

Garden, greenhouse or climate chamber? Experimental conditions influence whether genetic differences are phenotypically expressed

Pascal Karitter^{1*}, Martí March-Salas¹, Andreas Ensslin², Robert Rauschkolb^{3,4}, Sandrine Godefroid⁵, Hendrik Poorter^{6,7}, Johannes F. Scheepens¹

¹Plant Evolutionary Ecology, Institute of Ecology, Evolution and Diversity, Goethe University Frankfurt, Max-von-Laue-Str. 13, 60438 Frankfurt am Main, Germany

²Conservatory and Botanic Garden of the City of Geneva, Chemin de l'Impératrice 16 1, 1296 Chambésy, Geneva, Switzerland

³Institute of Ecology and Evolution with Herbarium Haussknecht and Botanical Garden, Department of Plant Biodiversity, Friedrich Schiller University Jena, Germany

⁴German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

⁵Meise Botanic Garden, Nieuwelaan 38, 1860 Meise, Belgium

⁶Horticulture and Product Physiology, Wageningen University, Droevendaalsesteeg 1, 6708PB Wageningen, the Netherlands

⁷ Department of Biological Natural Sciences, Macquarie University, North Ryde, NSW 2109, Australia

* Corresponding author: Pascal Karitter (p.karitter@gmail.com)

Conflict of interest

The authors have no conflict of interest to declare.

Author contribution

PK, MMS and JFS designed the study. SG and RR provided the seed material. PK conducted the experiment. PK, MMS and JFS analysed the data. PK wrote the first draft and all the authors contributed to further versions of the manuscript.

Abstract

1. Common-environment experiments are important to study genetically-based phenotypic variation within and among plant populations. Such experiments can be performed in an experimental garden, greenhouse or climate chamber. However, phenotypic expression may be strongly affected by the environmental conditions and influenced by parental and storage effects. Hence, it is unclear if results from common-environment experiments are reproducible across multiple experimental setups.
2. In this study, we assessed the effects of three different growth facilities – outdoor garden, greenhouse, and climate chamber –, on phenotypic expression. We compared ancestral and descendant genotypes of the same population of *Leontodon hispidus*. We also evaluated differences in phenotypic expression between plants grown after one (F1) vs. two (F2) intermediate generations.
3. We observed strong differences among plants growing in different growth facilities. Furthermore, we found that descendants had larger rosettes than ancestors only in the greenhouse and they flowered later than ancestors exclusively in the climate chamber. We did not find significant differences between intermediate generations within the growth facilities.
4. Overall, our study demonstrates that environmental variation among growth facilities can dictate the presence and magnitude of phenotypic differences. This implies that absence of evidence for phenotypic differences is not evidence of absence. Experimental systems should be carefully designed to provide meaningful conditions related to the research question. Finally, growing a second intermediate generation did not impact the genetic differences of ancestors and descendants within the facilities, supporting that only one

53 intermediate generation may be sufficient to reduce detectable parental and
54 storage effects.

55

56 **Keywords:** Controlled experimental conditions, common garden, experimental setup,
57 refresher generation, resurrection study

Introduction

Common-environment experiments have been a powerful approach for evolutionary biologists ever since the Swedish botanist Göte Turesson established this approach one century ago. The ability of a given genotype to express different phenotypes depending on the environment (i.e., phenotypic plasticity; Sultan, 2000) complicates studies on the genetic basis of phenotypic differences, since field observations cannot be used to unravel genetic from plastic effects on the phenotype. However, by growing plants of the same species from different habitats in the same developmental environment, Turesson (1922) was able to attribute phenotypic variation among plant populations to genetic differences and identify natural selection as the main driver. This concept was further developed to study local adaptation by transplanting individuals originating from different populations reciprocally among the habitats from which they were sampled (Clausen et al., 1940; Kawecki and Ebert, 2004). Common-environment experiments can even be used to study rapid evolution of plant populations to recent environmental changes by experimentally applying different selection pressures over multiple generations (Hill and Caballero, 1992) or by using the resurrection approach, in which seeds collected before a potential selection pressure are revived and compared to plants from recently collected of the same population (Franks et al., 2018). Comparisons of these phenotypes in a common environment can then uncover evolutionary changes that occurred between samplings (Franks et al., 2007).

It is important that the results of experimental approaches are reproducible and robust to ensure their validation and generalization, but this can be challenging (Drummond, 2009). Plant responses can be affected by the choice of the growth facility (i.e., outdoor garden, greenhouse or climate chamber; Poorter et al., 2016) or non-genetic biases such as parental effects (Latzel et al., 2023). For instance, Massonnet

and colleagues (2010) studied leaf growth variables and other traits of three *Arabidopsis thaliana* genotypes in 10 different laboratories and found that modest variations in growing conditions such as temperature, light quality and the handling of the plants can induce significant differences in molecular profiles and phenotypes. Consequently, relatively small differences in the growth facility can lead to significant differences in plant responses and results can be strongly affected by the choice of the common environmental conditions. Poorter and colleagues (2012) described experimental outdoor gardens as being relatively close to natural conditions in the field. They characterized outdoor gardens as commonly having low spatial heterogeneity but high temporal variations in temperature, light and water supply, high chance of plant damage (hail, herbivory, late frost) and episodes of extreme conditions (e.g., high irradiance causing high temperatures, and low precipitation). According to Poorter and colleagues (2012), greenhouses normally provide more buffered conditions with heating systems to protect against frost, adjustable shading screens to counteract high irradiance, and with a more controlled water supply. However, air temperatures can peak depending on the ventilation system, and the greenhouse can substantially shade the plants due to structural elements (Cabrera-Bosquet et al., 2016). In climate chambers, researchers can control the experimental conditions as reliable as possible, which would be optimal for reproducibility, but the conditions can be quite artificial and deviate strongly from field conditions with more heterogeneous light and humidity distributions. Especially light can have strong vertical gradients and heterogeneous horizontal gradients (Poorter et al., 2012). Furthermore, climate chambers provide limited space which can greatly affect the possible sample size for larger species and thus, have a trade-off between control over the environment and statistical power. These differences among growth facilities may be important, as certain environmental conditions may not elicit phenotypic variation, even though genetic variation is present

among genotypes. Accordingly, Stanton and colleagues (2000) found that, compared to near-optimal growing conditions, more stressful conditions, which better mimic natural conditions, tend to increase phenotypic variation among genotypes of wild mustard (*Sinapis arvensis*). Consequently, it is unclear if the results from common-environment experiments are consistent throughout different experimental environments and can be accurately reproduced. If not, the consequence would be that we may miss evidence of relevant genetic variation or, in contrast, that we may detect genetic variation that has currently no adaptive relevance in the field.

Studies on genetic differentiation can also be confounded by non-genetic or random variability induced by parental effects or storage. Parental effects occur, when the parental phenotype affects the phenotype of their offspring irrespective of the genes that are inherited (Badyaev and Uller, 2009; Auge et al., 2017). Seed provisioning is one major parental effect that can have a big impact on the offspring, because the resource availability and general environmental conditions (e.g., light quantity, quality and duration) of the mother plant can determine the amount of resources for the seeds and consequently affect seedling establishment and early life history traits (Herman and Sultan, 2011). Other parental effects include hormone-driven effects on physiology of the seedling or epigenetic processes through passing on distinct DNA methylations or chromatin changes (Blödner et al., 2007; Jablonka and Raz, 2009; Elwell et al., 2011; Herman and Sultan, 2011; Richards et al., 2017). Even though parental effects can have ecological and evolutionary significance (Latzel et al., 2023), their influence can be a source of bias in studies on genetic differentiation when parental environmental conditions differ among genotypes. Furthermore, long storage periods can affect seed viability and even post-emergence traits (Franks et al., 2018). In experiments where plants from the same population but largely different generation

are compared, so-called 'resurrection experiments', storage duration and conditions may well have been different. Consequently, phenotypes of seedlings may express varying plastic responses (Weis, 2018). These biases can be minimized by acclimating the experimental plants under common environmental conditions for one or more generations before the start of the experiment (Kawecki and Ebert, 2004). Parental effects caused by the environment usually disappear after one generation in a new environment (Agrawal, 2002; Gianoli, 2002), but have also been shown to persist over multiple generations (Wulff et al., 1999). Latzel (2015) recommends at least two intermediate generations before the start of an experiment, as it increases the chance of evening out epigenetic modifications. Still, the method of growing intermediate generations is not always implemented in common-environment studies as it is time- and labor-intensive, especially if the study plants do not flower in the first year (Bischoff and Müller-Schärer, 2010; Rauschkolb et al., 2023).

In this study, we investigated the consistency of phenotypic differences of a resurrection common-environment experiment among different growth facilities. Furthermore, we tested for differences between results from experimental plants grown after one vs. two intermediate (i.e., refresher) generations. Absence of any differences would suggest that one intermediate generation is sufficient to reduce detectable parental, epigenetic or storage effects. We worked with the perennial herb *Leontodon hispidus* and used seeds of ancestors sampled in 1995 and of descendants sampled in 2018 collected from the same population in a calcareous grassland in Belgium. We grew two intermediate generations (F1 and F2) from both ancestors and descendants under common conditions and cultivated them in three different growth facilities: in an outdoor garden, in a greenhouse and in a climate chamber. We measured eight traits regarding growth, leaf anatomy and flowering phenology to cover a wide spectrum of

functional traits. We hypothesized that (1) genetically-based phenotypic differences across traits tend to occur inconsistently across substantially different growth facilities. (2) We expected that phenotypic differences in evolved traits between ancestors and descendants would be strongest in less-controlled conditions such as the outdoor garden and weakest in the more optimal and constant conditions in the climate chamber. Finally, we hypothesized that (3) one intermediate generation is not enough to sufficiently reduce non-genetic differences between ancestors and descendants, which may be attributed to the storage or parental effects of ancestors.

Material and Methods

Study species

Leontodon hispidus (Asteraceae) is a perennial herbaceous herb and typically flowers from June to October (Kühn and Klotz, 2002). It is self-incompatible and pollinated by insects (Kühn and Klotz, 2002). It is widespread throughout Europe and commonly found in calcareous grasslands. Seed material was collected from two temporal origins, 1995 (ancestors) and 2018 (descendants), from a single population in a dry calcareous grassland in a Belgian nature reserve (50°47'35"N, 5°40'25"E). The distance to the nearest neighboring population is approximately 2 km, likely preventing the majority of gene flow into the population. The staff of the Meise Botanic Garden (Belgium) collected the ancestral seeds for conservation purposes and efforts were made to represent the genetic diversity of the population by collecting from as many individuals as possible dispersed throughout the population. The seed material from an unknown number of mother plants was cleaned, bulked, dried at 15% relative humidity, and stored at -20 °C at the seed repository of the Meise Botanic Garden. In the summer of 2018, we revisited the population and collected seeds from 20 mother plants. These seeds were cleaned, bulked and then stored at 4 °C. Rauschkolb et al. (2022a)

analyzed genomic relatedness and allelic richness among individuals within both temporal origins (ancestors and descendants) and found similar levels of relatedness without obvious kinship structure, which supports the comparability of the sampling procedures and confirms that sufficient seeds were collected.

Experimental design

Both ancestral and descendant seeds were grown for two consecutive intermediate generations (F1 and F2). For the first intermediate generation (F1), we sowed 300 seeds from each temporal origin and selected 15 random individuals from each temporal origin which were randomly pollinated by hand in net cages to prevent unintentional cross-pollination (Rauschkolb et al., 2022b). We used the seeds from the F1 intermediate generation for the F2 generation and grew them under similar conditions, this time using bumblebees (Natupol Seeds, Koppert GmbH, Straelen, Germany) as pollinators. Ultimately, seven maternal lines from the F1 intermediate generation and 8 maternal lines from the F2 intermediate generation yielded sufficient seed material for both temporal origins. For each maternal line, we used 12 seedlings grown individually in black 1.5 L pots that were prepared as follows: In July 2022, we placed 12 pots for each maternal line with cultivation soil (Spezial Substrat Typ T1b, Hawita GmbH, Vechta, Germany) in the greenhouse and sowed three seeds into each pot. All pots were watered three times a week to maximum soil capacity.

After the seedlings had developed their first true leaf, we thinned them to a single individual per pot, with the seedling moved to the center of the pot. We measured the initial size as rosette diameter and divided the pots randomly into three groups with four individuals from each maternal line. Each group was grown in a different growth facility for the rest of the experiment: outdoor garden, greenhouse or climate chamber. In total, we used 360 plants for this experiment (3 growth facilities × 2 temporal origins

× 2 generations × 7/8 maternal lines × 4 replicates). Pots in the garden were placed on gravel 2 m below a shading cloth to reduce radiation and temperature stress (Schattiergewebe 45%, Nitsch GmbH, Kreuztal, Germany). Plants were randomized every two weeks and watered weekly to soil capacity for the duration of the experiment. In the greenhouse, the plants were placed 1 m below lamps with a combination of two fluorescent tubes (Lumilux HO 80W-865, Berlin, Germany and Gro-Lux FH 80W, Sylvania, Erlangen, Germany). The lamps were programmed to switch on between 6 am and 20 pm (14 h photoperiod) whenever the natural light intensity went below 360 $\mu\text{mol m}^{-2} \text{s}^{-1}$ outside. To avoid extreme temperatures, sliding shutters and lamps were programmed activate once light intensity surpassed 1100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ outside. The climate chamber (ThermoTec GmbH, Weilburg, Germany) was set to 14h-10h day-night cycle with 21°C during the day and 18°C during night to simulate the start of the growing season. Air humidity was set to a constant 60% and the plants were placed 1 meter below halogen lamps (Radium HRI-BT 400W/D Pro Daylight, Lampenwerk GmbH, Wipperfürth, Germany).

In each growth facility, we inserted four temperature and soil moisture loggers (TMS-4 logger, Tomst, Prague, Czechia) in the center of black 1.5 L pots with the same cultivation soil and positioned them randomly among the pots with plants. The loggers monitored soil temperature at 5 cm depth, soil surface temperature, air temperature at 5 cm above the soil as well as soil moisture every 15 min. We used this data to calculate mean values of all four loggers from each environment and to derive daily mean values for all parameters (Fig. 1, Appendix S1; see Supplemental Data with this article). Furthermore, we calculated mean values for soil temperature, soil surface temperature and air temperature over the course of the whole experimental period (Table 1). On average, temperatures were intermediate in the garden with 21.2 – 22.6 °C, highest in

the greenhouse with 23.9 – 26.1 °C, and lowest in the climate chamber with 20.4 – 21.2 °C (Table 1).

At the start of the experiment at noon, we measured the light intensity using a light meter (Panlux electronic 2, GMC-Instruments, Nürnberg, Germany) in the outdoor garden and greenhouse at 12 spatially distributed pots at the same height 2 cm above the soil. After the end of the experiment, hourly solar radiation data (Wm^{-2}) was extracted for the whole experimental period from a weather station in the outdoor garden (iMetos 1, Pessl Instruments GmbH, Weiz, Austria). To assess, how much radiation the plants received in the outdoor garden, we multiplied the solar radiation data by 0.55 to account for the shading cloth (45 % shading) and converted the data into Photosynthetic Photon Flux Density (PPFD, $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) using the default function *Rg.to.PPFD()* from the *bigleaf* package (Knauer et al., 2018) in R (version 4.0.3, R Core Team, 2020). Fraction of incoming solar irradiance to photosynthetically active radiation (PAR) was set to 0.5 and the conversion factor was set to 4.6. According to the light intensity measurements at the start of the experiment, the greenhouse received 18.5 % of the light that the outdoor garden plants received. We used this ratio to approximate how much PPFD the plants in the greenhouse received over the course of the experiment by multiplying the PPFD data of the outdoor garden by 0.185. We summed all values per day to calculate the daily light integral (DLI) for each day and then calculated the mean DLI for the whole experimental period. For the climate chamber, we directly measured PPFD at 12 spatially distributed spots at pot height using a PAR sensor (PAR Special sensor SKP 210, Skye Instruments Ltd, Powys, UK). Since the photoperiod was constant in the climate chamber (14h), we multiplied the mean PPFD measurement ($297.5 \mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) by the photoperiod (in seconds) to calculate the DLI. The average DLI over the course of the experiment

(Table 1) was highest in the garden ($40.5 \text{ mol} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$), lowest in the greenhouse ($7.5 \text{ mol} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$) and intermediate in the climate chamber ($15.0 \text{ mol} \cdot \text{m}^{-2} \cdot \text{d}^{-1}$).

Measurements

During the experiment, we recorded onset of flowering three times per week, because it is an essential trait in the life history of a plant and has often been observed to evolve under changing environmental conditions (Pelletier et al., 2009; Franks, 2015). We regarded a plant as flowering when the first anther was visible. Plants started flowering in August and after three months, most plants had flowered and we measured the rosette diameter as a measure of growth and ability to capture light. We also recorded the number of flowering stems as a proxy for reproductive success. For each individual, we measured the chlorophyll content of four randomly selected healthy and fully developed leaves in SPAD units using a chlorophyll meter (SPAD-502 Plus, Konica Minolta, Neu-Isenburg, Germany). We measured leaf area of three randomly selected healthy and fully developed leaves per plant with the smartphone application “easy leaf area free” (Easlon and Bloom, 2014). These leaves were dried in a dry oven at 60°C for three days and then weighed together at a fine scale (CPA225D-0CE, $e = 1 \text{ mg}$, Sartorius AG, Göttingen, Germany). At the same day, we harvested flower heads and flowering stems as reproductive biomass and leaves as vegetative biomass and dried these using the same procedure. This allows us to detect whether plants invest their resources into vegetative growth or rather into reproductive structures. In order to investigate responses in leaf anatomy, we calculated specific leaf area (SLA) by dividing the combined leaf area of the three selected leaves by their dry weight and calculated the leaf dry matter content (LDMC) by dividing the dry weight by the fresh weight. The weight of the three selected leaves was added to the vegetative biomass.

Data analysis

All statistical analyses were performed using R (version 4.0.3, R Core Team, 2020). We used linear mixed effects models implemented in the *lme4* package (Bates et al., 2015) with temporal origin, generation, and environment as fixed factor, as well as their two- and three-way interactions. Maternal line was included as random factor and initial size as covariate. Rosette diameter, vegetative biomass, SLA, LDMC, onset of flowering, reproductive biomass, number of flowering stems and the SPAD measurements were used as response variables, and we applied appropriate transformations to these variables when necessary to improve normality and heteroscedasticity of model residuals (Table 2). All linear models were analysed using the function *Anova()*, and *P* values (Appendix S2) were adjusted for multiple testing with the method of False Discovery Rate (Benjamini and Hochberg, 1995) using the *fuzzySim* package (Barbosa, 2015). Tukey post-hoc tests were applied using the *emmeans* package (Lenth, 2021) whenever an explanatory factor with more than two levels was significant.

Results

The growth facility significantly affected all measured traits (Table 2). In the outside garden, plants had the lowest rosette diameter, and SLA, while those traits were highest in the greenhouse and intermediate in the climate chamber (Fig. 2AC). In contrast, vegetative biomass and LDMC were highest in the climate chamber followed by the garden, and lowest in the greenhouse (Fig. 2BD). A similar pattern can be observed in the onset of flowering (Fig. 3A): plants in the greenhouse and garden flowered at a similar time, but the onset of flowering of plants in the climate chamber was approximately 4 days delayed. The reproductive biomass, the number of flowering

stems and SPAD values (Fig. 3BCD) were highest in the garden, intermediate in the climate chamber and markedly low in the greenhouse.

The temporal origin significantly affected vegetative biomass, LDMC, and onset of flowering and we found significant interactions with the growth facility (Origin $\times$ Facility) in rosette diameter and onset of flowering (Table 2). Descendants had generally 13 % higher vegetative biomass (Appendix S3A) and 8 % higher LDMC (Appendix S3B) compared to ancestors, irrespective of the growth facility or intermediate generation. Regarding the rosette diameter however, descendants had 11 % higher rosette diameter (2.1 cm) than ancestors in the greenhouse, while values were similar in the garden and in the climate chamber (Fig. 4A). Descendants also flowered 2.6 days later in the climate chamber than ancestors but flowered at similar times in the garden and greenhouse (Fig. 4B).

The number of intermediate generations did not significantly affect any measured trait consistently across growth facilities, but we found significant interactions with the growth facility (Gen $\times$ Facility) in rosette diameter and SPAD values (Table 1). Although we did not find significant differences between the F1 and F2 generations within growth facilities for rosette diameter and SPAD values using Tukey post-hoc tests, the F2 had 11 % larger rosettes than the F1 both in the greenhouse and the climate chamber, while the F1 and F2 in the garden were similar (Fig. 4C). Concerning SPAD, the F1 generation showed 3 – 6 % larger values than F2 in the garden and climate chamber, while F1 and F2 did not differ in the greenhouse (Fig. 4D). We found no significant interactions between temporal origin and intermediate generation (Origin $\times$ Gen) and no significant three-way-interactions (Origin $\times$ Gen $\times$ Facility) in the measured traits (Table 2).

Discussion

We studied the effects of three growth facilities and one vs two intermediate generations on the phenotypic expression of ancestor and descendant genotypes of a single population of *Leontodon hispidus*. We found very strong phenotypic differences among the three growth facilities. Furthermore, we found significant temporal origin × facility interactions in two traits, indicating that the choice of the growth facility can affect detectability of phenotypic differences. Finally, we did not find differences between intermediate generations within the growth facilities, suggesting that there is no need for multiple intermediate generations to sufficiently reduce parental and storage effects for this species.

Differences in growth facilities

Plant responses in the different growth facilities varied greatly from each other and the observed patterns were trait specific. The outdoor garden proved to be the experimental environment where plants were most successful in their reproductive ability, as they had the most flowering stems and the highest reproductive biomass. Plants in the garden also had low SLA and high LDMC, which was probably caused by the high light availability present at that growth facilities (Table 1). It has been well established that light irradiance correlates negatively with SLA and positively with LDMC (Anten, 2005; Poorter et al., 2019). Plants in the garden also had the highest chlorophyll content (SPAD values) which could be the reason for the greater reproductive success. Previous studies also showed that leaf traits can strongly respond to water availability, with increasing dryness leading to decreasing SLA and increasing LDMC (Volaire, 2008; Poorter et al., 2009; Jung et al., 2014; Wellstein et al., 2017; Vitra et al., 2019). Accordingly, pots in the garden had the lowest soil moisture content (Fig. 1B) and high variability in soil moisture, which was caused by

the exposure to natural rain events, high temperature fluctuations (Fig. 1A), and potentially higher evaporation due to wind exposure.

In the greenhouse, plants grew a large rosette diameter, had high SLA, and low LDMC. This response is in line with a strategy to increase surface area to improve light capture in low-light environments (Poorter et al., 2019). Indeed, the greenhouse, where we expected more favourable conditions than the outside garden, had the lowest light availability (Table 2) due to shading by greenhouse structural elements, but also very high temperatures (Fig. 1A). These high temperatures are also likely to contribute to the high SLA, as they facilitate cell expansion and thus, reduce leaf density and number of cell layers (Atkin et al., 1996; Poorter et al., 2009). the number of flowering stems, reproductive biomass and SPAD values were very low, indicating that even with the adjustments in leaf anatomy, the plants in the greenhouse had the lowest performance and were least successful.

Plants growing in the climate chamber had intermediate rosette diameter, SLA, SPAD values, and reproductive traits, which correlates well with the intermediate light availability (Table 1). Interestingly, plants produced the most vegetative biomass in the climate chamber out of all growth facilities. The very controlled and stable conditions might support fast vegetative growth in the climate chamber without compromising reproductive investment.

The comparisons showed significant differences among growth facilities. The main drivers of these patterns are most likely the differences in light availability and temperature, but also the variability of environmental factors could have a significant impact (Hamann et al., 2021). We expected that the garden would be the least favourable environment due to higher temperatures and stronger environmental fluctuations, leading to decreased growth and fitness of plants. However, our results

indicate the opposite, which is probably due to the much higher light availability in outdoor gardens in summer compared to the greenhouse and climate chamber. Also, the garden conditions are much closer to natural conditions to which plants from the field are expected to be adapted (Lascoux et al., 2016).

Reproducibility among growth facilities

Although it can be expected that different growth facilities cause plants to differ in their overall performance, it may also be assumed that origin and treatment effects would show qualitatively similar results across environments. Under this assumption, if a common-environment experiment would be performed in a single environment – as in most studies – the expected patterns would also be observed irrespective of this environment. Two alternative scenarios are, however, possible. First, origin or treatment effects may be observed only under specific environmental conditions and not in others. This would imply that an experiment may not always reveal the expected patterns. Second, specific origin or treatment effects may be observed under the chosen experimental conditions, but contrasting origin or treatment effects could have been observed when alternative experimental conditions would have been chosen. A consequence would be that contrasting conclusions could have been drawn, depending on the experimental environment.

In our study, descendants consistently had higher vegetative biomass and higher LDMC compared to ancestors. These results are in line with the previous findings that this population of *L. hispidus* evolved faster growth in recent decades (Karitter et al., in press), which were observed in the greenhouse in autumn. Furthermore, high LDMC has been shown to increase drought survival chances (Bongers et al., 2017; De La Riva et al., 2017). LDMC correlates well with strong cell

walls and may be beneficial to maintain turgor under drought conditions (Monson and Smith, 1982). Therefore, high LDMC could have evolved in this population through selection by increasing drought events caused by climate change (IPCC, 2018). The phenotypic differences between ancestors and descendants of these two traits were consistent throughout the experimental environments, since we were able to reproduce them in the outdoor garden, greenhouse and in the climate chamber. In contrast, we found interactions of temporal origin with the experimental environment for rosette diameter and onset of flowering. Descendants had a larger rosette diameter in the greenhouse compared to ancestors, but temporal origins did not differ in any other environment. Given that the most prominent distinction of greenhouse was its low light irradiance, increasing rosette diameter may be a good strategy to increase the surface area of the leaves to capture more light. This interaction may be explained by evolution under more shaded conditions, which could have been caused by increased competition during the recent decades exacerbated by climate change (Parmesan and Hanley, 2015). At the collection site, we observed that *L. hispidus* naturally competes with grasses such as *Brachypodium pinnatum* and *Bromus erectus*, which can substantially shade this rosette species. Combined with high nutrient depositions from the atmosphere (Newman, 1995; Galloway et al., 2008; Bobbink et al., 2010) and surrounding agriculture, *L. hispidus* might have faced strong selection pressure through high competition and may have adapted its ability to plastically respond to increasingly shaded conditions. Hence, the response to high shading of the descendants was not triggered in the other environments, but this explanation needs further testing by additional experiments that include shading treatments.

We also observed later onset of flowering of descendants compared to ancestors exclusively in the climate chamber. Plants flowered generally later in the climate chamber compared to the other environments, but descendants delayed their

onset of flowering substantially more than ancestors. In another resurrection study on the same population conducted in a greenhouse, descendants also flowered later than ancestors (Rauschkolb, et al., 2022), which was explained by the introduction of sheep grazing in 2007, forcing plants to flower later to escape the grazing pressure. We could not reproduce this result in the greenhouse used in this experiment and also not in the outside garden. Reason for that may be moderately different conditions in the greenhouse and also different timing of the experiment. Therefore, our results also indicate that the experimental environments used in this study provide different environmental cues for the onset of flowering. The onset of flowering is strongly dependent on environmental cues and plants can accelerate or delay it depending on local conditions in order to guarantee seed production for the next generation (Coupland, 1995). Generally, plants tend to flower later at colder temperatures (Capovilla et al., 2015), and, indeed, the average temperature was significantly lower in the climate chamber compared to the garden or greenhouse. Furthermore, the photoperiod differed in the climate chamber with approximately 2.5 h less light per day compared to the garden. Finally, the climate chamber provides the most stable conditions compared to the other environments. Descendants showed a different phenology under the shorter photoperiod, more benign temperatures and more stable conditions in the climate chamber. High variability in environmental variables in general, and even more so under climate change, is the norm under natural conditions, and perennials such as *L. hispidus* might take advantage of a stable period to grow vegetatively in order to ensure survival and increase future reproductive output, thus delaying the onset of flowering time (Tun et al., 2021). However, following this explanation, we would also expect higher vegetative biomass of descendants in the climate chamber, which was not the case. If grazing at annually recurring times selected for later flowering plants, as suggested by Rauschkolb and colleagues (2022),

then the underlying process causing delayed flowering time might be through shifts in photoperiod requirement.

The majority of resurrection studies investigate the evolutionary responses of plant populations only in a single growth facility. As our study shows, depending on the trait, phenotypic differences are not guaranteed to be detected in a given experimental environment (i.e., growth facility), even if genetic differences are present. Especially when investigating evolution of plant populations, having the experimental conditions as close to their natural habitat as possible is desirable to detect evolution to the contemporary environmental conditions. With this approach, we can study the phenotypes that would occur in the field under natural conditions. However, creating deviating or even stressful environmental conditions may reveal genetic differences that are only expressed during a stressful period. The selection that caused genetic differences between ancestors and descendants could have been applied on plasticity under stressful conditions, which can only be observed under those stressful conditions. But care should be taken, because extreme conditions may also elicit responses that are not being expressed under natural extreme conditions (Ghalambor et al., 2007). When choosing between the three growth facilities we tested in this study, the best choice depends on the research aims. The garden seems to be the best option to mimic natural conditions, because of natural light, low spatial heterogeneity and contemporary weather conditions. The greenhouse seems to be the poorest choice for natural conditions, because of low light intensity and high temperatures, although it is the most used in resurrection studies (e.g., Anstett et al., 2021; Franks et al., 2007; Gay et al., 2022; Hamann et al., 2018; Kuester et al., 2016; Lambrecht et al., 2020; Nevo et al., 2012; O'Hara et al., 2021; Sultan et al., 2013; Vtipil & Sheth, 2020). Chiang and colleagues (2021) showed that environmental fluctuations can affect the phenotypic expression of multiple traits and are important to study natural-like plant

growth. Here, climate chambers present an intriguing option going forward if they are programmed very closely to field conditions. Using climate or weather data from the collection sites, one could program the average temperatures, moisture, light spectrum and their daily to seasonal variability to have the conditions very close to the field, while still having a high level of control (Poorter et al., 2016). This idea has already been successfully applied by Heuermann and colleagues (2023) who managed to simulate whole seasons in a reproducible manner in their specialized indoor growth facility. Additionally, extreme events can be modelled (e.g., heatwaves) as treatments to further investigate if differences are also expressed under such conditions, especially if these conditions are potential selection agents. Very controlled and homogenous conditions on the other hand, can be useful if only genetic differences per se are of interest and not how they relate to the actual natural conditions. Especially low temporal variation in experimental conditions is important to provide similar conditions throughout all ontogenetic stages, and to avoid interactions of ontology and the environmental conditions. Furthermore, it is possible to infer adaptive traits in the field from genotypes grown indoors by using modelling of the environmental conditions (Boudghaghen et al., 2023).

Intermediate generations

We found no differences between the two intermediate generations used in this study within each of the growth facilities, indicating either that no significant parental effects were present beforehand at all or that they were removed after the first intermediate generation. The magnitude and nature of parental effects can be strongly dependent on the environmental stresses experienced by the parents, as Latzel and colleagues (2023) showed that these stresses can affect the fitness of the offspring by up to 35

% in *Arabidopsis thaliana*. It is likely that environmental stresses differed between ancestors and descendants of our study population, as climate change increased the frequency and duration of droughts and heatwaves (IPCC, 2018). However, if there were any detectable parental effects, these have been eliminated in the first intermediate generation, which has also been found in other studies (Agrawal, 2002; Gianoli, 2002). Since we did not have seed material of the originally collected sample – i.e. before the intermediate generations –, we cannot quantify how much the first intermediate generation reduced parental and storage effects. Thus, we cannot know whether there were detectable parental effects in the first place. Theoretically, long-term storage of seeds may cause carry-over effects into the F1-generation in suboptimal storage conditions, with resurrected plants having lower fitness due to the storage and producing lower quality seeds (Gebeyehu, 2020). The seed material used in this study was stored at -20°C after drying at 15% RH which are optimal conditions to ensure viability for several decades (Solberg et al., 2020). The implications of potential carry-over effects can of course be very dependent on the storage condition and storage length, and other species might be affected more than *L. hispidus*. Thus, multispecies experiments are needed to advance our understanding of parental and storage effects and to make informed choices regarding the amount of required intermediate generations for a given species.

Conclusions

Our study shows that the choice of the growth facility in common-environment experiments can potentially impact the expression of phenotypic differences among genotypes, thereby affecting the conclusions. Thus, studying evolution of plant populations in only a single environment might result in incomplete or even deficient

interpretations for some traits. Hence, it is important to carefully choose the growth facility or even use multiple facilities. Outdoor garden experiments might be a good and simple option with regard to studying rapid evolution as plants will experience more natural conditions and the contemporary climate which they are expected to have evolved to. However, if environmental variables from the population are well known, using climate chambers might be a good alternative with a high level of control and detailed programming to encompass realistic natural conditions as well as less or more extreme conditions that may occur under natural conditions, e.g., as additional treatments. Finally, growing a second intermediate generation rather than only one intermediate generation did not impact the genetic differences of ancestors and descendants within growth facilities, suggesting that only one intermediate generation would be sufficient to reduce detectable parental and storage effects, if there were any in the first place. Overall, future studies should be aware of implications regarding reproducibility and wisely choose the experimental conditions.

Acknowledgements

The authors thank the Deutsche Bundesstiftung Umwelt (DBU) for their financial support by a PhD scholarship (20020/678) to PK. We thank Lutz Stübing, Susanne Pietsch, Robert Anton and the gardeners from the “Wissenschaftsgarten” of Goethe University Frankfurt for their support, and Charlotte Møller for her help with the data loggers.

Data archiving

The data that support the findings of this study will be made available at Dryad after this manuscript is accepted for publication.

References

- Agrawal, A. A. 2002. Herbivory and maternal effects: mechanisms and consequences of transgenerational induced plant resistance. *Ecology* 83: 3408–3415.
- Anstett, D. N., H. A. Branch, and A. L. Angert. 2021. Regional differences in rapid evolution during severe drought. *Evolution Letters* 5: 130–142.
- Anten, N. P. R. 2005. Optimal photosynthetic characteristics of individual plants in vegetation stands and implications for species coexistence. *Annals of Botany* 95: 495–506.
- Atkin, O. K., B. Botman, and H. Lambers. 1996. The causes of inherently slow growth in alpine plants: An analysis based on the underlying carbon economies of alpine and lowland *Poa* species. *Functional Ecology* 10: 698–707.
- Auge, G. A., L. K. Blair, H. Neville, and K. Donohue. 2017. Maternal vernalization and vernalization-pathway genes influence progeny seed germination. *New Phytologist* 216: 388–400.
- Badyaev, A. V, and T. Uller. 2009. Parental effects in ecology and evolution: mechanisms, processes and implications. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364: 1169–1177.
- Barbosa, A. M. 2015. fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution* 6: 853–858.
- Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57: 289–300.
- Bischoff, A., and H. Müller-Schärer. 2010. Testing population differentiation in plant species – how important are environmental maternal effects. *Oikos* 119: 445–454.
- Blödner, C., C. Goebel, I. Feussner, C. Gatz, and A. Polle. 2007. Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in *Arabidopsis thaliana* progenies. *Plant, cell & environment* 30: 165–175.
- Bobbink, R., K. Hicks, J. Galloway, T. Spranger, R. Alkemade, M. Ashmore, M. Bustamante, et al. 2010. Global assessment of nitrogen deposition effects on

- terrestrial plant diversity: A synthesis. *Ecological Applications* 20: 30–59.
- Bongers, F. J., M. Olmo, B. Lopez-Iglesias, N. P. R. Anten, and R. Villar. 2017. Drought responses, phenotypic plasticity and survival of Mediterranean species in two different microclimatic sites. *Plant Biology* 19: 386–395.
- Boudighaghen, J., L. Moreau, K. Beauchêne, R. Chapuis, N. Mangel, L. Cabrera-Bosquet, C. Welcker, et al. 2023. Robotized indoor phenotyping allows genomic prediction of adaptive traits in the field. *Nature Communications* 14: 1–14.
- Cabrera-Bosquet, L., C. Fournier, N. Brichet, C. Welcker, B. Suard, and F. Tardieu. 2016. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. *New Phytologist* 212: 269–281.
- Capovilla, G., M. Schmid, and D. Posé. 2015. Control of flowering by ambient temperature. *Journal of Experimental Botany* 66: 59–69.
- Chiang, C., D. Bânkestad, and G. Hoch. 2021. Effect of asynchronous light and temperature fluctuations on plant traits in indoor growth facilities. *Agronomy* 11: 1–13.
- Clausen, J., D. D. Keck, and W. M. Hiesey. 1940. Experimental studies on the nature of species. I. Effect of varied environments on western North American plants. *Experimental studies on the nature of species. I. Effect of varied environments on western North American plants*.
- Coupland, G. 1995. Genetic and environmental control of flowering time in Arabidopsis. *Trends in Genetics* 11: 393–397.
- Drummond, C. 2009. Replicability is not reproducibility: Nor is it good science. *Proceedings of the Evaluation Methods for Machine Learning Workshop at the 26th ICML*.
- Easlon, H. M., and A. J. Bloom. 2014. Easy Leaf Area: Automated digital image analysis for rapid and accurate measurement of leaf area. *Applications in Plant Sciences* 2: 1400033.
- Elwell, A. L., D. S. Gronwall, N. D. Miller, E. P. Spalding, and T. L. D. Brooks. 2011. Separating parental environment from seed size effects on next generation growth and development in Arabidopsis. *Plant, cell & environment* 34: 291–301.
- Franks, S. J. 2015. The unique and multifaceted importance of the timing of flowering. *American Journal of Botany* 102: 1401–1402.
- Franks, S. J., E. Hamann, and A. E. Weis. 2018. Using the resurrection approach to understand contemporary evolution in changing environments. *Evolutionary Applications* 11: 17–28.
- Franks, S. J., S. Sim, and A. E. Weis. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences of the United States of America* 104: 1278–1282.
- Galloway, J. N., A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. A. Martinelli, et al. 2008. Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. *Science* 320: 889–892.

- Gay, L., J. Dhinaut, M. Jullien, R. Vitalis, M. Navascués, V. Ranwez, and J. Ronfort. 2022. Evolution of flowering time in a selfing annual plant: Roles of adaptation and genetic drift. *Ecology and Evolution* 12: 1–17.
- Gebeyehu, B. 2020. Review on: Effect of seed storage period and storage environment on seed quality. *International Journal of Applied Agricultural Sciences* 6: 185.
- Ghalambor, C. K., J. K. McKay, S. P. Carroll, and D. N. Reznick. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* 21: 394–407.
- Gianoli, E. 2002. Maternal environmental effects on the phenotypic responses of the twining vine *Ipomoea purpurea* to support availability. *OIKOS* 99: 324–330.
- Hamann, E., D. Denney, S. Day, E. Lombardi, M. I. Jameel, R. MacTavish, and J. T. Anderson. 2021. Review: Plant eco-evolutionary responses to climate change: Emerging directions. *Plant Science* 304: 110737.
- Hamann, E., A. E. Weis, and S. J. Franks. 2018. Two decades of evolutionary changes in *Brassica rapa* in response to fluctuations in precipitation and severe drought. *Evolution*: 2682–2696.
- Herman, J., and S. Sultan. 2011. Adaptive transgenerational plasticity in plants: Case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science* 2.
- Heuermann, M. C., D. Knoch, A. Junker, and T. Altmann. 2023. Natural plant growth and development achieved in the IPK PhenoSphere by dynamic environment simulation. *Nature Communications* 14: 1–14.
- Hill, W. G., and A. Caballero. 1992. Artificial selection experiments. *Annual Review of Ecology and Systematics* 23: 287–310.
- IPCC. 2018. Summary for Policymakers. In: Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global.
- Jablonka, E., and G. Raz. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly review of biology* 84: 131–176.
- Jung, V., C. H. Albert, C. Violle, G. Kunstler, G. Loucougaray, and T. Spiegelberger. 2014. Intraspecific trait variability mediates the response of subalpine grassland communities to extreme drought events. *Journal of Ecology* 102: 45–53.
- Karitter, P., M. March-Salas, A. Ensslin, R. Rauschkolb, S. Godefroid, and J. F. Scheepens. 2023b. Combining the resurrection approach with transplant experiments to investigate adaptation of plant populations to environmental change. *Perspectives in Plant Ecology, Evolution and Systematics* in press.
- Kawecki, T. J., and D. Ebert. 2004. Conceptual issues in local adaptation. *Ecology Letters* 7: 1225–1241.
- Knauer, J., T. S. El-Madany, S. Zaehle, and M. Migliavacca. 2018. Bigleaf—An R package for the calculation of physical and physiological ecosystem properties from eddy covariance data. *PLOS ONE* 13: e0201114.

- Kuester, A., A. Wilson, S. M. Chang, and R. S. Baucom. 2016. A resurrection experiment finds evidence of both reduced genetic diversity and potential adaptive evolution in the agricultural weed *Ipomoea purpurea*. *Molecular ecology* 25: 4508–4520.
- Kühn, I., and S. Klotz. 2002. BIOLFLOR - a database on biological and ecological traits of vascular plants in Germany. *Verhandlungen der Gesellschaft für Ökologie*, 42.
- De La Riva, E. G., F. Lloret, I. M. Pérez-Ramos, T. Marañón, S. Saura-Mas, R. Díaz-Delgado, and R. Villar. 2017. The importance of functional diversity in the stability of Mediterranean shrubland communities after the impact of extreme climatic events. *Journal of Plant Ecology* 10: 281–293.
- Lambrecht, S. C., A. K. Gujral, L. J. Renshaw, and L. T. Rosengreen. 2020. Evolutionary and plastic changes in a native annual plant after a historic drought. *Ecology and Evolution* 10: 4570–4582.
- Lascoux, M., S. Glémin, and O. Savolainen. 2016. Local adaptation in plants. *eLS*: 1–7.
- Latzel, V. 2015. Pitfalls in ecological research – transgenerational effects. *Folia Geobotanica* 50: 75–85.
- Latzel, V., M. Fischer, M. Groot, R. Gutzat, C. Lampei, J. Ouborg, M. Parepa, et al. 2023. Parental environmental effects are common and strong, but unpredictable, in *Arabidopsis thaliana*. *New Phytologist* 237: 1014–1023.
- Lenth, R. V. 2021. emmeans: Estimated marginal means, aka least-squares means.
- Mänd, P., L. Hallik, J. Peñuelas, and O. Kull. 2013. Electron transport efficiency at opposite leaf sides: Effect of vertical distribution of leaf angle, structure, chlorophyll content and species in a forest canopy. *Tree Physiology* 33: 202–210.
- Massonnet, C., D. Vile, J. Fabre, M. A. Hannah, C. Caldana, J. Lisec, G. T. S. Beemster, et al. 2010. Probing the reproducibility of leaf growth and molecular phenotypes: A comparison of three *Arabidopsis* accessions cultivated in ten laboratories. *Plant Physiology* 152: 2142–2157.
- Monson, R. K., and S. D. Smith. 1982. Seasonal water potential components of Sonoran desert plants. *Ecology* 63: 113–123.
- Nevo, E., Y. B. Fu, T. Pavlicek, S. Khalifa, M. Tavasi, and A. Beiles. 2012. Evolution of wild cereals during 28 years of global warming in Israel. *Proceedings of the National Academy of Sciences of the United States of America* 109: 3412–3415.
- Newman, E. I. 1995. Phosphorus inputs to terrestrial ecosystems. *British Ecological Society* 83: 713–726.
- O’Hara, N. B., S. J. Franks, N. C. Kane, S. Tittes, and J. S. Rest. 2021. Evolution of pathogen response genes associated with increased disease susceptibility during adaptation to an extreme drought in a *Brassica rapa* plant population. *BMC Ecology and Evolution* 21: 1–11.
- Parmesan, C., and M. E. Hanley. 2015. Plants and climate change: Complexities and surprises. *Annals of Botany* 116: 849–864.
- Pelletier, F., D. Garant, and A. P. Hendry. 2009. Eco-evolutionary dynamics.

Philosophical Transactions of the Royal Society B: Biological Sciences 364: 1483–1489.

Poorter, H., F. Fiorani, R. Pieruschka, T. Wojciechowski, W. H. van der Putten, M. Kleyer, U. Schurr, and J. Postma. 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist* 212: 838–855.

Poorter, H., F. Fiorani, M. Stitt, U. Schurr, A. Finck, Y. Gibon, B. Usadel, et al. 2012. The art of growing plants for experimental purposes: A practical guide for the plant biologist. *Functional Plant Biology* 39: 821–838.

Poorter, H., Ü. Niinemets, N. Ntagkas, A. Siebenkäs, M. Mäenpää, S. Matsubara, and T. L. Pons. 2019. A meta-analysis of plant responses to light intensity for 70 traits ranging from molecules to whole plant performance. *New Phytologist* 223: 1073–1105.

Poorter, H., Ü. Niinemets, L. Poorter, I. J. Wright, and R. Villar. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist* 182: 565–588.

R Core Team. 2020. R: A Language and Environment for Statistical Computing.

Rauschkolb, R., W. Durka, S. Godefroid, L. Dixon, O. Bossdorf, A. Ensslin, and J. F. Scheepens. 2023. Recent evolution of flowering time across multiple European plant species correlates with changes in aridity. *Oecologia*.

Rauschkolb, R., L. Henres, C. Lou, S. Godefroid, L. Dixon, W. Durka, O. Bossdorf, et al. 2022a. Historical comparisons show evolutionary changes in drought responses in European plant species after two decades of climate change. *Basic and Applied Ecology* 58: 26–38.

Rauschkolb, R., Z. Li, S. Godefroid, L. Dixon, W. Durka, M. Májeková, O. Bossdorf, et al. 2022b. Evolution of plant drought strategies and herbivore tolerance after two decades of climate change. *New Phytologist* 235: 773–785.

Richards, C. L., C. Alonso, C. Becker, O. Bossdorf, E. Bucher, M. Colomé-Tatché, W. Durka, et al. 2017. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology letters* 20: 1576–1590.

Solberg, S. Ø., F. Yndgaard, C. Andreassen, R. von Bothmer, I. G. Loskutov, and Å. Asdal. 2020. Long-term storage and longevity of orthodox seeds: A systematic review. *Frontiers in Plant Science* 11: 1–14.

Stanton, M. L., B. A. Roy, and D. A. Thiede. 2000. Evolution in stressful environments. I. Phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. *Evolution* 54: 93–111.

Sultan, S. E. 2000. Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science* 5: 537–542.

Sultan, S. E., T. Horgan-Kobelski, L. M. Nichols, C. E. Riggs, and R. K. Waples. 2013. A resurrection study reveals rapid adaptive evolution within populations of an invasive plant. *Evolutionary Applications* 6: 266–278.

Tun, W., J. Yoon, J. S. Jeon, and G. An. 2021. Influence of climate change on flowering time. *Journal of Plant Biology* 64: 193–203.

- Turesson, G. 1922. The genotypical response of the plant species to the habitat. *Hereditas* 3: 211–350.
- Vitra, A., C. Delé Glise, M. Meisser, A. C. Risch, C. Signarbieux, L. Lamacque, S. Delzon, et al. 2019. Responses of plant leaf economic and hydraulic traits mediate the effects of early- and late-season drought on grassland productivity. *AoB PLANTS* 11: 1–15.
- Volaire, F. 2008. Plant traits and functional types to characterise drought survival of pluri-specific perennial herbaceous swards in Mediterranean areas. *European Journal of Agronomy* 29: 116–124.
- Vtipil, E. E., and S. N. Sheth. 2020. A resurrection study reveals limited evolution of phenology in response to recent climate change across the geographic range of the scarlet monkeyflower. *Ecology and Evolution* 10: 14165–14177.
- Weis, A. E. 2018. Detecting the “invisible fraction” bias in resurrection experiments. *Evolutionary Applications* 11: 88–95.
- Wellstein, C., P. Poschlod, A. Gohlke, S. Chelli, G. Campetella, S. Rosbakh, R. Canullo, et al. 2017. Effects of extreme drought on specific leaf area of grassland species: A meta-analysis of experimental studies in temperate and sub-Mediterranean systems. *Global Change Biology* 23: 2473–2481.
- Wulff, R. D., H. F. Causin, O. Benitez, and P. A. Bacalini. 1999. Intraspecific variability and maternal effects in the response to nutrient addition in *Chenopodium album*. *Canadian Journal of Botany* 77: 1150–1158.

Tables and figures

Table 1. Mean values and standard errors of environmental variables over the whole experimental period for all growth facilities.

Growth facility	Temperature [°C]			Daily light integral [mol·m ⁻² ·d ⁻¹]
	Soil (5cm)	Soil surface	Air	
Garden	21.2 ± 0.2	22.2 ± 0.3	22.6 ± 0.3	40.5 ± 1.6
Greenhouse	23.9 ± 0.2	25.5 ± 0.2	26.1 ± 0.2	7.5 ± 0.3
Climate chamber	20.4 ± 0.05	21.2 ± 0.03	20.7 ± 0.04	15.0 ± 1.0

794 **Table 2.** Results of the statistical models testing the effects of temporal origin (ancestors, descendants), intermediate generation (F1, F2), growth facility (garden,
795 greenhouse, climate chamber) and their interactions on the response variables (y) rosette diameter, vegetative biomass, specific leaf area (SLA), leaf dry matter
796 content (LDMC), onset of flowering, reproductive biomass, number of stems and SPAD measurements. We used linear mixed effects models with initial size as
797 covariate and maternal line as random factor followed by ANOVA's. Response variables were transformed if needed to fulfil parametric assumptions. Shown are
798 degrees of freedom (*df*), *F* values and adjusted *p* values using False Discovery Rates. Significant *p* values are shown with asterisks (*: $P \leq 0.05$; ***: $P \leq 0.001$)
799 and marginally significant values are shown in bold.

800

		Rosette diameter		Vegetative biomass		SLA		LDMC		Onset of flowering		Reproductive biomass		Number of stems		SPAD	
	df	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj	<i>F</i> value	<i>P</i> value adj
Initial size		65.42	***	25.00	***	8.41	*	17.28	***	20.74	***	2.09	0.3	0.14	0.8	0.78	0.6
Origin	1	2.61	0.3	8.55	*	0.18	0.8	9.39	*	7.01	*	0.00	1.0	1.65	0.4	2.90	0.3
Gen	1	2.26	0.3	0.69	0.6	1.70	0.4	0.23	0.8	1.75	0.4	0.18	0.8	0.25	0.8	2.45	0.3
Facility	2	155.48	***	301.09	***	199.31	***	91.46	***	50.73	***	457.85	***	280.86	***	210.71	***
Origin × Gen	1	0.31	0.8	2.78	0.3	0.23	0.8	1.30	0.5	0.71	0.6	0.20	0.8	1.12	0.5	0.95	0.6
Origin × Facility	2	3.91	0.067	1.55	0.4	0.52	0.8	0.53	0.8	5.57	*	2.07	0.3	2.71	0.2	0.11	0.9
Gen × Facility	2	4.77	*	0.39	0.8	1.00	0.6	0.68	0.7	0.41	0.8	0.09	0.9	1.98	0.3	4.57	*
Origin × Gen × Facility	2	0.10	0.9	0.32	0.8	0.24	0.9	1.52	0.4	0.30	0.8	1.47	0.4	1.75	0.4	0.20	0.9
Transformation		(x) ^{1.5}		(x)		log(x)		sqrt(x)		(x) ²		(x) ^{0.3}		(x) ^{0.9}		(x)	

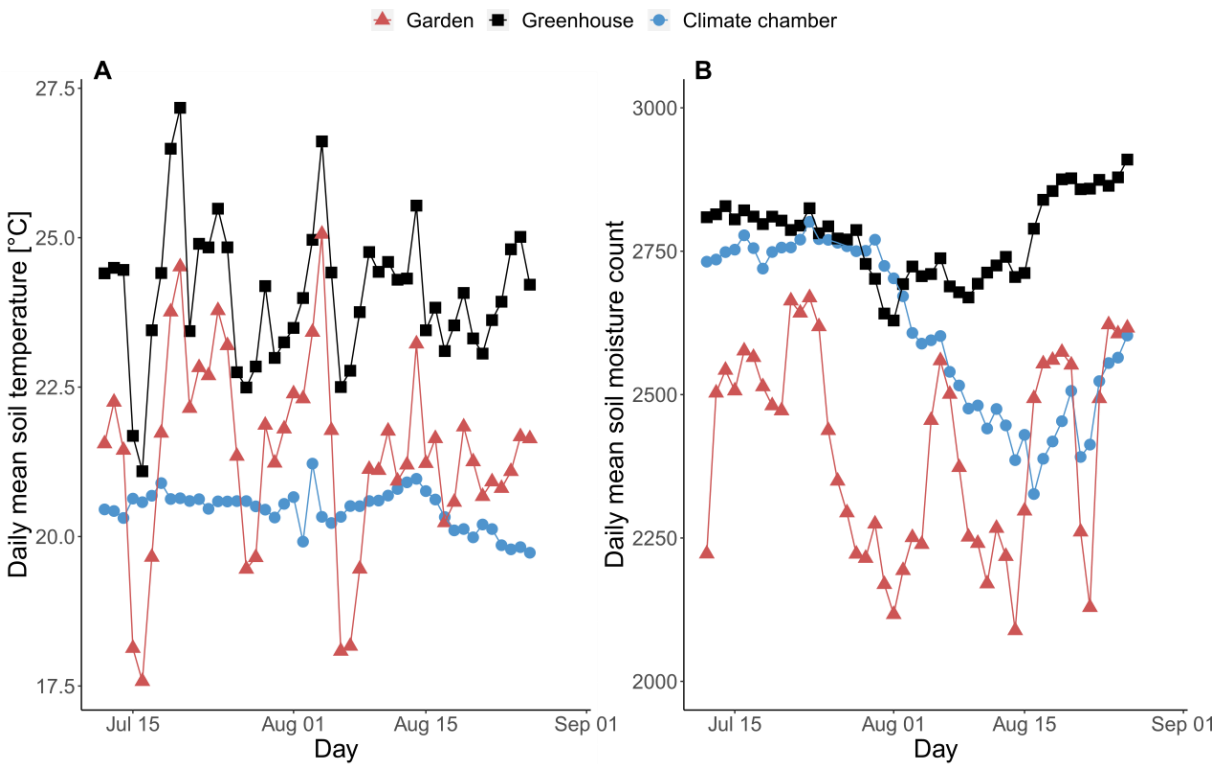


Figure 1. Daily mean soil temperature (A) and daily mean soil moisture count (B) of four random pots grown in different growth facilities. The growth facilities are garden (red triangles), greenhouse (black squares), and climate chamber (blue circles).

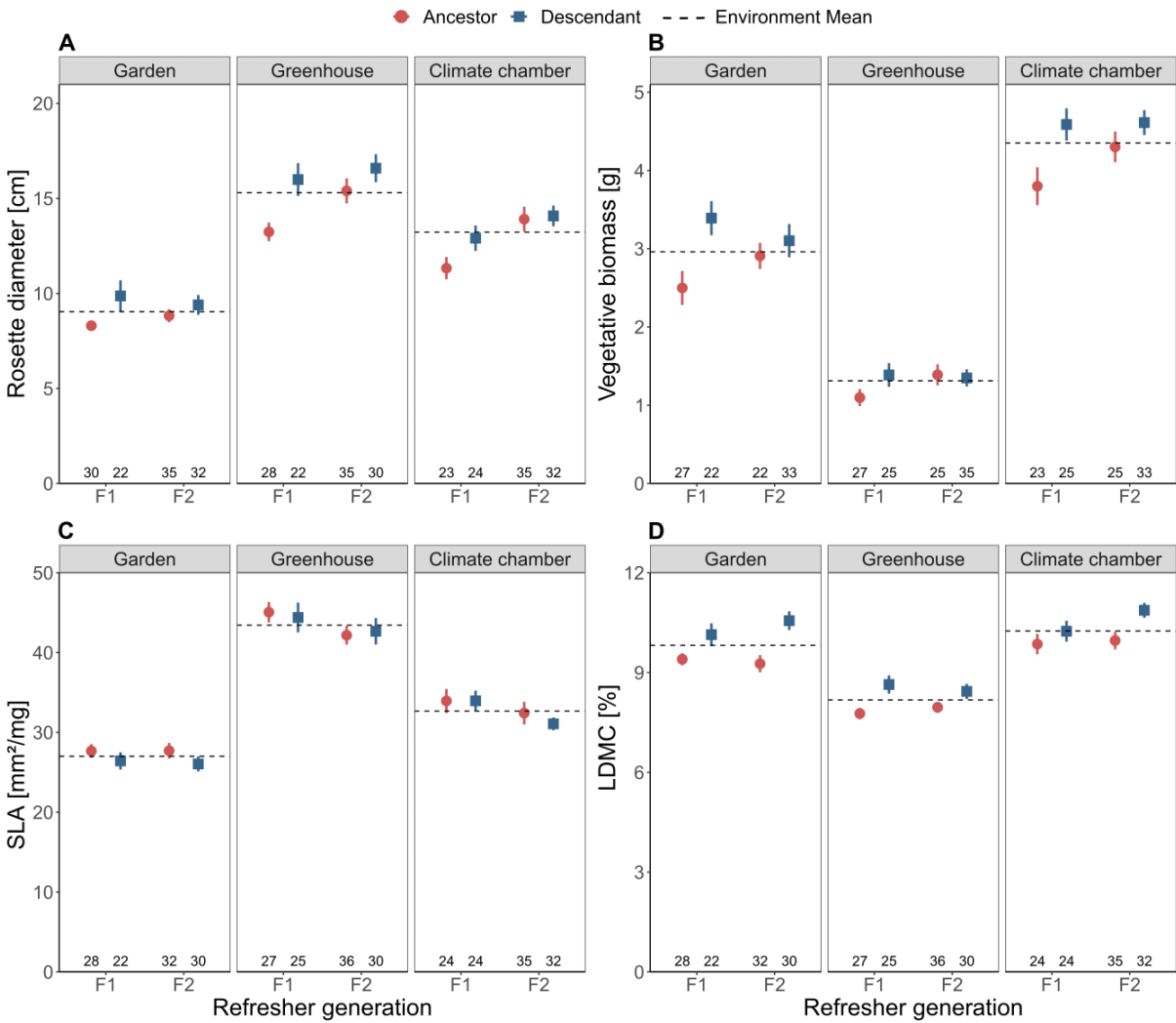


Figure 2. Rosette diameter (A), vegetative biomass (B), specific leaf area (C) and leaf dry matter content (D) of ancestors (blue) and descendants (red) after one intermediate generation (F1) and two intermediate generations (F2) grown in different growth facilities (garden, greenhouse, climate chamber). Shown are means and standard errors. The dotted line represents the overall mean value in each environment. Sample sizes are given at the bottom of the graph below their respective data point.

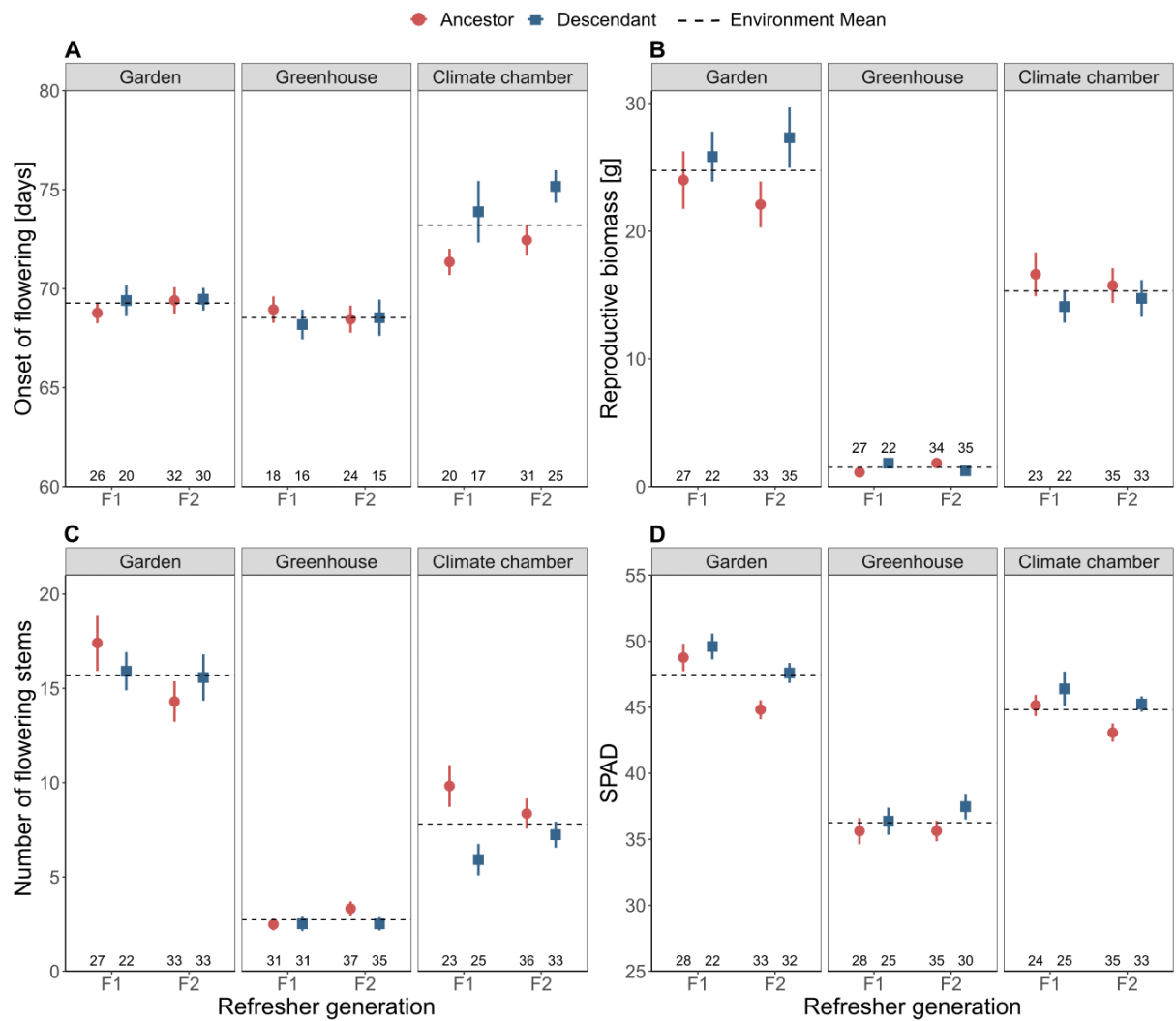


Figure 3. Onset of flowering (A), reproductive biomass (B), number of flowering stems (C) and SPAD values (D) of ancestors (blue) and descendants (red) after one intermediate generation (F1) and two intermediate generations (F2) grown in different growth facilities (garden, climate chamber, greenhouse). Shown are means and standard errors. The dotted line represents the overall mean value in each environment. Sample sizes are given at the bottom of the graph below their respective data point.

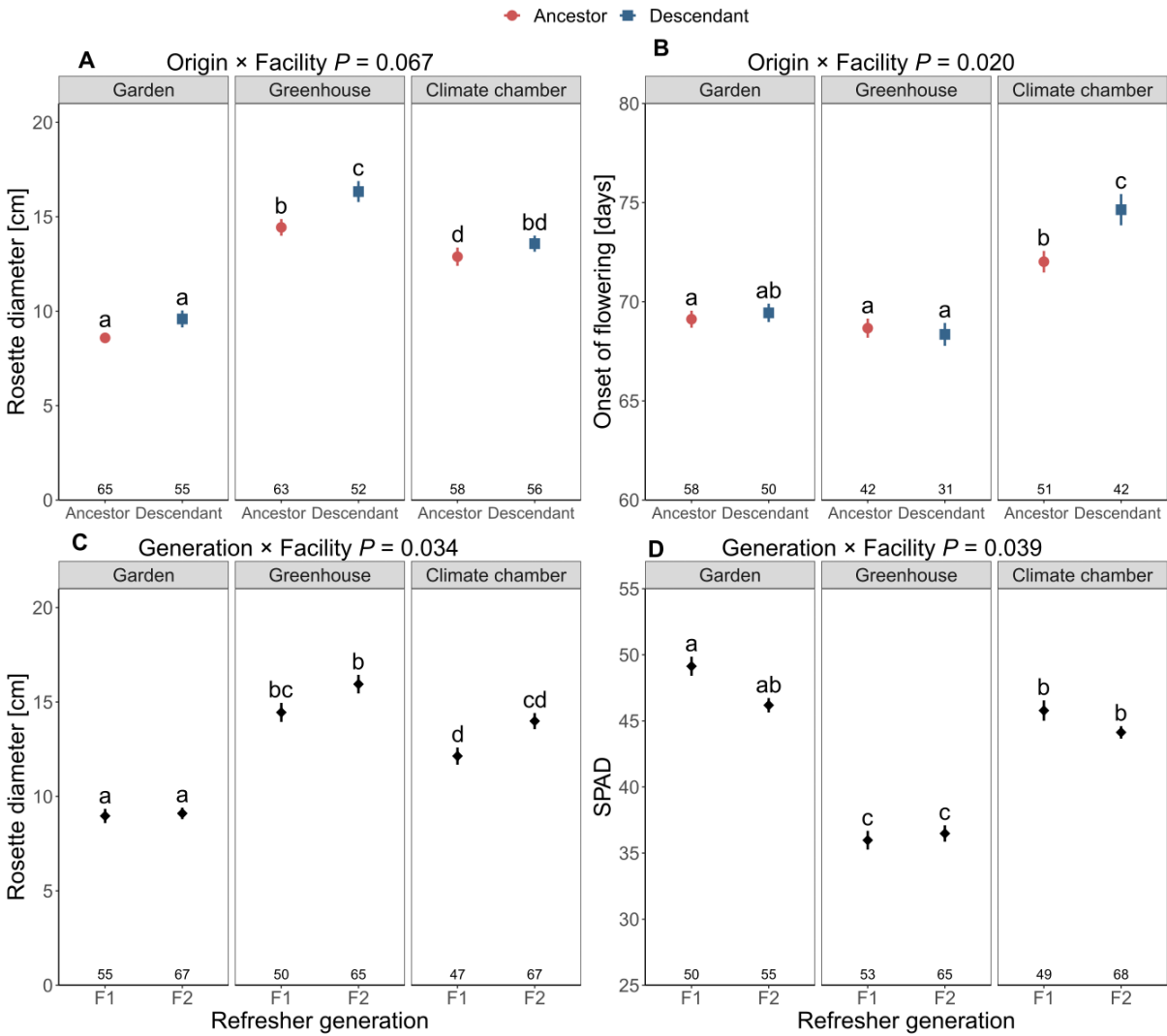
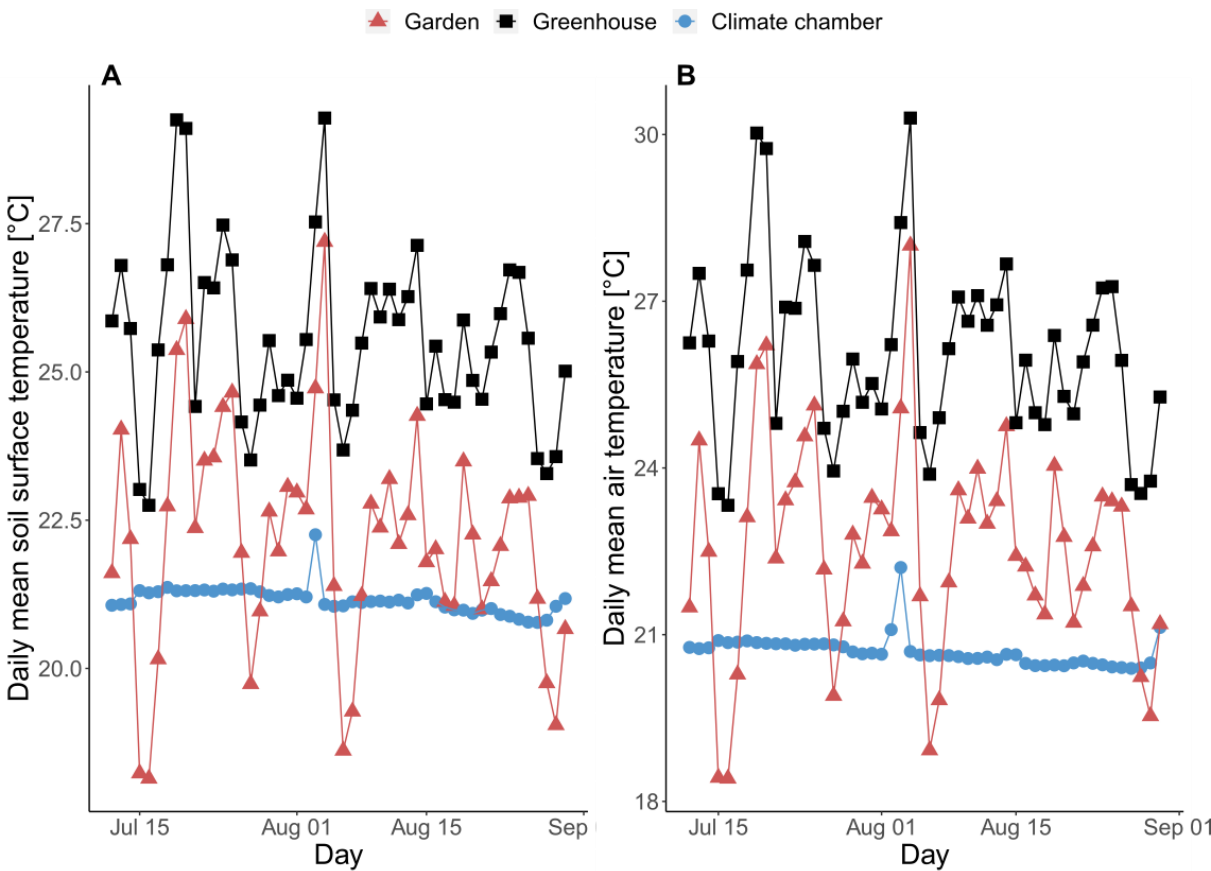


Figure 4. Rosette diameter (A) and onset of flowering (B) of ancestors (red) and descendants (blue) in the different environments (significant Origin × Facility effect). Rosette diameter (C) and SPAD values (D) of F1 and F2 intermediate generations in different facilities (significant Generation × Facility effect). Shown are means and standard errors. Sample sizes are given at the bottom of the graph below their respective data point.

Supporting information

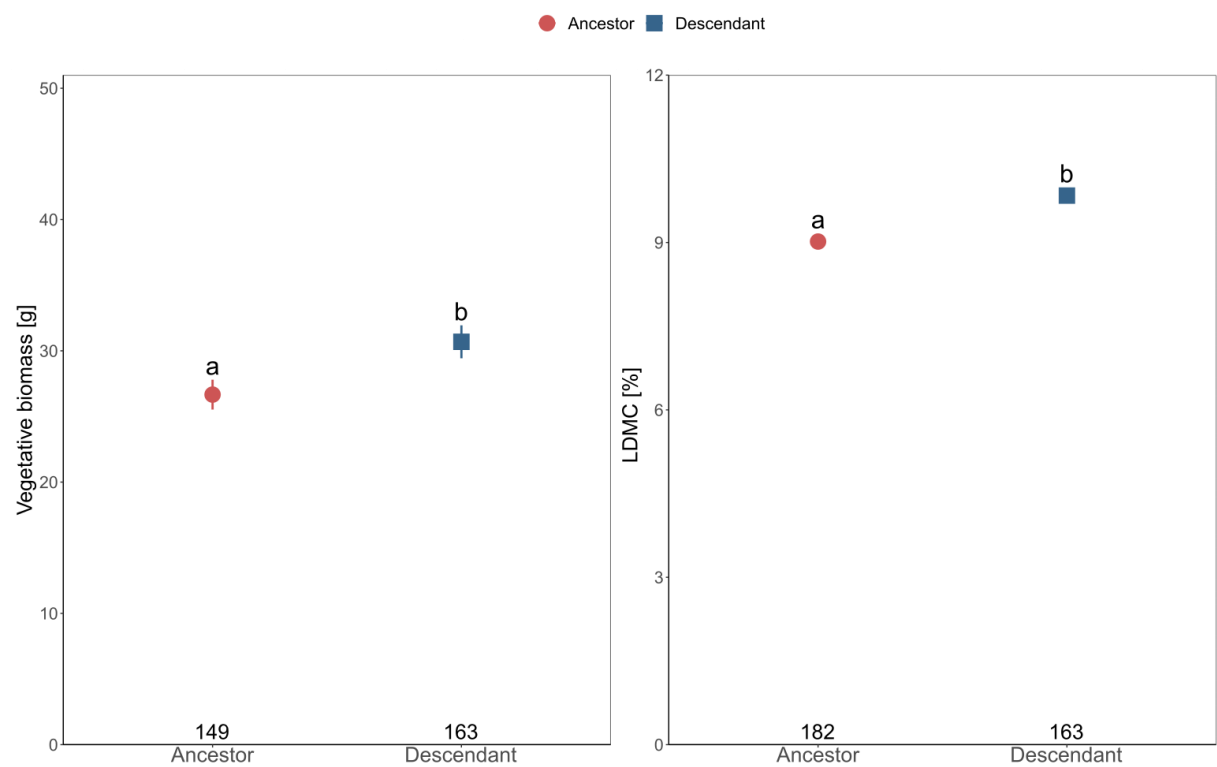


Appendix S1. Daily mean soil surface temperature (A) and daily mean air temperature (B) of four random pots grown in different growth facilities. The growth facilities are garden (red triangles), greenhouse (black squares), and climate chamber (blue circles).

845 **Appendix S2.** Results before adjustments of the statistical models testing the effects of temporal origin (ancestors, descendants), intermediate generation (F1,
846 F2), growth facility (garden, greenhouse, climate chamber) and their interactions on the response variables (y) rosette diameter, vegetative biomass, specific leaf
847 area (SLA), leaf dry matter content (LDMC), onset of flowering, reproductive biomass, number of stems and SPAD measurements. We used linear mixed effects
848 models with initial size as covariate and maternal line as random factor followed by ANOVA's. Significant values ($P < 0.05$) are shown in bold.

		Rosette diameter	Vegetative biomass	SLA	LDMC	Onset of flowering	Reproductive biomass	Number of stems	SPAD
	<i>df</i>	<i>P</i> value	<i>P</i> value	<i>P</i> value	<i>P</i> value	<i>P</i> value	<i>P</i> value	<i>P</i> value	<i>P</i> value
Initial size		<0.001	<0.001	0.004	<0.001	<0.001	0.15	0.708	0.379
Origin	1	0.118	0.007	0.676	0.005	0.014	0.969	0.210	0.101
Gen	1	0.145	0.414	0.204	0.639	0.198	0.673	0.623	0.130
Env	2	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Origin × Gen	1	0.584	0.106	0.633	0.264	0.408	0.659	0.299	0.339
Origin × Env	2	0.021	0.214	0.593	0.590	0.004	0.128	0.068	0.899
Gen × Env	2	0.009	0.678	0.371	0.507	0.661	0.912	0.140	0.011
Origin × Gen × Env	2	0.902	0.728	0.785	0.219	0.740	0.232	0.176	0.817

849



Appendix S3. Vegetative biomass (A) and LDMC (B) of ancestors and descendants (significant Origin effect). Shown are means and standard errors. Standard errors of LDMC are too small to be visible. Sample sizes are given at the bottom of the graph below their respective data point.