DIFFERENTIAL ROLE OF A PERSISTENT SEED BANK FOR GENETIC VARIATION IN EARLY VS. LATE SUCCESSIONAL STAGES

Benjamin Schulz1, Walter Durka2,3, Jiří Danihelka4,5, Rolf Lutz Eckstein6*

1 Institute of Landscape Ecology and Resource Management, Justus Liebig University Giessen, Giessen, Germany, 2 Department of Community Ecology (BZF), Helmholtz Centre for Environmental Research–UFZ, Halle (Saale), Germany, 3 German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany, 4 Department of Botany and Zoology, Masaryk University, Brno, Czech Republic, 5 Institute of Botany, The Czech Academy of Sciences, Průhonice, Czech Republic, 6 Department of Environmental and Life Sciences–Biology, Karlstad University, Karlstad, Sweden

* lutz.eckstein@kau.se

Abstract

Persistent seed banks are predicted to have an important impact on population genetic processes by increasing effective population size and storing past genetic diversity. Accordingly, persistent seed banks may buffer genetic effects of disturbance, fragmentation and/or selection. However, empirical studies surveying the relationship between aboveground and seed bank genetics under changing environments are scarce. Here, we compared genetic variation of aboveground and seed bank cohorts in 15 populations of the partially cleistogamous Viola elatior in two contrasting early and late successional habitats characterized by strong differences in light-availability and declining population size. Using AFLP markers, we found significantly higher aboveground than seed bank genetic diversity in early successional meadow but not in late successional woodland habitats. Moreover, individually, three of eight woodland populations even showed higher seed bank than aboveground diversity. Genetic differentiation among populations was very strong (ST = 0.8), but overall no significant differentiation could be detected between aboveground and seed bank cohorts. Small scale spatial genetic structure was generally pronounced but was much stronger in meadow (Sp-statistic: aboveground: 0.60, seed bank: 0.32) than in woodland habitats (aboveground: 0.11; seed bank: 0.03). Our findings indicate that relative seed bank diversity (i.e. compared to aboveground diversity) increases with ongoing succession and despite decreasing population size. As corroborated by markedly lower small-scale genetic structure in late successional habitats, we suggest that the observed changes in relative seed bank diversity are driven by an increase of outcrossing rates. Persistent seed banks in Viola elatior hence will counteract effects of drift and selection, and assure a higher chance for the species’ long-term persistence, particularly maintaining genetic variation in declining populations of late successional habitats and thus enhancing success rates of population recovery after disturbance events.
Introduction

Instead of relying solely on the spatial dispersal of their seeds, many plant species have developed the ability to disperse their offspring in time by accumulating long-lived, dormant seeds of several years in the soil or in aerial reservoirs. Such persistent seed banks are a common feature of plants to counteract the consequences of environmental or demographic stochasticity that can be found across a wide range of life history types, habitats and climate zones [1,2]. Particularly for rare species with disjunct populations or for species from highly dynamic or disturbed habitats, persistent seed banks may be crucial for population dynamics and stability, and potentially are capable of replacing standing individuals after bottlenecks or extinction events [3,4].

Theory predicts that, beyond affecting demography, seed banks could have an important impact on population genetic processes as they may consist of progeny produced by several and variable above-ground generations under varying selection regimes [5,6]. Consequently, persistent seed banks may increase effective population size [7,8] and store even more genetic diversity than present in the aboveground populations [5]. Thereby, they could buffer deleterious effects of random genetic drift and weaken patterns of genetic population structuring [5,9]. In addition, seed banks may enable gene flow from past generations stored in the soil, maintain genes in populations through periods in which they are selected against [6], and thus slow down adaptation processes and damp out directional selection in response to environmental fluctuations.

A number of studies have been carried out to empirically test the ecological and evolutionary impact of long-lived seed reservoirs. Some of these indicated that persistent seed banks might indeed increase effective population size, both in annuals [10,11] and perennials [12]. However, so far, it was impossible to conclusively confirm the potential of seed banks to accumulate genetic diversity and to serve as a genetic memory that might influence the evolutionary fate of populations [13]. Some studies found higher genetic diversity in seed banks [9,14], whereas most others showed higher diversity in standing plants (e.g. [6,15,16]) or no significant differences (e.g. [10–12,17,18]).

To shed some more light on this topic, Honnay et al. [19] conducted a meta-analysis using 13 studies that compared the genetic diversity of seed banks and aboveground populations. Interestingly, whereas levels of heterozygosity and percentage of polymorphic loci appeared to be similar in the two groups, allelic richness was significantly higher in the seed bank. Honnay et al. [19] concluded that differences in allelic richness are mainly driven by rare alleles and that selection might act as filter on seed bank alleles, preventing some of them to be established in standing plants. Moreover, the analysis showed significantly higher inbreeding and more homozygotes in the seed bank, substantiating the results of earlier studies, showing a gradual increase of heterozygosity towards the adult stage [16,20,21]. As discussed by Vitalis et al. [8], this pattern is most likely explained as an effect of selection that progressively eliminates less fit homozygotes in the course of seed germination and recruitment [16,22].

Overall, the study of Honnay et al. [19] thus gave no evidence that high levels of genetic diversity are accumulated in long-lived seed reservoirs. Instead, any difference in the genetic composition between seed bank and aboveground population may rather be the result of local selection than a buffering effect of stored seeds [8,19]. Consequently, the authors discouraged to continue surveying the genetic diversity of the two cohorts, unless this is performed under different selection regimes, in order to compare the outcome of the selection process [19]. However, until today the aspect of selection has been largely omitted in empirical research on seed bank genetics. To our knowledge, only three studies have yet compared seed bank and aboveground plants in contrasting environments and thus potentially under different selection
None of these showed clear differences in genetic diversity, neither between seed bank and aboveground plants nor among habitats. This might partially be attributed to comparatively small data sets and hence idiosyncratic characteristics of single populations.

To fill this gap in current knowledge, here we compared the genetic variation of seed bank and aboveground cohorts in the perennial flood plain species *Viola elatior* in two habitats representing the endpoints of a management-driven successional gradient, open grassland and alluvial forests fringes. The study species is a weak competitor for light [25]. Consequently, with increasing succession to closed forests, population sizes gradually decline and the species finally disappears from the aboveground vegetation [26]. Light availability is a strong environmental cue that influences an array of biotic and abiotic parameters (e.g., competition, water availability or temperature) and went along with epigenetic population differentiation in this species [27]. Therefore, the surveyed successional states appear well suited to study the outcome of local selection on the interplay of seed bank and aboveground genetics. Besides levels of standing genetic variation *per se*, also small-scale genetic structure may reveal patterns that give insight into spatio-temporal gene flow. We conducted a multi population study with 15 localities, covering two main distribution areas of *V. elatior* in Germany and the Czech Republic. Using amplified fragment length polymorphisms (AFLP) we asked the following questions: (1) Can the persistent seed bank of *V. elatior* maintain genetic diversity in different successional habitats with different selection regimes? (2) Do contrasting habitat types impact the small-scale spatial genetic structure of the seed bank and/or the aboveground plants?

**Materials and methods**

Permissions for field work and collection of plant tissues were issued by the Regional council Darmstadt (Regierungspräsidium Darmstadt) for the populations in the Rhine region (Germany) and by the South Moravian Regional Authority, Brno, Department of Environment (Krajský úřad Jihomoravského kraje, Brno, odbor životního prostředí) for populations in the Thaya/Morava region (Czech Republic).

**Study species**

*Viola elatior* (Violaceae) is a perennial iteroparous hemicryptophyte belonging to the section Viola, subsect. Rostratae [26]. The species’ distribution roughly covers the submeridional and temperate zone of western Eurasia ranging from the Parisian basin to southern Siberia and Central Asia [28]. In its core-area with summer-warm continental climates, *V. elatior* grows in steppe and forest-steppe vegetation, whereas in Central Europe, i.e., towards the western border of its distribution, the species is confined to large river corridors [26,29]. Here it becomes increasingly rare and occurs in different successional floodplain habitats ranging from managed open floodplain meadows to alluvial woodland fringes and gaps [26,30]. Populations are usually small, varying between tens and hundreds of individuals [26].

*Viola elatior* has an octoploid genome (2n = 40) and exhibits a mixed mating system with potentially cross-pollinated chasmogamous and obligatory self-pollinated cleistogamous flowers (for simplification hereafter just called “outcrossing” and “selfed” flowers, respectively). Whereas outcrossing flowers are only produced in the beginning of the growing season, i.e., May to early June, selfed flowers may develop from June to late October. Both flower types produce capsules with approximately equal numbers of seeds (~30). Nonetheless, seed production through selfed flowers strongly prevails, resulting in a very high overall selfing rate. In common garden experiments only around 4% of total capsule production resulted from outcrossing flowers ([30], Schulz, unpublished data).
As many other violets, *V. elatior* builds up persistent soil seed banks. Hözel & Otte [3] found maximum seed densities of up to 2660 germinable seeds/m² under a densely populated floodplain meadow, with more than 80% of all seeds in the upper 5 cm of the soil layer. Under strongly fluctuating conditions of floodplain habitats, the seed bank seems to be an important part of the species’ life strategy, which is illustrated by various reports about sudden emergence of plants in the course of disturbance events after long-term absence from the aboveground vegetation ([26] and reference therein).

**Study regions**

The study was conducted in two regions that represent strongholds of *Viola elatior* in Europe: The Upper Rhine floodplain south of the city Frankfurt am Main, Germany, and the Thaya/Morava floodplain around the town of Bréclav, the Czech Republic. In the densely populated and highly agriculturally influenced Upper Rhine area, the species occurs in nature reserves with a high share of open floodplain meadows that are regularly managed by mowing or grazing and thus provide high proportions of suitable early- and mid-successional habitats. In contrast, large parts of the Thaya/Morava floodplain are less influenced by settlements and intense land-use, and the landscape is characterized by a high share of alluvial forests and extensively managed patches of floodplain meadows. Here, populations of *V. elatior* are more widely scattered and mostly occur in late-successional habitats in gaps within forest stands or along forest fringes [31].

**Sampling design**

In both regions we surveyed stands (hereafter called populations) from each of the two extremes of the species’ environmental range, i.e. sunny floodplain meadows and shady alluvial woodland fringes. First, all known sites in the two regions inhabiting *V. elatior* were inspected and visually classified according to their light environment. The light environment of each population (Table 1) was measured indirectly with hemispherical photography as mean daily percentages of transmitted total photosynthetic active radiation (for details see [27]). Then, representative populations of both habitat types (Table 1) were selected and aboveground and seed bank samples were collected between May and June, i.e. after the spring germination peak and before the new seed rain.

Due to limited numbers of appropriate populations in the different habitat types with more than 20 aboveground individuals, we chose four meadow and three woodland sites in the Rhine region and three meadow and five woodland sites in the Thaya/Morava region (Table 1). Distances between populations ranged from 0.5 to 27 km in the Rhine region and from 1.6 to 70 km in the Thaya/Morava region.

To capture a maximum of allelic diversity and to detect potential spatial genetic structure within populations, we adopted a grid-based randomized sampling protocol. Therefore, in each population the presence of *Viola elatior* was mapped on a 1 m grid. For aboveground cohort sampling, we randomly selected 19–23 populated grid cells at each site and collected young, undamaged leaves from one plant per cell. Only for RM1, two to three individuals were sampled in 6 of the grid cells as the total number of populated grid cells did not exceed 14. Samples were immediately cooled to below 10°C, stored at -25°C and then freeze-dried for 48 h.

For seed bank cohort sampling, we applied the seedling emergence method. Using a soil corer of 5 cm in diameter and 4 cm in depth, we took five soil samples each in 30 randomly selected grid cells per population. The 150 soil samples represented 0.3 m² of soil surface and 11.8 l of soil volume. As previous studies showed that *V. elatior* seeds are strongly dormant during summer, with almost no germination between May and September [26], the soil cores were stored dry and dark until autumn. To reduce soil volume and optimize germination conditions,
we concentrated the soil samples by washing through two sieves with mesh sizes of 2.4 and 0.7 mm [32]. Seeds of *V. elatior* (diameter: 1.2–2.2 mm) thus accumulated in the middle fraction, whereas larger and smaller soil components were removed. Afterwards, the concentrated soil samples were spread in a ~0.5 cm thick layer on sterilized potting soil in 18 cm × 28 cm styrofoam trays. Due to logistic reasons, soil concentrates of two proximate grid cells were always pooled in one tray, respectively, resulting in 15 soil pools per population. For stratification, the trays were placed outside to expose them to natural conditions starting in late November. From the following spring to autumn the trays were watered regularly, and germinated *V. elatior* seedlings were identified and carefully transferred to individual pots once every month. The potted seedlings were grown in a greenhouse until they reached the four-leaf stage and were then harvested, stored at -25˚C and finally freeze-dried for 48 h. Depending on germination success, 12–23 seed bank individuals per population were chosen for genetic analyses. To obtain a comparable spatial sample resolution as for aboveground individuals, whenever possible only one to two samples per germination tray were selected (Table 1).

In summary, for each of the four region × habitat combinations we sampled between 3 and 5 populations, consisting of 19–23 aboveground cohort individuals and 12–23 seed bank cohort individuals, respectively (Table 1).

### Table 1. Overview of surveyed populations of *Viola elatior*.

<table>
<thead>
<tr>
<th>Population</th>
<th>Region</th>
<th>Latitude</th>
<th>Longitude</th>
<th>AG sample number</th>
<th>SB sample number</th>
<th>Populated grid cells (m²)</th>
<th>Habitat type</th>
<th>Mean transmitted PAR ± SD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM1</td>
<td>Upper Rhine</td>
<td>49°50'16&quot;N</td>
<td>8°24'00&quot;E</td>
<td>21</td>
<td>20 (11)</td>
<td>14</td>
<td>meadow</td>
<td>95.9 ± 0.2</td>
</tr>
<tr>
<td>RM2</td>
<td>Upper Rhine</td>
<td>49°50'01&quot;N</td>
<td>8°25'32&quot;E</td>
<td>19</td>
<td>20 (12)</td>
<td>-1400</td>
<td>meadow</td>
<td>87.7 ± 6.5</td>
</tr>
<tr>
<td>RM3</td>
<td>Upper Rhine</td>
<td>49°49'49&quot;N</td>
<td>8°28'03&quot;E</td>
<td>22</td>
<td>12 (8)</td>
<td>159</td>
<td>meadow</td>
<td>73.8 ± 8.3</td>
</tr>
<tr>
<td>RM4</td>
<td>Upper Rhine</td>
<td>49°36'08&quot;N</td>
<td>8°26'50&quot;E</td>
<td>22</td>
<td>19 (7)</td>
<td>170</td>
<td>meadow</td>
<td>80.5 ± 14.7</td>
</tr>
<tr>
<td>RW1</td>
<td>Upper Rhine</td>
<td>49°48'50&quot;N</td>
<td>8°24'57&quot;E</td>
<td>23</td>
<td>21 (10)</td>
<td>158</td>
<td>woodland</td>
<td>23.6 ± 14.2</td>
</tr>
<tr>
<td>RW2</td>
<td>Upper Rhine</td>
<td>49°35'49&quot;N</td>
<td>8°26'50&quot;E</td>
<td>20</td>
<td>20 (12)</td>
<td>31</td>
<td>woodland</td>
<td>12.5 ± 3.3</td>
</tr>
<tr>
<td>RW3</td>
<td>Upper Rhine</td>
<td>49°35'44&quot;N</td>
<td>8°25'55&quot;E</td>
<td>22</td>
<td>19 (10)</td>
<td>144</td>
<td>woodland</td>
<td>16.5 ± 9.4</td>
</tr>
<tr>
<td>TM1</td>
<td>Thaya/Morava</td>
<td>48°45'30&quot;N</td>
<td>16°51'57&quot;E</td>
<td>21</td>
<td>23 (14)</td>
<td>32</td>
<td>meadow</td>
<td>48.3 ± 21.2</td>
</tr>
<tr>
<td>TM2</td>
<td>Thaya/Morava</td>
<td>48°46'52&quot;N</td>
<td>16°51'48&quot;E</td>
<td>21</td>
<td>23 (13)</td>
<td>-1200</td>
<td>meadow</td>
<td>83.0 ± 4.7</td>
</tr>
<tr>
<td>TM3</td>
<td>Thaya/Morava</td>
<td>48°48'51&quot;N</td>
<td>16°49'53&quot;E</td>
<td>21</td>
<td>22 (14)</td>
<td>160</td>
<td>meadow</td>
<td>60.1 ± 14.9</td>
</tr>
<tr>
<td>TW1</td>
<td>Thaya/Morava</td>
<td>48°49'00&quot;N</td>
<td>16°27'08&quot;E</td>
<td>23</td>
<td>23 (14)</td>
<td>131</td>
<td>woodland</td>
<td>22.4 ± 11.7</td>
</tr>
<tr>
<td>TW2</td>
<td>Thaya/Morava</td>
<td>48°49'25&quot;N</td>
<td>16°46'26&quot;E</td>
<td>23</td>
<td>23 (14)</td>
<td>59</td>
<td>woodland</td>
<td>25.3 ± 20.2</td>
</tr>
<tr>
<td>TW3</td>
<td>Thaya/Morava</td>
<td>48°49'01&quot;N</td>
<td>16°47'41&quot;E</td>
<td>23</td>
<td>22 (13)</td>
<td>34</td>
<td>woodland</td>
<td>14.8 ± 7.4</td>
</tr>
<tr>
<td>TW4</td>
<td>Thaya/Morava</td>
<td>48°38'21&quot;N</td>
<td>16°57'19&quot;E</td>
<td>22</td>
<td>21 (12)</td>
<td>56</td>
<td>woodland</td>
<td>13.1 ± 3.2</td>
</tr>
<tr>
<td>TW5</td>
<td>Thaya/Morava</td>
<td>48°58'30&quot;N</td>
<td>17°23'09&quot;E</td>
<td>21</td>
<td>23 (15)</td>
<td>59</td>
<td>woodland</td>
<td>30.6 ± 13.4</td>
</tr>
</tbody>
</table>

* Numbers in brackets indicate germination trays that contributed with one or more emerged seedlings to the SB samples, respectively; Ger—Germany; Cz—Czech Republic; AG—aboveground; SB—seed bank; PAR—photosynthetic active radiation.

https://doi.org/10.1371/journal.pone.0209840.t001
AFLP genotyping

We investigated a total of 324 aboveground and 311 seed bank samples with amplified fragment length polymorphism (AFLP). Total genomic DNA was extracted from dried leaf tissue using the DNeasy 96 Plant extraction kit (QIAGEN). AFLP methodology followed Kloss et al. [33] and is described in detail in S1 Text. After an initial screening of 64 primer pairs, eight selective primer combinations (S1 Table) were chosen for AFLP analyses. Separation and visualization of fragments was done on an ABI 3130 capillary sequencer (Applied Biosystems, Foster City, USA) with Genescan 500(-250) LIZ internal size standard (Applied Biosystems). GENMAPPER version 3.7 (Applied Biosystems) was used to analyze the AFLP profiles. We binned fragments manually for all samples in one batch using a peak height threshold of 10 rfu. Then, peak height data were exported and for each fragment a specific peak height threshold was manually determined based on the peak height distribution which allowed scoring presence (1) and absence (0) of fragments. All loci that showed a monomorphic pattern or a deviation in only one individual were excluded from the data set to prevent biased parameter estimation. Overall error rate was 0.6%, based on 58 replicate samples (9%) that were repeated starting with DNA extraction.

Data analysis

We analyzed the binary AFLP data using a band- or marker-based strategy, i.e. without calculating allele frequencies. We estimated genetic diversity as band richness ($Br$) and as percentage of polymorphic loci at the 5% level ($PLP$) using AFLPDIV 1.1 [34]. To account for the unequal sample size of aboveground and seed bank individuals (Table 1), a rarefaction-based approach was employed with a standardized sample size equal to the smallest sample population (i.e. $n = 12$). To evaluate the number of bands within populations that are private to either aboveground or seed bank cohorts within each population, we calculated private band richness ($PBr$) with rarefaction analyses according to Kalinowski [35] separately for each population using ADZE 1.0 [36] and applying the same standardized sample size of $n = 12$. Finally, we estimated the proportion of unique AFLP-phenotypes $P_u = (eMLG - 1)/(N_{min} - 1)$ for each cohort using the R package poppr [37], where $eMLG$ is the expected number of multilocus genotypes based on rarefaction to the smallest sample size $N_{min}$. Statistical analyses comprised linear mixed models (LMM) testing the effects of the fixed factors region, habitat and cohort, and the interaction between habitat and cohort, while population served as random factor. Visual checks of model residuals led us to logit-transform $Br$, $PLP$ and $PBr$, whereas $P_u$ was not transformed before final analyses. Statistical analyses were done using the package lme4 and analysis of deviance tables with Wald II chi² tests using the package car in R version 3.3.3 [38]. Additionally, to improve comparability between populations and to address the relative diversity of seed bank in comparison to aboveground cohorts across regions, for all diversity parameters differences between cohorts were depicted using the natural-logarithmic response ratios (LnRR) as proposed by Goldberg and Scheiner [39]: $LnRR = Ln\left(\frac{P_{SB}}{P_{AG}}\right)$, where $P_{SB}$ is the mean value of seed bank cohorts and $P_{AG}$ is the mean value of aboveground cohorts. Differences between cohorts were considered significant when the 95% confidence interval did not overlap with zero.

We performed non-hierarchical and hierarchical analyses of molecular variance (AMOVA) using ARLEQUIN 3.5.1.2 [40] to quantify genetic variation among populations ($\phi_{ST}$), among groups of populations ($\phi_{CT}$) and among populations within groups ($\phi_{SC}$). To test for differentiation between aboveground and seed bank within each population, pairwise $\phi_{ST}$ values were calculated among cohorts. Significance levels were determined after 9999 permutations.
Clustering of individual samples was examined with principal component analysis (PCA) using the R package ADEGENET v1.4–2 [41].

To examine the small-scale spatial genetic structure of aboveground and seed bank cohorts within and between habitats we used spatial autocorrelation methods implemented in SPA-GeDi v1.4 [42]. We chose distance limits of 4, 8, 12, 16 and 20 m to assure a sufficient number of individual pairs per distance class. For seed bank cohorts, only samples originating from soil pools with a maximum distance of 4 m between the two soil sampling sites (see seed bank sampling strategy above) were considered, using the corresponding midpoint coordinates for the analyses, respectively. Thus, potential biases from the actual coordinates of the soil samples lie within the chosen distance classes, respectively. To construct spatial autocorrelograms, pairwise kinship coefficients ($F_{ij}$) for dominant markers [43] were calculated assuming an inbreeding coefficient of 0.5. Using higher inbreeding coefficients of up to 0.9 in additional trials had little effect on the results and did not change the general conclusions. Significance of mean $F_{ij}$ per distance class was tested with 9999 permutations of multilocus genotypes. We quantified small-scale spatial genetic structure for each population using restricted regression analyses (0–20 m) and calculating the $Sp$ statistic, representing the rate of decrease in pairwise kinship with distance [44]. $Sp$ was estimated as $-b_{log}/(1-F_{(1)})$, where $b_{log}$ is the regression slope of mean $F_{ij}$ on log geographic distance and $F_{(1)}$ is the mean $F_{ij}$ of the first distance class. To compare autocorrelation patterns, we furthermore pooled populations according to habitats and regions and tested for heterogeneous autocorrelation with heterogeneity tests for multiple populations subsets [45] using GenAlEx 6.5 [46] and applying the same distance classes as above. Number of permutations and bootstraps were set to 9999, respectively. Following Banks and Peakall [47] significance of the Heterogeneity Test is declared when $p<0.01$.

Results

Seedling emergence rates

Total numbers of seedlings that emerged from the concentrated soil cores strongly varied between populations (S2 Table) ranging from 12 in RM3 (40 seedlings/m$^2$) to 206 in TW2 (687.7 seedlings/m$^2$). Overall, mean values of germinated seedlings differed significantly between regions (Rhine = 24.9±9.4 and Thaya/Morava = 104.8±48.5; t-test, $p = 0.003$) but not between habitat types (meadow = 45.1±34.7 and woodland = 87±59.4; t-test, $p = 0.15$).

Genetic diversity

AFLP analysis resulted in a total of 528 scorable loci of which 128 (24%) were polymorphic which were used for all further calculations. Across the 635 surveyed samples we found 323 unique AFLP phenotypes. After rarefaction, measures of genetic diversity across populations (Table 2) revealed mean values of 1.17 (aboveground) and 1.15 (seed bank) for band richness ($Br$), 20% and 18% for percentage polymorphic loci ($PLP$), 0.06 and 0.04 for private band richness ($PBr$), and 0.71 and 0.58 for proportion of unique genotypes ($Pu$). For $PBr$ (LMM, $\chi^2 = 4.40$, df = 1, $p = 0.036$) and $Pu$ (LMM, $\chi^2 = 15.77$, df = 1, $p<0.001$) the values were significantly higher for aboveground cohorts (Fig 1). Additionally, for $Pu$ there was a significant interaction between habitat and cohort (LMM, $\chi^2 = 7.17$, df = 1, $p = 0.007$). The latter was furthermore corroborated by the comparison of log response ratios, indicating for $Br$, $PBR$ and $Pu$ significantly higher values for aboveground cohorts in meadow but not in woodland habitats (Fig 2). Individually, three populations had higher genetic diversity values for seed bank than aboveground cohorts for most or all diversity descriptors, notably all in woodland habitats (RW2, TW3, TW5; Table 2).
Genetic structure

Analyses of molecular variance resulted in global $\phi_{ST}$ values of 0.80 and 0.83 for aboveground and seed bank cohorts (Table 3), respectively, ranging for population pairwise $\phi_{ST}$ between 0.24 and 0.96 for aboveground cohorts and between 0.28 and 0.99 for seed bank cohorts (S3 Table). Hierarchical AMOVA showed that for both cohorts around 16% of genetic variance resided between regions, while most variation (65.8 and 68.7%, respectively) was partitioned among populations within regions (Table 3). Overall, we found very strong genetic differentiation among populations but no significant differentiation between aboveground and seed bank cohorts (S4 Table). Nonetheless, at the individual population level, 2 of 7 meadow populations and 4 of 8 woodland populations exhibited significant genetic differentiation between cohorts with pairwise $\phi_{ST}$ values ranging from 0.05 for TW1 to 0.16 for TW4 (S3 Table).

Principal component analysis corroborated the AMOVA results, revealing a very close clustering of cohort pairs for most populations (Fig 3). Overall, the first three components accounted for 20.4%, 11.4% and 9.3% of genetic variation. Regions were separated along the first axis, while there was no consistent structuring according to habitat types.

Small-scale spatial genetic structure

For aboveground and seed bank samples each 5 of 15 populations showed significant small-scale spatial genetic structure. $Sp$ values ranged from 0.186 to 2.198 for aboveground and from 0.073 to 0.253 for seed bank cohorts (Table 4). Although $Sp$ values were lower for seed bank
samples, the presence of small-scale spatial genetic structure was comparable among the two cohorts, and in all but one case, populations showing a significant spatial structure at the aboveground level also showed it at the seed bank level. Overall, meadow populations displayed a stronger degree of spatial genetic structure with markedly higher average $Sp$ values than woodland populations (aboveground: 0.60 vs. 0.11; seed bank: 0.32 vs. 0.03). This was
confirmed by the heterogeneity test with pooled data sets, revealing for both cohorts significantly stronger small-scale spatial genetic structure in meadow than in woodland populations, both overall (Fig 4) and at the regional scale (S1 Fig). Moreover, heterogeneity tests confirmed that small-scale spatial genetic structure was not significantly different between aboveground and seed bank (Fig 4). However, for meadow populations kinship coefficients of aboveground and seed bank samples showed significant differences in the second distance class.

Discussion

Genetic diversity

Genetic diversity of *Viola elatior* was comparably low [48,49]. This is consistent with earlier studies on this species [27,31] and with findings in other plants with predominant seed production through selfed cleistogamous flowers [50,51]. These typically show high levels of

![Figure 2](https://doi.org/10.1371/journal.pone.0209840.g002)

**Fig 2.** Mean (±95% CI) ln response ratio of Br, PLP, PBr, and Pu between seed bank and aboveground cohorts in meadow (M) and in woodland (W) habitats, respectively. Br—band richness, PLP—percentage of polymorphic loci at the 5% level; PBr—pairwise private band richness (aboveground vs. seed bank); Pu—proportion of unique AFLP-phenotypes. Negative ln response ratios denote higher values in the aboveground cohort. Differences between aboveground and seed bank cohorts were considered significant when 95% CI did not overlap with zero.

Table 3. Summary of hierarchical AMOVA results for aboveground and seed bank cohorts of the surveyed *Viola elatior* populations.

<table>
<thead>
<tr>
<th>Source</th>
<th>cohort</th>
<th>V</th>
<th>% total</th>
<th>P</th>
<th>( \phi ) statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among all populations</td>
<td>AG</td>
<td>14.62</td>
<td>80.06</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
<tr>
<td>Among all populations</td>
<td>SB</td>
<td>15.18</td>
<td>83.47</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
<tr>
<td>Among regions</td>
<td>AG</td>
<td>3.09</td>
<td>15.69</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
<tr>
<td>Among populations within regions</td>
<td>AG</td>
<td>12.97</td>
<td>65.83</td>
<td>&lt;0.001</td>
<td>( \phi_{SC} = 0.78 )</td>
</tr>
<tr>
<td>Within populations</td>
<td>AG</td>
<td>3.64</td>
<td>18.47</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
<tr>
<td>Among regions</td>
<td>SB</td>
<td>3.16</td>
<td>16.03</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
<tr>
<td>Among populations within regions</td>
<td>SB</td>
<td>13.53</td>
<td>68.71</td>
<td>&lt;0.001</td>
<td>( \phi_{SC} = 0.82 )</td>
</tr>
<tr>
<td>Within populations</td>
<td>SB</td>
<td>3.01</td>
<td>15.27</td>
<td>&lt;0.001</td>
<td>( \phi_{CT} = 0.16 )</td>
</tr>
</tbody>
</table>

AG—aboveground individuals; SB—seed bank individuals; V—variance components

https://doi.org/10.1371/journal.pone.0209840.t003
inbreeding, little or no genetic variability within populations and strong population differentiation [52].

It is widely assumed that decreasing population sizes lead to a loss of genetic variation through effects of increased random genetic drift, higher inbreeding rates and the accumulation of deleterious mutations (e.g. [53,54]). Furthermore, under changing environmental conditions during succession the loss of genetic variation might be aggravated by an increased probability of local extinction of certain genotypes due to selection [55]. Consequently, we expected to find a decrease of aboveground genetic diversity from meadow to woodland habitats, while seed bank genetic diversity should stay rather constant, given that the seed bank stores genotypes that are lost aboveground. The seed bank reservoir of genotypes might then help to regain former aboveground genetic diversity under more favorable conditions (e.g. after disturbance).
However, we did not find any clear difference of aboveground diversity (Fig 1) between habitats and only cohorts appeared to be significantly different for some of the diversity measures, showing higher aboveground than seed bank values for $P_{Br}$ and $P_{u}$. Overall, these differences can mainly be attributed to changes in meadow habitats. This is substantiated by the significant interaction between habitat and cohort for $P_{u}$, suggesting that the relationship of the two cohorts changes with ongoing succession. Indeed, using the log-response ratio approach to highlight the relative seed bank genetic diversity (compared to the aboveground cohort) corroborated the results of the linear model for $P_{u}$. For all diversity measures but PLP, values of meadow populations were significantly higher in aboveground than seed bank cohorts, whereas in woodlands the two cohorts did not differ. Our data thus imply that relative seed bank genetic diversity increases towards the later successional stage.

Generally, higher aboveground than seed bank genetic diversity, as found in meadow populations, might be explained by selection against homozygotes or inbred individuals during germination and recruitment as suggested by earlier seed bank genetic studies [8,16,19]. However, as we did not find reduced seed bank genetic diversity in late successional woodland stages, in which similar or even stronger selection pressure can be expected, we alternatively suggest that in woodland habitats an increase in the relative contribution of outcrossed seeds to the seed bank may be causal for the observed pattern, counteracting effects of selection and reduced population size. Temporal changes in seed bank genetic variation have already been observed for an aerial seed bank [56], presumably due to increased outcrossing rate.

In $V.\ elatior$, several linked processes may have fostered outcrossing in late successional habitats. First, a change in the balance between outcrossed and selfed flowers might have led to higher outcrossing rates in woodland sites. In partially cleistogamous species, the allocation to outcrossing and selfed flowers is controlled by environmental factors [57] and outcrossing often increases with decreasing plant density [58], decreasing light availability [58] and

---

**Table 4. Small-scale genetic structure for aboveground and seed bank individuals in populations of Viola elatior.**

<table>
<thead>
<tr>
<th>Population</th>
<th>n</th>
<th>$F_{(1)}$</th>
<th>$b_{log}$</th>
<th>$Sp$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM1</td>
<td>21</td>
<td>0.000</td>
<td>-0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>RM2</td>
<td>19</td>
<td>0.102</td>
<td>-0.079</td>
<td>0.088</td>
</tr>
<tr>
<td>RM3</td>
<td>22</td>
<td>0.678**</td>
<td>-0.707**</td>
<td>2.198</td>
</tr>
<tr>
<td>RM4</td>
<td>22</td>
<td>0.391**</td>
<td>-0.428**</td>
<td>0.703</td>
</tr>
<tr>
<td>WR1</td>
<td>23</td>
<td>0.329**</td>
<td>-0.291**</td>
<td>0.433</td>
</tr>
<tr>
<td>WR2</td>
<td>20</td>
<td>0.016</td>
<td>-0.021</td>
<td>0.021</td>
</tr>
<tr>
<td>WR3</td>
<td>22</td>
<td>0.147</td>
<td>0.358*</td>
<td>0.186</td>
</tr>
<tr>
<td>TM1</td>
<td>21</td>
<td>-0.021</td>
<td>0.007</td>
<td>0.001</td>
</tr>
<tr>
<td>TM2</td>
<td>21</td>
<td>0.041</td>
<td>-0.134</td>
<td>0.140</td>
</tr>
<tr>
<td>TM3</td>
<td>21</td>
<td>0.515**</td>
<td>-0.487**</td>
<td>1.003</td>
</tr>
<tr>
<td>TW1</td>
<td>23</td>
<td>-0.008</td>
<td>-0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>TW2</td>
<td>23</td>
<td>-0.001</td>
<td>-0.042</td>
<td>0.042</td>
</tr>
<tr>
<td>TW3</td>
<td>23</td>
<td>0.005</td>
<td>-0.065</td>
<td>0.066</td>
</tr>
<tr>
<td>TW4</td>
<td>22</td>
<td>0.014</td>
<td>-0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>TW5</td>
<td>21</td>
<td>0.160*</td>
<td>0.030</td>
<td>0.079</td>
</tr>
<tr>
<td>Average</td>
<td>21.6</td>
<td>0.158</td>
<td>-0.166</td>
<td>0.332</td>
</tr>
</tbody>
</table>

AG—aboveground individuals; SB—seed bank individuals; n—sample number; $F_{(1)}$—kinship coefficient of the first distance class; $b_{log}$—regression slope of spatial genetic autocorrelation; $Sp$—statistic. Significant values are presented in bold (*p<0.05, **p<0.01).

https://doi.org/10.1371/journal.pone.0209840.t004
Fig 4. Correlograms of spatial genetic autocorrelation in populations of Viola elatior. Comparison of correlogram homogeneity is shown for (A) aboveground populations grouped for habitats, and for aboveground and seed bank populations from either only (B) meadow or (C) woodland habitats; ω-test indicates overall significance (A: ω = 73.74, p = 0.0001; B: ω = 9.33, p = 0.497; C: ω = 19.85, p = 0.030). * p < 0.05 and ** p < 0.01 indicate significant differences for single distance classes. AG—aboveground; SB—seed bank.
increasing soil water availability [59], i.e. the same environmental differences that are present between meadow and woodland habitats. For the congeneric Viola mirabilis increasing outcrossing rates were indeed shown to be correlated with shading [60]. Second, also variation in abortion rates or quality of outcrossed seeds might be important. As in V. elatior capsules from outcrossed flowers mature only in the warmest period of the year, differences in water availability between meadow and woodland habitats are particularly distinct during the development of outcrossed seeds. This is corroborated by a survey in four populations from the Upper Rhine region (Schulz, unpublished data) indicating for seeds from outcrossing flowers higher abortion rates (30% vs. 6%) as well as lower seed mass (10 mg vs. 18 mg) in meadow than in woodland sites. In contrast, in the same populations differences for seeds from selfed flowers were less pronounced (seed abortion: 4.8% vs. 2.2%; seed mass: 15 mg vs. 17 mg). Additionally, seed longevity might be higher in woodland habitats, as soil parameters like high moisture content and constant temperatures are beneficial for seed survival. Thus, woodland seed banks potentially could be assembled from more seed generations, leading to higher relative genetic diversity. Finally, also anthropogenic effects may have an impact on the relative contribution of seeds from outcrossing flowers. As most of the floodplain meadows are regularly managed through mowing in early June, large amounts of ripening capsules from outcrossing flowers will be destroyed every year, whereas capsules from selfed flowers can freely develop from July to October, likely leading to a lower proportion of outcrossed seeds in meadow compared to woodland seed banks.

In accordance with Eckstein et al. [31], we found no significant differences in genetic diversity between German and Czech populations, neither for aboveground nor seed bank cohorts. This strongly implies that the detected habitat related differences seem to present a general effect that is independent of geographic location.

**Genetic structure**

Genetic differentiation in Viola elatior was very high, with 80.1% and 83.5% of genetic variation residing among populations for aboveground and seed bank cohorts, respectively. Similar differentiation (up to 82%) has been also reported in earlier studies on this species [27,31,61], reflecting the predominant selfing breeding system and relatively small population sizes and hence a strong influence of genetic drift on population structure. This is further substantiated by a lack of correlation between genetic differentiation and geographic distances in populations from the Upper Rhine Valley [27], which indicates that gene flow is much too low to counteract population divergence by genetic drift. Overall, population differentiation was virtually identical among seed bank and aboveground cohorts at all hierarchical levels (Table 3). In contrast, some other studies reported lower differentiation among seed bank cohorts than among aboveground cohorts [6,9,62], indicating that aboveground populations can become more differentiated than the homogeneous seed pools they derived from [9]. Accordingly, in V. elatior habitat related change in selection was not as high as expected and differentiation between cohorts is an exception that occurred only in few individual populations.

**Small-scale spatial genetic structure**

Spatial genetic structure within populations arises due to spatially restricted gene dispersal and is mainly related to the amount of gene flow by seeds and pollen [63]. Therefore, the observed significant decrease of small-scale spatial genetic structure from meadow to woodland habitats strongly supports the assumption that in Viola elatior the relative impact of outcrossing and thus pollen dispersal increases towards late successional stages. High selfing rates in meadow populations are furthermore corroborated by high Sp values that even exceeded mean values
reported for other predominantly selfing species ($Sp = 0.14$, [44]). In contrast, $Sp$ values in woodland populations were markedly lower and do match better with values for mixed mating species ($Sp = 0.04$, [44]). Thus, if in meadow populations almost all seeds result from self-fertilization, only seed dispersal (average distance ~1.3 m, [26]) does impact gene dispersal distances. In contrast, in woodland populations and assuming considerable outcrossing, both seed and pollen dispersal contribute to, and increase, gene dispersal [44]. Both, aboveground and seed bank cohorts showed significant small-scale spatial genetic structure and even though $Sp$ values were generally lower in seed bank cohorts, overall there were no significant differences between the two groups. Similar to our results, other studies also detected significant small-scale spatial genetic structure in both cohorts [15,17,64] suggesting that the spatial structure of aboveground and seed bank individuals is often mutually dependent ([17], but see [16]).

Taken together, the observed small-scale spatial genetic structure in *Viola elatior* likely reflects the proposed differences in outcrossing rates between habitats. However, generally spatial genetic structure appears to break down relatively fast with ongoing succession and is not stored over longer periods of time in seed bank cohorts. Otherwise much stronger spatial autocorrelation would have been present in the seed bank of woodland populations. Instead, the high correlation of aboveground and seed bank cohorts in both habitats suggests that turnover of seeds in soil and of adult plants is similar in *V. elatior* and that seeds do not persist in the seed bank for more than few adult generations [6].

**Conclusions**

To our knowledge, the present study is the first one that has compared aboveground and seed bank genetic diversity between successional stages. Surveying a relative large number of populations of the partially cleistogamous *Viola elatior* in two different regions, we could show that the contribution of outcrossing to reproduction seems to increase from early to late successional stages, leading to higher relative seed bank diversity and lower within population small-scale spatial genetic structure. This suggests that under favorable early successional conditions with high plant densities, populations maintain their approved genotypes mainly by selfing. Under less favorable conditions an increase in outcrossing may keep genetic diversity in *V. elatior* at a constant level and hence seems to compensate the detrimental effects of small population size. Ultimately, in late successional habitats with a higher risk of extinction, the resulting increased relative seed bank genetic diversity also assures a higher chance for population recovery and thus the long-term persistence of the species.

We suggest that the relationship between seed bank and above ground plants potentially can be driven by both, post-germination selection as seen in meadow habitats and genetic buffering through stored seeds that seems to counteract the effects of drift and selection as seen in woodland populations. However, to substantiate these results and to test if this is not a specific situation in partially cleistogamous plants, further studies surveying seed bank genetics along environmental gradients in species with other mating systems are needed.

**Supporting information**

S1 Text. AFLP protocol. (DOCX)

S1 Table. Adaptor- and primer sequences used for AFLP analyses. (DOCX)
S2 Table. Seedling emergence rates of *Viola elatior* in soil samples from the surveyed populations.

(DOCX)

S3 Table. Pairwise $\Phi_{ST}$-matrix for aboveground and seed bank cohorts of the surveyed populations.

(DOCX)

S4 Table. Summary of hierarchical AMOVA results for German and Czech populations grouped by cohort type.

(DOCX)

S1 Fig. Correlograms of genetic autocorrelation in populations of *Viola elatior*. Comparison of correlogram homogeneity is shown for (A) Czech and (B) German aboveground populations grouped for habitats, respectively; $\omega$-test indicates overall significance (A: $\omega = 67.18, p = 0.0001$; B: $\omega = 33.91, p = 0.0003$). *p*<0.05 and **p**<0.01 indicate significant differences for single distance classes. Cz—Czech Republic; Ger—Germany; AG—aboveground.

(TIF)

Acknowledgments

We thank Josef Scholz vom Hofe and Lena Föhr for help during field work, and Martina Herrmann and Ina Geier for assistance during laboratory work.

Author Contributions

Conceptualization: Walter Durka, Rolf Lutz Eckstein.

Formal analysis: Benjamin Schulz, Walter Durka, Rolf Lutz Eckstein.

Funding acquisition: Walter Durka, Rolf Lutz Eckstein.

Investigation: Benjamin Schulz, Jiří Danihelka.

Methodology: Benjamin Schulz, Walter Durka.

Project administration: Rolf Lutz Eckstein.

Resources: Walter Durka.

Supervision: Walter Durka, Rolf Lutz Eckstein.

Writing – original draft: Benjamin Schulz, Walter Durka, Jiří Danihelka, Rolf Lutz Eckstein.

Writing – review & editing: Benjamin Schulz, Walter Durka, Jiří Danihelka, Rolf Lutz Eckstein.

References


25. Moora M. Responses of a rare (Viola elatior) and a common (Viola mirabilis) congenic species to different management conditions in grassland—is different light competition ability responsible for different abundances? Acta Oecol. 2003; 24: 169–174.


