configuration and habitat composition of the landscape surrounding the plots: grassland-forest balance, forest-edge density and distance to the nearest forest edge. For this, maps of the habitats within a 564 m radius around each plot, covering 1 km^2, where the differences between forests, open habitats and artificial surfaces were based on satellite images. An area was considered to be forest if trees were clearly visible and their cover exceeded 50%. Open habitats were areas with none or very little tree cover and still dominated by natural vegetation (e.g. grasslands, scrublands, wetlands and abandoned agricultural fields). Human-altered areas included arable lands, settlements and roads. Grassland-forest balance was calculated as the difference in the ratio of open habitat to forest, divided by the sum of these two ratios. The index varies between –1 (there are no open habitats in the landscape) and +1 (there are no forests in the landscape). Forest-edge density was calculated as the ratio between the length of forest-open habitat boundary and the area of forest (Hesselbarth et al. 2019). The distance from the nearest forest edge was expressed as the square root of the distance of the plot centroid from the nearest forest edge (Table 1). Hence forth, this niche category will be referred to as the 'landscape-configurational niche'.

The microenvironment was defined based on variables related to the vegetation and the stand characteristics, recorded in the field at the plot level. Two types of variables were included. The first category included tree canopy cover, shrub cover, cover of the herbaceous layer and a topographic heat load index, which characterize light availability and the strength of competition for light and other resources such as nutrients and water. Heat load index was calculated based on slope inclination and aspect (Parker 1991) (Table 1). The second category included the species composition at plot level and was computed using five environmental indicator values describing the plant species' optimal environment. For this, we used the indicator values for temperature, soil moisture and soil reaction from Sanda et al. (2003) and light, N content and continentality from Borhidi (1995), which are adapted versions of the indicator values of the central-eastern European flora of Ellenberg et al. (1991). First, the species abundance matrix was log transformed, and then the weighted averages of the species indicator values were calculated for each plot, which will be referred to as the biotic niche components (Table 1). Hence forth, this will be referred to as the 'microenvironmental niche'.

Table 1. Environmental variables used in niche breadth and niche overlap estimates for the taxa of *Crataegus* studied

Scale	Class	Variable (abbreviation)	Scale	Details	Units	Statistics for sampling units (min) mean (max)	Relevance	Source	Reference
Macroenvironment	climatic	temperature seasonality (Bio4)	1 km^2	the standard deviation of monthly temperature averages ×100	%	(7425) 7636 (7776)	measure of continentality	http://www.worldclim.org	Hijmans et al. 2005
		minimum temperature in coldest month (Bio6)	1 km^2	calculated by selecting the minimum temperature value across all months within a given year.	°C	(–7.8) –5.8 (–4.1)	species have different tolerances to extremely cold temperatures		Hijmans et al. 2005
		precipitation seasonality (Bio15)	1 km^2	the ratio of the standard deviation of the monthly total precipitation to the mean monthly total precipitation ×100	%	(25.0) 35.2 (44.5)	measure of continentality		Hijmans et al. 2005
		climatic water balance in April (CWB4)	1 km^2	the difference between precipitation sum of April and the potential evapotranspiration for April	mm	(35.3) 39.8 (47.3)	water availability at budburst		Senf et al. 2020
		climatic water balance in August (CWB8)	1 km^2	the difference between precipitation sum of August and the potential evapotranspiration for August	mm	(–23.7) –10.7 (–1.6)	water availability before fruit ripening		Senf et al. 2020
	landscape configuration	grassland - forest balance (GR-FR)	1 km^2	proportion of open habitats in 1 km^2 buffer; proportion of forests in 1 km^2 buffer / proportion of open habitats in 1 km^2 buffer; proportion of forests in 1 km^2 buffer	%	(–1) 0.21 (1)	the ratio of open habitats and forest in the landscape, the fragmentation of these habitats and the edge-effect around habitat borders can influence intrusion of forest species in nearby grasslands, thus probability of hybridization	hand drawn habitat maps based on satellite images	
		forest-edge density (FED)	1 km^2	the length of forest edges in a 1 km buffer / area covered by forests in the 1 km^2 buffer	m/ha	(0) 25.6 (112.8)			
		distance to nearest forest-edge (DISTF)	1 km^2	the square root of the distance of the plot centroid from the nearest forest-edge (becomes negative if the plot is inside a forest)		(–59.5) 8.5 (79.7)			
Microenvironment		canopy cover	100 m^2	100-(the proportional cover in a horizontal plain of the sky by the tree layer)	%	(0) 31.7 (99)	sympatric species differentiate on a small scale regarding light environment, competition from herbaceous plants and other shrubs		
		herb layer cover	100 m^2	the proportional cover in a horizontal plain of the herbaceous layer	%	(1) 59 (100)			
		shrub layer cover	100 m^2	the proportional cover in a horizontal plain of the shrub layer	%	(0) 37 (95)			
		heat index	100 m^2	calculated from slope inclination and aspect based on Parker (1988)	%	(–0.43) 0.06 (1)			
		temperature	100 m^2	the average temperature optimum of plant species in a plot (weight for log-transformed abundance)	–	(2.8) 3.3 (4.2)	based on the assumption that vegetation composition is indicative of microenvironmental conditions		Borhidi 1995; Ellenberg et al. 1991
		soil moisture	100 m^2	the average moisture optimum of plant species in a plot (weight for log-transformed abundance)	–	(1.7) 2.6 (4.2)			
		soil reaction	100 m^2	the average soil reaction optimum of plant species in a plot (weight for log-transformed abundance)	–	(3.0) 3.6 (4.1)			
		nitrogen content	100 m^2	the average soil nitrogen optimum of plant species in a plot (weight for log-transformed abundance)	–	(2.0) 4.2 (6.7)			
		continentality	100 m^2	the average continentality optimum of plant species in a plot (weight for log-transformed abundance)	–	(3.8) 4.7 (6.2)			

Statistical analysis

To estimate the niche breadth and niche overlap between species of *Crataegus* and their hybrids in the area studied, dynamic range boxes (dynRB) were used, which is a robust nonparametric approach to quantifying size (a proxy for niche breadth) and overlap of n-dimensional hypervolumes, which do not require assumptions about the underlying distribution (Junker et al. 2016). In an n-dimensional hypervolume, each environmental variable represents one dimension and each *Crataegus* specimen is considered to be one point within this coordinate system. To reduce the effect of outliers, i.e. plots with exceptionally high numbers of specimens of *Crataegus* on the niche estimation procedure, the following calculations and modifications of the raw data were carried out: first, the *Crataegus* specimen counts were log transformed and values rounded up to the closest integer; second, each row was replicated using recorded values so as the number of rows per plot per taxa reflects the abundance of *Crataegus* taxa; third, environmental variables representing niche components were added as columns to the data. Due to low sample size, two *Crataegus* taxa were excluded from the analysis, *Crataegus ×kyrtostyla* with three records and *C. laevigata* subsp. *palmstruchii* with only one. The 'dynRB_Pa()' function from the 'dynRB' package in R was used with the geometric mean (gmean) aggregation method to calculate hypervolume size vol(Taxa1) as a proxy for relative niche breadth of taxa and hypervolume overlap port(Taxa1, Taxa2) as a proxy for relative niche overlap between taxa. The gmean aggregation method was selected because it assumes that the hypervolume becomes zero if size and overlap in just one of the dimensions is zero, while size and overlap are not biased by the number of dimensions, thus are comparable between hypervolumes with different numbers of dimensions. Both hypervolume size vol(Taxa1) and overlap port(Taxa1, Taxa2) are set between 0 and 1. Hypervolume size vol(Taxa1) represents the proportion of the total trait space in the data that is covered by the environmental niche space of Taxa1. The overlap port(Taxa1,Taxa2) represents the proportion of environmental niche space of Taxa2 covered by Taxa1, while port(Taxa2,Taxa1) represents the proportion of environmental niche space of Taxa1 covered by Taxa2, thus overlaps are asymmetrical (Junker et al. 2016), and port(Taxa1, Taxa2) does not equal port(Taxa2, Taxa1). Hypervolume size and overlap, being relative measures, are comparable between different groups of traits (e.g. climatic niche vs. landscape-configuration niche). Separate niche models were computed for climatic, landscape configuration and microenvironmental niche variables. Furthermore, because some of the selected environmental variables are correlated (Supplementary Fig. S2), the original dimensions were replaced with principal components, which were used in subsequent calculations (Junker et al. 2016).

To explore the underlying influence of single environmental variables on niche differentiation and the direction in which these variables influence *Crataegus* composition, separate constrained CCA models were developed for each of the three sets of environmental variables used in the niche models. As a dependent term the log-transformed *Crataegus* taxa count matrix was included, while the environmental variables are independent terms. In addition, type of habitat was included as a proxy covariate for the microenvironment, in order to explore the pure effect of the macroenvironment on *Crataegus* taxa composition, while partialling out the effect of the microenvironment. Alternatively, in the case of the CCA models with microenvironmental constraints as independent terms, study region was included as a proxy covariate to partial out the effect of the macroenvironment on *Crataegus* taxa composition. As our variable groups are correlated, a bidirectional model selection was carried out in order to exclude variables that do not have a significant effect on the model, while keeping the conditional term fixed. Furthermore, the pure effect of each of the significant variables included in the pCCA models on *Crataegus* taxa composition was also explored.

Niche breadth based on phytosociological affiliation of *Crataegus* taxa was investigated using divisive hierarchical clustering of the plot data. Species abundances were Hellinger-transformed, and the distance matrix obtained by clustering was obtained using the Bray-Curtis dissimilarity index (Pakgohar et al. 2021). Information regarding the vertical layering of the vegetation in the plot data was included in the analysis by merging the name of a particular layer with species names within the layer, thus avoiding duplicate species in the data. In this way, species present in more than one vegetation layer were considered statistically separate species. The optimal number of clusters was estimated using the gap statistical method with 500 permutations (Kassambara & Mundt 2020). In addition, plots were classified in a plant association based on the Romanian classification system (Sanda et al. 2008). Each plot located in either a forest or grassland ecosystem was classified into one syntaxon, while in the case of ecotonal plots, two associations were attributed to each plot, based on the two syntaxa meeting at the plot level. Both hierarchical relationships between the plots and the number of phytosociological associations were included in the estimate of the niche breadth of a particular taxon. The most important R-packages used were the ‘vegan’ package for the pCCA ordinations and calculation of the dissimilarity index (Oksanen et al. 2022) and the ‘cluster’ package used for divisive hierarchical clustering (Maechler et al. 2022). All statistical analyses were done using R (R Core Team 2021).

Results

Distribution patterns of Crataegus species and their hybrids

In total, 10 *Crataegus* taxa were identified in the area studied, five “pure” species, subspecies or varieties, and five nothospecies, nothosubspecies or nothovarieties (which will be referred to as “hybrids”) (Christensen 1992). The species identified were *Crataegus monogyna*, *C. rhipidophylla* var. *rhipidophylla*, *C. rhipidophylla* var. *lindmanii*, *C. laevigata* subsp. *laevigata* and *C. laevigata* subsp. *palmstruchii* and the hybrids *C.* ×*subsphaerica* (hybrid of *C. monogyna* and *C. rhipidophylla*), *C.* ×*media* (hybrid of *C. monogyna* and *C. laevigata*), *C.* ×*kyrtostyla* (hybrid of *C. monogyna* and *C. lindmanii*), *C.*×*macrocarpa* nothosubsp. *macrocarpa* (hybrid of *C. laevigata* and *C. rhipidophylla*) and *C.* ×*macrocarpa* nothosubsp. *hadensis* (hybrid of *C. laevigata* and *C. lindmanii*) (Supplementary Table S1, Supplementary Fig. S3). The names of infraspecific taxa are referred to as binomials.

Crataegus monogyna occurred relatively frequently in all four regions studied in open habitats such as grassland, scrubland and rarely wetlands, and it was also common in forest edges (Fig. 2). *Crataegus rhipidophylla* was also common, but mainly restricted to forest habitats and their periphery, sporadically also occurring in open habitats such as grassland and scrubland, however, usually in close proximity to forests (Fig. 2). *Crataegus*

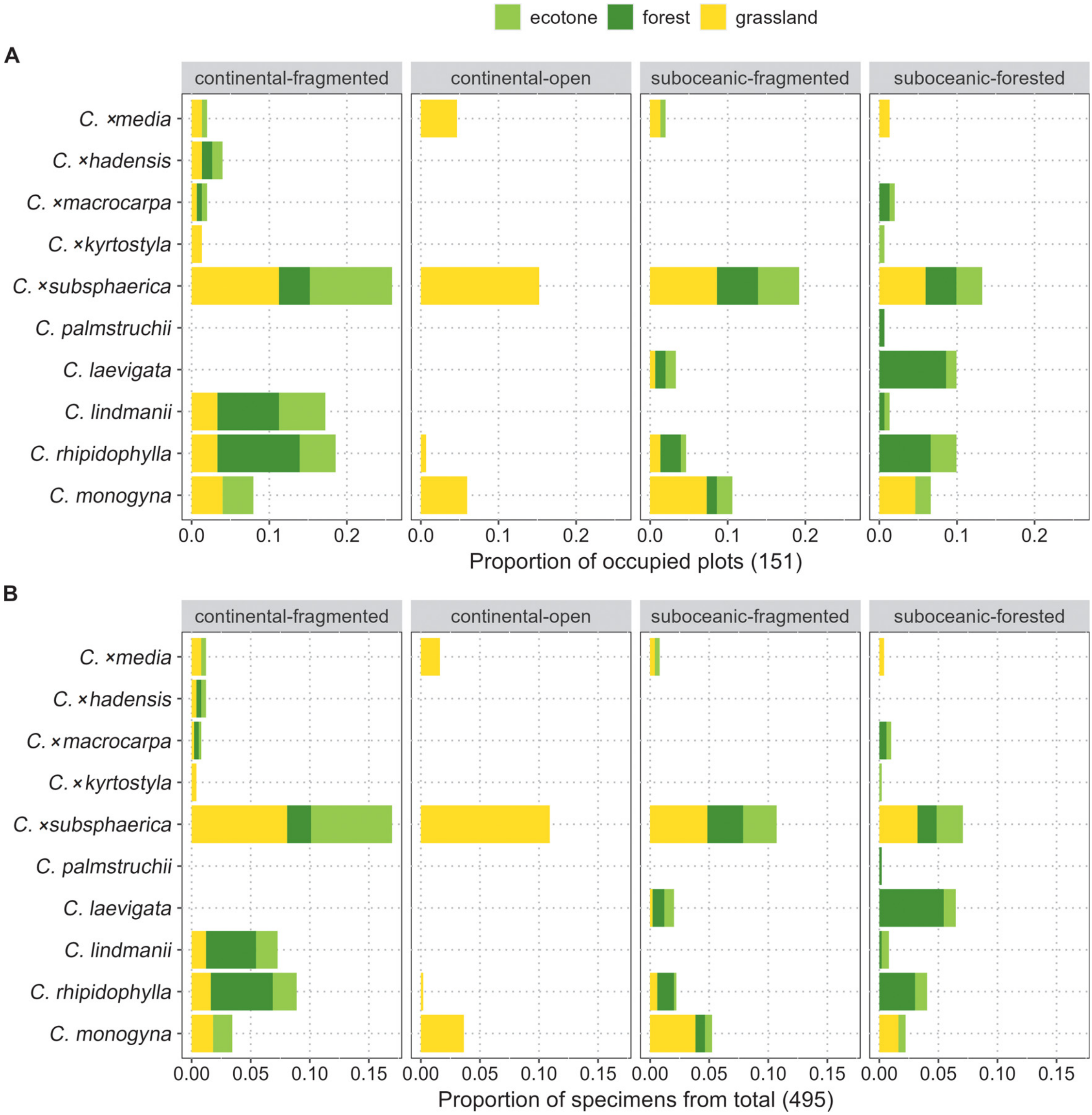


Fig. 2. Frequency (A – the proportion of occupied plots) and abundance (B – proportion of specimens in the total) of taxa of *Crataegus* in the study area, by region. Parentheses on the y-axis indicate total sampling effort relative to which the proportions were calculated.

laevigata was less frequent than *C. rhipidophylla*, and occurred in forest ecosystems mainly in the western part of the area studied, while *C. lindmanii* was sporadic in forests and forest edges in the continental fragmented region and the suboceanic forested region (Fig. 2). *Crataegus palmstruchii* was present in a single plot in the suboceanic forested region (Fig. 2). The most common hybrid and also the most common and widespread

Crataegus taxon identified was *C.* ×*subsphaerica*, which was present in every region studied in open habitats, at forest edges and in forest ecosystems, but was significantly less common in the latter (Fig. 2). Other hybrids were mostly sporadic or rare, with *C.* ×*media* occurring sporadically in open habitats and at forest edges in all four regions studied, *C.* ×*macrocarpa* and *C.* ×*hadensis* occurring rarely in open habitats, at forest edges and in forests in the continental fragmented region and suboceanic forested region, while *C.* ×*kyrtostyla* being very rare in open habitats and at forest edges in the continental fragmented and the suboceanic forested regions (Fig. 2).

Niche breadth

Crataegus taxa (both species and hybrids) varied greatly in niche breadth based on hypervolume sizes, with values ranging from 0.36 to 0.99 of the total available hypervolume (Fig. 3). Both macroenvironmental niche categories had greater between taxa variability in hypervolume compared to the microenvironmental niche, which also had a greater hypervolume (Fig. 3). In the case of climatic niche breadth, *C. monogyna* and its two hybrids, *C.* ×*subsphaerica* and *C.* ×*media* had the widest niches, followed by *C. rhipidophylla*, *C.* ×*macrocarpa*, *C. lindmanii*, and *C.* ×*hadensis* and *C. laevigata* had the narrowest (Fig. 3). In the case of the landscape-configuration niche, *C.* ×*subsphaerica*, *C. monogyna* and *C. laevigata* had the widest niches, followed by *C.* ×*macrocarpa*, *C.* ×*media* and *C. rhipidophylla*, and the smallest hypervolumes were estimated for *C.* ×*hadensis* and *C. lindmanii* (Fig. 3). Furthermore, in the case of the microenvironmental niche, hypervolume sizes were in general high, the widest niches were estimated for *C.* ×*subsphaerica*, *C. monogyna*, *C. lindmanii* and *C. rhipidophylla*, followed by *C.* ×*hadensis*, and the smallest for *C.* ×*macrocarpa* and *C.* ×*media* (Fig. 3).

Niche overlap

Crataegus monogyna, *C. rhipidophylla* and their hybrid *C.* ×*subsphaerica*

Niche overlap between *C. monogyna* and *C. rhipidophylla* was low to intermediate, with lower overlap values in the case of the landscape-configuration niche relative to the other two niche categories (Fig. 4A–C). Furthermore, the overlap between the hybrid *C.* ×*subsphaerica* and both of its parental species was in general higher than the niche overlap between the two parents, and the hybrid *C.* ×*subsphaerica* overlapped more with the niche of *C. monogyna* than that of *C. rhipidophylla* in the case of all three niche categories (Fig. 4A–C). In addition, the overlap between *C.* ×*subsphaerica* and *C. rhipidophylla* as well as *C.* ×*subsphaerica* and *C. monogyna* was asymmetrical, with port(*C.* ×*subsphaerica*, *C. rhipidophylla*) and port(*C.* ×*subsphaerica*, *C. monogyna*) having higher overlap values than port(*C. rhipidophylla*, *C.* ×*subsphaerica*) and port(*C.* ×*subsphaerica*, *C. rhipidophylla*) in the case of all three niche categories (Fig. 4A–C).

Crataegus monogyna, *C. laevigata* and their hybrid *C.* ×*media*

The overlap between *C. monogyna* and *C. laevigata* was low to intermediate, with the highest values for the climatic niche in the case of port(*C. monogyna*, *C. laevigata*). Furthermore, overlaps were relatively symmetrical for two of the three niche categories, but

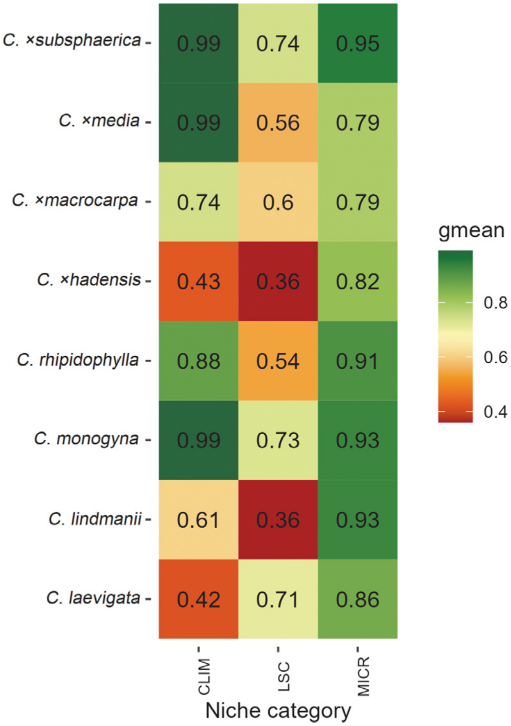
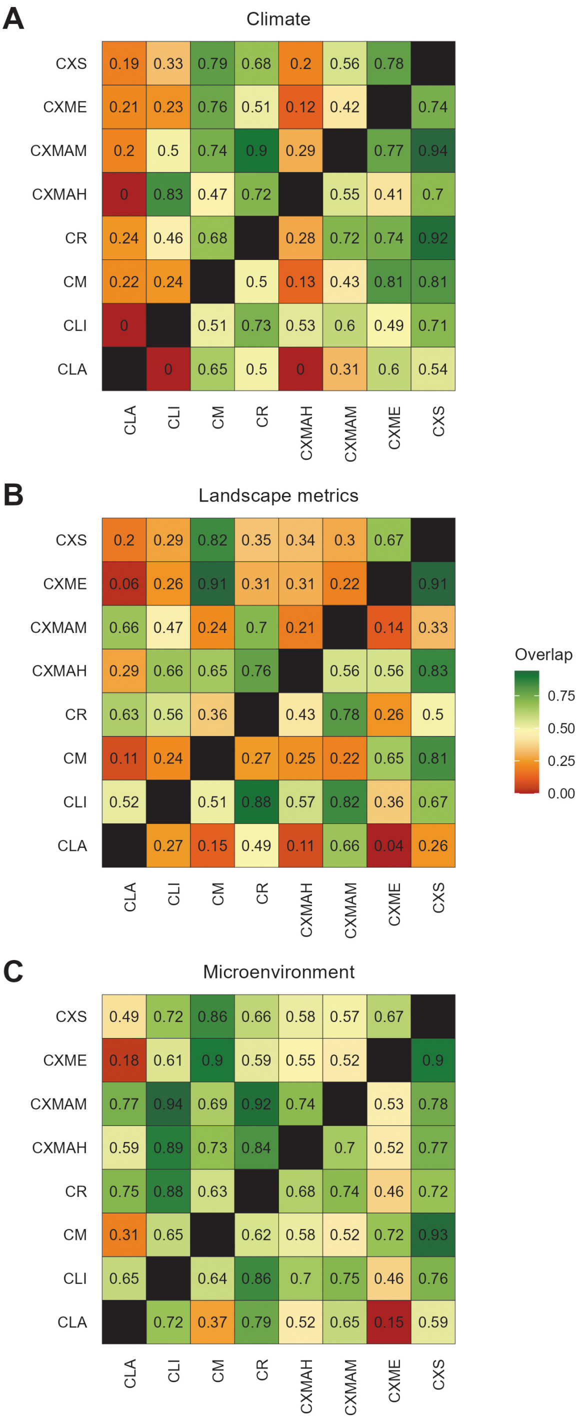


Fig. 3. Hypervolume size estimates with the gmeans-aggregation function for the taxa of *Crataegus* studied for each niche category (CLIM – climatic niche, LSC – landscape-configuration niche, MICR – microenvironmental niche; Table 2).

not for the climatic niche, where port(*C. monogyna*, *C. laevigata*) was much greater than port(*C. laevigata*, *C. monogyna*) (Fig. 4A–C). The hybrid *C.* ×*media* overlapped with *C. monogyna* to a higher degree than with *C. laevigata*, with lower overlap values for the latter in the case of landscape configuration and microenvironmental niche than for climatic niche (Fig. 4A–C). Furthermore, niche overlap asymmetry was detected between port(*C. monogyna*, *C.* ×*media*) and port(*C.* ×*media*, *C. monogyna*) and port(*C.* ×*media*, *C. laevigata*) and port(*C. laevigata*, *C.* ×*media*) (Fig. 4A–C).

Crataegus laevigata, *C. rhipidophylla* and their hybrid *C.* ×*macrocarpa*

The degree of overlap between *C. laevigata* and *C. rhipidophylla* was very variable and depended on the niche category and was also asymmetrical (Fig. 4A–C). The overlap between the two species is the highest for the microenvironment, followed by landscape configuration and the lowest for the climatic niche. Furthermore, the overlap between the hybrid *C.* ×*macrocarpa* and *C. rhipidophylla* was in general higher than the overlap between the two parental species in the case of the climatic and landscape-configuration niche, while in the case of the microenvironment the overlaps were more similar. In addition, in the case of the climatic and landscape-configuration niches, the overlap between the hybrid *C.* ×*macrocarpa* and *C. rhipidophylla* was higher than the overlap between *C.* ×*macrocarpa* and *C. laevigata* (Fig. 4A–C).



Climate

	CLA	CLI	CM	CR	CXMAH	CXMAM	CXME	CXS
CXS	0.19	0.33	0.79	0.68	0.2	0.56	0.78	
CXME	0.21	0.23	0.76	0.51	0.12	0.42		0.74
CXMAM	0.2	0.5	0.74	0.9	0.29		0.77	0.94
CXMAH	0	0.83	0.47	0.72		0.55	0.41	0.7
CR	0.24	0.46	0.68		0.28	0.72	0.74	0.92
CM	0.22	0.24		0.5	0.13	0.43	0.81	0.81
CLI	0		0.51	0.73	0.53	0.6	0.49	0.71
CLA		0	0.65	0.5	0	0.31	0.6	0.54

Landscape metrics

	CLA	CLI	CM	CR	CXMAH	CXMAM	CXME	CXS
CXS	0.2	0.29	0.82	0.35	0.34	0.3	0.67	
CXME	0.06	0.26	0.91	0.31	0.31	0.22		0.91
CXMAM	0.66	0.47	0.24	0.7	0.21		0.14	0.33
CXMAH	0.29	0.66	0.65	0.76		0.56	0.56	0.83
CR	0.63	0.56	0.36		0.43	0.78	0.26	0.5
CM	0.11	0.24		0.27	0.25	0.22	0.65	0.81
CLI	0.52		0.51	0.88	0.57	0.82	0.36	0.67
CLA		0.27	0.15	0.49	0.11	0.66	0.04	0.26

Microenvironment

	CLA	CLI	CM	CR	CXMAH	CXMAM	CXME	CXS
CXS	0.49	0.72	0.86	0.66	0.58	0.57	0.67	
CXME	0.18	0.61	0.9	0.59	0.55	0.52		0.9
CXMAM	0.77	0.94	0.69	0.92	0.74		0.53	0.78
CXMAH	0.59	0.89	0.73	0.84		0.7	0.52	0.77
CR	0.75	0.88	0.63		0.68	0.74	0.46	0.72
CM	0.31	0.65		0.62	0.58	0.52	0.72	0.93
CLI	0.65		0.64	0.86	0.7	0.75	0.46	0.76
CLA		0.72	0.37	0.79	0.52	0.65	0.15	0.59

Fig. 4. Niche overlaps between taxa of *Crataegus* for (A) climatic, (B) landscape-configuration and (C) microenvironmental niche. Taxa abbreviations follow Table 2.

Crataegus laevigata, *C. lindmanii* and their hybrid *C.*×*hadensis*

The degree of overlap between *C. laevigata* and *C. lindmanii* was very variable with higher overlap estimates for microenvironmental niche, intermediate values for the biotic niche and landscape-configuration niche and no overlap in the case of climatic niche. Overlaps were relatively symmetrical, except in the case of the landscape-configuration niche, where port(*C. laevigata*, *C. lindmanii*) was greater than port(*C. lindmanii*, *C. laevigata*) (Fig. 4A–C). The hybrid *C.* ×*hadensis* overlapped to a greater degree with the *C. lindmanii* parent, than the parental species with each other in the case of all three niche categories (Fig. 4A–C). This is, however, not the case for the overlap between the hybrid *C.* ×*hadensis* and *C. laevigata*, which overlapped two or three times less than that recorded for *C. laevigata* and *C. rhipidophylla* (Fig. 4A–C). Asymmetry in pairwise overlap was also detected, with greater overlap values for port(*C. lindmanii*, *C.* ×*hadensis*) than port(*C.* ×*hadensis*, *C. lindmanii*), and for port(*C. laevigata*, *C.* ×*hadensis*) than port(*C.* ×*hadensis*, C. *laevigata*) (Fig. 4A–C).

Influence of single environmental factors on niche differentiation between Crataegus taxa

Based on the results of the pCCA model, the explained constrained variance of *Crataegus* taxa composition was comparatively similar between the climatic and microenvironmental niche categories, while in the case of the landscape-configuration niche, the variance was about half that of the other two niche categories. In the case of the reduced pCCA models via model selection, the constrained variance explained was 8.3% for the two significant climatic variables (climatic water balance for April, climatic water balance for August), 4.3% for two significant landscape configurational variables (grassland-forest balance, forest-edge density) and 7.8% for canopy cover as the only significant microenvironmental variable. The conditional term ‘habitat’ in the macroenvironmental pCCA models accounted for 12.5%, while the conditional term ‘study region’ explained 13.5% of the variance in the microenvironmental pCCA models (Supplementary Table S2).

In the case of climatic niche components, climatic water balance in April and climatic water balance in August had a significant effect on *Crataegus* taxa composition. *Crataegus laevigata* and *C.* ×*macrocarpa* were associated with the highest values for climatic water balance in April, while *C.* ×*hadensis* was at the opposite end of this gradient, and *C. monogyna*, *C. rhipidophylla*, *C. lindmanii*, *C.* ×*subsphaerica* and *C.* ×*media* occupied a central position (Fig. 5A). In the case of climatic water balance in August, *C.* ×*macrocarpa* and *C. lindmanii* had the highest values for this variable, followed by *C.* ×*hadensis* and *C. laevigata*. Furthermore, *C. rhipidophylla*, *C.* ×*subsphaerica* and *C.* ×*media* had intermediate values, while *C. monogyna* had the lowest values for climatic water balance in August (Fig. 5B). In the case of landscape-configuration components, grassland-forest balance and forest-edge density had significant effects on taxa composition. *Crataegus* ×*hadensis* was associated with the highest, *C. laevigata* and *C.* ×*macrocarpa* with the lowest values for the grassland-forest balance, while *C.* ×*media*, *C. monogyna*, *C. rhipidophylla* and *C. lindmanii* had intermediate values for this variable (Fig. 5C). In case of the forest-edge density, the highest values were recorded for *C.* ×*subsphaerica* and *C. monogyna*, with *C. laevigata* and *C.* ×*macrocarpa* at the opposite end of this gradient, while *C.* ×*media*, *C. rhipidophylla* and *C.* ×*hadensis* occupy more intermediate positions (Fig. 5B). Furthermore, according to the CCA model containing

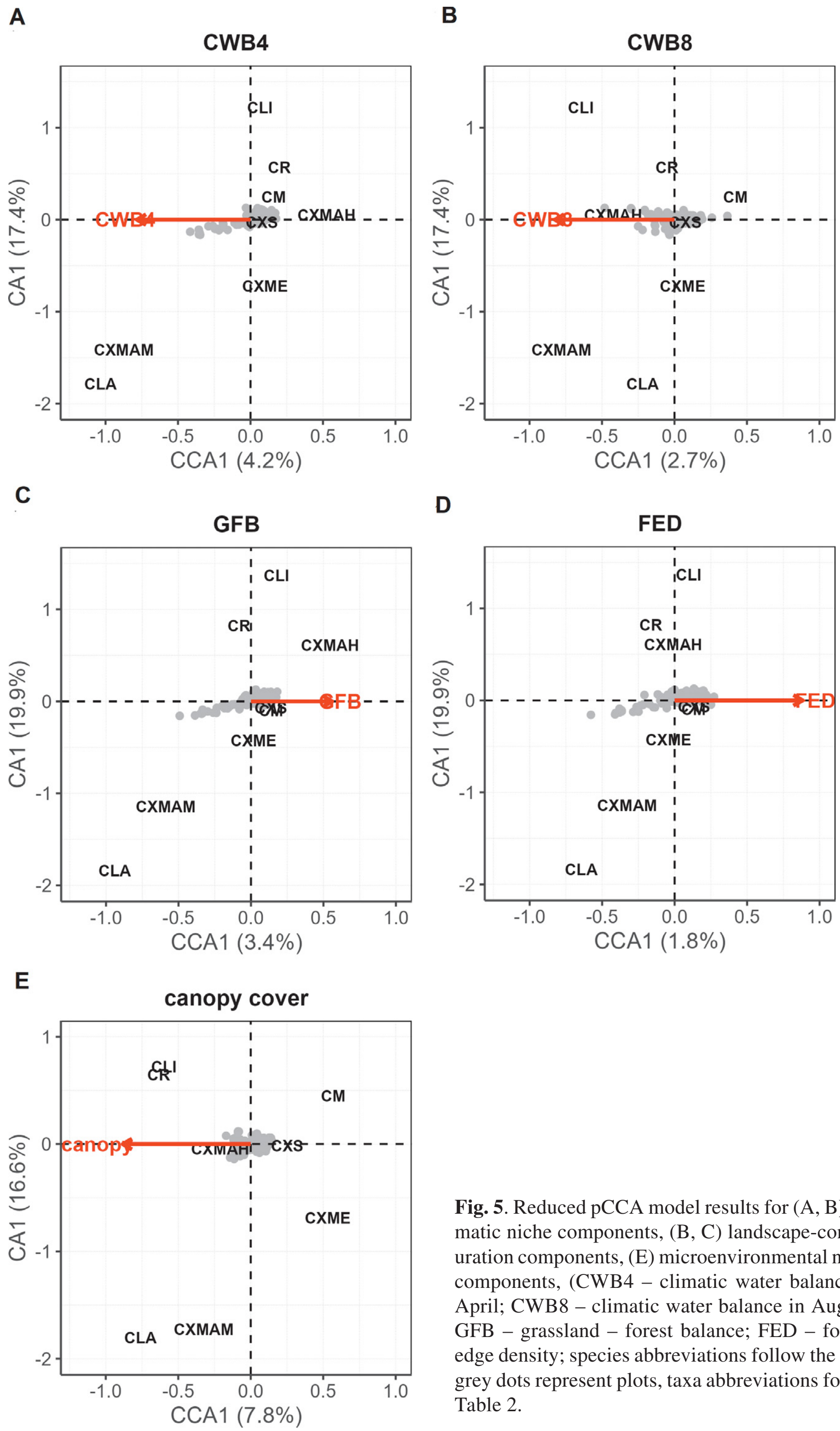


Fig. 5. Reduced pCCA model results for (A, B) climatic niche components, (B, C) landscape-configuration components, (E) microenvironmental niche components, (CWB4 – climatic water balance in April; CWB8 – climatic water balance in August; GFB – grassland – forest balance; FED – forest-edge density; species abbreviations follow the text; grey dots represent plots, taxa abbreviations follow Table 2.

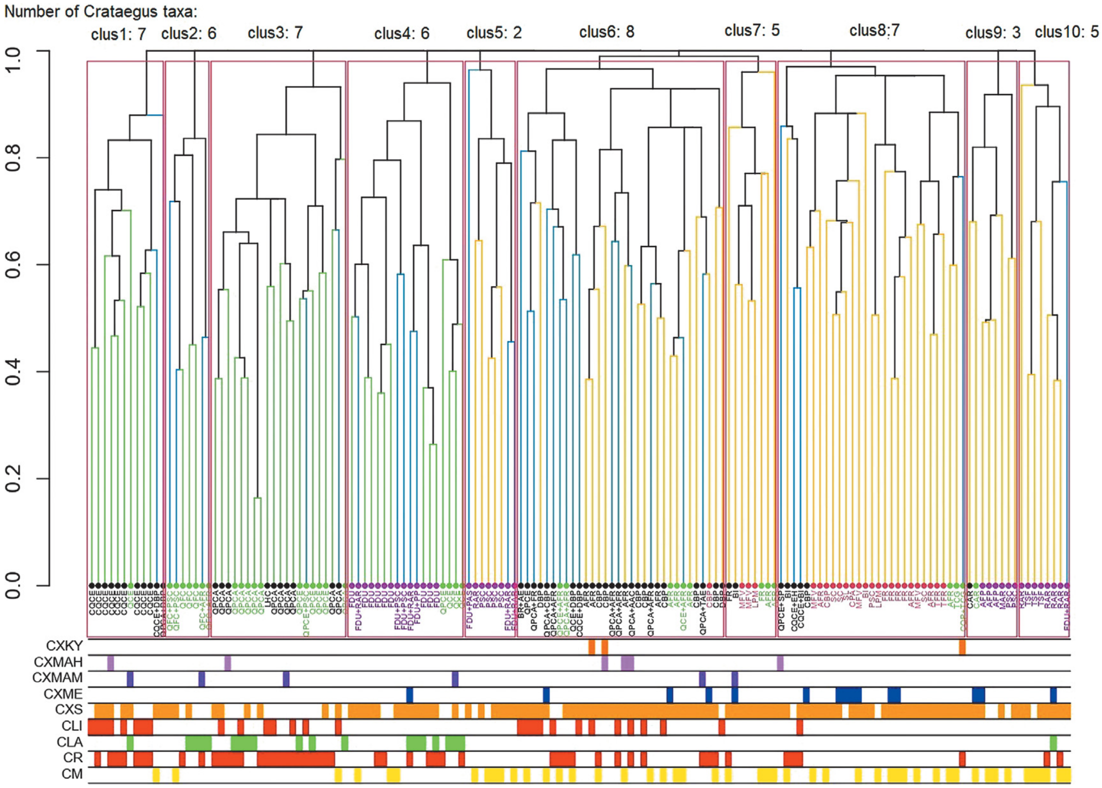


Fig. 6. Dendrogram resulting from the divisive hierarchical clustering of plot data. Red boxes show the 10 optimal clusters determined by the gap statistical method. Branch colours indicate type of habitat (green – forests, yellow – open habitats, blue – ecotone), leaf colour indicates study region (green – suboceanic forested, purple – suboceanic fragmented, black – continental fragmented, red – continental open). Plant associations to which the plots belong are listed below the leaves (abbreviations follow Table 2). Occurrence of taxa of *Crataegus* in plots is indicated by the coloured vertical lines at the bottom (taxa abbreviations follow Table 2).

microenvironmental components, only canopy cover had a significant effect on taxa composition. *Crataegus laevigata* had the highest canopy cover values closely followed by *C. rhipidophylla*, and *C. lindmanii*, while *C.* ×*media* and *C. monogyna* had the lowest values. *Crataegus* ×*subsphaerica*, *C.* ×*hadensis* and *C.* ×*macrocarpa* had intermediate cover values (Fig. 5C).

Phytosociological affiliation

The optimal number of clusters determined by the gap statistical method was 10 (Fig. 6, Supplementary Table S3). There was a clear distinction between forest and grassland clusters, with grassland plots classified into six neighbouring clusters, while forest plots were classified into four clusters (Fig. 6). Ecotonal plots were scattered across clusters of forest and grassland ecosystems (Fig. 6). Furthermore, plots in the different study regions differed clearly within the forest and grassland clusters, forming more or less compact sub clusters.

Table 2. Plant associations classified at the plot level (Code – association code used in the dendrogram in Fig. 6. Taxa abbreviations: CM – *C. monogyna*, CR – *C. rhipidophylla*, CLA – *C. laevigata*, CPA – *C. palmstruchii*, CLI – *C. lindmanii*, CXS – *C.* ×*subsphaerica*, CXME – *C.* ×*media*, CXMAM – *C.* ×*macrocarpa*, CXMAH – *C.* ×*hadensis*; Count – the first number represents the number of taxa of *Crataegus* present in a certain association while the second number followed by the letter 'e' represents the number of taxa found only in the ecotone of a certain association).

No.	Association	Class	Code	Ecosystem	Taxa	Count
1	*Convolvulo-Agropyretum repentis* Felföldy 1943	*Artemisietea vulgaris* Lohmeyer et al. in R. Tüxen 1950	CAR	open habitat (ruderal)	CM, CXME, CXS	4
2	*Agrostio-Festucetum rupicolae* Csűrös-Kaptalan 1964	*Festuco-Brometea* Br.-Bl. et R. Tüxen in Br.-Bl. 1949	AFR	open habitat (grassland)	CLI, CM, CXKY, CXMAH, CXME, CXS	6+3e
3	*Botriochloetum ischaemi* (Kristiansen 1937) Pop 1977		BI	open habitat (grassland)	CM, CR, CXMAM, CXME, CXS	5+1e
4	*Carici humilis-Brachypodietum pinnati* Soó (1942) 1947		CBP	open habitat (grassland)	CLI, CM, CR, CXKY, CXMAH, CXME, CXS	7
5	*Cynodonti-Poëtum angustifoliae* Rapaics ex Soó 1957		CPA	open habitat (grassland)	CXME, CXS	2
6	*Danthonio-Brachypodietum pinnati* Soó (1946) 1947		DBP	open habitat (grassland)	CLI, CXS	2+3e
7	*Elytrigietum hispidi* (Dihoru 1970) Popescu et Sanda 1988		EH	open habitat (grassland)	-	1e
8	*Festucetum rupicolae* Burduja et al. 1956		FR	open habitat (grassland)	CM, CR, CXME, CXS	4+1e
9	*Medicagini-Festucetum valesiacae* Wagner 1940		MFV	open habitat (grassland)	CM, CXME, CXS	3
10	*Stipetum capillatae* (Hueck 1931) Krausch 1961		SC	open habitat (grassland)	CXME, CXS	2
11	*Stipetum lessingianae* Soó (1927 n.n.) 1947		SL	open habitat (grassland)	CXME	1
12	*Stipetum pulcherrimae* Soó 1942		SP	open habitat (grassland)	-	2e
13	*Taraxaco serotinae-Festucetum valesiacae* (Burduja et al. 1956, Răvăruţ et al. 1956) Sârbu et al. 1999		TSF	open habitat (grassland)	CM, CXS	2
14	*Thymio pannonici-Chrysopogonetum grylli* Doniţă et al. 1992		TCG	open habitat (grassland)	CM, CXS	2+2e
15	*Thymo comosi-Festucetum rupicolae* (Csűrös et Gergely 1959) Pop et Hodişan 1985		TFR	open habitat (grassland)	CM, CXS	2
16	*Anthoxantho-Agrostietum capillaris* Sillinger 1933	*Molinio-Arrhenatheretea* R. Tüxen 1937	AAC	open habitat (grassland)	-	4e
17	*Medicagini lupulinae-Agropyretum repentis* Popescu et al. 1980		MAR	open habitat (ruderal)	CXS	1
18	*Poëtum pratensis* Răvăruţ et al. 1956		PP	open habitat (grassland)	-	2e
19	*Arrhenatheretum elatioris* Br.-Bl. ex Scherrer 1925		AE	open habitat (grassland)	-	2e
20	*Ranunculo repenti-Alopecuretum repentis* Ellmauer et Mucina in Mucina et al. 1993		RAR	open habitat (grassland)	CLA, CM, CR, CXME, CXS	5
21	*Lolio-Plantaginetum majoris* (Linkola 1921) Beger 1930 em. Sissingh 1969	*Plantaginetea majoris* R. Tüxen et Preising 1950	LPM	open habitat (ruderal)	CM, CXS	2
22	*Achilleo-Festucetum pseudovinae* (Magyar 1928) Soó (1933) 1945	*Puccinellio-Salicornietea* Ţopa 1939	AFP	open habitat (grassland)	CM, CXME, CXS	3
23	*Peucedano officinalis-Asteretum sedifolii* Soó 1947 corr. Borhidi 1996		PAS	open habitat (grassland)	-	1e

No.	Association	Class	Code	Ecosystem	Taxa	Count
24	*Bromo sterilis-Robinietum pseudacaciae* (Pócs 1954) Soó 1964	*Quercetea pubescenti-petraeae* (Oberdorfer 1948) Jakucs 1960	BR	forests	-	2e
25	*Corno-Quercetum pubescentis* Jakucs et Zsólyomi ex Máthé et Kovács 1962		CQP	forests	-	3e
26	*Lychno coronariae-Quercetum cerris* Sanda et al. 2003		QCE	forests	CLA, CR, CXMAM, CXS	4+1e
27	*Quercetum frainetto-cerris* Georgescu 1945, Rudski 1949		QFC	forests	CLA, CPA, CR, CXMAM, CXS	5+2e
28	*Quercetum petraeae-cerris* Soó 1963		QPCE	forests	CLA, CLI, CM, CR, CXS	5+1e
29	*Carpino-Quercetum cerris* Kilka 1938	*Querco-Fagetea* Br.-Bl. et Vlieger in Vlieger 1937 em. Borhidi 1996	CQCE	forests	CLI, CR, CXMAH, CXS	4+1e
30	*Fraxino danubialis-Ulmetum* Sanda et Popescu 1999		FDU	forests	CLA, CM, CR, CXME, CXS	4+1e
31	*Lathyro hallersteinii-Carpinetum* Coldea 1975		LHC	forests	CLI, CR	2
32	*Querco petraeae-Carpinetum* Soó et Pócs 1957		QPCA	forests	CLA, CLI, CM, CR, CXMAH, CXMAM, CXME, CXS	7+1e
33	*Pruno spinosae-Crataegetum* (Soó 1927) Hueck 1931	*Rhamno-Prunetea* Rivas Goday et Borja Carbonell 1961	PSC	open habitat (scrubland)	CM, CXS	2+2e
34	*Trifolio-Agrimonietum eupatoriae* Th. Müller (1961) 1962	*Trifolio-Geranietea* sanguinei Th. Müller 1961	TAE	open habitat (grassland)	-	3e

Crataegus monogyna was spread comparatively evenly throughout the six grassland-ecosystem clusters, and sporadically present also in the forest ecosystem clusters, but only in ecotonal plots (Fig. 6, Supplementary Table S3). *Crataegus rhipidophylla* was common in each forest cluster, but occurred sporadically in four grassland clusters (Fig. 6, Supplementary Table S3). The hybrid *C.* ×*subsphaerica* had the highest cluster diversity of all taxa, being present in all clusters, but with a higher frequency in grassland clusters (Fig. 6, Supplementary Table S3). *Crataegus laevigata* was sporadically present only in three forest clusters, with one occurrence in a fourth forest cluster in an isolated plot in the suboceanic forested region included between plots from the continental fragmented region. It was also present in a single plot in the last grassland cluster grouping the suboceanic fragmented region (Fig. 6, Supplementary Table S3). Despite its comparative rarity, *C.* ×*media* was spread throughout the grassland clusters, with only one occurrence in a forest cluster, in an ecotonal plot (Fig. 6, Supplementary Table S3). *Crataegus* ×*macrocarpa* was present in isolated plots in all four forest clusters, while occurring in two grassland clusters (Fig. 6, Supplementary Tab. S3). *Crataegus lindmanii* was sporadically present only in two forest clusters, and two grassland clusters, however, with only an isolated occurrence in one of these two clusters (Fig. 6, Supplementary Table S3). *Crataegus* ×*hadensis* was present in isolated plots only within two forest clusters and two grassland clusters (Fig. 6, Supplementary Table S3).

In total, 34 different plant associations were identified in which *Crataegus* taxa occurred, of which nine could be classified as forest associations in the class *Querco-Fagetea* and *Quercetea pubescenti-petraeae* and 25 were classified into various grassland, ruderal and scrubland associations included in the classes *Festuco-Brometea*, *Molinio-Arrhenatheretea*, *Rhamno-Prunetea*, *Trifolio-Geranietea sanguinei*, *Plantaginetea majoris* and *Artemisietea vulgaris* (Table 2, Fig. 6, Supplementary Tab. S3). *Crataegus monogyna* was present in 17 associations, which were mostly semi dry and mesic grasslands. In contrast, *C. rhipidophylla* was present only in 11 associations, mostly forests, but also some grassland associations. Their hybrid, *C.* ×*subsphaerica* occurred similarly in 23 grassland and forest associations. Furthermore, *C. laevigata* occurred in six mostly forest associations, while its hybrid with *C. monogyna*, *C.* ×*media*, was present in 11 grassland associations. *Crataegus* ×*macrocarpa* occurred only in three associations of which one was a type of grassland. *Crataegus lindmanii* occurred similarly in seven forest and grassland associations. Furthermore, *C.* ×*hadensis* was present in two grassland and two forest associations, while *C.* ×*kyrtostyla* was present in only two grassland associations. All taxa occurred frequently also in the ecotone between various forest and grassland associations.

Discussion

Distribution patterns of Crataegus species and their hybrids

The results revealed a clear geographic pattern in the distribution of some of the *Crataegus* taxa in the western part of Romania, with *C. laevigata* being common in forest and woodland ecosystems, and *C. lindmanii* restricted to forested landscapes in the hilly regions. However, it should be noted that while the samples did not include *C. laevigata* in the eastern sites, later investigations revealed that *C. laevigata* and *C. palmstruchii* were present in the western part of the Transylvanian Basin. However, their populations are highly scattered, and usually consist of a few isolated individuals. Literature on the distribution of *C. laevigata* indicate it reaches its eastern limit around the 24th longitude that crosses Romania (Christensen 1992, Thomas et al. 2021). *Crataegus monogyna* and *C. rhipidophylla* were more widespread species without any particular geographic pattern in their distribution in the area studied. In fact, both *C. monogyna* and *C. rhipidophylla* have a much wider geographical range than the other taxa identified in the area studied, with isolated populations reaching as far east as the Crimea and Caspian Sea (Christensen 1992, Fichtner & Wissemann 2021). Another important finding of the field study was that in the two regions studied with higher forest fragmentation (suboceanic fragmented and continental fragmented region), the more shade-tolerant species *C. laevigata*, *C. rhipidophylla* and *C. lindmanii* occurred in open habitats, thus the intermixing of *C. monogyna* and the shade-tolerant *Crataegus* populations extended beyond the forest-grassland ecotone, unlike in the case of the subsoceanic forested region with more extended and less fragmented forests. This confirms the importance of forest-grassland mosaics as a hybrid zone and a hotspot for hybrid formation. However, hybrids of *C. monogyna* and shade tolerant species were not restricted to these regions. Among the hybrids, only *C.* ×*subsphaerica* (hybrid between *C. monogyna* and *C. rhipidophylla*) was found to be widespread and common in each region studied, occurring in both grasslands, and at forest edges and in forests. In contrast, the other hybrid taxa *C.* ×*media*, *C.* ×*macrocarpa* and *C.* ×*hadensis*, were sporadic and *C.* ×*kyrtostyla* quite rare. *Crataegus* ×*subsphaerica* occurred even in the interior of forests, where one of the parental species, *C. monogyna*, was absent, or in extended grassland ecosystems with no forests in the surroundings, where the presence of the other parental species, *C. rhipidophylla*, could not be confirmed.

However, it should be noted that in the area studied, most specimens of *C.* ×*subsphaerica* and *C.* ×*media* were introgressive hybrids and are backcrosses with the *C. monogyna* parent, as confirmed in a previous study (Kuhn et al. 2021). Introgression in the case of these two hybrids is confirmed in other European studies as well as morphologically (Christensen 1992, Kerényi-Nagy 2015) and genetically (Fineschi et al. 2005). In contrast, *C.* ×*macrocarpa*, *C.* ×*hadensis* and *C.* ×*kyrtostyla* have a more intermediate combination of parental characters (Kuhn et al. 2021). Studies examining population structure and hybridization patterns in *Crataegus* taxa in western Europe, where *C. laevigata* is more abundant, indicate that hybrids of *C. laevigata* with other *Crataegus* species are also more abundant in western Europe than in Romania (Byatt 1975, 1976, Christensen 1982). However, the same studies also confirm that these hybrids are common outside the hybrid zone, where the *C. laevigata* and *C. rhipidophylla* parents are absent and introgressive forms backcrossed with *C. monogyna* predominate in open habitats (Byatt 1975, 1976, Christensen 1982, 1992). The study of Oklejewicz et al. (2013) in Poland confirmed a similar frequency pattern for *Crataegus* species as in Romania, with *C. laevigata* and *C. lindmanii* being rare, and *C. monogyna* and *C. rhipidophylla* more frequent. However, in contrast to the results presented here, in the south-eastern part of Poland, *C.* ×*macrocarpa* and *C.* ×*media* are much more common than their *C. laevigata* parent. This difference in the frequency of the two hybrids can be explained by the more fragmented landscape structure in Poland (Oklejewicz et al. 2013), compared e.g. to the Zărandului Mts study region in Romania, where *C. laevigata* was the most common.

Niche breadth

Based on the results presented, of the four *Crataegus* species included in the niche estimation, *C. laevigata* had the narrowest, while *C. monogyna* had the widest climatic niche, closely followed by *C. rhipidophylla*. This pattern can be explained by the predominantly western distribution of *C. laevigata* in the region studied, due to its suboceanic climatic requirements (Christensen 1992, Thomas et al. 2021). However, in the case of landscape configuration, *C. lindmanii* had the narrowest and *C. monogyna* the widest niche, suggesting that the distribution of *C. lindmanii* is determined by a more specific landscape configuration, and that considering the macroenvironment, *C. monogyna* is the species with the widest realized niche. Indeed, *C. monogyna* is considered to be the most common and widespread *Crataegus* species in Europe (Christensen 1992, Kerényi-Nagy 2015, Fichtner & Wissemann 2021). In contrast, while *C. lindmanii* is poorly studied, occurrence records suggest a narrower European distribution range and a more sporadic population distribution, compared to the other species of *Crataegus* studied (Christensen 1992, Kerényi-Nagy 2015). Similar to the findings presented, Oklejewicz et al. (2013) reveal that *C. rhipidophylla* and *C. monogyna* had the widest niches of the species studied. Furthermore, the results presented indicate that in general, hybrid *Crataegus* taxa do not have wider niches than their progenitors, with niche breadth estimates being intermediate or similar to that of the parental species. *Crataegus* ×*subsphaerica* was the only hybrid with a slightly wider niche breadth relative to both parental species in the case of landscape configuration and microenvironmental niches, probably due to the relatively small overlap in their niches compared to the other progenitor species. However, these values were only slightly higher than that of its *C. monogyna* parent. The fact that the taxa studied differed more in niche breadth defined based on macroenvironmental variables suggests a more similar degree of specialization in the case of the taxa of *Crataegus* studied in terms of the microenvironment compared to the macroenvironment. Furthermore, rarer *Crataegus* hybrids, such as *C.* ×*media*, *C.* ×*macrocarpa* and *C.* ×*hadensis* had a similarly wide niche in terms of microenvironmental constraints as the more frequent and widespread species or hybrids, *C. monogyna*, *C. rhipidophylla*, *C. laevigata* and *C.* ×*subsphaerica*. This indicates, that the smaller frequency of hybrids in the landscape is not the result of a lower tolerance of the hybrids of environmental stress. Furthermore, parental species are also relatively common, at least regionally, and the lack of encounter between populations should not limit interspecific gene transfer. Thus, it seems that frequency and geographic distribution of hybrids in the region studied is more likely to be a result of the likelihood of hybrid formation influenced by prezygotic barriers, rather than of environmental selection against hybrids, or the lack of parental population intermixing. Furthermore, there is no evidence that *Crataegus* hybrids in this study performed worse under certain environmental circumstances, or of a lower fitness than the parental species. Kuhn et al. (2022) indicate that seedlings of *C.* ×*subsphaerica* hybrids performance similarly under the same experimental light and moisture conditions as their parental species *C. monogyna* and *C. rhipidophylla*. Field observations also indicate that most hybrids have a similar stature to their parental species, are vigorous, have high fruit production and perform comparatively well in stressful environments, such as dry grasslands with shallow rocky soils or in deep shade in forests (Byatt 1975, Christensen 1992, Oklejewicz et al. 2013, Kuhn et al. 2021). While *C.* ×*kyrtostyla* and *C. palmstruchii* were not included in the analysis due to very low occurrence recorded in this study, occurrence data reported in other studies confirms their rarity in Europe and indicates a narrow distribution, thus a narrow climatic niche for both taxa (Christensen 1992, Kerényi-Nagy 2015). However, regarding landscape configuration and the microenvironment, further studies are needed to clarify ecological requirements of these taxa.

Niche overlap

Forest species overlapped more in terms of their niches than forest and grassland species in all the niche categories defined based on both macro- and microenvironmental variables. Furthermore, the niche overlap between the hybrid taxa and their parental species was in general greater than the overlap between the two parental species. However, hybrids usually overlapped with one of their parental species to a greater degree. Hybrids of *C. monogyna* overlapped more with *C. monogyna* than with the other parental species; in contrast, the hybrids of *C. laevigata* overlapped more with *C. rhipidophylla* or *C. lindmanii* than *C. laevigata* in the case of most of the niche categories. These results suggest that while hybrids are intermediate in environmental requirements relative to their parents, they tend to overlap more the niche of only one of their parents. This can be explained by asymmetric rehybridization and introgression towards one of the parental species. During the short period of overlap during flowering, in the case of protogynous species, such as *Crataegus*, the later flowering species has a greater chance of receiving pollen from the earlier flowering species, as the anthers do not release pollen when the stigma is receptive (Ruhsam et al. 2011, Pollegioni et al. 2013). Indeed, there is empirical evidence that in the case of the hybrid *C.* ×*subsphaerica*, the parent *C. monogyna* flowers later than *C. rhipidophylla*, for *C.* ×*media*, the parent *C. monogyna* flowers later than *C. laevigata*, while in the case of *C.* ×*macrocarpa*, the parent *C. rhipidophylla* flowers later than the parent *C. laevigata* (Fichtner & Wissemann 2021, Thomas et al. 2021, Kuhn & Ruprecht 2022).

Influence of single environmental factors on niche differentiation between *Crataegus taxa*

The results presented indicate that both micro- and macroenvironmental variables had a marked influence on the frequency and abundance of *Crataegus* taxa in the area studied. The optimal environmental conditions for *C. laevigata* were the highest values for climatic water balance in April, landscapes dominated by forests with little fragmentation and low light conditions due to high canopy cover. These results indicate that *C. laevigata* requires climatic conditions with oceanic influence. Distribution data published by Christensen (1992) and Thomas et al. (2021) confirm that *C. laevigata* is more widespread where oceanic and suboceanic climates prevail in central and western Europe. These studies also confirm that *C. laevigata* inhabits mostly old-growth, intact forests and woodlands. In contrast, *C. monogyna* seems to prefer a continental climate and tolerates landscape fragmentation, being equally common in open habitats and at forest edges, regardless of landscape structure (Christensen 1992, Fichtner & Wissemann 2021). Furthermore, *C. monogyna* almost never occurs under closed canopies, probably due to its sterility under low light conditions (Fichtner & Wissemann 2021). *Crataegus rhipidophylla* is more tolerant of the extremes of a continental climate than *C. laevigata*. Furthermore, *C. rhipidophylla* like *C. monogyna* occurs in a wide range of landscape configurations; however, contrary to *C. monogyna*, it tolerates higher canopy covers, thus deeper shade. *Crataegus lindmanii* seems to tolerate both oceanic and continental climates, whereas in terms of landscape configuration and light intensity, it is more associated with open habitats and woodlands than *C. rhipidophylla*. However, *C. lindmanii* still occurs predominantly in forest and woodlands like *C. rhipidophylla*. Indeed, occurrence data indicate, that *C. monogyna* is the most generalist in terms of its climatic requirements of the taxa of *Crataegus* studied, being a frequent species throughout Europe (Christensen 1992, Fichtner & Wissemann 2021). While hybrids were usually positioned between their two progenitors in terms of the environmental gradients investigated, their environmental optimum was shifted towards one of the progenitors. This was especially the case for hybrids between grassland and forest species (*C.* ×*subsphaerica* and *C.* ×*media*), which shifted towards the grassland species (*C. monogyna*). However, in the case of hybrids between two forest species, more complex patterns in hybrid environmental optimum relative to progenitors were observed. *Crataegus* ×*macrocarpa* inhabited climatically more similar areas to *C. laevigata*, and mosaic landscapes between the optimum for *C. laevigata*, *C. rhipidophylla* and *C.* ×*macrocarpa* had also intermediate light preferences relative to its parents. *Crataegus* ×*hadensis* was associated with environmental conditions more similar to its *C. lindmanii* parent. However, in the case of climatic water balance in August, *C.* ×*hadensis* could be positioned between the optimum of both parents. The habitat requirements of the hybrids investigated are mostly in accordance with the findings of other studies, which positioned hybrids in intermediate environments relative to parental species, while the penetration of hybrids into parent-like environments is also reported (Byatt 1975, 1976, Christensen 1982, 1992, Gosler 1990, Oklejewicz et al. 2013).

Phytosociological affiliation of Crataegus taxa

Our findings suggest a relatively weak phytosociological affiliation of the taxa of *Crataegus* studied. Thus, it appears that climate and local light availability are the primary determinants of the distribution of *Crataegus* taxa. The taxa studied seem to be less specialized in their affiliation to a certain plant association, and follow mostly major habitat categories, such as broad-leaved forests and dry or semi-dry grasslands. The diversity of plant associations in which a particular taxon has been found was also correlated with the frequency of particular taxa in the region studied, further confirming the low habitat specificity of *Crataegus* taxa within a major habitat category. These findings are partially confirmed by other, European-scale studies in the case of two well-known species, *C. monogyna* and *C. laevigata* (Fichtner & Wissemann 2021, Thomas et al. 2021). However, European-wide phytosociological affiliation of other species or hybrids of *Crataegus* is less clear due to the limited data. Nevertheless, Thomas et al. (2021) describe *C. laevigata* as an ancient woodland species and mention its occurrence in multiple high ranking syntaxonomical units, such as the class *Querco-Fagetea*, and the order *Quercetalia pubescentis*. Furthermore, Fichtner & Wissemann (2021) report *C. monogyna* as an open habitat generalist and associate this species also with woodlands. However, in case of woodlands, the phytosociological data on which their study is based, possibly contains misidentified *C. laevigata* or hybrids, or sterile individuals of *C. monogyna* that are difficult to identify.

Conclusions

Most hybrids of *Crataegus* were rare compared to their progenitors, however *C.* ×*subsphaerica* (the hybrid between *C. monogyna* and *C. rhipidophylla*) was more common than either of its progenitors. *Crataegus* ×*subsphaerica* was present even outside of the hybrid zone in geographic regions where *C. rhipidophylla* was absent, and was the most common *Crataegus* taxa throughout north-western Romania. In addition, it seems that the distribution of *Crataegus* taxa is determined primarily by climatic factors and light availability, and they are affiliated only to major habitat categories, such as broad-leaved forests and/or grasslands. Furthermore, most hybrids of *Crataegus* have narrower or similar niche breadths as their progenitors, while the frequency and range size of hybrid taxa do not correlate with their niche breadth, with rare hybrids like *C.* ×*hadensis*, *C.* ×*macrocarpa* and *C.* ×*media* having relatively wide niches. The lack of a linear relationship between relative niche breadth and hybrid frequency suggests that environmental selection is not important in preventing the spread of hybrids, while prezygotic hybridization barriers like phenological divergence of parental species and the probability of intermixing of parental populations is more decisive in restricting hybridization. Indeed, the most common hybrid recorded is between the two species with the greatest overlap in flowering time. There was a larger niche overlap between hybrid taxa and their progenitors than between the progenitors. However, the niche overlap with one of the progenitors was often greater. This can be explained by introgression directed towards one of the progenitor species, usually the later flowering one, as is the case for *C. monogyna* hybrids. In this case, there is an increased probability of pollen transfer towards the later flowering *C. monogyna* progenitor. Thus, the presumed intermediate ecological characters of the hybrids of *Crataegus* studied relative to their progenitors is not as evident as we hypothesized, probably due to introgression shift towards one of the progenitors. Furthermore, while significantly wider niches of hybrid taxa relative to their parents were not confirmed by the results presented, range expansion of *C.* ×*subsphaerica* and *C.* ×*media* hybrids was reflected in their distribution that confirmed the presence of these two hybrids in geographical regions, where at least one progenitor species is absent or very rare. Thus, considering the four hybrid zone models proposed by Arnold (1997), no single model describes the structure of the hybrid zone, hybrid fitness and distributions of progenitors and hybrids. However, depending on the hybridizing species, a combination of the 'mosaic model' and the 'evolutionary novelty model' provides a more reasonable explanation. The 'mosaic model' is supported by the reported mosaic-intermixing pattern of progenitor populations in the case of hybridizing forest and grassland species, where hybridization takes place in the ecotonal regions of the habitat mosaic, while the main evidence supporting the 'evolutionary novelty model' is the appearance of each of the five hybrids studied in novel habitats, outside of their hybrid zone (Arnold 1997). This expansion outside the hybrid zone is probably the result of the apomictic reproduction of these hybrids. Furthermore, the successful large-scale dispersal of hybrid diaspores by birds and mammals may also contributes to spatial expansion and an extensive distribution (Carlo et al. 2013, Fichtner & Wissemann 2021, Thomas et al. 2021).

Supplementary material

Fig. S1. Photographs of the regions studied.

Fig. S2. Correlations between environmental variables included in the modelling of niche breadth and overlap of species and hybrids of *Crataegus* in north-western Romania.

Fig. S3. Photographs of the taxa of *Crataegus* identified in the area studied.

Table S1. *Crataegus* taxa identified in the study area, their distribution, habitat preference and frequency.

Table S2. Conditional, constrained and unconstrained variance fractions explained by full and reduced pCCA models

Table S3. Characterization of the 10 vegetation clusters obtained using divisive hierarchical clustering and the gap statistical method.

Supplementary materials are available at www.preslia.cz

Acknowledgments

This research was funded by a Ph.D. grant from the Babeş-Bolyai University in Cluj Napoca. We are very grateful to Edina Gál and Boróka Jancsó for their help during the field investigation, and an anonymous reviewer for useful comments on a former version of this paper.

References

Abbott R. J. (2017) Plant speciation across environmental gradients and the occurrence and nature of hybrid zones: hybrid zones and plant speciation. – Journal of Systematics and Evolution 55: 238–258.

Arnold M. L. (1997) Natural hybridization and evolution. – Oxford University Press, New York.

Arnold M. L. & Hodges S. A. (1995) Are natural hybrids fit or unfit relative to their parents? – Trends in Ecology and Evolution 10: 67–71.

Barbu N., Georgescu F., Ştefanescu V. & Ştefan S. (2014) Large-scale mechanisms responsible for heat waves occurrence in Romania. – Romanian Journal of Physics 59: 1109–1126.

Beck H. E., Zimmermann N. E., McVicar T. R., Vergopolan N., Berg A. & Wood E. F. (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. – Scientific Data 5: 180214.

Blaine Marchant D., Soltis D. E. & Soltis P. S. (2016) Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. – New Phytologist 212: 708–718.

Borhidi A. (1995) Social behaviour types, the naturalness and relative ecological indicator values of the higher plants in the Hungarian Flora. – Acta Botanica Hungarica 39: 97-181.

Byatt J. I. (1975) Hybridization between *Crataegus monogyna* Jacq. and C. *laevigata* (Poiret) DC. in south-eastern England. – Watsonia 10: 253–264.

Byatt J. I. (1976) The structure of some *Crataegus* populations in north-eastern France and south-eastern Belgium. – Watsonia 11: 105–115.

Cai Q., Welk E., Ji C., Fang W., Sabatini F. M., Zhu J., Zhu J., Tang Z., Attorre F., Campos J. A., Čarni A., Chytrý M., Çoban S., Dengler J., Dolezal J., Field R., Frink J. P., Gholizadeh H., Indreica A., Jandt U., Karger D. N., Lenoir J., Peet R. K., Pielech R., De Sanctis M., Schrodt F., Svenning J., Tang C. Q., Tsiripidis I., Willner W., Yasuhiro K., Fang J. & Bruelheide H. (2021) The relationship between niche breadth and range size of beech (*Fagus*) species worldwide. – Journal of Biogeography 48: 1240–1253.

Carlo T. A., García D., Martínez D., Gleditsch J. M. & Morales J. M. (2013) Where do seeds go when they go far? Distance and directionality of avian seed dispersal in heterogeneous landscapes. – Ecology 94: 301–307.

Christensen K. I. (1982) A biometric study of some hybridizing *Crataegus* populations in Denmark. – Nordic Journal of Botany 2: 537–548.

Christensen K. I. (1992) Revision of *Crataegus* Sect. *Crataegus* and Nothosect. *Crataeguineae* (*Rosaceae-Maloideae*) in the Old World. – Systematic Botany Monographs 35: 1–199.

Coughlan J. M., Han S., Stefanović S. & Dickinson T. A. (2017) Widespread generalist clones are associated with range and niche expansion in allopolyploids of Pacific Northwest hawthorns (*Crataegus* L.). – Molecular Ecology 26: 5484–5499.

Curry C. M. (2015) An integrated framework for hybrid zone models. – Evolutionary Biology 42: 359–365.

Depypere L., Vander Mijnsbrugge K., De Cock K., Verschelde P., Quataert P., Van Slycken J. & Goetghebeur P. (2006) Indigenous species of *Crataegus* (*Rosaceae–Maloideae*) in Flanders (Belgium). An explorative morphometric study. – Belgian Journal of Botany 139: 139–152.

Dickinson T. A., Yan B. X., Han S. & Zarrei M. (2021) Niche shifts, hybridization, polyploidy and geographic parthenogenesis in Western North American hawthorns (*Crataegus* subg. *Sanguineae*, *Rosaceae*). – Agronomy 11: 2133.

Ellenberg H., Weber H. E., Dull R., Wirth V., Werner W. & Paulissen D. (1991) Zeigerwerte von Pflanzen in Mitteleuropas [Ecological indicator values of Central-European Plants]. – Scripta Geobotanica 18, Goltze Verlag, Göttingen.

Favre A. & Karrenberg S. (2011) Stress tolerance in closely related species and their first-generation hybrids: a case study of *Silene*. – Journal of Ecology 99: 1415–1423.

Fichtner A. & Wissemann V. (2021) Biological flora of the British Isles: *Crataegus monogyna*. – Journal of Ecology 109: 541–571.

Fick S. E. & Hijmans R. J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – International Journal of Climatology 37: 4302–4315.

Fineschi S., Salvini D., Turchini D., Pastorelli R. & Vendramin G. G. (2005) *Crataegus monogyna* Jacq. and *C. laevigata* (Poir.) DC. (*Rosaceae*, *Maloideae*) display low level of genetic diversity assessed by chloroplast markers. – Plant Systematics and Evolution 250: 187–196.

Gallego-Tévar B., Rubio-Casal A. E., de Cires A., Figueroa E., Grewell B. J. & Castillo J. M. (2018) Phenotypic plasticity of polyploid plant species promotes transgressive behaviour in their hybrids. – AoB PLANTS 10: ply055.

Gosler A. G. (1990) Introgressive hybridization between *Crataegus monogyna* Jacq. and *C. laevigata* (Poiret) DC. in the Upper Thames Valley, England. – Watsonia 18: 49–62.

Gramlich S. & Hörandl E. (2016) Fitness of natural willow hybrids in a pioneer mosaic hybrid zone. – Ecology and Evolution 6: 7645–7655.

Hesselbarth M. H. K., Sciaini M., With K. A., Wiegand K. & Nowosad J. (2019) landscapemetrics: an open-source R tool to calculate landscape metrics. – Ecography 42: 1648–1657.

Hijmans R. J., Cameron S. E., Parra J. L., Jones P. G. & Jarvis A. (2005) Very high resolution interpolated climate surfaces for global land areas. – International Journal of Climatology 25: 1965-1978.

Junker R. R., Kuppler J., Bathke A. C., Schreyer M. L. & Trutschnig W. (2016) Dynamic range boxes; a robust nonparametric approach to quantify size and overlap of n-dimensional hypervolumes. – Methods in Ecology and Evolution 7: 1503–1513.

Kassambara A. & Mundt F. (2020) factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.7. – URL: https://CRAN.R-project.org/package=factoextra.

Kerényi-Nagy V. (2015) A Kárpát-Pannon és Illir régió vadon termő galagonyáinak monográfiája [Monograph of wild hawthorns from the Carpathian, Pannon and Illyric regions]. – Szent István Egyetemi Kiadó, Gödöllo.

Kuhn T., Jancsó B. & Ruprecht E. (2021) Hawthorn (*Crataegus* L.) Taxa and their hybrids in North-Western Romania: a recommendation for national identification keys based on morphometric analysis. – Contribuţii Botanice 55: 7–26.

Kuhn T. & Ruprecht E. (2022) Flowering phenology may shape hybridization patterns of hawthorn (*Crataegus* L.) species. – Contribuţii Botanice 57: 95–107.

Kuhn T., Györfi O. & Ruprecht, E. (2022) Seedling performance, allocation patterns and phenotypic plasticity of two sympatric hawthorn species and their natural hybrid. – Flora 287: 151994.

Li Q., Grossenbacher D. L. & Angert A. L. (2018) The effect of range overlap on ecological niche divergence depends on spatial scale in monkeyflowers. – Evolution 72: 2100–2113.

Lopez-Alvarez D., Manzaneda A. J., Rey P. J., Giraldo P., Benavente E., Allainguillaume J., Mur L., Caicedo A. L., Hazen S. P., Breiman A., Ezrati S. & Catalan P. (2015) Environmental niche variation and evolutionary diversification of the *Brachypodium distachyon* grass complex species in their native circum-Mediterranean range. – American Journal of Botany 102: 1073–1088.

Maechler M., Rousseeuw P., Struyf A., Hubert M. & Hornik K. (2022) cluster: cluster analysis basics and extensions. R package version 2.1.4. – URL: https://svn.r-project.org/R-packages/trunk/cluster/.

Oklejewicz K., Chwastek E., Szewczyk M., Bobiec A. & Mitka J. (2013) Distribution of *Crataegus* (*Rosaceae*) in S-E Poland along a gradient of anthropogenic influence. – Polish Journal of Ecology 61: 683–691.

Oksanen J., Simpson G., Blanchet F., Kindt R., Legendre P., Minchin P., O'Hara R., Solymos P., Stevens M., Szoecs E., Wagner H., Barbour M., Bedward M., Bolker B., Borcard D., Carvalho G., Chirico M., De Caceres M., Durand S., Evangelista H., FitzJohn R., Friendly M., Furneaux B., Hannigan G., Hill M., Lahti L., McGlinn D., Ouellette M., Ribeiro Cunha E., Smith T., Stier A., Ter Braak C. & Weedon J. (2022) vegan: community ecology package. R package version 2.6–2. – URL: https://CRAN.R-project.org/package=vegan.

Pakgohar N., Eshaghi Rad J., Gholami G.H., Alijanpour A. & Roberts D. W. (2021) A comparative study of hard clustering algorithms for vegetation data. – Journal of Vegetation Science 32: e13042.

Parker K. C. (1991) Topography, substrate, and vegetation patterns in the northern Sonoran Desert. – Journal of Biogeography 18: 151–163.

Pollegioni P., Olimpieri I., Woeste K. E., De Simoni G., Gras M. & Malvolti M. E. (2013) Barriers to interspecific hybridization between *Juglans nigra* L. and *J. regia* L. species. – Tree Genetics & Genomes 91: 291–305.

R Core Team (2021) R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, URL: https://www.R-project.org/.

Ruhsam M., Hollingsworth P. M. & Ennos R. A. (2011) Early evolution in a hybrid swarm between outcrossing and selfing lineages in *Geum*. – Heredity 107: 246–255.

Sanda V., Biţă-Nicolae C. D. & Barabaş N. (2003) Flora cormofitelor spontane din România [The cormophyte flora of Romania]. – Editura Ion Borcea, Bacău.

Sanda V., Öllerer K. & Burescu P. (2008) Fitocenozele din România: sintaxonomie, structură, dinamică şi evoluţie [Plant communities of Romania: syntaxonomy, structure, dynamics and evolution]. – Ars Docendi, Bucureşti.

Senf C., Buras A., Zang C. S., Rammig A. & Seidl R. (2020) Excess forest mortality is consistently linked to drought across Europe. – Nature Communication 11: 6200.

Talent N. & Dickinson T. A. (2005) Polyploidy in *Crataegus* and *Mespilus* (*Rosaceae*, *Maloideae*): evolutionary inferences from flow cytometry of nuclear DNA amounts. – Canadian Journal of Botany 83: 1268–1304.

Thomas P. A., Leski T., La Porta N., Dering M. & Iszkuło G. (2021) Biological Flora of the British Isles: *Crataegus laevigata*. – Journal of Ecology 109: 572–596.

Vašková D. & Kolarčik V. (2019) Breeding systems in diploid and polyploid hawthorns (*Crataegus*): evidence from experimental pollinations of *C. monogyna*, *C. subsphaerica*, and natural hybrids. – Forests 10: 1059.

Wang H., McArthur E. D. & Freeman D. C. (1999) Narrow hybrid zone between two subspecies of big sagebrush (*Artemisia tridentata*: *Asteraceae*). IX. Elemental uptake and niche separation. – American Journal of Botany 86: 1099–1107.

Whitney K. D., Ahern J. R., Campbell L. G., Albert L. P. & King M. S. (2010) Patterns of hybridization in plants. – Perspectives in Plant Ecology, Evolution and Systematics 12: 175–182.

Šíře nik a jejich překryv u pseudogamních apomiktických hybridů rodu *Crataegus* a jejich rodičů v severozápadním Rumunsku

Hybridizace s následnou introgresí je u hlohů (rod *Crataegus*) dobře zdokumentována a přestože je známo mnoho široce rozšířených hybridů, faktory ovlivňující genový tok jsou stále nedostatečně prozkoumány. Porovnali jsme šíři a překryv nik u hybridů a jejich rodičů, abychom otestovali předpoklad, že relativně široká ekologická nika je výhodná pro šíření hybridů mimo hybridní zónu. Dále jsme zkoumali ekologické faktory, které potenciálně ovlivňují distribuci a frekvenci hybridů a jejich rodičů. Hlohy jsme studovali ve čtyřech oblastech severozápadního Rumunska, které se lišily v nadmořské výšce, hustotě zalesnění, fragmentaci stanovišť a míře antropogenního vlivu na krajinu. Pomocí ordinačních metod (CCA) jsme zkoumali vliv environmentálních faktorů na rozšíření taxonů rodu *Crataegus*. Fytocenologická příslušnosti studovaných taxonů byla hodnocena pomocí klastrové analýzy. Zjistili jsme deset taxonů rodu *Crataegus*, pět druhů a pět kříženců. Většina kříženců byla ve srovnání s jejich rodiči poměrně vzácná; výjimkou byl *Crataegus* ×*subsphaerica* (*C. monogyna* × *C. rhipidophylla*), nejčastější taxon ve zkoumané oblasti, přítomný i mimo hybridní zónu. Rozšíření studovaných taxonů primárně určují klimatické faktory, dostupnost světla a typ vegetace (listnaté lesy v nízkých nadmořských výškách, louky a křoviny). Výsledky také odhalily, že ekologická nika většina kříženců se podobá nikám rodičům a vzácní kříženci měli niky překvapivě široké. To naznačuje, že frekvence hybridů je ovlivněna spíše fenologickým překryvem mezi rodiči a účinným dálkovým šířením apomiktických semen než environmentální selekcí proti hybridům. Předpokládaná intermediární pozice hybridních taxonů mezi rodiči z hlediska ekologických preferencí nebyla tak zřejmá, jak by se dalo očekávat, pravděpodobně kvůli introgresi směrem k jednomu z rodičů. S ohledem na častý výskyt hybridů mimo hybridní zónu a mozaikovitý výskyt habitatů sympatrických druhů popisuje hybridní zónu zkoumaných druhů nejlépe kombinace mozaikového modelu a modelu evoluční novinky.

How to cite: Kuhn T. & Ruprecht E. (2023) Niche breadth and overlap of pseudogamous apomictic *Crataegus* hybrids and their progenitors in north-western Romania. – Preslia 95: 447–474.

Preslia, a journal of the Czech Botanical Society

www.preslia.cz