Evidence for rapid evolution in a grassland biodiversity experiment

Sofia J. van Moorsel*[1,6], Marc W. Schmid*[1,2,6], Niels C.A.M. Wagemaker[3], Thomas van Gurp[4], Bernhard Schmid[1] and Philippine Vergeer[3,5,6]

*Equal contribution

[1]Department of Evolutionary Biology and Environmental Sciences, University of Zürich, Winterthurerstrasse 190, CH-8057, Switzerland
[2]MWSchmid GmbH, Möhrlistrasse 25, CH-8006 Zürich
[3]Institute for Water and Wetland Research, Radboud University Nijmegen, Heyendaalseweg 135, NL-6500 GL Nijmegen
[4]Deena Bioinformatics, Haverlanden 237, NL-6708 GL Wageningen
[5]Department of Environmental Sciences, Plant Ecology and Nature Conservation group, Wageningen University, Droevendaalsesteeg 3, NL-6708 PB Wageningen

[6] Authors for correspondence:
Sofia J. van Moorsel, sofia.vanmoorsel@ieu.uzh.ch
Marc W. Schmid, contact@mwschmid.ch
Philippine Vergeer, philippine.vergeer@wur.nl

Summary

- In long-term grassland biodiversity experiments positive effects of biodiversity on plant productivity commonly increase with time. Previously it was shown that differential selection in monoculture and mixed-species grassland communities could lead to the rapid emergence of monoculture and mixture phenotypes. Underlying mechanisms for such rapid phenotypic responses are however still unclear.
- We hypothesized that in biodiversity experiments pre-adapted genotypes or epigenetic variants could be sorted out from the standing genetic or epigenetic variation.
- To test if biodiversity acted as a selective environment, we grew offspring from plants that were exposed for twelve years to a monoculture or mixture environment in a biodiversity experiment (Jena Experiment) under controlled greenhouse conditions. Using epiGBS, a genotyping-by-sequencing approach combined with bisulfite conversion to provide integrative genetic and epigenetic data, we showed that plants with a monoculture or mixture background were genetically distinct.
- Our data reveal strong genetic and epigenetic divergence within species according to selection history and suggest variation in epigenetic variation to be largely under genetic control. This pattern was consistently observed across six perennial grassland species. Our results suggest that selection of genetic variants caused the rapid emergence of monoculture and mixture types in the Jena Experiment.

Key words: biodiversity, epiGBS, epigenetic variation, genetic divergence, herbaceous plant species

Introduction

The Earth's biosphere is currently challenged by the impacts of anthropogenic environmental change and plant populations may encounter new abiotic or biotic environment due to climate-induced range shifts (Ouborg *et al.,* 2006). The unprecedented rate of environmental change raises the question whether natural communities can adapt fast enough to novel abiotic or biotic conditions. Biodiversity has been shown to buffer ecosystem towards climatic extremes and novel environmental conditions (Isbell *et al.,* 2015) and in addition it was shown that co-evolution among plants comprising a community could buffer the impact of an extreme climatic event (van Moorsel *et al.*, 2018a).

Adaptive responses of plant populations to environmental factors (e.g., Schmid 1985; Joshi *et al*., 2001) and biotic interactions (Gervasi & Schiestl 2017) are well studied, but little effort has been devoted to studying the influence of community diversity on population structure and productivity (but see Lipowsky *et al.,* 2011; Kleynhans *et al.,* 2016). In particular, the influence of multi-species interactions for the adaptive response of a species is largely unknown, despite a growing body of evidence pointing towards the importance of species-interaction networks for the maintenance of ecosystem stability (Bastolla *et al.,* 2009). It is conceivable that the feedback between species interactions and their adaptive responses shapes community-level ecosystem functioning (van Moorsel *et al.*, 2018).

In the 1960s it was proposed that there are large differences in the time scales between ecological and evolutionary processes (Slobodkin 1961), but now it is well known that micro-evolutionary and ecological processes can occur on the same temporal scale (reviewed by Hairston *et al.,* 2005; Schoener 2011). Thus, it appears that micro-evolutionary processes may allow for an evolutionary rescue in a rapidly changing environment. However, evidence of genetic adaptation to mild stresses in natural situations over short time scales of a few generations is still scarce. Understanding how biodiversity, i.e. interactions between species, shapes this evolutionary response will be instrumental to anticipate how ecosystems may change in response to global change.

Adaptation depends on several factors. For organisms with a short generation time and asexual reproduction, such as clonal populations of bacteria, mutations and horizontal gene transfer are the main sources of genetic variation (Anderson *et al.*,

2011). However, for species with longer generation times such as perennial plants selection more likely acts on standing genetic variation (Barrett & Schluter 2008), resulting in a sorting-out of suitable genotypes (Fakheran *et al.,* 2010). Furthermore, plants may adjust to a novel environment by phenotypic plasticity (Price *et al.,* 2003).

A possible indication for selection in grassland plant communities is the observed strengthening of biodiversity effects in field biodiversity experiments (Cardinale *et al.,* 2007; Fargione *et al.,* 2007; Reich *et al.,* 2012; Meyer *et al.,* 2016). An increase in complementarity between species resulted in an increasing positive effect of diversity on productivity over time (e.g. Meyer *et al.,* 2016). The increased complementarity could also have been a result of phenotypic plasticity (Ghalambour *et al.,* 2007). However, recent common garden experiments with plant material from a grassland biodiversity experiment (the Jena Experiment, Roscher *et al.,* 2004) gave a clear indication for genetic divergence between monoculture and mixture types in multiple plant species (Zuppinger-Dingley *et al.,* 2014; van Moorsel *et al.,* 2018b), suggesting that natural selection in response to community diversity had occurred in the field. As a consequence, mixture types showed stronger complementarity effects (Loreau & Hector 2001) than monoculture types when grown in mixture, i.e. community productivity was increased due to increased differences in functional traits between species for mixture types (Zuppinger-Dingley *et al.,* 2014; van Moorsel *et al.,* 2018b). The emergence of such monoculture and mixture types, with different growth performance and plant functional trait variation, thus suggested that community diversity in the field likely acted as a selective environment (Zuppinger-Dingley *et al.,* 2014; Rottstock *et al.,* 2017; van Moorsel *et al.,* 2018b). However, direct evidence for a genetic divergence between the different populations in the field experiment is still missing.

It should be noted that epigenetics, here defined as meiotically heritable changes in gene expression without changes to the underlying DNA sequence (Verhoeven *et al.,* 2016), has also been proposed to play a role (Bird 2007; Bossdorf *et al.,* 2008; Tilman & Snell-Rood 2014). In a comment accompanying the publication of Zuppinger-Dingley *et al.* (2014), Tilman and Snell-Rood wrote: "[…] chance that the differences between the high- and low-diversity selection groups were due to genetic divergence. However, it is possible that epigenetic factors […] could have had a simultaneous role" (Tilman & Snell-Rood 2014). Recent work on epigenetic recombinant inbred lines (epiRILs) of *Arabidopsis thaliana* suggests a considerable contribution of

induced epialleles to phenotypic variation, which is independent of genetic variation (Cortijo *et al.*, 2014; Kooke *et al.*, 2015). However, the importance of epigenetics in natural populations and whether it contributes to adaptation remains elusive (Quadrana & Colot 2016), in particular because it is difficult to separate epigenetic from genetic variation. A first step to disentangle genetic from epigenetic variation was achieved with apomictic clones of *Taraxacum officinale*, suggesting that differences in flowering time were mediated by differences in DNA methylation (Wilschut *et al.,* 2016). However, given the fundamental difference between apomixis (apomeiosis, parthenogenesis, autonomous endosperm formation) and sexual reproduction, the results from this study may not be directly transferable to non-apomictic plant species.

Here, we tested whether community diversity could act as a selective force leading to a divergence of phenotypes into monoculture and mixture types (Zuppinger-Dingley *et al.*, 2014; van Moorsel *et al.,* 2018b). In particular, we tested the hypothesis that the differentiation of plants into phenotypically distinct mixture (exhibiting stronger biodiversity effects) and monoculture types (exhibiting weaker biodiversity effects) within species was paralleled by genetic differentiation. Anticipating that the differentiation might also have been be due to epigenetic variation, we chose a method which allowed us to analyze genetic and epigenetic variation between plants of six species, all herbaceous perennials of European grassland.

Materials and Methods

Plant selection histories

To test whether plant types selected over eleven years in mixtures differ genetically or epigenetically from those types selected in monocultures, we chose six species grown in monoculture and mixture plots in the Jena Experiment (Jena, Thuringia, Germany, 51˚N, 11˚E, 135 m a.s.l., see Roscher *et al.* (2004) for experimental details). The following species belonging to three functional groups were selected: The three small herbs *Plantago lanceolata* L.*, Prunella vulgaris* L. and *Veronica chamaedrys* L., the tall herb *Galium mollugo* L.*,* and the two legumes *Lathyrus pratensis* L. and *Onobrychis viciifolia* Scop. For the experiment plants from three different selection histories were used (Fig. 1). Plants without a selection history in the experimental

148 field plots of the Jena experiment were obtained from commercial seed suppliers
149 (Rieger Hoffmann GmbH, Blaufelden-Raboldshausen, Germany and Otto Hauenstein
150 Samen AG, Switzerland), who also provided the seeds for the original set up of the
151 Jena Experiment in 2002. Plants with a selection history in either mixture or
152 monoculture had been growing in the Jena Experiment since 2002. In spring 2010,
153 plant communities of 48 plots (12 monocultures, 12 two-species mixtures, 12 four-
154 species mixtures and 12 eight-species mixtures) of the Jena Experiment were
155 collected as cuttings, multiplied by repeated further cutting and then transplanted into
156 an experimental garden in Zurich, Switzerland, in the identical plant composition as
157 the original experimental plots for the first controlled sexual reproduction among co-
158 selected plants (Zuppinger-Dingley *et al.*, 2014). In spring 2011, the seedlings
159 produced from the seeds of the first controlled sexual reproduction in Zurich were
160 transplanted back into those plots of the Jena Experiment from where the parents had
161 originally been collected. In these newly established plots, plant communities with an
162 identical composition to the original communities were maintained for three years
163 until 2014 (van Moorsel *et al*., 2018c).
164

165 ## Seed collection

166 To ensure a second sexual reproductive event for the collection of seed material,
167 entire plant communities from some of the experimental plots replanted in Jena in
168 2011 were excavated in March 2014 and used to established new 1-m^2 plots in the
169 experimental garden in Zurich. For this purpose, we excavated blocks of soil (0.5 m^2)
170 from plots in Jena including plant vegetation, seedlings and dormant seeds to make
171 sure we transferred the entire plant community. These blocks of soil were then
172 transferred and spread out into the plots in the experimental garden. We added a 30
173 cm layer of soil (1:1 mixture of garden compost and field soil, pH 7.4, commercial
174 name Gartenhumus, RICOTER Erdaufbereitung AG, Aarberg, Switzerland) to each
175 plot to make sure the plants established. As for the first sexual reproduction event in
176 2010, netting around each plot minimized the possibility of cross-pollination between
177 the same species from different selection histories. Seeds from this second controlled
178 sexual reproduction event were collected throughout the growing season of 2014 from
179 monoculture plots and 4- and 8-species mixture plots. The exact community
180 composition of the plots the seeds originated from is listed in Table S2. Seeds from

different mother plants were pooled together. The dry seeds were stored at 5° C for cold stratification until germination.

Common garden experiment

Seeds were germinated in December 2014 with those from the same species being planted within the same day. Germination was done in germination soil ("Anzuchterde", Ökohum, Herbertingen, Germany) under constant conditions in the glasshouse without additional light. Seedlings were planted in monocultures of four individuals and mixtures of two plus two individuals into 2-L pots filled with agricultural soil (50% agricultural sugar beet soil, 25% perlite, 25% sand; Ricoter AG, Aarberg, Switzerland). See van Moorsel *et al*. (2018b) for the experimental design and the planted combinations. In total, we planted 36 monoculture assemblies and 81 mixture assemblies from mixture history, 48 monoculture assemblies and 159 mixture assemblies from monoculture history and 33 monoculture assemblies and 100 mixture assemblies from seedlings without a common selection history. Every species combination was replicated, if possible, six times for each selection history (resulting in 457 pots, referred to as "assemblies", and 1828 plants). Seedlings that died within the first two weeks were replaced with seedlings of the same age.

The experiment was set up in six blocks with each block representing a replicate. Every block contained ca. 80 pots and within each block, pots were placed in the glasshouse in a randomized fashion without reference to selection history or species assembly. Single pots always contained four plants of a single selection history. During the experiment, plants were watered according to demand and grown at constant temperatures (17-20°C during the day, 13-17°C during the night) with no additional light added. The plants were not fertilized. The fungicide Fenicur (*Oleum foeniculi*, Andermatt Biocontrol) was applied on 5 May to control powdery mildew (*Podosphaera* spp.).

Sampling

Samples for subsequent epigenetic and genetic analysis were harvested between 18 and 28 May 2015, after twelve weeks of plant growth in the glasshouse. All four plants were sampled in each pot. One young leaf per plant was cut from the living

plant and immediately shock-frozen in liquid nitrogen. The samples were then stored at –80° C in Zurich before shipment to the Netherlands for further processing.

Genetic analysis

We measured both genetic and epigenetic variation in monoculture- and mixture-type plants when propagated in monoculture and two-species mixed assemblies using a novel reference-free bisulfite method (van Gurp *et al.*, 2016). This method allows measuring DNA cytosine methylation levels and identifying single nucleotide polymorphisms with merely one lane of sequencing per 96 samples.

Sample procession and library preparation

The epiGBS protocol used is a further developed protocol based on the protocol of van Gurp *et al.* (2016). The main improvements are the used enzyme combination, a "wobble" adapter facilitating the computational removal of PCR duplicates, and a conversion control nucleotide that allows easier Watson / Crick identification. For 348 samples, (Csp6I/NsiI) epiGBS libraries were created and sequenced on 4 Hiseq 2500 lanes. These samples were divided over six species and three selection histories (see Tables S1, S2).

DNA extraction

Plant material was disrupted by bead-beating frozen leaf tissue in a 2 mL eppendorf tube with 2-3 mm stainless steel beads. No more than 100 mg of fresh tissue was used per sample. DNA isolation was performed using the NucleoSpin® 8 Plant II Core Kit (740669.5 Macherey Nagel). We followed the manufacturers protocol with the following modifications. Cell lysis was performed using Cell lysis buffer PL1 for 30 instead of 10 min. After lysis and initial centrifugation, the lysate was carefully pipetted to a fresh 2.5 mL tubes, avoiding the cell debris. An extra centrifugation of 5 min at 18.000 g step was performed and the lysate transferred to a 96-well rack for the next steps. At the step were the washed columns were dried, we centrifuged 5 min at 4800 g to get rid of the last remaining wash buffer. The restriction enzymes of the first step of the epiGBS are very sensitive to ethanol and other contamination. DNA concentration was determined using Qubit® 2.0 Fluorometric dsDNA HS Assay Kit (Q32851 Life technologies).

DNA digestion and adapter ligation

Per individual, 30–300 ng genomic DNA was digested overnight (17 hrs) at 37° C in a volume of 40 µL containing 1x FD buffer (Thermo Scientific), and 2 µL of both *Csp*6I (FD0214, Thermo Scientific) and *Nsi*I (R0127S, NEB). Following digestion, barcoded "wobble" adapters were ligated to the fragments (Fig. S1). To minimize the possibility of misidentifying samples as a result of sequencing or adapter synthesis error, all pair-wise combinations of barcodes differed by a minimum of three mutational steps. Barcode lengths were modulated from 4 bp to 6 bp to maximize the balance of the bases at each position in the overall set. For the ligation, 4 µL of a sample specific barcode combination of both BA and CO adapters (600 pg/µL), 6 µL T4 DNA ligase buffer, 1 µL T4 DNA ligase (M0202M, NEB) and 5 µL of distilled water were added to the digestion mix to a total volume of 60 µL. Ligation was performed for 3 hours at 22° C followed by 4° C overnight.

Pooling, clean-up and nick translation

In order to assess the quality of libraries, the pooling was performed per species level in batches of around 12 samples per pool. When pooled, the total volume of the pool was reduced by Qiaquick PCR cleanup (28104, Qiagen) to 40 µL. The libraries were size selected by a 0.8x Agencourt AMPure XP (A63880, Beckman coulter) purification favouring > 200 bp DNA fragments and eluted in a total volume of 22 µL. To prevent the formation of adapter dimers, the barcoded adapters were not phosphorylated. Therefore, after the ligation, the insert DNA fragment–adapter connection was nicked at the 3' positions of the insert DNA fragment (Fig. S1). This nick is repaired by Nick translation that recreates the total non-(5mC) methylated adapter strand and during that process also "unwobbles" the adapters since the removed nucleotides are replaced by complementing nucleotides. This nick repair prevents the partial loss of the adapter during bisulfite treatment. The nick translation reaction (1 hour at 15° C) was performed in a reaction of 25 µL containing 19.25 µL of the purified library, 2.5 uL of 10 mM 5-methylcytosine dNTP Mix (D1030 Zymo research), 2.5 uL NEBuffer 2 and 0.75 uL DNA polymerase I (M0209, NEB).

Bisulfite conversion

For bisulfite conversion of non-methylated cytosines, 20 µL of the nick-translated library was used. Bisulfite treatment was performed using the EZ DNA Methylation-Lightning™ Kit (Zymo Research) with the following program according to the manufacturers protocol: 8 min at 98° C, 1 h at 54 °C followed by up to 20 h at 4° C.

epiGBS PCR

Library amplification was performed in four individual 10 µL reactions containing 1 µL ssDNA template, 5 µL KAPA HiFi HotStart Uracil+ ReadyMix (Kapabiosystems), 3 pmol of each illumina PE PCR Primer (5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3' and 5'-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-3'). Temperature cycling consisted of 95° C for 3 min followed by 18 cycles of 98° C for 10 s, 65° C for 15 s, 72° C for 15 s with a final extension step at 72° C for 5 min. Replicate PCR products were pooled and quantified using a Qubit® dsDNA HS Assay Kit (Life technologies). The quality of the Libraries was assessed by analyzing 1 µL on a High Sensitivity DNA chip on a 2100 Bioanalyzer system (Agilent). Libraries were considered suitable for sequencing if the majority of DNA fragments were between 150–400 bp. When the libraries passed quality control, they were pooled according to concentration and number of samples in the species pool, so that each individual sample was expected to yield an equal number of clusters on the Illumina flow cell. Before sequencing, the libraries were spiked with 10% PhiX control to increase the complexity of the libraries.

Sequencing

Finally, Paired-End sequencing was performed on a Hiseq2500 sequencer using the HiSeq v4 reagents and the latest version of the HiSeq Control Software (v2.2.38), which optimizes the sequencing of low-diversity libraries (http://res.illumina.com/documents/products/technotes/technote-hiseq-low-diversity.pdf). As the first five cycles of a sequencing run are used to calculate the color matrix, our barcode design achieves almost perfect balance of the first five nucleotides when equal numbers of sequences are obtained per "A" barcode. The "B" barcodes do not have this requirement; hence same length barcodes were used.

Sequencing reads were demultiplexed and deposited at SRA (SRA accession ID SRP132258). The barcode files can be accessed on Zenodo (DOI 10.5281/zenodo.1167563). In total, we generated 132'884'678, 130'721'280, 81'376'294, 140'883'779, 215'473'088, and 137'868'869 short reads for *G. mollugo*, *L. pratensis*, *O. viciifolia*, *P. lanceolata*, *P. vulgaris*, and *V. chamaedrys*, respectively.

Statistical analysis

De-multiplexing, *de novo* reference construction, trimming, alignment, strand-specific variant calling, and methylation calling were done for each species as described in van Gurp *et al.* (2016). *De novo* reference sequences were annotated with DIAMOND (protein coding genes; NCBI non-redundant proteins as reference; version 0.8.22; Buchfink, Xie, and Huson 2015) and RepeatMasker (transposons and repeats; *Embryophyta* as reference "species"; version 4.0.6; Smit, Hubley, and Green 2013–2015). We summarized the transposable element and repeat classes into "transposons" comprising DNA, LTR, LINE, SINE, and RC transposon, and "repeats" including satellite, telomeric satellite, simple, rRNA, snRNA, unknown, and unclassified repeats. The annotation was then used to classify the genetic and epigenetic variants into the different feature contexts (e.g., to identify whether a single nucleotide polymorphism is located in a gene or a transposon). A summary of the reference sequences is given in Table S3.

Genetic variation: visualization of genetic distances with single nucleotide polymorphisms (SNPs)

The following individuals with a SNP calling rate below 30% were *a priori* removed from the analysis of genetic variation: "pool_pla_lan_15", "pool_pla_lan_46", "pla_lan_81", "pla_lan_82", "pla_lan_52", "pla_lan_83", "pla_lan_95", "pla_lan_111", "pru_vul_60", "pru_vul_79", "pru_vul_80", "pru_vul_87", "pool_pru_vul_22", "pool_pru_vul_24", "ono_vic_15", and "ver_cha_73". These samples were well distributed across the experimental group, i.e., one or two for a single experimental group, except for the mixture-type mixture assembly of *P. lanceolata* for which three individuals were removed. For each species, we filtered the genetic variation data for single nucleotide polymorphisms (SNPs) sequenced in all

individuals with a total coverage between 5 and 200. SNPs homozygous for either the reference or the alternative allele in more than 90 % of all individuals were removed as uninformative SNPs. To reduce the impact of false positive SNP calls, we removed all SNPs located in contigs with more than 1 SNP per 50 basepairs because higher rates appear unlikely and may also originate from spurious alignments to the wrong reference contig (considering that the reference contigs represent only a minor fraction of the entire genome, there may be many reads originating from other locations not represented with a reference contig which are still similar enough to (wrongly) align to the reference). This also avoids that few contigs with high SNP rates drive the tests for genetic differentiation.

SNP allele frequencies were scaled with the function "scaleGen" from adegenet (version 2.0.1; Jombart 2008) and genetic distances between the individuals were visualized with t-SNE (van der Maaten and Hinton 2008, van der Maaten 2014). We calculated 100 maps starting from different random seeds and selected the map with the lowest final error. Individual maps were calculated in R with the package Rtsne (version 0.13; van der Maaten and Hinton 2008, van der Maaten 2014). Parameters for the function Rtsne were pca = FALSE, theta = 0, perplexity = 10 (except for *O. viciifolia* for which perplexity = 5). To select the SNPs with the highest differentiation between the populations, we calculated Jost's D (Jost 2008) with the function "basic.stats" from hierfstat (version 0.04-22; Goudet & Jombart 2015) and only included the top 5% in the visualization.

Genetic variation: test for genetic differentiation between populations with single nucleotide polymorphisms (SNPs)

SNP data were processed and filtered as described before. Considering that our design only included two factors with two and three levels (assembly in the glasshouse and selection history respectively; and incomplete for all species except *P. vulgaris*), we did not use a hierarchical model (with assembly nested within) to test for genetic differentiation. Instead, we tested each factor within all levels of the other factor for genetic differentiation. Taking *P. vulgaris* as an example, we tested for genetic differentiation between plant histories within monoculture and mixture assemblies

(between all three histories and between monoculture and mixture types), and between assemblies within the naïve, monoculture, and mixture-type selection histories. For each test, we extracted the corresponding individuals and tested for genetic differentiation with the G-statistic test (Goudet *et al.,* 1996, function gstat.randtest implemented in the package hierfstat, version 0.04-22; Goudet & Jombart 2015). We did not correct for multiple testing because *P*-values were derived from 999 simulations and were thus limited to a minimal value of 0.001. Instead, we used a significance threshold of 0.01 and provide plots showing histograms of permuted G-statistics and the observed G-statistics. This analysis was carried out with the (1) entire data set, (2) SNPs located within genes, and (3) SNPs located within transposons. We chose to separately test SNPs in genes and transposons because we expected that selection more likely acted on genes and that selection of transposons would primarily occur due to genetic linkage to an advantageous gene. In addition, we expected that SNP calls are more reliable within genes because many transposon families tend to be highly repetitive.

To estimate the extent to which the genetic variation was caused by the differentiation between populations, we calculated average (i.e., across all tested SNPs) pairwise F_{ST} values with the function pairwise.fst from the package adegenet (version 2.0.1; Jombart 2008). Assuming that only few loci were under selection, many SNPs have F_{ST} values close to zero (only SNPs under selection have F_{ST} values clearly larger than zero). To estimate the maximal divergence between the populations, we therefore also calculated the F_{ST} of each individual SNP and extracted the 99th percentile (we chose the 99th percentile because this is more robust than the highest value).

Epigenetic variation: characterization of genome-wide DNA methylation levels

For each species, we filtered the epigenetic variation data for cytosines sequenced in at least three individuals per population (i.e., experimental group) with a total coverage between 5 and 200. To provide an overview of the genome-wide DNA methylation levels of the six species or each experimental group per species, we

visualized the DNA methylation levels of all cytosines averaged across all individuals with violin plots. We also visualized the average DNA methylation level within genes, transposons, repeats, and unclassified reference contigs with heat maps. Both were done either using all sequence contexts (CG, CHG, CHH) at once or separately for each sequence context.

Epigenetic variation: identification of differentially methylated cytosines (DMCs)

DNA methylation data was processed and filtered as described before. Variation in DNA methylation at each individual cytosine was then analyzed with a linear model in R with the package DSS (version 2.24.0; Park and Wu 2016), according to a design with a single factor comprising all different experimental groups (similar to the approach described for RNA-Seq, Schmid 2017 and the testing procedure in Schmid *et al.,* 2018). Specific groups were compared with linear contrasts and *P*-values for each contrast were adjusted for multiple testing to reflect false discovery rates (FDR). Taking *P. vulgaris* as an example, we compared the three plant histories across both assemblies and within each assembly to each other. Likewise, we compared the two assemblies across all plant histories and within each to each other. A cytosine was defined as differentially methylated (DMC) if the FDR was below 0.01 for any of the contrasts.

Results

Visualization of genetic distances between the plant individuals using 5% of the loci with the strongest divergence between the populations clearly separated the individuals according to their selection history in five out of six species (Fig. 2). In addition to the separation by the selection history, individuals from *L. pratensis* also clustered according to the current diversity level (i.e., assembly). However, considering that the plants were assigned randomly to the assembly treatment, we expected a genetic differentiation according to the selection history but not the assembly. We therefore tested for a significant genetic divergence between the

selection histories and the assemblies with the G-statistics test (Fig. 3, S3 and S4, Goudet *et al.,* 1996).

Genetic differentiation was significant in all data sets ($P <= 0.001$), with two exceptions. First, the plant histories of *P. lanceolata* did not exhibit any significant genetic differentiation (only the test with all plant histories within monoculture assemblies in the data set with all SNPs was significant). Second, the test including only the monoculture and mixture types within the mixture assemblies was not significant for *L. pratensis*. In contrast, the tests comparing the monoculture and mixture assemblies within each of the plant histories were never significant ($P > 0.01$).

To estimate the amount of genetic variation explained by the plant histories, we calculated average pairwise F_{ST} values (Table S4) and the 99th percentiles of the SNP-wise F_{ST} values (Table 1 and Tables S5, S6). Average pairwise F_{ST} values for the different plant histories were between 0.014 (naïve *vs.* monoculture type within the monoculture assemblies of *L. pratensis*) and 0.1 (naïve *vs.* monoculture or mixture type within monoculture assemblies of *P. vulgaris*). The 99th percentiles were markedly higher and between 0.13 (monoculture vs. mixture types within mixture assemblies of *G. mollugo*) and 0.67 (all plant histories within mixture assemblies of *P. vulgaris*). Thus, overall, 1.4% to 10% of the genetic variation was explained by plant histories. However, for individual SNPs, plant histories could explain up to 67% of the genetic variation.

To get an overview of the DNA methylation data, we visualized overall DNA methylation levels for each plant species, sequence context (CG, CHG, CHH), and genomic feature context (genes, transposons, repeats, and unclassified contigs, Fig. 4). For all species, DNA methylation was generally highest in the CG context (80.3%), lower in the CHG context (59.7%), and lowest in CHH context (13.2%). Differences between species were most pronounced in the CHG context in which *L. pratensis* (77.1%) and *P. lanceolata* (85.8%) exhibited markedly higher methylation levels than the other four species (56.6%, 52.3%, 43.6%, and 53.1% in *G. mollugo*, *O. viciifolia*, *P. vulgaris*, and *V. chamaedrys*, respectively). Within each species and context, DNA methylation was highest in transposons and lowest in genes (Fig. 4B). Overall, these patterns are within the range of what was reported previously for other angiosperms (Niederhuth *et al.,* 2016).

For an initial comparison between the experimental groups, we visualized the overall DNA methylation levels as we did about for the genetic variation, but for each experimental group separately (Fig. S4). Given that the overall methylation levels appeared to be similar, we tested for significant differences in DNA methylation levels at each individual cytosine. On average, 1.8% of all tested cytosines were significant in at least one of the tested contrasts (FDR < 0.01, "DMCs" for differentially methylated cytosines, Table 1 for all contexts and Tables S5, S6, and S7 for each context separately). Relative to the total number of cytosines tested, differences between plant histories (tested within or across both assemblies) were between 0.14% and 1.22% on average across all species and between 0.03% and 1.99% per individual species. Differences between the two assemblies (tested within or across all plant histories) were between 0.04% and 0.15% on average across all species and between 0.03% and 0.35% per individual species. Thus, the fraction of differentially methylated cytosines between the plant histories was generally larger than differences between the two assemblies.

Within the plant histories, differences between the monoculture types and the naïve plants were between 0.20% and 1.22% within species. Differences between mixture types and naïve plants were between 0.13% and 1.27% within species. Differences between monoculture and mixture types were between 0.03% and 1.99% within species. However, if compared within each species separately, there were always more DMCs in the comparisons between plants from Jena and the naïve plants than in the comparison between monoculture and mixture types. The sole exception was *O. viciifolia* for which there were more differences between the monoculture and mixture types than between these and the naïve plants.

To further characterize the differences in DNA methylation, we calculated the average change in DNA methylation at the DMCs for each contrast, across and within all sequence contexts (CG, CHG, and CHH) and feature types (genes, transposons, repeats, and unclassified) and visualized these differences (Fig. 5). We could not identify clear patterns between the different comparisons with one exception: differences in the comparisons between plants from Jena and the naïve plants within genes (all sequence contexts) were mostly biased towards a higher methylation in the naïve plants. Thus, plants in the Jena Experiment lost on average DNA methylation at DMCs within genes (*O. viciifolia* was the exception with changes close to zero).

Discussion

We aimed at measuring to what extent differential selection in monoculture and mixed species grassland communities had influenced the genetic and epigenetic make-up of six grassland species and whether these mechanisms could have mediated the rapid phenotypic responses that were previously observed in plants with different diversity backgrounds (see Zuppinger-Dingley *et al*., 2014; van Moorsel *et al.*, 2018b). Strong genetic divergence within species in response to selection history was found in five out of six species, suggesting that in these five species different genotypes may have been sorted out from standing genetic variation that was present at the start of the diversity experiment. Less likely, recombination and mutation could also have contributed.

For *P. lanceolata,* no significant differentiation was found between monoculture and the mixture types within mixture assemblies. This may be due to the panmictic nature of *P. lanceolata*. Movement of pollen and seeds between plots together with high germination and establishment rates in this species may have leveled out genetic differentiation between plots of different species diversity. Furthermore, coverage was low and only a small proportion of the genome was covered, resulting in only 202 SNPs in the dataset. We therefore cannot exclude the possibility that genetic variation was present but that our epiGBS screening missed the genomic regions under selection. For future research, we propose that increasing the sequencing coverage should be prioritized over adding more samples.

Only in *O. viciifolia* we found more genetic differences between the monoculture and mixture types than between these and naïve plants. This may be due to the fact that some subplots that were transplanted from the field site in Jena to the experimental plots contained only few individuals of *O. viciifolia*. The results indicate that *O. viciifolia* plant material originating from monoculture and mixture selection history could have originated from a singly ancestor, which resulted in a high degree of relatedness within each population. In contrast, the seeds sourced from the commercial supplier likely originated from many mother plants.

Overall, only 1.4% to 10% of genetic variation was explained by plant histories. However, for individual SNPs, plant histories could explain up to 67% of the genetic variation. This may indicate that only few loci were under selection (high

differentiation) and that most of the genome segregated randomly (low differentiation). Interestingly, this would suggest that the traits underlying the differentiation within species into monoculture and mixture types are not highly polygenic, thus making them accessible to breeding.

Our work supports the hypothesis that perennial grassland species in field experiments evolve as a response to their biotic environment (Zuppinger-Dingley *et al.*, 2014; van Moorsel *et al.*, 2018b; van Moorsel *et al.*, 2018c). Recent experiments showed stronger complementarity effects (Loreau & Hector 2001) in communities with a selection history in species mixtures as opposed to plant individuals with a selection history in monocultures (Zuppinger-Dingley *et al*., 2014; van Moorsel *et al*., 2018b). It has been under debate whether epigenetic or genetic mechanisms are the main driver of observed phenotypic variation in such scenarios (Rapp & Wendel 2005; Bird 2007; Bossdorf *et al.*, 2008; Tilman & Snell-Rood 2014). In our experiment, we found genetic and epigenetic variation. However, genetic variation can have a major effect on epigenetic variation not only in *cis* but also in *trans* (Dubin *et al.,* 2015; Kawakatsu *et al.,* 2016). However, *cis* associations can also evolve independently of underlying genetic variation, when spontaneous epimutations are inherited through epigenetic inheritance (Taudt *et al.,* 2016). Thus, given the observed and the additional unobserved (only parts of the genome were sequenced) genetic variation, it seems plausible that most of the observed epigenetic variation was caused by the underlying genetic variation. This is in line with quantitative genetics studies in *A. thaliana* in which DNA methylation differences were found to be mostly associated with underlying genetic variation (e.g., Li *et al.,* 2014).

We did, however, also observe methylation differences between the assemblies in absence of genetic differentiation, even though these differences were small when put into context. A recent study using the same statistical method but a more stringent significance threshold (FDR < 0.001 instead of FDR < 0.01) compared DNA methylation profiles of different tissue types of the liverwort *Marchantia polymorpha*, covering important stages of the life cycle (Schmid *et al.,* 2018). With less than 1% DMCs, the differences between the two assemblies of this study correspond to the difference between two individual plants in Schmid *et al.* (2018). However, around 42% of all cytosines varied significantly across the entire life cycle in that study. Thus, differences within individuals were much more pronounced than between individuals. Several other studies provide direct or indirect evidence for clear

differences in DNA methylation between different tissue, cell types or developmental stages (Jullien *et al.,* 2012; Calarco *et al.,* 2012; Ibarra *et al.,* 2012; Park *et al.,* 2016; Ingouff *et al.*, 2017; Bouyer *et al.,* 2017). Clearly, this does not exclude the possibility that epigenetics may play a role in adaptation (see Hauser *et al.,* 2011, Kronholm & Collins 2016 and Quadrana & Colot 2016 for examples of natural epialleles), but it makes it less likely that the strengthening of the biodiversity effect in the populations assessed in this study was due to epigenetic differences that were not a consequence of the underlying genetic differentiation.

Conclusions

We used novel genomic tools to study the importance of epigenetics for adaptation to novel biotic conditions in six non-model plant species differentially selected in a field biodiversity experiment. Future research should continue to integrate molecular tools with studies on non-model species in order to push the field forward (Richards *et al*., 2017; Heer *et al.*, 2018). Our findings suggest that selection on standing genetic variation is a powerful driver of evolution even in the absence of many generations of plant growth. In addition, we propose that community diversity had the selective power to differentiate plant populations within species into mixture and monoculture types in only a few generations. Molecular tools and the integration of evolutionary concepts into plant community ecology can open up new alleys of research, which should be exploited in order to understand the community evolutive processes (Shafer *et al.*, 2015) that lead to the plant community compositions and structures as we see them today.

Acknowledgements

We thank T. Zwimpfer, M. Furler, D. Trujillo and D. Topalovic for technical assistance. Thanks to T. Hahl for help with the glasshouse experiment. This study was supported by the Swiss National Science Foundation (grants number 147092 and 166457 to B. Schmid) and the University Research Priority Program Global Change and Biodiversity of the University of Zurich. S.J.V.M. was furthermore supported by a travel grant from the ESF Congenomics network (reference number 7167).

Author contributions

S.J.V.M, P.V. and B.S. planned and designed the study, S.J.V.M. carried out the pot experiment and collected plant material, C.A.M.W. performed the lab work and created the sequencing library and T.V.G. created the bioinformatics pipeline. M.W.S. analyzed the data and produced the figures. S.J.V.M and M.W.S. wrote the first draft of the manuscript. All authors contributed to revisions.

References

Anderson JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. Trends in Genetics 27:258–266.

Barrett R, Schluter D. 2008. Adaptation from standing genetic variation. Trends in Ecology & Evolution 23:38–44.

Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J. 2009. The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature 458:1018–1020.

Bird A. 2007. Perceptions of epigenetics. Nature 447:396–398.

Bossdorf O, Richards CL, Pigliucci M. 2008. Epigenetics for ecologists. Ecology letters 11:106–15.

Bouyer D, Kramdi A, Kassam M, Heese M, Schnittger A, Roudier F, Colot V. 2017. DNAmethylation dynamics during early plant life. Genome Biology 18:179.

Calarco JP, Borges F, Donoghue MTA, Van Ex F, Jullien PE, Lopes R, Gardner R, Berger F, Feijó JA, Becker JD. 2012. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell 151:194–205.

Cardinale BJ, Wright JP, Cadotte MW, Carroll IT, Hector A, Srivastava DS, Loreau M, Weis JJ. 2007. Impacts of plant diversity on biomass production increase through time because of species complementarity. Proceedings of the National Academy of Sciences 104:18123–18128.

Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury J-M, Wincker P, *et al.* 2014. Mapping the epigenetic basis of complex traits. Science 343:1145–1148.

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Paolo Casale F, Drewe P, Kahles A, Jean G, Vilhjálmsson B, *et al.*,2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. Elife 4:e05255.

Fakheran S, Paul-Victor C, Heichinger C, Schmid B, U. Grossniklaus U, Turnbull L. 2010. Adaptation and extinction in experimentally fragmented landscapes. Proceedings of the National Academy of Sciences 107:19120–19125.

Fargione J, Tilman D, Dybzinski R, Lambers JHR, Clark C, Harpole WS, Knops JM, Reich PB, Loreau M. 2007. From selection to complementarity: shifts in the causes of biodiversity-productivity relationships in a long-term biodiversity experiment. Proceedings of the Royal Society B: Biological Sciences 274:871–876.

Gervasi DDL, Schiestl FP. 2017. Real-time divergent evolution in plants driven by pollinators. Nature communication 8:14691.

Ghalambour CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology 21:394–407.

Goudet, J, Raymond, M, Meeüs, T, Rousset, F. 1996. Testing differentiation in diploid populations. Genetics 144:1933-1940.

Goudet J, Jombart T. 2015. hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.04-22. https://CRAN.R-project.org/package=hierfstat

van Gurp TP, Wagemaker NCAM, Wouters B, Vergeer P, Ouborg JNJ, Verhoeven KJF. 2016. epiGBS: reference-free reduced representation bisulfite sequencing. Nature Methods 13:322–324.

Hairston NG, Ellner SP, Geber MA, Yoshida T, Fox JA. 2005. Rapid evolution and the convergence of ecological and evolutionary time. Ecology Letters 8:1114–1127.

Hauser M-T, Aufsatz W, Jonak C, Luschnig C. 2011. Transgenerational epigenetic inheritance in plants. Biochim Biophys Acta 1809:459–468.

Heer K, Mounger J, Boquete MT, Richards CL, Opgenoorth L. 2018. The diversifying field of plant epigenetics. New Phytologist 217:988–992.

667 **Ibarra CA, Feng X, Schoft VK, Hsieh T-F, Usawa R, Rodrigues JA, Zemach A,**
668 **Chumak N, Machlicova A, Nishimura T, *et al.*,2012.** Active DNA
669 demethylation in plant companion cells reinforces transposon methylation in
670 gametes. Science 337:1360–1364.
671 **Ingouff M, Selles B, Michaud C, Vu TM, Berger F, Schorn AJ, Autran D, Van**
672 **Durme M, Nowack MK, Martienssen RA, *et al.*,2017.** Live-cell analysis of
673 DNA methylation during sexual reproduction in *Arabidopsis* reveals context
674 and sex-specific dynamics controlled by noncanonical RdDM. Genes &
675 Development 31:72–83.
676 **Isbell F, Craven D, Connolly J, Loreau M, Schmid B, Beierkuhnlein C, Bezemer**
677 **TM, Bonin C, Bruelheide H, de Luca E, *et al.*,2015.** Biodiversity increases
678 the resistance of ecosystem productivity to climate extremes. Nature 526:574–
679 577.
680 **Jombart T. 2008.** Adegenet: A R Package for the Multivariate Analysis of Genetic
681 Markers. Bioinformatics 24: 1403–5.
682 **Joshi J, Schmid B, Caldeira MC, Dimitrakopoulos PG, Good J, Harris R, Hector**
683 **A, Huss-Danell K, Jumpponen A, Minns A *et al.*,2001.** Local adaptation
684 enhances performance of common plant species. Ecology Letters 4:536–544.
685 **Jost L. 2008.** G_{ST} and its relatives do not measure differentiation. Molecular Ecology
686 17:4015–4026.
687 **Jullien PE, Susaki D, Yelagandula R, Higashiyama T, Berger F. 2012.** DNA
688 methylation dynamics during sexual reproduction in *Arabidopsis thaliana*.
689 Current Biology 22:1825–1830.
690 **Kawakatsu T, Huang SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R,**
691 **Nery JR, Barragan C, He Y, *et al.*,2016.** Epigenomic Diversity in a Global
692 Collection of Arabidopsis thaliana Accessions. Cell 166:492-505.
693 **Kleynhans EJ, Otto SP, Reich PB, Vellend M. 2016.** Adaptation to elevated CO2 in
694 different biodiversity contexts. Nature Communications 7:12358.
695 **Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V,**
696 **Vreugdenhil D, Keurentjes, JJB. 2015.** Epigenetic basis of morphological
697 variation and phenotypic plasticity in *Arabidopsis thaliana*. Plant Cell 27:
698 337–348.
699 **Kronholm I, Collins S. 2016.** Epigenetic mutations can both help and hinder
700 adaptive evolution. Molecular Ecology 25: 1856–1868.

Li Y, Cheng R, Spokas KA, PalmerAA, Borevitz JO. 2014. Genetic variation for life history sensitivity to seasonal warming in Arabidopsis thaliana. Genetics 196: 569–577.

Lipowsky A, Schmid B, Roscher C. 2011. Selection for monoculture and mixture genotypes in a biodiversity experiment. Basic and Applied Ecology 12:360–371.

Loreau M, Hector A. 2001. Partitioning selection and complementarity in biodiversity experiments. Nature 412:72–76.

de Meeûs T, Goudet J. 2007. "A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels." *Infection, Genetics and Evolution* 7: 731-735.

Meyer ST, Ebeling A, Eisenhauer N, Hertzog L, Hillebrand H, Milcu A, Pompe S, Abbas M, Bessler H, Buchmann N, *et al.*,2016. Effects of biodiversity strengthen over time as ecosystem functioning declines at low and increases at high biodiversity. Ecosphere 7:e01619.

van der Maaten L, Hinton G. 2008. Visualizing data using t-SNE. Journal of Machine Learning Research 9:2579–2605.

van der Maaten L. 2014. Accelerating t-SNE using tree-based algorithms. Journal of machine learning research 15:3221–3245.

van Moorsel SJ, Hahl T, Petchey OL, Ebeling A, Eisenhauer N, Schmid B, Wagg C. 2018a. Evolution increases ecosystem temporal stability and recovery from a flood in grassland communities. bioRxiv.

van Moorsel SJ, Schmid MW, Hahl T, Zuppinger-Dingley D, Schmid B. 2018b. Selection in response to community diversity alters plant performance in and functional traits. bioRxiv.

van Moorsel SJ, Hahl T, Wagg C, De Deyn GB, Flynn DFB, Zuppinger-Dingley D, Schmid B. 2018c. Community evolution increases plant productivity at low diversity. Ecology Letters 21:128–137.

Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, *et al.*,2016. Widespread natural variation of DNA methylation within angiosperms. Genome Biology 17:194.

Ouborg NJ, Vergeer P, Mix C. 2006. The rough edges of the conservation genetics paradigm for plants. Journal of Ecology 94:1233–1248.

734 **Park K, Kim MY, Vickers M, Park J-S, Hyun Y, Okamoto T, Zilberman D,**
735 **Fischer RL, Feng X, Choi Y, *et al.*,2016.** DNA demethylation is initiated in
736 the central cells of *Arabidopsis* and rice. PNAS 113:15138–15143.
737 **Park Y, Wu H. 2016.** Differential methylation analysis for BS-seq data under general
738 experimental design. Bioinformatics 32: 1446–1453.
739 **Price TD, Qvarnstrom A, Irwin DE. 2003.** The role of phenotypic plasticity in
740 driving genetic evolution. Proceedings of the Royal Society B: Biological
741 Sciences 270:1433–1440.
742 **Quadrana L, Colot V. 2016.** Plant transgenerational epigenetics. Annual Review of
743 Genetics 50:467–491.
744 **Rapp RA, Wendel JF. 2005.** Epigenetics and plant evolution. New Phytologist
745 168:81–91.
746 **Reich PB, Tilman D, Isbell F, Mueller K, Hobbie SE, Flynn DFB, Eisenhauer N.**
747 **2012.** Impacts of biodiversity loss escalate through time as redundancy fades.
748 Science 336:589–592.
749 **Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M,**
750 **Durka W, Engelhardt J, Gaspar B, Gogol-Döring A, *et al.*,2017.**
751 Ecological plant epigenetics: Evidence from model and non-model species,
752 and the way forward. Ecology Letters 20:1576–1590.
753 **Roscher C, Schumacher J, Baade J, Wilcke W, Gleixner G, Weisser WW,**
754 **Schmid B, Schulze E-D. 2004.** The role of biodiversity for element cycling
755 and trophic interactions: an experimental approach in a grassland community.
756 Basic and Applied Ecology 5:107–121.
757 **Rottstock T, Kummer V, Fischer M, Joshi J. 2017.** Rapid transgenerational effects
758 in *Knautia arvensis* in response to plant community diversity. Journal of
759 Ecology 105:714–725.
760 **Schmid B. 1985.** Clonal Growth in Grassland Perennials: III. Genetic Variation and
761 Plasticity Between and Within Populations of Bellis Perennis and Prunella
762 Vulgaris. Journal of Ecology 73:819–830.
763 **Schmid MW. 2017.** RNA-Seq data analysis protocol: Combining in-house and
764 publicly available data. In: Plant germline development: Methods and
765 protocols. Schmidt A, editor. New York, NY: Springer New York; 2017. pp.
766 309–35.

Schmid MW, Giraldo-Fonseca A, Rövekamp M, Smetanin D, Bowman JL, Grossniklaus U. 2018. Extensive epigenetic reprogramming during the life cycle of *Marchantia polymorpha*. Genome Biology 19:9.

Schoener TW. 2011. The newest synthesis: Understanding the interplay of evolutionary and ecological dynamics. Science 331:426–429.

Shafer ABA, Wolf JBW, Alves PC, Bergström L, Bruford MW, Brännström I, Colling G, Dalén L, De Meester L, Ekblom R. *et al.*,2015. Genomics and the challenging translation into conservation practice. Trends in Ecology & Evolution 30:78–87.

Slobodkin LB. 1961. Growth and regulation of animal populations. Holt, Rinehart and Winston, New York.

Smit AFA, Hubley R, Green P. 2013–2015. RepeatMasker Open-4.0. http://www.repeatmasker.org/.

Taudt A, Colomé-Tatché M, Johannes F. 2016. Genetic sources of population epigenomic variation. Nature Reviews Genetics 17: 319–332.

Tilman D, Snell-Rood EC. 2014. Ecology: Diversity breeds complementarity. Nature 515:44–45.

Verhoeven KJF, vonHoldt BM, Sork VL. 2016. Epigenetics in ecology and evolution: what we know and what we need to know. Molecular Ecology 25:1631–1638.

Wilschut RA, Oplaat C, Snoek LB, Kirschner J, Verhoeven KJF. 2016. Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage. Molecular Ecology 25:1759–1768.

Zuppinger-Dingley D, Schmid B, Petermann JS, Yadav V, De Deyn GB, Flynn DFB. 2014. Selection for niche differentiation in plant communities increases biodiversity effects. Nature 515:108–111.

Data Accessibility Statement

Data will be made publicly available on Zenodo (DOI 10.5281/zenodo.1167563) and SRA (accession ID SRP132258) at time of acceptance.

Supporting information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Adapter overview.

Fig. S2 Results from the G-statistic tests given all SNPs within genes.

Fig. S3 Results from the G-statistic tests given all SNPs within transposons.

Fig. S4 DNA methylation levels in % at individual cytosines across all or within each individual sequence context (CG, CHG, CHH) for each experimental group of each species used.

Table S1 Sample overview.

Table S2 Community diversity and composition of the plots the seeds originated from.

Table S3 Reference sequences generated in this study.

Table S4 Average pairwise F_{ST} values.

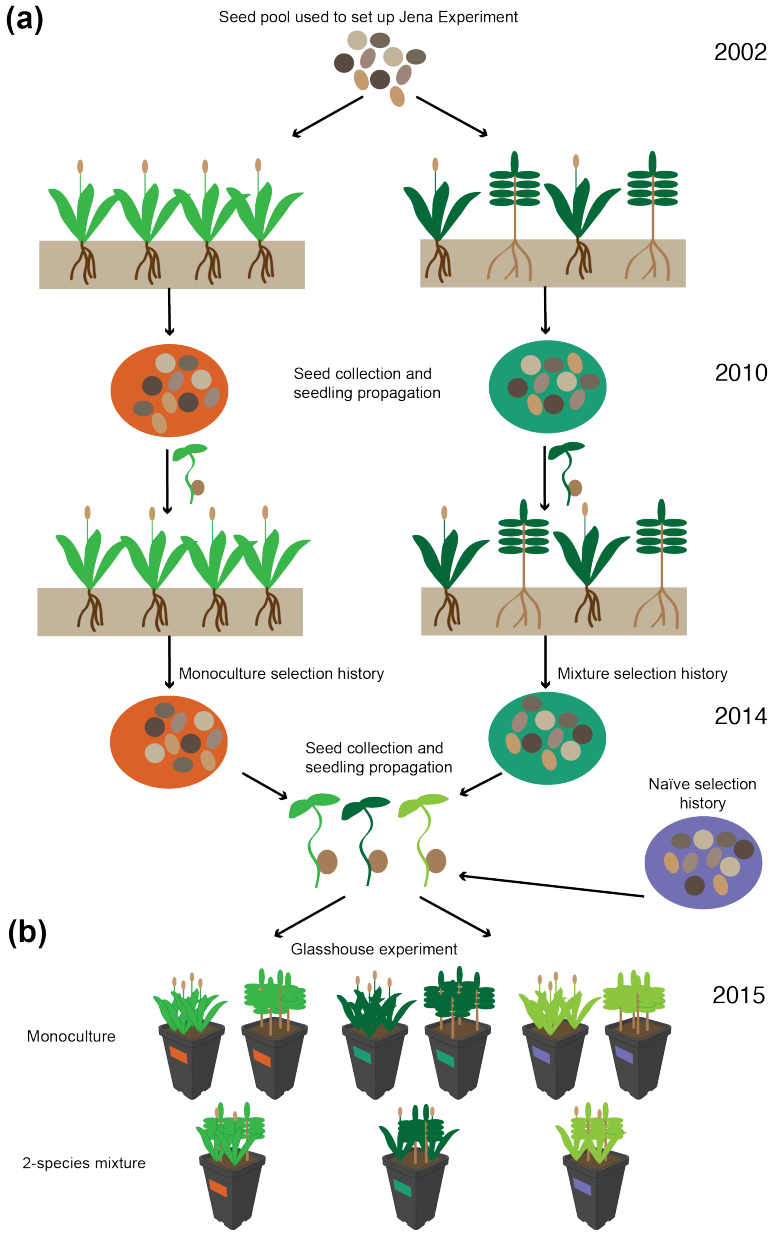
Table S5 99th percentiles of F_{ST} values in the SNPs within genes.

Table S6 99th percentiles of F_{ST} values in the SNPs within transposons.

Table S7 Number of cytosines with significant differences (FDR < 0.01) in DNA methylation between selection-history treatments and assemblies in the CG sequence context.

Table S8 Number of cytosines with significant differences (FDR < 0.01) in DNA methylation between selection-history treatments and assemblies in the CHG sequence context.

Table S9 Number of cytosines with significant differences (FDR < 0.01) in DNA methylation between selection-history treatments and assemblies in the CHH sequence context.

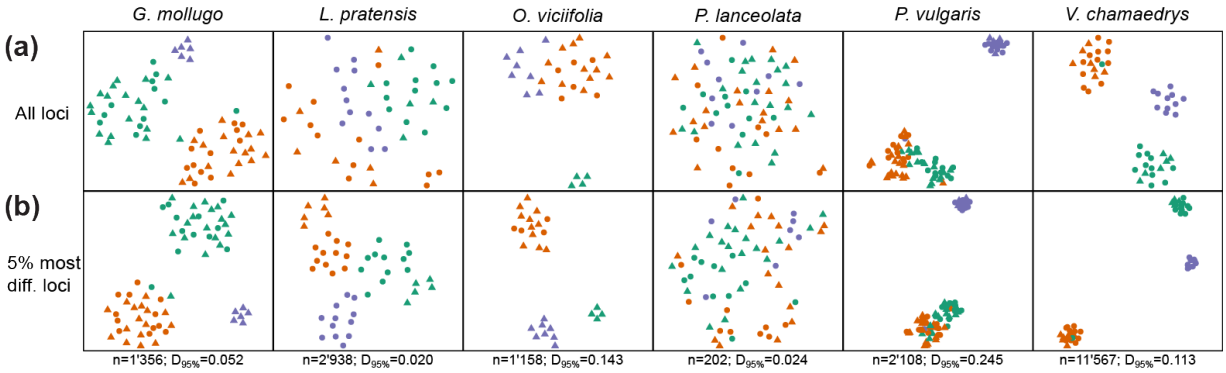


825

826 **Fig. 1** (a) The origin of seeds used for the glasshouse experiment and genetic analysis.
827 Seedlings were planted in mixtures and monocultures in Jena in the year 2002. Two
828 reproduction events occurred when seeds were collected and subsequently new
829 seedlings were produced and planted again in the same community composition. (b)
830 Schematic representation of the glasshouse experiment. Monoculture assemblies and
831 2-species mixture assemblies were planted with either plants with mixture selection
832 history (green), monoculture selection history (orange) or naïve plants originating
833 from a commercial seed supplier (blue). Figure modified after van Moorsel *et al.*
834 (2018b).

835

836



837

838 **Fig. 2** Genetic distance between individuals of the different populations for the six
839 species. (a) All loci included in the analyses. (b) Only the 5% most differentiated loci
840 between the experimental groups included. Green: selection history in mixture,
841 orange: selection history in monocultures, blue: naïve selection history. Triangles:
842 current assembly monoculture, circles: current assembly mixture. $D_{95\%}$ is the
843 threshold to choose the 5% most differentiated loci based on Jost's D (the 95%
844 percentile in the entire data set, Jost 2008).

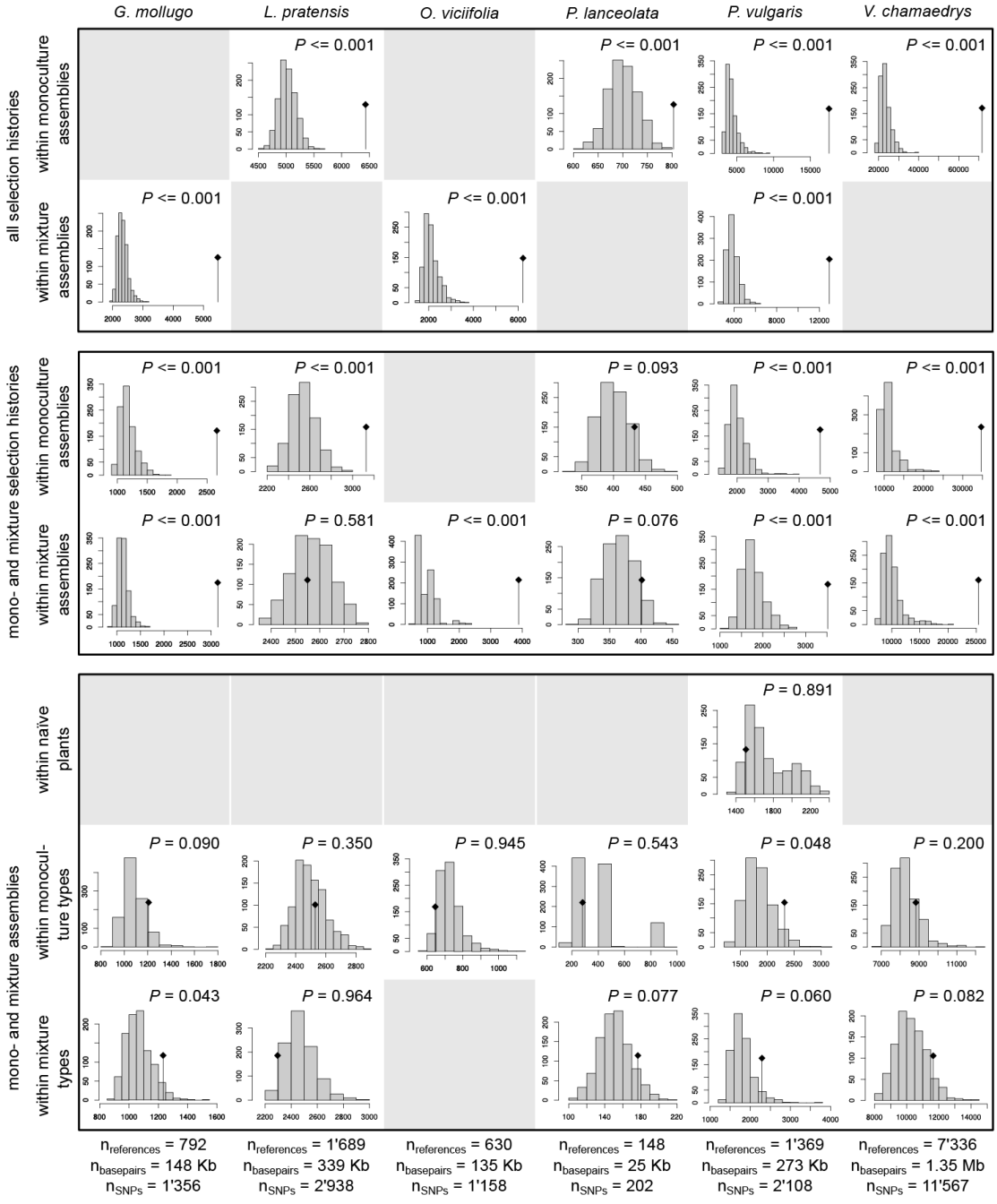


Fig. 3 Results from the G-statistic tests given all SNPs. Each panel shows a histogram of permuted test statistics (999 permutations) and indicates the observed statistics by a black dot and a segment. Test statistics are on the x-axis, frequencies on the y-axis. Grey boxes where data was not available (experimental group missing).

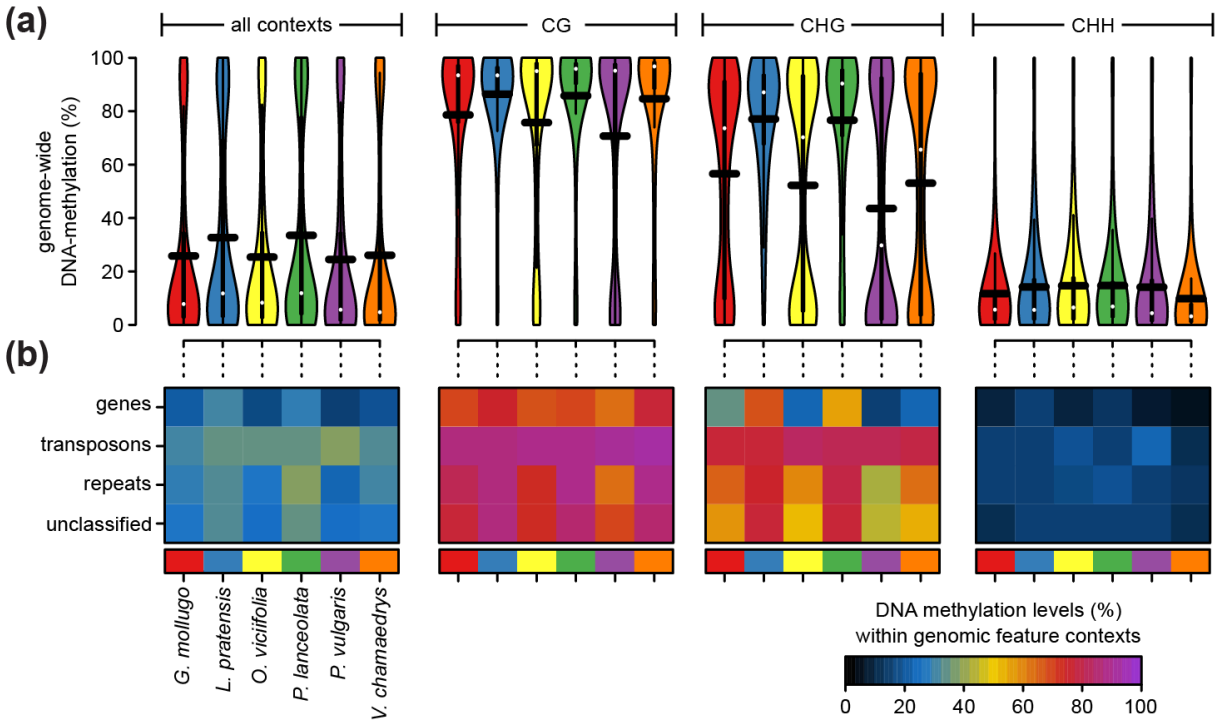


Fig. 4 (a) DNA methylation levels in percent at individual cytosines across all or within each individual sequence context (CG, CHG, CHH) for each species used in this study shown as violin plots. The horizontal black bars correspond to the means. (b) Average DNA methylation levels in percent for each sequence context, genomic feature, and species shown as a heat map.

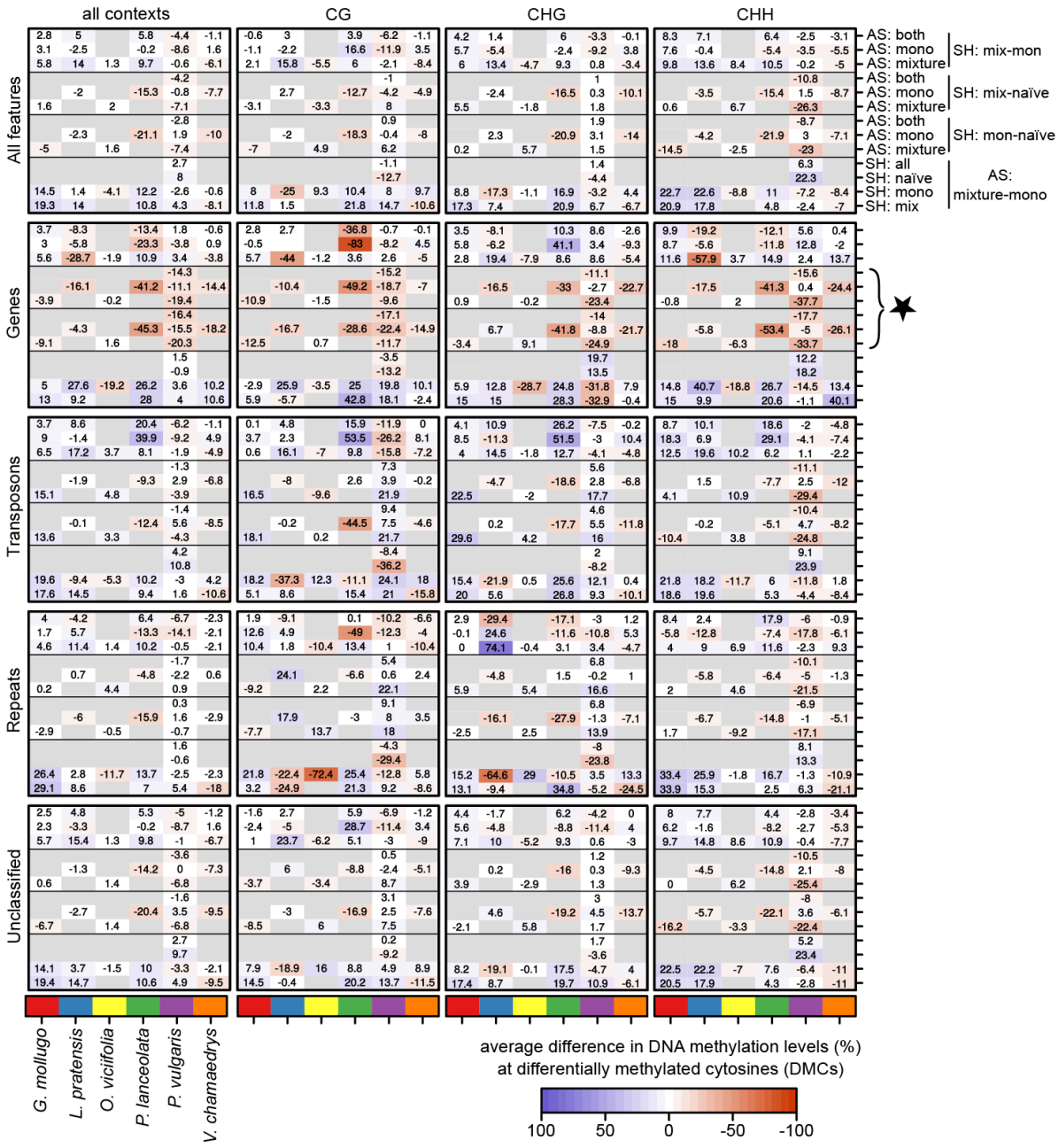


Fig. 5 Average differences in DNA methylation at significantly differentially methylated cytosines (DMCs; FDR < 0.01) within a given sequence (all, CG, CHG, and CHH) and feature (all, genes, transposons, repeats, unclassified) context are shown for all contrasts. The average differences are shown as colour gradient. The numbers within the heat map are the average differences. The asterisk marks the rows showing that plants in the Jena field lost on average DNA methylation at DMCs within genes compared to naïve plants.

Table 1 99th percentile of FST values in the data set with all SNPs.

	G. mollugo	*L. pratensis*	*O. viciifolia*	*P. lanceolata*	*P. vulgaris*	*V. chamaedrys*
SH in monoculture	NA	0.10641	NA	0.1413	0.6032	0.29453
SH in mixture	0.2354	NA	0.4	NA	0.67471	NA
SH contrast mono vs. mix in monoculture	0.1579	0.1	NA	0.09282	0.25	0.2571
SH contrast mono vs. mix in mixture	0.13332	0.1462	0.3333	0.05899	0.15592	0.3333
AS for naïve history	NA	NA	NA	NA	0.13182	NA
AS for mono history	0.0755	0.0909	0.08469	0.0762	0.10173	0.14305
AS for mix history	0.08355	0.0926	NA	0.09214	0.12	0.1586

AS = assembly, SH = selection history. For SNPs within genes or transposons see Tables S5 and S6.

868 **Table 2** Number of cytosines with significant differences (FDR < 0.01) in DNA
869 methylation between selection-history treatments and assemblies.

	G. mollugo	*L. pratensis*	*O. viciifolia*	*P. lanceolata*	*P. vulgaris*	*V. chamaedrys*	average %
SH: mixture vs. monoculture	6365 (0.55%)	239 (0.03%)	-	582 (0.11%)	6578 (0.28%)	12966 (0.60%)	0.31%
% in genes	13.68	4.6		4.47	12.82	11.51	
% in transposons	10.65	30.54		8.25	11.28	8.18	
% in repeats	4.41	4.18		8.93	6.67	4.63	
% in unclassified contigs	71.25	60.67		78.35	69.23	75.68	
>> within monoculture assembly	2332 (0.20%)	360 (0.05%)	-	216 (0.04%)	2507 (0.11%)	6507 (0.30%)	0.14%
% in genes	12.22	4.72		6.48	10.33	10.7	
% in transposons	11.11	28.61		7.87	11.61	8.16	
% in repeats	3.82	4.72		12.04	5.78	4.84	
% in unclassified contigs	72.86	61.94		73.61	72.28	76.3	
>> within mixture assembly	4617 (0.40%)	288 (0.04%)	10353 (1.99%)	2573 (0.47%)	8875 (0.37%)	5538 (0.25%)	0.59%
% in genes	11.63	4.17	12.13	3.26	11.14	9.73	
% in transposons	10.92	34.38	15.94	11.31	14.6	8.22	
% in repeats	4.7	5.21	6.93	7.5	6.49	4.66	
% in unclassified contigs	72.75	56.25	65.01	77.92	67.76	77.39	
SH: mixture vs. naïve	-	-	-	-	28970 (1.21%)	-	(1.21%)
% in genes					9.69		
% in transposons					11.84		
% in repeats					6.73		
% in unclassified contigs					71.73		
>> within monoculture assembly	-	893 (0.13%)	-	1040 (0.19%)	18439 (0.77%)	14692 (0.68%)	0.44%
% in genes		4.26		7.69	9.75	10.22	
% in transposons		26.32		9.71	12.55	7.83	
% in repeats		4.82		5.1	6.84	4.27	
% in unclassified contigs		64.61		77.5	70.87	77.68	
>> within mixture assembly	7012 (0.61%)	-	6574 (1.27%)	-	15019 (0.63%)	-	0.84%
% in genes	11.52		12.64		9.54		
% in transposons	10.95		16.06		11.85		
% in repeats	4.08		6.77		6.78		
% in unclassified contigs	73.45		64.53		71.83		
SH: monoculture vs. naïve	-	-	-	-	29052 (1.22%)	-	(1.22%)
% in genes					9.25		
% in transposons					12.52		
% in repeats					6.83		
% in unclassified contigs					71.4		
>> within monoculture assembly	-	1388 (0.20%)	-	620 (0.11%)	15877 (0.67%)	13614 (0.63%)	0.40%
% in genes		5.04		6.29	9.01	10.13	

% in transposons		28.6		7.1	12.36	7.28	
% in repeats		6.63		8.06	7.26	4.73	
% in unclassified contigs		59.73		78.55	71.37	77.86	
>> within mixture assembly	5336 (0.46%)	-	5464 (1.05%)	-	18419 (0.77%)	-	0.76%
% in genes	10.91		11.79		9.98		
% in transposons	9.07		15.98		12.24		
% in repeats	3.77		6.9		6.59		
% in unclassified contigs	76.26		65.34		71.19		
AS: mixture vs. monoculture	-	-	-	-	937 (0.04%)	-	(0.04%)
% in genes					8		
% in transposons					11.31		
% in repeats					8.86		
% in unclassified contigs					71.82		
>> within naïve selection history	-	-	-	-	2988 (0.13%)	-	(0.13%)
% in genes					9.64		
% in transposons					14.63		
% in repeats					7.83		
% in unclassified contigs					67.9		
>> within monoculture selection history	948 (0.08%)	191 (0.03%)	216 (0.04%)	506 (0.09%)	2264 (0.09%)	1588 (0.07%)	0.06%
% in genes	9.07	7.85	8.8	10.87	7.91	5.79	
% in transposons	10.34	31.94	18.52	6.92	14.31	12.22	
% in repeats	5.38	3.14	3.7	11.86	7.99	6.11	
% in unclassified contigs	75.21	57.07	68.98	70.36	69.79	75.88	
>> within mixture selection history	4032 (0.35%)	708 (0.10%)	-	944 (0.17%)	1857 (0.08%)	1624 (0.07%)	0.15%
% in genes	6.92	4.38		3.6	7.7	9.3	
% in transposons	12.7	29.66		11.55	16.05	8.37	
% in repeats	6.42	6.36		9	6.41	5.11	
% in unclassified contigs	73.96	59.6		75.85	69.84	77.22	
Total (percentage DMCs of tested cytosines)	21086 (1.8%)	3291 (0.5%)	16544 (3.2%)	5122 (0.9%)	65507 (2.7%)	36076 (1.7%)	1.80%

AS = assembly, SH= selection history. For data on separate sequence contexts see Tables S7 (CG), S8 (CHG), and S9 (CHH).