Sublethal effects of metal toxicity and the measure of plant fitness in ecotoxicological experiments

Julien Nowak, Nathalie Faure, Cédric Glorieux, Denis Vile, Maxime Pauwels, Hélène Frérot

To cite this version:

Julien Nowak, Nathalie Faure, Cédric Glorieux, Denis Vile, Maxime Pauwels, et al.. Sublethal effects of metal toxicity and the measure of plant fitness in ecotoxicological experiments. Environmental Pollution, 2022, 304, pp.119138. 10.1016/j.envpol.2022.119138. hal-03678102

HAL Id: hal-03678102
https://hal.inrae.fr/hal-03678102v1

Submitted on 15 Apr 2024

Sublethal effects of metal toxicity and the measure of plant fitness in ecotoxicological experiments

Julien Nowak[1], Nathalie Faure[1], Cédric Glorieux[1], Denis Vile[2], Maxime Pauwels[1,3$], Hélène Frérot[1,3$*]

1. Univ. Lille, CNRS, UMR 8198 - EEP - Laboratoire Evolution Ecologie Paléontologie F-59000 Lille, France
2. LEPSE, Univ Montpellier, INRA, SupAgro, Montpellier, France
3. Univ. Lille, CNRS, UMR 8516 - LASIRE - Laboratoire Avancé de Spectroscopie pour les Interactions la Réactivité et l'Environnement, F-59000 Lille, France
4. **[$]These authors contributed equally to this work**

[*] Correspondence: Hélène Frérot

Email: helene.frerot@univ-lille.fr

Tel: + 33 3 20 43 49 29

Abstract

Anthropogenic pollution is a major driver of global environmental change. To be properly addressed, the study of the impact of pollutants must consider both lethal effects and sublethal effects on individual fitness. However, measuring fitness remains challenging. In plants, the total number of seeds produced, *i.e.* the seed set, is traditionally considered, but is not readily accessible. Instead, performance traits related to survival, e.g., vegetative biomass and reproductive success, can be measured, but their correlation with seed set has rarely been investigated.

To develop accurate estimates of seed set, relationships among 15 vegetative and reproductive traits were analyzed. For this purpose, *Noccaea caerulescens* (Brassicaceae), a model plant to study local adaptation to metal-contaminated environments, was used. To investigate putative variation in trait relationships, sampling included several accessions cultivated in contrasting experimental conditions. To test their applicability, selected estimates were used in the first generation of a Laboratory Natural Selection (LNS) experiment exposing experimentally plants to zinc soil pollution.

Principal component analyses revealed statistical independence between vegetative and reproductive traits. Traits showing the strongest positive correlation with seed set were the number of non-aborted silicles, and the product of this number and mean silicle length. They thus appeared the most appropriate to document sublethal or fitness effects of environmental contaminants in plant ecotoxicological studies. The relevance of both estimates was confirmed by using them to assess the fitness of parental plants of the first generation of an LNS experiment: the same families consistently displayed the highest or the lowest performance values in two independent experimental metal-exposed populations. Thus, both these fitness estimates could be used to determine the expected number of offspring and the composition of successive generations in further LNS experiments investigating the impact of multi-generational exposure of a plant species to environmental pollution.

Keywords: Ecological Risk Assessment, Experimental evolution, Evolutionary ecotoxicology, Fitness, Performance, Sublethal effects

Introduction

Anthropogenic environmental pollution is listed among the five direct drivers of the current global decline in biodiversity (Díaz et al., 2019). Indeed, environmental stress associated with contaminant exposure can cause severe damage to populations and lead to local extinction (Tovar-Sánchez, 2018). Mitigation efforts require appropriate evaluation of the environmental impact of pollutants, following ecological risk assessment (ERA) processes, to develop relevant conservation programs. However, current ERA schemes may be unsuitable when they estimate environmental risks by focusing solely on organism survival despite evidence that sublethal effects may also have severe consequences (Straub et al., 2020). For example, metal exposure of earthworms can significantly reduce growth and reproduction without causing mortality (Pauwels et al., 2013). Sublethal effects can affect many species, from plants to animals, including habitat-forming species, and this can also impact communities and ecosystem functioning (Mayer-Pinto et al., 2020).

Growth and reproduction are major components of "fitness"—a fundamental factor allowing understanding of population fate in evolutionary biology (Frank, 2011; Haldane, 1937; Hartl, 1989; Krimbas, 2004; Orr, 2009; Pence and Ramsey, 2015). For example, a reduction in individual fitness, such as that caused by environmental stressors, can result in a progressive population demographic decline and increased risk of local extinction (Carlson et al., 2014). Alternatively, population exposure to contaminants can reveal varying capacities of individuals to maintain fitness despite environmental toxicity, through genetic tolerance mechanisms (e.g., varying offspring number). In that sense, contaminant exposure can act as a selective agent that can provoke population genetic changes, and *in fine* restore population growth. This process, known as "evolutionary rescue" (Carlson et al., 2014), is particularly studied in the subdiscipline of ecotoxicology known as "evolutionary ecotoxicology" (Coutellec & Barata, 2011). Thus, because studying fitness allows temporal changes in the genetic composition of populations to be discussed (Orr, 2009), ecotoxicological tests could be improved by including fitness measurements rather than focusing on mortality alone (Brady et al., 2017; Straub et al., 2020).

Fitness was first defined theoretically as an evolutionary model parameter—presumably measurable—assessing the capacity to have offspring and to contribute to the next generation (Krimbas, 2004). In practice, however, due to conceptual issues measuring individual fitness highly challenging (Pence and Ramsey, 2015). For ecological studies, the concept of "whole-organism performance" has been suggested as a relevant proxy (Arnold, 1983). Broadly, performance can be defined at the whole-organism level as the ability to conduct an ecologically relevant task (Irschick et al., 2008; Lailvaux and Husak, 2014). Performance can thus be measured by following so-called “performance traits" that may influence fitness directly and may rely on lower-level functional traits, including morphological, physiological and life-history traits (Violle et al, 2007).

In plants, performance and underlying traits should involve survival, vegetative biomass and reproductive success, representing the three major components of fitness (survival, growth and fecundity, respectively; Violle et al., 2007). However, in most situations, for practical reasons, only one or a limited set of traits are selected for phenotyping. It follows that traits directly contributing to the number of offspring, e.g., total seed production (i.e., seed set), should be preferred (Hereford et al., 2009). However, precise assessment of seed set is often unrealistic in the wild or, more generally, in outdoor conditions. Indeed, climatic events (wind, rainfall, storms, etc.) may disperse seeds from the mother plant before they can be collected. Collecting all seeds requires time-consuming daily surveillance, which is clearly inappropriate for large-scale experiments. To tackle this question, a phenotyping method allowing the estimation, rather than the measurement, of seed production and fitness is required. For example, Roux et al. (2004) used some reproductive traits expected to show a statistical relationship with seed set. On the other hand, Younginger et al. (2017) instead considered biomass production as a reliable estimate of plant fitness. The adoption of such widely differing methodologies to estimate performance or fitness is a consequence of our relatively poor knowledge of the relationship between seed set and other vegetative and reproductive traits (but see, for example Kleyer et al., 2019).

In the present study, we aimed to investigate which functional or performance traits best correlate with seed set—considered a relevant surrogate for individual fitness. A secondary goal was to provide an empirical tool to measure fitness in Laboratory Natural Selection (LNS) experiments, which represent a powerful strategy with which to investigate the consequences of contaminant exposure on the evolution of populations. As a model plant, we used *Noccaea caerulescens*, a self-compatible Brassicaceae, because it is a well-known pseudometallophyte species, occurring on non-metalliferous, calamine (zinc- and cadmium-enriched) and serpentine (nickel-enriched) soils, showing adaptive phenotypic divergence among metallicolous and non-metallicolous populations (Escarré et al., 2000; Gonneau et al., 2014; Meerts and Van Isacker, 1997; Sterckeman et al., 2017). From various biological materials, including natural accessions and the progenies of controlled crosses, cultivated in either non-polluted or Zn-polluted conditions, we performed a phenotyping survey of a large set of morphological traits of leaves (vegetative traits) and floral stems/silicles (reproductive traits). We also ensured that we were able to collect all seeds produced by individual plants to measure seed set directly. The statistical relationships between traits were then measured so as to define estimates of plant performance that should be both reliable and suitable for large-scale ecotoxicological experiments. As a test application, we used selected estimates to assess fitness variation among individuals in an LNS experiment exposing experimental populations to toxic concentrations of zinc in mesocosms.

Materials and methods

Designing a plant performance estimate

Plant origin. In June 2013, *N. caerulescens* seed families, defined as bulks of seeds collected from a single mother plant, were sampled from natural populations (Table S1). Two sampled sites were long-studied metalliferous (M) smelting sites in Belgium, abandoned several decades ago (Assunção et al., 2003), named "La Calamine" (LC – 25 families) and "Prayon" (PR – 25 families). Three sampled sites were among the closest non-metalliferous (NM) sites in Luxembourg (Reeves et al., 2001): "Lellingen" (LE – 15 families), "Wilwerwiltz" (WIL – 15 families) and "Winseler" (WIN – 20 families). Four of the five sites were also sampled in June 2014: LE (25 families), WIN (28 families), LC (25 families) and PR (23 families).

Plant cultivation and crossing program. In September 2013, seeds from the first sampling (30 seeds per plant family) were sown in seedling trays containing peat-based compost, placed for 1 week in a cold room at 4°C then transferred to a greenhouse for 6 weeks. Then, 84 individuals (LE = 17, WIL = 17, WIN = 17, LC = 16 and PR = 17) were transplanted into individual pots containing 1 kg of a peat-and-clay-based compost (70%)/zeolite (30%) mix and, after pot randomization, placed in a tunnel greenhouse to induce flowering. Of these 84 individuals, 72 (LE = 15, WIL = 15, WIN = 15, LC = 13 and PR = 14) were used in the phenotypic survey described below (Table S2), and the remaining 12 (LE = 2, WIL = 2, WIN = 2, LC = 3 and PR = 3) were involved in controlled crosses (Tables S2 and S3). The inbred progenies corresponded to self-crosses, and the outbred progenies to inter-edaphic crosses, namely MxNM or NMxM crosses (Table S3). All crosses were performed manually by rubbing the open anthers of male parents onto the stigmas of female parents.

In September 2014, 30 seeds harvested from each controlled cross were sown in seedling trays containing peat-based compost. Seedling trays were placed for 1 week in a cold room at 4 °C then transferred to a greenhouse for 6 weeks. Then, 15 representatives from 12 inbred and 12 outbred progenies were transferred into individual pots containing 1 kg of peat and clay-based compost (70%)/zeolite (30%) mix. For 10 of these 15 individuals, the soil mix was polluted with 750 mg of Zn (kg fresh compost)$^{-1}$ (20 mL of a solution at 0.57 M of $ZnSO_4$, 7 H_2O). This moderate level of zinc exposure was chosen to induce toxicity without causing plant mortality, which would have prevented phenotyping and seed harvesting (Nowak et al., 2018). The 360 pots were randomized and placed in a ventilated and unheated greenhouse to induce flowering.

Phenotypic survey. In both 2014 and 2015, the 72 plants that were not involved in controlled crosses (hereafter referred to as “wild” plants because they represented the variability captured from natural populations) and the 360 progenies were phenotyped for several vegetative and reproductive traits throughout their life cycle.

The following vegetative traits were measured at the first flower bud emergence for each plant: the surface of the plant (SP), calculated as the surface of an ellipse from two orthogonal diameters of the rosette (one being the largest diameter); the number of leaves (NL); leaf length (LL); leaf width (LW) from the mean of the three longest leaves; estimated leaf shape ($LS = \frac{LL}{LW}$); and estimated leaf area ($LA = \frac{LL}{2} * \frac{LW}{2} * \pi$). Plants were bagged at the first silicle onset in order to harvest the entire seed set (SS). At the end of seed production, SS was measured with a seed counter (Elmor C3 Seed Counter, Elmor Ltd.). The plants were then harvested and the following reproductive traits were measured: number of flower stems (NFS), length of the longest stem (maxLFS), total number of silicles (NS), number of aborted and non-aborted silicles (NAS and NNAS, respectively), ratio between NNAS and NAS (RS) and mean silicle length (SL) (determined by measuring the length of five silicles per flower stem). The length of silicles corresponded to the length of the septum in open silicles. The estimated total length of silicles (totalLS) was calculated by multiplying the total number of non-aborted silicles by the mean length of silicles ($totalLS = NNAS * SL$). To test the underlying hypothesis that the length of a silicle is related to the number of seeds inside the silicle, 140 silicles were chosen randomly among the 72 wild plants to measure their length and to count the seeds they contained.

Statistical analyses. From datasets obtained from the 72 wild plants and the 360 progenies, separate principal component analyses (PCA) were performed, followed by Pearson's correlations corrected for multiple tests by applying a Benjamini-Hochberg adjustment. The specific aim was to determine which traits, or combinations of traits, showed the highest correlation with seed set, and thus could be used as performance estimates.

Next, we performed hierarchical analysis of variance (ANOVA) to identify and compare the main sources of phenotypic variance for seed set and correlated traits. For the 72 wild plants, a hierarchical ANOVA was performed with the edaphic origin (M or NM) as a fixed effect and the population as random effect nested in edaphic origin. For the 360 progenies, a hierarchical ANOVA was also performed with the pollution level as a fixed effect, the cross type as a fixed effect (M self, NM self, M*NM and NM*M), the cross origin as a random effect nested in cross type, the cross pedigree as a random effect nested in cross origin (see Table S3), and the interactions between the pollution level and the other three factors.

Comparison tests, correlations and linear regressions were performed with R 3.3.2 (R Core Team, 2016), with packages “factoMineR” and “ggplot2” for PCA and graphical representations.

Application to experimental populations

Laboratory Natural Selection design. Seed families from NM populations sampled in June 2014 were used. In September 2014, 15 seeds per family were sown in seedling trays containing peat-based compost, placed for a week in a cold room at 4°C and transferred to a greenhouse for 6 weeks. Depending on germination rates and seedling survival, 36 distinct families were finally available, originating from the Lellingen and Winseler populations. A plant from each family was included once or twice in 49 positions at the center of equal area cells of a 7*7 grid. These 49 plants were placed in a device called a “mesocosm” (see technical description below). One experimental population (EP) actually comprised two mesocosms to obtain a total of 98 individuals per EP (Table S4). Three EPs were finally built by always using two mesocosms per EP, and the same families were represented in the same proportions in all mesocosms. These EPs were thus named EP1 (EP1.1 + EP1.2), EP2 (EP2.1 + EP2.2) and EP3 (EP3.1 + EP3.2).

In late October 2014, EPs were transferred to mesocosms and placed outdoors so that the plants could fulfill their life cycle under natural climatic conditions. In such experimental conditions, it is generally not possible to determine the total number of seeds produced, so this had to be estimated.

Experimental culture conditions. Mesocosms were constructed using square tubs designed using a lysimeter model (Ruttens et al., 2006). Tubs had an area of 0.436 m² enclosed by a 13-cm-wide buffer zone. Buffer zones were filled with compost in order to restrain border effects and create a thermic buffer (Figure S1). Tubs were filled with a mix of 140 kg of peat-and-clay-based compost and 70 kg of zeolite. In EP2.1, EP2.2, EP3.1 and EP3.2, zinc sulfate ($ZnSO_4$, $7H_2O$) was introduced, in powder form, to obtain a pollution level of 750 mg (kg fresh substrate)$^{-1}$ (659.59 g of $ZnSO_4$, $7H_2O$), while EP1.1 and EP1.2 were non-contaminated environments. Therefore, EP2 and EP3 were considered as biological replicates and EP1 was considered as a control. In order to obtain homogeneous substrates, mixes were prepared in a cement mixer. During the life cycle, the same traits as those used for the 72 wild plants and the progenies (see above) were followed. In addition, we randomly chose 236 silicles distributed over EP1.2 to measure their length and count the number of seeds in each silicle.

Statistical analyses. A PCA was performed to assess the data structure, followed by Pearson's correlation coefficients corrected for multiple tests by applying a Benjamini-Hochberg adjustment. In order to finely compare results among EPs, the average relative performances (RPs) based on selected performance estimates were calculated for each individual and also each family. RPs for each individual were calculated by dividing the individual trait value by the sum of the trait values for the whole dataset. RPs for each family were calculated by summing the individual trait values in a given family and dividing this total by the sum of the trait values for the whole dataset. RPs were used to properly compare plant ranking using the different performance estimates by performing Wilcoxon tests for paired data. Tests were performed with R 3.3.2 (R Core Team, 2016), with packages "factoMineR" and "ggplot2".

Results

Correlations among traits

PCA performed on wild plants showed a clear separation between vegetative and reproductive traits (Figure 1). The first component represented 44.8% of the phenotypic variance (Figure 1A,D) and reproductive traits contributed most to that dimension (Figure 1B). The second component represented 20.3% of the phenotypic variance (Figure 1A,D) and was explained mostly by vegetative traits (Figure 1C). The projection of individuals on the first factorial plane was structured mainly along the first dimension according to edaphic origin (M vs NM), with NM individuals showing greater values of reproductive traits, but slightly lower values of vegetative traits (Figure 1E and Figure 2). Pearson's correlations confirmed that reproductive and vegetative traits were slightly, but significantly, negatively correlated (Figure S2). Accordingly, the traits showing the strongest positive correlations with seed set were NS ($r = 0.93$, $p < 0.001$), NNAS ($r = 0.94$, $p < 0.001$) and totalLS ($r = 0.92$, $p < 0.001$).

PCA on inbred and outbred progenies showed an even stronger independence between vegetative and reproductive traits (Figure 3). The first component represented 38% of the phenotypic variance (Figure 3A,D) and was explained mainly by reproductive traits (Figure 3B). The second component represented 25% of the phenotypic variance (Figure 3A,D) and was explained mainly by vegetative traits (Figure 3C). Edaphic origin did not discriminate individual plants in the first factor plane as strongly as for wild plants. However, individuals were discriminated according to zinc pollution level (Figure 3F). The effect of pollution was also visible on trait values, in particular for M self and outbred crosses (Figure 4). Independence between reproductive and vegetative traits was confirmed by Pearson's correlations (Figure S3). As with wild plants, the traits most correlated with seed set were NS ($r = 0.71$, $p < 0.001$), NNAS ($r = 0.83$, $p < 0.001$) and totalLS ($r = 0.85$, $p < 0.001$).

Thus, whatever the dataset used, seed set was highly correlated with NS, NNAS (both being very highly correlated, Figures S2 and S3) and totalLS. Therefore, linear regressions were performed between SS and NNAS on the one hand, and between SS and totalLS on the other (Table 1, datasets A and B); all showed quite high R^2 values ($0.68 < R^2 < 0.87$). NNAS and totalLS were then selected as the best estimates for plant performance.

In addition, regardless of the dataset, a significant but moderate positive correlation was obtained between the length of a silicle and the number of seeds in that silicle, with low to moderate R^2 values ($0.20 < R^2 < 0.46$, see Table 1, datasets C and D).

Source of phenotypic variance in seed set and plant performance estimates (NNAS and totalLS)

For the 72 wild plants, significant differences were observed mostly between edaphic origins for seed set (F=16, df=1,61, P=0.024), NNAS (F=11, df=1,61, P=0.039) and totalLS (F=12, df=1,61, P=0.037), with higher values for plants of non-metallicolous origin (SS= 3873.6 ± 1872.1, NNSA= 397.9 ± 200.6, totalLS= 2834 ± 1569) compared with those of metallicolous origin (SS= 1297.8 ± 1278.2, NNAS= 163.3 ± 158.9, totalLS= 902.7 ± 917.3) (Figure 2). The effect of population within edaphic origin was slightly significant only for totalLS (F=2.8, df=3,61, P=0.049) compared with seed set (F=2.4, df=3,61, P=0.08) and NNAS (F=2.3, df=3,61, P=0.09), but *post hoc* tests did not support this difference (Figure 2F).

For self and outbred progenies, the seed set and its two estimates again showed the same sources of phenotypic variation, with significant differences obtained mainly according to contamination level (Table 2). Overall, individuals grown on polluted soil produced higher numbers of seeds, more non-aborted silicles and displayed higher total length of silicles (SS = 1604.5 ± 1290.4, NNAS =421.1 ± 264.9, totalLS=1932.1 ± 1308.9) than individuals grown on non-polluted soil (SS = 953.6 ± 1056.8, NNAS = 253 ± 194.5, totalLS = 1140.2 ± 1041.3). A strong interaction between cross origin and pollution level, due mostly to the response of LC*LE cross (C5) in polluted conditions, was also observed (Table 2 and Figure S4). Among self-crosses, those from non-metallicolous origin showed higher traits values than crosses from metallicolous origin in non-polluted conditions, in particular for seed set (Figure 4E). In contrast, trait values for crosses from metallicolous origin were largely increased by zinc pollution (Figure 4C–E).

Application of performance estimates to experiment populations

PCA seemed to show a slight tendency towards positive correlations between vegetative and reproductive traits. However, Pearson's correlation coefficients corrected for multiple tests did not confirm such positive correlations (Figure S5). The two first axes of the PCA explained 63% of the variance (Figure 5A,D), with the first axis more represented by vegetative variables and the second axis more represented by reproductive variables (Figure 5B,C). A projection of the individuals according to these axes was performed, showing no clear discrimination according to pollution level, but suggesting more variability among individuals cultivated in non-polluted conditions (Figure 5E).

Comparing the observed ranks of RPs, no difference was detected between RP_{NNAS} and $RP_{totalLS}$, either by comparing EP1, EP2 and EP3 (Figure S6), or by comparing all the replicates (Figure S7). Only EP3.2 showed a marginal probability ($V = 797$, $p = 0.06691$) with a tendency for a better performance estimated through totalLS (Figure S7).

The families with the best and worst relative performance were the same in both contaminated EPs for NNAS as well as for totalLS (Figure 6 and Figure S8). Thus, some families (W103, W105, W106, W107, W109, W112 and L105) were present in the quartile of the highest performing individuals in the contaminated EPs (EP2 and EP3) but not necessarily in the non-contaminated EP (EP1). Similarly, several families (L102, L103, L114, L124) were present in the quartile of the worst performing plants in EP2 and EP3, but not necessarily in EP1.

Discussion

The relationship between vegetative and reproductive traits is weak

In plants, analysis of the ability to survive despite soil metal toxicity has focused mostly on vegetative traits, particularly root growth, and interpreted as plant survival (see, for example Pauwels et al., 2006). This could suggest that vegetative traits are a suitable proxy for the measurement of the performance of plants facing sublethal concentrations of metals in their environment. However, our results revealed no positive correlation among vegetative and reproductive traits. This was seen whatever the biological material used (i.e., wild plants, controlled crosses or experimental populations) and whatever the culture conditions (plastic tunnel, greenhouse or mesocosms; on polluted or unpolluted substrate). Indeed, our results suggest that, at least in *N. caerulescens*, reproductive effort is not correlated positively with vegetative growth at the population level, despite high variability of rosette sizes among individuals at the time of vegetative-to-reproductive phase transition. This argues against the use of vegetative traits alone to assess plant performance.

Our results contrast with reviews suggesting that a positive allometric relationship may exist between vegetative and reproductive biomass in plants (Weiner et al., 2009; Younginger et al., 2017). However, the literature on phenotypic integration, namely “the pattern of functional, developmental and/or genetic correlation among different traits in a given organism” (Pigliucci, 2003), puts forward an original viewpoint. To estimate the degree of phenotypic integration, traits can be gathered in “modules” of correlations (clusters of highly correlated traits). A comparative study of floral traits versus vegetative modules showed that correlation between modules could be very weak (Murren, 2002). In addition, Murren, Pendleton, & Pigliucci (2002) revealed that, in two out of seven *Brassica* species, none of the measured vegetative traits was related to fruit number (used as a measure of fitness). In the five remaining species, the measured vegetative traits that were related to fruit number were variable among species, and no general pattern could be identified.

Instead of a positive relationship, a slightly negative relationship between reproductive and vegetative traits was found in our study, in wild plants only, whereas no correlation at all was detected in control crosses. The projection of wild plants on the first PCA axis (Figure 1E) suggested that this could be explained by different developmental strategies between M and NM ecotypes (performing PCAs on M and NM individuals separately confirmed independence between reproductive and vegetative traits, Figures S9 and S10). Indeed, M individuals were characterized by higher vegetative biomass and lower reproductive biomass than NM individuals (Figure 2). Self M and NM progenies cultivated in non-polluted conditions gave similar results (Figure 4). In polluted conditions, however, this pattern was somehow modified, with increased vegetative and reproductive performance for M progenies only. Consequently, difference among M and NM plants was higher for vegetative traits, but lower (always non-significant) for reproductive traits. Such contrasted developmental strategies in interaction with levels of metal exposure had already been suggested in previous studies in which metallicolous and non-metallicolous populations of *N. caerulescens* were cultivated in non-polluted or low-polluted conditions (as determined by following total seed production and growth in Jiménez-Ambriz *et al.* 2007, and by following number of fertile fruits and number of leaves in Dechamps *et al.* 2008). Dechamps et al. (2008) also observed that the fitness of metallicolous individuals increased in polluted soils. In contrast, Dechamps et al. (2007) and Jiménez-Ambriz et al. (2007) observed maintenance of vegetative and reproductive performances of metallicolous populations despite increasing levels of Zn pollution in soil, whereas performances of non-metallicolous populations decreased; in particular, reproductive performances were qualified as more sensitive to Zn pollution (Dechamps et al., 2007).

Contrasted developmental strategies between M and NM ecotypes, together with a response of vegetative and reproductive performance traits to variation in levels of Zn exposure, suggest that changes in allometric relationships between vegetative and reproductive traits may be involved in local adaptation to metal pollution. In the traditional perspective of plant resource allocation, a positive relationship between vegetative and reproductive biomass results from the fact that vegetative biomass mostly represents an energy resource to produce reproductive structures. Our results suggest that the energy stored in vegetative biomass may not be fully available for reproduction allocation, but that increased biomass may be selected *per se* with no direct quantitative effect on reproductive output. Under environmental constraints, it is reasonable to assume that increasing biomass may first ensure survival, and then minimum reproduction. Although costly in terms of allocation economy, such a strategy would be preferable to "death without sex" (Aarssen, 2008). *N. caerulescens* is a metal-hyperaccumulator species that store metals at high concentration in its shoots. In such species, increased vegetative biomass in metallicolous populations may have evolved as an adaptation to metal toxicity, allowing a dilution of metals in shoots, and thus detoxification, assuming that reduced shoot metal concentration may increase tolerance to soil metal (Nowak et al., 2018).

Fecundity is best estimated using fruit traits

Because vegetative and reproductive traits were not strongly correlated, fitness—expressed here as seed set—could not be estimated by following vegetative traits. In our study, seed set was related strongly and positively to the NNAS, and to the product of the number of non-aborted silicles and the mean length of the silicles (totalLS). The latter estimate was similar to the “total silique length” established for *A. thaliana* (Roux et al., 2004; Vasseur et al., 2018). Roux et al (2004) indeed demonstrated that seed set was linearly correlated with the product of silique mean length by silique number. In addition, the partitioning of variance for seed set, NNAS and totalLS, was largely consistent (Tables 2 and 3). Such results are in accordance with methodological choices for fecundity estimation in *N. caerulescens* used in previous studies, involving measuring the number of fertile fruits, multiplying the number of fruits per plant by the mean number of seeds per fruit, or by calculating the product of flowering stalk number, fruit number per flowering stalk and seed number per fruit (Dechamps et al., 2008, 2007; Jiménez-Ambriz et al., 2007). However, none of these studies properly assessed the relationship between seed set and the chosen estimates.

Our combined estimate (totalLS) relies on the hypothesis of a positive correlation between silicle length and the number of seeds in a silicle, as is the case in *Arabidopsis thaliana* (Roux et al., 2004). In our case, this correlation was significant but weak, with a large part of variance remaining unexplained (Table 1, datasets C and D). A plausible explanation, for a given silicle length, may be that some ovules could be not pollinated or some seeds could have aborted. Thus, even if the parts of the variance explained by NNAS and totalLS were similar (Table 1, datasets A and B), taking into account silicle length may imply additional sources of variability. However, if a choice must be made, we would suggest using a fitness estimate like totalLS, which is based on both fruit and seed traits. Indeed, although this should be investigated further, totalLS should be closer to the number of offspring. In addition, it has been demonstrated recently in *A. thaliana* that considering the number of fruits only for fitness estimation, instead of including seed production, could lead to biased fitness estimation and underestimation of adaptive differentiation between ecotypes (Ellis et al., 2021).

Interestingly, Dechamps et al. (2011) suggested that, *in natura*, fecundity (estimated as seed output per plant, itself estimated from a relationship previously established with the cumulated length of flowering stems) was significantly increased with rosette size. Nevertheless, the authors divided rosette size into categories (“small”, “medium”, “large”) and assumed that each category of rosette size was associated to different life cycles (monocarpic winter annual, monocarpic perennial, polycarpic perennial). In contrast, in our experimental conditions, all plants exhibited a winter annual monocarpic life cycle, as seed germination was realized in autumn and rosettes showed senescence at the end of the reproductive period. This life cycle may be, at least partly, genetically and not only environmentally induced, because Dechamps et al. (2011) showed that, even in the field conditions of the Belgium/Luxembourg region, 95% to 100% of plants exhibited a monocarpic life cycle in all populations studied. In addition, these authors also showed evidence that seedling recruitment peaked at the beginning of autumn; thus, our experimental conditions were consistent with the natural tendency.

Therefore, even though *N. caerulescens* individuals were cultivated in non-lethal conditions, variations in plant performance could be observed through appropriate fecundity estimates. Such sublethal, rather than lethal, effects are probably due to the fact that *N. caerulescens* is adapted to metal toxicity, as can be the case for any ecotoxicological assessments from individuals sampled in the field and thus potentially adapted to contaminated environments (Pauwels et al., 2013).

Application of fitness estimates to experimental populations

Laboratory Natural Selection experiments can be used either to empirically "scrutinize fundamental predictions of evolutionary theory" or to "test hypotheses about the selective agents responsible for maintaining particular patterns of variation in nature" (Fuller et al., 2005). In the latter, the approach consists of isolating a putative selective agent and testing its capacity to provoke adaptive evolution of related traits, comparing exposed and control populations. In the context of evolutionary ecotoxicology, LNS is an original approach that should allow the impact of a long-term contaminant exposure to be properly tested.

In LNS experiments, a major challenge consists of building successive generations representing performance variation among parental individuals, in particular if population size or density must not vary in time. Then, experimenters can choose between either randomly selecting a subset of seed from the entire set of seeds produced by all parental individuals, hoping that sampling will be representative of fitness variation among parents, or, to limit random effects, adjusting the relative contribution of each parent to its relative fitness. The latter obviously requires being able to estimate relative fitness, i.e. to measure individual performance.

To test the ability of our estimators (NNAS and totalLS) to be used in LNS approaches, we benefit from the setup of an LNS experiment involving *N. caerulescens* to test whether soil Zn could act as a directional selective pressure in the evolution of metal-related traits (see Nowak et al., 2018 for a detailed presentation of the scientific hypotheses). Absolute NNAS and totalLS values of parental plants of the first generation of the experiment were calculated and related to the sum of values of the corresponding EP to assess relative fitness. The objective was to confirm that we were able to evidence differences in individual fitness, without which no adaptive evolution can be expected. Both relative performance estimates (RP_{NNAS} and $RP_{totalLS}$) were ranked equally, suggesting that they provided similar information (Figures S6 and S7). For both estimates, the range of values was higher in the experimental population exposed to Zn pollution (EP2 and EP3) than in the control population (EP1). In addition, the same families were distinguished as showing either the highest or the lowest performance values in EP2 and EP3, whereas they were mixed in an unordered way with other families in EP1 (Figure 6 and Figure S8). This suggests that both estimates revealed nonrandom variation in plant performance in EP2 and EP3, corresponding to what is expected if selection favors the fittest genotypes and provokes directional evolution. Therefore, we conclude that both estimates are suitable as fitness estimates to build the following generation. Then, they just need to be multiplied by the desired populations size (in this study: 98) to calculate an expected number of descendants (Nowak et al., 2018). More specifically, we highlight that special attention should be paid to the consequences of the contrasted developmental strategies between metallicolous and non-metallicolous *N. caerulescens* populations we evidenced in our study (see above). Our LNS experiment involved only NM progenies. Any experiment including individuals from metallicolous and non-metallicolous populations should keep in mind that M individuals tend to show comparatively lower reproductive performance than NM populations. This precludes direct comparison of reproductive outputs among ecotypes.

Being able to estimate individual fitness differences is a prerequisite for evolutionary experiments seeking to limit the effect of chance events, but it does not ensure that adaptive evolution will occur accordingly in subsequent generations. Several phenomena may occur during the transition between seed production, seedling emergence, and offspring survival and reproduction that may result in an evolution that would not be explained by difference in seed sets of parental individuals. For example, producing more seeds may results in lower seed quality and germination rate, higher competition among seedlings and lower survival of offspring, so that differences in individual fitness differ from those calculated from differences in seed set only (Campbell et al., 2017, and references therein).

Conclusion

We showed that differences among the fitness of individual plants could be assessed accurately by following easy-to-track reproductive traits, such as the number of fruits and fruit length. This could be very helpful for experiments that intend to better understand the ecotoxicological effects of environmental contaminants on the evolution of populations. Following stressed experimental populations over several generations should confirm that fitness differences observed at the stage of seed production are a good predictor of population evolution. Following traits at additional stages of the life cycle, such as seed germination, seedling survival, etc., could improve fitness estimation. Such a multi-generational LNS experiment would also address the question of whether the observed evolution is due to transgenerational plasticity or local adaptation. For this purpose, using a fitness estimator like that established in this study in a dedicated experiment allowing parent-offspring comparisons would be helpful.

Funding

This research was supported by the French National Research Agency (ANR ELOCANTH, grant number ANR-12-JSV7-0010). Julien Nowak was funded by French Ministry of Higher Education, Research and Innovation.

Acknowledgements

The authors thank the Région Hauts-de-France and the European Fund for Regional Economic Development for their financial support (Contrat de plan Etat-Région “Climibio”).

Conflict of interest

The authors have no conflict of interest to declare.

References

Aarssen, L.W., 2008. Death without sex—the ‘problem of the small’ and selection for reproductive economy in flowering plants. Evol. Ecol. 22, 279–298. https://doi.org/10.1007/s10682-007-9170-z

Arnold, S.J., 1983. Morphology, Performance and Fitness1. Am. Zool. 23, 347–361. https://doi.org/10.1093/icb/23.2.347

Assunção, A.G.L., Ten Bookum, W.M., Nelissen, H.J.M., Vooijs, R., Schat, H., Ernst, W.H.O., 2003. A cosegregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator Thlaspi caerulescens. New Phytol. 159, 383–390. https://doi.org/10.1046/j.1469-8137.2003.00758.x

Brady, S.P., Richardson, J.L., Kunz, B.K., 2017. Incorporating evolutionary insights to improve ecotoxicology for freshwater species. Evol. Appl. 10, 829–838. https://doi.org/10.1111/eva.12507

Campbell, D.R., Brody, A.K., Price, M. V, Waser, N.M., Aldridge, G., 2017. Is Plant Fitness Proportional to Seed Set? An Experiment and a Spatial Model. Am. Nat. 190, 818–827. https://doi.org/10.1086/694116

Carlson, S.M., Cunningham, C.J., Westley, P.A.H., 2014. Evolutionary rescue in a changing world. Trends Ecol. Evol. 29, 521–530. https://doi.org/10.1016/j.tree.2014.06.005

Dechamps, C., Elvinger, N., Meerts, P., Lefèbvre, C., Escarré, J., Colling, G., Noret, N., 2011. Life history traits of the pseudometallophyte Thlaspi caerulescens in natural populations from Northern Europe. Plant Biol. 13, 125–135. https://doi.org/10.1111/j.1438-8677.2010.00387.x

Dechamps, C., Lefèbvre, C., Noret, N., Meerts, P., 2007. Reaction norms of life history traits in response to zinc in Thlaspi caerulescens from metalliferous and nonmetalliferous sites. New Phytol. 173, 191–198. https://doi.org/10.1111/j.1469-8137.2006.01884.x

Dechamps, C., Noret, N., Mozek, R., Escarré, J., Lefèbvre, C., Gruber, W., Meerts, P., 2008. Cost of adaptation to a metalliferous environment for Thlaspi caerulescens: a field reciprocal transplantation approach. New Phytol. 177, 167–177. https://doi.org/10.1111/j.1469-8137.2007.02245.x

Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneth, A., Balvanera, P., Brauman, K.A., Butchart, S.H.M., Chan, K.M.A., Lucas, A.G., Ichii, K., Liu, J., Subramanian, S.M., Midgley, G.F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razzaque, J., Reyers, B., Chowdhury, R.R., Shin, Y.J., Visseren-Hamakers, I., Willis, K.J., Zayas, C.N., 2019. Pervasive human-driven decline of life on Earth points to the need for transformative change. Science (80-.). 366. https://doi.org/10.1126/science.aax3100

Ellis, T.J., Postma, F.M., Oakley, C.G., Ågren, J., 2021. Life-history trade-offs and the genetic basis of fitness in Arabidopsis thaliana. Mol. Ecol. n/a. https://doi.org/https://doi.org/10.1111/mec.15941

Escarré, J., Lefèbvre, C., Gruber, W., Leblanc, M., Lepart, J., Rivière, Y., Delay, B., 2000. Zinc and cadmium hyperaccumulation by Thlaspi caerulescens from metalliferous and nonmetalliferous sites in the Mediterranean area: implications for phytoremediation. New Phytol. 145, 429–437. https://doi.org/10.1046/j.1469-8137.2000.00599.x

Frank, S.A., 2011. Natural selection. I. Variable environments and uncertain returns on investment. J. Evol. Biol. 24, 2299–2309. https://doi.org/10.1111/j.1420-9101.2011.02378.x

Fuller, R.C., Baer, C.F., Travis, J., 2005. How and when selection experiments might actually be useful. Integr. Comp. Biol. 45, 391–404. https://doi.org/10.1093/icb/45.3.391

Gonneau, C., Genevois, N., Frérot, H., Sirguey, C., Sterckeman, T., 2014. Variation of trace metal accumulation, major nutrient uptake and growth parameters and their correlations in 22 populations of Noccaea caerulescens. Plant Soil 384, 271–287. https://doi.org/10.1007/s11104-014-2208-4

Haldane, J.B.S., 1937. The Effect of Variation of Fitness. Am. Nat. 71, 337–349. https://doi.org/10.1086/280722

Hartl, D.L., 1989. Principles of population genetics / Daniel L. Hartl, Andrew G. Clark. Sinauer Associates, Sunderland, Mass.

Hereford, J., Elle, A.E.E., Geber, E.M.A., 2009. A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. Am. Nat. 173, 579–588. https://doi.org/10.1086/597611

Irschick, D.J., Meyers, J.J., Husak, J.F., Le Galliard, J.F., 2008. How does selection operate on whole-organism functional performance capacities? A review and synthesis. Evol. Ecol. Res. 10, 177–196.

Jiménez-Ambriz, G., Petit, C., Bourrié, I., Dubois, S., Olivieri, I., Ronce, O., 2007. Life history variation in the heavy metal tolerant plant Thlaspi caerulescens growing in a network of contaminated and noncontaminated sites in southern France: Role of gene flow, selection and phenotypic plasticity. New Phytol. 173, 199–215. https://doi.org/10.1111/j.1469-8137.2006.01923.x

Kleyer, M., Trinogga, J., Cebrián-Piqueras, M.A., Trenkamp, A., Fløjgaard, C., Ejrnæs, R., Bouma, T.J., Minden, V., Maier, M., Mantilla-Contreras, J., Albach, D.C., Blasius, B., 2019. Trait correlation network analysis identifies biomass allocation traits and stem specific length as hub traits in herbaceous perennial plants. J. Ecol. 107, 829–842. https://doi.org/https://doi.org/10.1111/1365-2745.13066

Krimbas, C.B., 2004. On fitness. Biol. Philos. 19, 185–203. https://doi.org/10.1023/B:BIPH.0000024402.80835.a7

Lailvaux, S.P., Husak, J.F., 2014. The life history of whole-organism performance. Q. Rev. Biol. 89, 285—318. https://doi.org/10.1086/678567

Mayer-Pinto, M., Ledet, J., Crowe, T.P., Johnston, E.L., 2020. Sublethal effects of contaminants on marine habitat-forming species: a review and meta-analysis. Biol. Rev. 95, 1554–1573. https://doi.org/https://doi.org/10.1111/brv.12630

Meerts, P., Van Isacker, N., 1997. Heavy metal tolerance and accumulation in metallicolous and non-metallicolous populations of Thlaspi caerulescens from continental Europe. Plant Ecol. 133, 221–231. https://doi.org/10.1023/A:1009717619579

Murren, C.J., 2002. Phenotypic integration in plants. Plant Species Biol. 17, 89–99. https://doi.org/10.1046/j.1442-1984.2002.00079.x

Murren, C.J., Pendleton, N., Pigliucci, M., 2002. Evolution of phenotypic integration in Brassica (Brassicaceae). Am. J. Bot. 89, 655–663. https://doi.org/10.3732/ajb.89.4.655

Nowak, J., Frérot, H., Faure, N., Glorieux, C., Liné, C., Pourrut, B., Pauwels, M., 2018. Can zinc pollution promote adaptive evolution in plants? Insights from a one-generation selection experiment. J. Exp. Bot. 69, 5561–5572. https://doi.org/10.1093/jxb/ery327

Orr, H.A., 2009. Fitness and its role in evolutionary genetics. Nat. Rev. Genet. 10, 531–9. https://doi.org/10.1038/nrg2603

Pauwels, M., Frérot, H., Bonnin, I., Saumitou-Laprade, P., 2006. A broad-scale analysis of population differentiation for Zn tolerance in an emerging model species for tolerance study: Arabidopsis halleri (Brassicaceae). J. Evol. Biol. 19, 1838–1850. https://doi.org/10.1111/j.1420-9101.2006.01178.x

Pauwels, M., Frérot, H., Souleman, D., Vandenbulcke, F., 2013. Using biomarkers in an evolutionary context: Lessons from the analysis of biological responses of oligochaete annelids to metal exposure. Environ. Pollut. 179, 343–350. https://doi.org/10.1016/j.envpol.2013.05.005

Pence, C.H., Ramsey, G., 2015. Is Organismic Fitness at the Basis of Evolutionary Theory? Philos. Sci. 82, 1081–1091. https://doi.org/10.1086/683442

Pigliucci, M., 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Ecol. Lett. 6, 265–272. https://doi.org/10.1046/j.1461-0248.2003.00428.x

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput.

Reeves, R.D., Schwartz, C., Morel, J.L., Edmondson, J., 2001. Distribution and metal-accumulating behavior of Thlaspi caerulescens and associated metallophytes in France. Int. J. Phytoremediation 3, 145–172. https://doi.org/10.1080/15226510108500054

Roux, F., Gasquez, J., Reboud, X., 2004. The dominance of the herbicide resistance cost in several Arabidopsis thaliana mutant lines. Genetics 166, 449–460. https://doi.org/10.1534/genetics.166.1.449

Ruttens, A., Mench, M., Colpaert, J. V., Boisson, J., Carleer, R., Vangronsveld, J., 2006. Phytostabilization of a metal contaminated sandy soil. I: Influence of compost and/or inorganic metal immobilizing soil amendments on phytotoxicity and plant availability of metals. Environ. Pollut. 144, 524–532. https://doi.org/10.1016/j.envpol.2006.01.038

Sterckeman, T., Cazes, Y., Gonneau, C., Sirguey, C., 2017. Phenotyping 60 populations of Noccaea caerulescens provides a broader knowledge of variation in traits of interest for phytoextraction. Plant Soil 418, 523–540. https://doi.org/10.1007/s11104-017-3311-0

Straub, L., Strobl, V., Neumann, P., 2020. The need for an evolutionary approach to ecotoxicology. Nat. Ecol. Evol. 4, 895. https://doi.org/10.1038/s41559-020-1194-6

Tovar-Sánchez, E., 2018. Heavy Metal Pollution as a Biodiversity Threat, in: Hernández-Plata, I. (Ed.), . IntechOpen, Rijeka, p. Ch. 21. https://doi.org/10.5772/intechopen.74052

Vasseur, F., Bresson, J., Wang, G., Schwab, R., Weigel, D., 2018. Image-based methods for phenotyping growth dynamics and fitness components in Arabidopsis thaliana. Plant Methods 14, 63. https://doi.org/10.1186/s13007-018-0331-6

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., Garnier, E., 2007. Let the concept of trait be functional! Oikos 116, 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x

Weiner, J., Campbell, L.G., Pino, J., Echarte, L., 2009. The allometry of reproduction within plant populations. J. Ecol. 97, 1220–1233. https://doi.org/10.1111/j.1365-2745.2009.01559.x

Younginger, B.S., Sirová, D., Cruzan, M.B., Ballhorn, D.J., 2017. Is Biomass a Reliable Estimate of Plant Fitness? Appl. Plant Sci. 5, 1600094. https://doi.org/10.3732/apps.1600094

Table 1: Linear regressions performed with seed set (SS) as dependent variable (Y), and either the number of non-aborted silicles (NNAS), the estimated total length of silicles (totalLS) or the silicle length (SL) as independent variable (X). The results of the t-test refer to the significance of the slope. t= Student statistics, df= degrees of freedom.

Dataset	**A) Wild plants (n = 72 plants)**		**B) Progenies (n = 360 plants)**		**C) EP1.2 (n = 236 silicles)**	**D) Wild plants (n = 140 silicles)**
Variables	X=NNAS; Y=SS	X=totalLS; Y=SS	X=NNAS; Y=SS	X=totalLS; Y=SS	X=SL; Y=SS	X=SL; Y=SS
Adjusted R^2	0.876	0.849	0.686	0.724	0.458	0.201
Intercept	149.332	480.647	65.723	155.942	-4.114	-0.521
Slope	8.985	1.159	3.972	0.817	1.585	1.018
t-test results	t=21.42, df=64, $p<0.0001$	t=19.145, df=64, $p<0.0001$	t=26.402, df=318, $p<0.0001$	t=28.916, df=317, $p<0.0001$	t=14.137, df=234, $p<0.0001$	t=6.005, df=138, $p<0.0001$

Table 2: Analyses of variance for seed set (SS), number of non-aborted silicles (NNAS) and total length of silicles (totalLS) in the self progenies and outbred progenies of *Noccaea caerulescens* subjected to two pollution levels (non-polluted soil and soil polluted with 750 mg.kg^{-1} of zinc sulfate).

	SS				NNAS				(totalLS)			
	df	MS	F	**P value**	df	MS	F	**P value**	df	MS	F	**P value**
Cross type	2	18361847	1.28	**0.3341**	2	710885	3.15	**0.0956**	2	16908453	2.57	**0.1389**
Pollution	1	28728797	7.68	**0.0253**	1	1382836	9.11	**0.0151**	1	32145693	8.05	**0.0204**
Cross type*Pollution	2	13917302	1.83	**0.2250**	2	695561	2.25	**0.1634**	2	12824083	1.58	**0.2609**
Cross origin within Cross type	8	66627471	1.33	**0.2937**	8	1112259	0.62	**0.7439**	8	33539490	0.64	**0.7353**
Cross pedigree within Cross origin	14	59019240	4.18	**0.0036**	14	1321379	2.38	**0.0632**	14	440992859	3.04	**0.0265**
Pollution*(Cross origin within Cross type)	8	28195048	3.71	**0.0179**	8	1367928	4.32	**0.0128**	8	36169031	4.35	**0.0124**
Pollution*(Cross pedigree within Cross origin)	12	12352906	1.09	**0.3701**	12	474937	0.96	**0.4841**	12	12451695	1.06	**0.3922**
Residuals	272	266771249			272	11171890			271	264656369		

Figure legends

Figure 1: Results of the principal component analysis performed on the 72 wild plants including the correlation circle based on dimension 1 and dimension 2 (A), the variable contributions to dimension 1 (B), the variable contribution to dimension 2 (C), the scree plot of eigenvalues (D) and the projection of individuals based on dimension 1 and dimension 2 according to the edaphic origin of plants (E). M=metallicolous; NM=non-metallicolous. Vegetative traits: SP=surface of the plant, LL=leaf length, LW=leaf width, LS=leaf shape, LA=leaf area, NL=number of leaves. Reproductive traits: SS=seed set, NFS=number of flower stems, maxLFS=length of the longest stem, NS=number of silicles, NAS=number of aborted silicles, NNAS=number of non-aborted silicles, RS=ratio between NNAS and NAS, SL=silicle length, totalLS=total length of silicles. Ellipse surfaces grouped 90% of individuals for each origin. For B and C, the dotted line represents the contribution if all the variables contributed equally.

Figure 2: Boxplots and statistical results for some vegetative (A, B, C, D) and reproductive (E, F, G) traits in the 72 wild plants from the two metalliferous sites (red boxes) and the three non-metallicolous sites (blue boxes). Vegetative traits: SP=surface of the plant, LS=leaf shape, LA=leaf area, NL=number of leaves. Reproductive traits: SS=seed set, NNAS=number of non-aborted silicles, totalLS=total length of silicles. The statistical results above each graph correspond to those of the Kruskal-Wallis test. K= Kruskal-Wallis statistics, df= degree of freedom, p= p-value. Letters above each boxplot indicate significant differences at the 5% threshold for the *post hoc* Conover test.

Figure 3: Results of the principal component analysis performed on self- and outbred progenies including the correlation circle based on dimension 1 and dimension 2 (A), the variable contributions to dimension 1 (B), the variable contribution to dimension 2 (C), the scree plot of eigenvalues (D), the projection of individuals based on dimension 1 and dimension 2 according to the cross type (E) and the projection of individuals based on the dimension 1 and dimension

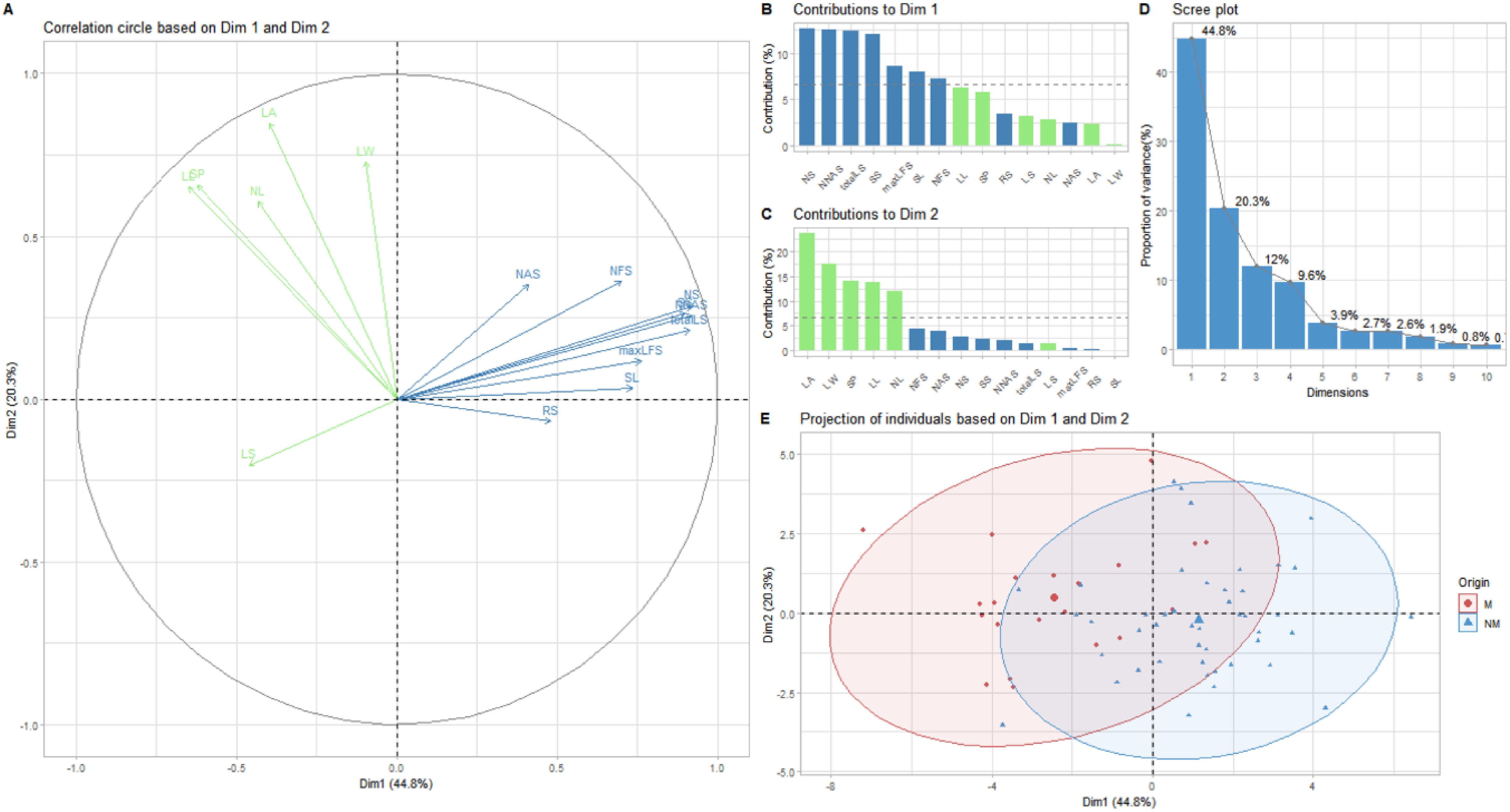
2 according to pollution level (F). M=metallicolous self-crosses; NM=non-metallicolous self-crosses; C=outbred crosses. Vegetative traits: SP=surface of the plant, LL=leaf length, LW=leaf width, LS=leaf shape, LA=leaf area, NL=number of leaves. Reproductive traits: SS=seed set, NFS=number of flower stems, maxLFS=length of the longest stem, NS=number of silicles, NAS=number of aborted silicles, NNAS=number of non-aborted silicles, RS=ratio between NNAS and NAS, SL=silicle length, totalLS=total length of silicles. Ellipse surfaces grouped 90% of individuals for each origin and each pollution level. For B and C, the dotted line represents the contribution if all the variables contributed equally.

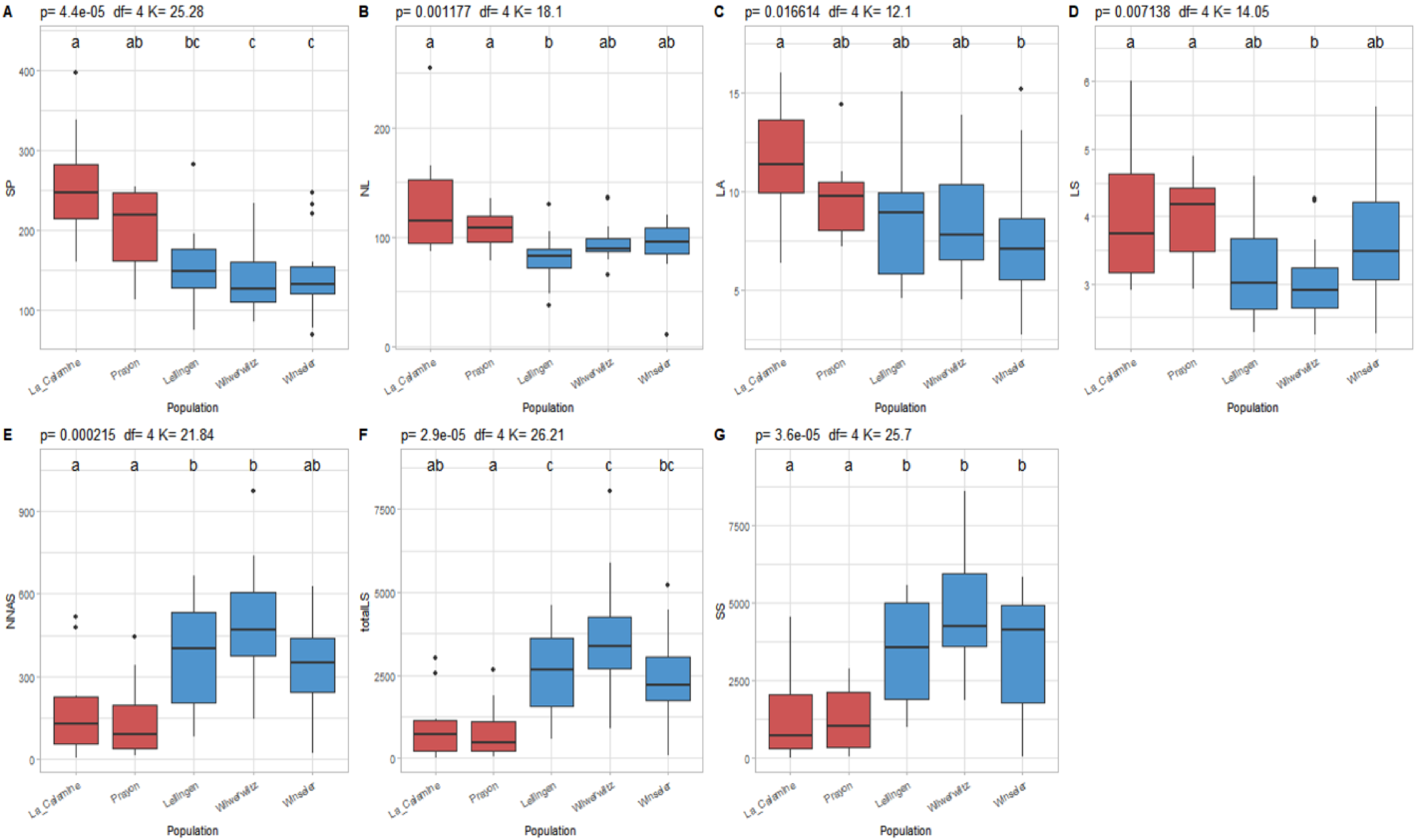
Figure 4: Boxplots and statistical results for some vegetative (A, B) and reproductive (C, D, E) traits in self- and outbred progenies in non-polluted (blue boxes) and polluted (red boxes) soils. Vegetative traits: SP=surface of the plant, NL=number of leaves. Reproductive trait: NNAS=Number of non-aborted silicles, totalLS=total length of silicles, SS=seed set. The statistical results above each graph correspond to those of the Kruskal-Wallis test for the non-polluted soil (blue boxes) and the polluted soil (red boxes). K= Kruskal-Wallis statistics, df= degree of freedom, p= p-value.

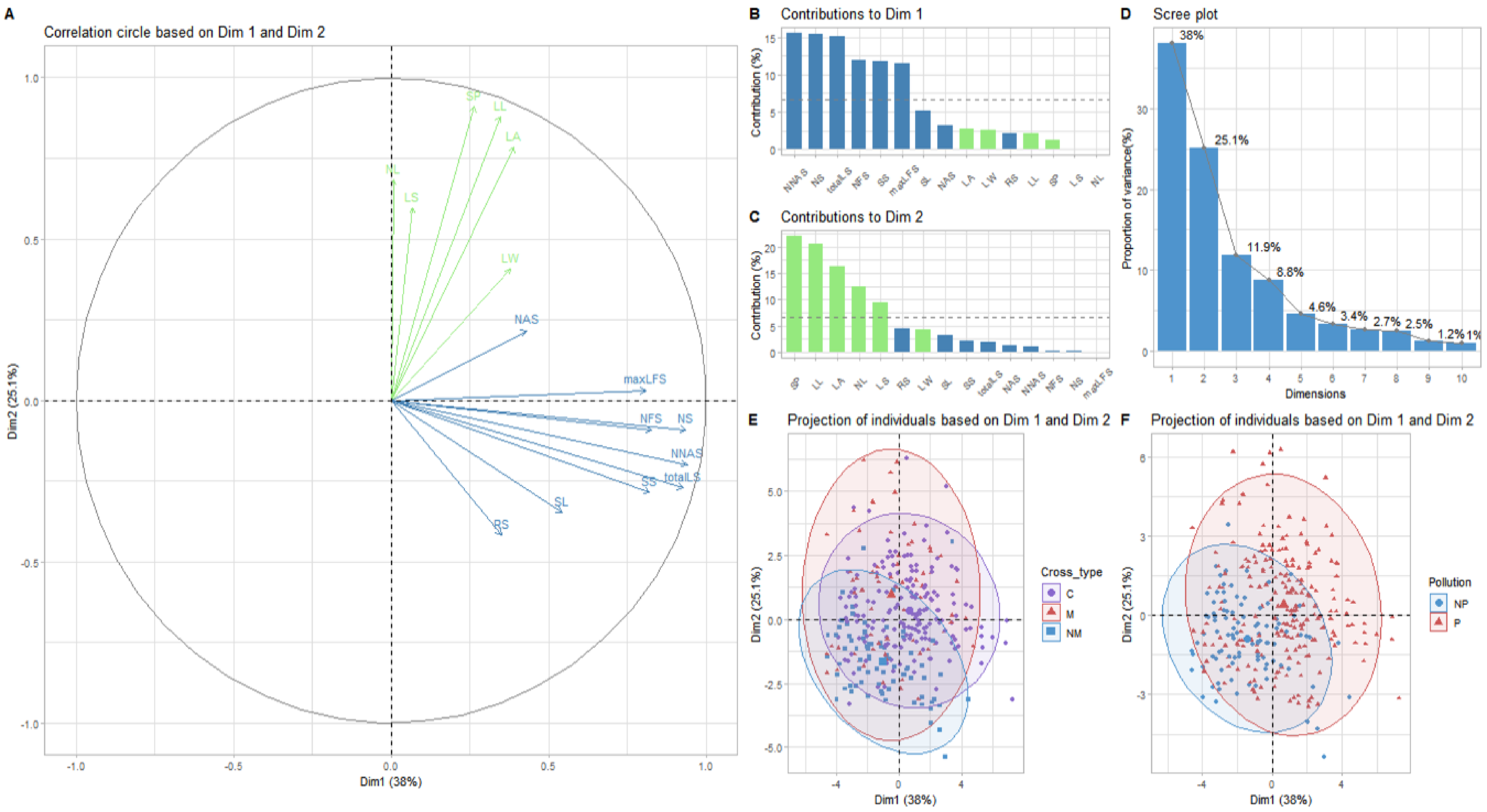
Figure 5: Results of the principal component analysis performed on experimental populations including the correlation circle based on dimension 1 and dimension 2 (A), the variable contributions to dimension 1 (B), the variable contribution to dimension 2 (C), the scree plot of eigenvalues (D), and the projection of individuals based on dimension 1 and dimension 2 according to pollution level (E). P=polluted condition (EP2 and EP3); NP=non-polluted condition (EP1). Vegetative traits: SP=surface of the plant, LL=leaf length, LW=leaf width, LS=leaf shape, LA=leaf area, NL=number of leaves. Reproductive traits: SS=seed set, NFS=number of flower stems, maxLFS=length of the longest stem, NS=number of silicles, NAS=number of aborted silicles, NNAS=number of non-aborted silicles, RS=ratio between NNAS and NAS, SL=silicle length, totalLS=total length of silicles. Ellipse surfaces grouped

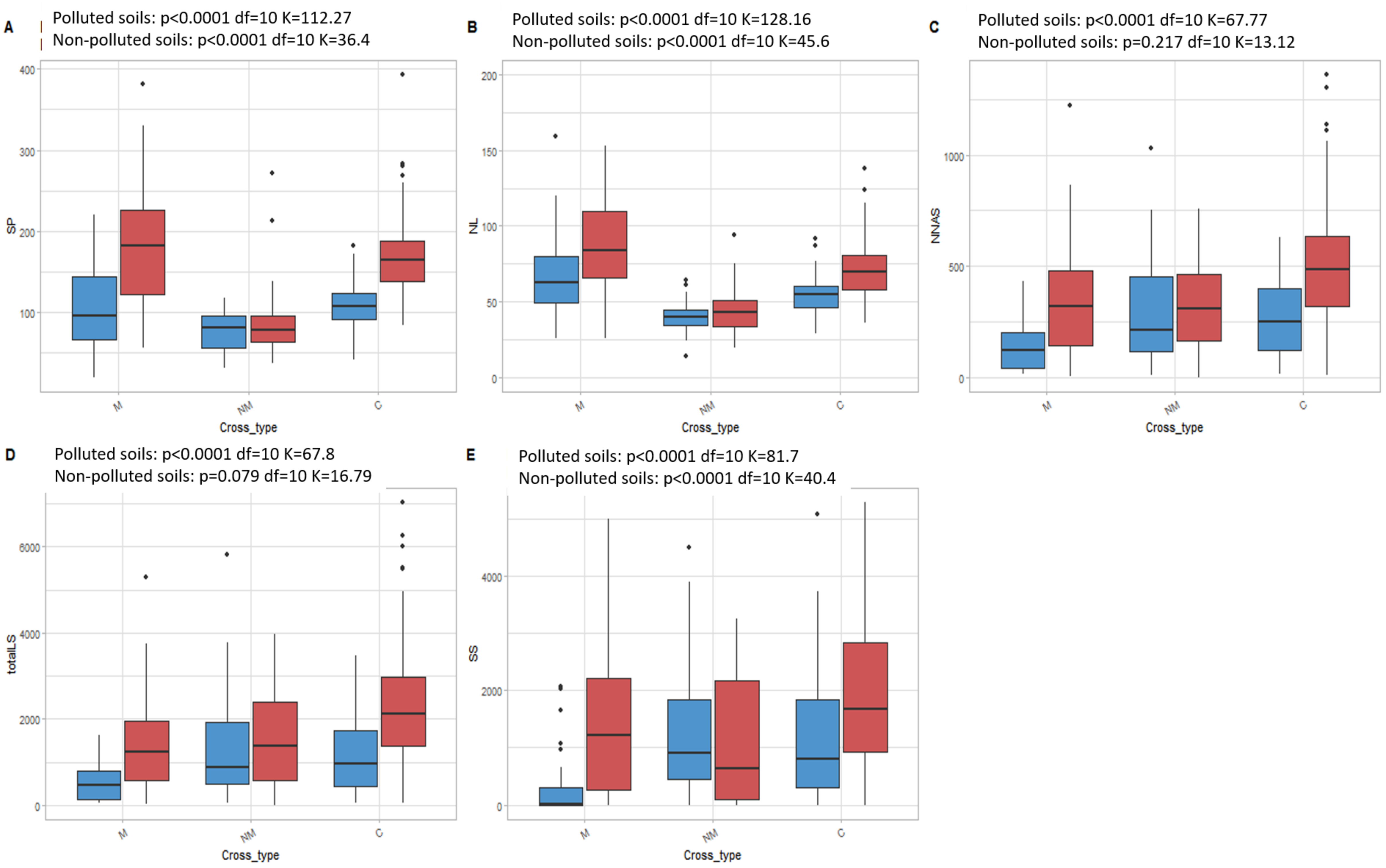
90% of individuals for each pollution level. For B and C, the dotted line represents the contribution if all variables contributed equally.

Figure 6: Rank comparison between relative performances based on the estimated total length of silicles (*totalLS)* calculated from the mean of each family in EP1, EP2 and EP3. Blue points represent families of the first quartile of the most successful plants present in the two polluted EPs (EP2 and EP3). Yellow points represent families of the last quartile of the less successful plants present in the two polluted EPs (EP2 and EP3). Grey points represent the other families.


A
Correlation circle based on Dim 1 and Dim 2
Dim1 (44.8%)
Dim2 (20.3%)
-1.0
-0.5
0.0
0.5
1.0
LA
LW
LL
SP
NL
LS
NAS
NFS
NS
NNAS
SS
totalLS
maxLFS
SL
RS
B
Contributions to Dim 1
Contribution (%)
0
5
10
NS
NNAS
totalLS
SS
maxLFS
SL
NFS
LL
SP
RS
LS
NL
NAS
LA
LW
C
Contributions to Dim 2
Contribution (%)
0
5
10
15
20
LA
LW
SP
LL
NL
NFS
NAS
NS
SS
NNAS
totalLS
LS
maxLFS
RS
SL
D
Scree plot
Proportion of variance(%)
Dimensions
0
10
20
30
40
1
2
3
4
5
6
7
8
9
10
44.8%
20.3%
12%
9.6%
3.9%
2.7%
2.6%
1.9%
0.8%
0.7
E
Projection of individuals based on Dim 1 and Dim 2
Dim1 (44.8%)
Dim2 (20.3%)
-8
-4
0
4
-5.0
-2.5
0.0
2.5
5.0
Origin
M
NM


A
p= 4.4e-05 df= 4 K= 25.28
a ab bc c c
SP
100
200
300
400
B
p= 0.001177 df= 4 K= 18.1
a a b ab ab
NL
0
100
200
C
p= 0.016614 df= 4 K= 12.1
a ab ab ab b
LA
5
10
15
D
p= 0.007138 df= 4 K= 14.05
a a ab b ab
LS
3
4
5
6
E
p= 0.000215 df= 4 K= 21.84
a a b b ab
NNAS
0
300
600
900
F
p= 2.9e-05 df= 4 K= 26.21
ab a c c bc
totalLS
0
2500
5000
7500
G
p= 3.6e-05 df= 4 K= 25.7
a a b b b
SS
0
2500
5000
7500
La_Calamine
Prayon
Lellingen
Wiwerwiltz
Winseler
Population


A
Correlation circle based on Dim 1 and Dim 2
SP
LL
LA
NL
LS
LW
NAS
maxLFS
NFS
NS
NNAS
SS
totalLS
SL
RS
Dim2 (25.1%)
Dim1 (38%)
1.0
0.5
0.0
-0.5
-1.0
B
Contributions to Dim 1
Contribution (%)
15
10
5
0
NNAS
NS
totalLS
NFS
SS
maxLFS
SL
NAS
LA
LW
RS
LL
SP
LS
NL
C
Contributions to Dim 2
Contribution (%)
20
15
10
5
0
SP
LL
LA
NL
LS
RS
LW
SL
SS
totalLS
NAS
NNAS
NFS
NS
maxLFS
D
Scree plot
Proportion of variance(%)
38%
25.1%
11.9%
8.8%
4.6%
3.4%
2.7%
2.5%
1.2%
1%
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Dimensions
E
Projection of individuals based on Dim 1 and Dim 2
Dim2 (25.1%)
5.0
2.5
0.0
-2.5
-5.0
-4
0
4
Dim1 (38%)
Cross_type
C
M
NM
F
Projection of individuals based on Dim 1 and Dim 2
Dim2 (25.1%)
6
3
0
-3
-4
0
4
Dim1 (38%)
Pollution
NP
P



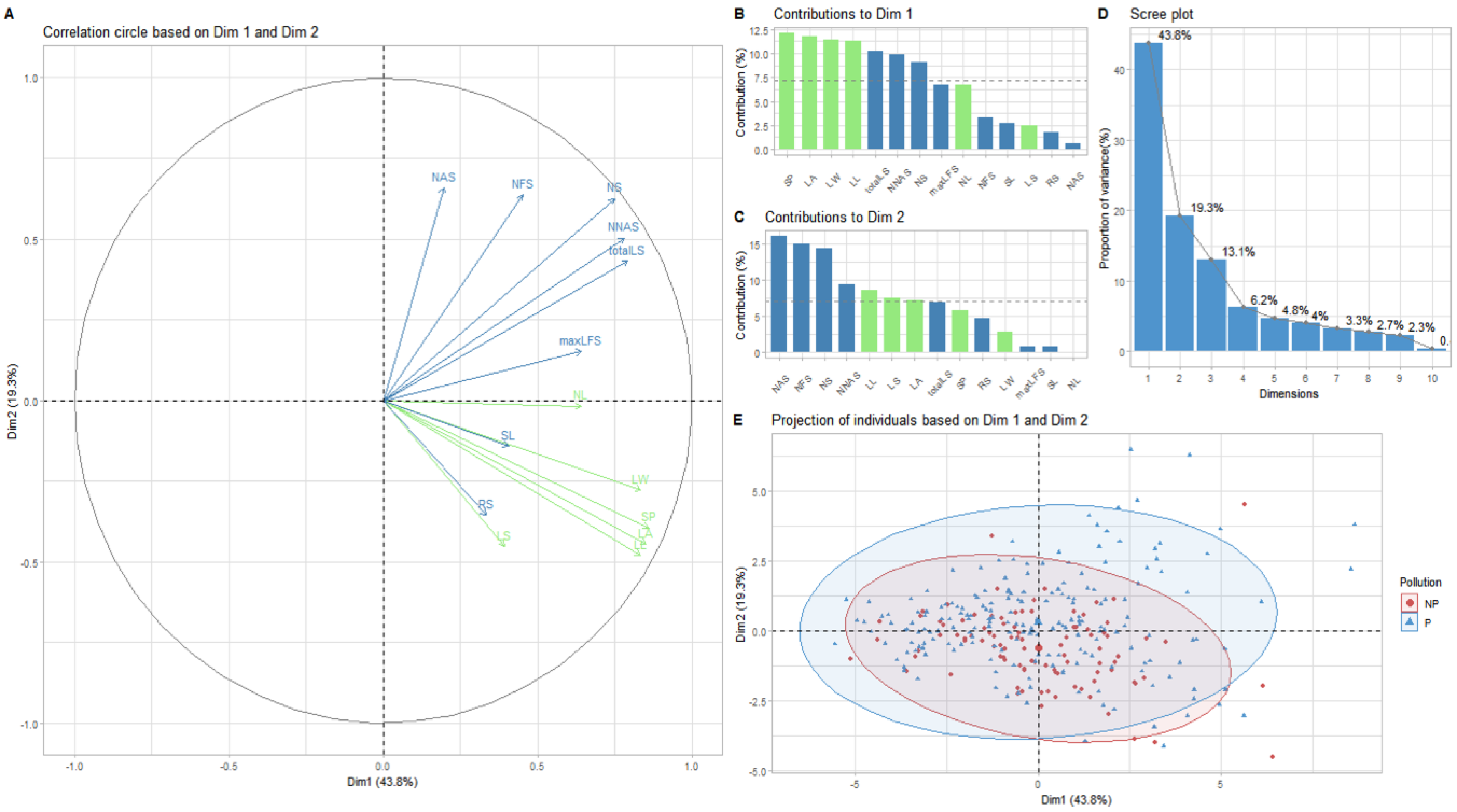
A
Polluted soils: p<0.0001 df=10 K=112.27
Non-polluted soils: p<0.0001 df=10 K=36.4

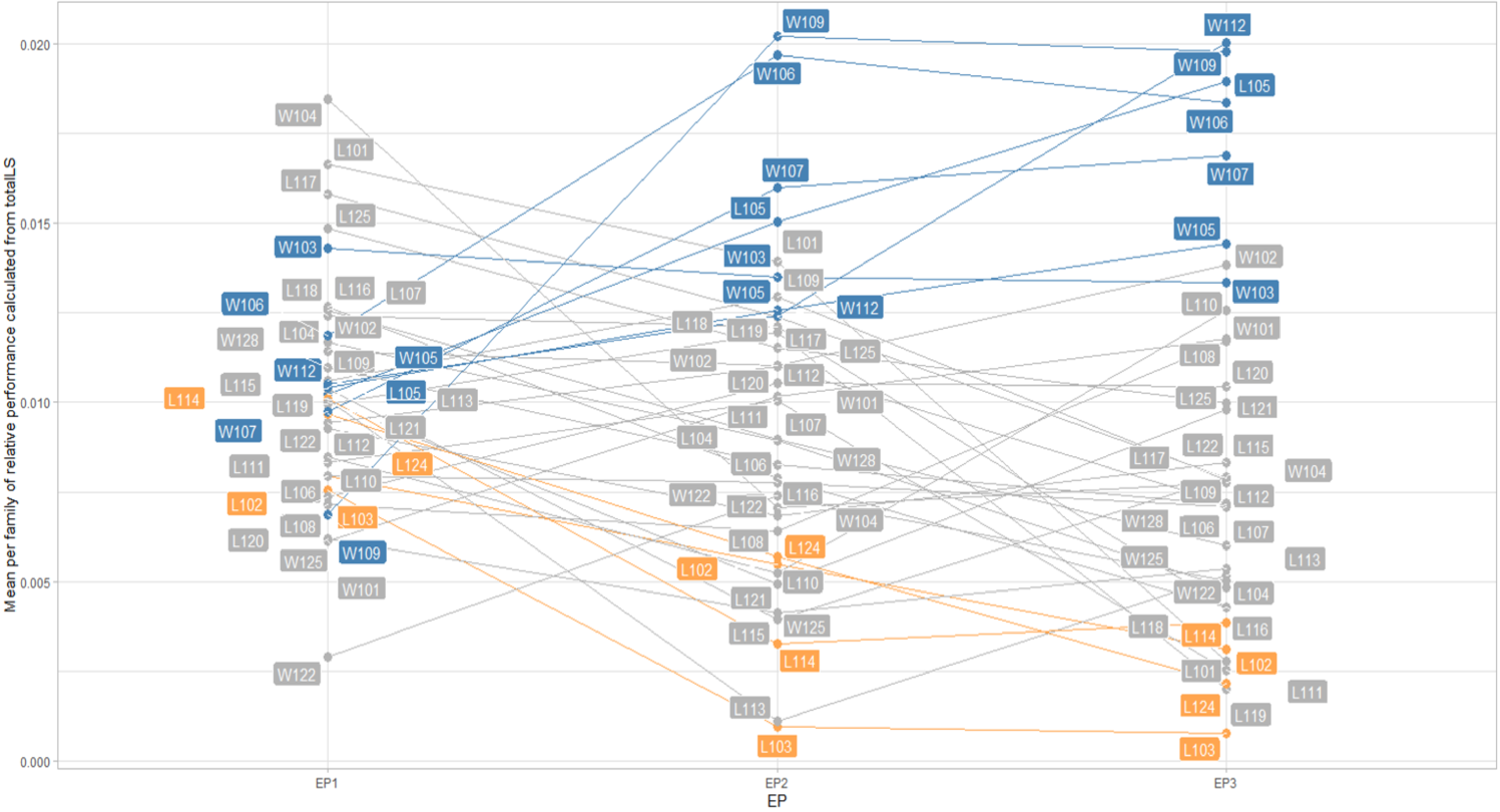
B
Polluted soils: p<0.0001 df=10 K=128.16
Non-polluted soils: p<0.0001 df=10 K=45.6

C
Polluted soils: p<0.0001 df=10 K=67.77
Non-polluted soils: p=0.217 df=10 K=13.12

D
Polluted soils: p<0.0001 df=10 K=67.8
Non-polluted soils: p=0.079 df=10 K=16.79

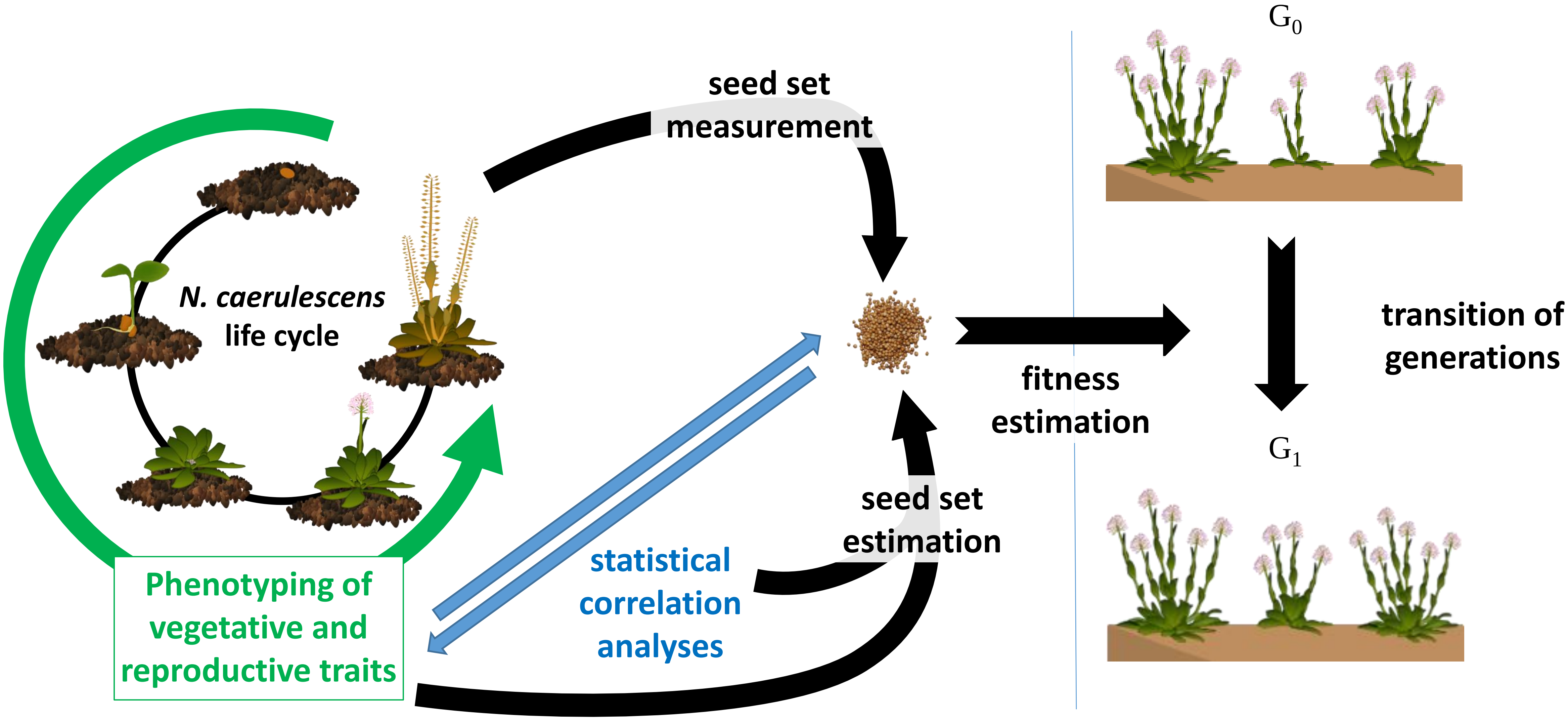
E
Polluted soils: p<0.0001 df=10 K=81.7
Non-polluted soils: p<0.0001 df=10 K=40.4


A
Correlation circle based on Dim 1 and Dim 2
NAS
NFS
NS
NNAS
totalLS
maxLFS
NL
SL
LW
RS
SP
LA
LS
LL
1.0
0.5
0.0
-0.5
-1.0
Dim2 (19.3%)
-1.0
-0.5
0.0
0.5
1.0
Dim1 (43.8%)
B
Contributions to Dim 1
Contribution (%)
12.5
10.0
7.5
5.0
2.5
0.0
SP
LA
LW
LL
totalLS
NNAS
NS
maxLFS
NL
NFS
SL
LS
RS
NAS
C
Contributions to Dim 2
Contribution (%)
15
10
5
0
NAS
NFS
NS
NNAS
LL
LS
LA
totalLS
SP
RS
LW
maxLFS
SL
NL
D
Scree plot
Proportion of variance(%)
40
30
20
10
0
43.8%
19.3%
13.1%
6.2%
4.8%
4%
3.3%
2.7%
2.3%
0.
1
2
3
4
5
6
7
8
9
10
Dimensions
E
Projection of individuals based on Dim 1 and Dim 2
5.0
2.5
0.0
-2.5
-5.0
Dim2 (19.3%)
-5
0
5
Dim1 (43.8%)
Pollution
NP
P


Mean per family of relative performance calculated from totalLS
0.020
0.015
0.010
0.005
0.000
EP1
EP2
EP3
EP
W104
L101
L117
L125
W103
L118
L116
L107
W106
L104
W102
W128
W105
L109
W112
L115
L105
L113
L114
L119
L121
W107
L122
L112
L124
L111
L110
L106
L102
L103
L108
L120
W109
W125
W101
W122
W109
W106
W107
L105
L101
W103
L109
W105
W112
L118
L119
L117
L125
W102
L112
L120
W101
L111
L107
L104
W128
L106
L116
W122
L122
W104
L108
L124
L102
L110
L121
W125
L115
L114
L113
L103
W112
W109
L105
W106
W107
W105
W102
W103
L110
W101
L108
L120
L125
L121
L122
L115
L117
W104
L109
L112
W128
L106
L107
W125
L113
W122
L104
L118
L114
L116
L101
L102
L111
L124
L119
L103

DETERMINATION OF A SEED SET ESTIMATOR FOR LARGE SAMPLE-SIZE EXPERIMENTS

LABORATORY NATURAL SELECTION EXPERIMENT



G_0

seed set measurement

N. caerulescens life cycle

transition of generations

fitness estimation

G_1

seed set estimation

statistical correlation analyses

Phenotyping of vegetative and reproductive traits