**Supplementary Material**

# Traces of genetic but not epigenetic adaptation in the invasive goldenrod *Solidago canadensis* despite the absence of population structure

Silvia Eckert, Jasmin Herden, Marc Stift, Walter Durka, Mark van Kleunen, Jasmin Joshi

**Supplementary Material**

**Method S1**

**Changes applied on the MSAP protocol**

- In the restriction/ligation step, the reaction mixture contained 50 pmol/µl of the *Hpa*II/*Msp*I adapter instead of the *Mse*I adapter, 10 U /µl of the *Hpa*II/*Msp*I enzymes instead of the *Mse*I enzyme. The extracted DNA was mixed with 5.2 µl instead of 5 µl of the reaction mixture and the mixture was diluted at a ratio of 1:3 instead of 1:5.
- In the pre-selective amplification step, 30 ng/µl of *Hpa*II/*Msp*I pre-selective primer were used instead of the *Mse*I pre-selective primer and the PCR product was diluted at a ratio of 1:6 instead of 1:10.
- In the selective amplification step, the *Hpa*II/*Msp*I pre-selective primers were used instead of the *Mse*I selective primer and the PCR program was updated to the following: 2 min at 95 °C; 10 cycles (20 sec at 94 °C; 30 sec at 66 °C, decrease of 1 °C per cycle, 2 min at 72 °C); 20 cycles (20 sec at 94 °C; 30 sec at 56 °C; 2 min at 72 °C); 30 min at 60 °C.
- The agarose gel concentration was updated to 1.5 % and, for the second gel, 10 µl of the PCR product was used instead of 6 µl.
- In order to select appropriate primers, we tested 16 primers per dye for AFLP and 10 primers per dye for MSAP analysis and chose the corresponding primers based on visual inspection of peak quality in the resulting electropherograms after sequencing.

**Method S2**

**Details of the AFLP/MSAP genotyping approach**

Peak heights from binned data were exported and the peak quality for each bin was assessed manually using a custom Excel file. For this, we set a minimum peak height for each bin as a threshold to accept a peak as present. This was assisted by two error rates in defining the optimal threshold for ambiguous cases, the overall error rate (defined as the number of mismatches per number of sample-replicate pairs) and a confined error rate (defined as the number of mismatches per number of comparisons; Gáspár et al., 2019). Additionally, non-polymorphic peaks were defined according to Lynch & Milligan (1994) and ignored in further analyses. For AFLP scoring, peaks that showed an overall error rate lower than 5% or at least a three-fold higher number of sample-replicate comparisons compared to the number of mismatches were included in the AFLP binary table for

subsequent statistical analyses. Samples from both MSAP enzyme combinations, HpaII + EcoRI and MspI + EcoRI, were binned jointly to assure consistent peak detection across both enzyme combinations. However, peak quality was assessed separately to take into account the different distributional characteristics of the scored signals based on their methylation status. Corresponding peaks from both sets were selected such that the scoring of at least one of the two peaks was below the error threshold. Therefore, the MSAP error rates were relaxed because the peak quality differed between the isoschizomer measurements.

**Method S3**

**Details of genome-scan approaches**

- **BAYESCAN** determines the evidence of whether a locus is under selection by calculating posterior odds (PO), i.e., comparing the model with selection to the neutral model. Following Jeffreys (1939) and Pérez-Figueroa et al. (2010), we checked for false positives and retained loci as outliers when exceeding a threshold of $\log_{10}(PO) > 1.3$. This follows the scale of Jeffreys (1939), were loci with values exceeding this threshold are interpreted as having *substantial evidence* for selection (Figure S7). We ran BAYESCAN using 100,000 iterations with 10 repetitions and a burn-in of 50,000 following 20 pilot runs.
- A principal components analysis (**PCA**) was used to account for collinearity of WorldClim 2.0 environmental variables applied on BAYESCENV and latent factor mixed models (LMM). The first three principal components were retained for further analyses according to the broken-stick criterion (MacArthur, 1957; Figure S8). Given the first three variables with the highest relative [%] above-average contribution (> 4.76%; Figure S8), PC1 had the highest negative association with solar radiation (10.1%) and wind speed (8.4%), and the highest positive association with precipitation seasonality (BIO15; 7.9%). PC2 had the highest negative associations with precipitation-related variables, i.e., precipitation of both the coldest (BIO19; 12.7%) and the driest quarter (BIO17; 10.1%), and annual precipitation (BIO12; 8.6%). PC3 had the highest positive association with temperature-related variables, i.e., mean temperature of the coldest quarter (BIO11; 30.4%), annual mean temperature (BIO1; 18.4%), and minimum temperature of coldest month (BIO6; 17.1%). Together, all three PCs explained 87.4% of climatic variation along the sampled latitudinal gradient (PC1: 40.3%; PC2: 32.7%; PC3: 14.4%).

**Supplementary Material**

- In **BAYESCENV**, allele frequency and environmental variation are considered associated for a given locus if the posterior probability of the neutral model is lower than that of the local adaptation model. Following Aguirre-Liguori et al., (2017), we used the posterior error probability (PEP) calculated by BAYESCENV for both models and defined any locus as outlier whose PEP was lower for the local adaptation model compared to the neutral model (Figure S9). We ran BAYESCENV using a thinning interval of 20, a final sample size of 5,000, and a burn-in of 50,000 following 20 pilot runs.
- For **LFMM**, we first ran an admixture coefficient estimation (**LEA**) on a range of $K$ = 1-11 ancestral populations using 20 repetitions and 100,000 iterations. We chose the final $K$ based on the cross-entropy criterion and according to the LEA package vignette, i.e., either a $K$ at the minimum of the cross-entropy values (i.e., MSAP-n and MSAP-m) or, if there was no clear minimum value for the cross-entropy, at the beginning of the plateau (i.e., AFLP; Frichot & François, 2015; Figures S10–S12). The final $K$ was then used in LFMMs to assess outlier loci associated with environmental variables using the *lfmm*-function with 10 repetitions, 50,000 iterations, and a burn-in of 5,000. LFMM computes $p$-values for each locus indicating whether a given locus is associated with the environmental variable of interest. A final $K$ = 2 was applied for AFLPs (Figure S10). For both MSAP-m and MSAP-n data, however, a $K$ = 1 was used because no pronounced population structuring was visible along the sampled latitudinal gradient using the *snmf*-function (Figures S11 and S12). To correct for multiple testing, we used the false discovery rate (FDR) and applied a threshold of 95% where loci were defined as outliers when $\log_{10}(\text{FDR}) > 1.3$ (Figure S13).
- **RDA** was found to detect loci even under moderate-to-weak selection (Forester et al., 2018). The significance of the global model was assessed using the *anova.cca*-function in the VEGAN v2.5-6 package ($p < 0.05$ based on n = 9,999 permutations) and the first three axes were retained in the analysis according to the broken-stick criterion (MacArthur, 1957). Loci loading with a 3-times standard deviation from the center of the loadings distribution of each retained RDA axis (corresponding to a two-tailed $p$-value of $p < 0.001$) were defined as outlier loci (Figures 3 and S14). The proportion of the variance explained by the environmental predictors (adjusted $R^2$) was 0.3% for the MSAP-m dataset and 0.4% for the MSAP-n dataset, respectively.

## Supplementary Material

### Tables

**Table S1.** Details of sampled *Solidago canadensis* source populations**.** Information includes seed-sampling sites, population IDs, coordinates given in decimal degrees, elevation [m a.s.l.], estimated population sizes based on the estimated number of shoots at the sampling sites, and the number of sampled maternal lines along the latitudinal south-north gradient in Central Europe (47–54°N). Additionally, the number of individuals per population is given that were used for genotyping (AFLP) and epigenotyping (MSAP-m, MSAP-n). Abbreviations: NA – information not available; AFLP – Amplified fragment length polymorphisms; MSAP – Methylation-sensitive amplified polymorphisms.

| ***S. canadensis*** | **ID** | **Coordinates** | | **Elevation** | **Estimated** | **Maternal** | **AFLP** | **MSAP** |
|---|---|---|---|---|---|---|---|---|
| | | **°N** | **°E** | **[m a.s.l.]** | **shoots** | **lines** | | |
| *Rhäzüns (CH)* | 1 | 46.80153 | 9.39882 | 656.0 | 1000 | 19 | 19 | 17 |
| *Landquart (CH)* | 2 | 46.95745 | 9.55380 | 521.3 | 1000 | 17 | 17 | 14 |
| *Rheineck (CH)* | 3 | 47.47613 | 9.57967 | 401.3 | 30 | 13 | 13 | 12 |
| *Konstanz (D)* | 4 | 47.67258 | 9.16095 | 399.0 | 1000+ | 18 | 18 | 16 |
| *Radolfzell (D)* | 5 | 47.76432 | 8.98473 | 434.8 | 100 | 20 | 19 | 18 |
| *Engen (D)* | 6 | 47.85763 | 8.79533 | 539.5 | 100 | 16 | 16 | 14 |
| *Pleidelsheim (D)* | 7 | 48.96087 | 9.19872 | 189.4 | 1000 | 19 | 19 | 17 |
| *Heilbronn (D)* | 8 | 49.14628 | 9.19737 | 159.8 | 80–100 | 14 | 14 | 13 |
| *Tauberbischofsheim (D)* | 9 | 49.63172 | 9.65733 | 179.6 | 15–20 | 15 | 15 | 13 |
| *Tauberbischofsheim (D)* | 10 | 49.64210 | 9.64187 | 200.3 | 15000 | 18 | 18 | 15 |
| *Volkach (D)* | 11 | 49.86363 | 10.22380 | 196.7 | 200 | 19 | 18 | 14 |
| *Rödelmaier (D)* | 12 | 50.52465 | 10.42545 | 310.5 | 150 | 12 | 12 | 10 |
| *Breitungen (D)* | 13 | 50.77387 | 10.32783 | 259.5 | 600 | 19 | 18 | 15 |
| *Eisenach (D)* | 14 | 50.97552 | 10.32697 | 225.7 | 30–50 | 10 | 9 | 10 |
| *Hoheneiche (D)* | 15 | 51.12350 | 9.97555 | 197.3 | 1000+ | 18 | 17 | 14 |
| *Wollrode (D)* | 16 | 51.36967 | 9.91143 | 242.6 | 500 | 19 | 19 | 17 |
| *Kassel (D)* | 17 | 51.50665 | 9.91338 | 154.8 | 300 | 18 | 18 | 15 |
| *Bad Gandersheim (D)* | 18 | 51.86658 | 10.03582 | 147.4 | 40 | 15 | 15 | 13 |
| *Potsdam (D)* | 19 | 52.47797 | 13.01649 | 37.2 | NA | 15 | 15 | 14 |
| *Kaltenweide (D)* | 20 | 52.47997 | 9.74547 | 47.7 | 300 | 18 | 17 | 16 |
| *Walsrode (D)* | 21 | 52.84990 | 9.60088 | 38.2 | 40–50 | 8 | 8 | 7 |
| *Nützen (D)* | 22 | 53.85155 | 9.92888 | 20.3 | 1000+ | 12 | 12 | 11 |
| *Neumünster (D)* | 23 | 54.09808 | 9.98737 | 25.4 | 20–30 | 9 | 9 | 6 |
| *Neumünster (D)* | 24 | 54.11280 | 9.99352 | 29.1 | 800 | 16 | 16 | 12 |
| *Flensburg (D)* | 25 | 54.76182 | 9.44657 | 37.0 | 100–150 | 14 | 14 | 13 |

## Supplementary Material

**Table S2.** Oligo-sequences from AFLP/MSAP analyses. Combination of selective primers applied in the AFLP analysis: E1+*Mse*I_2, E2+*Mse*I_2, E3+*Mse*I_3, E4+*Mse*I_1. Combination of selective primers applied in the MSAP analysis: E1+*H*_1/*M*_1, E2+*H*_2/*M*_2, E3+*H*_3/*M*_2 and E4+*H*_3/*M*_3.

| Oligo name | Sequence (‘5–‘3) |
|---|---|
| **Adapters in the restriction/ligation step** | |
| *Eco*RI | CTCGTAGACTGCGTACC |
| | AATTGGTACGCAGTCTAC |
| *Mse*I | GACGATGAGTCCTGAG |
| | TACTCAGGACTCAT |
| *Hpa*II / *Msp*I | GACGATGAGTCTAGAA |
| | CGTTCTAGACTCATC |
| **Pre-selective primers** | |
| *Eco*RI | GACTGCGTACCAATTCA |
| *Mse*I | GATGAGTCCTGAGTAAC |
| *Hpa*II (H) / *Msp*I (M) | ATCATGAGTCCTGCTCGG |
| **Selective primers** | |
| *Eco*RI-FAM (E1) | FAM-GACTGCGTACCAATTC-ACT |
| *Eco*RI-VIC (E2) | VIC-GACTGCGTACCAATTC-ACA |
| *Eco*RI-NED (E3) | NED-GACTGCGTACCAATTC-ACC |
| *Eco*RI-PET (E4) | PET-GACTGCGTACCAATTC-AGC |
| *Mse*I_1 | GATGAGTCCTGAGTAA-CAGT |
| *Mse*I_2 | GATGAGTCCTGAGTAA-CTGT |
| *Mse*I_3 | GATGAGTCCTGAGTAA-CTAT |
| *H/M*_1 | ATCATGAGTCCTGCTCGGTCAGT |
| *H/M*_2 | ATCATGAGTCCTGCTCGGTCTGT |
| *H/M*_3 | ATCATGAGTCCTGCTCGGTCTAT |

## Supplementary Material

**Table S3.** Model coefficients from logistic mixed-effects models based on climate-related principle components. Models included principal components axes (PC1, PC2, PC3; Table S5; Figure S8) and zebularine-treatment as fixed effects and population as well as maternal lines nested within populations as random effects. PCs were determined via the broken-stick criterion. Outlier loci were derived from genome-scan approaches outlined in the main manuscript (see Figure 3). Statistical significance of explanatory variables was assessed in likelihood-ratio tests (Table 3). Abbreviations: *SE* – Standard error of model estimate; *z* – model-based z-score; *p* – the *p*-value for each model term (significant *p*-values given in bold).

| Locus | Term | *Estimate* | *SE* | *z* | *p* |
|---|---|---|---|---|---|
| loc58 (FAM-AAC-CCT-58; AFLP) | *Intercept* | -1.27 | 0.65 | -1.94 | 0.052 |
| | *Zebularine* | 0.04 | 0.75 | 0.05 | 0.958 |
| | *PC1* | 0.55 | 0.20 | 2.76 | **0.006** |
| | *PC2* | 0.03 | 0.20 | 0.15 | 0.884 |
| | *PC1 x Zebularine* | -0.03 | 0.22 | -0.15 | 0.879 |
| | *PC2 x Zebularine* | -0.15 | 0.22 | -0.67 | 0.504 |
| | *SD Maternal lines* | 0.00 | | | |
| | *SD Population* | 0.60 | | | |
| | *AICc* = 512.97; $R^2m$ = 0.07; $R^2c$ = 0.15 | | | | |
| loc286 (FAM-AAC-CCT-286; AFLP) | *Intercept* | 0.64 | 0.79 | 0.81 | 0.416 |
| | *Zebularine* | -0.14 | 0.73 | -0.19 | 0.848 |
| | *PC1* | 0.02 | 0.26 | 0.06 | 0.955 |
| | *PC2* | -0.35 | 0.27 | -1.31 | 0.191 |
| | *PC1 x Zebularine* | 0.06 | 0.24 | 0.24 | 0.807 |
| | *PC2 x Zebularine* | -0.03 | 0.24 | -0.13 | 0.895 |
| | *SD Maternal lines* | 0.76 | | | |
| | *SD Population* | 0.97 | | | |
| | *AICc* = 510.44; $R^2m$ = 0.02; $R^2c$ = 0.29 | | | | |
| loc189 (VIC-ACG-CAT-189; MSAP-m) | *Intercept* | 2.53 | 1.06 | 2.39 | **0.017** |
| | *Zebularine* | -1.76 | 1.10 | -1.60 | 0.110 |
| | *PC2* | -0.04 | 0.32 | -0.13 | 0.894 |
| | *PC3* | -0.11 | 0.30 | -0.36 | 0.719 |
| | *PC2 x Zebularine* | 0.12 | 0.34 | 0.34 | 0.731 |
| | *PC3 x Zebularine* | 0.49 | 0.34 | 1.46 | 0.144 |
| | *SD Maternal lines* | 0.76 | | | |
| | *SD Population* | 0.91 | | | |
| | *AICc* = 308.02; $R^2m$ = 0.01; $R^2c$ = 0.18 | | | | |

**Supplementary Material**

**Table S4.** Model coefficients from logistic mixed-effects models based on spatial genetic neighborhoods. Models included spatial genetic variation (MEMGENE1, MEMGENE2, MEMGENE3; Table S5; Figure S2) and zebularine-treatment as fixed effects and population as well as maternal lines nested within populations as random effects. MEMGENEs were determined via *mgQuick*-function in the MEMGENE v1.0.1 R-package. Outlier loci were derived from genome-scan approaches outlined in the main manuscript (see Figure 3). Statistical significance of explanatory variables was assessed in likelihood-ratio tests (Table 4). Abbreviations: *SE* – Standard error of model estimate; *z* – model-based z-score; *p* – the *p*-value for each model term (significant *p*-values given in bold).

| Locus | Term | *Estimate* | *SE* | *z* | *p* |
|---|---|---|---|---|---|
| loc282 (FAM-AAC-CCT-282; MSAP-m) | *Intercept* | 1.27 | 0.82 | 1.55 | 0.122 |
| | *Zebularine* | 0.41 | 0.99 | 0.41 | 0.681 |
| | *MEMGENE1 (MG1)* | 0.14 | 0.25 | 0.56 | 0.575 |
| | *MEMGENE2 (MG2)* | -0.18 | 0.25 | -0.70 | 0.483 |
| | *MEMGENE3 (MG3)* | 0.26 | 0.25 | 1.04 | 0.297 |
| | *MG1 x Zebularine* | -0.16 | 0.34 | -0.48 | 0.631 |
| | *MG2 x Zebularine* | 0.03 | 0.31 | 0.09 | 0.930 |
| | *MG3 x Zebularine* | 0.10 | 0.32 | 0.32 | 0.748 |
| | *SD Maternal lines* | 0.71 | | | |
| | *SD Population* | 0.60 | | | |
| | *AICc* = 325.73; *R2m* = 0.02; *R2c* = 0.13 | | | | |
| loc176 (FAM-AAC-CCT-176; MSAP-n) | *Intercept* | -0.80 | 0.67 | -1.19 | 0.233 |
| | *Zebularine* | 0.87 | 0.70 | 1.25 | 0.213 |
| | *MEMGENE1 (MG1)* | -0.56 | 0.34 | -1.64 | 0.101 |
| | *MEMGENE2 (MG2)* | 0.25 | 0.23 | 1.08 | 0.281 |
| | *MEMGENE3 (MG3)* | 0.34 | 0.33 | 1.05 | 0.294 |
| | *MG1 x Zebularine* | 0.42 | 0.39 | 1.08 | 0.282 |
| | *MG2 x Zebularine* | -0.36 | 0.25 | -1.45 | 0.148 |
| | *MG3 x Zebularine* | -0.36 | 0.35 | -1.01 | 0.314 |
| | *SD Maternal lines* | 0.59 | | | |
| | *SD Population* | 0.73 | | | |
| | *AICc* = 462.46; $R^2m$ = 0.02; $R^2c$ = 0.20 | | | | |

## Supplementary Material

## Continuation of Table S4

| Locus | Term | *Estimate* | *SE* | *z* | *p* |
|---|---|---|---|---|---|
| loc222 (PET-AGG-CGA-222; MSAP-n) | *Intercept* | 0.60 | 0.51 | 1.18 | 0.237 |
| | *Zebularine* | -0.97 | 0.74 | -1.32 | 0.188 |
| | *MEMGENE1 (MG1)* | -0.21 | 0.15 | -1.37 | 0.171 |
| | *MEMGENE2 (MG2)* | 0.02 | 0.15 | 0.15 | 0.883 |
| | *MEMGENE3 (MG3)* | -0.06 | 0.16 | -0.36 | 0.719 |
| | *MG1 x Zebularine* | 0.20 | 0.23 | 0.86 | 0.392 |
| | *MG2 x Zebularine* | 0.12 | 0.22 | 0.55 | 0.583 |
| | *MG3 x Zebularine* | 0.37 | 0.23 | 1.61 | 0.107 |
| | *SD Maternal lines* | 0.00 | | | |
| | *SD Population* | 0.00 | | | |
| | *AICc* = 470.90; *R2m* = 0.03; *R2c* = 0.03 | | | | |
| loc135 (PET-AGG-CGA-135; MSAP-n) | *Intercept* | 2.83 | 0.76 | 3.74 | 0.000 |
| | *Zebularine* | -1.26 | 0.95 | -1.32 | 0.188 |
| | *MEMGENE1 (MG1)* | -0.51 | 0.23 | -2.26 | 0.024 |
| | *MEMGENE2 (MG2)* | -0.10 | 0.20 | -0.50 | 0.615 |
| | *MEMGENE3 (MG3)* | -0.12 | 0.24 | -0.52 | 0.601 |
| | *MG1 x Zebularine* | 0.65 | 0.31 | 2.10 | 0.036 |
| | *MG2 x Zebularine* | -0.03 | 0.28 | -0.10 | 0.922 |
| | *MG3 x Zebularine* | 0.15 | 0.30 | 0.49 | 0.624 |
| | *SD Maternal lines* | 0.75 | | | |
| | *SD Population* | 0.25 | | | |
| | *AICc* = 359.00; $R^2m$ = 0.03; $R^2c$ = 0.12 | | | | |

**Supplementary Material**

**Table S5.** Jointly detected outlier loci and their applied corresponding genome-scan approaches. For each marker, the applied fluorescence dye and the cutting sequence are given. Environmental variables that helped detecting outlier loci, are given in bold and were included as fixed factors in logistic mixed-effects models (see Tables S3 and S4). Details of the genome-scan approaches are given in the main manuscript. Abbreviations: bp – base pair size of outlier locus.

| Marker | Locus ID | Dye | Sequence | bp | Genome scan approach | Environmental variables |
|---|---|---|---|---|---|---|
| AFLP | loc58 | FAM | AAC-CCT | 58 | BayeScEnv and LEA | **PC1, PC2,** PC3 |
| | loc286 | FAM | AAC-CCT | 286 | BayeScEnv and LEA | **PC1, PC2,** PC3 |
| MSAP-m | loc189 | VIC | ACG-CAT | 189 | BayeScEnv and LEA | PC1, **PC2, PC3** |
| | loc282 | FAM | AAC-CCT | 282 | RDA and LEA | **MEMGENE1, MEMGENE2, MEMGENE3** |
| MSAP-n | loc135 | PET | AGG-CGA | 135 | RDA and LEA | **MEMGENE1, MEMGENE2, MEMGENE3** |
| | loc176 | FAM | AAC-CCT | 176 | RDA and LEA | **MEMGENE1, MEMGENE2, MEMGENE3** |
| | loc222 | PET | AGG-CGA | 222 | RDA and LEA | **MEMGENE1, MEMGENE2, MEMGENE3** |

**Supplementary Material**

**Figures**

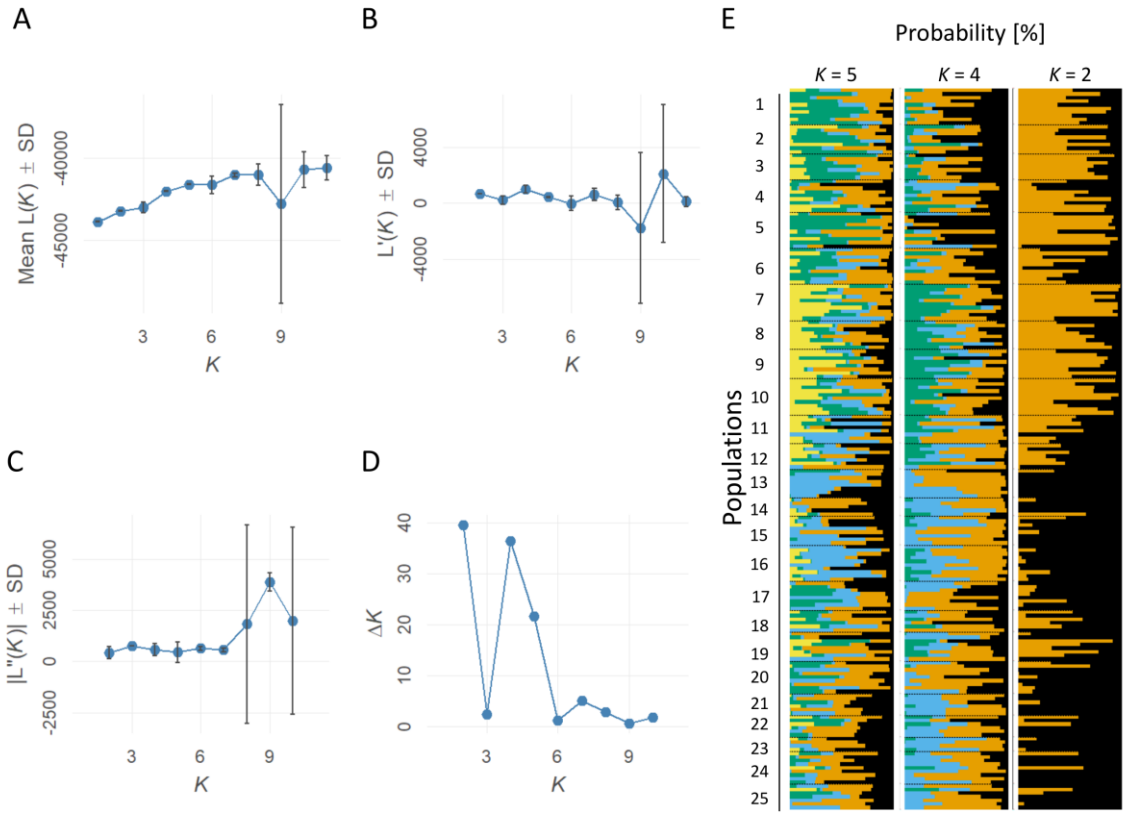


**Figure S1.** STRUCTURE analysis of AFLP loci. STRUCTURE analysis was applied using the admixture model with LOCPRIOR parameter, a burn-in of 1,000,000 and 500,000 repetitions. Diagnostics were based on Evanno et al., (2005), i.e., (**A**) the likelihood distribution as the mean L(*K*) (± SD), (**B**) the rate of change of the likelihood distribution as L'(*K*) = L(*K*) – L(*K*-1), (**C**) the absolute values of the second-order rate of change of the likelihood distribution as |L''(*K*)| = |L'(*K*+1) – L'(*K*)|, and (**D**) Δ*K* from Δ*K* = m| L''(*K*)| / s[L(*K*)], for each simulation of *K*. Based on these diagnostics, (**E**) *K* = 5, *K* = 4, and *K* = 2 were estimated as the most probable number of genetic clusters present in *S. canadensis* populations along the latitudinal gradient.

**Supplementary Material**

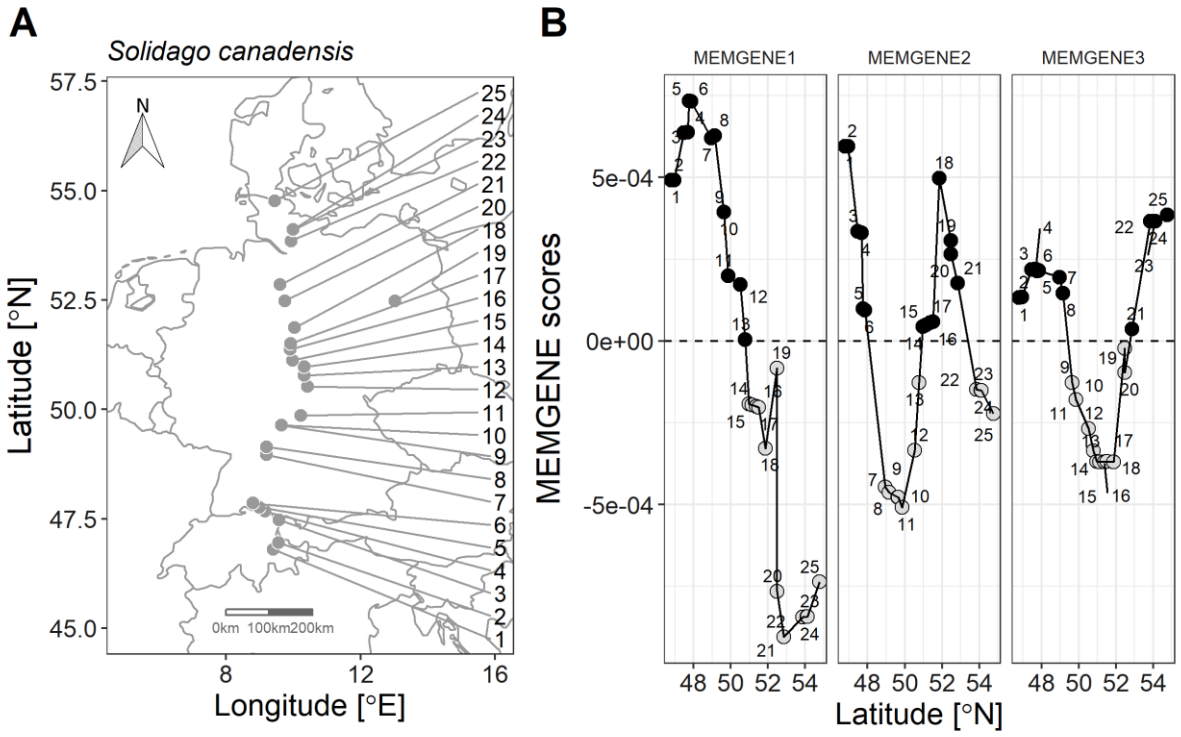


**Figure S2.** Spatial genetic neighborhoods based on Moran's eigenvector maps (MEM). Three autocorrelation (spatial neighborhoods) axes were significantly associated with genetic variation in redundancy analysis using MEMs and transformed to MEMGENE axes using principal components analysis. MEMGENE selection based on (**A**) population coordinates (longitude and latitude in decimal degrees) and (**B**) similar size and color of circles represent similar MEMGENE values for the corresponding populations. In total, MEMGENE analysis of spatial genetic neighborhoods explained 13.2% of genetic variation.

**Supplementary Material**

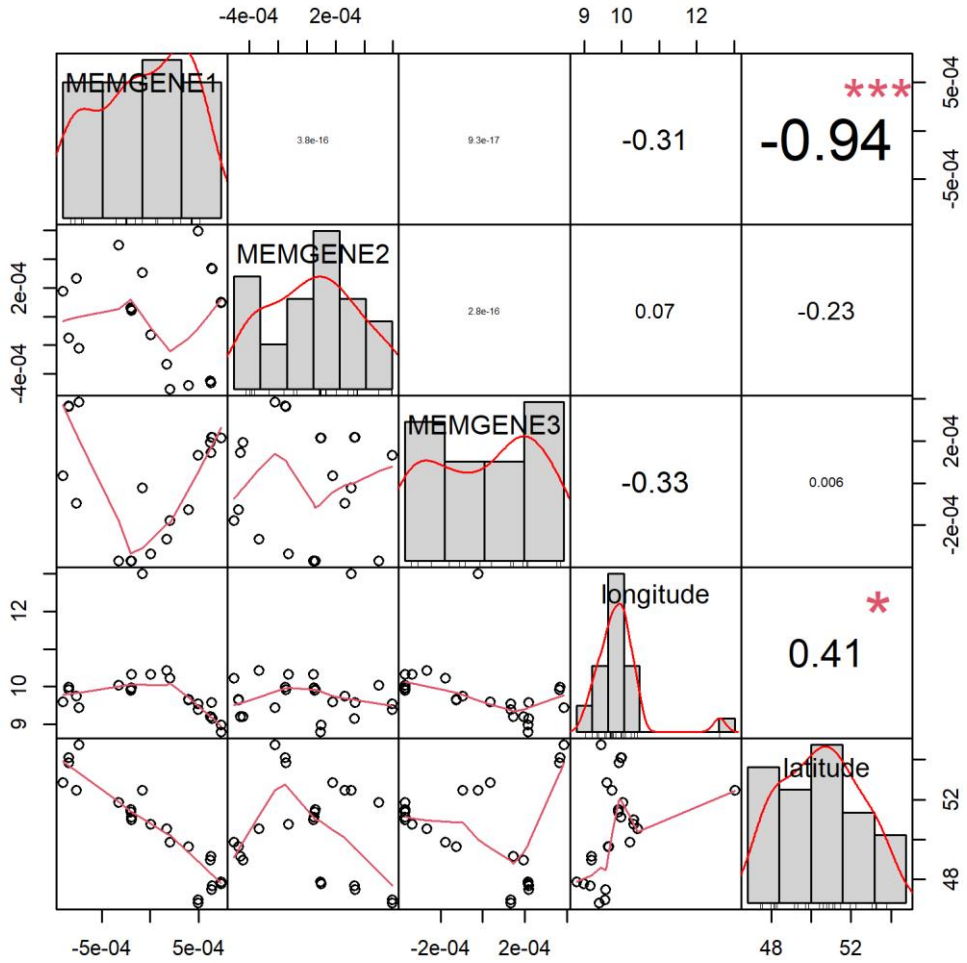


**Figure S3.** Correlation matrix of MEMGENE axes versus population coordinates. The strength and direction of each correlation is given as $R^2$ and was estimated using the Pearson correlation coefficient. Significance level: $p < 0.05$ *; $p < 0.01$ **; $p < 0.001$ ***. Latitude and longitude values are given as decimal degrees.

**Supplementary Material**

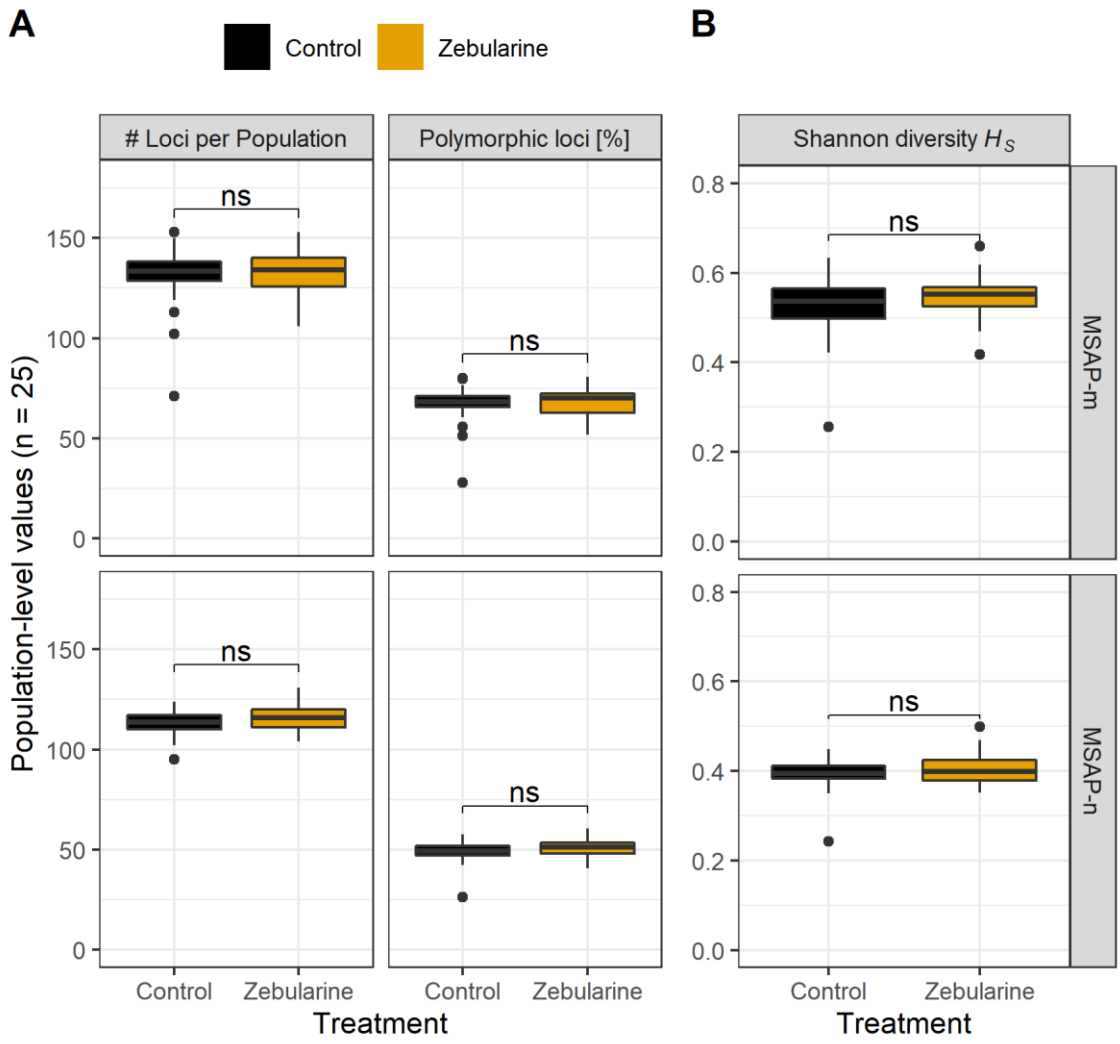


**Figure S4.** MSAP population-level diagnostics using all available samples. Population-level paired *t*-test comparing control versus zebularine-treated individuals was conducted separately on the sets of methylated (MSAP-m) and non-methylated (MSAP-n) loci for (**A**) the total number of loci and the percentage of polymorphic loci, (**B**) the Shannon information criterion ($H_S$). Significance level: $p < 0.05$. Abbreviation: ns – not significant.

**Supplementary Material**

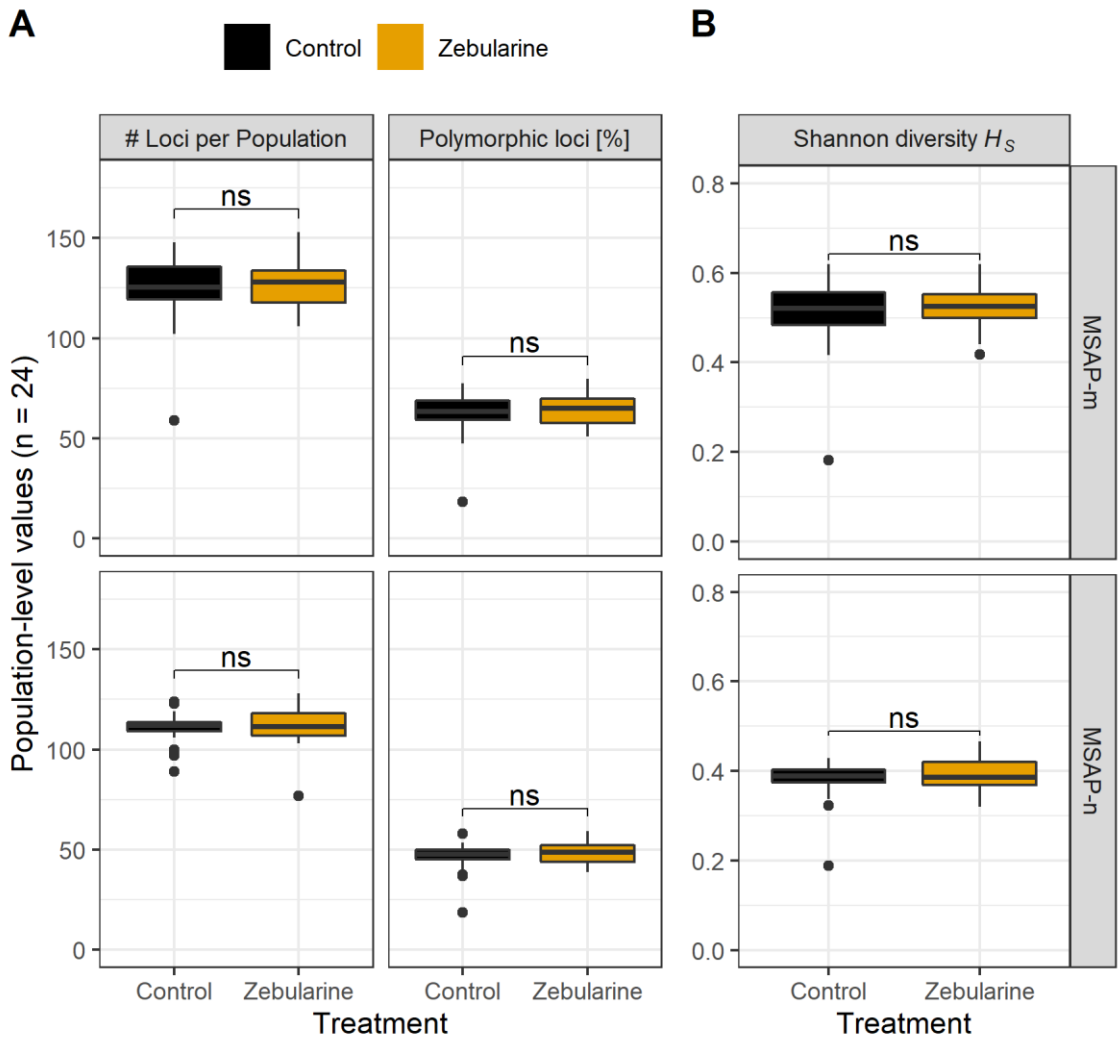


**Figure S5.** MSAP population-level diagnostics using only samples from siblings present in control and zebularine treatment groups. Population-level paired *t*-test comparing control versus zebularine-treated individuals was conducted separately on the sets of methylated (MSAP-m) and non-methylated (MSAP-n) loci for (**A**) the total number of loci and the percentage of polymorphic loci, (**B**) the Shannon information criterion ($H_S$). Please note that population 21 was not included because no individual of this population was present in both treatment groups. Significance level: $p < 0.05$. Abbreviation: ns – not significant.

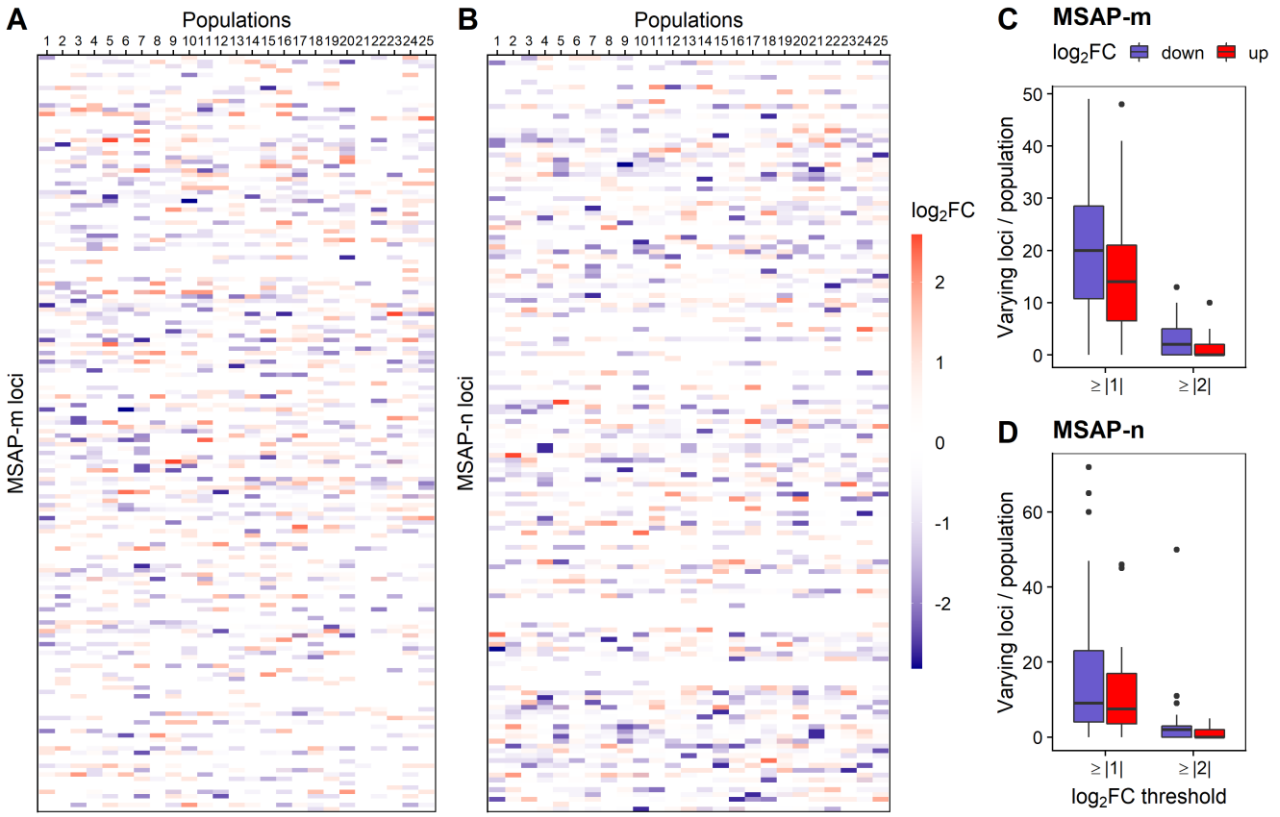


**Figure S6.** Heatmaps of population-level change ($log_2FC$) in MSAP loci due to the zebularine treatment. The increase and decrease, respectively, are shown color-coded as $log_2FC$ per population separately for (**A**) methylated (MSAP-m) and (**B**) non-methylated (MSAP-n) loci. Positive values denote that loci occur more frequently in zebularine-treated individuals compared to control plants and vice versa. A $log_2FC$ value of 2 (or -2) means that the corresponding locus occurs four times more (or less) frequent in the population-level zebularine subgroup. Both **(C)** MSAP-m and (**D**) MSAP-n datasets were filtered separately for loci changing in frequency per population, see box-and-whisker plots. For comparison, loci were filtered for two $log_2FC$ thresholds, i.e., two-fold variance ($log_2FC \geq |1|$) and four-fold variance ($log_2FC \geq |2|$).

**Supplementary Material**

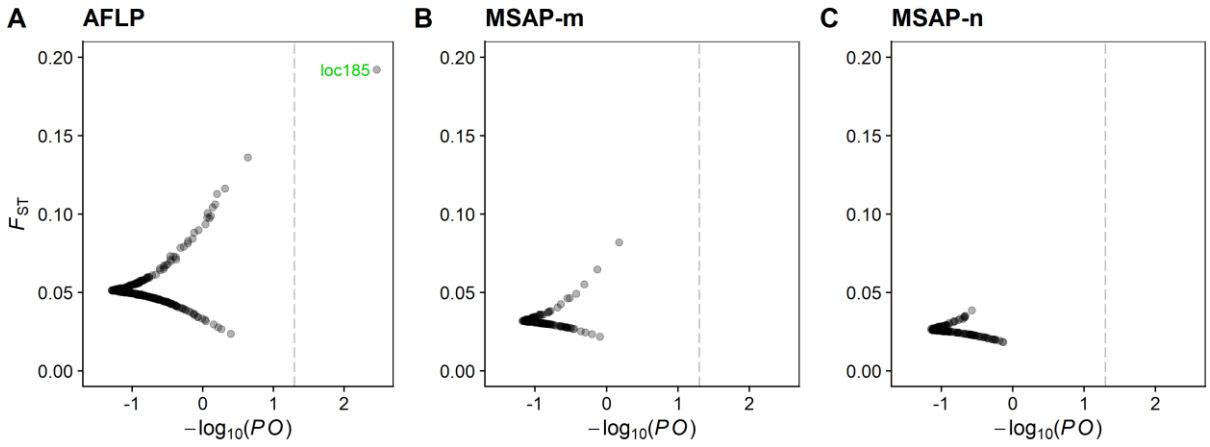


**Figure S7.** BAYESCAN analysis of AFLP/MSAP loci. Outlier screening was conducted separately for (**A**) AFLP, (**B**) MSAP-m, and (**C**) MSAP-n datasets. BAYESCAN was run with 100,000 iterations, 10 repetitions and a burn-in of 50,000 following 20 pilot runs (see Method S2). The dashed line marks the threshold of 1.3 for the false discovery rate.

**Supplementary Material**

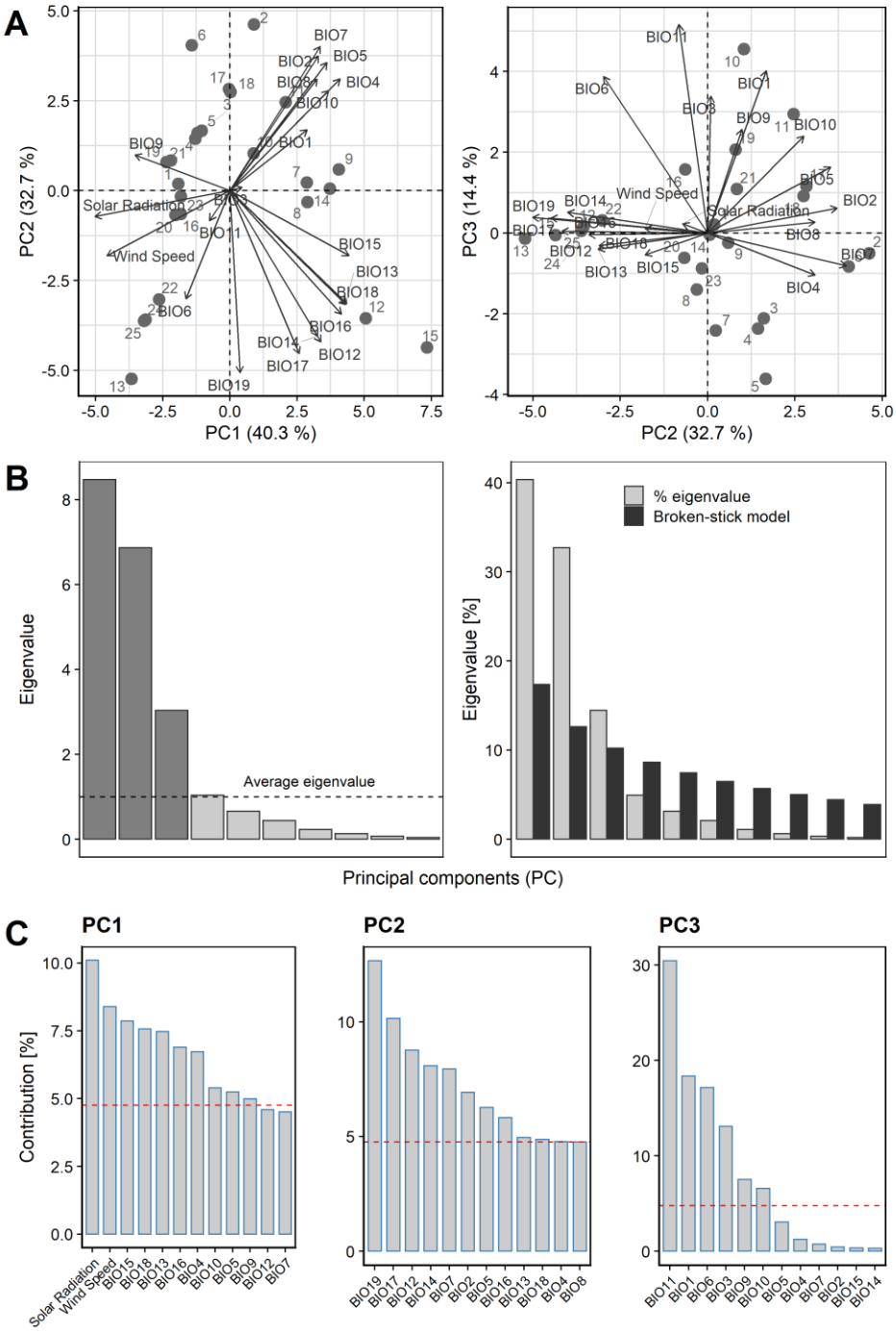


**Figure S8.** Principal components analysis (PCA) of WorldClim 2.0 climate variables. (**A**) Biplots of first-versus-second (PC1, PC2) and second-versus-third (PC2, PC3) PCA axes. Numbers denote the population IDs (see Table S1). (**B**) Eigenvalues of PCs, their corresponding percentage of explained variation [%] against the broken-stick criterion, and (**C**) the WorldClim 2.0 climate variables with highest loadings on each PC axis.

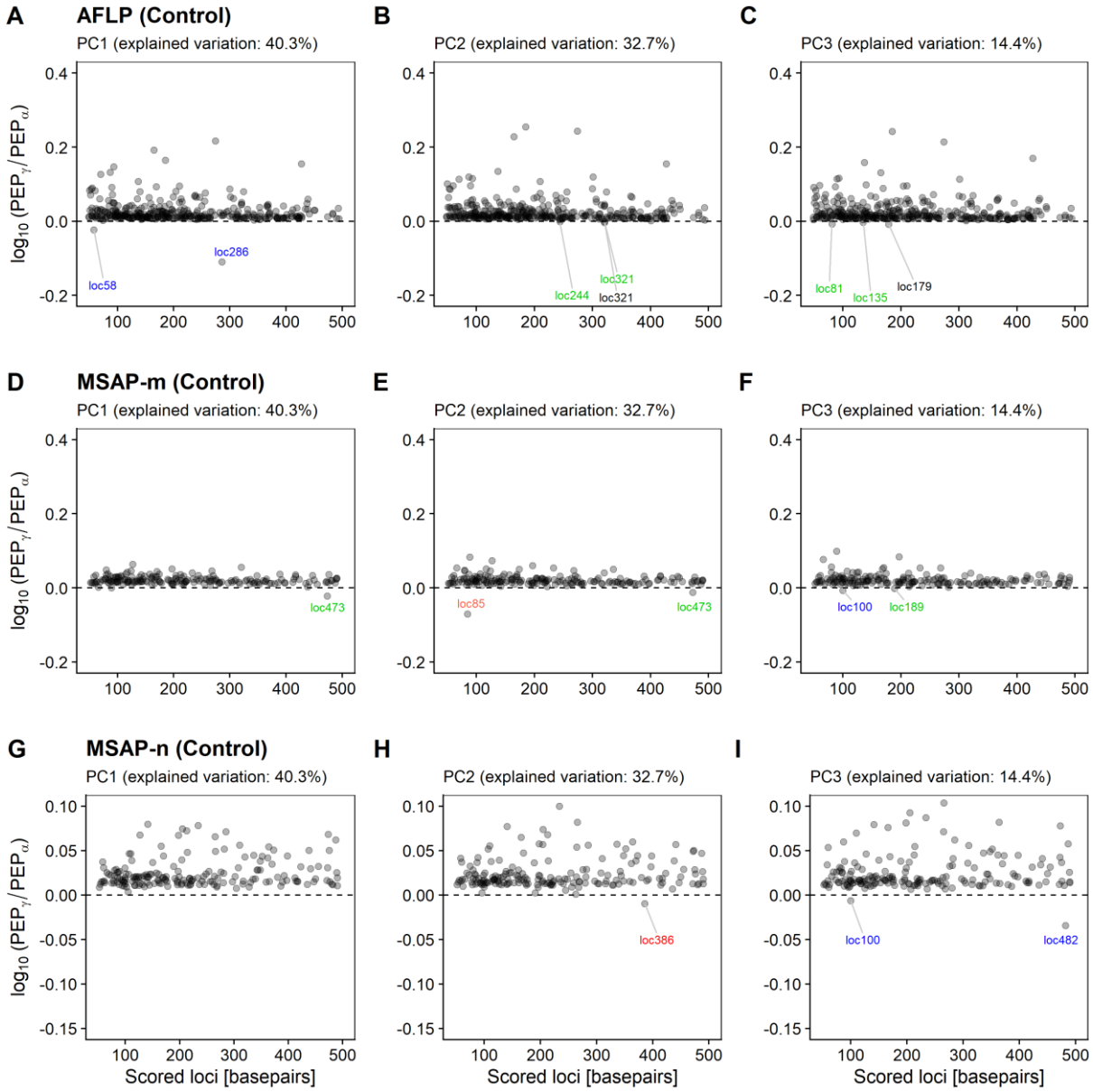


**Figure S9.** BAYESCENV analysis of AFLP/MSAP loci. Colors of loci IDs correspond to the fluorescent dyes from the GENEMAPPER software. Outlier screening was conducted using three principal component axes (**A, D, G**: PC1, **B, E, H**: PC2, **C, F, I**: PC3) separately based on principal components analysis of WorldClim 2.0 variables (Figure S8). BAYESCENV analysis was applied separately on (**A-C**) AFLP, (**D-F**) MSAP-m, and (**G-I**) MSAP-n datasets (see Method S2). In total, the applied PCs explained 87.1% of climatic variation among sampling locations of *S. canadensis* source populations. Only non-treated control plants were used for analysis and detected outlier loci were pooled for subsequent statistical analyses.

**Supplementary Material**

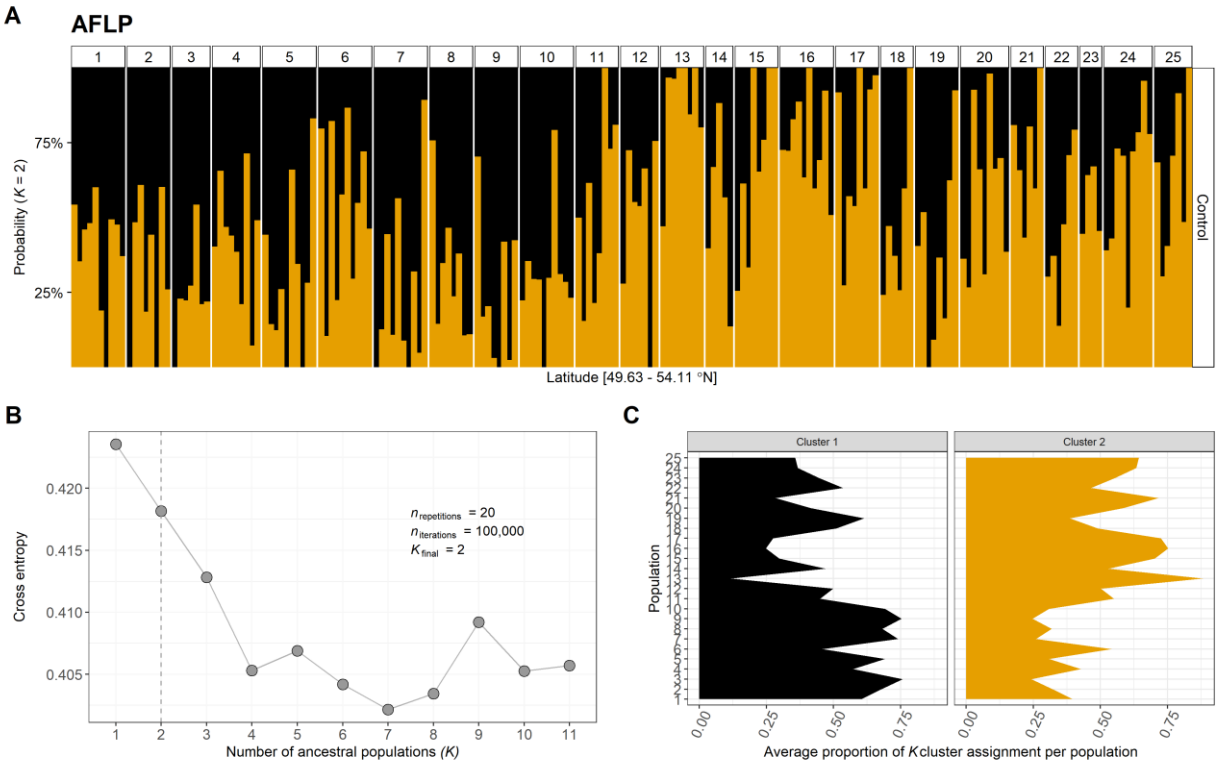


**Figure S10.** LEA analysis of genetic (AFLP) population structure. (**A**) Membership proportions of *K*=2 estimated genetic clusters along the latitudinal gradient under study, (**B**) cross-entropy used for selecting the number of *K* clusters used in subsequent outlier screening with LFMM (Figure S13 and Method S2), (**C**) distribution of the averaged population-level proportion of each *K* cluster assignment along the latitudinal gradient. Only non-treated control plants were used to analyze genetic population structure.

**Supplementary Material**

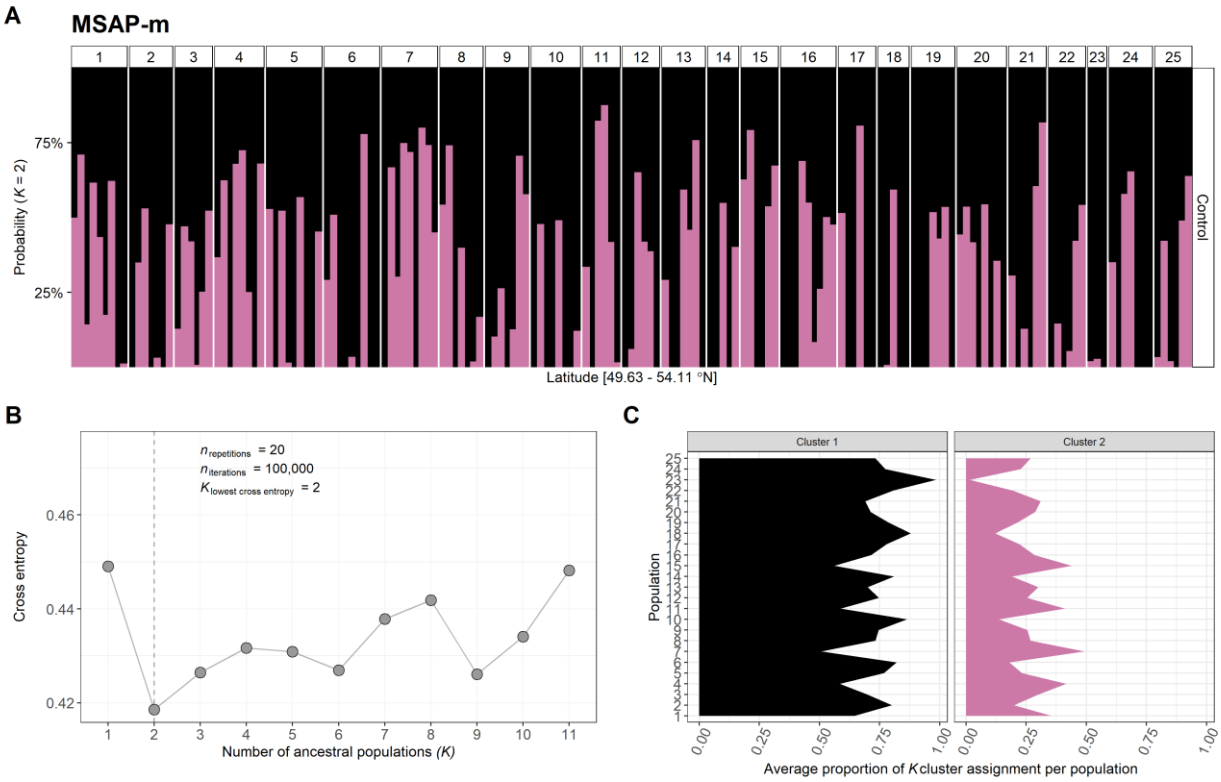


**Figure S11.** LEA analysis of epigenetic (MSAP-m) population structure. (**A**) Membership proportions of *K*=2 estimated genetic clusters along the latitudinal gradient under study, (**B**) cross-entropy used for selecting the number of *K* clusters, (**C**) distribution of the averaged population-level proportion of each *K* cluster assignment along the latitudinal gradient. Only non-treated control plants were used to analyze genetic population structure. Because no pronounced population structure was found with LEA, a final *K* =1 was used in subsequent outlier screening with LFMM (Figure S13 and Method S2).

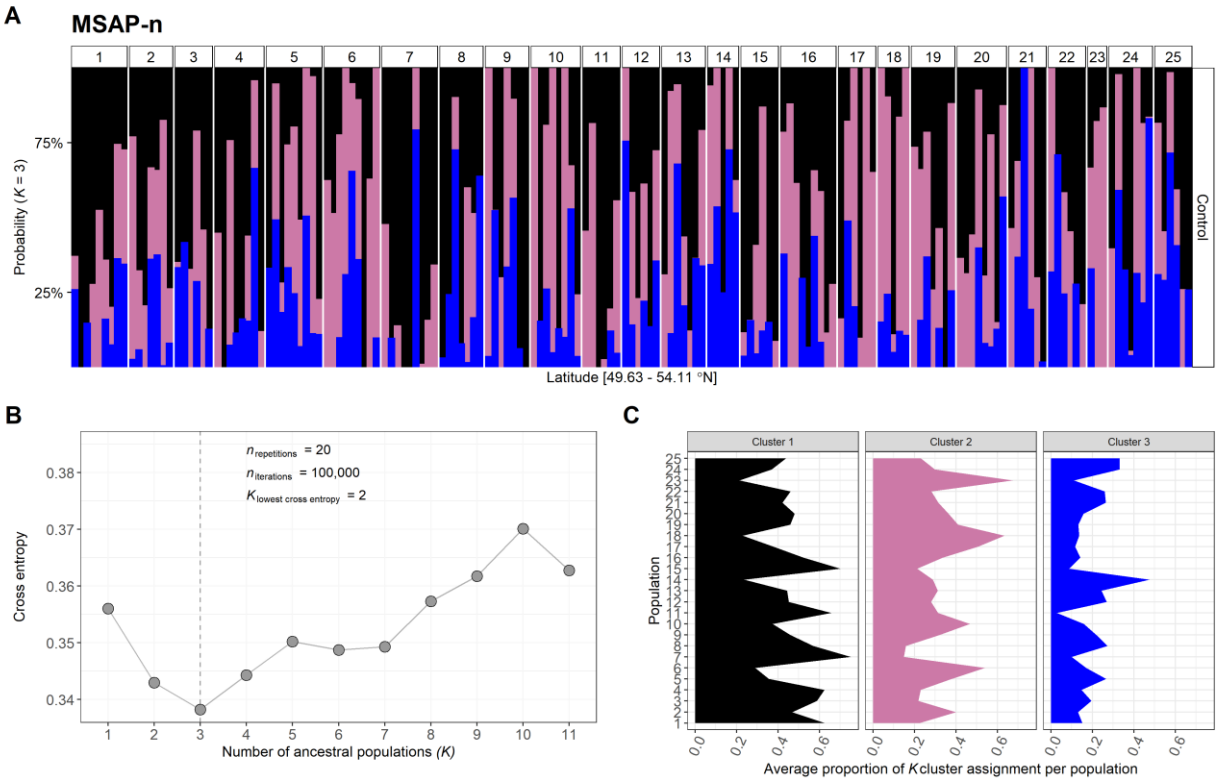


**Figure S12.** LEA analysis of epigenetic (MSAP-n) population structure. (**A**) Membership proportions of *K*=3 estimated genetic clusters along the latitudinal gradient under study, (**B**) cross-entropy used for selecting the number of *K* clusters, (**C**) distribution of the averaged population-level proportion of each *K* cluster assignment along the latitudinal gradient. Only non-treated control plants were used to analyze genetic population structure. Because no pronounced population structure was found with LEA, a final *K* =1 was used in subsequent outlier screening with LFMM (Figure S13 and Method S2). Please note that the colors do not represent the same or similar clusters as in Figure S11 and were chosen only to facilitate visualization in cases of color-deficiency.

**Supplementary Material**

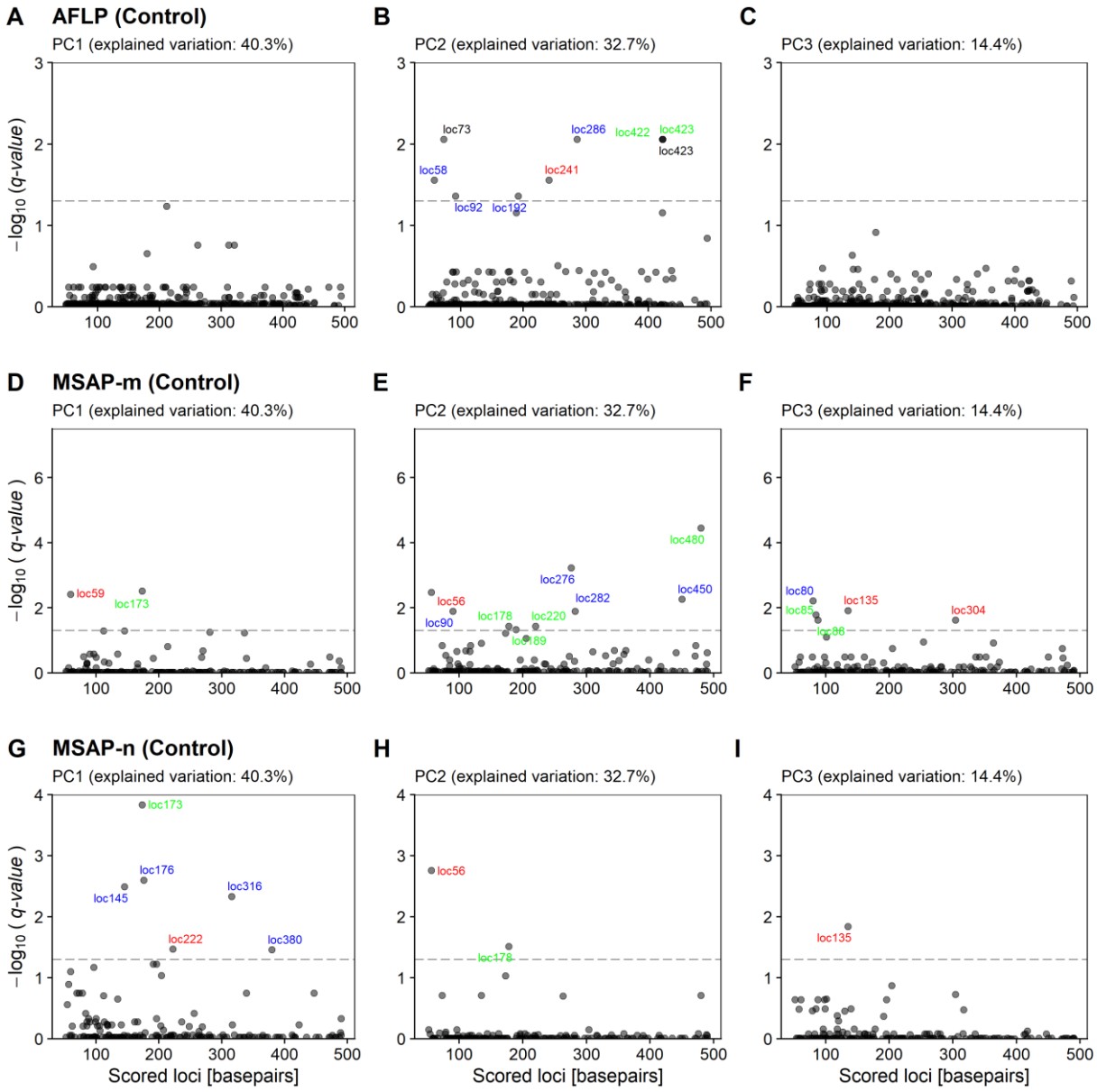


**Figure S13.** LFMM analysis of AFLP/MSAP loci. Colors of loci IDs correspond to the fluorescent dyes represented in the GENEMAPPER software. Outlier screening was conducted using three principal component axes (**A, D, G**: PC1, **B, E, H**: PC2, **C, F, I**: PC3) separately based on principal components analysis of WorldClim 2.0 variables (Figure S8 and Method S2). LFMM analysis was applied separately on (**A-C**) AFLP, (**D-F**) MSAP-m, and (**G-I**) MSAP-n datasets. In total, the applied PCs explained 87.1% of climatic variation among sampling locations of *S. canadensis* source populations. Only non-treated control plants were used for analysis and detected outlier loci were pooled for subsequent analyses.

**Supplementary Material**

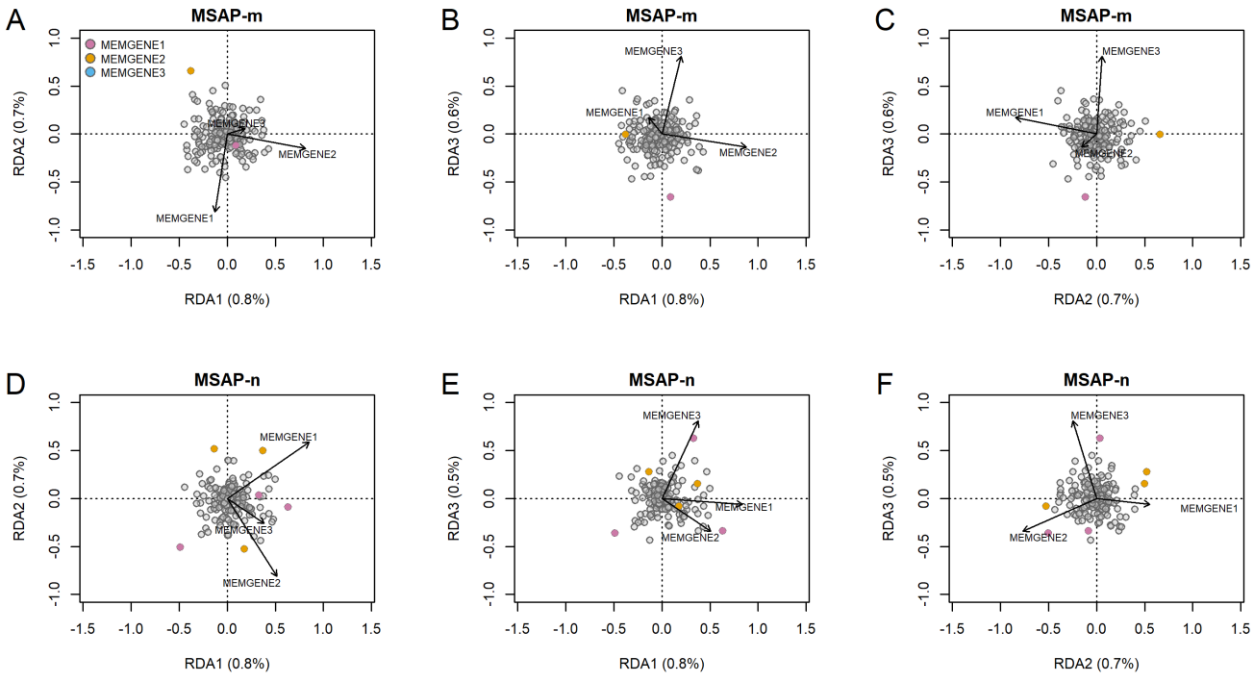


**Figure S14.** Redundancy analysis (RDA) of MSAP loci. RDA was conducted on methylated MSAPs (MSAP-m: **A**-**C**) and non-methylated MSAP (MSAP-n: **D**-**F**) separately using spatial genetic neighborhoods from MEMGENE analysis(Figure S2 and Method S2). Outlier loci were colored based on their highest loading on corresponding MEMGENE variables. Biplots showing the percentage of explained variation of (**A, D**) the first-versus-second (**B, E**) first-versus-third, and (**C, F**) second-versus-third RDA axis separately for MSAP-m (RDA1: 0.8%; RDA2: 0.7%; RDA3: 0.6%) and MSAP-n RDA1: 0.8%; RDA2: 0.7%; RDA3: 0.5%).

**Supplementary Material**

**References for Supplementary Material**


Aguirre-Liguori, J. A., Tenaillon, M. I., Vázquez-Lobo, A., Gaut, B. S., Jaramillo-Correa, J. P., Montes-Hernandez, S., Souza, V., & Eguiarte, L. E. (2017). Connecting genomic patterns of local adaptation and niche suitability in teosintes. Mol. Ecol. 26, 4226-4240. doi: 10.1111/mec.14203

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. Mol. Ecol. 14, 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Mol. Ecol. 27, 2215-2233. doi: 10.1111/mec.14584

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. Methods Ecol. Evol. 6, 925-929. doi: 10.1111/2041-210X.12382

Gáspár, B., Bossdorf, O., & Durka, W. (2019). Structure, stability and ecological significance of natural epigenetic variation: A large-scale survey in Plantago lanceolata. New Phytol. 221, 1585–1596. doi: 10.1111/nph.15487

Jeffreys, H. (1939). The Theory of Probability. 3rd Edition. Oxford: Clarendon Press

Lynch, M., & Milligan, B. G. (1994). Analysis of population genetic structure with RAPD markers. Mol. Ecol. 3, 91-99. doi: 10.1111/j.1365-294X.1994.tb00109.x

MacArthur, R. H. (1957). On the relative abundance of bird species. Proc. Natl. Acad. Sci. U. S. A. 43, 293-295. doi: 10.1073/pnas.43.3.293

Pérez-Figueroa, A., García-Pereira, M. J., Saura, M., Rolán-Alvarez, E., & Caballero, A. (2010). Comparing three different methods to detect selective loci using dominant markers. J. Evol. Biol. 23, 2267-2276. doi: 10.1111/j.1420-9101.2010.02093.x