MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2016) **16**, 80–90 doi: 10.1111/1755-0998.12426

MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species

CONCHITA ALONSO,* RICARDO PÉREZ,† PILAR BAZAGA,* MÓNICA MEDRANO* and CARLOS M. HERRERA*

*Estación Biológica de Doñana, CSIC, Avenida Américo Vespucio s/n, Sevilla 41092, Spain, †Instituto de Investigaciones Químicas, Centro de Investigaciones Científicas Isla de La Cartuja, CSIC-US, Avenida Américo Vespucio s/n, Sevilla 41092, Spain

Abstract

Methylation of DNA cytosines affects whether transposons are silenced and genes are expressed, and is a major epigenetic mechanism whereby plants respond to environmental change. Analyses of methylation-sensitive amplification polymorphism (MS-AFLP or MSAP) have been often used to assess methyl-cytosine changes in response to stress treatments and, more recently, in ecological studies of wild plant populations. MSAP technique does not require a sequenced reference genome and provides many anonymous loci randomly distributed over the genome for which the methylation status can be ascertained. Scoring of MSAP data, however, is not straightforward, and efforts are still required to standardize this step to make use of the potential to distinguish between methylation at different nucleotide contexts. Furthermore, it is not known how accurately MSAP infers genome-wide cytosine methylation levels in plants. Here, we analyse the relationship between MSAP results and the percentage of global cytosine methylation in genomic DNA obtained by HPLC analysis. A screening of literature revealed that methylation of cytosines at cleavage sites assayed by MSAP was greater than genome-wide estimates obtained by HPLC, and percentages of methylation at different nucleotide contexts varied within and across species. Concurrent HPLC and MSAP analyses of DNA from 200 individuals of the perennial herb *Helleborus foetidus* confirmed that methyl-cytosine was more frequent in CCGG contexts than in the genome as a whole. In this species, global methylation was unrelated to methylation at the inner CG site. We suggest that global HPLC and context-specific MSAP methylation estimates provide complementary information whose combination can improve our current understanding of methylation-based epigenetic processes in nonmodel plants.

Keywords: DNA methylation, epigenetics, *Helleborus foetidus*, high-performance liquid chromatography, methyl-cytidine, methylation-sensitive amplification polymorphism, plant adaptation, population ecology, Ranunculaceae, stress

Received 17 October 2014; revision received 21 April 2015; accepted 22 April 2015

Introduction

The DNA sequence does not carry all the information required to determine the phenotype of an organism. Epigenetic regulation involves a variety of reversible chemical modifications that occur on the DNA and on its interacting proteins, and impinge on chromatin structure. As a result, epigenetic mechanisms may largely affect an organism phenotype without altering its DNA sequence (Grant-Downton & Dickinson 2005; Jablonka & Raz 2009; Hirsch *et al*. 2013). In plants, the covalent methylation at carbon 5′ position of cytosine residues of DNA is an important epigenetic mechanism. This contributes to the control of genomic integrity, regulation of gene expression and cell differentiation, individual development and growth, and plant responses to biotic and abiotic stresses (Finnegan *et al*. 1998; Grant-Downton & Dickinson 2006; Chinnusamy & Zhu 2009; Finnegan 2010). Understanding the role of DNA methylation in plant adaptation and evolution requires further study of patterns of methylation variability across multiple species and wild populations (Bossdorf *et al*. 2008; Flatscher *et al*. 2012; Grativol *et al*. 2012; Hirsch *et al*. 2013; Diez *et al*. 2014).

Currently, several techniques discriminate between methylated and nonmethylated cytosines, allowing to quantify or locate them in extracted DNAs, and even mapping in situ their position along chromosome

Correspondence: Conchita Alonso, Fax: +34 954621125;
E-mail: conalo@ebd.csic.es

structures (Fraga & Esteller 2002; Finnegan 2010; Laird 2010). For quantification of genome-wide cytosine methylation, high-performance liquid chromatography (HPLC) is recommended because of its global assessment, accuracy and reproducibility (Fraga & Esteller 2002; Lisanti *et al.* 2013). Selection of specific protocols will depend on available expertise, equipment and amount of sample (Johnston *et al.* 2005; Lopez-Torres *et al.* 2011). Such analyses do not differentiate between coding and noncoding sequences of DNA and are unable to detect subtle differences in the methylation status of individual genes. Notwithstanding, the HPLC technique is successfully employed in human cancer analysis (Lisanti *et al.* 2013) and has been recently used to infer macroevolutionary patterns in vertebrates (Varriale 2014) and plants (Alonso *et al.* 2015). As regards analyses focusing on the genomic location of cytosine methylation, they generally consist of two distinct steps (Laird 2010). In the first step, DNA samples are treated to discriminate methylated sites by digestion with methylation-sensitive restriction enzymes, affinity purification or bisulphite conversion. The second step involves different analytical equipment and procedures aimed at identifying fragment size polymorphisms or locus-specific methylation changes (Finnegan 2010; Laird 2010; Schrey *et al.* 2013). Finally, simultaneous quantification and location of cytosine methylation are nowadays possible by high-throughput sequencing methods that combine bisulphite conversion and deep sequencing. These methods, however, still require a high-quality reference sequence for whole-genome scanning (BC-seq and WGSBS), which generally precludes their application to nonmodel plants (Laird 2010; Schrey *et al.* 2013; Kim *et al.* 2014).

The methylation-sensitive amplification polymorphism (MS-AFLP or MSAP) technique does not require a sequenced reference genome and is useful to assess cytosine methylation state in a large number of anonymous loci randomly distributed over the genome. For this reason, the method remains as the most widely used tool in the analysis of changes in cytosine methylation in ecological epigenetics (Schrey *et al.* 2013). Within this subdiscipline, MSAP has been used to assess DNA methylation variability and epigenetic structure of wild plant populations, as well as to detect correlates between ecological conditions, phenotypic traits and methylation status at certain loci (e.g. Paun *et al.* 2010; Verhoeven *et al.* 2010; Herrera & Bazaga 2011; Schulz *et al.* 2013, 2014; Medrano *et al.* 2014). Further, MSAP results have been also used to confirm hypomethylation of plant mutants deficient in specific methylation enzymes (e.g. Kakutani *et al.* 1999; Papa *et al.* 2001; Rozhon *et al.* 2008) or after application of demethylation agents (Akimoto *et al.* 2007; Amoah *et al.* 2012). They can also be useful to relate specific methylation responses to function, by cloning and sequencing differentially methylated fragments extracted from acrylamide gels, and detecting putative location through BLAST homology searches (e.g. Greco *et al.* 2012; Cicatelli *et al.* 2014). Despite its broad use, however, there is not a unique way of running MSAP and interpreting the results obtained (reviewed by Schulz *et al.* 2013; Fulneček & Kovařík 2014), and the method's potential for distinguishing between methylation at different nucleotide contexts has been rarely exploited (but see Schulz *et al.* 2014). In addition, no attempt has been made to investigate the accuracy of different MSAP scoring methods to infer the global percentage of cytosine methylation in plants. Motivation for addressing these two issues is outlined in the next section.

MSAP analyses and global cytosine methylation

MSAP is a modification of the AFLP method in which two or more endonuclease isoschizomers that recognize the same restriction site but show differential sensitivity to DNA methylation are employed in parallel as 'methyl-sensitive cutter' in combination with the same 'indifferent cutter', and their respective band patterns compared (Schulz *et al.* 2013). For instance, the most widely used combination in ecological epigenetics involves the methylation-sensitive *Hpa*II and *Msp*I pair. These two enzymes recognize the same motif (5′-CCGG-3′), cleaving in the two enzymes is blocked when both cytosines are methylated, *Hpa*II cuts when only the external cytosine is hemi- (single strand) methylated, and *Msp*I cuts when only the internal cytosine is hemi- or fully (double strand) methylated (Fraga & Esteller 2002; Schulz *et al.* 2013). No consensus exists on the interpretation and scoring of the four possible outcomes obtainable from the combined MSAP banding patterns (11, 10, 01 and 00, where 1 denotes the presence of fragment; Schulz *et al.* 2013; Fulneček & Kovařík 2014), and it is frequent to find conflicting interpretations (C. Alonso, unpublished data). Although Schulz *et al.* (2013) found that different scoring strategies of MSAP products provided similar patterns of epigenetic diversity and differentiation of wild plant populations, separate consideration of different fragment types should provide more detailed insights into different epigenetic processes. To explore this possibility, we searched the literature for studies where MSAP was used to infer methylation changes after stress treatments and explored whether there was some consistent relationship between methylation levels at different nucleotide contexts.

Quantitative interpretations of MSAP analyses will depend on the distribution of potential cleavage sites across the genome and the relative frequency of CG, CHG and CHH (where H = A, C, T) nucleotide contexts (Fraga & Esteller 2002), both of which vary greatly

among species (Matassi *et al.* 1992; Feng *et al.* 2010) and genomic regions of the same species (Messeguer *et al.* 1991, and references therein). In addition, the *Hpa*II and *Msp*I combination will not detect cytosine methylation in other nucleotide contexts, such as CHH, CTG and CAG, or at fully methylated external CCGG sites (Kato *et al.* 2003; Schmitz *et al.* 2011). Thus, although relative hyper/hypomethylation can be derived in a preplanned comparison (see, e.g., Lira-Medeiros *et al.* 2010; Cicatelli *et al.* 2014), it would also be worth investigating the accuracy of different MSAP scorings to infer genome-wide DNA cytosine methylation levels, that is the proportion of total cytosines that are methylated irrespective of their specific sequence context. To fill this gap, we searched the literature for studies in which concurrent estimates of cytosine methylation obtained by enzymatic and chromatographic methods were available for the same samples and compared their results.

To gain a better understanding of the relationship between different MSAP scorings and genome-wide cytosine methylation levels, a study was performed on leaf DNA samples from wild-growing individuals of the herbaceous perennial *Helleborus foetidus*, for which we assessed global cytosine methylation in genomic DNA by means of HPLC analysis. These data allowed to explore the relationship between global HPLC quantification and estimates of cytosine methylation at the specific (5′-CCGG-3′) cleavage sites obtained using MSAP. We specifically tested for the possible influence on this relationship of two different scoring issues. The first one involves treating the simultaneous absence of the fragment after digestion with either of the two methyl-sensitive enzymes (i.e. the 00 result) as informative, missing or null data. The second one is related to the decision of using MSAP scorings of combined internal methylated plus hemimethylated external cytosines (i.e. 01 + 10), just internal methylated cytosines (i.e. 01) or fully unmethylated loci (i.e. 11 result). Rather than offering a criticism on the MSAP method, our purpose is to contribute to its expansion and improvement by highlighting the value of combining MSAP and global methylation data to gain better insights on the role of cytosine methylation in the ecology and evolution of natural plant populations.

Methods

Literature survey

We screened the literature for studies providing simultaneous estimates of genome-wide cytosine methylation for angiosperm species obtained by the application of quantitative chromatographic methods based on HPLC techniques and some variant of methylation-sensitive restriction endonucleases protocol, without attempting to cover exhaustively the extensive literature available. We also used ISI Web of Science (www.webofknowledge.com) to search for articles published in English between 1990 and 2014 (last accessed 17th July 2014) using the following keywords combination [(plant methyl* cytosine) AND (MSAP) AND (stress OR azacyt* OR zebularine) NOT (callus OR 'in vitro' OR calli)]. This excluded studies conducted with cell cultures, which were out of our scope.

Study species and field sampling

We used *Helleborus foetidus* L. (Ranunculaceae), a perennial evergreen understory herb widely distributed in western Europe, as our case study. Recent investigations on this species revealed extensive natural variability in methylation status of cytosines at 5′-CCGG sites across individual plants that are more variable than the traditional AFLP markers in this species, spatially differentiated among populations and inherited transgenerationally to a considerable extent (Herrera *et al.* 2013, 2014; Medrano *et al.* 2014). In addition, both global cytosine methylation and specific methylation status at certain MSAP loci have been associated with individual functional traits in this species (Alonso *et al.* 2014; Herrera *et al.* 2014; Medrano *et al.* 2014). These characteristics make *H. foetidus* a suitable system to test for the relationship between global and specific cytosine methylation.

The study was conducted in the Sierra de Cazorla, a mountainous area in Jaén province, southeastern Spain where the species is widely distributed across elevations and environments. In the spring of 2012, 20 inflorescence-bearing plants were selected for study at each of 10 sites (N = 200 plants) chosen to encompass the entire ecological range of the species in the region (see Medrano *et al.* 2014 for field sampling details). A sample of young leaves was collected from each plant at each population's flowering peak. Leaves were dried at ambient temperature in silica gel and subsequently homogenized to a fine powder using a Retsch MM 200 mill. Total genomic DNA was extracted from dry leaf samples using Qiagen DNeasy Plant Mini Kit. Two aliquots from the same DNA extract were used for HPLC and MSAP analyses.

MSAP method and scoring

MSAP analyses of *H. foetidus* samples used four *Mse*I + 3/*Hpa*II–*Msp*I + 2 primer combinations (Table S1, Supporting information). We used *Mse*I instead of the most commonly used *Eco*RI mainly because it provided better repeatability in our study species (Medrano *et al.*

2014). *Mse*I recognizes (5′-TTAA-3′) and, because of its shorter recognition sequence relative to *Eco*RI, it tends to cut more frequently, which reduces the probability of having blind internal CCGG targeted sites (Fulneček & Kovařík 2014), and also importantly, cleavage site does not contain any C residue. Additional details of laboratory methods can be found in Medrano *et al.* (2014). Fragment separation and detection were made using an ABI PRISM 3130xl DNA sequencer (Applied Biosystems, Foster City, CA, USA). Only fragments >150 bp in size were considered to reduce the potential impact of size homoplasy. We obtained 287 loci in the size range 150–500 bp in the methylation raw data matrix of 200 individuals with four primer combinations. Presence–absence of *Mse*I–*Hpa*II and *Mse*I–*Msp*I fragments in each individual plant was scored automatically with GENEMAPPER 3.7 software (Applied Biosystems) with the sum of signal normalization method and by fixing a common absolute peak height threshold (sensu Whitlock *et al.* 2008). Cytosine methylation was estimated based on just the 'methylation-susceptible' loci, that is those that did not exhibited the same status in >95% of total study samples, as performed by Pérez-Figueroa (2013) and commonly applied in ecological epigenetic studies (Schulz *et al.* 2013). The number of 'methylation-susceptible' loci obtained was 241 for a fixed peak height threshold of 50 relative fluorescence units (rfu), which was found to be the best threshold for our data (see Appendix S1, Supporting information).

Special care was taken to ensure reproducibility of MSAP analyses. To check the consistency of individual results, 17 samples (8.5%) from different sites were replicated for all primer combinations, and genotyping error rates were computed separately for each fragment and estimated as the ratio of the number of discordances to the number of samples scored twice (Table S1, Supporting information). Before purging, the mean error rate across loci was 10.7% and the median was 8.8% ($N = 287$). A conservative purging was adopted. Only the fragments with error rates equal or lower than the median of the error distribution for the whole set of fragments were retained ($N = 155$), and mean genotyping error rates were then determined separately for each primer combination (see Medrano *et al.* 2014, for details). The retained loci exhibited an average genotyping error rate of 3.7%, and 125 of 155 loci were methylation susceptible (Table S1, Supporting information).

Analyses of MSAP results were based on element-wise comparisons of fragment presence–absence matrices for individual plants obtained with *Mse*I–*Hpa*II and *Mse*I–*Msp*I primer combinations. For every individual and particular fragment, it was first determined whether the fragment was (I) present in both *Mse*I–*Hpa*II and *Mse*I–*Msp*I products; (II) present only in *Mse*I–*Msp*I products; (III) present only in *Mse*I–*Hpa*II products; and (IV) absent from both *Mse*I–*Hpa*II and *Mse*I–*Msp*I products. Condition (I) denotes a nonmethylated state, condition (II) corresponds to full- or hemimethylated internal cytosine, condition (III) corresponds to hemimethylation of external cytosine, and condition (IV) is uninformative, as it could be caused by either restriction target absence or hypermethylation (Schulz *et al.* 2013; but see also Fulneček & Kovařík 2014). Summing up the number of loci per category, we obtained the relative proportions of loci at each four conditions per individual. The process was accomplished twice using MSAP package for R (Pérez-Figueroa 2013; downloadable from http://cran.r-project.org/) applied to all loci and filtered loci, respectively.

HPLC method

A 100 ng aliquot of DNA extract was digested with 3 U of DNA Degradase Plus™ (Zymo 71 Research, Irvine, CA, USA), a nuclease mix that degrades DNA to its individual nucleoside components. Digestion was carried out in a 40 μL volume at 37 °C for 3 h and terminated by heat inactivation at 70 °C for 20 min. Two independent replicates of DNA hydrolysate were prepared for each sample; all samples (200 plants × 2 replicates) were stored at −20 °C until analysis and processed in randomized order.

DNA cytosine methylation was determined for each sample by reversed-phase HPLC with a spectrofluorimetric detection technique modified after Lopez-Torres *et al.* (2011). This technique was selected because high selectivity and sensitivity of spectrofluorimetric detection reduces the detection levels and allows accurate quantification in small amounts of DNA (Lopez-Torres *et al.* 2011). Selective derivatization of cytosine moieties with 2-bromoacetophenone was conducted under anhydrous conditions, the 2-bromoacetophenone solution (0.5 M in DMF anhydrous) was discarded every 4–5 days, and samples were derivatized just before running the analyses to improve repeatability. HPLC quantification was accomplished in a Waters equipment (Waters 2695 Separations Module, Waters 2475 FDL) with a SunFire C18 column controlled by EMPOWER™ software (Waters Corporation, Milford, MA, USA). Fluorimetric detection was carried out at excitation/emission wavelengths of 306/378 nm, respectively. Equipment and detector were stabilized for >3 h, column temperature was maintained at 30 °C, and each derivatized sample was automatically diluted (1:1) with water immediately prior to injection.

The chromatographic separation was achieved within 30 min at a fixed flow rate of 0.59 mL/min. We applied a 5-min delay between injections. Four mobile phases

were used [A: water, B: acetonitrile, C: TFA 0.4% m/v and D: methanol], with the following gradient programme: 0–25 min 49% A, 10% B, 13% C, 28% D; 25–26 min 12% A, 15% B, 13% C, 60% D; 26–30 min 49% A, 10% B, 13% C, 28% D. Double-distilled water and HPLC-quality solvents were used for the analyses. The position of the peak corresponding to each nucleoside was determined using commercially available standards (Sigma-Aldrich), which varied <7 s across dates. The two standards elute under isocratic conditions at 11.69 (±0.006) min and 14.65 (±0.006) min for 2′-deoxycytidine (dC) and 5-methyl-2′-deoxycytidine (5mdC), respectively. The method allows also the distinction of 5-hydroxy-methyl-2′-deoxycytidine, which should appear in an intermediate position, but was not detected in our samples. The response measured was linear for the 5mdC and dC in concentrations from <10 nM up to 1000 nM. The method was particularly suitable to estimate the relative proportion of the two nucleosides in real samples differing in the amount of DNA because the relationship between relative concentration and relative signal was linear in the range of 1.2% to 48.55% of 5mdC, regardless of absolute concentration of the two nucleosides (see also Lopez-Torres *et al.* 2011).

Data analyses

Percentage of total cytosine methylation by HPLC was estimated for each sample as $100 \times 5\,\text{mdC}/(5\,\text{mdC} + \text{dC})$, where 5mdC and dC are the integrated areas under the peaks for 5-methyl-2′-deoxycytidine and 2′-deoxycytidine, respectively. Differences among populations and plants nested within populations were analysed by generalized linear models with REML estimation (Procedure Mixed in SAS 9.2; SAS Institute 2008). For the analyses of correlation, each individual was characterized by the average of the two replicates, except for two cases where single replicates were available. Spearman rank correlations were used to account for nonnormality of data.

Obtaining the equivalent individual methylation percentage from MSAP is not so straightforward (see Schulz *et al.* 2013 for a thorough discussion). In addition to the band scoring errors typical from AFLP (Bonin *et al.* 2004; see Appendix S1, Supporting information for analyses), the MSAP results will also depend on (i) the interpretation of condition IV loci, as fully methylated (e.g. Marconi *et al.* 2013), null data (i.e. treated as 0 and keeping the total number of loci used per sample constant; Schulz *et al.* 2013) or uninformative data (i.e. treated as missing, which leads to a variable number of scored loci per sample; Herrera & Bazaga 2010); and (ii) whether internal and external cytosine methylation states (conditions II and III) are combined to obtain a unique methylation scoring (e.g. Herrera & Bazaga 2010) or analysed separately (Schulz *et al.* 2014). Here, we will evaluate the relationships between the global HPLC quantification and three MSAP estimates, namely the summed methylation scoring (MS = conditions II + III), internal cytosine methylation (ICM = condition II) and the specific non-methylation (NM = condition I) percentages obtained treating ambiguous condition IV loci either as absence of the two fragments (0) or missing values.

Results

Literature survey

Cytosine methylation estimates based on enzymatic and chromatographic methods were gathered for 16 species and consistently confirmed DNA cytosine methylation in these plants (Table 1). HPLC estimates of percentage of total cytosines in DNA that were methylated varied between 4.6% and 40% (Table 1). At intraspecific level, qualitative estimates of cytosine methylation by enzymatic methods largely agreed with HPLC quantitative global estimates. However, in the few cases where HPLC and quantitative reports of MSAP were simultaneously available, including this study (see below), methylation at the specific restriction sites probed was always considerably higher than indicated by the global HPLC estimate (Table 1). Depending on species, methylation varied widely among the cytosine contexts assayed (Table 1). In particular, *Helleborus foetidus* and *Lycopersicon esculentum* showed higher methylation at CHG positions, *Gossypium hirsutum*, *Gossypium barbadense* and *Brassica oleracea* at CG positions, and *Arabidopsis thaliana* exhibited similar methylation percentages in the two nucleotide contexts considered.

Regarding plant methylation responses to environmental stress, we gathered 49 study cases involving 18 species and different stress treatments, including heavy metals pollution, osmotic regulation, temperature, soil nutrient and light availability, and specific demethylating agents (Table S2, Supporting information). MSAP confirmed changes in the methylation status of specific cytosine loci (i.e. variability in the presence/absence of specific bands) in most cases. In contrast, in 22% of instances, results did not support the hypothesis that stress elicits changes in global methylation of DNA cytosines, and the sign of overall methylation changes varied among studies reporting a global effect (Table 2). Some caveats must be explicitly recognized in relation to the heterogeneity of results (Table S2, Supporting information). The studies reviewed varied widely regarding (i) band scoring technology (gels vs. capillary sequencers); (ii) number of methylation-susceptible loci obtained; and (iii) interpretation of the absence/absence (type IV)

Table 1 Literature survey results for comparison of chemical and enzymatic methods suitable to discriminate between C and mC

		Methylation estimate (%mC)					
			Endonucleases				
Sp	Comparison	HPLC global	CHG	CG	Others	Global	References
Helleborus foetidus	10 populations	29.4*	38.4	16.8		55.1	This study
Lycopersicon esculentum	None	23–25	85	55		21	Messeguer *et al*. (1991)
Gossypium hirsutum	8 genotypes	24.8 ns	5.2 *	37.8 ns	6.7 ns	49.7 *	Osabe *et al*. (2014)
Gossypium barbadense	2 genotypes	24.2	7.7	37.2	7.5	51.9	Osabe *et al*. (2014)
Brassica oleracea	30 phenotypes	16†	17–27	30–41		52–67	Salmon *et al*. (2008)
Arabidopsis thaliana	2 phenotypes	4.6–7.5†	18.9	21			Salmon *et al*. (2008)
A. thaliana	None	4.6		Ambiguous			Leutwiler *et al*. (1984)
Pennisetum purpureum	None	38–40	Consistent	Consistent			Morrish & Vasil (1989)
Stellaria longipes (alpine)	Environment	18.1–12.3 *		Consistent	Inverse	50	Cai & Chinnappa (1999)
S. longipes (prairie)	Environment	19.1–18.0 ns		Consistent	Inverse	66.3	Cai & Chinnappa (1999)
Zea mays (inbreds)	Environment	27.5–29.2 *		Consistent			Tsaftaris & Polidoros (2000)
Z. mays (hybrids)	Environment	26.1–26.4		Consistent			Tsaftaris & Polidoros (2000)
Medicago sativa	Environment	19.2–21.7		Consistent			Rozhon *et al*. (2008)
A. thaliana	Wild vs. ddm1–5 mutant	6–3.8		Consistent			Rozhon *et al*. (2008)
A. thaliana	Wild vs. ddm1 mutant	6.6–2.2		Consistent			Kakutani *et al*. (1999)
Z. mays	Wild vs. Zmet2 mutant	24.8–21.7	<30–50	0			Papa *et al*. (2001)
Millets	6 species	14.5–30.9 *	Ambiguous	Ambiguous			Kumar *et al*. (1990)

Estimates obtained by chromatographic quantification (HPLC) of individual nucleosides are typically expressed as percentage of global cytosines that are methylated [%mC = 100 * mC/(C + mC)]. Analyses of fragment size polymorphisms obtained by endonucleases differing in mC sensitivity (MSAP) were more variable in their way of reporting results. Qualitative estimates were categorized as consistent, inverse or ambiguous when banding pattern was clearly consistent, inconsistent or doubts were expressed by the authors, respectively. Quantitative estimates are presented as average %mC for the specific nucleotide contexts revealed by different enzymes (i.e. mCG by *Hpa*II, mCHG by *Msp*I or *Bst*NI) and the global %mC estimation. In studies comparing groups of samples, we distinguished between types of comparisons regarding the use of different phenotypes (e.g. populations, morphotypes), genotypes, mutants or plants growing under different experimental environmental conditions. Significant heterogeneity among analysed samples is marked as significant (*) or nonsignificant (ns) when reported in the original publication.
†HPLC and endonuclease methods were not applied to exactly the same DNA material.

output. Even with these caveats, a reliable result seems to be that methylation response to stress (increase or decrease) varied across loci and sequence contexts. Its global sign and magnitude changed not only with the specific stress and plant species, but also across subspecies, lines or genotypes, and even between tissues of the same individuals (Table S2, Supporting information).

Helleborus foetidus: *global cytosine methylation level by HPLC*

Mean (±SE) percentage of cytosine methylation in young leaves of *H. foetidus* was 29.4% (±0.1), and ranged between 25.5% and 36.7% in our $N = 200$ samples. Average population figures ranged between 28.8 (±0.1) and 30.9 (±0.2). Methylation estimates differed significantly among populations ($F_{9,199} = 29.69$, $P < 0.0001$) and individuals ($F_{191,199} = 4.90$, $P < 0.0001$). Variation across individual plants accounted for 75.5% of explained variation.

Helleborus foetidus: *MSAP analyses and comparisons with HPLC*

Interpretation of the condition IV (absence from both *Mse*I–*Hpa*II and *Mse*I–*Msp*I products) as informative, missing or null data modified the relationship between HPLC estimates and the quantitative interpretations of MSAP results.

Table 2 Summary of literature survey on methylation changes associated with five main groups of abiotic stress factors obtained by application of the MSAP technique to different plant species and tissues

Stress factor	Without change *	With changes across loci		
		Equal methylation†	Net increase	Net decrease
Heavy metals addition	2	0	3	4
Osmotic stress (water, salt, alkali, propilenglycol)	7	5	8	6
Temperature	2	1	1	1
Nutrients and light availability	0	4	1	1
Specific mutagens	0	1	0	2

We show the number of study cases on each category. See Table S2 (Supporting information) for full data set and further details.
*None or very few loci (<10%) changed status in response to stress treatment.
†A similar number of loci gained and lost methylation after treatment.

Table 3 Spearman rank correlation coefficients (*P*-value in parentheses) for the relationship between percentages of cytosine methylation obtained for individual *Helleborus foetidus* plants by HPLC and different MSAP scoring methods [methylation scoring (MS); internal cytosine methylation (ICM); nonmethylation (NM)], calculated before (all methylation-susceptible loci, $N = 241$ loci) and after filtering MSAP data by genotyping error of replicated samples ($N = 125$ loci)

MSAP scoring	All (prior to error-filtering)	After error-filtering
MS	**0.218** (0.002)	0.126 (0.07)
MS*	−0.002 (*ns*)	−0.007 (*ns*)
ICM	0.020 (*ns*)	0.049 (*ns*)
ICM*	−0.099 (*ns*)	−0.065 (*ns*)
NM	0.088 (*ns*)	0.060 (*ns*)
NM*	0.002 (*ns*)	0.007 (*ns*)

Scoring methods bearing an asterisk treated condition IV (fragment absence in the two enzyme products) as missing values. statistically significant coefficients are highlighted in bold, ns denotes $P > 0.10$.

The number of condition IV loci averaged 25.6 (±0.2) % across individuals. Fragment filtering by scoring errors in replicated samples increased this percentage to 28.0 (±0.2) % (t = 13.28, $P < 0.0001$, $N = 200$). Individual quantitative estimates with and without fragment filtering were positively correlated ($r_s = 0.69$, $P < 0.0001$, $N = 200$). Moreover, error-filtering increased the estimates of relative frequencies of internal methylated cytosine loci (2.9 ± 0.1% increase) and condition IV uninformative loci (2.4 ± 0.2%) in the global sample ($P < 0.0001$ Student's paired *t*-test, $N = 200$), suggesting that these two categories are not inflated by processing errors.

If most condition IV loci were informative, that is derived from full methylation at the external or the two adjacent cytosines in the CCGG motif, then a significant positive correlation should be expected between their frequency and HPLC estimates. This prediction was clearly falsified by our data for *H. foetidus*, as the relationship was negative and highly significant ($r_s = -0.28$; $P < 0.0001$; $N = 200$) and remained significant after genotyping error-filtering ($r_s = -0.18$; $P = 0.008$; $N = 200$). Remarkably, this was the strongest, clearest relationship between cytosine methylation estimates obtained by HPLC and MSAP (see Table 3 for comparison). In addition, methylation estimates yielded by MSAP scoring methods that treated condition IV as missing data were not significantly related to genome-wide cytosine methylation estimates from HPLC (Table 3). Thus, for quantitative interpretation of MSAP, condition IV loci should be better considered null data, as suggested by Schulz *et al*. (2013).

Individual estimates based on the number of methylation-susceptible loci in the sample averaged 16.8% (±0.1) for internal cytosines, 38.4% (±0.3) for hemimethylated external cytosines and 55.2% (±0.3) for their sum, that is total methylation. Methylation scoring (MS = conditions II + III) correlated positively to HPLC estimates across individuals, but internal cytosine scoring (ICM = condition II) did not (Table 3). Filtering loci by scoring error reduced the magnitude of correlation between the two methods, but the changes were similar to those obtained by randomly decreasing the number of loci to half ($P > 0.05$ in all correlations after 10 000 randomizations) and hence cannot be attributed to quality filtering but rather to the reduction in number of loci.

Prevalence of this positive relationship between HPLC global cytosine methylation and some quantitative estimate based on MSAP results should be ascertained by replicating the study in more cases. Here, we explored the predictive accuracy of the relationship between HPLC and MS method across populations and found that HPLC methylation score was not similarly related to MS across populations ($F_{9,180} = 2.74$, $P = 0.0051$; for the population * MS interaction). Remarkably, not only the significance but also the sign of the relationship varied across sites (Fig. 1).

Discussion

Methylation of cytosine DNA is a major epigenetic mechanism through which environmental factors can alter the

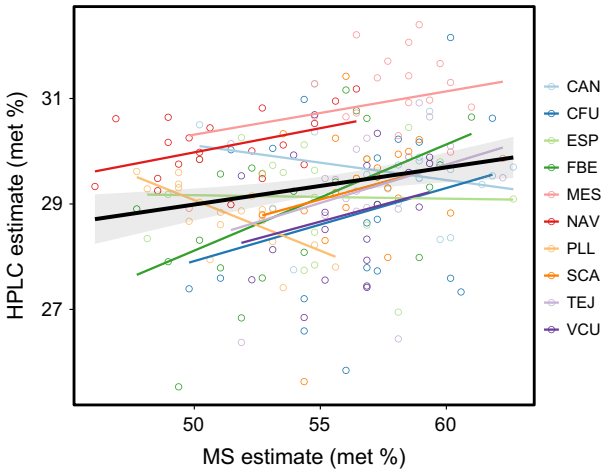

Fig. 1 Relationship between cytosine methylation estimates in young leaves of *Helleborus foetidus* obtained using HPLC and the percentage of cytosine methylation in CCGG sites obtained by methylation scoring of MSAP data. Each symbol represents an individual plant, and the 10 populations sampled are coded with different colours. The black thick line and grey area show the linear regression and 95% confidence interval obtained for the full sample, respectively (N = 197; 3 outliers were omitted to improve readability).

expression of genetically based traits (Finnegan 2010). In plants, global cytosine methylation varies widely between species, and interspecific differences show a strong phylogenetic signal and are evolutionarily correlated with genome size (Alonso *et al.* 2015). Within species, global cytosine methylation may also vary across individuals and populations, as shown here for the perennial herb *H. foetidus* (see also Alonso *et al.* 2014). These findings highlight the potential of this genomic feature for gaining a better understanding of the implications of DNA methylation at both intra- and interspecific scales (Flatscher *et al.* 2012; Richards *et al.* 2012; Alonso *et al.* 2014).

In this study, we investigated the relationships between global methylation estimates obtained by HPLC and the more widely used MSAP analyses, which have generally focused on the detection of changes in methylation status of cytosines in specific contexts distributed across the genome (Laird 2010; Schrey *et al.* 2013). Both the literature survey and the data presented here for the first time suggested that methylation estimates based on MSAP should not be directly used as synonym of, or valid surrogate for, global cytosine methylation level. In particular, in studies looking for methylation responses to stress, frequency of methylation-susceptible MSAP loci was very variable. This suggests that information obtained from the analysis of global DNA methylation and MSAP variability should be taken as complementary, an aspect apparently dismissed so far in ecological epigenetics studies of nonmodel plants (Schrey *et al.* 2013). Furthermore, reported changes in methylation status were frequently locus and context dependent (see, e.g., Labra *et al.* 2002; Lira-Medeiros *et al.* 2010; Uthup *et al.* 2011; Rico *et al.* 2014), a finding that emphasizes the value of analysing methylation at CG and CHG contexts separately (see also Schulz *et al.* 2014). In the following paragraphs, we discuss our findings obtained by comparing different MSAP scoring strategies and HPLC results in the perennial herb *H. foetidus*.

MSAP analyses and global cytosine methylation estimates in Helleborus foetidus

Scoring of MSAP data affected the relationship with global methylation estimates obtained by HPLC. Our results supported some of the scoring strategies previously proposed by Schulz *et al.* (2013). In particular, condition IV should be considered uninformative and remain as a null result, mainly for two reasons. First, because its interpretation as full methylation frequently found in the literature (see, e.g., Karan *et al.* 2012; Marconi *et al.* 2013) may often be incorrect, as exemplified here by results for *H. foetidus*. Remarkably, frequency of condition IV loci was not positively but negatively related to HPLC estimates, suggesting that such *informative* interpretation should be taken with caution when site mutation cannot be discarded. And second, treating condition IV cases as missing values may be appropriate to interpret variability (see, e.g., Herrera & Bazaga 2010, 2011), but this procedure tended to blur correlations between HPLC and MSAP methylation estimates, turning them statistically nonsignificant. Including a third schizomer (*Bsi*SI) with the same cleavage site but insensible to any methylation and/or the addition of the combined digestion of the two isoschizomers (*Hpa*II + *Msp*I) may help to reduce uncertainty of some ambiguous results and improve reliability of MSAP interpretations if subsequent increased costs are affordable (Fulneček & Kovařík 2014; Osabe *et al.* 2014).

Regarding the adequacy of combining condition II and III markers, it is important to emphasize that the relative contribution of CG, CHG and CHH contexts to global cytosine content estimates varies across plant species (Gruenbaum *et al.* 1981; Belanger & Hepburn 1990; Kumar *et al.* 1990; Kovařík *et al.* 1997). In addition, we found that percentages of methylation at CG and CHG contexts varied among species (Table 1; see also Feng *et al.* 2010; Zhong *et al.* 2013), thus suggesting that analyses of methylation at different contexts may be useful to reveal species-specific aspects. In the particular case of *H. foetidus*, global methylation was unrelated to frequency of methylation at the inner cytosine. Also, hemimethylation of external cytosine was more frequent

than methylation of inner cytosine(s) at 5′-CCGG sites, a pattern shared with tomato (Messeguer *et al.* 1991) but apparently not with *Arabidopsis thaliana*, rice or poplar (Feng *et al.* 2010). Interestingly, we found that the relationship between global HPLC and methylation at the specific cleavage sites was not constant among populations, which suggests the possibility that the relative importance of the methylation in these two specific nucleotide contexts may even vary at the intraspecific level, as shown, for instance, by Li *et al.* (2011) in relation to ploidy changes.

Analysing methylation at CG and CHG contexts separately is a valuable feature of MSAPs that could be profitably exploited to predict or interpret epigenetic changes in natural plant populations (Schulz *et al.* 2014). The separate analysis is particularly interesting because changes in methylation status at CG and CHG contexts depend on different families of methyltransferases and, thus, they could be independently regulated (Finnegan 2010). In brief, DNA METHYLTRANSFERASE family (METs) methylates specifically CG sequences; the plant specific family of CHROMOMETHYLASES (CMTs) are involved primarily in the maintenance of symmetrical CHG methylation (where H = A, C, T) and also in de novo methylation; and the DOMAINS REARRANGED METHYLTRANSFERASES (DRMs) maintain asymmetric CHH methylation through persistent de novo methylation (for details, see Finnegan 2010 and references therein). Furthermore, whole-genome bisulphite sequencing in a few model species indicated that gene body methylation is almost exclusively restricted to CG sites, which is in marked contrast to the methylation typically seen for repeat and transposon sequences, which includes CG, CHG and CHH sites (Teixeira & Colot 2009; Feng *et al.* 2010; Gent *et al.* 2012; Takuno & Gaut 2013; Zhong *et al.* 2013; but see Uthup *et al.* 2011). Such whole methylome studies also confirmed that methylation percentages in each specific context (CG, CHG, CHH) varied widely among the few model species assayed to date (tomato, maize, *A. thaliana*), which clearly supports the conclusion based on MSAPs noted above that the pattern of methylation distribution is not universal in angiosperms (Gent *et al.* 2012; Zhong *et al.* 2013). In particular, maize and tomato, two species with larger genomes and higher frequency of transposons than *A. thaliana*, exhibited higher global methylation levels and higher prevalence of methylation at CHG and CHH sites (Gent *et al.* 2012; Zhong *et al.* 2013). Additional studies are required to reveal the specific nucleotide contexts more relevant for different plant species and adaptive processes of interest, distinguishing whether they depend more on gene (in)activation at CG sites or transposon activity regulation occurring in different nucleotide contexts (Gent *et al.* 2012; Takuno & Gaut 2013; Zhong *et al.* 2013). The separate analysis of MSAP conditions II and III combined with overall estimation by HPLC may well serve this aim in the meantime until WGBS becomes affordable for nonmodel species.

Conclusions

MSAP is a powerful technique to investigate diversity of cytosine methylation in species without a reference sequenced genome. It is currently used to assess methylation correlates with individual phenotypic traits, environmental conditions and species divergence, to identify loci under selection or to quantify transmissibility across generations of the methylation status of specific loci (Schrey *et al.* 2013; Herrera *et al.* 2013, 2014; Medrano *et al.* 2014). However, methylation estimates based on MSAP alone should not be directly used as valid surrogates for global cytosine methylation percentage, mainly because the relative abundance of the specific contexts assayed by different endonucleases does actually change across species. The analyses of HPLC global cytosine methylation illustrate variability in global methylation of cytosines across plant species (Alonso *et al.* 2015; see also Varriale 2014 for vertebrates). They could also be useful to evaluate the magnitude of variation at different hierarchical levels of intraspecific variation (e.g. within and among populations). Combined use of the two techniques may help to distil all information contained in MSAP of wild-growing plants and its relationship with species-specific traits, individual phenotypes and ecological conditions.

Acknowledgements

Esmeralda López provided laboratory assistance. We thank Axios Review editors Tim Vines and Jeffrey Ross-Ibarra, and four anonymous referees, for the critical review of the manuscript, and Kathy L. Thomas for kindly revising the writing. Permission to work in Sierra de Cazorla was provided by Consejería de Medio Ambiente, Junta de Andalucía. Financial support for this study was provided by the Spanish Ministerio de Economía y Competitividad through the Research Project CGL2013-43352-P, and Estación Biológica de Doñana (EBD-CSIC) through the Severo Ochoa Program for Centres of Excellence (SEV-2012-0262).

References

Akimoto K, Katakami H, Kim H-J *et al.* (2007) Epigenetic inheritance in rice plants. *Annals of Botany*, **100**, 205–217.

Alonso C, Pérez R, Bazaga P, Medrano M, Herrera CM (2014) Individual variation in size and fecundity is correlated with differences in global DNA cytosine methylation in the perennial herb *Helleborus foetidus* (Ranunculaceae). *American Journal of Botany*, **101**, 1309–1313.

Alonso C, Pérez R, Bazaga P, Herrera CM (2015) Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. *Frontiers in Genetics*, **6**, 4.

Amoah S, Kurup S, Lopez CMR *et al.* (2012) A hypomethylated population of *Brassica rapa* for forward and reverse epi-genetics. *BMC Plant Biology*, **12**, 193.

Belanger FC, Hepburn AG (1990) The evolution of CpNpG methylation in plants. *Journal of Molecular Ecology*, **30**, 26–35.

Bonin A, Bellemain E, Bronken Eidesen P *et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecology Letters*, **11**, 106–115.

Cai Q, Chinnappa C (1999) Cytosine methylation levels in the genome of *Stellaria longipes*. *Journal of Bioscience*, **1**, 27–33.

Chinnusamy V, Zhu JK (2009) Epigenetic regulation of stress responses in plants. *Current Opinion in Plant Biology*, **12**, 133–139.

Cicatelli A, Todeschini V, Lingua G, Biondi S, Torrigiani P, Castiglione S (2014) Epigenetic control of heavy metal stress response in mycorrhizal versus non-mycorrhizal poplar plants. *Environmental Science and Pollution Research*, **21**, 1723–1737.

Diez CM, Roessler K, Gaut BS (2014) Epigenetics and plant genome evolution. *Current Opinion in Plant Biology*, **18**, 1–8.

Feng S, Cokus SJ, Zhang X *et al.* (2010) Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences, USA*, **107**, 8689–8694.

Finnegan EJ (2010) DNA methylation: a dynamic regulator of genome organization and gene expression in plants. In: *Plant Developmental Biology—Biotechnological Perspectives* (eds Pua E-C, Davey MR), vol. 2, pp. 295–323. Springer-Verlag, Berlin.

Finnegan EJ, Genger RK, Peacock WJ, Dennis ES (1998) DNA methylation in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**, 223–247.

Flatscher R, Frajman B, Schönswetter P, Paun O (2012) Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation? *Genetics Research International*, **2012**, doi:10.1155/2012/698421.

Fraga ME, Esteller M (2002) DNA methylation: a profile of methods and applications. *BioTechniques*, **33**, 632–649.

Fulneček J, Kovařík A (2014) How to interpret Methylation Sensitive Amplified Polymorphism (MSAP) profiles? *BMC Genetics*, **15**, 2–9.

Gent JI, Ellis NA, Guo L *et al.* (2012) CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Research*, **23**, 628–637.

Grant-Downton RT, Dickinson HG (2005) Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Annals of Botany*, **96**, 1143–1164.

Grant-Downton RT, Dickinson HG (2006) Epigenetics and its implications for plant biology. 2. The "Epigenetic Epiphany": epigenetics, evolution and beyond. *Annals of Botany*, **97**, 11–27.

Grativol C, Hemerly AS, Gomes Ferreira PC (2012) Genetic and epigenetic regulation of stress responses in natural plant populations. *Biochimica et Biophysica Acta*, **1819**, 176–185.

Greco M, Chiappetta A, Bruno L, Bitonti MB (2012) In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *Journal of Experimental Botany*, **63**, 695–709.

Gruenbaum Y, Navehmany T, Cedar H, Razin A (1981) Sequence specificity of methylation in higher plant DNA. *Nature*, **292**, 860–862.

Herrera CM, Bazaga P (2010) Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytologist*, **187**, 867–876.

Herrera CM, Bazaga P (2011) Untangling individual variation in natural populations: ecological, genetic and epigenetic correlates of long-term inequality in herbivory. *Molecular Ecology*, **20**, 1675–1688.

Herrera CM, Medrano M, Bazaga P (2013) Epigenetic differentiation persists after male gametogenesis in natural populations of the perennial herb *Helleborus foetidus* (Ranunculaceae). *PLoS One*, **8**, e70730.

Herrera CM, Medrano M, Bazaga P (2014) Variation in DNA methylation transmissibility, genetic heterogeneity and fecundity-related traits in natural populations of the perennial herb *Helleborus foetidus*. *Molecular Ecology*, **23**, 1085–1095.

Hirsch S, Baumberger R, Grossniklaus U (2013) Epigenetic variation, inheritance, and selection in plant populations. *Cold Spring Harbor Symposia on Quantitative Biology*, **77**, 97–104.

Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Quarterly Review of Biology*, **84**, 131–176.

Johnston J, Harding K, Bremner DH *et al.* (2005) HPLC analysis of plant DNA methylation: a study of critical methodological factors. *Plant Physiology and Biochemistry*, **43**, 844–853.

Kakutani T, Munakata K, Richards EJ, Hirochika H (1999) Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. *Genetics*, **151**, 831–838.

Karan R, DeLeon T, Biradar H, Subudhi PK (2012) Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PLoS One*, **7**, e40203.

Kato M, Miura A, Bender J, Jacobsen SE, Kakutani T (2003) Role of CG and non-CG methylation in immobilization of transposons in *Arabidopsis*. *Current Biology*, **13**, 421–426.

Kim KD, El-Baidouri M, Jackson SA (2014) Accessing epigenetic variation in the plant methylome. *Briefings in Functional Genomics*, **13**, 318–327.

Kovařík A, Matyasek R, Leitch A *et al.* (1997) Variability in CpNpG methylation in higher plant genomes. *Gene*, **204**, 25–33.

Kumar LS, Hendre RR, Ranjekar PK (1990) 5-Methylcytosine content and methylation status in six millet DNAs. *Journal of Biosciences*, **15**, 47–52.

Labra M, Ghiani A, Citterio S *et al.* (2002) Analysis of cytosine methylation pattern in response to water deficit in pea root tips. *Plant Biology*, **4**, 694–699.

Laird PW (2010) Principles and challenges of genomewide DNA methylation analysis. *Nature Reviews Genetics*, **11**, 191–203.

Leutwiler LS, Hough-Evans BR, Meyerowitz EM (1984) The DNA of *Arabidopsis thaliana*. *Molecular and General Genetics*, **194**, 15–23.

Li A, Hu B-Q, Xue Z-Y *et al.* (2011) DNA methylation in genomes of several annual herbaceous and woody perennial plants of varying ploidy as detected by MSAP. *Plant Molecular Biology Reporter*, **29**, 784–793.

Lira-Medeiros CF, Parisod C, Fernandes RA *et al.* (2010) Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS One*, **5**, e10326.

Lisanti S, Omar W, Tomaszewski B *et al.* (2013) Comparison of methods for quantification of global DNA methylation in human cells and tissues. *PLoS One*, **8**, e79044.

Lopez-Torres A, Yanez-Barrientos E, Wrobel K, Wrobel K (2011) Selective derivatization of cytosine and methylcytosine moieties with 2-bromoacetophenone for submicrogram DNA methylation analysis by reversed phase HPLC with spectrofluorimetric detection. *Analytical Chemistry*, **83**, 7999–8005.

Marconi G, Pace R, Traini A *et al.* (2013) Use of MSAP markers to analyse the effects of salt stress on DNA methylation in rapeseed (*Brassica napus* var. *oleifera*). *PLoS One*, **8**, e75597.

Matassi G, Melis R, Kuo K *et al.* (1992) Large-scale methylation patterns in the nuclear genomes of plants. *Gene*, **122**, 239–245.

Medrano M, Herrera CM, Bazaga P (2014) Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology*, **23**, 4926–4938.

Messeguer R, Ganal MW, Steffens JC, Tanksley SD (1991) Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. *Plant Molecular Biology*, **16**, 753–770.

Morrish FM, Vasil IK (1989) DNA methylation and embryogenic competence in leaves and callus of napiergrass (*Pennisetum purpureum* Schum.). *Plant Physiology*, **90**, 37–40.

Osabe K, Clement JD, Bedon F *et al.* (2014) Genetic and DNA methylation changes in cotton (Gossypium) genotypes and tissues. *PLoS One*, **9**, e86049.

Papa CM, Springer NM, Muszynski MG, Meeley R, Kaeppler SM (2001) Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation. *Plant Cell*, **13**, 1919–1928.

Paun O, Bateman RM, Fay MF *et al.* (2010) Stable epigenetic effects impact adaptation in allopolyploid orchids (*Dactylorhiza*: Orchidaceae). *Molecular Biology and Evolution*, **27**, 2465–2473.

Pérez-Figueroa A (2013) msap: a tool for the statistical analysis of methylation sensitive amplified polymorphism data. *Molecular Ecology Resources*, **13**, 522–527.

Richards CL, Verhoeven KJF, Bossdorf O (2012) Evolutionary significance of epigenetic variation. In: *Plant Genome Diversity* (eds Wendel JF, Greilhuber J, Dolezel J, Leitch IJ), vol. 1, pp. 257–274. Springer, Vienna.

Rico L, Ogaya R, Barbeta A, Peñuelas J (2014) Changes in DNA methylation fingerprint of *Quercus ilex* trees in response to experimental field drought simulating projected climate change. *Plant Biology*, **16**, 419–427.

Rozhon W, Baubec T, Mayerhofer J, Scheid OM, Jonak C (2008) Rapid quantification of global DNA methylation by isocratic cation exchange high-performance liquid chromatography. *Analytical Biochemistry*, **375**, 354–360.

Salmon A, Clotault J, Jenczewski E, Chable V, Manzanares-Dauleux MJ (2008) *Brassica oleracea* displays a high level of DNA methylation polymorphism. *Plant Science*, **174**, 61–70.

SAS Institute (2008) *SAS for Windows*. SAS Institute, Cary, North Carolina.

Schmitz RJ, Schultz MD, Lewsey MG *et al.* (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science*, **334**, 369–373.

Schrey AW, Alvarez M, Foust CM *et al.* (2013) Ecological epigenetics: beyond MS-AFLP. *Integrative and Comparative Biology*, **53**, 340–350.

Schulz B, Eckstein RL, Durka W (2013) Scoring and analysis of methylation-sensitive amplification polymorphisms for epigenetic population studies. *Molecular Ecology Resources*, **13**, 642–653.

Schulz B, Eckstein RL, Durka W (2014) Epigenetic variation reflects dynamic habitat conditions in a rare floodplain herb. *Molecular Ecology*, **23**, 3523–3537.

Takuno S, Gaut BS (2013) Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proceedings of the National Academy of Sciences, USA*, **110**, 1797–1802.

Teixeira FK, Colot V (2009) Gene body DNA methylation in plants: a means to an end or an end to a means? *EMBO Journal*, **28**, 997–998.

Tsaftaris AS, Polidoros AN (2000) DNA methylation and plant breeding. *Plant Breeding Reviews*, **18**, 87–176.

Uthup TK, Ravindran M, Bini K, Thakurdas S (2011) Divergent DNA methylation patterns associated with abiotic stress in *Hevea brasiliensis*. *Molecular Plant*, **4**, 996–1013.

Varriale A (2014) DNA methylation, epigenetics, and evolution in vertebrates: facts and challenges. *International Journal of Evolutionary Biology*, **2014**, 475981. doi:10.1155/2014/475981.

Verhoeven KJF, Jansen JJ, van Dijk PJ, Biere A (2010) Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist*, **185**, 1108–1118.

Whitlock R, Hipperson H, Mannarelli M, Butlin RK, Burke T (2008) An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error. *Molecular Ecology Resources*, **8**, 725–735.

Zhong S, Fei Z, Chen Y-R *et al.* (2013) Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology*, **31**, 154–159.

C.A., R.P. and C.M.H. conceived and designed the study; R.P. developed the HPLC method; C.A., C.M.H., M.M. and P.B. analysed the data; C.A. and C.M.H. wrote the manuscript; all authors contribute to sample processing and writing refinement.

Data accessibility

Sampling details can be found as supporting information in Medrano *et al.* (2014) that can be downloaded from the online version doi:10.1111/mec.12911.

MSAP data from the Medrano *et al.* (2014) study are archived at Dryad doi:10.5061/dryad.fr2k8.

HPLC data and methylation scoring obtained from MSAP for this study are archived at Dryad doi:10.5061/dryad.04d0d.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Primer combinations, number of fragments and estimates of scoring error rates for MSAP analyses.

Table S2 Results of the literature survey on methylation changes associated to stress treatments obtained by MSAP in different plant species and tissues.

Appendix S1 Analysis of the effects of scoring MSAP fragments with different semi-automated thresholds on the relationship between global methylation and MSAP estimates.