Reduced-representation bisulfite sequencing finds epigenome-wide response to oil pollution in the foundation plant *Spartina alterniflora*

Mariano Alvarez*
Department of Biology, Duke University, Durham, NC, 27708, USA;
marianoferdinandalvarez@gmail.com

Marta Robertson*
IRHS-INRA 49071 Beaucouzé Cedex, France; marta.h.robertson@gmail.com

Thomas van Gurp
Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, the Netherlands;

Cornelis A. M. Wagemaker
Department of Experimental Plant Ecology, Radboud University, Nijmegen, the Netherlands**;**
n.wagemaker@science.ru.nl

Delphine Giraud
UMR CNRS 6553 ECOBIO, OSUR, Université de Rennes 1, Campus Scientifique de Beaulieu, 35042 Rennes, France; delphine.giraud@univ-rennes1.fr

Malika L. Ainouche
UMR CNRS 6553 ECOBIO, OSUR, Université de Rennes 1, Campus Scientifique de Beaulieu, 35042 Rennes, France ; malika.ainouche@univ-rennes1.fr

Armel Salmon
UMR CNRS 6553 ECOBIO, OSUR, Université de Rennes 1, Campus Scientifique de Beaulieu, 35042 Rennes, France; armel.salmon@univ-rennes1.fr

Koen J. F. Verhoeven
Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, the Netherlands; k.verhoeven@nioo.knaw.nl

Christina L. Richards
Department of Integrative Biology, University of South Florida, Tampa, FL 33620, USA; clr@usf.edu

*indicates equal contribution by authors

Running title: Epigenetic response to pollution
Keywords: AFLP, DNA methylation, epiGBS, epigenetics, foundation species, oil pollution, salt marsh, *Spartina alterniflora*
Type of article: Letter
147 words in abstract, 4989 words in main text
80 references, 3 figures and 2 tables; 1 supplemental table, 4 supplemental figures

Correspondence: Christina L. Richards, Department of Integrative Biology, University of South Florida, 4202 East Fowler Avenue, Tampa, FL 33617, USA; phone: +001 813 974 5090; fax: +001 813 974 3263; clr@usf.edu

Author Contributions

CLR & KJFV conceived the study. CLR, KJFV, MA, MR, MLA and AS designed the experiments and analyses. MA, MR, CAMW and TVG did the laboratory work. MA, MR, and TVG analyzed the epiGBS data. CLR, MA, DG, and AS analyzed the transcriptome and gene expression data. CLR, MA and MR wrote the first draft of the manuscript. All co-authors provided input and revisions to the manuscript.

Data accessibility

Raw data files are available on Dryad at XXX. Processed data and R scripts are available on github.com/alvarezmf/DWH_epigbs.

Abstract

The application of genomics technology in ecological contexts allows for examination of how rapid environmental change may shape standing molecular level variation and organismal response. We previously demonstrated an effect of oil pollution on gene expression patterns and genetic variation, but not methylation variation, in oil-exposed populations of the foundation salt marsh grass, *Spartina alterniflora*. Here, we used a reduced representation bisulfite sequencing approach, epigenotyping by sequencing (epiGBS), to examine relationships among DNA sequence, DNA methylation, gene expression, and exposure to oil pollution. With the increased resolution of epiGBS, we document genetic and methylation differentiation between oil-exposed and unexposed populations, and a correlation of genome-wide methylation patterns and gene expression, independent of population genetic structure. Overall, these findings demonstrate that variation in DNA methylation is abundant, responsive, and correlated to gene expression in natural populations, and may represent an important component of the response to environmental stress.

Introduction

The application of molecular techniques to ecological questions can provide insight into the mechanisms that govern ecological interactions at the most basic levels of biological organization. Studies across a diversity of organisms have described the association of genetic variation with environmental factors (Feder & Mitchell-Olds, 2003; Andrew *et al.*, 2013). More recently, transcriptomics studies in natural populations have identified gene expression differences that underlie phenotypic plasticity, genotype-by-environment interactions, and local adaptation, and that some of these differences are only elicited in natural environments (Nicotra *et al.*, 2010; Andrew *et al.*, 2013; Alvarez *et al.*, 2015). Hence, genetic variation and gene expression variation can translate into trait variation that contributes to organismal performance with important population- and community-level ecological effects (Whitham *et al.*, 2006; Hughes *et al.*, 2008; Schoener, 2011; Alvarez *et al*., 2015). Additional layers of regulatory variation, including chromatin modifications, small RNAs, and other non-coding variants, can mediate these changes in genotypic expression and phenotype, but are infrequently studied in ecological settings (Nicotra *et al.*, 2010; Richards *et al.*, 2017).

Like genetic variation and gene expression variation, chromatin modifications such as DNA methylation can also vary among individuals within populations, and contribute to phenotypic variation through a variety of regulatory roles (Becker & Weigel, 2012; Jackson, 2017; Richards *et al.*, 2017). For example, DNA methylation may affect phenotype and subsequent ecological interactions by modulating the expression of genes (Alvarez *et al*., 2015, 2018), the types of transcripts that genes produce (Maor *et al.*, 2015), the movement of mobile elements (Matzke & Mosher, 2014), and the production of structural variants (Yelina *et al.*, 2015; Underwood *et al.*, 2018). On the other hand, changes in genetic sequence or gene expression may cause variation in patterns of DNA methylation, creating a relationship that is bidirectional and heterogeneous across the genome (Secco *et al.*, 2015; Meng *et al.*, 2016; Niederhuth & Schmitz, 2017). DNA methylation can vary in response to environment, and has been correlated to habitat types, exposure to stress, and shifts in species range (Verhoeven *et al*., 2010; Richards *et al*., 2012a; Liebl *et al*., 2013; Xie & Li *et al*., 2015; Foust *et al*., 2016). While some studies have provided insight into how the complex interaction between genetic variation, DNA methylation, and the regulation of gene expression translates into organismal response (e.g Wibowo *et al.* 2016), these mechanisms are only understood in a few model species, partly

because of the limited genomic resources available for most species (Ainouche *et al.*, 2009; Richards *et al.*, 2017). Increasing application of high-resolution methods to understand genome-level response to environmental conditions in natural populations is one of the central current tasks of ecological genomics.

In 2010, the *Deepwater Horizon* (*DWH*) oil spill developed into the largest marine oil spill in history (National Commission on the BP *Deepwater Horizon* oil spill, 2011), and became an opportunity to apply emerging genomics technologies in natural populations to examine responses to this recurrent anthropogenic stress in a diversity of organisms (e.g. Hazen *et al*. 2010; Whitehead *et al*. 2012; Kimes *et al*. 2013, 2014; Alvarez *et al.* 2018; DeLeo *et al.* 2018). A mixture of crude oil made and dispersants made landfall along 1,773 kilometers on the shorelines of Louisiana, Mississippi, Alabama and Florida (Mendelssohn *et al.*, 2012; Michel *et al.,* 2013). Nearly half of the habitat affected was salt marsh, which provides valuable ecosystem functions such as providing nurseries for birds and fish, and buffering storm and wave action (Day *et al.*, 2007; Mendelssohn *et al.*, 2012; Michel *et al.,* 2013). Gulf of Mexico salt marshes are dominated by the foundation species *Spartina alterniflora,* which is remarkably resilient to a variety of environmental stressors (Pennings & Bertness, 2001; Baisakh *et al.*, 2008; Baisakh & Subudhi, 2009; Silliman *et al.*, 2012; Bedre *et al.*, 2016). Crude oil exposure from the *DWH* oil spill resulted in reduced carbon fixation, reduced transpiration, and extensive above-ground dieback in *S. alterniflora* populations (Lin & Mendelssohn, 2012; Silliman *et al.,* 2012), but oil-affected populations showed partial to complete recovery within seven months of the spill (Lin *et al.,* 2016). However, the genomic and regulatory mechanisms that underlie this remarkable recovery have been poorly characterized.

Despite the tremendous resilience of *S. alterniflora* to the *DWH* spill, we discovered genetic divergence of individuals from oil-exposed areas and nearby uncontaminated areas (Robertson *et al.*, 2017). We also found that pollution tolerance may be modulated by a diverse set of candidate genes (Alvarez *et al.*, 2018). In particular, we provided support for the role of an epigenetic regulator (a homolog of the histone methyltransferase SUVH5), which significantly altered growth and flower production under crude oil stress in experimental studies (Alvarez *et al.*, 2018). Considering that histone modifications may regulate and be regulated by DNA methylation (reviewed in Du *et al.*, 2015), we expected to find a strong correlation between patterns of DNA methylation and crude oil exposure. However, while a few DNA methylation

loci (measured via methylation sensitive amplified fragment length polymorphism; MS-AFLP) were correlated with oil exposure, we did not find genome-wide patterns in DNA methylation correlated with oil exposure in *S. alterniflora* (Robertson *et al.*, 2017). Our findings supported genetic differentiation in response to exposure to crude oil, but suggested that our ability to detect changes in DNA methylation may have been limited by the number of MS-AFLP markers (39 polymorphic loci). Further, although our previous study identified population-level patterns of divergence, we were unable to associate the anonymous AFLP loci with gene function (Schrey *et al.*, 2013).

In this study, we used a recently developed reduced representation bisulfite sequencing technique, epigenotyping by sequencing (epiGBS), to generate a more robust DNA sequence and DNA methylation data set (van Gurp *et al.*, 2016). We expected that the increased resolution, both in number and in detail of the markers, provided by this sequencing approach would allow us to identify fine scale DNA methylation structure that was not apparent in our previous MS-AFLP study. We further expected that biasing toward coding regions of the genome with the use of methylation sensitive enzymes in the epiGBS protocol (van Gurp *et al.*, 2016) would allow us to find evidence that DNA methylation was correlated with changes in gene function since some fragments might overlap with the promoter or 5' end of the coding regions of genes (Niederhuth & Bewick *et al.*, 2016). Therefore, we hypothesized that genome-wide expression, as well as expression of candidate genes, could be correlated to changes in genetic variation and DNA methylation.

Materials and Methods

Sample Collection

We collected individuals from the leading edge of the marsh at three contaminated and three uncontaminated sites near Grand Isle, Louisiana and Bay St. Louis, Mississippi in August 2010, four months after the *DWH* oil spill as described in previous studies (Table 1; Robertson *et al.,* 2017; Alvarez *et al.,* 2018). These sites were naturally highly variable in conditions, but all sites supported monocultures of *S. alterniflora*. Contaminated sites were identified by the visual presence of oil on the sediment, and substantial above-ground dieback of *S. alterniflora* on the leading edge of the marsh with *S. alterniflora* plants growing up through the dead wrack. Nearby

uncontaminated sites did not have any visual signs of the presence of oil or noticeable dieback of the above ground portions of *S. alterniflora* (contamination status confirmed via NRDA databases, Robertson *et al.*, 2017). To standardize age and minimize developmental bias in sampling, we collected the third fully expanded leaf from each of eight individuals, spaced 10 meters apart at each of the six sites (N=48). Leaf samples were immediately frozen in liquid nitrogen to prevent degradation, and kept frozen during transport to the University of South Florida for processing and analysis.

DNA extractions and library prep

We isolated DNA from each sample (48) using the Qiagen DNeasy plant mini kit according to the manufacturer's protocol. We prepared epiGBS libraries *sensu* van Gurp *et al.* (2016). Briefly, isolated DNA was fragmented with the enzyme PstI, which is sensitive to CHG methylation and biases resulting libraries toward coding regions (van Gurp *et al.,* 2016). After digestion, adapters with variable barcodes were ligated to either end of the resulting fragments. Adapters contained methylated cytosines to ensure their sequence fidelity through the subsequent bisulfite treatment. We used the Zymo EZ Lightning methylation kit to bisulfite treat and clean the DNA. Libraries were then amplified with the KAPA Uracil Hotstart Ready Mix with the following PCR conditions: an initial denaturation step at 98ºC for 1 min followed by 16 cycles of 98ºC for 15s, 60ºC for 30s, and 72ºC for 30s, with a final extension of 72ºC for 5 min. Paired-end reads were sequenced at the University of Florida Interdisciplinary Center for Biotechnology Research on the Illumina HiSeq 5000.

Data pre-processing and mapping to transcriptome

We used the epiGBS pipeline (van Gurp *et al.,* 2016) to demultiplex samples, trim adapter sequences, and assemble the *de novo* reference sequence (https://github.com/thomasvangurp/epiGBS). To improve variant calling while accounting for polyploidy, we realigned individuals to the *de novo* reference (generated from the epiGBS pipeline), separately using BWAmeth (Pedersen *et al.,* 2014), and called variants using the GATK pipeline (DePristo *et al.,* 2011) with a specified ploidy of six. Methylation was tabulated using the existing epiGBS analysis pipeline. Both SNP and methylation data were filtered separately to include only loci that were present in more than 90% of individuals, with no more

than 50% missing from any one individual. Thus, each locus had, at most, 10% missing data, while each individual had less than 50% missing data across all loci. We removed loci with less than 5x depth of coverage and remaining missing data were imputed via a k-nearest neighbors approach, yielding a matrix of "common loci" (impute, Hastie *et al.*, 2018).

All fragments were mapped to the published *S. alterniflora* transcriptome (Boutte *et al.,* 2016) and the *Oryza sativa* genome (Michigan State University version 7, Kawahara *et al.*, 2013) using BLAST (Altschul *et al.,* 1997). We estimated the number and order of exons by mapping the *S. alterniflora* transcriptome to *Oryza sativa*. We also tested for random sampling across exons among the epiGBS fragments that occurred within coding regions using a Kolmogorov-Smirnov test. We used BLAST (Altschul *et al.,* 1997) and RepeatExplorer (Novak *et al.,* 2013) to compare our sequenced fragments to the *S. alterniflora* transcriptome (Ferreira de Carvalho *et al.,* 2013, 2017; Boutte *et al.*, 2016) and known repeat elements, respectively.

Population genetics

Our RRBS did not provide sufficient sequencing depth to estimate hexaploid genotype likelihoods with confidence, particularly considering the high levels of heterozygosity measured in *S. alterniflora* populations (Richards *et al.*, 2004; Travis & Hester, 2005) and the lack of a high-quality reference genome (Dufresne *et al.,* 2014; Boutte *et al.,* 2016). We therefore used the frequency of the most common allele within an individual at each polymorphic locus to estimate allelic diversity, and to estimate Nei's genetic distances among individuals and sites (StAMPP, Pembleton *et al.,* 2013).

We obtained pairwise F_{ST} values between populations to test for significance of pairwise differences. Following recommendations from Meirmans (2015) to minimize false positives, we used distance-based redundancy analysis (RDA; Vegan, Oksanen *et al.* 2017) to assess isolation by distance by regressing the genetic relatedness matrix against latitude and longitude (genetic distance ~ latitude * longitude) for each site using constrained analysis of principal coordinates (Vegan, Oksanen *et al.,* 2017).

To quantify the relationship between genome-wide variation and environmental conditions, we used RDA with surrogate variables to represent the complex differences among sites with the formula (genetic distance ~ oil exposure + surrogate variables). Redundancy analysis allowed us to assess the joint influence of all SNPs simultaneously (Lasky *et al.*, 2015;

Forester et al. 2016). Although we attempted to control for variation among samples and sites with a replicated sampling strategy, rather than using a single term for "site", we included parameters identified by surrogate variable analysis (SVA), which provides a method to detect residual variation due to unmeasured differences among populations including environmental variation, life history variation, and geographical separation, between samples that may distort analyses (Leek *et al.*, 2017). We estimated surrogate variables by fitting a full and null model to the appropriate (i.e. DNA sequence allele frequency or DNA methylation frequency) high dimensional data matrix and examining the principal components of the residuals (Leek & Storey, 2007). We used a permutational test (999 permutations) to assess the significance of differences between the treatment and control groups, given the covariates. We also tested for differences in the amount of dispersion around centroids (multivariate variance) between oil-exposed or unexposed populations (PERMDISP2, Vegan, Oksanen *et al.* 2017). We corrected the resulting P values for false discovery using a Holm correction, and visualized results using multidimensional scaling (MDS, Vegan, Oksanen *et al.* 2017).

To identify SNPs that were significantly correlated with oil contamination, we used binomial linear mixed modeling (MACAU; Lea *et al.,* 2015) including a genetic relatedness matrix to account for population structure and the surrogate variables as covariates. All locus-by-locus association scans were corrected for family-wise false discovery using the Q-value package in R (Storey *et al.,* 2015).

Methylation analysis

During the filtering process, loci were annotated with their methylation context, but all contexts were pooled for distance-based analyses as well as family-wise error rate corrections after locus-by-locus modeling. We tabulated methylation frequency at each locus, defined as the fraction of methylated cytosines observed out of the total number of cytosines measured at a given locus (methylated cytosines/(methylated+unmethylated cytosines)), and used methylation frequencies to generate a methylation distance matrix in R.

To identify signatures of DNA methylation variation that were correlated with oil exposure and not genetic structure, we used principal components analysis to reduce the genetic data to orthogonal components, and identified 8 of the 42 possible principal components axes that were significantly correlated with methylation. We then modeled the impact of oil exposure

to genome-wide patterns of DNA methylation by jointly assessing the influence of all polymorphic methylation loci concurrently via RDA (Vegan, Oksanen *et al.* 2017) with the formula (methylation distance ~ oil exposure + surrogate variables + significant principal components of allele frequency data). We corrected the resulting P values for false discovery using a Holm correction, and visualized results using multidimensional scaling (MDS, Vegan, Oksanen *et al.* 2017).

To identify differentially methylated positions (DMPs) between contaminated and uncontaminated samples, we used binomial linear mixed modeling (MACAU; Lea *et al.,* 2015), using the genetic relatedness matrix and surrogate variables as covariates to control for population structure. We corrected locus-specific P-values for multiple testing (qvalue, Storey *et al.,* 2015), and tested for overrepresentation of cytosine contexts (CG, CHG, CHH) using an exact binomial test, implemented in R. Because our epiGBS fragments rarely exceeded 200bp, we did not identify differentially methylated regions.

Relationships to gene expression variation

We used transcriptome data from our previous analyses of three mRNA pools of three individuals from each of the six sites to assess the relationship between variation in gene expression, genetic variation, and methylation variation (Alvarez *et al.*, 2018). We found 3,334 genes that responded to crude oil exposure in *S. alterniflora* out of 15,950 genes that were assessed. In order to make the epiGBS data comparable to our pools of mRNA, we concatenated SNPs and methylation polymorphisms from individuals into *in silico* sample pools by averaging values at individual loci across samples within pools of the same individuals that were used in the gene expression analysis. We then calculated genetic, expression, and methylation distances between sample pools and used multiple regression on distance matrices to assess the relationship between all three data types (MRM; ecodist, Goslee & Urban, 2007).

We assessed the correlation between gene expression, genetic variation, and methylation variation within the same fragment (i.e. in *cis*) gene-by-gene using only microarray probes and epiGBS fragments that could be mapped to the *S. alterniflora* transcriptome. We regressed the expression values of each gene against the methylation frequencies and most common allele frequencies at corresponding epiGBS fragments using the *lm* function in R with the formula (gene expression ~ *cis* methylation + *cis* SNP). We summarized variation across each epiGBS

fragment for multiple SNP or methylation loci by using the first principal component of each data type. We interpreted significant effects of SNPs or DNA methylation to mean that the expression of a given gene depended on the frequency of the nucleotide or methylation polymorphism. We corrected for false discovery using qvalue (Storey *et al.*, 2015).

Results

epiGBS yields informative genetic and methylation loci

Following library preparation, we removed four samples due to stochastic under-sequencing. The libraries for 44 individuals (Table 1) generated 6,809,826 total raw sequencing reads, of which 3,833,653 (56.3%) could be matched to their original mate strand. *De novo* assembly using the epiGBS pipeline resulted in 36,131 contiguous fragments (19-202 basepairs) with a total length of 5,441,437 bp (Figs. **S1, S2**). Given the size of the *S. alterniflora* genome (2C values = 4.3 pg; Fortune *et al.*, 2008), our epiGBS approach assayed approximately 0.3% of the genome. However, fragments that were >90% similar were merged, and polyploid homeologs may have been concatenated. With BLAST, we found 10,103 fragments mapped to 2,718 transcripts in *S. alterniflora* transcriptome. We found no bias in mapping to particular gene regions (Fig. **S3**), but 1,571 transcripts (57.8%) contained multiple epiGBS fragments that align to the same place, and 296 (10.9%) contained multiple epiGBS fragments that mapped to different places within the same gene. We suspect that multiple epiGBS fragments map to the same location because some epiGBS contiguous fragments represent different homeologs of the same region, which mapped to the same location. Only 1% of reads map to repetitive elements, confirming that *Pst1*-fragmented libraries were biased away from highly methylated, repetitive regions (van Gurp *et al.,* 2016).

Genetic differentiation

Our initial sequencing run yielded 171,205 SNPs across all individuals. After filtering to common loci and removing invariant sites, we imputed 10-25% missing genetic data and 10-50% methylation data per sample (Fig. **S4**), resulting in 6,521 SNP loci. We hypothesize that this significant data loss after filtering was due to insufficient sampling depth across an exceptionally large hexaploid genome, and a preponderance of restriction sites. Of these 6,521 loci, 243 SNPs occurred in transcripts. Pairwise F_{ST} calculations showed that all sites were significantly different

323 from each other (Table 3), with no evidence of isolation by distance (P>0.05 for latitude,
324 longitude, and interaction). As in our AFLP study, we found significant genetic differentiation
325 that was correlated to oil exposure: oil exposure explained 7.9% of the variance in DNA
326 sequence (P=0.035, Table 2). These results are recapitulated in the MDS visualization, which
327 shows partial separation of oil-exposed and unexposed sites along MDS axis 1 (Fig. **1**). We
328 found 42 SNPs that were significantly associated with oil exposure, including 6 that overlapped
329 with the *S. alterniflora* transcriptome. Four of these genes were annotated in *O. sativa* (Table
330 S1). After testing for homogeneity of group dispersion, we found no evidence of change in
331 variance in oil-exposed populations (P=0.696).
332
333 *DNA methylation differentiation*
334 Our bisulfite sequencing yielded 1,402,083 cytosines that were polymorphic for
335 methylation across our samples before filtering. After filtering our data to common loci as
336 described above, we analyzed 19,577 polymorphic methylated loci, 5,836 of which occurred in
337 the CG context, 5,069 in the CHG context, and 8,673 in the CHH context. These proportions of
338 polymorphic loci did not change significantly due to oil exposure (Fig. **2**). Methylation calls
339 were collapsed for symmetric CG and CHG loci across "watson" and "crick" strands so that
340 methylation on either one or both strands was considered as a single locus. A quarter of cytosines
341 that showed methylation polymorphism (26.3%) were methylated in fewer than 5 individuals,
342 suggesting that many cytosines were infrequently methylated or methylated only on some sub-
343 genomes. We found significant methylation differentiation by oil exposure after controlling for
344 population structure with principal components of genetic data (P=0.022). Oil explained 9.2% of
345 the variation in methylation.
346 Using locus-by-locus tests, we found 14 differentially methylated positions (DMPs
347 between exposure types (Q<0.05, Fig. **3**; Table S1). These loci occurred nearly equally in CG (4
348 loci), CHG (5 loci), and CHH (5 loci) contexts, and no contexts were overrepresented relative to
349 their distribution across all loci (P>0.2 for all tests). Among these significant loci, one was
350 located within a fragment that mapped to the *S. alterniflora* transcriptome. In addition, four were
351 located in one of three fragments that mapped to the *Oryza sativa* transcriptome (Table S1). Most
352 of these 14 loci had small differences in the magnitude of methylation frequency changes
353 (average 2.8% change between exposed and unexposed populations). However, one locus

354 experienced frequency changes of 30%, and was located in an actin filament depolymerization
355 protein (Os03g56790; ADF2). We found no DMPs that mapped to the first 250 nucleotides of a
356 fragment in exon 1, where DNA methylation may be particularly relevant to the differential
357 expression of genes (Secco *et al.*, 2015).
358
359 *Correlations with gene expression*
360 Although 2,718 epiGBS fragments mapped to the *Spartina* transcriptome, none of these
361 fragments overlapped with transcripts that showed expression differences correlated with
362 exposure to the *DWH* oil spill (Alvarez *et al.*, 2018). Across all genes within a sample, we found
363 no significant relationship between genetic distance with gene expression distance (P=0.185,
364 Mantel's R=0.1489), but there was a correlation between patterns of methylation variation and
365 genome wide gene expression (P=0.009, Mantel's R=0.2319). We had expression data and either
366 SNPs or methylation polymorphisms for only 54 genes, and we found no association between
367 individual SNPs or methylation polymorphisms with expression in *cis*.
368
369 **Discussion**
370 *Spartina alterniflora* displays high levels of genetic and methylation variation across
371 environmental conditions in its native range (Richards *et al.*, 2004; Hughes & Lotterhos, 2014;
372 Foust *et al.,* 2016; Robertson *et al.*, 2017), providing substrate for both genetic and epigenetic
373 response to pollution. We previously found that genetic structure and expression of 3,334 genes
374 were correlated to exposure to the *DWH* oil spill, but genome-wide methylation variation was
375 not (Robertson *et al.,* 2017; Alvarez *et al.*, 2018). Higher resolution epiGBS suggests that both
376 genetic sequence and DNA methylation are correlated with crude oil exposure in *S. alterniflora*,
377 and that changes in gene expression are more strongly correlated to genome-wide patterns of
378 DNA methylation than population structure.
379
380 *Increased resolution of genetic and epigenetic response to the DWH*
381 Our study confirmed significant genetic differentiation between oil-exposed and
382 unexposed sites. This type of genetic structure may reflect either stochastic mortality in oil-
383 exposed areas that may result from a severe bottleneck, or a signature of selection for oil
384 tolerance in affected populations. *Spartina alterniflora* is a highly stress tolerant plant, and

populations have persisted after exposure to the *DWH* oil spill, even after extensive dieback (Lin & Mendelssohn, 2012; Silliman *et al.*, 2012; Lin *et al.*, 2016). However, our genetic data suggest that at least some *S. alterniflora* genotypes were susceptible to crude oil stress, and either had not regrown at the time of sampling or experienced mortality as a result of oil exposure. Our genetic marker studies, combined with initial losses in live aboveground and belowground biomass, support the hypothesis of differential mortality (Lin *et al*., 2016). Although we found no evidence for a reduction in genetic variation, which may indicate selection for tolerant genotypes, the polyploidy of *Spartina* makes accurate quantification of total genetic variation challenging. Further investigations will be required to confirm whether mortality varied by genotype, and whether the *DWH* oil spill was truly a selective event.

The DNA methylation differences that were correlated to oil exposure may reflect either the downstream effects of different genetic variants, a plastic regulatory response to oil exposure, or both. In another study of *S. alterniflora* populations, patterns of DNA methylation were more strongly correlated than genetic structure with microhabitat, and some response of DNA methylation to environment was independent of population structure (Foust *et al.,* 2016). In the oil contaminated populations, we identified a differentially expressed gene that was homologous to the histone methyltransferase SUVH5, which may modulate fitness effects during oil exposure (Alvarez *et al.,* 2018). Histone methylation is intimately linked to DNA methylation through the regulation of *DNA (cytosine-5)-methyltransferase CMT3* activity (Stroud *et al.*, 2013). Therefore, we hypothesized that the differential expression of SUVH5 in response to crude oil exposure would result in differences in DNA methylation variation. With the increase in resolution offered by the epiGBS data set (36,131 epiGBS fragments yielding 19,615 methylation loci compared to 39 polymorphic MS-AFLP fragments), we found a previously unidentified genome-wide signature of methylation differentiation between oil affected and unaffected sites after controlling for population structure.

Epigenetic variation is correlated to genome wide patterns of gene expression

We expected that our fragments would be biased toward coding regions of the genome, and that we would find DNA methylation was correlated with changes in gene function since fragments could overlap with the promoters or the 5’ end of the coding regions of genes. We found that almost 1/3 (10,103 of the 36,131) epiGBS fragments mapped to the *S. alterniflora*

transcriptome. However, many transcripts contained multiple epiGBS fragments such that only 2,718 transcripts were represented. None of these 2,718 genes overlapped with the 3,334 oil responsive genes from our previous transcriptome analysis (Alvarez *et al.*, 2018), and few fragments overlapped the 5' end of genes. Therefore, our expectation that we could correlate local methylation polymorphisms to expression of oil responsive genes was not fulfilled in this study. However, genome-wide patterns of DNA methylation are more strongly correlated than genetic structure with gene expression.

Combined, our results suggest that genetic variation and regulatory mechanisms may play a role in the response of *S. alterniflora* populations to crude oil pollution. Although we are unable to dissect causality, our data suggest one of two possibilities. In the first scenario, crude oil exposure may have generated genetic differentiation between oil exposed and unexposed populations, which in turn shaped patterns of gene expression. Differential transcription may then have generated divergent patterns of methylation, either through genetic variation in DNA methyltransferases (Dubin *et al.*, 2015), or through changes in gene expression affected by neighboring transposable element activity (*sensu* Secco *et al.*, 2015). In this scenario, DNA methylation patterns are generated by genetic variation and/or reflect expression variation (*sensu* Meng *et al.,* 2016; Niederhuth & Schmitz, 2017). In the second scenario, crude oil exposure may have affected genetic variation while simultaneously but independently inducing patterns of DNA methylation e.g. as a response to the genomic "shock" of crude oil exposure (Rapp & Wendel, 2005; Richards *et al.*, 2012b). This variation may modulate the plastic response of gene expression variation in this system, and reflect an induced regulatory shift in *S. alterniflora*, regardless of genetic background. Of course, these scenarios are not mutually exclusive. While more complex models of causation will require substantially more genomic resources than are currently available in *S. alterniflora*, reverse genetic screens and higher resolution surveys may help to discern the relative contributions and causal roles of methylation and genetic variation in genes of interest.

Reduced representation sequencing compared to AFLP

As the field of ecological genomics matures, there is a pressing need to develop robust assays and statistically sound measures of population regulatory variation. Reduced representation methylation sequencing techniques are attractive for ecological genomics because

447 they can be used to infer genome wide patterns of both genetic and methylation variation without
448 a high-quality reference genome (Robertson & Richards, 2015; Richards *et al.*, 2017; Paun *et al.*,
449 2018). However, considering the small portion of the genome that is sampled in these
450 techniques, they still have serious limitations particularly for species that do not yet have a fully
451 sequenced reference genome (Paun *et al.*, 2018).

452 We assessed the utility of epiGBS compared to MS-AFLP, and expected that the
453 substantial increase in informative markers would confirm our results from our MS-AFLP study
454 and lend greater resolution to detect patterns of DNA methylation variation. Our epiGBS survey
455 detected significant differentiation in both genetic variation and DNA methylation that was
456 correlated to oil exposure, suggesting that epiGBS provides increased resolution over MS-AFLP
457 to detect genome-wide methylation differences in studies of similar sample sizes. However,
458 despite the much larger data set generated by epiGBS compared to MS-AFLP, we only found 14
459 differentially methylated positions, which is not substantial improvement over the five loci
460 detected by MS-AFLP (Robertson *et al.*, 2017). The comparable number of DMPs and lack of
461 power to evaluate *cis* association between methylation, genetic variation, and gene expression
462 suggests that the ability to characterize individual functional elements remains obscured by
463 substantial variation in methylation across individuals, and is limited by the small fraction (an
464 estimated 0.3%) of the genome that is assayed. Further, although we may have biased sampling
465 to coding regions of the genome, we found few fragments that overlapped the functionally
466 relevant 5' end of transcripts. Without a reference genome for *S. alterniflora,* we were unable to
467 identify fragments that overlapped promoter regions.

468 Sequencing-based techniques provide the potential to identify functional genomic
469 regions, but correct annotations rely on genomic resources in a relevant reference. In polyploid
470 species like *S. alterniflora*, a greater number of paralogs and the potential for
471 neofunctionalization among them creates additional uncertainty for using annotations from other
472 species (Primmer *et al.*, 2013). *Spartina alterniflora* has various levels of duplicated gene
473 retention and homeologous expression (Ainouche *et al.*, 2003; Fortune *et al.*, 2007; Ferreira de
474 Carvalho, 2013, 2017; Boutte *et al.*, 2016), which may result in more opportunities for gene
475 diversification and subfunctionalization (Chen *et al.*, 2015; Shimizu-Inatsugi *et al.* 2017), and
476 complicates gene annotation. Therefore, while studies with RRBS techniques in non-model
477 plants offer increased power to detect broad, genome-wide patterns of variation that may be

478 correlated to ecology, they are still limited for the detection of specific gene function. Future
479 studies will benefit from optimizing protocols that enrich for specific portions of the genome, but
480 generating a draft reference genome will be imperative to locate the promoter regions and allow
481 for better exploitation of RRBS data (Paun *et al.*, 2018).
482
483 **Acknowledgments**
484 We thank Steve Pennings, Brittney DeLoach McCall, Aaron Schrey, Christina Meals,
485 and Ashley Shayter for access to field sites and assistance with plant sampling. We thank Acer
486 Van Wallendael, Kieran Samuk, and Kate Ostevik for constructive feedback. This work was
487 supported by funding from the National Science Foundation (U.S.A.) DEB-1419960 and IOS-
488 1556820 (to CLR) and through the Global Invasions Network Research Exchange (Grant No.
489 0541673 for MR), the Franco-American Fulbright Commission (to CLR), and the Netherlands
490 Organisation for Scientific Research (NWO-ALW No. 820.01.025 to KJFV).

491 **Tables and Figures**

492

493 **Table 1** Sampling information across all sites after filtering.

Site Location	Site Code	Exposure	No. of individuals
Grand Isle, LA	GIN1	None	7
Grand Isle, LA	GIN2	None	7
Barataria Bay, LA	GIO1	Heavily Oiled	8
Barataria Bay, LA	GIO2	Heavily Oiled	8
Bay St. Louis, MS	MSN	None	7
Bay St. Louis, MS	MSO	Heavily Oiled	7

494

495 Site information includes location and whether oil was observed at each site (exposure).

Table 2 Association between oil exposure, genetic distance, and methylation distance across tests.

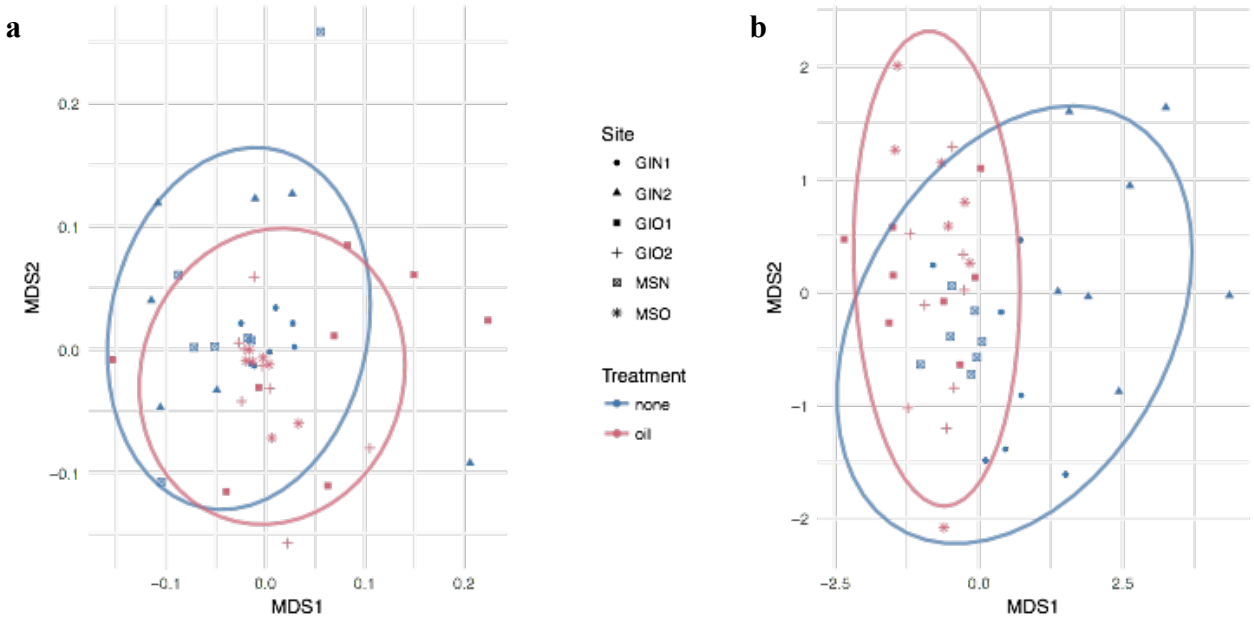
		Genetic		Methylation	
Factor	df	F	Variance explained	F	Variance explained
oil	1	1.395 *	0.079	1.960*	0.092

Test statistics and significance, determined through RDA. For clarity, we do not show significance for estimated surrogate variables (7 for methylation data, 16 for genetic data) or the 8 principal components of allele frequencies for the methylation data. P * ≤ 0.05, P ** ≤ 0.01, P *** ≤ 0.0001

503 **Table 3** Pairwise Fst among three oil contaminated and three uncontaminated sites. Bold (i.e. all
504 entries) indicates significance at P<0.001.

	GIN1	GIN2	MSN	GIO1	GIO2
GIN2	**0.1304**				
MSN	**0.1329**	**0.1421**			
GIO1	**0.1451**	**0.1264**	**0.1382**		
GIO2	**0.0966**	**0.1020**	**0.1198**	**0.1158**	
MSO	**0.1171**	**0.1471**	**0.1497**	**0.1580**	**0.1008**

505



506 **Figure 1** Multidimensional scaling (MDS) of **a**) genetic distance matrix and **b**) DNA
507 methylation distance matrix. Circles represent 95% confidence intervals around the centroid of
508 each group: no oil exposure ("none") or oil-exposed ("oil").
509

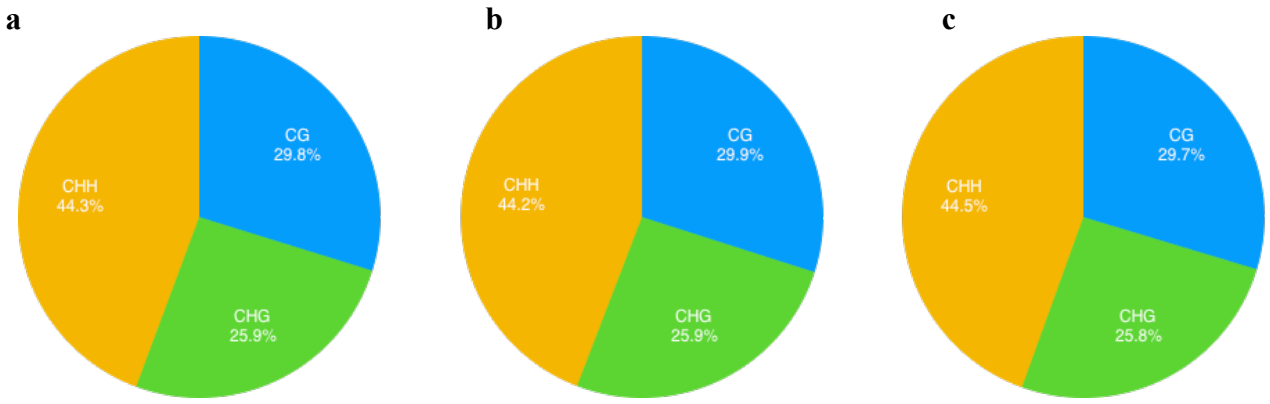
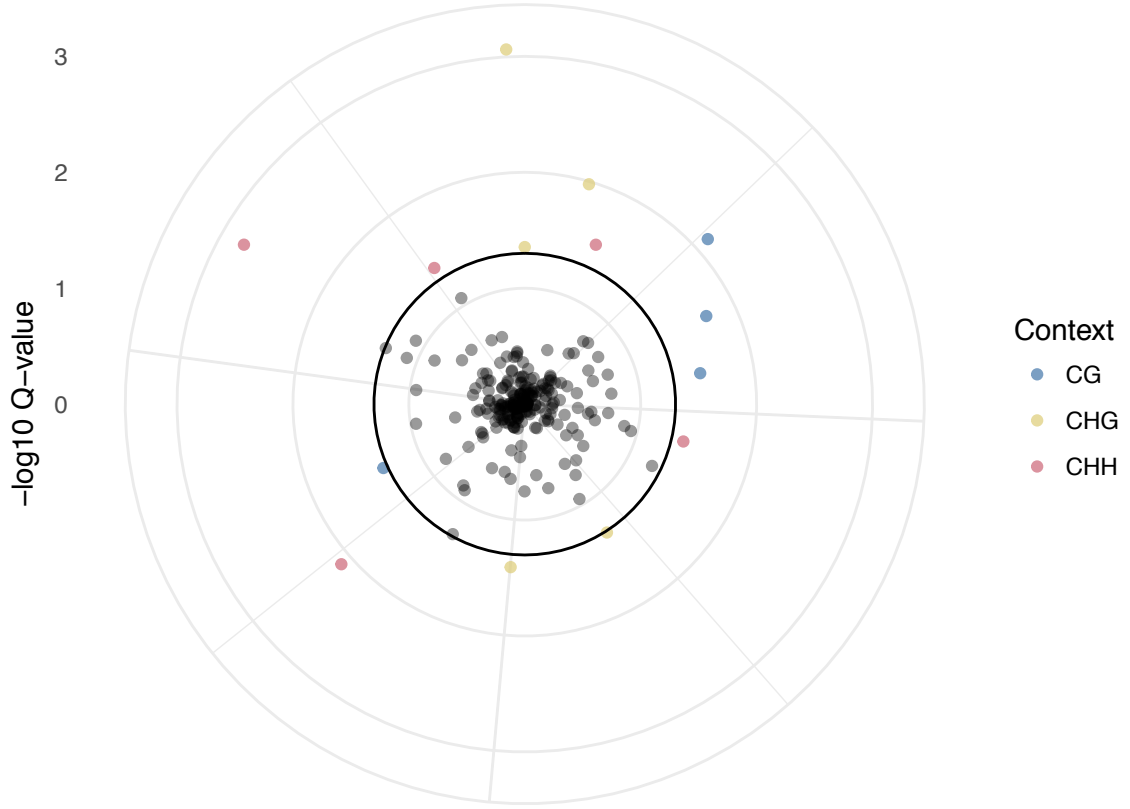


Figure 2 Proportion of methylated sites by context across a) all samples, b) in uncontaminated sites and c) oil-contaminated sites.

515



516

517

518 **Figure 3** Circular manhattan plot showing 14 significant outlier methylation loci, as determined
519 by Q-value. Only significant loci (Q<0.05) are colored, where colors represent trinucleotide
520 context. Plot is shown as circular to accommodate unknown positions along a chromosome, as *S.*
521 *alterniflora* does not have a reference genome.

522

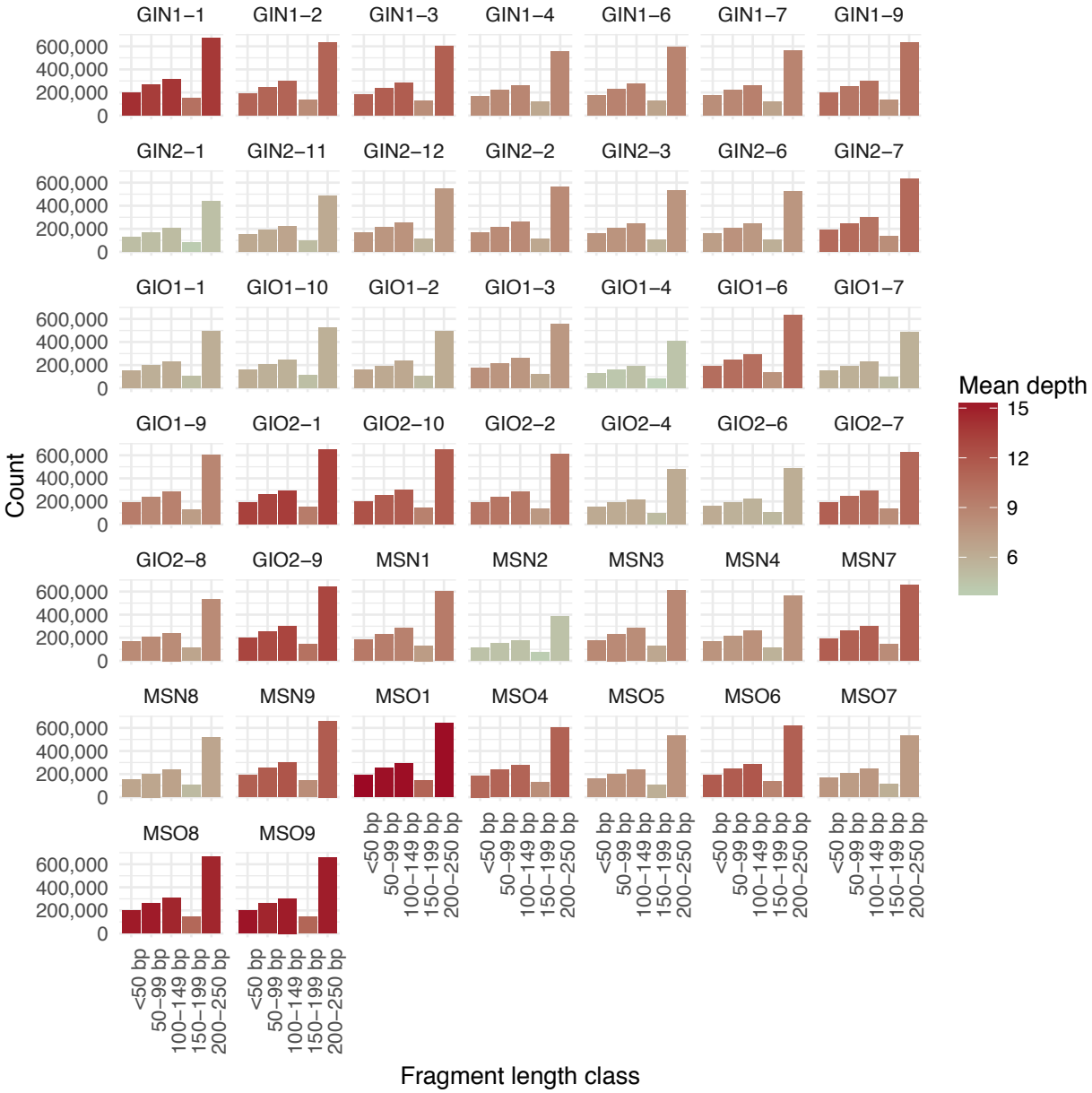
523 **Table S1** Genetic and methylation loci significantly differentially associated with oil exposure.

Contiguous fragment	Position	Context	Q-value	*O. sativa* annotation	Gene name
3386	141	-	0.0056		
5449	67	-	0.0399		
7242	26	-	0.0447		
8158	95	-	0.016		
9488	45	-	0.0339		
9488	51	-	0.0421		
9488	94	-	0.0385		
10931	43	-	0.0385		
11792	111	-	0.0339		
11939	123	-	0.0033	Os06g02570	Syntaxin, putative
11972	49	-	0.0327		
11974	102	-	0.012		
11974	131	-	0.0327		
12084	26	-	0.0447		
12418	86	-	0.0033		
13976	47	-	0.0385		
14656	84	-	0.0385		
14804	93	-	0.0001		
14957	105	-	0.012		
19335	34	-	0.0032		
19759	76	-	0.0447		
26805	111	-	0.0466		
26906	41	-	0.0009		
27016	91	-	0.0421	Os03g18220	pyruvate decarboxylase isozyme 2, putative
27227	103	-	0.0307		
28167	59	-	0.0443		
28937	126	-	0.0421		
28937	133	-	0.0413		
28937	147	-	0.0307		
30433	105	-	0.013		
31126	85	-	0.0181		
31126	147	-	0.0174		
31604	78	-	0.0421		
32215	52	-	0.0033		
33738	102	-	0.0218		
34288	53	-	0.0001	Os11g39230	retrotransposon protein, putative
34398	88	-	0.0013		
34744	123	-	0.0366		

35646	61	-	0	Os01g43070	psbP-related thylakoid lumenal protein 4, chloroplast precursor, putative
36084	92	-	0.0009		
36084	96	-	0.0447		
36084	115	-	0.0308		
673	43	CG	0.0461		
8098	139	CG	0.0291		
11959	13	CG	0.0182	Os01g69980	TCP family transcription factor, putative
34711	123	CG	0.0075		
183	10	CHG	0.0443		
527	19	CHG	0.0388	Os03g56790	actin-depolymerizing factor, putative
4191	107	CHG	0.0105		
5540	171	CHG	0.0482		
7334	97	CHG	0.0009		
45	82	CHH	0.0388	Os03g51920	peptidase, M50 family, putative
5540	34	CHH	0.008		
6013	69	CHH	0.0016	Os10g34480	cytochrome P450, putative
14227	11	CHH	0.0312	Os08g38560	receptor-like protein kinase 2 precursor, putative
14776	15	CHH	0.0394		

Only loci with Q-value <0.05 are shown. Trinucleotide context (CG, CHG, CHH) shown for methylated loci. Gene numbers from *O. sativa* are shown for loci on fragments that were able to annotated.

528



529

530

531 **Figure S1** Depth of sampling by fragment size in each sample across all fragments. Fragments
532 ranged from 19-202 basepairs (bp) for a total length of 5,441,437 bp (approximately 0.3% of the
533 2117 Mbp *S. alterniflora* genome (2C= 4.33pg); Fortune *et al.,* 2008). Fragments shown in
534 classes on the X axis.

535

536

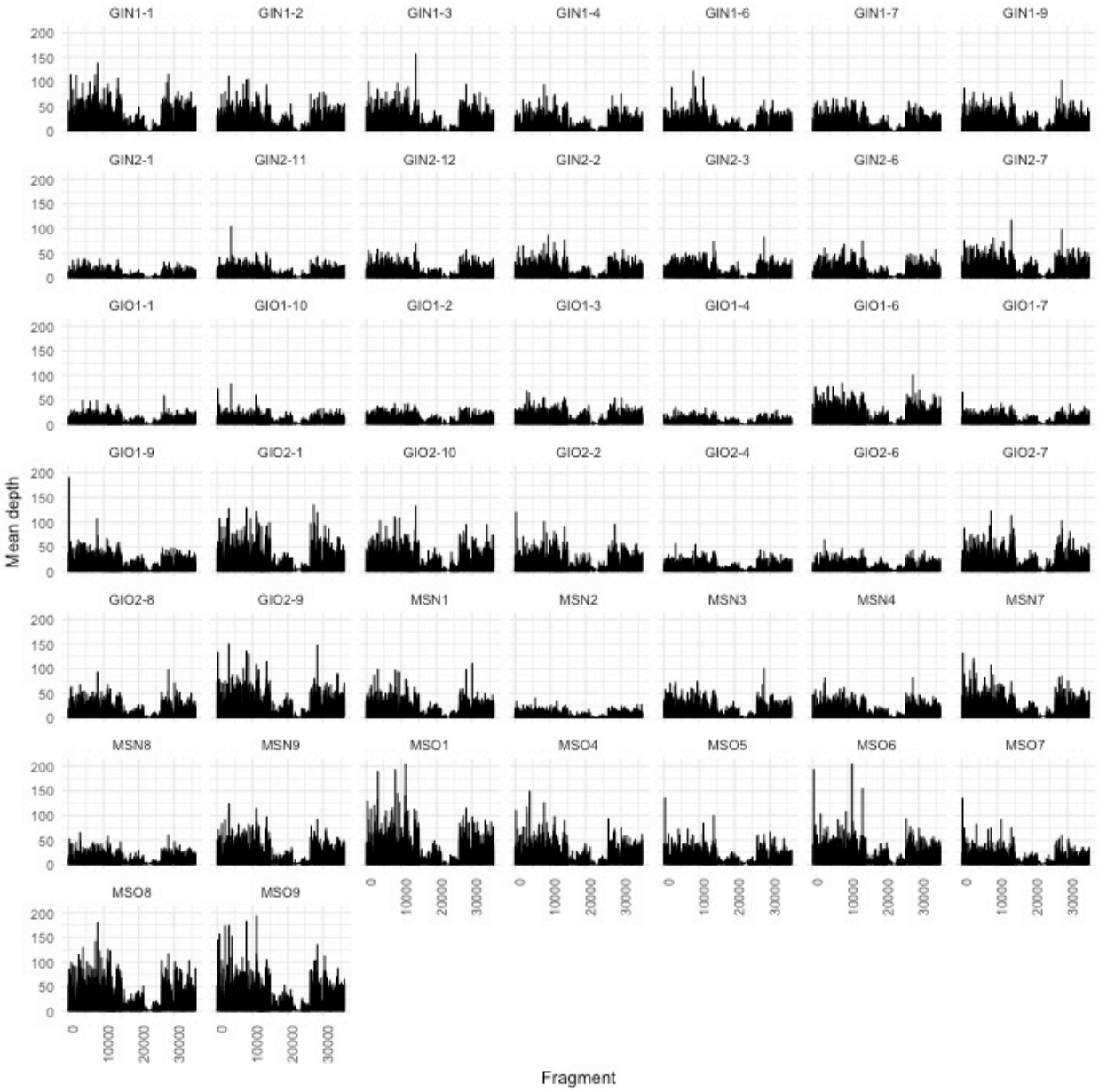


Figure S2 Sequencing depth for each sample for each of the 36,131 fragments.

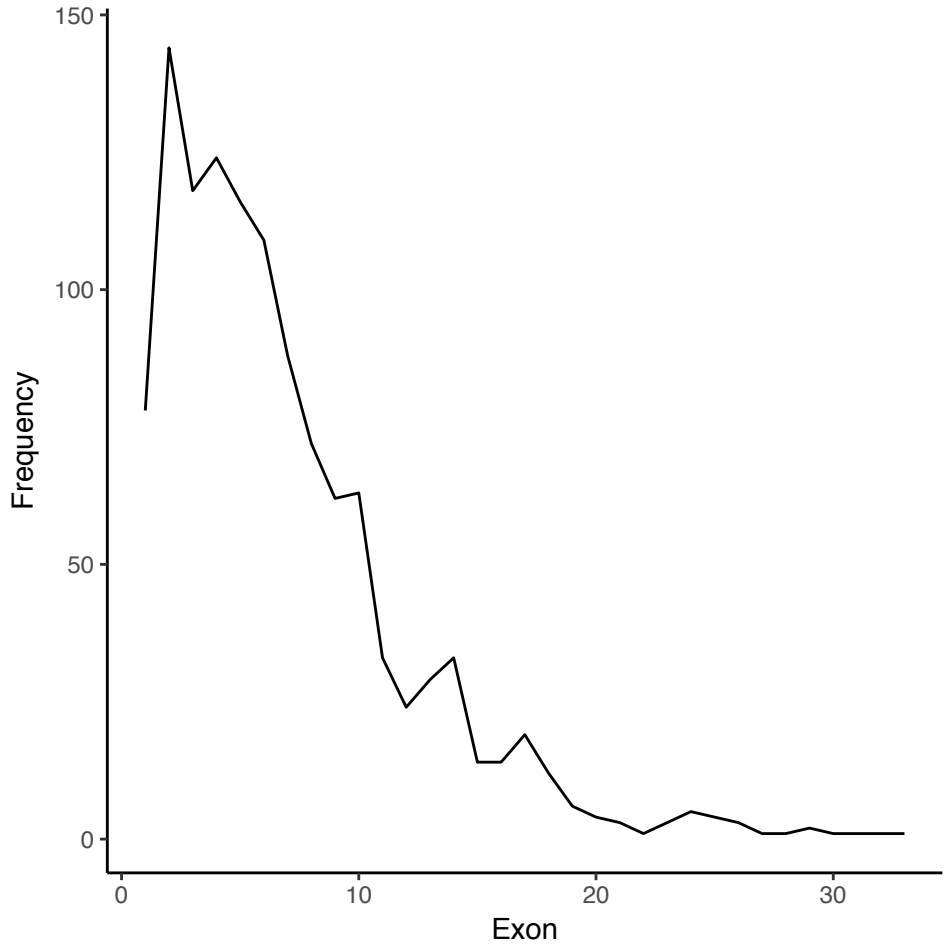


Figure S3. Distribution of epiGBS hits across exons. Exon order is predicted by BLAST hits to *Oryza sativa*.

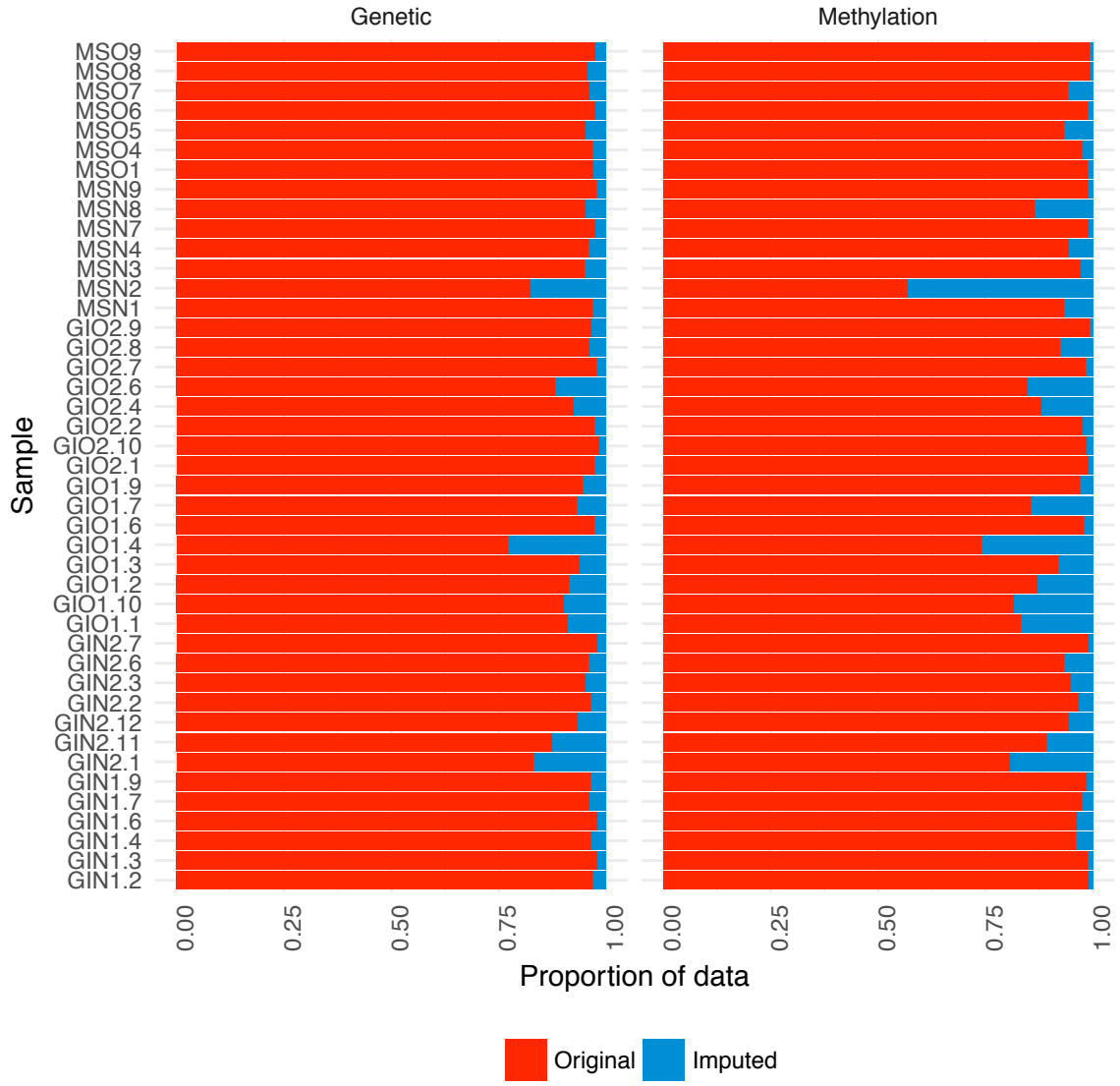


Figure S4. Missing and imputed data by sample after filtering. Genetic and methylation data were both filtered to include only loci present in >90% of samples and samples with less than 50% missing data.

549 **References**

550

551 **Ainouche ML, Baumel A, Salmon A, Yannic G. 2003.** Hybridization, polyploidy and
552 speciation in Spartina (Poaceae). *New Phytologist* **161**: 165-172.

553

554 **Aïnouche ML, Fortune PM, Salmon A, Parisod C, Grandbastien MA, Fukunaga K, Ricou**
555 **M, Misset MT. 2009.** Hybridization, polyploidy and invasion: lessons from Spartina (Poaceae).
556 *Biological invasions* **11**: 1159. doi: 10.1007/s10530-008-9383-2

557

558 **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997.**
559 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.
560 *Nucleic Acids Research* **25**: 3389-3402.

561

562 **Alvarez M, Ferreira de Carvalho J, Salmon A, Ainouche ML, Cavé-Radet A, El Amrani A,**
563 **Foster TE, Moyer S, Richards CL. 2018.** Transcriptome response to the *Deepwater Horizon*
564 oil spill identifies novel candidate genes for oil tolerance in natural populations of the foundation
565 plant *Spartina alterniflora*. *Molecular Ecology* doi: 10.1111/mec.14736

566

567 **Alvarez M, Schrey AW, Richards CL. 2015.** Ten years of transcriptomics in wild populations:
568 what have we learned about their ecology and evolution?. *Molecular Ecology* **24**: 710-25. doi:
569 10.1111/mec.13055

570

571 **Andrew RL, Bernatchez L, Bonin A, Buerkle CA, Carstens BC, Emerson BC, Garant D,**
572 **Giraud T, Kane NC, Rogers SM *et al*. 2013.** A road map for molecular ecology. *Molecular*
573 *Ecology* **22**: 2605-2626. doi: 10.1111/mec.12319

574

575 **Baisakh N, Subudhi PK. 2009.** Heat stress alters the expression of salt stress induced genes in
576 smooth cordgrass (*Spartina alterniflora* L.). *Plant Physiology and Biochemistry:PPB* **47**: 232–
577 235. doi: 10.1016/j.plaphy.2008.11.010

578

579 **Baisakh N, Subudhi PK, Varadwaj P. 2008.** Primary responses to salt stress in a halophyte,
580 smooth cordgrass (*Spartina alterniflora* Loisel.). *Functional & Integrative Genomics* **8**: 287–
581 300. doi: 10.1007/s10142-008-0075-x
582
583 **Becker C, Weigel D. 2012.** Epigenetic variation: origin and transgenerational
584 inheritance. *Current Opinion in Plant Biology* **15**: 562-567.
585
586 **Bedre R, Mangu VR, Srivastava S, Sanchez LE, Baisakh N. 2016.** Transcriptome
587 analysis of smooth cordgrass (*Spartina alterniflora* Loisel), a monocot halophyte, reveals
588 candidate genes involved in its adaptation to salinity. *BMC Genomics* **17**: 657. doi:
589 10.1186/s12864-016-3017-3
590
591 **Boutte J, Ferreira de Carvalho J, Rousseau-Gueutin M, Poulain J, Da Silva C, Wincker P,**
592 **Ainouche M, Salmon A. 2016.** Reference Transcriptomes and Detection of Duplicated Copies
593 in Hexaploid and Allododecaploid Spartina Species (Poaceae). *Genome Biology and Evolution* **8**:
594 3030–3044. doi: 10.1093/gbe/evw209
595
596 **Chen X, Ge X, Wang J, Tan C, King GJ, Liu K. 2015.** Genome-wide DNA methylation
597 profiling by modified reduced representation bisulfite sequencing in Brassica rapa suggests that
598 epigenetic modifications play a key role in polyploid genome evolution. *Frontiers in Plant*
599 *Science* **6**: 836.
600
601 **Day JW, Boesch DF, Clairain EJ, Kemp GP, Laska SB, Mitsch WJ, Orth K, Mashriqui H,**
602 **Reed DJ, Shabman L *et al.* 2007.** Restoration of the Mississippi Delta: lessons from hurricanes
603 Katrina and Rita. *Science* **23**: 1679-84.
604
605 **DeLeo DM, Herrera S, Lengyela SD, Quattrinia AM, Kulathinala RJ, Cordes EE. *In press.***
606 Gene expression profiling reveals deep-sea coral response to the *Deepwater Horizon* oil spill.
607 *Molecular Ecology* doi: 10.1111/mec.14847.
608

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, Del Angel G, Rivas MA, Hanna M *et al.* 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**: 491-498.

Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**: 519-532. doi: 10.1038/nrm4043

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Casale FP, 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.

Dufresne F, Stift M, Vergilino R, Mable BK. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state□of□the□art molecular and statistical tools. *Molecular Ecology* **23**: 40-69.

Feder ME, Mitchell-Olds T. 2003. Evolutionary and ecological functional genomics. *Nature Reviews Genetics* **4**: 649.

Ferreira de Carvalho J, Poulain J, Da Silva C, Wincker P, Michon-Coudouel S, Dheilly A, Naquin D, Boutte J, Salmon A, Ainouche M. 2013. Transcriptome de novo assembly from next-generation sequencing and comparative analyses in the hexaploid salt marsh species Spartina maritima and Spartina alterniflora (Poaceae). *Heredity* **110**: 181-193.

Ferreira de Carvalho J, Boutte J, Bourdaud P, Chelaifa H, Ainouche K, Salmon A, Ainouche M. 2017. Gene expression variation in natural populations of hexaploid and allododecaploid Spartina species (Poaceae). *Plant Systematics and Evolution* **303**: 1061-1079.

Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR. 2016. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. Molecular ecology, 25(1), 104-120.

Fortune PM, Schierenbeck KA, Ainouche AK, Jacquemin J, Wendel JF, Ainouche ML. 2007. Evolutionary dynamics of Waxy and the origin of hexaploid Spartina species (Poaceae). *Molecular Phylogenetics and Evolution* **43**: 1040–1055. doi: 10.1016/j.ympev.2006.11.018

Foust CM, Preite V, Schrey AW, Alvarez M, Robertson MH, Verhoeven KJF, Richards CL. 2016. Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. *Molecular Ecology* **25**: 1639-1652.

Goslee SC, Urban DL. 2007. The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* **22**: 1-19.

Hastie T, Tibshirani R, Narasimhan B, Chu G. 2018. impute: Imputation for microarray data. R package version 1.54.0. doi: 10.18129/B9.bioc.impute

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, et al. 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* 330: 204–208. doi: 10.1126/science.1195979

Hughes AR, Inouye BD, Johnson MT, Underwood N, Vellend M. 2008. Ecological consequences of genetic diversity. *Ecology Letters* **11**: 609-23.

Hughes AR, Lotterhos KE. 2014. Genotypic diversity at multiple spatial scales in the foundation marsh species, *Spartina alterniflora*. *Marine Ecology Progress Series* **497**: 105–117. doi: 10.3354/meps10565

Jackson SA. 2017. Epigenomics: dissecting hybridization and polyploidization. *Genome Biology* **18**: 117.

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S *et al.* 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* **6**: 4.

Kimes NE, Callaghan AV, Aktas DF, Smith WL, Sunner J, Golding BT, et al. 2013. Metagenomic analysis and metabolite profiling of deep-sea sediments from the Gulf of Mexico following the *Deepwater Horizon* oil spill. *Front. Microbiol*. 4:50. doi: 10.3389/fmicb.2013.00050

Kimes NE, Callaghan AV, Suflita JM, Morris PJ. 2014. Microbial transformation of the *Deepwater Horizon* oil spill—past, present, and future perspectives. *Front. Microbiol*. 5:603. doi: 10.3389/fmicb.2014.00603.

Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE, Hyma K, Acharya C, Mitchell SE *et al.* 2015. Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances* **6**: e1400218. doi: 10.1126/sciadv.1400218

Lea AJ, Tung J, Zhou X. 2015. A flexible, efficient binomial mixed model for identifying differential DNA methylation in bisulfite sequencing data. *PLoS Genetics* **11**: e1005650.

Leek J, McShane BB, Gelman A, Colquhoun D, Nuijten MB, Goodman SN. 2017. Five ways to fix statistics. *Nature* **551**: 557-559.

Leek JT, Storey JD. 2007. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genetics* **3**: e161.

Liebl AL, Schrey AW, Richards CL, Martin LB. 2013. Patterns of DNA methylation throughout a range expansion of an introduced songbird. *Integrative and Comparative Biology* **53**: 351-358. doi: 10.1093/icb/ict007

Lin Q, Mendelssohn IA. 2012. Impacts and recovery of the Deepwater Horizon oil spill on vegetation structure and function of coastal salt marshes in the northern Gulf of Mexico. *Environmental Science & Technology* **46**: 3737-43.

Lin Q, Mendelssohn IA, Graham SA, Hou A, Fleeger JW, Deis DR. 2016. Response of salt marshes to oiling from the Deepwater Horizon spill: Implications for plant growth, soil surface-erosion, and shoreline stability. *Science of the Total Environment* **1:** 369-77.

Maor GL, Yearim A, Ast G. 2015. The alternative role of DNA methylation in splicing regulation. *Trends in Genetics* **31**: 274-80.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**: 394-408.

Meirmans, P. G. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular ecology*, **24**: 3223-3231.

Mendelssohn IA, Andersen GL, Baltz DM, Caffey RH, Carman KR, Fleeger JW, Joye SB, Lin Q, Maltby E, Overton EB, Rozas LP. 2012. Oil impacts on coastal wetlands: implications for the Mississippi River Delta ecosystem after the Deepwater Horizon oil spill. *BioScience* **62**: 562-74.

Meng D, Dubin M, Zhang P, Osborne EJ, Stegle O, Clark RM, Nordborg M. 2016. Limited contribution of DNA methylation variation to expression regulation in Arabidopsis thaliana. *PLoS Genetics* **12**: e1006141.

Michel J, Owens EH, Zengel S, Graham A, Nixon Z, Allard T, Holton W, Reimer PD, Lamarche A, White M *et al.* 2013. Extent and degree of shoreline oiling: Deepwater Horizon oil spill, Gulf of Mexico, USA. *PloS One* **12**: e65087.

National Commission on the BP Deepwater Horizon Oil Spill. 2011. Graham B, Reilly WK, Beinecke F, Boesch DF, Garcia TD, Murray CA, Ulmer F, eds. *Deep Water: The Gulf Oil Disaster and the Future of Offshore Drilling. Report to the President.* doi: 10.1111/jols.12003/full

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F *et al.* 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**: 684-92.

Niederhuth CE*, Bewick AJ*, Ji L, Alabday M, 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.

Niederhuth CE, Schmitz RJ. 2017. Putting methylation in context: from genomes to gene expression in plants. *Biochemica et Biophysica Acta* **1**: 149–156. doi: 10.1016/j.bbagrm.2016.08.009

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. 2013. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* **29**: 792-793.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P *et al.* 2017. vegan: Community Ecology Package. R package version 2.4-4. https://CRAN.R-project.org/package=vegan

Paun O, Verhoeven KJF, Richards CL. 2018. Opportunities and limitations of reduced representation bisulphite sequencing in plant ecological epigenomics. *New Phytologist* **doi**: 10.1111/nph.15388

Pedersen BS, Eyring K, De S, Yang IV, Schwartz DA. 2014. Fast and accurate alignment of long bisulfite-seq reads. *arXiv preprint* arXiv:1401.1129.

Pembleton LW, Cogan NO, Forster JW. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**: 946-952.

764 **Pennings SC, Bertness MD. 2001.** Salt Marsh Communities. In: Bertness MD, Gaines SD, Hay
765 M, eds. *Marine Community Ecology*. Sunderland, Massachusetts: Sinauer Associates, 289-316.
766
767
768 **Primmer CR, Papakostas S, Leder EH, Davis MJ, Ragan, MA. 2013.** Annotated
769 genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and
770 evolution research. *Molecular Ecology* **22**: 3216–3241. doi: 10.1111/mec.12309
771
772 **Rapp RA, Wendel JF. 2005.** Epigenetics and plant evolution. *New Phytologist*. **168**: 81-91.
773
774 **Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colomé-Tatché M, Durka W,**
775 **Engelhardt J, Gaspar B, Gogol-Doring A *et al.* 2017.** Ecological plant epigenetics: Evidence
776 from model and non-model species, and the way forward. *Ecology Letters* **20**: 1576-1590.
777
778 **Richards CL, Hamrick JL, Donovan LA, Mauricio R. 2004.** Unexpectedly high clonal
779 diversity of two salt marsh perennials across a severe environmental gradient. *Ecology Letters* **7**:
780 1155–1162. doi: 10.1111/j.1461-0248.2004.00674.x
781
782 **Richards CL, Schrey AW, Pigliucci M. 2012a.** Invasion of diverse habitats by few Japanese
783 knotweed genotypes is correlated with high epigenetic differentiation. *Ecology Letters* **15**: 1016-
784 1025.
785
786 **Richards CL, Verhoeven KJF, Bossdorf O. 2012b.** Evolutionary significance of epigenetic
787 variation. In: Wendel JF, Greilhuber J, Dolezel J, Leitch IJ, eds. *Plant Genome Diversity*.
788 Springer, 257-274.
789
790 **Robertson M, Richards C. 2015.** Opportunities and challenges of next□generation sequencing
791 applications in ecological epigenetics. *Molecular Ecology* **24**: 3799-3801.
792

Robertson M, Schrey A, Shayter A, Moss CJ, Richards, C. 2017. Genetic and epigenetic variation in Spartina alterniflora following the Deepwater Horizon oil spill. *Evolutionary Applications* **10**: 792-801.

Schoener TW. 2011. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science* **331**: 426-429.

Schrey AW, Alvarez M, Foust CM, Kilvitis HJ, Lee JD, Liebl AL, Martin LB, Richards CL, Robertson M. 2013. Ecological Epigenetics: Beyond MS-AFLP. *Integrative and Comparative Biology* **53**: 340-50.

Secco D, Wang C, Shou H, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R. 2015. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *Elife* **4**: e09343

Shimizu-Inatsugi R, Terada A, Hirose K, Kudoh H, Sese J, Shimizu KK. 2017. Plant adaptive radiation mediated by polyploid plasticity in transcriptomes. *Molecular Ecology* **26**: 193-207. doi: 10.1111/mec.13738

Silliman BR, van de Koppel J, McCoy MW, Diller J, Kasozi GN, Earl K, Adams PN, Zimmerman AR. 2012. Degradation and resilience in Louisiana salt marshes after the BP-Deepwater Horizon oil spill. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 11234–11239. doi: 10.1073/pnas.1204922109

Storey JD, Bass AJ, Dabney A, Robinson D. 2015. qvalue: Q-value estimation for false discovery rate control. R package version 2.6.0. http://github.com/jdstorey/qvalue

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**: 352-64.

824 **Travis SE, Hester MW. 2005.** A space-for-time substitution reveals the long-term decline in
825 genotypic diversity of a widespread salt marsh plant, *Spartina alterniflora,* over a span of 1500
826 years. *Journal of Ecology* **93**: 417-430. doi: 10.1111/j.0022-0477.2005.00985.x
827
828 **Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E,**
829 **Jacob Y, Henderson IR *et al.* 2018.** Epigenetic activation of meiotic recombination near
830 *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome*
831 *Research* **28**: 1-13.
832
833 **van Gurp TP, Wagemaker NC, Wouters B, Vergeer P, Ouborg JN, Verhoeven KJ. 2016.**
834 epiGBS: reference-free reduced representation bisulfite sequencing. *Nature Methods* **13**: 322-
835 324.
836
837 **Verhoeven KJ, Jansen JJ, van Dijk PJ, Biere A. 2010.** Stress-induced DNA methylation
838 changes and their heritability in asexual dandelions. *New Phytologist* **185**: 1108-1018.
839
840 **Wibowo AT, Becker C, Marconi G, Durr J, Price J, Hagmann J. et al. 2016.** Hyperosmotic
841 stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome
842 and is restricted in the male germline by DNA glycosylase activity. *eLife* **5**: e13546.
843
844 **Whitehead A, Dubansky B, Bodinier C, Garcia TI, Miles S, Pilley C, Walter RB. 2012.**
845 Genomic and physiological footprint of the Deepwater Horizon oil spill on resident marsh fishes.
846 *Proceedings of the National Academy of Sciences* **109**: 20298–20302.
847
848 **Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, LeRoy CJ, Lonsdorf**
849 **EV, Allan GJ, DiFazio SP, Potts BM *et al.* 2006.** A framework for community and ecosystem
850 genetics: from genes to ecosystems. *Nature Reviews Genetics* **7**: 510.
851
852 **Xie HJ*, Li H*, Liu D, Dai WM, He JY, Lin S, Duan H, Liu LL, Chen SG, Song XL *et al.***
853 **2015.** *ICE1* demethylation drives the range expansion of a plant invader through cold tolerance
854 divergence. *Molecular Ecology* **24**: 835-850. doi: 10.1111/mec.13067

Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. 2015. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes & Development* **29**: 2183-2202.