Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multiple fitness-related QTLs control serpentine adaptation in *Arabidopsis arenosa*

Veronika Lipánová
veronika.konecna@ibot.cas.cz
Czech Academy of Sciences https://orcid.org/0000-0001-9568-342X
Vít Bureš
Faculty of Science, Charles University
Timothée Lamotte
Faculty of Science, Charles University
Miloš Duchoslav
Faculty of Science, Charles University
Gabriela Šrámková
Faculty of Science, Charles University
Josselin Clo
CNRS, Univ. Lille
Filip Kolar
Charles University https://orcid.org/0000-0002-8793-7992

Article

Keywords: bulk segregant analysis, adaptation, QTL, Arabidopsis, serpentine soils

Posted Date: June 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9557974/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: There is no duality of interest

Abstract

Linking genomic signatures of selection to fitness remains a major challenge in the studies of local adaptation. In contrast with the rapidly increasing reports of candidate genes from selection scanning studies, there is a shortage of studies associating genes underlying ecologically important traits with their fitness consequences in natural systems. To fill this gap, we identified candidate loci conferring fitness advantage of *Arabidopsis arenosa* in stressful serpentine soil using bulk segregant analysis (BSA) combined with low-coverage individual whole-genome sequencing. We generated the F2 population from serpentine and non-serpentine parental populations, quantified fitness-related phenotypic variation in serpentine soil, and compared two bulks of individuals from the extreme tails of the fitness distribution. We detected five quantitative trait loci (QTL) covering with a total of 475 genes. These genes were enriched for processes and functions related to ion homeostasis and detoxification and cell-wall organization. Importantly, we identified a significant overlap of these QTL candidates with 15 candidate genes previously identified by the genome-wide divergence scans of the natural populations, providing further evidence that targets of selection also likely contribute to fitness variation. Our results establish a link between genomic differentiation and fitness in a natural system and highlight serpentine soils as a powerful model for studying polygenic adaptation to extreme environments.

Introduction

Local adaptation to extreme environments provides a powerful system for studying the genetic architecture of complex traits (Savolainen et al. 2013). Edaphically challenging habitats, representing a heterogeneous landscape mosaic, can impose strong selection leading to local adaptation (Jain and Bradshaw 1966). Although empirical studies linking fitness to environmental variation in natural populations are now relatively abundant (Espeland and Rice 2007; Hereford 2009; Exposito-Alonso et al. 2018; Dittmar and Schemske 2023), the underlying genomic architecture of adaptation at fine spatial and genetic scales remains poorly understood (Friedman and Willis 2013; Hoban et al. 2016; Selby and Willis 2018; James et al. 2023). Serpentine soils represent a particularly powerful model for addressing these gaps, as their extreme chemical and physical properties impose strong, persistent, and quantifiable selective pressures (Brady et al. 2005; Rajakaruna 2018; Konečná et al. 2020). These soils are typically nutrient-poor, with low Ca/Mg ratios, elevated metal concentrations, and low water-holding capacity (Harrison and Rajakaruna 2011). Successful establishment in serpentine environments, therefore, requires coordinated responses to multiple interacting stressors — chemical toxicity, nutrient imbalance, and water limitation (Brady et al. 2005).

The multifactorial nature of serpentine soils and the resulting diversity of adaptive phenotypes suggest a predominantly complex, polygenic genetic basis. On the one hand, the few quantitative trait locus (QTL) studies conducted on serpentine ecotypes to date have provided evidence for relatively simple genetic architectures underlying individual traits, such as nickel tolerance in *Silene vulgaris* (Bratteler et al. 2006) and *Caulanthus amplexicaulis* (Burrell et al. 2012). However, studies in two wild, outcrossing *Arabidopsis* species — *A. arenosa* (Arnold et al. 2016; Konečná et al. 2021) and *A. lyrata* (Turner et al. 2008, 2010) — point to a polygenic basis of serpentine adaptation. Consistent with this view, Toll and Willis (2024) proposed that the trait divergence in serpentine adaptation is driven by a combination of a few loci of large effect and many loci of small effect. However, because serpentine environments impose multiple, concurrent abiotic stressors, the extent to which identified genomic candidates translate into fitness differences in natural populations remains unresolved.

Quantitative trait locus (QTL) mapping offers a powerful approach for dissecting the genetic basis of adaptive traits (Bomblies and Peichel 2022). One efficient QTL-mapping approach is bulk segregant analysis (BSA; Michelmore et al. 1991), which compares allele frequencies between pools of individuals with the extreme phenotypes, enabling identification of QTLs underlying complex traits (Yang et al. 2013; Han et al. 2015; Mansfeld and Grumet 2018). Despite extensive population genomic work on serpentine adaptation, there is a virtual lack of follow-up studies linking candidate loci to fitness variation in serpentine soil. The exception is e.g. the work of Selby and Willis (2018), which identified a major-effect QTL underlying approximately 33% of survival differences between serpentine and non-serpentine populations of *Mimulus guttatus* in field-transplant experiments. Beyond this example, however, experimental validation of candidate loci identified from genome-wide divergence scans linking them with fitness remains rare.

Arabidopsis arenosa provides a suitable system for addressing this gap. Konečná et al. (2021, 2022) have detected repeated local adaptation to serpentine soils in multiple Central European populations of the species, and demonstrated the polygenic basis of adaptation by identifying candidate genes showing serpentine-related allele frequency differences in multiple populations. Reciprocal transplant experiment further showed significantly better performance of serpentine populations in their native soils, assessed by cumulative fitness estimates (Konečná et al. 2022). However, the causal link between candidate loci and fitness variation has not yet been experimentally tested. To address this gap, we selected one population pair growing on serpentine (S) and non-serpentine (N) soils previously studied in (Konečná et al. 2021, 2022) to test for the association between the 494 candidate genes identified by genome-wide divergence scan, and fitness.

Here, we studied the genetic architecture of serpentine adaptation in *Arabidopsis arenosa*. We aimed to refine the broad list of candidate genes to those that confer a fitness advantage in serpentine soil. To do so, we combined bulk segregant analysis with low-coverage whole-genome sequencing of high- and low-fitness individuals from F2 mapping populations derived from crossing of serpentine (S) and non-serpentine (N) parental populations of *A. arenosa* (Konečná et al. 2021). We specifically ask: (i) Does the fitness of the F2 mapping population vary to a different extent under different soil conditions? (ii) What is the number and genomic position of loci underlying fitness differences of plants in the serpentine soil? (iii) Which candidate loci, identified by the genome-wide divergence scans of natural populations, are associated with the fitness of plants? Together, our results provide the first direct experimental test of whether genomic candidates identified by natural divergence scans in *A. arenosa* can be linked with fitness advantages in serpentine soil conditions.

Materials and Methods

Plant material and soil sampling

We collected seeds from the autotetraploid serpentine (S) and non-serpentine (N) populations in Central Bohemia in the Czech Republic (see Konečná et al. (2021) for details about the studied populations, named as S1 and N1 in the original studies). Seeds were collected from ~30 mother plants/population from

multiple sub-sites within each locality. Further, right before the main cultivation experiment we collected ~500l of serpentine soil (S) from multiple sub-sites within the locality to cover the habitat heterogeneity. In order to ensure sufficient survival of plants from both populations, the original S soil was mixed 90:10 with garden soil, reflecting our pilot cultivation experiment with different ratios of S and garden soil.

F1 and F2 production/cultivation and experimental design

Seeds of S and N populations (9 mother plants/population) were germinated in garden soil in the growth chamber (Conviron) under these conditions: 12 h light at 20 °C and 12 h dark at 10 °C. Seedlings were transferred to multi-pots ~three weeks after sowing and were cultivated under these conditions: 16 h light at 23 °C and 8 h dark at 18 °C. Further, seedlings were vernalized (16 h dark at 5°C and 8 h light at 10 °C) for ~six weeks to induce synchronised flowering following (Baduel et al. 2016). Eleven plants from each population were paired and reciprocally crossed to create F1 hybrids between the N-S populations. The F1 seeds were germinated in Petri dishes under these conditions: 12 h light at 23 °C and 12 h dark at 13 °C. After three weeks, they were transferred to multi-pots and kept under the same conditions in the growth chambers for two months and then subjected to vernalization as in the previous generation. We germinated seeds from 15 F1 hybrid seed families and used them for reciprocal crosses, to account for potential cytoplasmic effects, to create 30 variable F2 seed families.

These 30 F2 seed families were cultivated in the final experiment focussed on scoring fitness in the F2 segregating population. Five seeds per each F2 seed family were sown in garden soil, and 45 seeds per each F2 seed family in S soil. After four weeks of cultivation (12 h light at 23 °C, 12 h night at 13 °C), germination was scored as cotyledon appearance. Due to poor germination in the S soil (no seedlings survived two weeks after germination), we germinated 55 seeds per each F2 seed family for the further experimental cultivation in small multipots filled with the garden soil for four weeks. Then we transplanted young seedlings to 1l pots filled with either garden soil (control treatment) and to the 90:10 blended S soil (selective treatment with S soil) and we cultivated them for another four weeks under standard greenhouse conditions and watered them regularly. In total, we cultivated 730 F2 plants in the S soil treatment and 89 F2 plants and 89 F1 plants in the control treatment (Fig. 1A).

Phenotyping and bulk segregant analysis

Following the work of (Konečná et al. 2021, 2022), which identified phenotypic traits differentiating the parental N and S populations - including the observation that plants reached their maximum rosette diameter after three months of cultivation (four months after germination). We measured rosette areas as it was shown as a relevant proxy to the fitness in this system (Konečná et al. 2021, 2022) and we used this trait to define low-fitness and high-fitness bulks. From the 730 phenotyped F2 individuals, we selected 10% of individuals with the smallest rosette areas (low-fitness bulk) and 10% of individuals with the largest rosette areas (high-fitness bulk). We collected leaves and stored them in silica gel for subsequent sequencing. Due to an unsuccessful DNA extraction in eleven plants of the low-fitness bulk, the final bulk sizes were adjusted from 73 to 62 plants per bulk.

DNA extraction, library preparations, sequencing, and raw data processing

We used bulk segregant analysis (BSA; Michelmore et al. 1991) combined with low-coverage individual whole-genome sequencing to identify loci underlying fitness differences of plants in S soil. The genomic DNA was extracted from dried leaves in silica gel using the sorbitol extraction method (Štorchová et al. 2000) and then purified using 0.4X AMPure XP (Beckman Coulter Inc., Brea, California, USA). WGS libraries were prepared following the genome sequencing protocol LITE of Perez-Sepulveda et al. (2021), with samples indexed individually and pooled for sequencing. The libraries were sequenced in 300 cycles (2 × 150 bp paired-end reads) on the HiSeqX Illumina platform. We used TRIMMOMATIC v.0.36 (Bolger et al. 2014) to remove adaptor sequences and low-quality base pairs. Trimmed reads >100 bp were mapped to the reference genome of closely related North American *A. lyrata* (Hu et al. 2011) by BWA v. 0.7.15 (Li and Durbin 2009) with default settings. Duplicated reads were identified by Picard v.2.8.1 (2019) and discarded. The overview of mapping statistics is presented in Supplementary Data S1. The average depth of coverage was 2.25x for low-fitness bulk and 2.41x for high-fitness bulk (Supplementary Data S1).

Variant calling and filtrations

BAM files from 62 individuals from the high-fitness bulk and 62 individuals from the low-fitness bulk were first processed using Picard Tools (Picard Toolkit, 2019) to add read group information. We used a PoolSeq approach for variant calling as the alternative genotype-likelihood based inference of individual autotetraploid genotypes from low-coverage data is challenging due to higher uncertainty given by additional heterozygous genotypes in tetraploids. The BAM files were thus merged and sorted into one BAM file per bulk using SAMtools v1.14 (Danecek et al. 2021). Variants were then identified jointly on the two bulk-level BAM files using FreeBayes v.9.9.2 (Garrison and Marth 2012). Variant calling was used to identify polymorphic sites and assign reference and alternative alleles. However, downstream analyses did not rely on genotype calls. Instead, allele frequencies were estimated directly from read counts (allele depth, AO, and total depth, DP) extracted from the VCF file. The resulting VCF file was evaluated for quality metrics available from Freebayes. Variants were filtered using BCFtools (Danecek et al. 2021) to retain only biallelic SNPs with read depth supporting site between 4 and 1000 per bulk, mapping quality ≥20, strand bias metrics SRP (Strand Read Position) and SAP (Strand Allele Position) <100, and allele balance between 0.1 and 0.9. This filtered data set contained 2,441,424 SNPs with an average depth of coverage of 169x in high-fitness bulk (0.04% missing data) and 153x in low-fitness bulk (0.05% missing data).

QTL analysis

We conducted bulk segregant analysis (BSA) using the QTLseqr v.0.7.5.2 package (Mansfeld and Grumet, 2018) to detect the quantitative trait loci (QTL) applying the (i) QTL-seq approach (Takagi et al. 2013) and (ii) G'statistics approach (Magwene et al. 2011). We applied the following additional filtering steps for each bulk according to Mansfeld and Grumet (2018). We retained only SNPs with a minimum depth of 8 in both bulks and excluded SNPs with a depth difference >100 between bulks to reduce coverage bias. We also removed SNPs where the reference allele frequency across both bulks was either >0.95 or <0.05.

QTL-seq analysis was performed using the runQTLseqAnalysis() function to calculate the SNP index for each bulk as follows: the proportion of reads supporting the alternative allele over the total depth at that position (Takagi et al. 2013). Then we calculated the ΔSNP index as the difference between the SNP indices of the two bulks. The analysis was conducted with a window size of 500 kb, using 10,000 permutations to simulate the null distribution. The population structure was specified as F2, and bulk sizes were set to 248 chromosomes per bulk, corresponding to 62 tetraploid individuals. Allele frequency estimation in the BSA framework was based exclusively on read counts, with the effective number of chromosome copies incorporated only in the QTLseqr simulations. Confidence intervals were set at the 95th and 99th percentiles of the simulated ΔSNP index distribution.

We applied G'statistics approach using the runGprimeAnalysis() function, which calculated the G statistic for each SNP. Then, it estimated the tricube-smoothed G' and ΔSNP index values of each SNP within the window of the same size, i.e., 500 kb, as in the previous approach (recommended by Mansfeld and Grumet (2018). This approach provides a robust statistical framework for detecting significant deviations in allele frequency between bulks while accounting for local SNP density and reducing the influence of single-SNP outliers. Visualization of QTL-seq results was performed using the plotting functions implemented in QTLseqr: plotGprimeDist() to display the distribution of G' statistics with Hampel and ΔSNP outlier filters, and plotQTLStats() to generate genome-wide profiles of G', ΔSNP index, and $-\log_{10}(p)$ values with significance thresholds. To visualize ΔSNP and SNP-index patterns within significant QTL regions, we generated smoothed allele frequency profiles using custom R plotting scripts. The ΔSNP index and SNP index values were smoothed over a sliding window of 100 SNPs (Mok et al. 2023). For each significant QTL, we plotted the absolute ΔSNP values together with the SNP indices of both bulks. Because the direction of allele frequency change between bulks does not directly indicate which allele is adaptive, we additionally visualized and quantified the absolute values of ΔSNP, plotting absolute ΔSNP together with SNP indices of both bulks for each significant QTL region.

Simulations

To test whether the allele-frequency differences observed between low- and high-fitness bulks are consistent with realistic genetic architectures of serpentine adaptation, we performed forward simulations mimicking the BSA experimental design. We simulated obligately outcrossing autotetraploid F2 populations with additive fitness effects across *L* independent biallelic loci (*L* = 6, 100, and 494), representing major-effect, intermediate, and polygenic architectures. Beneficial alleles contributed additively to fitness, with their frequencies in the S parental population varying from 0.05 to 0.95 across simulations, while remaining rare in the N parental population, with its frequency being fixed at 0.1. For each combination of genetic architecture and parental allele frequency, 100 independent simulations were performed to estimate mean allele-frequency differences between low- and high-fitness bulks and their 95% confidence intervals (see Supplementary Methods 1 for more details).

Gene annotation

Genes located within significant QTL regions were identified based on the Arabidopsis lyrata genome annotation available from NCBI (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/004/255/GCF_000004255.2_v.1.0/). All genes containing at least one significant SNP were retained for downstream analyses. Gene-level summaries were derived by aggregating SNP-level statistics within each gene. Identified *A. lyrata* genes were first translated to corresponding NCBI protein identifiers using a custom R pipeline. These protein identifiers were subsequently mapped to *Arabidopsis thaliana* orthologs using a curated orthology table from the Brassicaceae orthology resource (https://github.com/mduchoslav/Brassicaceae_orthology/blob/2.0/R_analysis/Arabidopsis_lyrata_NCBI_At_orthologues_and_functional_annotations.tsv). This standardized conversion enabled direct comparison of BSA - QTL candidate genes with the candidate loci, identified by the F_{ST} scans candidate genes of S and N populations, from Konečná et al. (2021) which were also reported as *Arabidopsis thaliana* orthologs. Lastly, we tested if the overlap between these two datasets is higher than expected by chance using Fisher's exact test implemented in the SuperExactTest R package (Wang et al. 2015). For each gene identified in the overlap between BSA - QTL and F_{ST} candidate sets, we visualized absolute ΔSNP values together with SNP indices for both bulks across the surrounding genomic region, using the same sliding window size and smoothing parameters as applied in the QTL analysis, implemented in a custom R script.

Gene ontology enrichment and tissue-specific gene expression analyses

To assess the functional relevance of genes located within significant QTL regions (BSA - QTL candidates), we performed (i) gene ontology (GO) enrichment analysis and (ii) tissue-specific gene expression analysis. For GO enrichment analysis, we used a Bioconductor package topGO v2.50.0 (Alexa and Rahnenfuhrer, 2025), using GO terms for *Arabidopsis lyrata* genes compiled from several sources, available at https://github.com/mduchoslav/Brassicaceae_orthology/blob/3.0/functional_annotation/1_only_At_orthologues_and_GO/Arabidopsis_lyrata_NCBI_At_orthol Because the annotation table linked GO terms to *A. lyrata* NCBI protein identifiers, these were converted to the corresponding *A. lyrata* NCBI gene identifiers using the associated gene–mRNA–protein translation table. Therefore, only protein-coding BSA - QTL candidate genes were included in the analysis (N = 430). To identify significantly enriched biological processes (BP) and molecular functions (MF), we applied Fisher's exact test ($p < 0.05$) with the "elim" algorithm, which accounts for hierarchical GO dependencies by excluding genes already mapped to significant child terms (Alexa and Rahnenfuhrer, 2025; Grossmann et al. 2007). Results from both ontologies were merged, and broad terms, i.e., those associated with more than 500 genes, or terms represented by a single gene, were excluded. Fold enrichment (FE) was calculated as the ratio of observed to expected candidate genes per GO term. To highlight functional categories relevant to serpentine adaptation, we selected significantly enriched GO terms based on a manual keyword search (transport, chromium, cobalt, nickel, detoxification, homeostasis, oxidative stress, root, nutrient, calcium, magnesium, potassium, stress, toxic, ion, microtubule, vacuole, cellulose, cellulase, dessication, wall).

Furthermore, we examined the tissue-specific expression of genes located within significant QTL regions using the *Arabidopsis* tissue atlas (Mergner et al. 2020). We downloaded expression data across all *A. thaliana* genes (available athttps://www.ebi.ac.uk/gxa/experiments/E-MTAB-7978/Downloads). Our analysis focused on leaf tissues (rosette, cauline, and adult vascular leaves), as well as root and root tip tissues, as these collectively represent both direct root–soil interactions and above-ground physiological responses relevant to serpentine adaptation (Konečná et al. 2020). We summarized expression per

gene and tissue as mean transcripts per million (TPM) values using a custom R script. Within each tissue, genes were classified into high, medium, or low expression categories based on the upper (75th percentile) and lower (25th percentile) quantiles of mean TPM values. Candidate genes (N = 376 *A. thaliana* orthologs) in significant QTL regions were then merged with tissue-specific expression data to assess their relative expression levels.

Results

Variation in phenotypic traits

The overall seed germination in the S soil was 21% in N population, 80% in S population, and 73% in F2 population. We observed quantitative variation in rosette diameter in garden soil. In F1 plants, rosette diameter ranged from 4.5 cm to 25 cm (mean = 11.2 cm), whereas in F2 plants it ranged from 1.63 cm to 13.5 cm (mean = 6.78 cm; Table S1; Fig. S1). Rosette diameter differed significantly between F2 plants grown in garden soil and those grown in S soil (Kruskal–Wallis test: $X^2 = 222.61$, df = 1, $p < 0.001$; Fig. S2). F2 plants grown in S soil were markedly smaller, with rosette diameters ranging from 0.2 cm to 6.34 cm (mean = 1.49 cm, on average 3.6x smaller than in control treatment; Table S1). Following the initial measurements of rosette diameter, we defined the low-fitness and high-fitness bulk based on rosette area, described in Konečná et al. (2022) as rosette diameter and rosette area were positively correlated (Spearman's correlation test: R = 0.76, $p < 0.001$; Fig. S3). We defined low-fitness and high-fitness bulks as 10% outliers from the lower and upper tails of the rosette area distribution of the total of 730 plants (Fig. 1B; see Methods). The low-fitness bulk consisted of 62 F2 individuals with the smallest rosette areas (< 0.14 cm²; mean = 0.1 cm²; Table S1), and the high-fitness bulk consisted of 62 F2 individuals with the largest rosette areas (> 3.05 cm²; mean = 4.79 cm²; Fig. 1B-C; Table S1).

Simulation-based expectations for allele-frequency differences between fitness bulks in an autotetraploid population

Simulated allele-frequency differences between high- and low-fitness bulks depended strongly on the underlying genetic architecture assumed in the forward simulations (Fig. S4). Genetic architecture with a few major-effect loci (L = 6) resulted in large shifts in beneficial-allele frequency, with ΔSNP values ranging from ~ 0.2 to 0.4. In contrast, intermediate polygenic architecture (L = 100) led to smaller differentiation ~0.08–0.1, whereas highly polygenic architecture (L = 494) yielded only subtle allele-frequency differences between bulks, with ΔSNP values on the order of ~0.02–0.04.

Quantitative trait loci associated with serpentine adaptation

We identified loci underlying the fitness differences of F2 plants in stressful S soil by performing bulk segregant analysis (BSA). We found no significant QTL based on the ΔSNP index (QTL-seq approach) suggesting that allele-frequency differences between bulks were generally modest and not driven by large-effect loci. However, five significant QTLs were detected using the G'statistic approach (FDR = 0.05; Tab. 1, Fig. 2A). These QTLs spanned five regions of 103-733 kbp at four chromosomes and encompassed between 1,200 and 8,605 SNPs. Within these QTL regions, we identified a total of 453 genes that potentially contribute to serpentine adaptation (named BSA - QTL candidates hereafter, Supplementary Data S2). Out of the 453 candidate genes, 430 were protein-coding and 23 long non-coding RNA (lncRNA) genes. One of the strongest signals was detected on chromosome 6, where an elevated absolute ΔSNP (0.136; smoothed across 100 SNPs) coincided with a significant G' statistic (Fig. 2B, Table 1). This region contains genes encoding members of the Cystatin/Monellin superfamily and papain-like cysteine proteases (PLCPs; C1A/papain family). These proteases are in general key regulators of proteolysis under stress conditions (Alomrani et al. 2021; Coppola et al. 2024).

Functions of candidate genes in serpentine adaptation

We inspected the putative functions of 430 protein-coding BSA - QTL candidates using Gene Ontology (GO) enrichment analysis. This analysis identified 108 significantly enriched (FDR < 0.05) GO terms from biological process and molecular function categories (Fig. 3A, Supplementary Data S3). Enriched GO terms included multiple categories related to ion homeostasis and detoxification, such as sequestering of zinc ion (GO:0032119), vacuolar sequestering (GO:0043181), and magnesium ion transport (GO:0015693). In addition, significant enrichment was observed for terms connected to cell wall and its modifications, including plant-type primary cell wall biogenesis (GO:0009833), plant-type cell wall loosening (GO:0009828), cellulose synthase activity (GO:0016760), cellulose catabolic process (GO:0030245), regulation of microtubule nucleation (GO:0010968), microtubule motor activity (GO:0003777) and lignin metabolic process (GO:0009808). Further, expression analyses revealed that *A. thaliana* orthologs of 37% of BSA-QTL candidate genes were classified as highly expressed in root tips, 36% in root, and 35% in leaf tissues in the *A. thaliana* expression atlas, falling within the upper quantile (> 75th percentile) of mean TPM values relative to all genes. In contrast, 16% of candidate genes were classified as low expressed in root tips, leaf tissues, and 15% in root tissue (< 25th percentile; Supplementary Data S4).

Overlap of quantitative trait loci with F_{ST} candidates from natural populations

We explored whether candidate loci previously identified by genome-wide divergence scans between the parental N and S populations (F_{ST} candidates, Konečná et al. 2021) are associated with fitness-related variation detected in the QTL analysis (BSA - QTL candidates). 471 out of the 494 candidate loci, reported by Konečná et al. (2021), could be translated to unique *A. thaliana* gene identifiers. These were compared with 376 unique BSA - QTL candidates with assigned *A. thaliana* gene identifiers (out of 430 protein-coding genes) (Fig 3B). We found the significant overlap of 15 genes (Fisher's exact test; p = 0.005; Fig. 3B) that were independently identified by both approaches. The list of overlapping genes (Supplementary Data S2) include several promising candidates for serpentine adaptation (Table 2, Fig. 3C) associated with ion homeostasis, cell wall modification, and higher-level regulatory processes. Furthermore, three out of the 15 genes (*BIG1*, *CSLD4*, and *SKS9*) were found as candidates for adaptation also in at least one additional serpentine/non-serpentine population pair (N2–S2 to N5–S5) from Konečná et al. (2021); the overlap between the 92 repeated candidate genes and BSA-QTL candidates is significant (Fisher's exact test, $p = 0.005$).

Discussion

Serpentine adaptation is predominantly polygenic

We identified five significant QTL regions associated with fitness differences in serpentine soil containing 475 candidate genes. Together with weak ΔSNP signals and simulation-based predictions, this indicates that serpentine adaptation in *A. arenosa* is polygenic, presumably dominated by small to moderate effect rather than a few major-effect loci, as no large-effect loci were detected by BSA (QTL-seq approach). Importantly, our phenotyping captured only a subset of fitness components (rosette growth), while other key traits such as germination and flowering time were not assessed. These life-history traits are expected to contribute to fitness differences and likely involve further genetic loci pointing to polygenic architecture. This interpretation is consistent with the previous studies of serpentine adaptation in *A. arenosa*, which identified ~160 (Arnold et al. 2016) or ~60-500 (Konečná et al. 2021, 2022) candidate genes. Moreover, forward simulations mimicking our experimental design showed that such polygenic architectures produce only small allele-frequency differences between bulks, matching our empirical observations. Thus, when fitness is controlled by more small-effect loci, high fitness is achieved through subtle allele-frequency changes rather than large shifts at individual loci (Zwaenepoel 2025).

Evidence from other serpentine systems supports a polygenic architecture

Although early QTL studies of serpentine populations identified genetic architectures dominated by large-effect loci, e.g., Ni tolerance in *Silene vulgaris* (Bratteler et al. 2006), metal tolerance in *Caulanthus amplexicaulis* (Burrell et al. 2012); flowering and leaf traits in *Microseris* (Gailing et al. 2004), genome-wide analyses consistently reveal a more complex genetic basis (Arnold et al. 2016; Konečná et al. 2021). In *Mimulus guttatus*, bulk segregant mapping identified a major QTL (Tol1) with a strong effect on survival in serpentine soils (Selby and Willis 2018); however, this QTL spans a large genomic region containing many genes (on the order of hundreds), making it difficult to determine how many loci actually contribute to the observed effect. Another study in *M. guttatus* showed that a few large-effect loci interact with numerous small-effect variants to generate the multivariate trait differences distinguishing serpentine and non-serpentine ecotypes (Toll and Willis 2024). Beyond serpentine systems, numerous QTL mapping and association studies in forest trees have identified loci of small to moderate effect contributing to local adaptation (Savolainen et al. 2013).

Although the results of QTL analysis did not reveal any large effect loci, one of the strongest QTL signals was detected on chromosome 6 and encompassed genes encoding papain-like cysteine proteases (PLCPs), key regulators of proteolysis and stress-related signaling pathways (Liu et al. 2018). PLCPs have a role in the modulation of cellular responses to abiotic stress, making them compelling candidates for contributing to fitness differences under serpentine conditions. Another compelling candidate for handling heavy metal stress on serpentine soils detected by QTL analysis are heavy metal-associated proteins such as HIPP26, HIPP21 (Barth et al. 2009; de Abreu-Neto et al. 2013). GO enrichment analysis revealed significant enrichment of genes related to ion homeostasis and detoxification, further supporting the interpretation that multiple genes within the identified QTL regions contribute to adaptation to serpentine soils. Several enriched terms were also associated with the cell wall and its modification (Fig. 3A, Supplementary Data S3). In addition, several enriched terms were related to the cell wall and its modification (Fig. 3A, Supplementary Data S3). This is consistent with previous evidence that the cell wall plays an important role in plant responses to heavy metals (Berni et al. 2019; Cheng et al. 2025).

Experimental support for divergence-based candidates

A subset of 15 BSA - QTL candidates overlapped with F_{ST} candidates, previously identified in genome-wide divergence scans (Konečná et al. 2021, 2022). This significant overlap provides a direct link between the genotype and fitness by showing loci consistently differentiated between populations and loci producing detectable allele-frequency shifts under experimental serpentine conditions. This provides the support that some of the genomic regions confer a fitness advantage under stressful serpentine conditions. We lack such experimental validations in general, and besides some examples, e.g. Price et al. (2018), James et al. (2023), the integrative studies are rare.

The functions of the overlapping candidates are consistent with the GO enrichment results obtained for the full set of BSA-QTL candidates, particularly in pointing to an important role of cell wall-related processes. Several overlapping genes are associated with cell wall biosynthesis, remodeling, or secretion. These include *PMEI-PME47* (*AT5G04970*), *SKS9* (*AT4G38420*), and *GAUT3* (*AT4G38270*), all likely involved in pectin-related cell wall modification (Sedbrook et al. 2002; Bayer et al. 2006; Louvet et al. 2006; Caffall et al. 2009). In particular, pectin methylesterases enhance pectin methylesterification, reducing apoplastic negative charge and limiting metal binding (Coculo and Lionetti 2022). Members of this family are implicated in Zn, Cd, and Al stress responses (Weber et al. 2013; Geng et al. 2017; Li et al. 2022). Other cell wall-associated candidates are cellulose synthase *CSLD4* (*AT4G38190*) (Wang et al. 2011); *AUG7* (*AT5G17620*), which binds microtubules and contributes to phragmoplast formation (Schmidt-Marcec et al. 2024); and *BIG1* (*AT4G38200*), a guanine-nucleotide exchange factor involved in post-Golgi trafficking and secretion (Richter et al. 2014).

Another notable overlapping candidate is *MTN1* (*AT4G38800*), encoding methylthioadenosine nucleosidase 1, which interacts with calcium sensors involved in the regulation of magnesium homeostasis (Tang et al. 2015; Ok et al. 2015). This is particularly relevant because high magnesium concentrations, and especially the low Ca:Mg ratio, are among the major chemical stressors of serpentine soils (Proctor and Woodell 1975; Brady et al. 2005).

Several overlapping genes also point to a role for higher-level regulatory processes in serpentine adaptation. These include *ROPGEF1* (*AT4G38430*), which is involved in the regulation of polar auxin transport (Liu et al. 2017), as well as *FIONA1* (*AT2G21070*) and *SHORT LIFE* (*SHL*, *AT4G39100*), both implicated in flowering-time regulation (López-González et al. 2014; Cai et al. 2023). Together, these genes suggest that adaptation to serpentine conditions may involve not only ion homeostasis and cell wall-related traits, but also broader developmental and regulatory pathways.

Complementarity and limitations of BSA and genome-wide divergence scans

Genome-wide divergence scans detect loci with elevated genetic differentiation between populations, which may result from selection but can also arise from demographic processes. They do not estimate effect sizes and are prone to false positives (Hoban et al., 2016; Lotterhos and Whitlock, 2014; Tiffin and Ross-Ibarra, 2017). In contrast, BSA identifies loci contributing to fitness differences between experimentally defined bulks (Michelmore et al. 1991; Magwene et al. 2011; Takagi et al. 2013). However, BSA is biased toward loci of moderate to large effect and is particularly limited in polyploids due to allele dosage uncertainty and polysomic inheritance reduce allele-frequency contrasts between bulks (Oswald and Nuismer 2011; Baduel et al. 2018). As a result, BSA is not expected to capture the full set of loci contributing to serpentine adaptation. The partial overlap between the two approaches is therefore informative: genome-wide divergence scans identify loci showing elevated differentiation consistent with long-term divergent selection, whereas BSA identifies loci whose effects are large enough to generate detectable allele-frequency differences under experimental conditions.

Study limitations and future directions

Several limitations constrain our ability to fully resolve the genetic architecture of serpentine adaptation in *A. arenosa*. The detection of numerous small-effect loci contributing to fitness advantage may be limited by the cross design (e.g. higher recombinant mapping population, i.e. F3 or F4, will be necessary) and sequencing depth. Higher recombinant mapping populations and/or deeper genomic coverage would potentially improve power to identify these loci by e.g. narrowing the associated region. More comprehensive phenotyping of fitness-related traits could also reveal additional QTL, for example those associated with germination. Future work would benefit from more fine-scale QTL mapping integrated with transcriptomic and ionomic analyses. Genome-wide association studies (GWAS) could also provide a valuable complementary approach by offering higher mapping resolution and helping to identify additional variants contributing to adaptation. Together these approaches will help clarify how root development, cell-wall modification, and intracellular trafficking pathways jointly shape adaptation to serpentine soils in *A. arenosa*. The strongest candidate genes identified across complementary methods represent promising targets for functional validation. CRISPR/Cas9-mediated allele editing or transgenic approaches, combined, e.g., with ionomic profiling, offer a powerful approach to directly test causal relationships between genotype, phenotype, and fitness. Together, these efforts will help fill the gap between genomic signatures of selection and the ecological mechanisms that drive adaptation to extreme environments.

Declarations

Competing interests

None declared.

Data availability

All sequencing data are available at the European Nucleotide Archive (ENA) under accession code PRJEB111765.

References

1. Alexa A, Rahnenführer J (2025). topGO: Enrichment Analysis for Gene Ontology. R package version 2.62.0. doi:10.18129/B9.bioc.topGO. https://bioconductor.org/packages/topGO
2. Alomrani S, Kunert KJ, Foyer CH (2021) Papain-like cysteine proteases are required for the regulation of photosynthetic gene expression and acclimation to high light stress. J Exp Bot 72:3441–3454. https://doi.org/10.1093/jxb/erab101
3. Arnold BJ, Lahner B, DaCosta JM, et al (2016) Borrowed alleles and convergence in serpentine adaptation. Proc Natl Acad Sci 113:8320–8325. https://doi.org/10.1073/pnas.1600405113
4. Baduel P, Arnold B, Weisman CM, et al (2016) Habitat-Associated Life History and Stress-Tolerance Variation in Arabidopsis arenosa. Plant Physiol 171:437–451. https://doi.org/10.1104/pp.15.01875
5. Baduel P, Bray S, Vallejo-Marin M, et al (2018) The "Polyploid Hop": Shifting Challenges and Opportunities Over the Evolutionary Lifespan of Genome Duplications. Front Ecol Evol 6:. https://doi.org/10.3389/fevo.2018.00117
6. Barth O, Vogt S, Uhlemann R, et al (2009) Stress induced and nuclear localized HIPP26 from Arabidopsis thaliana interacts via its heavy metal associated domain with the drought stress related zinc finger transcription factor ATHB29. Plant Mol Biol 69:213–226. https://doi.org/10.1007/s11103-008-9419-0
7. Bayer EM, Bottrill AR, Walshaw J, et al (2006) Arabidopsis cell wall proteome defined using multidimensional protein identification technology. PROTEOMICS 6:301–311. https://doi.org/10.1002/pmic.200500046
8. Berni R, Luyckx M, Xu X, et al (2019) Reactive oxygen species and heavy metal stress in plants: Impact on the cell wall and secondary metabolism. Environ Exp Bot 161:98–106. https://doi.org/10.1016/j.envexpbot.2018.10.017
9. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
10. Bomblies K, Peichel CL (2022) Genetics of adaptation. Proc Natl Acad Sci 119:e2122152119. https://doi.org/10.1073/pnas.2122152119
11. Brady KU, Kruckeberg AR, Bradshaw Jr. HD (2005) Evolutionary Ecology of Plant Adaptation to Serpentine Soils. Annu Rev Ecol Evol Syst 36:243–266. https://doi.org/10.1146/annurev.ecolsys.35.021103.105730
12. Bratteler M, Lexer C, Widmer A (2006) Genetic architecture of traits associated with serpentine adaptation of Silene vulgaris. J Evol Biol 19:1149–1156. https://doi.org/10.1111/j.1420-9101.2006.01090.x
13. Broad Institute (2019). Picard Toolkit. Broad Institute, GitHub repository. https://broadinstitute.github.io/picard/

14. Burrell AM, Hawkins AK, Pepper AE (2012) Genetic analyses of nickel tolerance in a North American serpentine endemic plant, Caulanthus amplexicaulis var. barbarae (Brassicaceae). Am J Bot 99:1875–1883. https://doi.org/10.3732/ajb.1200382
15. Bürstenbinder K, Waduwara I, Schoor S, et al (2010) Inhibition of 5′-methylthioadenosine metabolism in the Yang cycle alters polyamine levels, and impairs seedling growth and reproduction in Arabidopsis. Plant J 62:977–988. https://doi.org/10.1111/j.1365-313X.2010.04211.x
16. Caffall KH, Pattathil S, Phillips SE, et al (2009) *Arabidopsis thaliana* T-DNA Mutants Implicate *GAUT* Genes in the Biosynthesis of Pectin and Xylan in Cell Walls and Seed Testa. Mol Plant 2:1000–1014. https://doi.org/10.1093/mp/ssp062
17. Cai J, Hu J, Amara U, et al (2023) Arabidopsis N6-methyladenosine methyltransferase FIONA1 regulates floral transition by affecting the splicing of FLC and the stability of floral activators SPL3 and SEP3. J Exp Bot 74:864–877. https://doi.org/10.1093/jxb/erac461
18. Cheng Z, Wang C, Tang F, et al (2025) The cell wall functions in plant heavy metal response. Ecotoxicol Environ Saf 299:118326. https://doi.org/10.1016/j.ecoenv.2025.118326
19. Coculo D, Lionetti V (2022) The Plant Invertase/Pectin Methylesterase Inhibitor Superfamily. Front Plant Sci 13:. https://doi.org/10.3389/fpls.2022.863892
20. Coppola M, Mach L, Gallois P (2024) Plant cathepsin B, a versatile protease. Front Plant Sci 15:. https://doi.org/10.3389/fpls.2024.1305855
21. Danecek P, Bonfield JK, Liddle J, et al (2021) Twelve years of SAMtools and BCFtools. GigaScience 10:giab008. https://doi.org/10.1093/gigascience/giab008
22. de Abreu-Neto JB, Turchetto-Zolet AC, de Oliveira LFV, et al (2013) Heavy metal-associated isoprenylated plant protein (HIPP): characterization of a family of proteins exclusive to plants. FEBS J 280:1604–1616. https://doi.org/10.1111/febs.12159
23. Dittmar EL, Schemske DW (2023) Temporal Variation in Selection Influences Microgeographic Local Adaptation. Am Nat 202:471–485. https://doi.org/10.1086/725865
24. Espeland EK, Rice KJ (2007) Facilitation Across Stress Gradients: The Importance of Local Adaptation. Ecology 88:2404–2409. https://doi.org/10.1890/06-1217.1
25. Exposito-Alonso M, Becker C, Schuenemann VJ, et al (2018) The rate and potential relevance of new mutations in a colonizing plant lineage. PLOS Genet 14:e1007155. https://doi.org/10.1371/journal.pgen.1007155
26. Friedman J, Willis JH (2013) Major QTLs for critical photoperiod and vernalization underlie extensive variation in flowering in the Mimulus guttatus species complex. New Phytol 199:571–583. https://doi.org/10.1111/nph.12260
27. Gailing O, Macnair MR, Bachmann K (2004) QTL Mapping for a Trade-Off between Leaf and Bud Production in a Recombinant Inbred Population of Microseris douglasii and M. bigelovii (Asteraceae, Lactuceae): A Potential Preadaptation for the Colonization of Serpentine Soils. Plant Biol 6:440–446. https://doi.org/10.1055/s-2004-817958
28. Garrison E, Marth G (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907.
29. Geng X, Horst WJ, Golz JF, et al (2017) LEUNIG_HOMOLOG transcriptional co-repressor mediates aluminium sensitivity through PECTIN METHYLESTERASE46-modulated root cell wall pectin methylesterification in Arabidopsis. Plant J 90:491–504. https://doi.org/10.1111/tpj.13506
30. Grossmann S, Bauer S, Robinson PN, Vingron M (2007) Improved detection of overrepresentation of Gene-Ontology annotations with parent–child analysis. Bioinformatics 23:3024–3031. https://doi.org/10.1093/bioinformatics/btm440
31. Han Y, Lv P, Hou S, et al (2015) Combining Next Generation Sequencing with Bulked Segregant Analysis to Fine Map a Stem Moisture Locus in Sorghum (Sorghum bicolor L. Moench). PLOS ONE 10:e0127065. https://doi.org/10.1371/journal.pone.0127065
32. Harrison SP, Rajakaruna N (2011) Serpentine: The Evolution and Ecology of a Model System. University of California Press
33. Hereford J (2009) A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. Am Nat 173:579–588. https://doi.org/10.1086/597611
34. Hoban S, Kelley JL, Lotterhos KE, et al (2016) Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. Am Nat 188:379–397. https://doi.org/10.1086/688018
35. Hu TT, Pattyn P, Bakker EG, et al (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet 43:476–481. https://doi.org/10.1038/ng.807
36. Jain SK, Bradshaw AD (1966) Evolutionary divergence among adjacent plant populations I. The evidence and its theoretical analysis. Heredity 21:407–441. https://doi.org/10.1038/hdy.1966.42
37. James ME, Allsopp RN, Groh JS, et al (2023) Uncovering the genetic architecture of parallel evolution. Mol Ecol 32:5575–5589. https://doi.org/10.1111/mec.17134
38. Konečná V, Bray S, Vlček J, et al (2021) Parallel adaptation in autopolyploid Arabidopsis arenosa is dominated by repeated recruitment of shared alleles. Nat Commun 12:4979. https://doi.org/10.1038/s41467-021-25256-5
39. Konečná V, Šustr M, Požárová D, et al (2022) Genomic basis and phenotypic manifestation of (non-)parallel serpentine adaptation in Arabidopsis arenosa. Evolution 76:2315–2331. https://doi.org/10.1111/evo.14593
40. Konečná V, Yant L, Kolář F (2020) The Evolutionary Genomics of Serpentine Adaptation. Front Plant Sci 11:. https://doi.org/10.3389/fpls.2020.574616
41. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324
42. Li Z, Rao MJ, Li J, et al (2022) CRISPR/Cas9 Mutant Rice Ospmei12 Involved in Growth, Cell Wall Development, and Response to Phytohormone and Heavy Metal Stress. Int J Mol Sci 23:. https://doi.org/10.3390/ijms232416082

43. Liu H, Hu M, Wang Q, et al (2018) Role of Papain-Like Cysteine Proteases in Plant Development. Front Plant Sci 9:1717. https://doi.org/10.3389/fpls.2018.01717
44. Liu Y, Dong Q, Kita D, et al (2017) RopGEF1 Plays a Critical Role in Polar Auxin Transport in Early Development. Plant Physiol 175:157–171. https://doi.org/10.1104/pp.17.00697
45. López-González L, Mouriz A, Narro-Diego L, et al (2014) Chromatin-Dependent Repression of the Arabidopsis Floral Integrator Genes Involves Plant Specific PHD-Containing Proteins. Plant Cell 26:3922–3938. https://doi.org/10.1105/tpc.114.130781
46. Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Mol Ecol 23:2178–2192. https://doi.org/10.1111/mec.12725
47. Louvet R, Cavel E, Gutierrez L, et al (2006) Comprehensive expression profiling of the pectin methylesterase gene family during silique development in Arabidopsis thaliana. Planta 224:782–791. https://doi.org/10.1007/s00425-006-0261-9
48. Magwene PM, Willis JH, Kelly JK (2011) The Statistics of Bulk Segregant Analysis Using Next Generation Sequencing. PLOS Comput Biol 7:e1002255. https://doi.org/10.1371/journal.pcbi.1002255
49. Mansfeld BN, Grumet R (2018) QTLseqr: An R Package for Bulk Segregant Analysis with Next-Generation Sequencing. Plant Genome 11:180006. https://doi.org/10.3835/plantgenome2018.01.0006
50. Mergner J, Frejno M, List M, et al (2020) Mass-spectrometry-based draft of the Arabidopsis proteome. Nature 579:409–414. https://doi.org/10.1038/s41586-020-2094-2
51. Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828–9832. https://doi.org/10.1073/pnas.88.21.9828
52. Mok S, Yeo T, Hong D, et al (2023) Mapping the genomic landscape of multidrug resistance in Plasmodium falciparum and its impact on parasite fitness. Sci Adv 9:eadi2364. https://doi.org/10.1126/sciadv.adi2364
53. Ok SH, Cho JH, Oh S-I, et al (2015) Calcineurin B-like 3 calcium sensor associates with and inhibits 5'-methylthioadenosine nucleosidase 2 in Arabidopsis. Plant Sci 238:228–240. https://doi.org/10.1016/j.plantsci.2015.06.016
54. Oswald BP, Nuismer SL (2011) A Unified Model of Autopolyploid Establishment and Evolution. Am Nat 178:687–700. https://doi.org/10.1086/662673
55. Perez-Sepulveda BM, Heavens D, Pulford CV, et al (2021) An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes. Genome Biol 22:349. https://doi.org/10.1186/s13059-021-02536-3
56. Price N, Moyers BT, Lopez L, et al (2018) Combining population genomics and fitness QTLs to identify the genetics of local adaptation in Arabidopsis thaliana. Proc Natl Acad Sci 115:5028–5033. https://doi.org/10.1073/pnas.1719998115
57. Proctor J, Woodell SRJ (1975) The Ecology of Serpentine Soils. In: Advances in Ecological Research. Elsevier, pp 255–366
58. Rajakaruna N (2018) Lessons on Evolution from the Study of Edaphic Specialization. Bot Rev 84:39–78. https://doi.org/10.1007/s12229-017-9193-2
59. Richter S, Kientz M, Brumm S, et al (2014) Delivery of endocytosed proteins to the cell–division plane requires change of pathway from recycling to secretion. eLife 3:e02131. https://doi.org/10.7554/eLife.02131
60. Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. Nat Rev Genet 14:807–820. https://doi.org/10.1038/nrg3522
61. Schmidt-Marcec S, Parish A, Smertenko T, et al (2024) The microtubule-nucleating factor MACERATOR tethers AUGMIN7 to microtubules and governs phragmoplast architecture. Plant Cell 36:1072–1097. https://doi.org/10.1093/plcell/koad304
62. Sedbrook JC, Carroll KL, Hung KF, et al (2002) The Arabidopsis SKU5 Gene Encodes an Extracellular Glycosyl Phosphatidylinositol–Anchored Glycoprotein Involved in Directional Root Growth[W]. Plant Cell 14:1635–1648. https://doi.org/10.1105/tpc.002360
63. Selby JP, Willis JH (2018) Major QTL controls adaptation to serpentine soils in Mimulus guttatus. Mol Ecol 27:5073–5087. https://doi.org/10.1111/mec.14922
64. Štorchová H, Hrdličková R, Chrtek Jr. J, et al (2000) An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. TAXON 49:79–84. https://doi.org/10.2307/1223934
65. Takagi H, Abe A, Yoshida K, et al (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J 74:174–183. https://doi.org/10.1111/tpj.12105
66. Tang R-J, Zhao F-G, Garcia VJ, et al (2015) Tonoplast CBL–CIPK calcium signaling network regulates magnesium homeostasis in Arabidopsis. Proc Natl Acad Sci 112:3134–3139. https://doi.org/10.1073/pnas.1420944112
67. Tiffin P, Ross-Ibarra J (2017) Advances and Limits of Using Population Genetics to Understand Local Adaptation: (Trends in Ecology & Evolution 29, 673–680; 2014). Trends Ecol Evol 32:801–802. https://doi.org/10.1016/j.tree.2017.08.001
68. Toll K, Willis JH (2024) The genetic basis of traits associated with the evolution of serpentine endemism in monkeyflowers. Evolution 78:111–126. https://doi.org/10.1093/evolut/qpad196
69. Turner TL, Bourne EC, Von Wettberg EJ, et al (2010) Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. Nat Genet 42:260–263. https://doi.org/10.1038/ng.515
70. Turner TL, Wettberg EJ von, Nuzhdin SV (2008) Genomic Analysis of Differentiation between Soil Types Reveals Candidate Genes for Local Adaptation in Arabidopsis lyrata. PLOS ONE 3:e3183. https://doi.org/10.1371/journal.pone.0003183
71. Wang M, Zhao Y, Zhang B (2015) Efficient Test and Visualization of Multi-Set Intersections. Sci Rep 5:16923. https://doi.org/10.1038/srep16923

72. Wang W, Wang L, Chen C, et al (2011) Arabidopsis CSLD1 and CSLD4 are required for cellulose deposition and normal growth of pollen tubes. J Exp Bot 62:5161–5177. https://doi.org/10.1093/jxb/err221
73. Weber M, Deinlein U, Fischer S, et al (2013) A mutation in the Arabidopsis thaliana cell wall biosynthesis gene pectin methylesterase 3 as well as its aberrant expression cause hypersensitivity specifically to Zn. Plant J 76:151–164. https://doi.org/10.1111/tpj.12279
74. Yang Z, Huang D, Tang W, et al (2013) Mapping of Quantitative Trait Loci Underlying Cold Tolerance in Rice Seedlings via High-Throughput Sequencing of Pooled Extremes. PLOS ONE 8:e68433. https://doi.org/10.1371/journal.pone.0068433
75. Zwaenepoel A (2025) Autopolyploid establishment through polygenic adaptation. Evolution 79:1911–1922. https://doi.org/10.1093/evolut/qpaf124

Tables

Tables 1 and 2 are available in the Supplementary Files section.

Figures

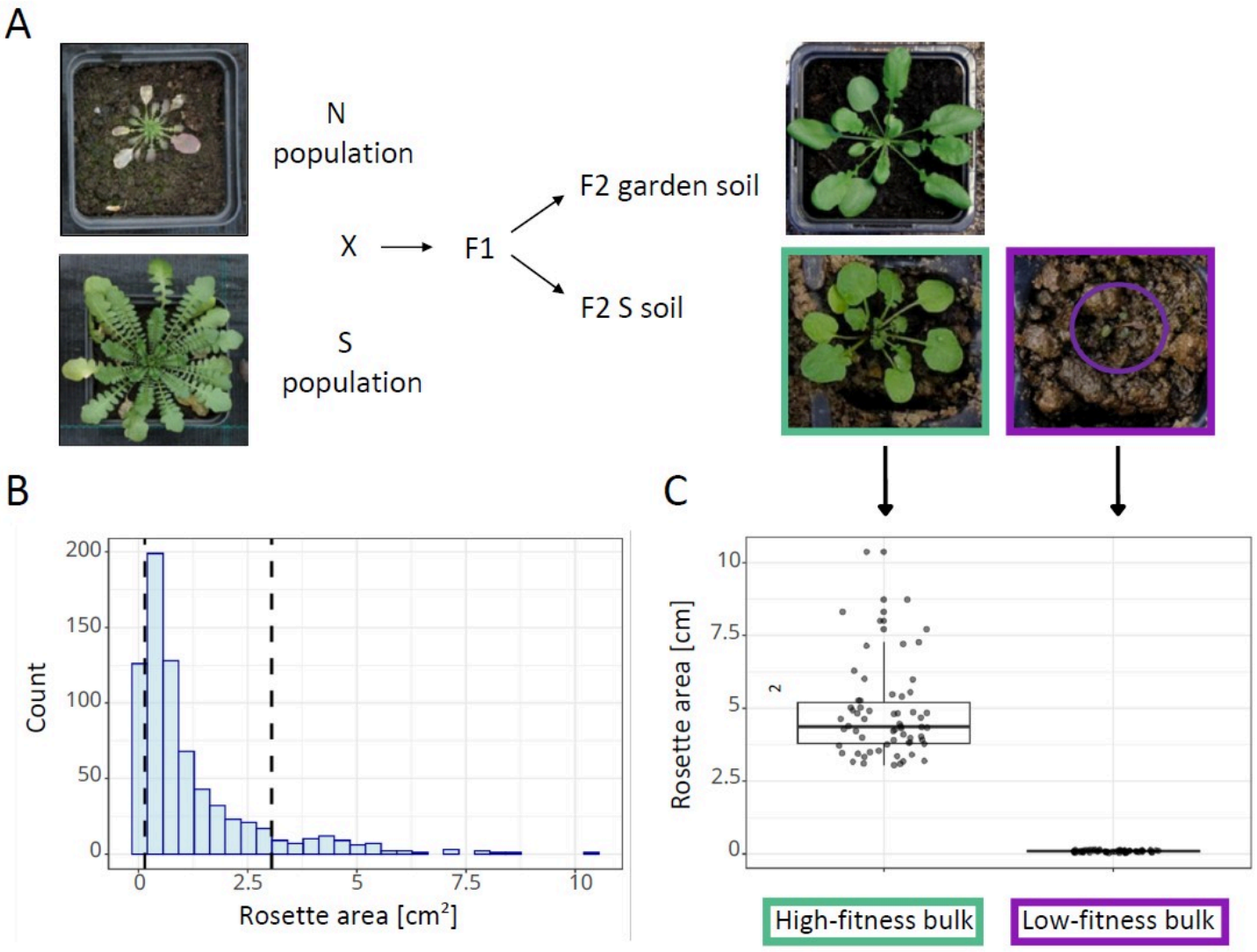


Figure 1

Bulk segregant analysis experiment. A) Production of F1 and F2 plants, photos illustrate N and S plants cultivated in S soil (adopted from Konečná et al. 2021) and F2 plants cultivated in garden soil and selective treatment with S soil. B) Distribution of rosette area values in F2 plants grown in S soil. Low- and high-fitness bulks were defined as the lower and upper 10% tails of the distribution, respectively (marked by dashed line). The low-fitness bulk comprised individuals with rosette areas < 0.14 cm², and the high-fitness bulk comprised individuals with rosette areas > 3.05 cm² (see Methods and Table S1). C) Variation in the rosette areas of F2 plants between low-fitness (N = 62 individuals) and high-fitness bulk (N = 62 individuals). Photos illustrate representative plants from low-fitness (violet frame) and high-fitness (green frame) bulk.

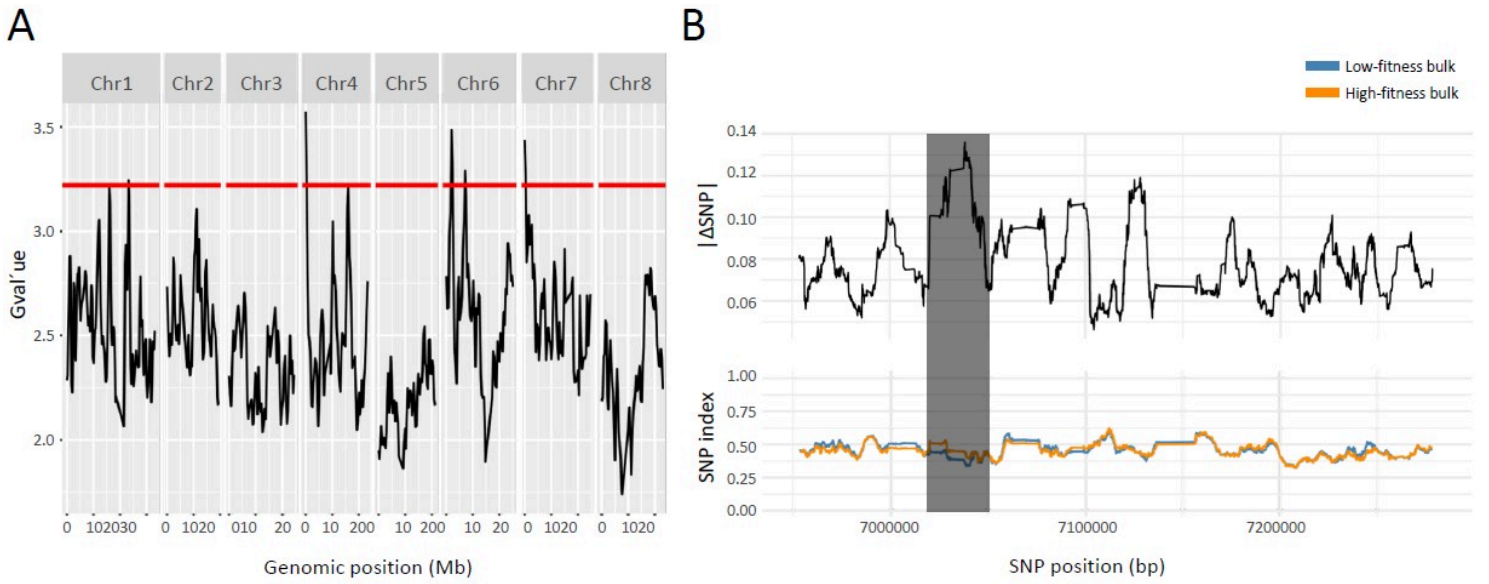


Figure 2

Identification and visualization of quantitative trait loci (QTL) using bulk segregant analysis (BSA) in A. arenosa F2 population segregating for serpentine adaptation. (A) Genome-wide distribution of tricube-smoothed G' values calculated in 500 kb sliding windows using QTLseqr. The horizontal red line indicates the genome-wide false discovery rate threshold (FDR) of 0.05 used to identify the five significant QTL regions. (B) Detailed view of a significant QTL (QTL4) on chromosome 6. The upper panel shows the smoothed absolute ΔSNP index (black line), calculated from SNP indices of the high- and low-fitness bulks and smoothed across a window of 100 SNPs. The lower panel shows smoothed SNP indices for the low-fitness (blue) and high-fitness (orange) bulks. The shaded region indicates the genomic interval containing genes putatively encoding members of the Cystatin/Monellin superfamily and papain-like cysteine proteases (PLCPs; C1A/papain family).

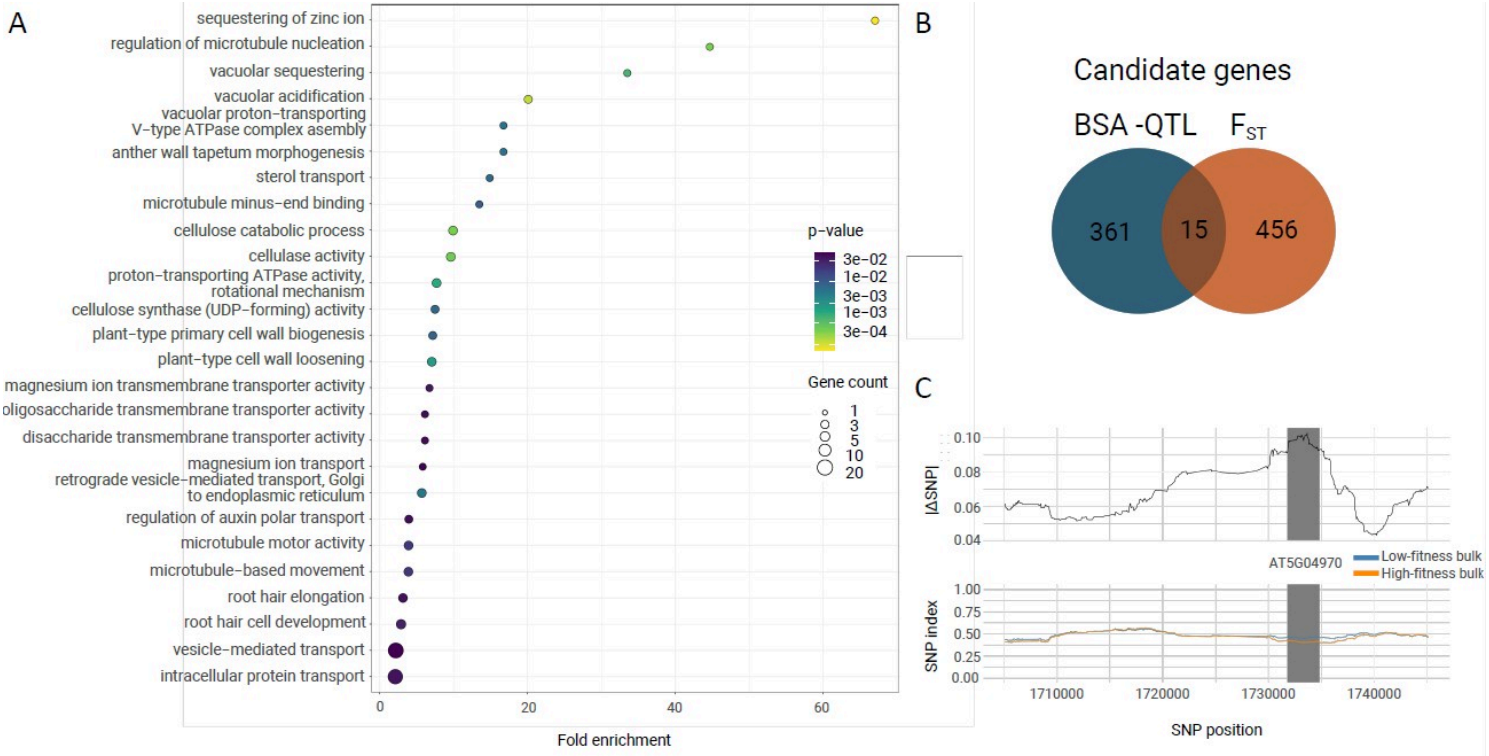


Figure 3

BSA-QTL candidate genes and their overlap with F_{ST} candidates. (A) Gene Ontology (GO) enrichment analysis of BSA-QTL candidates (n = 430, Arabidopsis thaliana orthologs). To highlight functional categories relevant to serpentine adaptation, significantly enriched GO terms (FDR < 0.05) were filtered using a manual keyword search targeting processes associated with transport, chromium, cobalt, nickel, detoxification, homeostasis, oxidative stress, root, nutrient, calcium, magnesium, potassium, stress, toxic, ion, microtubule, vacuole, cellulose, cellulase, dessication, wall. (B) Overlap between BSA-QTL and F_{ST} candidates identified

from genome-wide divergence scans of the parental N and S populations (Konečná et al. 2021). (C) Detailed view of a representative shared candidate gene PMEI-PME47 *(AT5G04970) on chromosome 6 (shown region: −26.7 kb to +10.3 kb relative to the gene). The upper panel shows smoothed absolute ΔSNP values, and the lower panel shows smoothed SNP indices for the high- and low-fitness bulks across the region (smoothing window of 100 SNPs). The shaded region indicates the gene interval.*

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tab1.docx
- Tab2.docx
- Supplementarydata.xlsx
- Supplementaryinformation.docx