



OPEN

## Gene body DNA methylation in seagrasses: inter- and intraspecific differences and interaction with transcriptome plasticity under heat stress

Laura Entrambasaguas<sup>1</sup>, Miriam Ruocco<sup>1</sup>, Koen J. F. Verhoeven<sup>2</sup>, Gabriele Procaccini<sup>1,4</sup>✉ & Lazaro Marín-Guirao<sup>1,3,4</sup>

The role of DNA methylation and its interaction with gene expression and transcriptome plasticity is poorly understood, and current insight comes mainly from studies in very few model plant species. Here, we study gene body DNA methylation (gbM) and gene expression patterns in ecotypes from contrasting thermal environments of two marine plants with contrasting life history strategies in order to explore the potential role epigenetic mechanisms could play in gene plasticity and responsiveness to heat stress. In silico transcriptome analysis of CpG<sub>O/E</sub> ratios suggested that the bulk of *Posidonia oceanica* and *Cymodocea nodosa* genes possess high levels of intragenic methylation. We also observed a correlation between gbM and gene expression flexibility: genes with low DNA methylation tend to show flexible gene expression and plasticity under changing conditions. Furthermore, the empirical determination of global DNA methylation (5-mC) showed patterns of intra and inter-specific divergence that suggests a link between methylation level and the plants' latitude of origin and life history. Although we cannot discern whether gbM regulates gene expression or vice versa, or if other molecular mechanisms play a role in facilitating transcriptome responsiveness, our findings point to the existence of a relationship between gene responsiveness and gbM patterns in marine plants.

Human-induced climate change is challenging the persistence of natural ecosystems. The rapid environmental changes that organisms are currently facing, are forcing them to acclimatize, adapt or migrate to avoid becoming extinct. Unlike slow genetic adaptation of populations through natural selection, phenotypic plasticity of individual genotypes facilitates a fast response to environmental change and it is currently considered key for the ecological success of organisms in a rapidly changing world<sup>1–3</sup>. An important component of the ability of an organism to change its phenotype is the modulation of gene expression (e.g.<sup>4,5</sup>). The mechanisms by which gene expression, and thus phenotypic plasticity, are influenced and regulated remain poorly understood in most organisms. Epigenetic changes (e.g. DNA methylation and histone modifications) have been suggested as key candidates for regulating gene expression and generating/regulating phenotypic plasticity<sup>6–10</sup> and adaptive responses to environmental change (e.g.<sup>11</sup>).

Cytosine methylation within DNA is an evolutionary widespread epigenetic modification found in most eukaryotes, including plants and animals<sup>12</sup>. In plants, DNA methylation is found at cytosines (C) in three sequence contexts: CG, CHG, and CHH, where H corresponds to C, T, or A, and it is established and maintained by different pathways<sup>13</sup>. Methylated CG (mCG) is the most abundant form of mC in plant genomes and predominantly occurs in noncoding regions such as transposable elements and other repetitive DNA regions<sup>14,15</sup>. Protein coding regions are also methylated but its establishment, maintenance, possible function and evolutionary consequences remain unclear. Gene body DNA methylation (gbM) refers to genes with enriched mCG within the transcribed regions and depletion at the transcriptional start and termination sites<sup>16,17</sup>. GbM is a common

<sup>1</sup>Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli, Italy. <sup>2</sup>Terrestrial Ecology Department, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands. <sup>3</sup>Seagrass Ecology Group, Oceanographic Center of Murcia, Spanish Institute of Oceanography, C/Varadero, 30740 San Pedro del Pinatar, Spain. <sup>4</sup>The authors equally contributed: Gabriele Procaccini, Lazaro Marín-Guirao ✉email: gpro@szn.it

feature of eukaryotic genomes and remains well conserved among plants and animals<sup>13,18</sup>. Furthermore, gbM genes are often long (bp), slowly evolving and evolutionary conserved when compared to genes with low gbM; in contrast to CHG or CHH methylation, CG methylation is the only correlated across orthologues although the cause of this correlation remains unknown<sup>19–21</sup>.

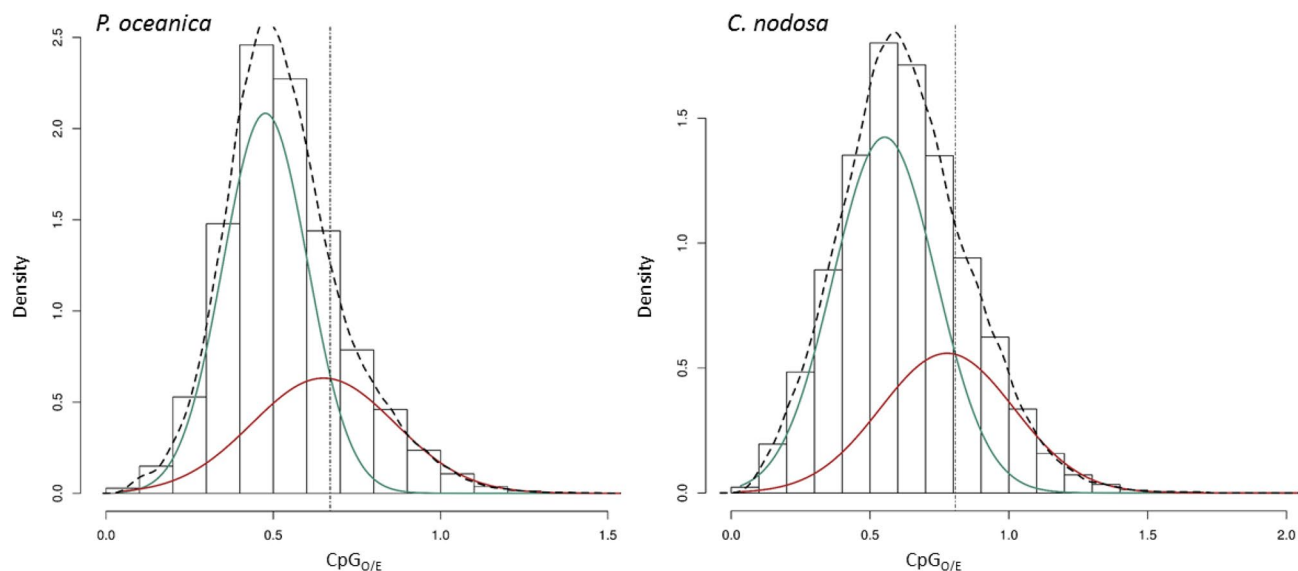
Although the causes and potential functional consequences of gbM for gene expression are not fully understood<sup>22</sup>, genic CG methylation was reported to be associated with gene expression in both plants and animals<sup>17</sup>. Several studies have found that methylation of gene bodies varies according to gene function, indicating that highly conserved genes with housekeeping functions tend to be more heavily methylated than those with inducible functions<sup>15,21,23–26</sup>. Thereby, contrary to the transcriptionally repressive effects of other chromatin modifications within gene bodies, gbM genes are typically moderately or constitutively expressed. One hypothesis is that the main function of gbM is most likely homeostatic e.g. by enhancing splicing accuracy, reducing the accumulation of histone variant H2A.Z and/or establishing constitutive expression patterns within housekeeping genes<sup>15,17,26–29</sup>. Yet, the absence of methylation within gene bodies could allow a variety of transcriptional opportunities in genes involved in stress and environmental responses<sup>24,30</sup>. In invertebrates, low gbM in environmental responsive genes seem to increase phenotypic plasticity and the adaptive potential of species<sup>18,24,25,31,32</sup>. Nevertheless, gbM may also lack functional consequences for gene expression<sup>21</sup>, and its complete loss in some basal plants<sup>21</sup> and in the angiosperm *Eutrema salsugineum*<sup>16</sup> demands for further studies on the general function of this type of DNA methylation in plants.

In the absence of direct methylation data, the normalized CpG content (i.e. the ratio of observed to expected CpG,  $CpG_{O/E}$ ) is a robust measure of the strength of DNA methylation levels on an evolutionary time scale<sup>33</sup>. The correlation between (historic) DNA methylation and normalized CpG content is based on the mutagenic properties of mC, which undergoes spontaneous deamination to thymine more readily than normal cytosine<sup>33</sup>. Because of this hypermutability, sequences that are heavily methylated in the germline become deficient in CpGs over evolutionary time<sup>34</sup>. Consequently, genomic regions that are subject to heavy germline DNA methylation (hypermethylated) lose CpG dinucleotides over time and have lower-than-expected  $CpG_{O/E}$ . In contrast, regions that undergo little germline DNA methylation (hypomethylated) maintain high  $CpG_{O/E}$ . In addition to capturing historic information on DNA methylation,  $CpG_{O/E}$  levels can be used as a proxy for *current* DNA methylation in the genomic region of interest<sup>35–37</sup> because  $CpG_{O/E}$  strongly correlates with direct measures of somatic DNA methylation<sup>38</sup>. This has been shown in many animals<sup>18,24,25,37,39,40</sup> and plants<sup>20,21</sup>.

Seagrasses are a polyphyletic group of angiosperms fully adapted to develop their life cycle completely submerged in marine waters, for which they have lost some biological characteristics of their terrestrial ancestors and acquired others from seaweeds<sup>41</sup>. This group of marine clonal plants forms extensive meadows in coastal bottoms of almost all continents, where they provide high-value ecological and economic services<sup>42</sup>, including coastal protection, nutrient cycling, water quality improvement, fishery maintenance and carbon sequestration, among others<sup>43</sup>. One of the central questions in seagrass ecology is the understanding of how seagrass ecosystems will respond to climate change, in particular to the increased incidence of severe climatic events such as summer heatwaves<sup>44,45</sup>. Predictions have initially challenged their long-term survival<sup>46–48</sup>, although some recent studies are glimpsing a more optimistic future for these marine plants. These studies have demonstrated that seagrasses are phenotypically flexible organisms that display intraspecific variability in morphology and physiology in response to environmental changes (e.g.<sup>49–51</sup>). Furthermore, current findings on gene expression patterns in seagrasses revealed a wide phenotypic plasticity, which can play a key role in determining how species and phenotypes withstand climate change (e.g.<sup>52–54</sup>). So far, molecular studies on seagrasses remain limited and the complete genome has been sequenced only for two species of the *Zostera* genus<sup>41,55</sup>.

The epigenetic research on marine plants is still at the infancy to date. Just two studies have characterized methylation levels at specific gene loci of the species *Posidonia oceanica* in response to specific environmental conditions<sup>56,57</sup>. The level of global DNA methylation has also been seen to vary with tissue age, likely resulting from the interplay of developmental and light cues, and under heat stress and chronic low light<sup>58–60</sup>. Just very recently, methylome variation among ramets of the same genet have been reported in the seagrass *Zostera marina*, together with co-variation between DNA methylation and photosynthetic performance under experimental conditions, which suggests a link between epigenetic variations and phenotypic plasticity of the species<sup>61</sup>.

The general objective of the present study is to evaluate potential relationships between DNA methylation and environmentally-driven gene expression and phenotypic plasticity in marine plants. By exploring two species with different ecological and biological attributes, we expanded our analysis to contrasting life history strategies of seagrasses. In both species, we addressed historic gbM patterns and present-day whole genome methylation levels. The transcriptomes of the two species were examined through *in silico* analyses and the ratio of observed to expected CpG dinucleotides ( $CpG_{O/E}$ ) was used to predict methylation status of specific genes, on the basis of CpG evolutionary changes induced by the hypermutability of mC. In particular, we aimed: (i) to explore differences in the level of predicted gene body DNA methylation between housekeeping and environmental/stress responsive genes; (ii) to examine the potential link between gbM and transcriptional regulation, by comparing the transcriptome profiles of different species and ecotypes in relation to their thermal origin and to their response to heat stress. We hypothesized that environmentally flexible gene-expression patterns are associated with signatures of weak gbM, and that ecotypes with different thermo-tolerance would show distinct relationships between DNA methylation and gene-expression profiles. Finally, to complement the *in silico* analysis, (iii) we empirically characterize intra- and inter-specific differences in DNA methylation via colorimetric quantification of global DNA methylation.



**Figure 1.** Density plots of CpG<sub>O/E</sub> values from gene bodies of the seagrasses *P. oceanica* (left panel) and *C. nodosa* (right panel). Black dashed lines represent the unimodal CpG<sub>O/E</sub> distribution, whereas the green and red solid lines represent low-CpG<sub>O/E</sub> and high-CpG<sub>O/E</sub> components respectively, as derived from the bimodal distribution provided by the 2-component Gaussian model. Dotted vertical lines indicate the intersection of the two fitted component curves.

Species	Lambda		Mu		Sigma		Log-likelihood (k=1)	Log-likelihood (k=2)
	Comp-1	Comp-2	Comp-1	Comp-2	Comp-1	Comp-2		
<i>P. oceanica</i>	0.661	0.339	0.476	0.650	0.127	0.214	18,741	21,138
<i>C. nodosa</i>	0.663	0.337	0.553	0.778	0.186	0.241	1712	2156

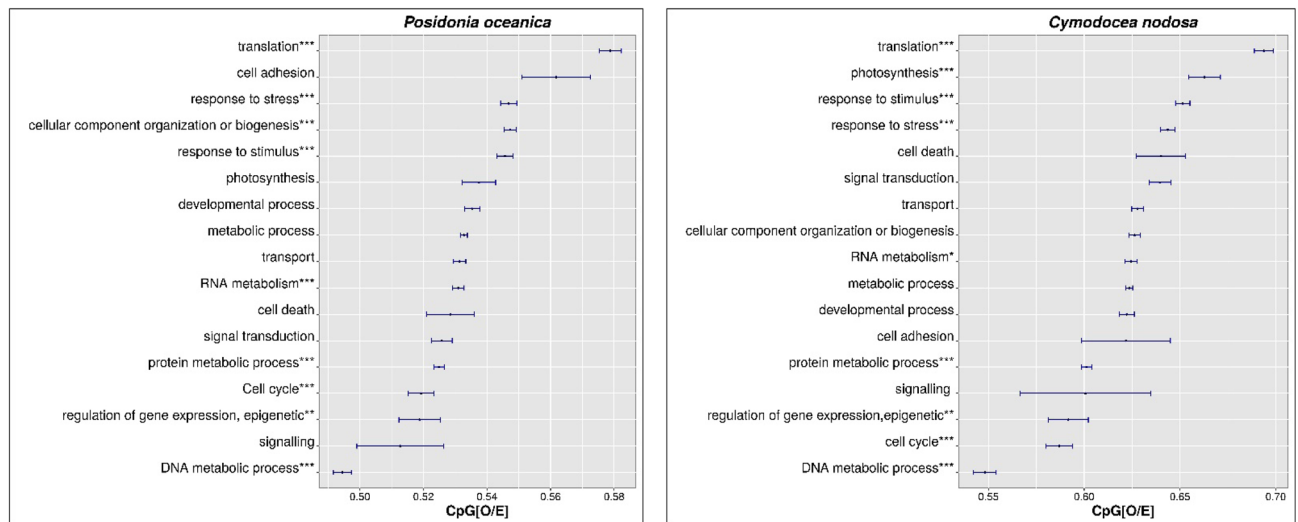
**Table 1.** Distribution parameters of the mixture models of CpG<sub>O/E</sub> in the two seagrass transcriptomes. Higher log-likelihood statistics indicates the model that better fits the density distribution.

## Results

**Evaluation of predicted gbM in seagrasses.** Our in silico transcriptome analysis showed that predicted patterns of gene body DNA methylation, as measured by CpG<sub>O/E</sub>, were similar for the two seagrass species (Fig. 1). Bimodal mixture models showed a significant improvement in fit compared to the normal null model indicating that the distribution of CpG<sub>O/E</sub> values for both species is better described by a mixture of two distinct Gaussian distributions (Table 1), although the visual examination of density plots suggests that bimodality is not very strong in either of the species (Fig. 1). The intersection point of the 2-component density curves was used to objectively attribute genes into two components: the ‘low-CpG component’ (biased toward hypermethylation; green line) and the ‘high-CpG component’ (biased toward hypomethylation; red line). The intersection of the two fitted component density curves was at 0.67 for *P. oceanica* and 0.80 for *C. nodosa*. A large number of *P. oceanica* (80.06%) and *C. nodosa* (78.94%) annotated genes were depleted in CpG dinucleotides, suggesting a considerably higher number of predicted hypermethylated genes (CpG<sub>O/E</sub> depletion).

**Relationship between gbM and gene function.** Both species showed similar patterns of gbM in relation to broad classes of biological processes (Fig. 2). Two distributions were comprised of functionally distinct gene classes in *P. oceanica* and *C. nodosa* and several biological processes showed CpG<sub>O/E</sub> ratios that were significantly different (Fisher’s exact test; FDR < 0.05), indicating that genes with lower-CpG<sub>O/E</sub> and higher-CpG<sub>O/E</sub> were involved in different biological functions. In both species, the top five biological processes with lower CpG<sub>O/E</sub> were protein metabolic process, cell cycle, DNA metabolic process (all  $p < 0.0001$ ), regulation of gene expression ( $p < 0.001$ ) and signaling (ns). High-ranked biological processes with the highest mean CpG<sub>O/E</sub> were a bit more variable between both species. However, genes with higher CpG<sub>O/E</sub> ratios showed significant enrichment of terms associated to translation, stress response and response to stimulus in both seagrasses ( $p < 0.0001$ ). Photosynthesis was also included within the high-ranked biological processes, ranking second and sixth in the list of *P. oceanica* and *C. nodosa*, respectively (Fig. 2).

**Relationships between gbM and gene expression patterns.** *Effects of plants’ thermal origin.* Differences in gene expression that were distinctive of the thermal origin (warm vs cold) of *P. oceanica* and *C. nodosa* plants were explored by grouping the samples of each species according to their origin, regardless of the



**Figure 2.** Variation of CpG<sub>O/E</sub> among genes assigned to different GO Slim terms (biological processes) in *P. oceanica* (left panel) and *C. nodosa* (right panel). Points represent mean CpG<sub>O/E</sub> for the corresponding biological processes ( $\pm$  standard error). Asterisks indicate significance of enrichment in the low- or high-CpG components (\*  $< 0.05$ , \*\*  $< 0.01$ , \*\*\*  $< 0.001$ ; Fisher's exact test).

	<i>P. oceanica</i>	<i>C. nodosa</i>
<b>Annotated genes with expression</b>	20,083 (14.89%)	12,230 (14.90%)
<b>Thermal origin</b>	4,465 (3.31%)	1,009 (1.23%)
<b>Experimental treatment</b>		
All samples	390 (0.29%)	1,870 (2.28%)
Cold-plants only	860 (0.64%)	1,614 (1.97%)
Warm-plants only	413 (2.06%)	1,071 (8.76%)

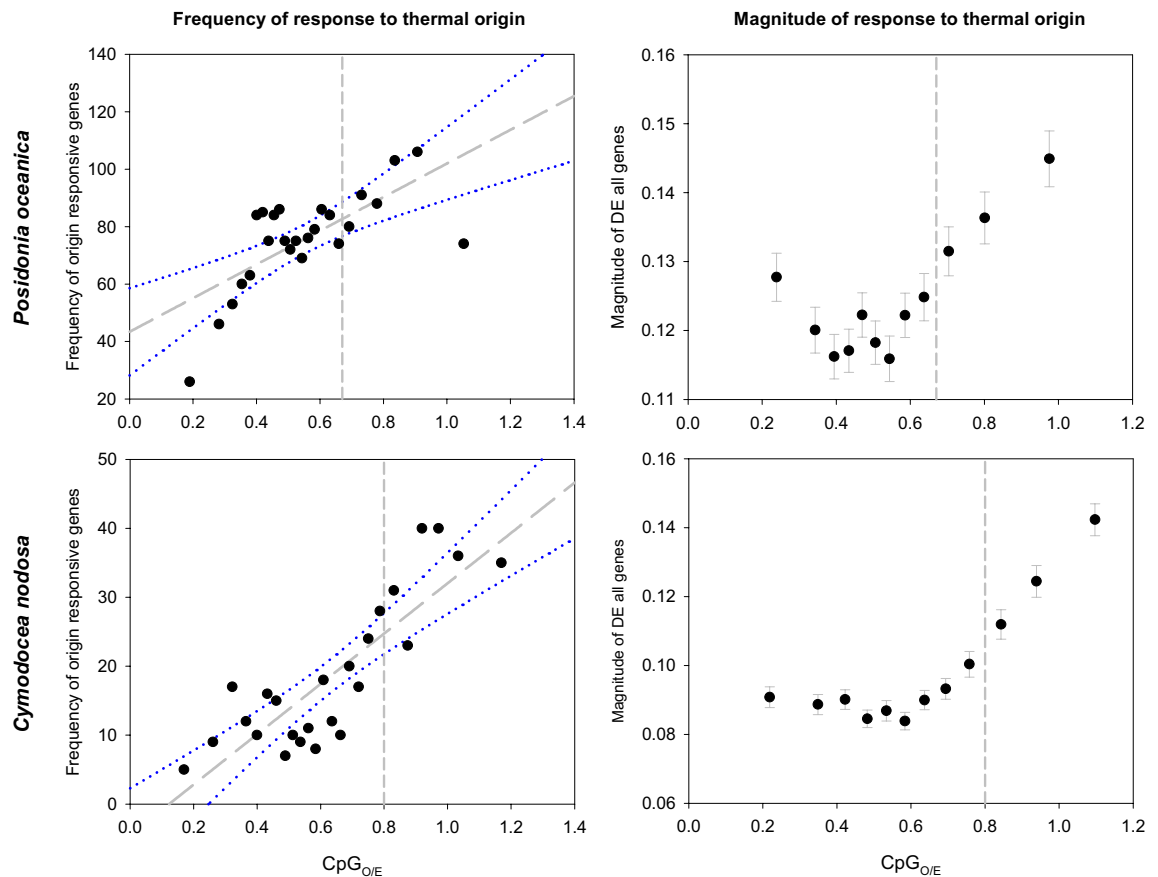
**Table 2.** Number of differentially expressed genes (FDR  $< 0.05$ ) in response to thermal origin of plants (warm-plants vs cold-plants), experimental treatment (i.e. warming; controls vs heated) and population-specific response to experimental warming (controls vs heated of cold- and warm-plants, separately). In parenthesis is shown the percentage that the number DEGs represent against the total number of genes of each species' transcriptome.

experimental treatment. The number of differentially expressed genes (DEGs) in the comparison between warm and cold populations was much higher in *P. oceanica* (4,465) than in *C. nodosa* (1,009), representing 3.31% and 1.23% of the total number of genes included in the species transcriptomes, respectively (Table 2).

Both species showed a progressive increase in the frequency of DEGs with the increase in CpG<sub>O/E</sub> values (Fig. 3, left panels) as supported by their significant linear correlations (Pearson;  $N = 25$ ,  $p < 0.001$ ,  $R = 0.699$  and  $0.844$  for *P. oceanica* and *C. nodosa*). Indeed, the number of DEGs in lower-CpG<sub>O/E</sub> bins was significantly lower than in higher-CpG<sub>O/E</sub> bins (t-test;  $p < 0.05$  for *P. oceanica* and  $p < 0.001$  for *C. nodosa*). In *P. oceanica*, the average number of DE genes was  $71 (\pm 15; SD)$  for “low-CpG”, and  $90 (\pm 12; SD)$  for “high-CpG” genes; while the corresponding values for *C. nodosa* were  $14 (\pm 6; SD)$  and  $34 (\pm 6; SD)$  respectively.

The magnitude of the expression difference between samples from different thermal origin also generally increased with the increase of CpG<sub>O/E</sub> in both species (Fig. 3, right panels), although this relationship was observed mainly for CpG<sub>O/E</sub> values  $> 0.5$ . Mean expression levels varied significantly across CpG<sub>O/E</sub> (ANOVA;  $p < 0.001$ ) and were positively correlated with CpG<sub>O/E</sub> (Spearman rank;  $N = 12$ ,  $p < 0.01$ ,  $R = 0.755$  and  $0.843$  for *P. oceanica* and *C. nodosa*). The differential expression due to plant origin increased sharply within the high-CpG<sub>O/E</sub> component. This component showed a log mean magnitude of  $0.138$  and  $0.126$  in *P. oceanica* and *C. nodosa* respectively, while the corresponding magnitudes of genes from the low-CpG<sub>O/E</sub> component were  $0.120$  and  $0.090$ . Mean magnitudes of expression of genes from the bins of each component were significantly different in both species (t-test;  $p < 0.001$ ).

**Species specific effects of experimental warming.** When sample of each species were grouped according to the experimental treatment regardless the thermal origin of plants, 390 and 1,870 DEGs (i.e. 0.29% and 2.28% of the total number of genes in the corresponding transcriptomes) were found between control and heated plants of *P. oceanica* and *C. nodosa*, respectively (Table 2). Heat-responsive genes were overrepresented within the high-CpG<sub>O/E</sub> component in both species, with average values that were the double that of values observed in the low-CpG<sub>O/E</sub> component (Fig. 4, left panels). In particular, the average frequency of *P. oceanica* heat-responsive



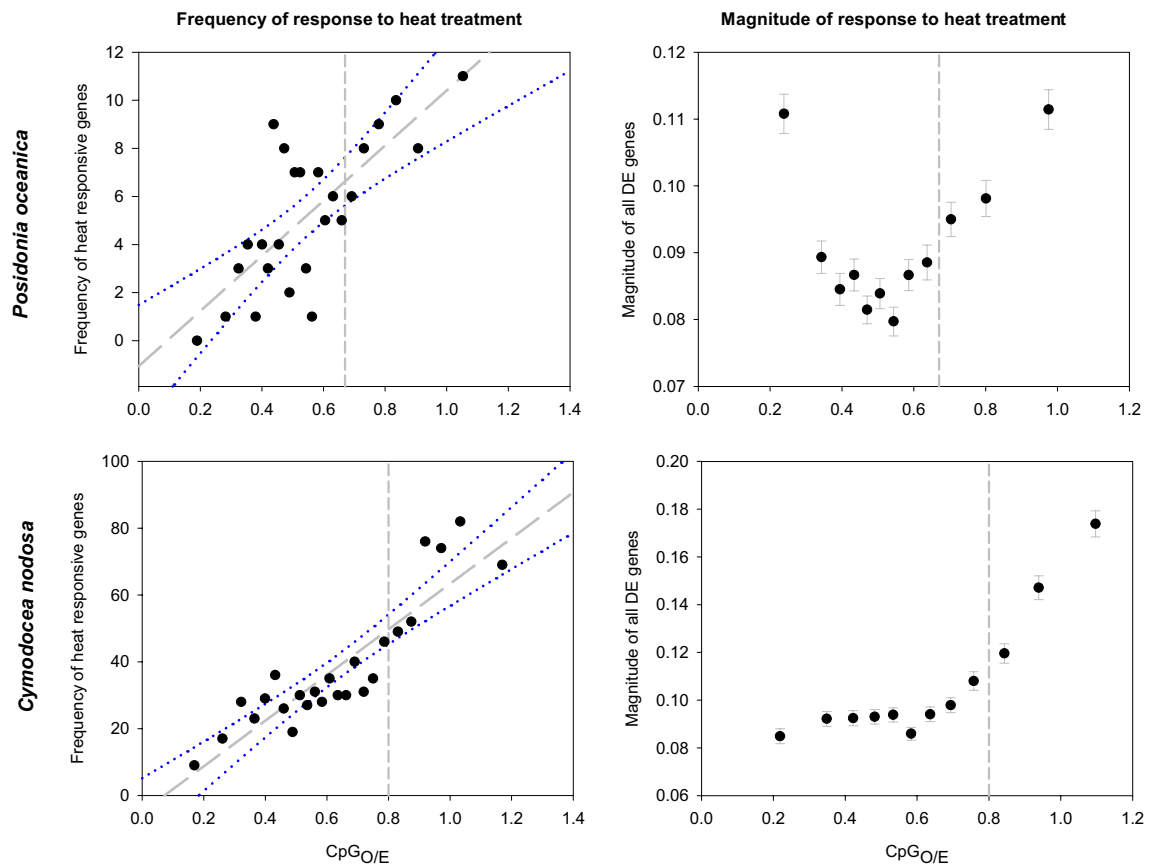
**Figure 3.** Relationship between  $\text{CpG}_{\text{O/E}}$  and gene responsiveness associated to the thermal origin of *P. oceanica* (upper panels) and *C. nodosa* (lower panels) plants. Frequency of plant origin flexible genes (left panels): each data point represents the number of DEGs (FDR < 0.05) responding to the thermal origin of plants within each of the 25  $\text{CpG}_{\text{O/E}}$  bins and the mean  $\text{CpG}_{\text{O/E}}$  for the bin. Magnitude of response to thermal origin of plants (right panels): magnitude of differential expression for each annotated gene. Magnitude points represent the log mean differential expression value and standard error of the genes divided in 12 bins. Dashed vertical lines represent the intersection of the two fitted component curves of the bimodal distribution of  $\text{CpG}_{\text{O/E}}$  in each species. Solid grey lines and dotted blue lines represent the Pearson correlation line and the 95% confident intervals respectively.

genes per bin was  $4 (\pm 2; \text{SD})$  for the low- $\text{CpG}_{\text{O/E}}$  component and  $9 (\pm 2; \text{SD})$  for the high- $\text{CpG}_{\text{O/E}}$  component. In *C. nodosa*, the number of genes significantly responding to the heat treatment was much higher with respect to *P. oceanica*, and showed an average of 29 per bin ( $\pm 8; \text{SD}$ ) for the low- $\text{CpG}_{\text{O/E}}$  component and 67 per bin ( $\pm 13; \text{SD}$ ) for the high- $\text{CpG}_{\text{O/E}}$  component. Both species showed a significant positive correlation between  $\text{CpG}_{\text{O/E}}$  values and frequency of DEGs per bin ( $N = 25$ ,  $p < 0.001$ ,  $R = 0.752$  and  $0.892$  for *P. oceanica* and *C. nodosa*) and mean DEGs frequency was significant different between bins from the low- and high- $\text{CpG}$  components (t-test;  $p < 0.001$ ).

The magnitude of differential expression in response to heat (i.e. experimental treatment) varied significantly across  $\text{CpG}_{\text{O/E}}$  in both seagrasses (ANOVA;  $p < 0.001$ ), being significantly higher in genes within the high- $\text{CpG}_{\text{O/E}}$  component than in genes from the low- $\text{CpG}_{\text{O/E}}$  component (t-test;  $p < 0.05$  for *P. oceanica* and  $p < 0.001$  for *C. nodosa*) (Fig. 4, right panels). In both species, the magnitude of expression of warming-responsive genes increased steeply within the high- $\text{CpG}_{\text{O/E}}$  component. However, the correlation between  $\text{CpG}_{\text{O/E}}$  and magnitude of expression was only significant in *C. nodosa* ( $N = 12$ ,  $R = 0.872$ ,  $p < 0.001$ ), since in *P. oceanica* ( $N = 12$ ,  $R = 0.316$ ,  $p = 0.317$ ) the average expression of genes from the first bin (i.e. lowest  $\text{CpG}_{\text{O/E}}$ ) was similar to the magnitude of genes from the highest  $\text{CpG}_{\text{O/E}}$  bin.

**Population-specific response to warming treatment.** At a smaller scale, we also examined the  $\text{CpG}_{\text{O/E}}$  composition of gene bodies in relation to differential gene expression under heat stress in plants of both species from the two different thermal origins (cold-plants vs warm-plants). In the four cases, the frequency of DEGs progressively increased with the increase of  $\text{CpG}_{\text{O/E}}$  levels (Fig. 5), and both parameters showed a significant linear tendency. In *P. oceanica*, DEGs frequency of cold-plants was significantly higher than in warm-plants (t-test,  $p < 0.001$ ), but similar in the case of *C. nodosa* (t-test,  $p = 0.070$ ). The slope of the linear adjustment of *P. oceanica*

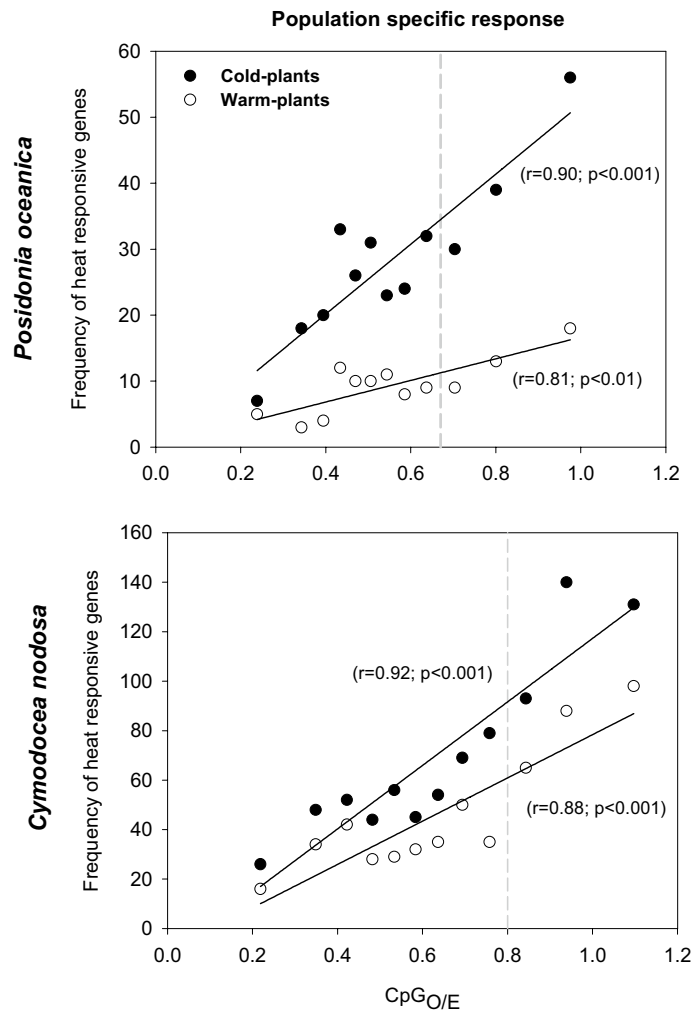




**Figure 4.** Relationship between  $CpG_{O/E}$  and gene responsiveness associated to the heat treatment in *P. oceanica* (upper panels) and *C. nodosa* (lower panels). Frequency of heat flexible genes (left panels): each data point represents the number of DEGs (FDR < 0.05) responding to the heat treatment within each of the 25  $CpG_{O/E}$  bins. Magnitude of response to heat exposure (right panels): magnitude of differential expression for each annotated gene. Magnitude points represent the log mean differential expression value and standard error of the genes divided in 12 bins. Dashed vertical lines represent the intersection of the two fitted component curves of the bimodal distribution of  $CpG_{O/E}$  in each species. Solid grey lines and dotted blue lines represent the Pearson correlation line and the 95% confident intervals respectively.

cold-plants, in fact, was 3.2 fold higher than the slope of warm-plants, while the difference was only of 1.5 for *C. nodosa* cold- and warm-plants.

The functional analysis of DEGs from the high- $CpG_{O/E}$  component of *P. oceanica* showed a higher number of enriched GO terms (i.e. biological processes; FDR < 0.001) in cold-plants with respect to warm-plants (176 vs 38; Supplementary Table S1a, b). The ten most significant Biological Processes (i.e. lower FDR values) were common in plants from both populations; however, only cold-plants were enriched in the GO terms “response to external stimulus” (GO: 0009605), “response to abiotic stimulus” and (GO:0009628) and “response to stress” (GO:0006950), which match with the higher level of physiological stress experienced by these plants during the warming exposure<sup>62</sup>. These plants were also exclusively enriched in the biological processes “generation of precursor metabolites and energy” (GO:1901135) and “carbohydrate derivative metabolic process” (GO:0006091), reflecting the impaired energetic status (i.e. 20% reduced starch content) only reported in this experiment for *P. oceanica* cold-plants<sup>62</sup>. In addition, “cell wall organization or biogenesis” (GO:0071554) and “cellular lipid metabolic processes” (GO:0044255) were other exclusive enriched processes in cold-plants, which coincides with their incomplete lipid readjustment for cell wall and membrane fluidity acclimation to warming<sup>63</sup>. For its part, the functional analysis of DEGs from the high- $CpG_{O/E}$  component of *C. nodosa* showed more uniform results between cold- and warm-plants in terms of number of GO-terms (170 and 128, respectively), in accordance with the similar DEGs frequency commented above. The enriched biological processes were quite similar between plants from both populations, with 84 shared GO-terms (Table S1c, d), likely reflecting the higher heat tolerance of the species and the similar heat responses of both *C. nodosa* ecotypes<sup>62</sup>. The main difference between both *C. nodosa* ecotypes was the inability of cold-plants to compensate heat-enhanced respiratory activity through a proportional increase in photosynthetic rates to avoid plant carbon imbalances<sup>62</sup>. Accordingly, “photosynthesis” (GO:0015979) and “photosynthesis, light reaction” (GO:0019684) were much more significantly enriched in cold-plants (5.46E–23 and 5.11E–18, respectively) than in warm-plants (8.91E–08 and 2.76E–04, respectively).



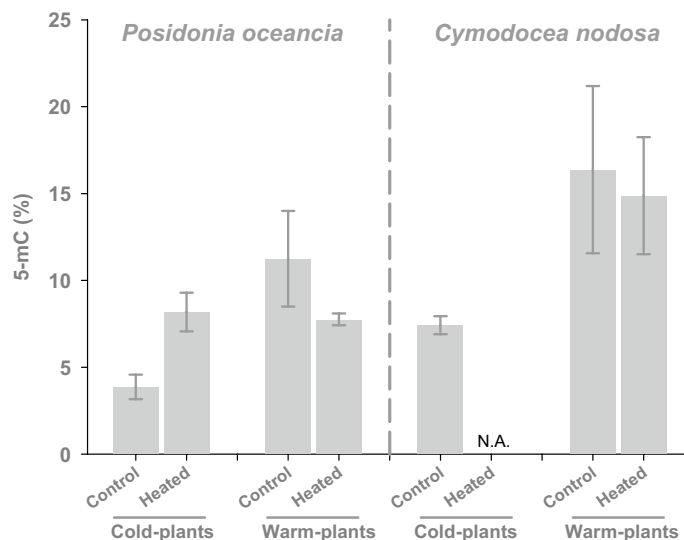
**Figure 5.** Relationship between CpG<sub>O/E</sub> and frequency of heat responsive genes of *P. oceanica* (upper panels) and *C. nodosa* (lower panels) cold- (solid circles) and warm-plants (empty circles). Frequency of heat flexible genes: each data point represents the number of DEGs (FDR < 0.05) responding to the thermal treatment within each of CpG<sub>O/E</sub> bins and the mean CpG<sub>O/E</sub> for the bin.

**Global DNA methylation.** The analysis of global DNA methylation (5-mC) showed constitutive differences between plants of both species from different origins (Fig. 6). The percentage of methylated-C in total genomic DNA of control warm-plants was significantly higher than that of control cold-plants ( $p < 0.05$ ). In response to heat stress, contrasting responses were observed between *P. oceanica* cold- and warm-plants. While cold-plants doubled the percentage of 5-mC ( $p < 0.05$ ), warm-plants did not significantly modify their global DNA methylation levels. Regarding *C. nodosa* warm-plants, they did not modify their percentage of 5-mC in response to heat stress exposure, and unfortunately the response of cold-plants could not be characterized due to the lack of plant material available for analysis in this particular treatment.

## Discussion

This study provides one of the first insights into the potential relationship between DNA methylation patterns and gene expression plasticity in marine plants. The two species here studied, *P. oceanica* and *C. nodosa*, have shown high levels of gene body methylation (gbM), differences in the level of total DNA methylation (5-mC) according to the origin of plants (e.g. latitudinal gradient of distribution) and the ability to modify global methylation levels in response to environmental changes (i.e. heat stress). We also found significant relationships between gbM and gene functionality, as well as between gbM and transcriptome profile under different environmental conditions, suggesting a possible link between epigenetic mechanisms and seagrass stress responses, as previously described in other organisms (e.g.<sup>64–67</sup>).

Our *in silico* transcriptome analyses of CpG<sub>O/E</sub> ratios have shown that genes of both *P. oceanica* and *C. nodosa* could be divided into two groups, according to their patterns of DNA methylation. Different CpG<sub>O/E</sub> ratio distributions have been recently described in eukaryotes<sup>38</sup>, including a highly marked bimodality in diverse invertebrate taxa, through similar analyses<sup>13,24,25,40</sup>. Gene G + C content seems to be heterogeneous in plants and bimodal



**Figure 6.** Global genomic methylation levels in *P. oceanica* and *C. nodosa* cold- and warm-adapted ecotypes from the control and heated treatments at the end of the warming exposure. Methylation values were calculated as percentage of the methylated-C in total genomic DNA.

and unimodal distributions have been observed through different genome-wide epigenetic methods (e.g. bisulfite sequencing, methylation-sensitive markers)<sup>68,69</sup>. Based on the consideration that gbM is evolutionarily conserved, it has been proposed that gbM may cause the bimodality of G + C content among grass (and monocot) genes<sup>20</sup>.

Our results are in accordance with the high levels of genic DNA methylation described within plant genomes<sup>14,17,21,70</sup>. Nevertheless, the percentage of gbM genes varies widely within and among species due to unknown reasons<sup>14,21,26</sup>. Most of *P. oceanica* and *C. nodosa* transcripts were depleted in CpG dinucleotides (i.e. CpG<sub>O/E</sub> < 1.0), suggesting that the bulk of their genes possess high levels of methylation.

Our analysis also indicated that in both seagrass species, genes with lower CpG<sub>O/E</sub> (predicted to be hyper-methylated) were significantly enriched for GO terms related to general metabolic functions, while high CpG<sub>O/E</sub> genes (predicted to be hypo-methylated) showed a significant enrichment of terms associated to dynamic biological functions. These results are consistent with distinctive functional enrichment patterns observed in other plant species<sup>15,20,23,26,70</sup> and invertebrates<sup>18,24,25,32</sup>. Hyper-methylation of ubiquitously expressed genes has been proposed as a homeostatic mechanism for limiting transcriptional regulation opportunities in critical genes (i.e. housekeeping). These genes are typically constitutively expressed and function to maintain basic cellular functions (e.g. <sup>17,28,30</sup>). On the contrary, genes with significantly lower levels of methylation tend to be the ones with inducible expression and more likely to exhibit tissue and/or environmental condition specificity<sup>24,30–32</sup>. Unmethylated or sparsely methylated genes have been proposed to passively facilitate a greater variety of transcriptional opportunities including access to alternative transcription start sites, increasing sequence mutations, exon skipping and greater epigenetic flexibility through transient methylation<sup>24,25,30,71</sup>. Therefore, the fact that genes differentially expressed in response to experimental warming were highly related to the high-CpG<sub>O/E</sub> component in *P. oceanica* and *C. nodosa* is in line with the hypothesis that weak gbM facilitates environmentally responsive expression. However, the functional consequence of gbM on gene expression is controversial and still needs to be demonstrated since several other factors, including gene length, the number of transcription factors regulating the genes and the chromatin environment are potentially involved in facilitating expression variability<sup>72,21</sup>.

Genes with differential expression depending on plants' origin tended toward high CpG<sub>O/E</sub>, suggesting that putative hypo-methylated genes are more likely to be differentially expressed between plants from different native thermal environments. Similar results have been also shown in natural ecological contexts of the reef-building coral *Acropora millepora*, where genes with differential expression based on coral's origin were enriched in the low methylation component<sup>32</sup>, reinforcing the possible link between weak gbM and gene expression (phenotypic) plasticity. As suggested in corals, one possible explanation for the origin-specific expression differences could be the genetic divergence among the studied populations<sup>32</sup>. A strong genetic structure along the latitudinal gradient of distribution of *P. oceanica* in the western Mediterranean has been reported<sup>73</sup> and accordingly, our experimental *P. oceanica* populations were genetically distinct<sup>62</sup>. In *C. nodosa*, however, such genetic pattern is not as strong within the western Mediterranean<sup>74</sup>.

Populations of the two species from the different thermal regimes showed contrasting levels of global genomic cytosine methylation (5-mC), with higher levels in warm-adapted with respect to cold-adapted plants. DNA methylation variations can result from stochastic epimutations, environmental induction or genetic control, which at the same time can be further shaped by natural selection and genetic drift<sup>22,75–77</sup>. Therefore, the interpretation of these epigenetic patterns is not straightforward in the studied populations since at the same time they naturally experience contrasting environmental conditions (e.g. temperature) and are genetically differentiated. In natural populations of the model species *Arabidopsis thaliana*, DNA methylation has shown striking correlations



with the place of origin and its climate. In particular, the latitudinal distribution of plants was strongly correlated with methylation levels in transposable elements<sup>66</sup> and in gene bodies<sup>65</sup>. This might suggest that temperature compensation has evolved in the natural range of the species favoring local adaptation. In accordance with this hypothesis, thermal adaptation along the latitudinal gradient of the species distribution has been evidenced in *P. oceanica*<sup>78</sup>, and the existence of local adaptation was recently suggested for *C. nodosa*, though in particular between Atlantic (Canary Archipelago) and Mediterranean populations<sup>79</sup>. In our analysis, this is further supported by the high number of genes differentially expressed on the basis of plants' origin in *P. oceanica* with respect to *C. nodosa* (4,465 vs 1,009). Moreover, in our *P. oceanica* plants, thermal tolerance was previously shown to be higher in ecotypes from warm waters (i.e. southern latitudes) in respect to ecotypes from cold waters, which was interpreted as the result of local adaptation<sup>62,63</sup>. In *C. nodosa*, instead, plants from contrasting thermal environments exhibited similar tolerance to experimental warming. These differences in thermal tolerance were also further supported by our analysis since the response to warming was markedly different between the two *P. oceanica* (warm- and cold-plants) but similar between the two *C. nodosa* ecotypes (Fig. 5).

After the warming exposure, warm-adapted ecotypes of both species did not significantly modify their global methylation levels, but interestingly methylation levels increased in cold-adapted *P. oceanica* ecotypes (no data from *C. nodosa* cold-plants). These results suggest that plants from different thermal origins show different epigenetic responses likely to cope better with increased temperatures. Several studies have revealed that environmental stress can result in an increase or decrease in cytosine methylation throughout the genome and at specific loci, to mediate environmentally-responsive and stress-responsive gene expression<sup>56,80–83</sup>. Nevertheless, the interaction can also be the other way round, since it is plausible that changes in cytosine methylation, as stated above, were actually the result of changes in gene expression under stressful conditions<sup>22,84</sup>. In accordance with our results, cytosine methylation increased more in heat-sensitive genotypes of *Brassica napus* under increased temperatures than in heat-tolerant genotypes, likely to respond and adapt to heat-stress<sup>80</sup>. In seagrasses, global DNA hyper-methylation and methylation changes in target functional genes of *P. oceanica* have been shown under increased cadmium levels and reduced light conditions<sup>56,57,60</sup>, and have been proposed to be at the basis of the species acclimation to heat stress<sup>54,58</sup>. *P. oceanica* cold-plants not only changed DNA methylation levels during the warming exposure, but they also showed the modulation of genes involved in epigenetic processes commonly used by plants for the transcriptional stimulation of stress tolerance and flowering as a possible mechanism to survive and optimize reproductive success under stress conditions<sup>85</sup>. A very recent paper showed that the methylome of the seagrass *Z. marina* is flexible and responds directly to environmental changes induced by plant cultured under laboratory conditions and to heat stress<sup>61</sup>. Authors suggest that the co-variation in DNA methylation and photosynthetic performance may be linked via gene expression because methylation patterns varied in functionally relevant genes involved in photosynthesis, and in the repair and prevention of heat-induced protein damage. In consequence, this suggests that in marine plants DNA is amenable to epigenetic modulation through methylation. Although we cannot discriminate whether this is the cause or the consequence of the activation of heat-stress responsive genes, we can assume that it could play a key role in seagrasses to cope with and adapt to seawater warming. In fact, the modulation of methylation-related genes in response to heat stress has recently been suggested to be at the basis of thermal stress memory in seagrasses with the potential to accelerate the adaptation of marine plants to the ongoing climate change<sup>86</sup>.

The increase of global DNA methylation (empirical determination) after warming exposure in cold-adapted *P. oceanica* ecotypes could be somehow in contrast with results obtained from in silico normalized CpG content (CpG<sub>O/E</sub>) analysis. Indeed, most of DEGs under heat stress belonged to the high-CpG component in both species, and thus they are predicted to exhibit low gbM. Nevertheless, CpG<sub>O/E</sub> does not necessarily correspond to the actual level of whole genome methylation, since cytosine methylation occurs within CG, CHG and CHH motifs<sup>87</sup>. Moreover, the global methylation analysis presented here does not discriminate between intra and inter-genic cytosine methylation, and the increase in DNA methylation could have occurred in inter-genic heterochromatic regions (e.g. repetitive, transposable elements (TEs)), which actually represent the bulk of methylated genomic sites in plants<sup>87</sup> and/or in untranscribed regulative regions (e.g. gene promoters). In particular, a higher level of cytosine methylation changes around TEs, in respect to other genomic regions, has been described in plants following abiotic stress events (e.g.<sup>88</sup>). Overall, our data may suggest that plant's stress response could occur in connection with both hypomethylation of responsive genes (gene body methylation) and hypermethylation of regulative regions and/or TEs<sup>84</sup> (non-genic methylation).

In conclusion, and in agreement with some previous observations (e.g.<sup>24,29–32</sup>) our results support the hypothesis that gene body methylation is associated with gene expression flexibility: low-gbM genes are much more likely to be differentially expressed in response to environmental stimuli than high-gbM genes. This is consistent with the idea that high gbM may function to suppress flexibility in transcriptional regulation. However, we acknowledge that our evidence is not based on direct estimations of DNA methylation at specific loci or throughout the genome, and that the role of gbM in regulating gene expression is still under an intensive debate. GbM could also be passively induced and other molecular processes could be responsible of its regulation<sup>22,72</sup>. Interestingly, we show that the same association with gbM exists for genes that are differentially expressed between populations from different origins. This suggests that low-gbM genes could not only be more inducible, but their expression may also evolve more rapidly, pointing to a more relevant role in population divergence and local adaptation. However, further experiments are certainly needed to infer about the role of gbM in phenotypic plasticity and adaptive capacity of marine plants. Experimental removal of DNA methylation using specific enzymes<sup>89</sup> or experimental modification of methylation patterns through genome editing methods (e.g. via CRISPR-Cas system<sup>90</sup>), could be critically useful to identify eventual causal relationships between methylation changes under stress and adaptive phenotypic changes.

## Methods

**Target species and experimental plants.** In this study, we compared an opportunistic and a persistent seagrass species that differ in plant size, life expectancy and reproductive characteristics. *P. oceanica* is one of the slowest-growing and longest-living plants on earth, with predominant clonal growth and whose extensive and permanent populations are considered one of the main climax stages of the Mediterranean coastal environment<sup>91</sup>. In contrast, *C. nodosa* is a fast-growing colonizer species of medium-size, characterized by high rates of both sexual reproduction and clonal propagation through rhizome elongation, and whose populations show a marked demographic seasonal pattern<sup>92</sup>. Based on these striking inter-specific differences, *C. nodosa* is considered to be more plastic to environmental changes, including light reduction<sup>93</sup>, salinity increase<sup>51</sup> and temperature rise<sup>50</sup>.

The experimental plants used in this study for the analysis of genic CG methylation patterns, total DNA methylation and transcriptomic changes came from the experiment described in<sup>62,63</sup>. Briefly, early summer plants of each species were collected from two different populations that were ca. 700 km apart and that experienced different natural thermal environments (up to 6 °C of difference in summer sea-surface temperature). Apical plant fragments bearing at least 10 interconnected vertical shoots were individually transplanted in plastic pots and four of these pots were randomly distributed in each tank of the experimental mesocosm systems of the Spanish Institute of Oceanography. Twelve independent 500-L tanks were used for *P. oceanica* and twelve 100-L tanks for *C. nodosa*, with half of the tanks containing plants from the cold-environment population (cold-plants) and the other half plants from the warmer-environment population (warm-plants). After a 3-week acclimation period, and for each of the species and ecotypes, the temperature in three of the six tanks was progressively increased of 4 °C (rate of 0.5 °C day<sup>-1</sup>) above mean summer levels (control temperature) and maintained during 6 weeks to simulate a marine summer heatwave. Thus, the experiment followed a factorial design with population (cold versus warm) and experimental treatment (control versus high temperature) as factors and three independent replicates per group. After the warming period, a 7 cm leaf segment from the middle portion of the youngest and fully developed leaf was collected (14:00 h) from one randomly selected shoot per pot. Leaf segments were conserved in RNAlater for RNA-seq analysis or flash frozen in liquid nitrogen for total DNA methylation analysis. All tissues were stored at – 80 °C until RNA and DNA extraction. Leaf pieces of similar biomass (25 mg W.W.) from the four pots of each tank were pooled ( $n = 3$  per treatment) before extractions.

**DNA extraction and global DNA methylation analysis.** Genomic DNA was extracted from leaf tissues with NucleoSpin® Plant II kit (Macherey–Nagel) and the quality assessed through 1.0% (w/v) agarose 0.5X TBE gel (0.5 mg/mL EtBr). DNA purity was measured using a NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific) and concentration determined by using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Global DNA methylation was assessed colorimetrically in duplicate by an ELISA-like reaction with the MethylFlash™ Methylated DNA Quantification Kit (Epigentek Inc.), and reported as % of 5-mC methylated DNA relative to the input DNA quantity. Fifty nanograms of DNA per sample were analyzed. Absorbance at 450 nm was assayed using a Multiskan™ FC Microplate Photometer (Thermo Fisher Scientific). Significant differences in the level of global methylation among samples were explored through ANOVA analysis using the statistical package STATISTICA (StatSoft, Inc.v. 10).

**RNA extraction and sequencing.** Leaf RNA was extracted using the Aurum Total RNA kit (Bio-Rad). The quality and quantity of the extractions were assessed by using Nanodrop (Thermo Fisher Scientific), 1% agarose gel electrophoresis and a 2100 BioAnalyzer (Agilent). Twelve libraries of each species, six from each thermal-population, half of which were from control tanks and half from heated tanks, were sequenced on an Ion Proton™ sequencer and a total of 106,210,104 and 89,086,102 single-end reads of 85 bp in average were generated for *P. oceanica*<sup>85</sup> and *C. nodosa*, respectively.

**Seagrass transcriptomic resources.** We used transcriptome assemblies as a basis for gene body methylation analysis, since no genomic references are currently available for *P. oceanica* and *C. nodosa* (see Table 3 and Supplementary information). For *P. oceanica*, we used a transcriptome assembly that was recently generated<sup>85</sup>, which is the most complete assembly currently available. A description of the *P. oceanica* transcriptome construction and functional annotation can be found in<sup>85</sup>. For building the *C. nodosa* transcriptome, we updated a previously published transcriptome assembly<sup>94</sup> with the transcriptome reads obtained in this study (77,222,158 (86.68% of raw reads) HQ single-end cleaned reads, see Supplementary information). The same procedure as in<sup>85</sup> was followed for the functional annotation and characterization of the *C. nodosa* transcript set (Supplementary information).

**Evaluation of predicted germline DNA methylation patterns in seagrasses.** We characterized CpG density in gene bodies through the CpG observed/expected ratio (CpG<sub>O/E</sub>). The CpG<sub>O/E</sub> for each gene was defined as:

$$CpG_{O/E} = \frac{\text{number of CpG}}{\text{number of Cx number of G}} \times \frac{l^2}{l-1}$$

where  $l$  is the number of nucleotides in the gene to account for the effect of gene length on gbM<sup>31,72</sup>. In silico analysis was limited to annotated sequences in order to be confident that sequences were reported in the 5' to 3' direction.

	<i>P. oceanica</i>	<i>C. nodosa</i>
Number of transcripts	225,579	117,740
Number of genes	134,863	82,095
GC	41.02	41.18
N50	1,993	1,819
Mean (bp)	990	933.55
Median (bp)	466	455
Minimum contig length (bp)	201	201
Maximum contig length (bp)	17,138	18,275
Number of transcripts > 1,000 bp	90,311 (40.04%)	45,845 (38.94%)
RMBT	74.95%	64.09%
<b>BUSCO</b>		
Complete orthologs	88.2%	79.24%
Single-copy	34.9%	52.36%
Duplicated	53.3%	26.88%
Fragmented orthologs	5.1%	8.61%
Missing orthologs	97.67%	12.15%
Total orthologs	1,440	1,440
Number of annotated transcripts	65,388 (48.11%)	39,802 (33.80%)

**Table 3.** Summary statistics of *Posidonia oceanica*<sup>85</sup> and *Cymodocea nodosa* de novo transcriptomes used in the present study. GC = the proportion of guanine and cytosine nucleotides among total nucleotides; N50 = the length of the longest contig such that all contigs of at least that length compose at least 50% of the bases of the assembly; RMBT = percentage of reads that mapped back to transcripts in each final assembly; BUSCO: Benchmarking Universal Single-Copy Orthologs.

To evaluate if CpG<sub>O/E</sub> values of both transcriptomes were best described as a single distribution or as a mixture of different distributions, we used the package Mclust<sup>95</sup> in R v3.4.1<sup>96</sup>. Bayesian information criterion (BIC) was used to compare the likelihood of Gaussian mixture models with different numbers of components. Subsequently, the normalmixEM function in the mixtools R package<sup>97</sup> was used to fit and estimate the mixture distribution parameters. Mixture models were evaluated against the null single-component model by comparison of log-likelihood statistics.

**Relationship between gbM and gene function.** To evaluate the variation of CpG<sub>O/E</sub> within and among functional biological process of *P. oceanica* and *C. nodosa* genes, we summarized this set of annotations to 19 high level Gene Ontology (GO) biological processes list using the Map2Slim option in OWLTools (<https://github.com/owlcollab/owltools>). The GO terms were selected to match a subset of those analyzed in previous works<sup>25,31,32</sup> as well as to better demonstrate the spread of “housekeeping” versus dynamic biological processes.

In order to assess if low-CpG and high-CpG genes were functionally distinct, we performed a GO biological process term enrichment using the fisher.test function in R v3.4.1, by comparing low-CpG and high-CpG genes separately with a background reference composed by all genes.

**Relation between gbM and transcription.** Reads generated from the experimental plants of this study were individually mapped to the corresponding transcriptome using the Bowtie v.1.2 aligner<sup>98</sup> and expression levels of each gene were quantified using the Expectation–Maximization (EM) method of RSEM v1.3.0<sup>99</sup>. Before differential expression (DE) analysis, very lowly expressed genes were removed keeping those genes that have at least a cpm (read/count per million) of 1 or greater for at least three samples (the size of the smallest group of replicates); subsequently, data were normalized to scale the raw library sizes. Statistical inference of differential expression at the gene level was performed by edgeR v3.16.5<sup>100</sup> under a generalized linear model (GLM) approach.

For each species, DE analysis was firstly performed between ecotypes (i.e. cold-plants vs warm-plants, irrespective of their thermal treatment) to test for differences in gene expression that were distinctive of the populations of origin. DE comparison was also performed between control and heated plants of each species (irrespective of their thermal origin) to test the impact of seawater warming on gene expression. Finally, DE analysis was conducted for each of the ecotypes independently (control vs heated plants) to explore the potential role of gbM on the transcriptomic response (i.e. plasticity) of cold- and warm-plants to heat stress. Genes were considered differentially expressed at a significance of a false discovery rate (FDR) threshold corrected p value < 0.05. The magnitude of the expression difference for each DEG was calculated as log (mean expression Warm-plants/mean expression Cold-plants) for the plants origin comparison and, log (mean expression Heated plants/mean expression control plants) for the warming treatment comparison.

Trends between differential expression and CpG composition of genes were explored by plotting the frequency (i.e. number of DEGs) and mean differential expression (i.e. magnitude) values against predicted methylation. For each gene, expression was averaged across all samples. These data were then joined with CpG<sub>O/E</sub> data and

genes were divided into 25 (803 and 489 genes per quantile for *P. oceanica* and *C. nodosa*, respectively) and 12 (1.672 and 1.019 genes per quantile for *P. oceanica* and *C. nodosa*, respectively) equally sized quantiles based on CpG<sub>O/E</sub> for the frequency and magnitude analysis, respectively. Mean expression for all genes in each quantile was then plotted against the mean CpG<sub>O/E</sub> value for the quantile.

Relationships between CpG<sub>O/E</sub> values and the frequency and magnitude of expression of DEGs were explored through Pearson product-moment correlation. Analysis of variance (ANOVA) was conducted to test if the magnitude of expression significantly differed among CpG<sub>O/E</sub> bins. Significant differences in the number of DEGs or the magnitude of expression between the low-CpG component and the high-CpG component were assessed through *t*-test. A *p*-value cutoff of 0.05 was used in the statistical analyses that were performed using the statistical package STATISTICA (StatSoft, Inc.v. 10).

## Data availability

This study made use of publicly available transcriptome data. *P. oceanica* data are available at the Sequence Read Archive (SRA) at the National Center for Biotechnology Information (NCBI) under the Accession Number SRP126951 (RNA-Seq reads) and at DDBJ/EMBL/GenBank under the accession GGFN00000000. *C. nodosa* data are available at the ENA (European Nucleotide Archive) repository (<http://www.ebi.ac.uk/ena>) under Accession nos. HADH01000001-HADH01059478. RNA-Seq reads generated for this study are accessible at the Sequence Read Archive (SRA) at the National Center for Biotechnology Information (NCBI) under BioProject ID PRJNA548629.

Received: 19 October 2020; Accepted: 28 June 2021

Published online: 12 July 2021

## References

- Merilä, J. & Hendry, A. P. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.* **7**, 1–14. <https://doi.org/10.1111/eva.12137> (2014).
- Reusch, T. B. Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. *Evol. Appl.* **7**, 104–122. <https://doi.org/10.1111/eva.12109> (2014).
- Pazzaglia, J., Reusch, T. B., Terlizzi, A., Marín-Guirao, L. & Procaccini, G. Phenotypic plasticity under rapid global changes: the intrinsic force for future seagrasses survival. *Evol. Appl.* (2021).
- Lopez-Maury, L., Marguerat, S. & Baehler, J. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat. Rev. Genet.* **9**, 583–593 (2008).
- Mäkinen, H., Papakostas, S., Vøllestad, L. A., Leder, E. H. & Primmer, C. R. Plastic and evolutionary gene expression responses are correlated in European grayling (*Thymallus thymallus*) subpopulations adapted to different thermal environments. *J. Hered.* **107**, 82–89 (2016).
- Alonso, C., Pérez, R., Bazaga, P., Medrano, M. & Herrera, C. M. MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species. *Mol. Ecol. Resour.* **16**, 80–90 (2016).
- Jeremias, G. *et al.* Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. *Mol. Ecol.* **27**, 2790–2806 (2018).
- Nicotra, A. B. *et al.* Adaptive plasticity and epigenetic variation in response to warming in an Alpine plant. *Ecol. Evol.* **5**, 634–647 (2015).
- Kelly, S., Panhuis, T. & Stoehr, A. (2012).
- Thorson, J. L. *et al.* Epigenetics and adaptive phenotypic variation between habitats in an asexual snail. *Sci. Rep.* **7**, 1–11 (2017).
- Rey, O., Danchin, E., Mirouze, M., Loot, C. & Blanchet, S. Adaptation to global change: a transposable element–epigenetics perspective. *Trends Ecol. Evol.* **31**, 514–526. <https://doi.org/10.1016/j.tree.2016.03.013> (2016).
- Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
- Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
- Niederhuth, C. E. *et al.* Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* **17**, 1–19 (2016).
- Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* **126**, 1189–1201 (2006).
- Bewick, A. J. *et al.* On the origin and evolutionary consequences of gene body DNA methylation. *Proc. Natl. Acad. Sci.* **113**, 9111–9116 (2016).
- Bewick, A. J. & Schmitz, R. J. Gene body DNA methylation in plants. *Curr. Opin. Plant Biol.* **36**, 103–110 (2017).
- Sarda, S., Zeng, J., Hunt, B. G. & Yi, S. V. The evolution of invertebrate gene body methylation. *Mol. Biol. Evol.* **29**, 1907–1916 (2012).
- Takuno, S. & Gaut, B. S. Body-methylated genes in Arabidopsis thaliana are functionally important and evolve slowly. *Mol. Biol. Evol.* **29**, 219–227 (2012).
- Takuno, S. & Gaut, B. S. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl. Acad. Sci.* **110**, 1797–1802 (2013).
- Takuno, S., Ran, J.-H. & Gaut, B. S. Evolutionary patterns of genic DNA methylation vary across land plants. *Nat. Plants* **2**, 1–7 (2016).
- Wendte, J. M. *et al.* Epimutations are associated with CHROMOMETHYLASE 3-induced *de novo* DNA methylation. *Elife* **8**, e47891 (2019).
- Aceituno, F. F., Moseyko, N., Rhee, S. Y. & Gutiérrez, R. A. The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics* **9**, 438 (2008).
- Elango, N., Hunt, B. G., Goodisman, M. A. & Soojin, V. Y. DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc. Natl. Acad. Sci.* **106**, 11206–11211 (2009).
- Gavery, M. R. & Roberts, S. B. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). *BMC Genomics* **11**, 1–9 (2010).
- Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. & Henikoff, S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61–69 (2007).
- Coleman-Derr, D. & Zilberman, D. in *Cold Spring Harbor symposia on quantitative biology*. 147–154 (Cold Spring Harbor Laboratory Press).
- Kim, M. Y. & Zilberman, D. DNA methylation as a system of plant genomic immunity. *Trends Plant Sci.* **19**, 320–326 (2014).



29. Muyle, A. & Gaut, B. S. Loss of gene body methylation in *Eutrema salsugineum* is associated with reduced gene expression. *Mol. Biol. Evol.* **36**, 155–158 (2019).
30. Roberts, S. B. & Gavery, M. R. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates?. *Front. Physiol.* **2**, 116 (2012).
31. Dimond, J. L. & Roberts, S. B. Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. *Mol. Ecol.* **25**, 1895–1904 (2016).
32. Dixon, G. B., Bay, L. K. & Matz, M. V. Bimodal signatures of germline methylation are linked with gene expression plasticity in the coral *Acropora millepora*. *BMC Genomics* **15**, 1–11 (2014).
33. Bird, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8**, 1499–1504 (1980).
34. Sved, J. & Bird, A. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl. Acad. Sci.* **87**, 4692–4696 (1990).
35. Suzuki, M. M., Kerr, A. R., De Sousa, D. & Bird, A. CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res.* **17**, 625–631 (2007).
36. Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466 (2007).
37. Glastad, K., Hunt, B. G., Yi, S. & Goodisman, M. DNA methylation in insects: on the brink of the epigenomic era. *Insect Mol. Biol.* **20**, 553–565 (2011).
38. Aliaga, B., Bulla, I., Mouahid, G., Duval, D. & Grunau, C. Universality of the DNA methylation codes in Eucaryotes. *Sci. Rep.* **9**, 1–11 (2019).
39. Asselman, J., De Coninck, D. I., Pfrender, M. E. & De Schampelaere, K. A. Gene body methylation patterns in *Daphnia* are associated with gene family size. *Genome Biol Evol* **8**, 1185–1196 (2016).
40. Park, J. *et al.* Comparative analyses of DNA methylation and sequence evolution using *Nasonia* genomes. *Mol. Biol. Evol.* **28**, 3345–3354 (2011).
41. Olsen, J. L. *et al.* The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* **530**, 331–335. <https://doi.org/10.1038/nature16548> (2016).
42. Costanza, R. *et al.* Changes in the global value of ecosystem services. *Glob. Environ. Chang.* **26**, 152–158. <https://doi.org/10.1016/j.gloenvcha.2014.04.002> (2014).
43. Nordlund, L. M., Koch, E. W., Barbier, E. B. & Creed, J. C. Correction: Seagrass ecosystem services and their variability across genera and geographical regions. *PLoS ONE* **12**, e0169942 (2017).
44. Orth, R. J. *et al.* A global crisis for seagrass ecosystems. *Bioscience* **56**, 987–996. [https://doi.org/10.1641/0006-3568\(2006\)56\[987:agcfse\]2.0.co;2](https://doi.org/10.1641/0006-3568(2006)56[987:agcfse]2.0.co;2) (2006).
45. Waycott, M. *et al.* Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proc. Natl. Acad. Sci.* **106**, 12377–12381. <https://doi.org/10.1073/pnas.0905620106> (2009).
46. Koch, M., Bowes, G., Ross, C. & Zhang, X. H. Climate change and ocean acidification effects on seagrasses and marine macroalgae. *Glob. Change Biol.* **19**, 103–132. <https://doi.org/10.1111/j.1365-2486.2012.02791.x> (2013).
47. Marbà, N. & Duarte, C. M. Mediterranean warming triggers seagrass (*Posidonia oceanica*) shoot mortality. *Glob. Change Biol.* **16**, 2366–2375. <https://doi.org/10.1111/j.1365-2486.2009.02130.x> (2010).
48. Thomson, J. A. *et al.* Extreme temperatures, foundation species, and abrupt ecosystem change: an example from an iconic seagrass ecosystem. *Glob. Change Biol.* **21**, 1463–1474. <https://doi.org/10.1111/gcb.12694> (2014).
49. Maxwell, P. S. *et al.* Phenotypic plasticity promotes persistence following severe events: physiological and morphological responses of seagrass to flooding. *J. Ecol.* **102**, 54–64 (2014).
50. Marín-Guirao, L., Ruiz, J. M., Dattolo, E., Garcia-Munoz, R. & Procaccini, G. Physiological and molecular evidence of differential short-term heat tolerance in Mediterranean seagrasses. *Sci. Rep.* **6**, 28615. <https://doi.org/10.1038/srep28615> (2016).
51. Sandoval-Gil, J. M., Ruiz, J. M., Marín-Guirao, L., Bernardeau-Esteller, J. & Sanchez-Lizaso, J. L. Ecophysiological plasticity of shallow and deep populations of the Mediterranean seagrasses *Posidonia oceanica* and *Cymodocea nodosa* in response to hypersaline stress. *Mar. Environ. Res.* **95**, 39–61. <https://doi.org/10.1016/j.marenvres.2013.12.011> (2014).
52. Franssen, S. *et al.* Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a marine foundation species. *Proc. Natl. Acad. Sci. USA* **108**, 19276–19281 (2011).
53. Jueterbock, A. *et al.* Phylogeographic differentiation versus transcriptomic adaptation to warm temperatures in *Zostera marina*, a globally important seagrass. *Mol. Ecol.* **25**, 5396–5411 (2016).
54. Marín-Guirao, L., Entrambasaguas, L., Dattolo, E., Ruiz, J. M. & Procaccini, G. Molecular mechanisms behind the physiological resistance to intense transient warming in an iconic marine plant. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2017.01142> (2017).
55. Lee, H. *et al.* The genome of a southern hemisphere seagrass species (*Zostera muelleri*). *Plant Physiol.* (2016).
56. Greco, M., Chiappetta, A., Bruno, L. & Bitonti, M. B. Effects of light deficiency on genome methylation in *Posidonia oceanica*. *Mar. Ecol. Prog. Ser.* **473**, 103–114 (2013).
57. Greco, M., Chiappetta, A., Bruno, L. & Bitonti, M. B. In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *J. Exp. Bot.* **63**, 695–709. <https://doi.org/10.1093/jxb/err313> (2012).
58. Ruocco, M., De Luca, P., Marín-Guirao, L. & Procaccini, G. Differential leaf age-dependent thermal plasticity in the keystone seagrass *Posidonia oceanica*. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2019.01556> (2019).
59. Ruocco, M., Marín-Guirao, L. & Procaccini, G. Within- and among-leaf variations in photo-physiological functions, gene expression and DNA methylation patterns in the large-sized seagrass *Posidonia oceanica*. *Mar. Biol.* **166**, 24. <https://doi.org/10.1007/s00227-019-3482-8> (2019).
60. Ruocco, M. *et al.* A king and vassals' tale: Molecular signatures of clonal integration in *Posidonia oceanica* under chronic light shortage. *J. Ecol.* (2020).
61. Jueterbock, A. *et al.* The seagrass methylome is associated with variation in photosynthetic performance among clonal shoots. *Front. Plant Sci.* **11** (2020).
62. Marín-Guirao, L. *et al.* Carbon economy of Mediterranean seagrasses in response to thermal stress. *Mar. Pollut. Bull.* **135**, 617–629 (2018).
63. Beca-Carretero, P. *et al.* Effects of an experimental heat wave on fatty acid composition in two Mediterranean seagrass species. *Mar. Pollut. Bull.* **134**, 27–37 (2018).
64. Angers, B., Castonguay, E. & Massicotte, R. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. *Mol. Ecol.* **19**, 1283–1295 (2010).
65. Dubin, M. J. *et al.* DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *Elife* **4**, e05255 (2015).
66. Kawakatsu, T. *et al.* Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell* **166**, 492–505 (2016).
67. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
68. Serres-Giard, L., Belkhir, K., David, J. & Glémin, S. Patterns and evolution of nucleotide landscapes in seed plants. *Plant Cell* **24**, 1379–1397 (2012).
69. Tatarinova, T., Elhaik, E. & Pellegrini, M. Cross-species analysis of genic GC3 content and DNA methylation patterns. *Genome Biol. Evol.* **5**, 1443–1456 (2013).



70. Vining, K. J. *et al.* Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics* **13**, 27 (2012).
71. Lyko, F. *et al.* The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol* **8**, 1506 (2010).
72. Cortijo, S., Aydin, Z., Ahnert, S. & Locke, J. C. Widespread inter-individual gene expression variability in *Arabidopsis thaliana*. *Mol. Syst. Biol.* **15**, e8591 (2019).
73. Procaccini, G., Olsen, J. L. & Reusch, T. B. H. Contribution of genetics and genomics to seagrass biology and conservation. *J. Exp. Mar. Biol. Ecol.* **350**, 234–259. <https://doi.org/10.1016/j.jembe.2007.05.035> (2007).
74. Alberto, F. *et al.* Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean-Atlantic transition region. *J. Biogeogr.* **35**, 1279–1294 (2008).
75. Becker, C. *et al.* Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**, 245–249 (2011).
76. Schmitz, R. J. *et al.* Patterns of population epigenomic diversity. *Nature* **495**, 193–198 (2013).
77. Yi, S. V. Insights into epigenome evolution from animal and plant methylomes. *Genome Biol. Evol.* **9**, 3189–3201 (2017).
78. Jahnke, M. *et al.* Adaptive responses along a depth and a latitudinal gradient in the endemic seagrass *Posidonia oceanica*. *Heredity* <https://doi.org/10.1038/s41437-018-0103-0> (2018).
79. Tuya, F. *et al.* Biogeographical scenarios modulate seagrass resistance to small-scale perturbations. *J. Ecol.* **107**, 1263–1275 (2019).
80. Gao, G. *et al.* Comparison of the heat stress induced variations in DNA methylation between heat-tolerant and heat-sensitive rapeseed seedlings. *Breed. Sci.* **64**, 125–133 (2014).
81. Downen, R. H. *et al.* Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci.* **109**, E2183–E2191 (2012).
82. Wada, Y., Miyamoto, K., Kusano, T. & Sano, H. Association between up-regulation of stress-responsive genes and hypomethylation of genomic DNA in tobacco plants. *Mol. Genet. Genomics* **271**, 658–666 (2004).
83. Yaish, M. W., Colasanti, J. & Rothstein, S. J. The role of epigenetic processes in controlling flowering time in plants exposed to stress. *J. Exp. Bot.* **62**, 3727–3735 (2011).
84. Secco, D. *et al.* Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *Elife* **4**, e09343 (2015).
85. Marin-Guirao, L., Entrambasaguas, L., Ruiz, J. M. & Procaccini, G. Heat-stress induced flowering can be a potential adaptive response to ocean warming for the iconic seagrass *Posidonia oceanica*. *Mol. Ecol.* **28**, 2486–2501. <https://doi.org/10.1111/mec.15089> (2019).
86. Nguyen, H. M. *et al.* Stress memory in seagrasses: first insight into the effects of thermal priming and the role of epigenetic modifications. *Front. Plant Sci.* **11**, 494 (2020).
87. Pikaard, C. S. & Scheid, O. M. Epigenetic regulation in plants. *Cold Spring Harbor Perspect. Biol.* **6**, a019315 (2014).
88. Yu, Y. *et al.* Cytosine methylation alteration in natural populations of *Leymus chinensis* induced by multiple abiotic stresses. *PLoS ONE* **8**, e55772 (2013).
89. Liu, R. & Lang, Z. The mechanism and function of active DNA demethylation in plants. *J. Integr. Plant. Biol.* **62**, 148–159 (2020).
90. Xu, X. *et al.* A CRISPR-based approach for targeted DNA demethylation. *Cell Discovery* **2**, 1–12 (2016).
91. Arnaud-Haond, S. *et al.* Implications of extreme life span in clonal organisms: millenary clones in meadows of the threatened seagrass *Posidonia oceanica*. *PLoS ONE* **7**, e30454. <https://doi.org/10.1371/journal.pone.0030454> (2012).
92. Mascaró, O., Romero, J. & Pérez, M. Seasonal uncoupling of demographic processes in a marine clonal plant. *Estuar. Coast. Shelf Sci.* **142**, 23–31 (2014).
93. Olesen, B., Enríquez, S., Duarte, C. M. & Sand-Jensen, K. Depth-acclimation of photosynthesis, morphology and demography of *Posidonia oceanica* and *Cymodocea nodosa* in the Spanish Mediterranean Sea. *Mar. Ecol. Prog. Ser.* **236**, 89–97. <https://doi.org/10.3354/meps236089> (2002).
94. Ruocco, M. *et al.* Genomewide transcriptional reprogramming in the seagrass *Cymodocea nodosa* under experimental ocean acidification. *Mol. Ecol.* **26**, 4241–4259. <https://doi.org/10.1111/mec.14204> (2017).
95. Fraley, C. & Raftery, A. E. Model-based methods of classification: using the mclust software in chemometrics. *J. Stat. Softw.* **18**, 1–13 (2007).
96. R Core Team (ISBN 3-900051-07-0, 2012).
97. Benaglia, T., Chauveau, D., Hunter, D., Young, D. mixtools: an R package for analyzing finite mixture models (2009).
98. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
99. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
100. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140. <https://doi.org/10.1093/bioinformatics/btp616> (2010).

## Acknowledgements

The experiment conducted in this study was funded by the European Union under the Seventh Framework (FP7) Program (HEATGRASS project, Tolerance to heat stress induced by climate change in the seagrass *Posidonia oceanica*, #624035) and by the Spanish Government (RECCAM project, Seagrass Meadows resilience to global warming: an analysis based on responses at eco-physiological, population and ecosystem levels, CTM2013-48027-C3-2-R). The work has also been partially supported by the project Marine Hazard, PON03PE\_00203\_1, Italian Ministry of Education, University and Research (MIUR). We are grateful to the Molecular Biology and Sequencing Service of the Stazione Zoologica Anton Dohrn, in particular to Elio Biffali, Pasquale De Luca, Raimondo Pannone and Elvira Mauriello, for transcriptome sequencing and technical support. Special thanks are given to Juan M. Ruiz, Rocio Garcia, Aranzazu Ramos and Jaime Bernardeau for their help in plant sampling and in conducting the mesocosm experiment. LM-G was supported by a Marie-Curie Fellowship (FP7-PEOPLE-IEF-2013; HEATGRASS Project); MR was supported by an SZN Ph.D. fellowship via the Open University.

## Author contributions

LE and LM-G planned and designed the study. LE conducted bioinformatics and together with LM-G data analysis. LM-G conducted the mesocosm experiment. MR performed the analysis of global DNA methylation. All authors participated to the interpretation of data. LE and LM-G led the writing of the paper with contributions from the rest of authors. All Authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93606-w>.

**Correspondence** and requests for materials should be addressed to G.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021