Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

Article

Atlas of Age- and Tissue-specific DNA Methylation during Early Development of Barley (*Hordeum vulgare*)

Moumouni Konate [1,6], Mike J. Wilkinson [2*], Benjamin T. Mayne [3], Eileen S. Scott [4], Bettina Berger [4,5] and Carlos M. Rodríguez López [4,6,*]

[1] Institut de l'Environnement et de Recherche Agricole (INERA), Station de Farako-Ba, 01 BP 910 Bobo-Dioulasso 01, Burkina Faso.

[2] Pwllpeiran Upland Research Centre, Institute of Biological, Environmental and Rural Sciences, Penglais Campus, Aberystwyth, Ceredigion, SY23 3FG, UK.

[3] Robinson Research Institute, School of Medicine, The University of Adelaide, SA 5005 Australia.

[4] School of Agriculture, Food and Wine, Waite Research Institute, The University of Adelaide, Waite Campus, PMB 1, Glen Osmond, SA 5064, Australia.

[5] The Plant Accelerator, Australian Plant Phenomics Facility, School of Agriculture, Food and Wine, Waite Research Institute, The University of Adelaide, Waite Campus, PMB 1, Glen Osmond, SA 5064, Australia.
[6] Environmental Epigenetics and Genetics Group; Department of Horticulture, College of Agriculture, Food and Environment, University of Kentucky, Lexington, KY 40546*

[*] Correspondence: carlos.rodriguezlopez@adelaide.edu.au, Phone: +61 8 8313 0774; mjw19@aber.ac.uk, Phone: +44 (0) 1970 823204

Abstract: The barley (*Hordeum vulgare*) genome comprises over 32,000 genes, with differentiated cells expressing only a subset of genes; the remainder being silent. Mechanisms by which tissue-specific genes are regulated are not entirely understood, although DNA methylation is likely to be involved. DNA methylation patterns are not static during plant development, but it is still unclear whether different organs possess distinct methylation profiles. Methylation-sensitive GBS was used to generate DNA methylation profiles for roots, leaf-blades and leaf-sheaths from five barley varieties, using seedlings at the three-leaf stage. Differentially Methylated Markers (DMMs) were characterised by pairwise comparisons of roots, leaf-blades and leaf-sheaths of three different ages. While very many DMMs were found between roots and leaf parts, only a few existed between leaf-blades and leaf-sheaths, with differences decreasing with leaf rank. Organ-specific DMMs appeared to target mainly repeat regions, implying that organ differentiation partially relies on the spreading of DNA methylation from repeats to promoters of adjacent genes. Furthermore, the biological functions of differentially methylated genes in the different organs correlated with functional specialisation. Our results indicate that different organs do possess diagnostic methylation profiles and suggest that DNA methylation is important for both tissue development and differentiation and organ function.

Keywords: epigenomics, *hordeum vulgare*, leaf, root, tissue-specific methylation, developmental epigenomics

---

1. Introduction

DNA methylation is an important characteristic of plant genomes [1,2], and can occur in all cytosine contexts (CG, CHG and CHH, where H = A, C or T) [3]. The effect of DNA methylation variants on plant development has been demonstrated through methylation alteration tests, which led to plant abnormalities [4,5]. Furthermore, DNA methylation has been reported to vary from tissue

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

to tissue in many species [6-10], and these methylation changes seemed to be essential for normal plant development [11,12].

Additionally, tissue-specific methylation was proposed to have a strong correlation with the differential expression of some tissue-specific genes. Examples include tissue-specific pigmentation in maize, which is reported to be under epigenetic control [13], and differential gene expression between organs attributed to differentially methylated regions in soybean [14] and sorghum [10]. These studies extended our understanding of the functional importance of tissue-specific DNA methylation, including its role in setting developmental trajectories [9,13,15].

A substantial proportion of developmentally expressed genes have alternative promoters (multiple promoters that regulate the same gene) which are under different regulatory programmes [16]. Maunmakea et al. [17] proposed that alternative promoters are, at least sometimes, controlled by intragenic DNA methylation. This form of developmental gene regulation is reasoned to be dependent on transposon activity [16] and by implication would mean that silencing of transposons due to DNA methylation may be central to tissue-specific gene expression. Also, tissue-specific gene expression has been associated with methylation changes in promoter regions [2,18,19], especially CG islands within promoters [20]. These studies indicate that tissue-specific gene expression does not rely on a single methylation pattern in the genome but, probably, on a combination of variable DNA methylation features.

The magnitude of differential methylation between tissues has been the subject of controversy. It was believed that significant distinctive DNA methylation existed only between specialised tissues such as endosperm, pollen, leaves and roots [9,10,21,22]. Nevertheless, many of these studies also showed that differential DNA methylation between organs, such as roots and leaves, was minor in rice [23], maize [24], sorghum [10] and Arabidopsis [9]. DNA methylation differences between roots and leaves were small in both ^{m}CG and ^{m}CHG contexts [9,10], with about 1% and 5% divergence, respectively, reported in Arabidopsis [9]. While these studies of differential DNA methylation between tissues generally compared the overall methylation levels [9,10,24], these results differ from comparisons made with differentially methylated markers (DMMs) between the same tissues [10], probably due to differences in methylation profiling methods, making it difficult to compare results from different studies. Therefore, it is difficult to know whether differences in the results concerning tissue-specific DNA methylation are due to the plant species or to the approach taken. The study of DNA methylation patterns in plant tissues is important for a better understanding of how these epigenetic markers determine tissue differentiation. Thus, further investigation is warranted to clarify organ specificity of cytosine methylation and the distribution patterns of tissue-specific DNA methylation markers in the plant genome.

To undertake such an investigation, we used barley, a globally important cereal crop, the genome of which has been sequenced recently [25]. The availability of a reference genome made barley a model for the study of cereal crops such as wheat, oats or rye. In this study, we assessed differential DNA methylation between two barley (*Hordeum vulgare*) organs (roots and leaves), using methylation-sensitive genotyping by sequencing (ms-GBS) on five genetically distinct varieties.. For the sake of simplicity and consistency with the literature, roots and leaves or leaf parts (sheath, blade) may be referred to here as tissues and not organs.

2. Results

Methylation-sensitive genotyping by sequencing

The sequencing of the 170 samples of barley tissue which met DNA quality requirements yielded over 900 million raw reads, with more than 91% bases above Q30 (99.9% accuracy of base call [26]) across all samples (Table 1). Of these reads, 99.27% contained the barcode and *Eco*RI/*Msp*I adapters ligated during library construction. Further filtering was performed to retain reads that strictly aligned with the barley reference genome. In this way, we obtained nearly 450 million reads (50.10%), with a mean of 2,637,916 high quality reads per sample. These high-quality reads accounted for 913,697 sequence tags, representing 32.30% of the 2,828,642 CCGG sites in the barley genome. Of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

95 these sequence tags, 748,594 (80.62%) showed some form of polymorphism for methylation between
96 samples.

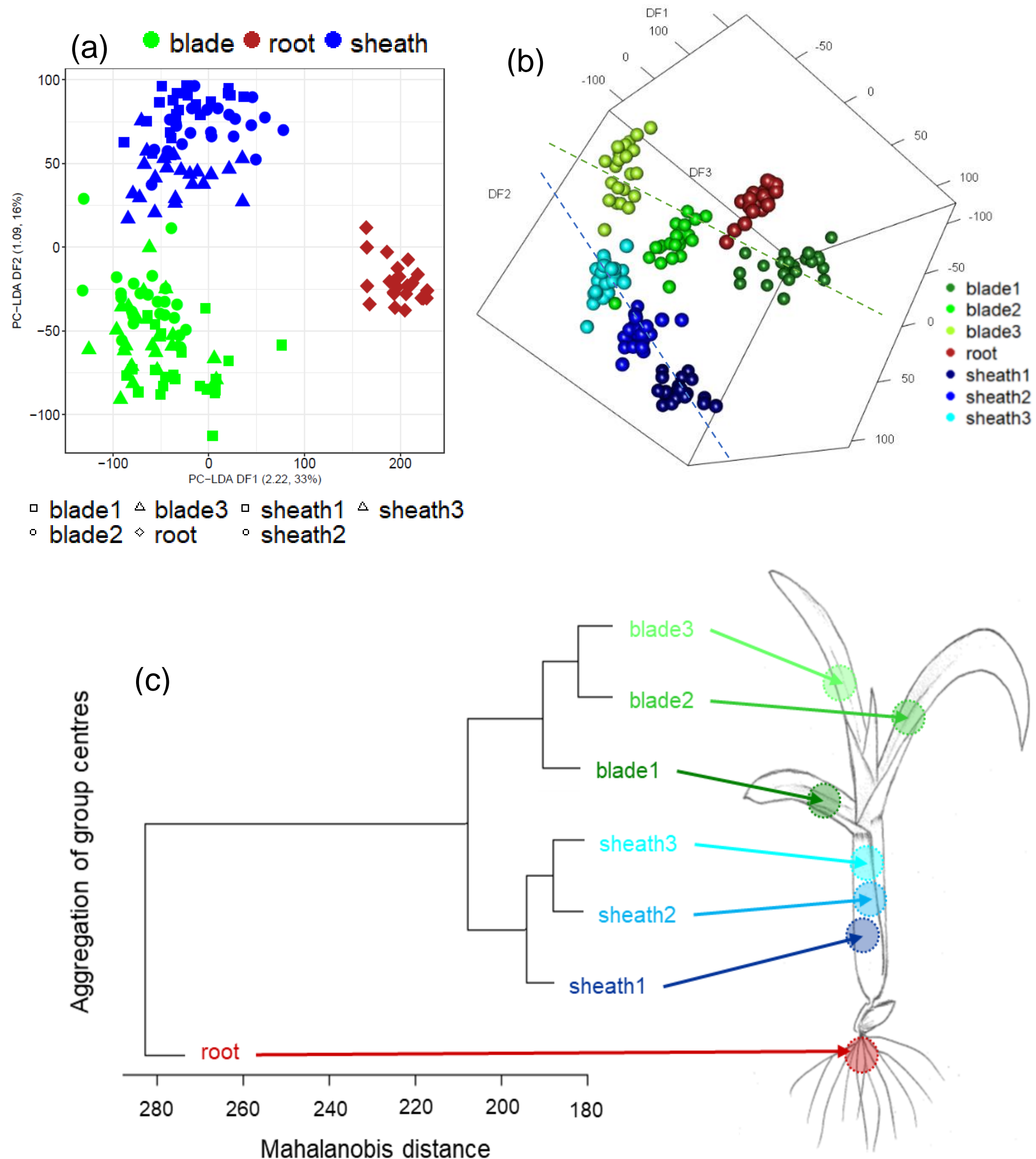
97 **Table 1**: Data yields from ms-GBS, generated using the Illumina HiSeq 2500 platform.

Raw reads	901,617,058
Reads that matched barcodes	895,013,295
Reads aligned to barley reference genome	448,445,748
Samples	170
Average reads per sample	2,637,916
Total unique tags	913,697
Polymorphic tags	748,594

98 *Estimation of tissue- and tissue rank- dependent epigenetic differentiation*

99 The PC-LDA plots revealed clear evidence of structuring of methylation between samples (Fig.
100 1a). A 3D plot using the first three discriminant factors (DF1, DF2 and DF3) revealed that blades and
101 sheaths were further grouped according to the rank of the leaf from which they were harvested. The
102 distance between blades and sheaths seems to shrink with leaf rank (Fig. 1b). This leaf rank-
103 dependent grouping was also supported by hierarchical cluster analysis (HCA) of the distances
104 between sample group centres (Fig. 1c), based on the Mahalanobis distance [27,28], and sample
105 clusters matched the leaf developmental age (Fig. 1c)). Leaf rank-dependent DNA methylation
106 differences were further assessed between tissues by comparing the methylation profiles of blades
107 and sheaths for each rank of leaf appearance. No DMMs were observed between the three leaf blades,
108 whereas sheaths 1 and 3 presented 18 DMMs (Table 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1



109

110 **Figure 1**: Analysis of the differentiation of DNA methylation profiles of barley roots, leaf sheaths and
111 leaf blades. (a) Scatter plot of the first two discriminant factors of the Principal Component - Linear
112 Discriminant Analysis (PC-LDA) (DF1 and DF2) using 913,697 ms-GBS markers generated from
113 genomic DNA of roots, leaf sheaths and leaf blades, collected from 25 barley plants at the three-leaf
114 stage (21 days after sowing), comprising five varieties (Barque 73, Flagship, Hindmarsh, Schooner
115 and Yarra). (b) Three-dimensional plot of the first three discriminant factors of the PC-LDA of the
116 same ms-GBS data. (c) Hierarchical cluster of the distances between sample group centres, based on
117 Mahalanobis distance. Blade 1-3 and sheath 1-3 indicate the rank of the organ type, first, second and
118 third leaf of seedlings, respectively.

119

120 **Table 2**: Number of Differentially Methylated Markers in barley tissues of different ages.
121 Differentially methylated markers (FDR <0.05) were obtained from 913,697 ms-GBS tags generated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

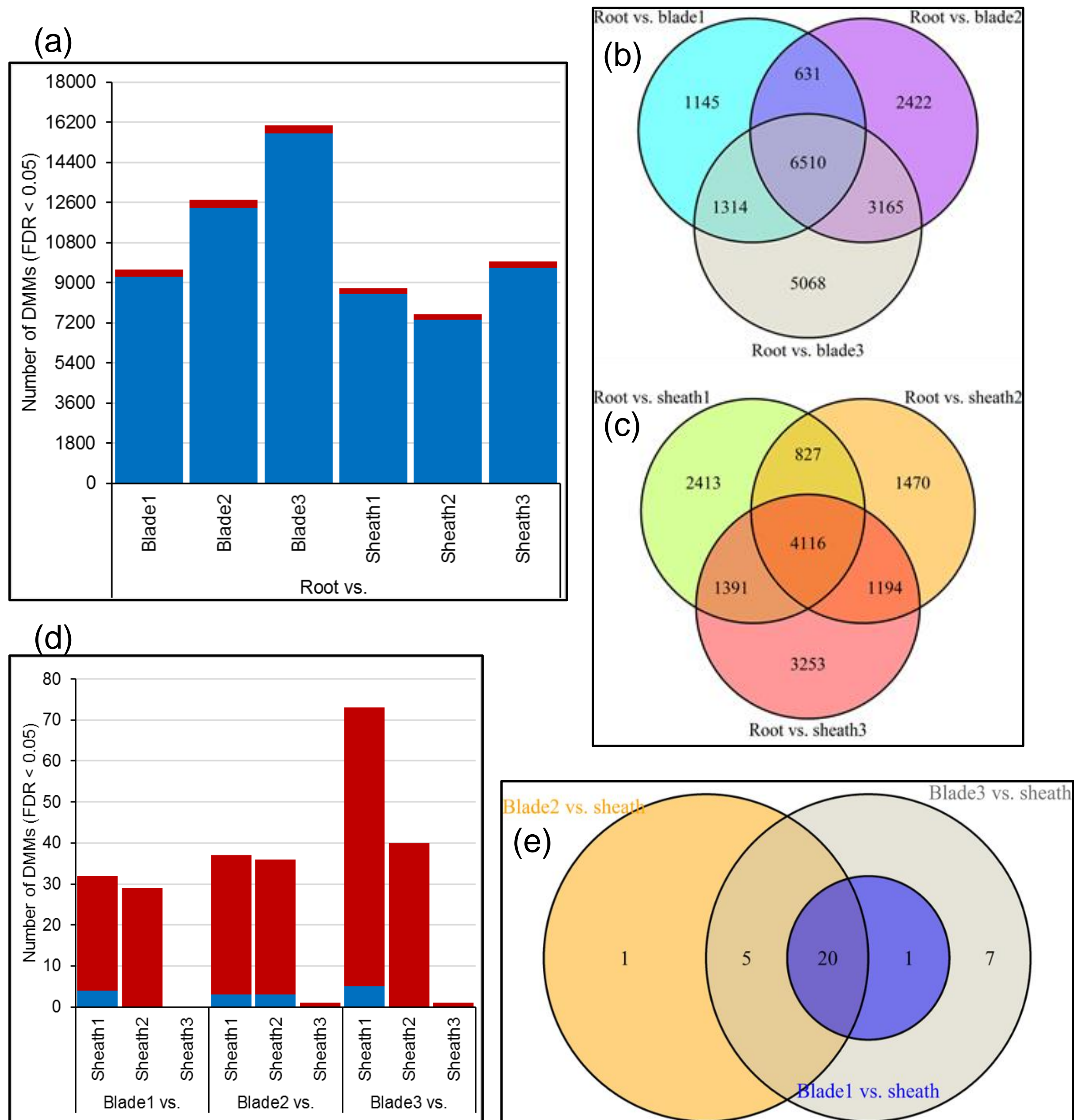
122 from genomic DNA of barley roots, leaf sheaths and leaf blades, collected from 25 plants at three-leaf
123 stage (21 days after sowing) of five barley varieties (Barque 73, Flagship, Hindmarsh, Schooner and
124 Yarra). Blade 1-3 and sheath 1-3 indicate the rank of the leave, first, second and third, respectively, on
125 seedlings.

	Blade 1	Blade 2	Blade 3	Sheath 1	Sheath 2	Sheath 3
Blade 1	-					
Blade 2	0	-				
Blade 3	0	0	-			
Sheath 1	32	37	73	-		
Sheath 2	29	36	40	0	-	
Sheath 3	0	1	1	18	0	-

126 *Differentially methylated DNA markers between roots and leaves*

127 DMMs between barley roots and leaves were obtained through comparison of the read count
128 per million of tissue types, independently of genotypes. This comparison revealed substantial DMMs
129 between both roots vs blades and roots vs. sheaths (Fig. 2a), and there were more DMMs between
130 roots and blades (6510 DMMs, Fig. 2b) than between roots and sheaths (4116 DMMs, Fig. 2c). Of these
131 markers, 3266 DMMs were present in both blades and sheaths when compared to roots, and their
132 methylation changed consistently in the same direction in each comparison (Fig. 3a). The number of
133 DMMs between roots and leaf blades increased with leaf-rank, whereas DMMs between roots and
134 leaf sheaths did not show any relationship with rank (Fig. 2a). Tissue-specific DMMs were
135 predominantly hypomethylated (95-98%) in leaf parts (sheath or blade) compared to roots (Fig. 2a).
136 This result was in line with the median of the fold-changes of DMMs, which indicated an overall
137 DNA hypomethylation in leaves (Fig. 4a,b). From here on, DMMs consistently present in roots vs.
138 sheaths and roots vs. blades will be designated as stable markers between roots and leaves.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1



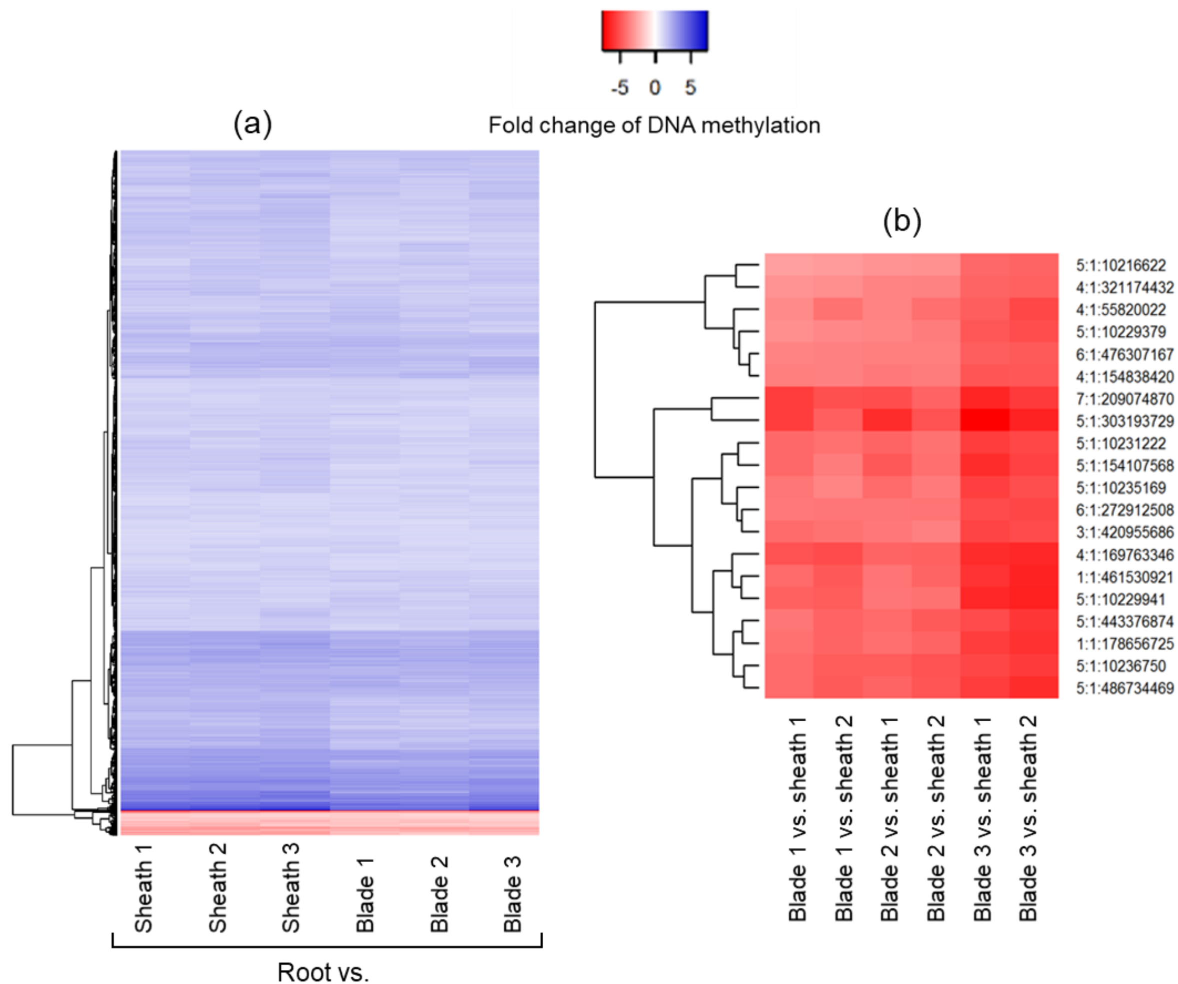
139

140 Figure 2: Analysis of the number of DMMs among three barley tissues. (a) Number of DMMs between
141 roots and leaf blades (Root vs. blade) and roots and sheaths (Roots vs. sheaths). Histogram colour
142 indicates whether the DMMs are hypomethylated (blue) or hypermethylated (red) in leaf parts
143 compared to roots. (b-c) Venn diagram showing the number of DMMs stable between root and blade
144 tissues (b) and between root and sheath tissues (c). (d) Number of DMMs from pairwise comparison
145 between leaf blades 1-3 and sheaths 1-3. Histogram colour indicates whether the DMMs are
146 hypomethylated (blue) or hypermethylated (red) in sheaths compared to blades. (e) Venn diagram
147 showing the number of DMMs common in pairwise comparisons between leaf blades 1-3 and sheaths
148 1-2. Tissue samples were collected from seedlings at the three-leaf stage of five barley varieties grown
149 in five replicates for 21 days after sowing. Blade 1-3 and sheath 1-3 indicate the rank of the organ type,
150 first, second and third, respectively, on seedlings. DMMs were selected based on the significance of
151 the false discovery rate, FDR, < 0.05. DMMs present in both sheaths and blades when compared with
152 roots, have been designated as markers between roots and leaves.

153 *Differentially methylated DNA markers between the leaf blade and sheath*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

154 There was only a small number of DMMs between leaf blades and sheaths (0 to 73 DMMs, Table
155 2; Fig. 2d). These DMMs were basically between leaf blades and sheaths 1 and 2; and there was none
156 between blade 1 and sheath 3. There was only 1 DMM between sheath 3 and blades 2 and 3 (Table 2;
157 Fig. 2d). Pairwise comparisons between blades 1-2 and sheaths 1-2 revealed 20 common DMMs,
158 which were all hypermethylated in sheaths compared to blades (Fig. 2e and Fig. 4b). Half of the 20
159 common DMMs between blades and sheaths were located on chromosome 5H. Furthermore, there
160 were no DMMs in pairwise comparisons among blades 1-3 and among sheaths 1-3, except between
161 sheath 1 and sheath 3 which had 18 DMMs (Table 2). However, comparing blades and sheaths of the
162 same leaf rank showed 32 DMMs between blade 1 and sheath 1, 36 DMMs between blade 2 and
163 sheath 2 and 1 DMM between blade 3 and sheath 3.
164

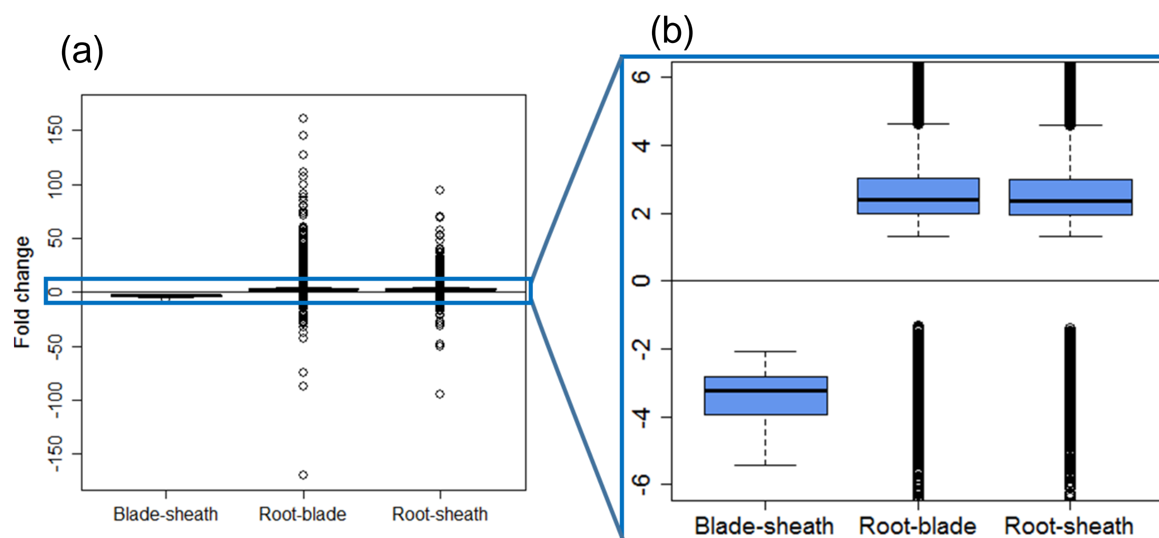


165

166 **Figure 3**: Hierarchical clustering analysis of the DMMs. (a) the 3266 common DMMs between roots
167 and all leaf parts (sheath 1-3, blade 1-3). The colours in the heat map indicate whether the DMM is
168 hypomethylated (blue) or hypermethylated (red) in leaf parts compared to roots. (b) Hierarchical
169 clustering of the 20 stable DMMs between blades and sheaths. In this heat map the red colour shows
170 hypermethylation of DMMs in sheaths compared to blades. Blade and sheath samples were collected
171 from seedlings at three-leaf stage of five barley varieties grown in five replicates for 21 days after
172 sowing. Blade 1-3 and sheath 1-3 indicate the rank of the leaf on seedlings, first, second and third,
173 respectively. The first number of the marker label on the y axis indicates the chromosome number on
174 which the marker is located.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

175

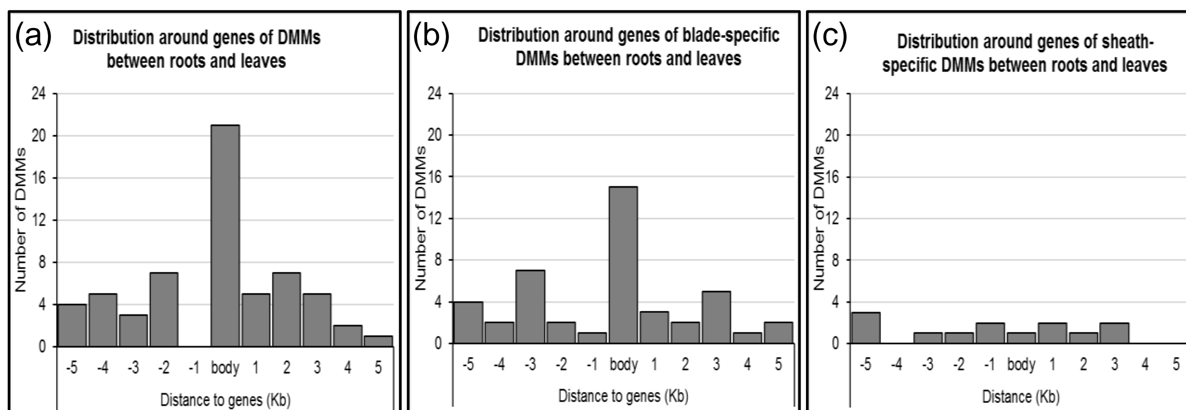


176 **Figure 4**: Directionality of the methylation in tissue-specific DNA methylation markers. (a) Boxplots
177 showing the spread of the fold-change of read counts of DMMs between blades and sheaths, roots
178 and blades, and roots and sheaths. (b) Detail of boxplots, highlighting the median of methylation fold-
179 change of all samples in each comparison. The fold-change of DNA methylation was estimated by
180 computing $2^{(log2FC)}$, with *log*2FC = logarithm 2 of fold-change in read counts per DMM between
181 pairwise comparisons of tissues collected from three-leaf stage barley seedlings. Leaf blades were the
182 reference state for blade-sheath comparison, whereas roots were the reference for root-blade and root-
183 sheath comparisons. Negative and positive values on the y axis indicate respectively,
184 hypermethylation and hypomethylation of the tissue that is compared to the reference.

185 *Distribution of tissue-specific DMMs around genes*

186 Relatively few of the tissue-specific DMMs were located around gene exons. Indeed, of the 3266
187 stable DMMs between root and leaf samples, only 60 (1.8%) were located within 5 Kb of a gene,
188 including 21 overlaps with genes and 39 DMMs that were spread within 5 Kb upstream and
189 downstream of genes (Fig. 5a). Apart from the absence of DMMs within 1 Kb upstream of
190 transcription start sites, there was no obvious tissue-specific DMM distribution pattern around the
191 genes (Fig. 5a). The same assessment process showed that, as with common DMMs, only a small
192 proportion of blade-specific DMMs (44 of 3246, 1.3%) was positioned close to a gene (Fig. 5b). Of
193 these, 15 DMMs overlapped with a gene transcript, whereas the remaining 29 DMMs were
194 distributed within 5 Kb of the gene without any clear pattern (Fig. 5b), except that the number of
195 DMMs located between 2 and 3 Kb bins was higher both upstream and downstream, than any other
196 1 Kb bin within the 5 Kb flanking regions (Fig. 5b). There were fewer sheath-specific methylation
197 markers within 5 Kb from genes than blade-specific markers (13 of 2391 DMMs, 0.5%) (Fig. 5c). The
198 majority of these (10 out of 13 DMMs) were sited within 3 Kb of a gene, and no DMMs were present
199 3-5 Kb from transcription margins (Fig. 5c). Of 37 gene-body DMMs detected across all comparisons
200 (Fig. 5a-c), 27 overlapped with an exon and the remaining 10 markers were in intergenic regions, 70
201 to 604 bp upstream of exons, except 1 DMM, which was 62 bp downstream an exon (Table S1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1



202

203 **Figure 5**: Distribution of tissue-specific differentially methylated markers (DMMs) around genes. (a)
204 DMMs between roots and leaves, present in both blades and sheaths as in Figure 2b-c; (b) Blade-
205 specific DMMs between roots and leaves and (c) Sheath-specific DMMs between roots and leaves.
206 The y axis indicates the distance to genes in kilo base pairs (Kb) on both flanking regions. Negative
207 and positive values indicate upstream and downstream of genes, respectively. DMMs overlapping
208 with genes are considered as changes in gene-body methylation (body). The x axis shows the number
209 of DMMs per 1 Kb window.

210 *Distribution of tissue-specific DMMs near repeat regions*

211 Many more tissue-specific DMMs were detected near repeats than near genes. The DMMs
212 around repeat regions (as defined in the Ensembl database
213 (http://plants.ensembl.org/biomart/martview/) were concentrated either within the repeats or within
214 1Kb of their margins (Fig. 6a). A similar distribution pattern was obtained with both blade-specific
215 and sheath-specific DMMs when contrasted with roots, with more DMMs overlapping with the
216 repeats themselves than in the 1 Kb stretches flanking their margins (Fig. 6b,c). The few markers that
217 were differentially methylated between blades and sheaths (20 DMMs in total) were all located within
218 1 Kb of a repeat (Fig. 6d). Therefore, stable tissue-specific DMMs appeared to occur preferentially
219 within repeats and 1 Kb flanking regions, with higher frequency within 1 Kb downstream than within
220 1 Kb upstream, regardless of tissue types (Fig. 6a-d).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

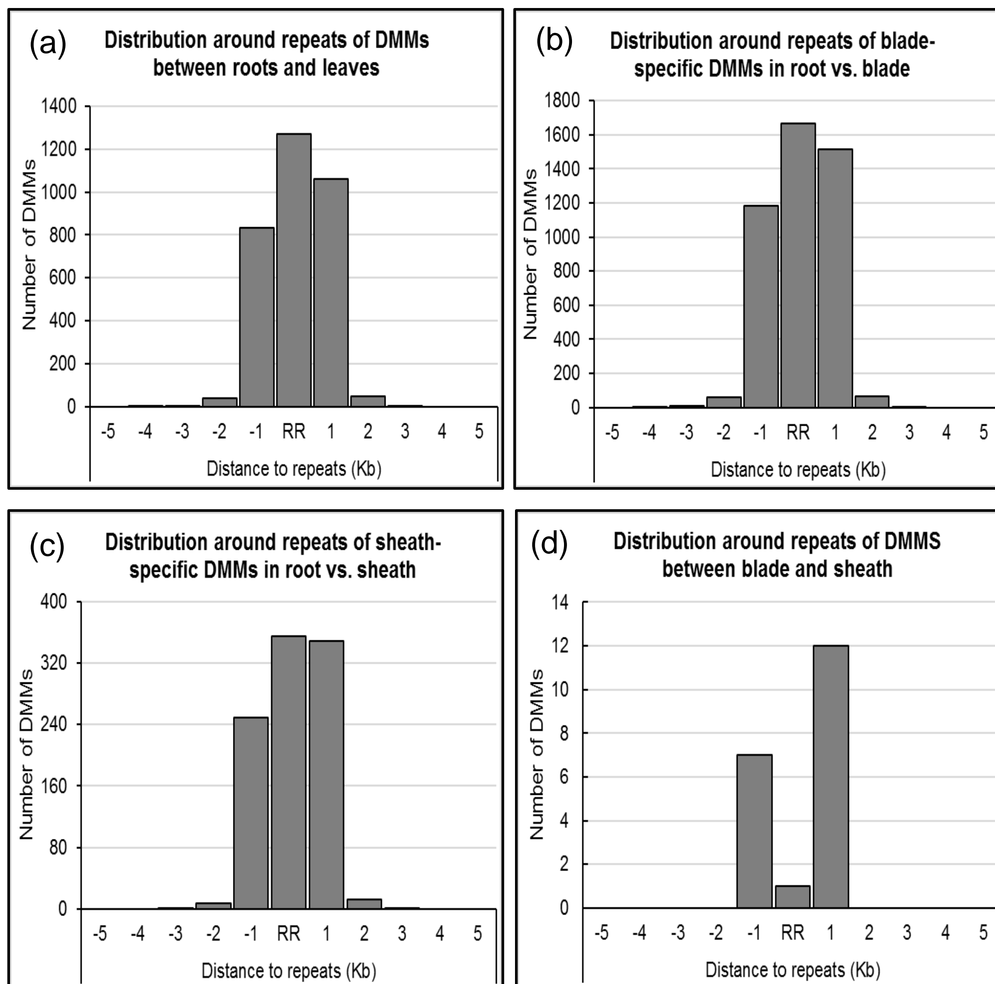
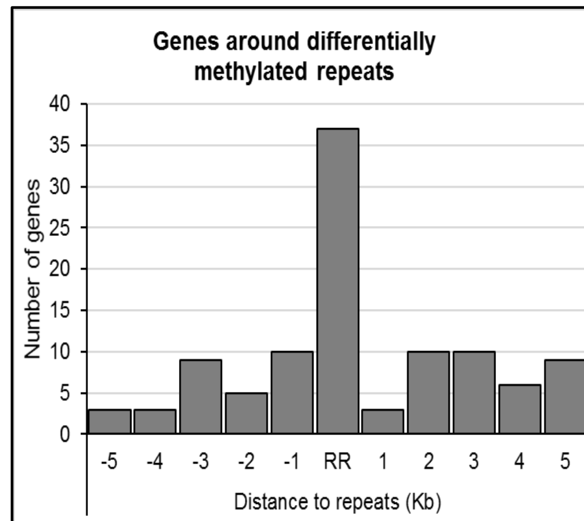


Figure 6: Distribution of tissue-specific differentially methylated markers (DMMs) around repeats. (a) DMMs between roots and leaves, present in both blades and sheaths as in Figure 2bc; (b) blade-specific DMMs between roots and leaves; and (c) sheath-specific DMMs between roots and leaves; (D) DMMs between blades and sheaths. The x axis indicates the distance to repeats in kilo base pairs (Kb) on both flanking regions. Negative and positive values indicate upstream and downstream repeat regions, respectively. RR, repeat regions. The y axis shows the number of DMMs per 1 Kb window.

Distribution of genes around differentially methylated (DM) repeats

To investigate a possible interaction between differentially methylated (DM) repeats and genes, the distance of genes from DM repeats between root and leaf samples was evaluated. In this way, we found 105 genes near repeats (up to 5 Kb either side), of which 37 overlapped with a repeat and the remaining genes were scattered up- and downstream from the repeat (Fig. 7). The number of DM repeats surrounded by genes thus represented only a tiny proportion of the total repeats that were differentially methylated between roots and leaves (105 out of 3266 DM repeats, 3.21%). Genes around DM repeats are listed in Table S2. About half of these genes near DM repeats (52 of 105 genes) were also differentially methylated, whereas the remainder (53 genes) were not.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1



238

239 **Figure 7**: Distribution of genes around differentially methylated repeat regions. The x axis indicates
240 the distance to repeats in kilo base pairs (Kb) on both flanking regions. Negative and positive values
241 indicate upstream and downstream repeat regions, respectively. RR, repeat regions. The y axis shows
242 the number of genes per 1 Kb window.

243 *Gene ontology of differentially methylated genes*

244 The 107 genes differentially methylated between root and leaf samples were described by 213
245 GO terms within the three main categories; "biological process", "cellular component" and
246 "molecular function", of which 121 GO terms were shared by both leaf parts, 88 were specific to
247 blades and 4 to sheaths (Table 3). Thus, it appeared that although most functions were shared by both
248 parts, leaf-blades carried a greater diversity of specific functions than did the leaf blades.
249 The GO analysis provided a picture of the role of DM genes in barley physiology and
250 metabolism. Genes that were differentially hypermethylated in leaves compared to roots related to
251 GO terms predominantly represented by "organonitrogen compound metabolism" and "generation
252 of precursor metabolites and energy" (Fig. 8a, Table S3-4). The top five GO term representatives of
253 hypomethylated genes in leaves relative to roots were; "organophosphate biosynthesis", "peptide
254 metabolism", "monovalent inorganic cation transport", "electron transport chain", and "generation
255 of precursor metabolites and energy". Similarly, it is also worth mentioning that photosynthesis-
256 associated GO terms (GO:0015979) were enriched among genes hypomethylated in leaf tissues. Also,
257 cellular components that set apart roots and leaves concerned chloroplast thylakoid (GO:0009534 and
258 GO:0009579) and cytochrome complex (GO:0070069), which are part of the photosynthetic
259 machinery, and were all derived from genes hypomethylated in leaves (Figure S1a, Supplemental
260 Data 3.S3). Furthermore, there was a high frequency of the GO term "plastid" (GO:0009536), which
261 was enriched in both hypermethylated and hypomethylated genes (Figure S1b-e, Supplemental Data
262 3.S3-4).
263 Some of the GO terms from differentially hypomethylated genes in leaves, were related to
264 molecular functions represented by; "tetrapyrrole binding", "monovalent inorganic cation
265 transmembrane transporter activity", "transition metal ion binding", "hydrolase activity" and
266 "quinone activity" (Fig. S2a, Table S2). While no GO term belonging to molecular function was
267 enriched by DM genes specific to sheaths, blade-specific DM genes enriched GO terms around
268 "monovalent inorganic cation transmembrane transporter activity", "ATPase activity coupled", and
269 "adenyl-ribonucleotide binding" (Fig. S2b, Table S3).

270 **Table 3**: Number of differentially methylated DM genes and associated gene ontology (GO) terms.
271 DM genes between roots and leaves common to both blade and sheath (Root vs. blade + sheath),
272 specific to blade (Root vs. blade specific) and specific to sheath (Root vs. sheath specific). Hyper and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

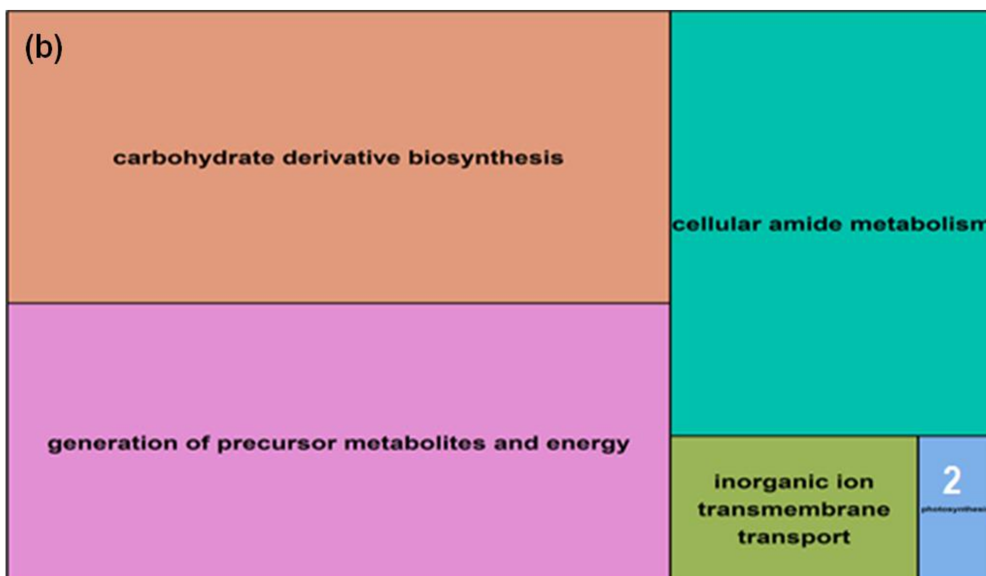
273 hypo refer to hypermethylation and hypomethylation in roots compared with the other tissue (blade
274 and sheath, respectively). GO terms were selected based on difference between their frequency in DM
275 genes and non-DM genes, with adjusted P-value < 0.01.

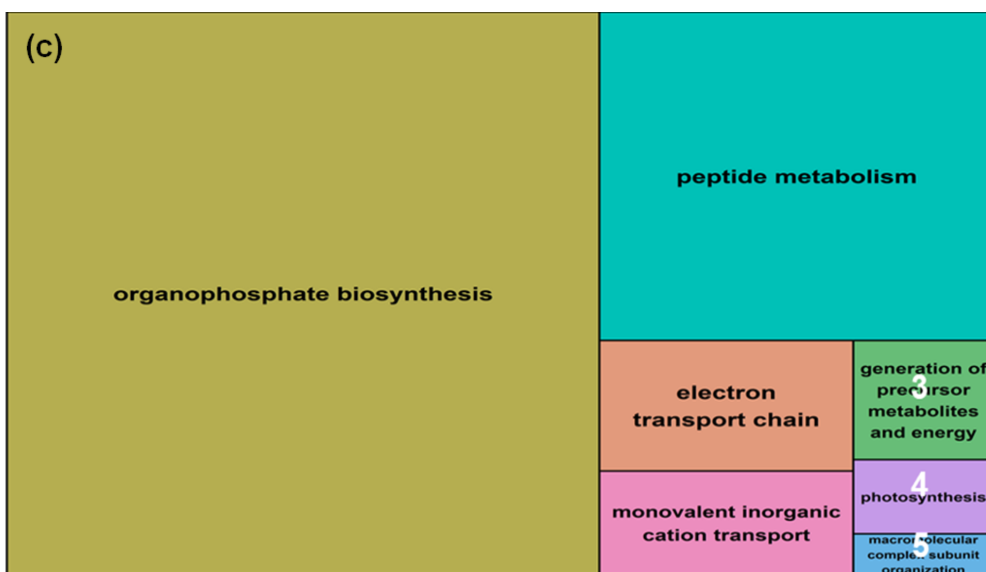
276

	DM genes			GO terms		
	Hyper	Hypo	*Total	Hyper	Hypo	*Total
Root vs. blade + sheath	10	51	61	23	100	123
Root vs. blade specific	3	41	44	3	85	88
Root vs. sheath specific	1	1	2	2	2	4
*Total	14	92	107	28	187	215

277 *Totals may include duplicates, since the same gene can be both hypo- and hypermethylated, and the same GO
278 term may be present in both groups.

doi:10.20944/preprints201804.0328.v1





Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

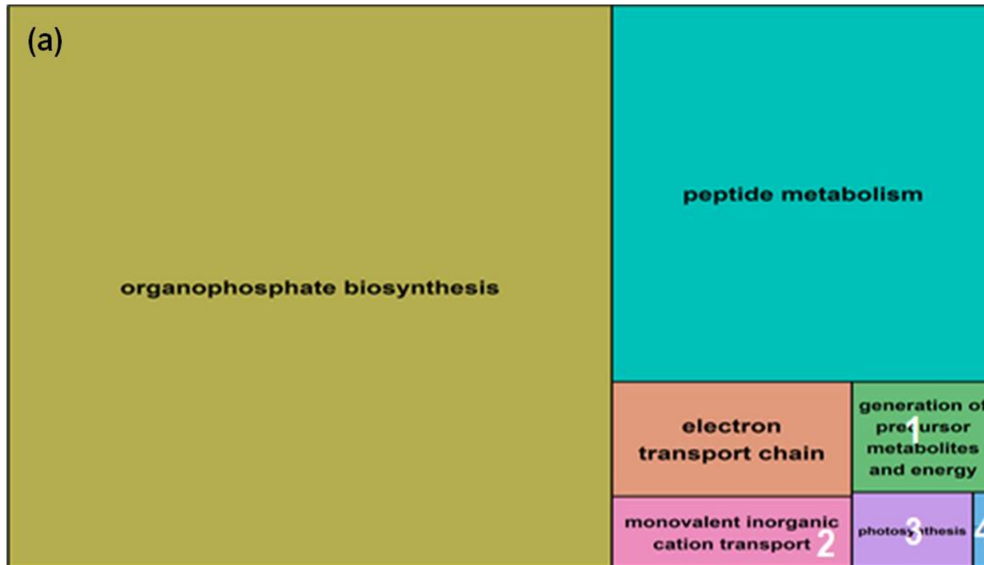
279 **Figure 8**: Summary treemaps of GO (gene ontology) term representatives for the category "biological
280 process" obtained from differentially methylated genes between roots and leaves. (a) Representatives
281 of GO terms enriched by common differentially hypermethylated genes in blades and sheaths; 1 =
282 generation of precursor metabolites and energy; (b) Representatives of GO terms enriched by
283 common differentially hypomethylated genes in blades and sheaths; 2= photosynthesis; (c)
284 Representatives of GO terms enriched by blade-specific differentially hypomethylated genes; 3 =
285 generation of precursor metabolites and energy, 4 = photosynthesis, 5 = macromolecular complex
286 subunit. Treemaps were constructed using R scripts produced by the REVIGO server
287 (http://revigo.irb.hr/). The detailed list of terms in the background of GO representatives is provided
288 in the Table S2-4.

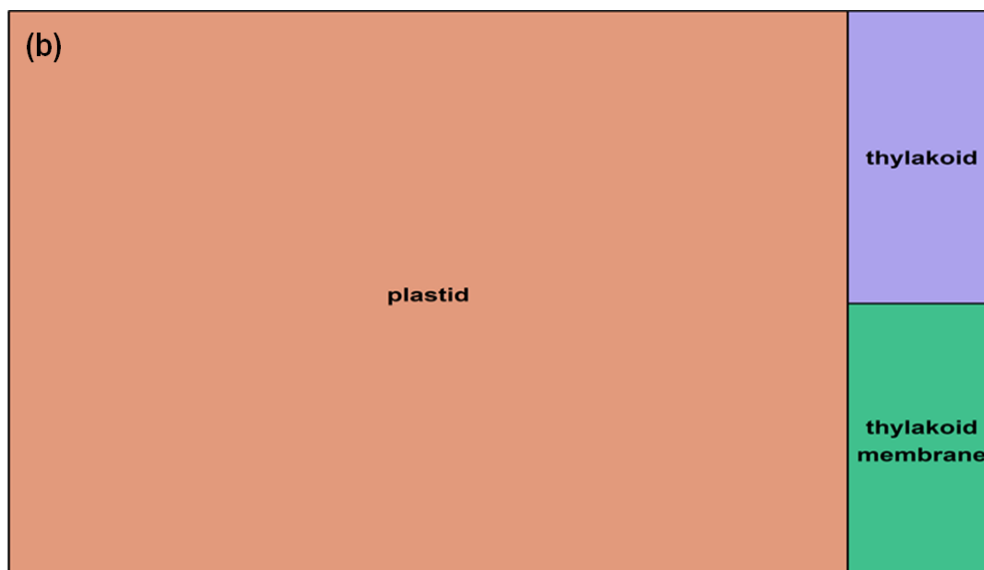
289

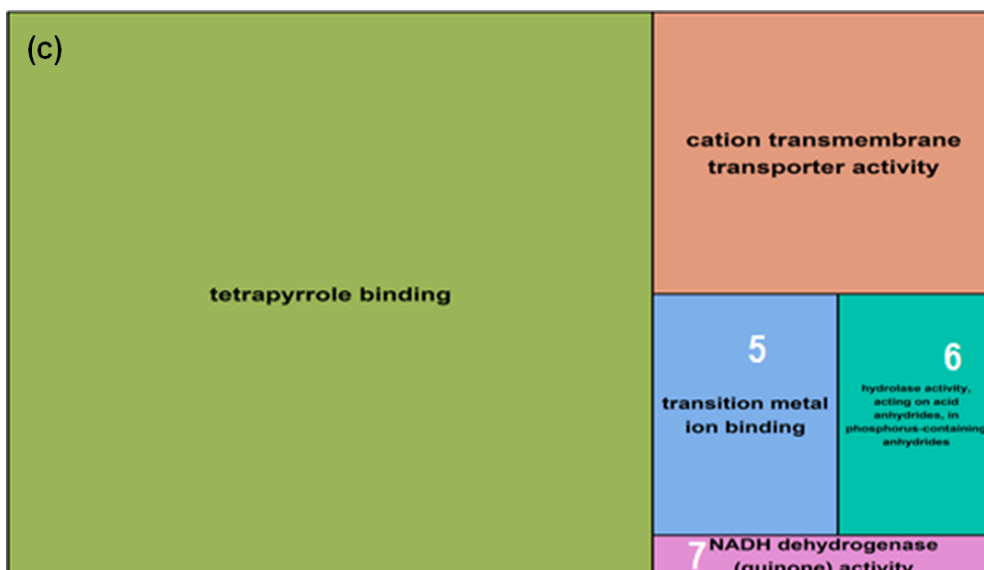
290 *Gene ontology of genes near differentially methylated repeats*

291 Since some of the genes around DM repeats were also differentially methylated and analysed as
292 such for GO enrichment above, only non-DM genes around DM repeats (53 of 105) were used for
293 further GO analysis. This analysis generated 97 significantly enriched GO terms in the three
294 categories; "biological process", "molecular function" and "cellular component". Strikingly, most of
295 the GO terms enriched by non-DM genes around DM repeats (93 of 97 GO terms, 95.88%) were also
296 enriched in DM genes.
297 The top GO term representatives in the "biological process" category were; "organophosphate
298 biosynthesis", "peptide metabolism", electron transport chain", "monovalent inorganic cation
299 transport", "generation of metabolites and energy" and "photosynthesis" (Fig. 9a). In the GO
300 category "cellular component", the GO term "plastid" predominated, along with "thylakoid" and
301 "thylakoid membrane" (Fig. 9b). GO terms enriched in the category "molecular function" belonged
302 to the following five sub-categories, in order of importance; "tetrapyrrole binding", "cation
303 transmembrane transporter activity", "transition metal ion binding", "hydrolase activity" and
304 "NADH dehydrogenase (quinone) activity" (Fig. 9c).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018
doi:10.20944/preprints201804.0328.v1







Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

305 **Figure 9**: Representative GO enrichment summary treemaps obtained from genes near DM repeats
306 between roots and leaves. (a) Representatives of GO terms enriched in the category "biological
307 process"; 1 = generation of precursor metabolites and energy; 2 = monovalent inorganic cation
308 transport; 3 = photosynthesis; 4 = macromolecular complex subunit organisation; (b) Representatives
309 of GO terms enriched in the category "cellular component"; (c) Representatives of GO terms enriched
310 in the category "molecular function"; 5 = transition metal ion binding; 6 = hydrolase activity, acting
311 on acid anhydrides, in phosphorus containing anhydrides; 7 = NADH dehydrogenase (quinone)
312 activity. Treemaps were constructed using R scripts produced by the REVIGO server
313 (http://revigo.irb.hr/).

314

315 **3. Discussion**

316 *Extensive epigenetic differentiation between roots and leaves*

317 In this study, we detected large numbers of DDMs between roots and leaves that were conserved
318 across diverse array of barley genotypes, and so were deemed far more likely to be organ-specific
319 than genotype-dependent. Of these, hypomethylation of the ^{m}CCGG motif predominated in leaves
320 (Fig. 2b,c, Fig. 3b and Fig. 4a). More surprisingly, we also detected similarly conserved DMMs
321 between leaf-blades and leaf-sheaths (Fig. 2e and Fig. 4b). The number of conserved DMMs between
322 blades and sheaths (20 DMMs), all hypermethylated in sheaths, was relatively consistent with the
323 closeness of these structures in position and function. These findings are broadly congruent with
324 previous studies, which reported differential DNA methylation between variable tissues (e.g.
325 endosperm, pollen, leaves, roots) in diverse plant species [7-10], but additionally hint that the
326 developmental closeness of structures being compared may also be reflected in the distinctiveness of
327 their methylation profiles. However, controversy over the extent and validity of organ-specific
328 DMMs [9,10,21-23] could cast doubt over their utility for organ diagnosis or as a tool to gain greater
329 insight into the genes responsible for organ development/identity. Here, we sought to mitigate
330 against the possibility of type I errors in DMM assignment through the unprecedented use of five
331 diverse varieties and five biological replicates of each variety in the identification of these marks. In
332 contrast to our findings, previous workers have reported little difference in the methylation levels of
333 both ^{m}CG and ^{m}CHG motifs between roots and leaves in Arabidopsis [9] and sorghum [10]. Further,
334 no significant difference was detected at all for ^{m}CG and ^{m}CHG methylation levels between tissues
335 in cotton [29]. These divergences may simply reflect genuine biological differences between
336 taxonomic groups. However, it is also important to recognise that such differences may also arise
337 from the approach used to identify organ-specific DMMs. Variability in the techniques used to assess
338 plant methylation profiles may introduce different forms of bias and preclude or complicate
339 comparison among studies. DMM detection can be influenced by factors such as 1) the genome
340 coverage of the methylation profiling method (low coverage methods such as MSAP are likely to
341 miss many markers [7], and 2) the data analysis approach used, which can compare either global
342 methylation levels (e.g. percent methylation [9] or methylated loci (e.g. DMMs [30]. We contend that
343 relying solely on global methylation levels can be misleading in comparing tissue profiles, because
344 similar methylation levels may show completely different patterns and so vital information content
345 is lost.
346 The current study revealed that tissue-specific DNA methylation occurred abundantly in the
347 ^{m}CHG context (at least ^{m}CCGGs) (Fig. 2a-c). This concurs with reports of the CHG context similarly
348 dominating differential DNA methylation between organs in *Brachypodium distachyon* [8] and
349 sorghum [10]. Although tissue-specific methylation also occurs in other cytosine contexts [10], our
350 results and other studies [10,22] suggest that ^{m}CCGG is a primary motif of epigenetic distinctiveness
351 of plant organs. Additionally, while tissue-specific DMMs were mostly hypomethylated in leaves
352 compared to roots in the present study (Fig. 3b), in Arabidopsis, Widman *et al.* [9] found that
353 hypermethylation prevailed in leaves compared with roots. This apparent contradiction in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

directionality of methylation in DMMs between roots and leaves may be a reflection a difference in the polarity of early divisions in the monocotyledonous barley and the dicotyledonous Arabidopsis embryos or else the methylation profiling method implemented.

DNA methylation flux is tissue specific during barley seedlings development

In addition to tissue-specificity of methylation profiles, one notable finding in the current study was that leaf cohorts exhibited a strong tendency to co-cluster. This suggests that the nature of methylation divergence between organs is not absolutely fixed and instead appears to change with developmental progression. This observation accords with previous reports that genome-wide methylation patterns are not static during plant development [31]. Additionally, a considerable portion of DMMs between roots and leaves was also specific to the leaf rank, due to the steady decrease in the number of DMMs between roots and leaf blades with the rank of the latter (Fig. 2a-c). In this case, therefore, the slow but progressive accumulation of additional methylation marks in the leaves increases their divergence from root profiles and enables the separation of leaf cohorts. However, the small number of DMMs distinguishing between leaf blades and leaf sheaths ran counter to this trend such that there were no DMMs capable of discrimination between these leaf parts among the oldest cohort studied (leaf 1) (Fig. 2d, Table 2). It seems intuitively improbable that older cohorts of leaves would simply lose differentiation between structurally distinct parts, especially if these marks had a functional role in defining function. Perhaps the most plausible biological explanation for the apparent erosion of divergence lies in the different chronological ages of the leaf cohorts that were sampled. Put simply, the third leaves were the least mature of the three cohorts collected and so it is entirely possible that the blade-sheath differential marks had yet to appear in these samples. Thus, it is important to consider the developmental and ageing progression chronology when assigning DMMs and that some organ- or structure-specific marks may only become organ-specific late in their development. Such late-emerging developmental DMMs should mean that the cumulative number of tissue-specific markers increases and so the organs or structures become more distinct, through leaf growth stages [32], each of which may carry a specific epigenetic profile. Certainly, others have noted that methylation profiles vary progressively as the organ develops [3,33,34] before reaching, at maturity, a "default" methylome, which may be conserved across varieties [24]. These results suggest that, once leaves are differentiated and mature, they do not show significant differences in DNA methylation profiles, regardless of their rank of appearance. Additionally, the location of half of the 20 common DMMs between blades and sheaths on chromosome 5H implies that this chromosome carries loci important for blade and sheath identities.

Tissue-specific DNA methylation preferably occurs in repeat regions of the barley genome

One noteworthy feature of the organ-specific DMMs identified here is that far more were associated with repeat regions than in or around genes. Nevertheless, the fact that 27 DMMs overlapped with exons and 10 were located in introns (Table S1) contradicts previous claims that CHG methylation marks are exclusively restricted to repeat regions and intergenic regions [20,21,35,36]. The possible regulatory significance of such gene body CHG methylation marks requires further investigation [37]. However, it is already well-established that tissue-specific DMMs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

can influence gene expression by enhancing gene transcription [9] and alternative splicing [38] or through repression due to immediate proximity to transcription start site [39].

The predominance of DMMs around and within repeats leads us to speculate that they could play an important role in defining organ identity in barley, and accords with previous findings in *Brachypodium distachyon* [8]. This flux of DNA methylation patterns in repeats [8,34,40] has been proposed to regulate [35] developmental shifts during plant growth and development [11,31]. Nevertheless, the association between DMMs in/around repeat regions and organ identity described here does not establish a causal link between the two. However, there are grounds for reasoning that this may be the case and that the possibility warrants further study. First, repeat regions were previously proposed to be involved in alternative promoters, a substantial proportion of which (>40%) was reported to shape tissue differentiation [16]. Therefore, tissue-specific DMMs in repeats may contribute to alternative promoters, and thus determine organ identity. Second, differential gene expression between roots and leaves [25,41] implies a firm regulatory system, including epigenetic mechanisms to guarantee tissue-specific cell development. Tissue-specific DMMs in repeats show that repeats are not that so-called "selfish parasites" of the genome [42], but can directly or indirectly affect tissue-specific gene expression [34,43,44]. Finally, it has been suggested that transposons coordinate splice variants, a genomic event that occurs in more than 60% of plant genes [45,46], thus generating multiple mRNA transcripts from a single gene [47,48]. Many splice variants are tissue-specific [49], suggesting that it is entirely possible that tissue-specific DMMs in repeats affect alternative splicing and subsequent gene expression. Also, some DM genes might potentially be regulated simultaneously by their own methylation and that of repeats [44,50], due to proximity with DM repeats.

Tissue-specific DMMs target genes relevant to plant tissue function

The widely noted observation that differential DNA methylation correlates with differential gene expression [10,14,51], led us to perform ontology analysis on DM genes and genes around DM repeats in the plant tissues. We reasoned that if a significant proportion of these marks were indeed regulatory, then their Gene Ontology classification profiles should broadly match those reported previously for these organs by transcript analysis. Our findings supported this tenet, with marked similarities noted between the overall GO classifications among DM-associated genes found here and those previously associated with expression divergence between roots and leaves. For instance, photosynthesis-associated GO terms (chloroplast thylakoid and cytochrome complex) were enriched in genes that were hypomethylated in leaves, which host photosynthesis (Simpson *et al.*, 1986). Similarly, genes involved in organonitrogen compound metabolism were hypomethylated in roots, and relate to processes occurring in roots, such as nitrogen assimilation [52,53]. The cellular component "plastid", known as tissue-specific in green plants [54], was highly enriched in DM genes.

Conclusions

This study has identified a series of DMMs that allow for the discrimination of leaf parts and roots in young barley plants. Considered collectively, our data illustrates that during tissue development, DNA methylation evolves to reach a default profile once the tissue is completely differentiated at maturity. It is possible that the plant organ formation and maturation is under at least partial control of DNA methylation changes. In addition, repeats could play an important role in tissue definition. Tissue specific mCCGG sites suggests that this context carries important factors of tissue differentiation. Expression analysis of tissue samples would perfectly demonstrate the role of tissue-specific DMMs in gene regulation. Nevertheless, mapping indexed markers generated from the ms-GBS with barley reference gene transcripts, allowed us to identify between tissues DM genes, which Ontology indicated that their tight correlation with expected tissue-specific functions is not coincidental.

4. Materials and Methods

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

446 *Plant material and growth conditions*

447 Five spring barley varieties (Barque 73, Flagship, Hindmarsh, Schooner and Yarra) were grown
448 in potting mix comprising 50% UC (University of California at Davis), 35% coco-peat and 15%
449 clay/loam (v v^{-1}) in 3.3 L pots, 17.5 cm deep, free-draining and placed on saucers. The experiment was
450 conducted from 30th January to 20th February 2015 in a greenhouse at the Waite Campus, University
451 of Adelaide, South Australia (34°58'11"S, 138°38'19"E). The seedlings were grown under natural
452 photoperiod while temperatures were set at 22°C/15°C (day/night). The experiment consisted of five
453 randomized blocks of five varieties (25 seedlings per block). Pots were watered to weight every 2
454 days to a gravimetric water content of 16.8% (w w^{-1}) (0.8 × field capacity) [55] until sampling 21 days
455 after sowing, when seedlings were at three-leaf stage (Zadok stage 13 [56]). Blades and sheaths of
456 leaves 1-3 were sampled separately. Leaves 1 and 2 were fully expanded prior to sampling, whilst
457 leaf 3 had just completed growth. About 50 mg of plant material was cut from the middle section of
458 each leaf blade and each leaf sheath and snap frozen in liquid nitrogen in 2 ml micro tubes. Roots
459 were cut from the seedlings and washed using tap water to remove soil particles, then blotted dry
460 with paper towels before sampling 50 mg of root tissue. Root samples were also frozen in liquid
461 nitrogen, and then all samples were stored at -80°C until DNA extraction.

462 *DNA isolation*

463 Prior to DNA extraction, frozen plant material was homogenized in a bead beater (2010-
464 Geno/Grinder, SPEX SamplePrep®, USA). DNA isolation was performed from pulverised plant
465 samples using a Qiagen DNeasy kit and following the manufacturer's instructions. DNA samples
466 were quantified using a NanoDrop® 1000 Spectrophotometer (V 3.8.1, ThermoFisher Scientific Inc.;
467 Australia) and concentrations were standardized to 10 ng μl^{-1} for subsequent library preparation.

468 *Methylation Sensitive genotyping by sequencing (ms-GBS)*

469 The ms-GBS was performed using a modified version [30,57] of the original GBS technique [58].
470 Genomic DNA was digested using the combination of a methylation-insensitive rare cutter, *Eco*RI
471 (GAATTC), and a frequent and methylation-sensitive cutter, MspI (CCGG). Each sample of DNA
472 was digested in a reaction volume of 20 µl containing 2 µl of NEB Smartcut buffer, 8 U of HF-*Eco*RI
473 (High-Fidelity) and 8 U of *Msp*I (New England BioLabs, Australia). The reaction was performed in a
474 BioRad 100 thermocycler at 37°C for 2 hours, followed by enzyme inactivation at 65°C for 10 min.
475 Then, the ligation of adapters to individual samples was achieved in the same plates by adding
476 0.1 pmol of the respective barcoded adapters with an *Msp*I cut site overhang, 15 pmol of the common
477 Y adapter with an *Eco*RI cut site overhang, 200 U of T4 Ligase and T4 Ligase buffer (New England
478 BioLabs, Australia) in a total volume of 40 µl. Ligation was carried out at 24°C for 2 hours followed
479 by an enzyme inactivation step at 65°C for 10 min.
480 DNA samples were allocated to plates, 81 samples each, including the negative control, water.
481 Prior to pooling plate samples into a single 81-plex library, the ligation products were individually
482 cleaned up to remove excess adapters using an Agencourt AMPure XP purification system (Beckman
483 Coulter, Australia) at a ratio of 0.85 (AMPure magnetic beads/ligation product), following the
484 manufacturer's instructions. Individual GBS libraries were produced by pooling 25 ng of DNA from
485 each sample. Each constructed library was then amplified in eight separate PCR (25 µl each)
486 containing 10 µl of library DNA, 5 µl of 5x Q5 high fidelity buffer, 0.25 µl polymerase Q5 high fidelity,
487 1 µl each of Forward and Reverse common primers at 10 µM, 0.5 µl of 10 µM dNTP and 7.25 µl of
488 sterile pure water. PCR amplification was performed in a BioRad T100 thermocycler, consisting of
489 DNA denaturation at 98°C (30 s) and ten cycles of 98°C (30 s), 62°C (20 s) and 72°C (30 s), followed
490 by 72°C for 5 min. PCR products were next pooled to reconstitute libraries. DNA fragments between
491 200 and 350 bp in size were captured using AMPure XP magnetic beads following the manufacturer's
492 instructions. Bead-captured fragments were eluted in 35 µl of water, of which 30 µl were collected in
493 a new labelled microtube. Libraries were next paired-end sequenced in an Illumina HiSeq 2500
494 (Illumina Inc., USA) at the Australian Genome Research Facility (AGRF, Melbourne node, Australia).

495 *Principal component – linear discriminant analysis*

496 Grouping of organ type samples was explored by performing a principal component – linear
497 discriminant analysis (PC-LDA) and a hierarchical cluster analysis using the R package *FIEmspro 1.1-*
498 *0* [28]. To visualise the results, we performed a scatter plot of the first two discriminant factors (DFs),
499 a 3D plot using the first three DFs and a hierarchical cluster tree based on Mahalanobis distance [27].

500 *Detection of DMMs in barley*

501 Differentially methylated DNA was assessed in ^{m}CCGG motifs (recognised by *Msp*I), between
502 barley leaf parts (blade and sheath) and roots. To do so, samples were grouped according to organ
503 type (root, blade and sheath) regardless of the genotype of origin, making 25 samples per organ. This
504 approach aimed to minimise genotype-dependent methylation markers. DMMs were identified
505 using the package, *msgbsR,* developed by Mayne *et al.* [59]. DMMs were selected based on Bonferroni
506 adjusted P-values with a threshold of 0.05 [60,61]. The significance of the marker also fulfilled the
507 condition that the read counts reached at least 1 CPM (count per million reads) and was present in at
508 least 20 samples per organ type (maximum sample per group = 25). The *log*FC (logarithm 2 of fold-
509 change) was computed to estimate the intensity and directionality of differential DNA methylation
510 between tissues. Determining the directionality of DNA methylation uses the fold change as an
511 inverse proxy for change in the methylation level. That is, higher methylation levels on a specific
512 locus will reduce the number of *Msp*I restriction products and therefore reduce the number of
513 sequences generated for that locus [62].

514 *Distribution of DMMs around genomic features and gene ontology*

515 To test whether there was a relationship between tissue-specific DMMs and particular genomic
516 features (e.g., genes and repeat regions as defined in Ensembl database
517 (http://plants.ensembl.org/biomart/martview/), DMM distribution was assessed in the barley
518 genome. Therefore, DMMs stable between tissues were mapped to the barley reference genome.
519 Then, the number of DMMs within genomic features (repeats, genes, exons, UTRs and tRNA genes)
520 and per 1 Kb bins within 5 Kb flanking regions [24,30] was tallied, using the shell module, *bedtools*
521 */2.22.0* [63]. Furthermore, the functions of tissue-specific DM genes and genes near DM repeats were
522 explored by performing gene ontology analysis. Genes within 5 Kb of a DMM were deemed
523 differentially methylated genes (DM) genes, and were grouped in hypermethylated and
524 hypomethylated genes per tissue. These genes were next used separately for Gene Ontology (GO)
525 terms enrichment, using two R packages: *GO.db* and *annotate* [64,65]. Significant GO terms were
526 selected based on Bonferroni adjusted P-values [60] at a significance threshold of 0.05 and a total GO
527 enrichment of DM and non-DM genes at least equal to 10. To visualize the results of GO analysis,
528 treemaps were generated in REVIGO [66], using default settings.

529 **Supplementary Materials:** The following are available online at www.mdpi.com/link, Figure S1: Summary
530 treemaps for GO category "cellular components" from DM genes between roots and leaves. Figure S2: Summary
531 treemaps for GO category "molecular function" from DM genes between roots and leaves. Table S1: List of
532 differentially methylated exons. Table S2: List of GO terms enriched by differentially methylated genes between
533 roots and leaves. Table S3: List of GO terms enriched by differentially methylated genes between roots and leaves
534 and specific to blades. Table S4: List of GO terms enriched by differentially methylated genes between roots and
535 leaves and specific to sheaths.

536 **Acknowledgments:** The authors are grateful to the Australian Agency for International Development (AusAID)
537 for supporting MK with an Australian Awards Scholarship. MJW was paid by the BBSRC (BBS/E/W/0012843C).

538 **Author Contributions:** M.K. conceived and performed the experiments, analysed the data and wrote the
539 manuscript; B.J.M. performed ms-GBS data alignments; M.J.W., E.S.S., B.B., C.M.R.L. conceived the experiments
540 and supervised the work. All authors read and commented on the manuscript.

541 **Conflicts of Interest:** The authors declare that they have no conflicts of interest.

542 **References**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

543 1. Bird, A. Perceptions of epigenetics. *Nature* 2007, *447*, 396-398.
544 2. Zilberman, D.; Henikoff, S. Genome-wide analysis of DNA methylation patterns. *Development* 2007, *134*,
545 3959-3965.
546 3. Cokus, S.J.; Feng, S.; Zhang, X.; Chen, Z.; Merriman, B.; Haudenschild, C.D.; Pradhan, S.; Nelson, S.F.;
547 Pellegrini, M.; Jacobsen, S.E. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA
548 methylation patterning. *Nature* 2008, *452*, 215-219.
549 4. Bossdorf, O.; Arcuri, D.; Richards, C.; Pigliucci, M. Experimental alteration of DNA methylation affects the
550 phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. *Evolutionary Ecology* 2010, *24*, 541-
551 553.
552 5. Finnegan, E.J.; Peacock, W.J.; Dennis, E.S. Reduced DNA methylation in *Arabidopsis thaliana* results in
553 abnormal plant development. *Proceedings of the National Academy of Sciences* 1996, *93*, 8449-8454.
554 6. Aceituno, F.F.; Moseyko, N.; Rhee, S.Y.; Gutiérrez, R.A. The rules of gene expression in plants: organ
555 identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*.
556 *BMC Genomics* 2008, *9*, 1-14.
557 7. Rodríguez López, C.M.; Wetten, A.C.; Wilkinson, M.J. Progressive erosion of genetic and epigenetic
558 variation in callus-derived cocoa (*Theobroma cacao*) plants. *New Phytologist* 2010, *186*, 856-868.
559 8. Roessler, K.; Takuno, S.; Gaut, B.S. CG methylation covaries with differential gene expression between Leaf
560 and Floral Bud Tissues of *Brachypodium distachyon*. *PLoS One* 2016, *11*, e0150002.
561 9. Widman, N.; Feng, S.; Jacobsen, S.E.; Pellegrini, M. Epigenetic differences between shoots and roots in
562 *Arabidopsis* reveals tissue-specific regulation. *Epigenetics* 2014, *9*, 236-242.
563 10. Zhang, M.; Xu, C.; von Wettstein, D.; Liu, B. Tissue-specific differences in cytosine methylation and their
564 association with differential gene expression in sorghum. *Plant Physiology* 2011, *156*, 1955-1966.
565 11. Ay, N.; Janack, B.; Humbeck, K. Epigenetic control of plant senescence and linked processes. *Journal of*
566 *Experimental Botany* 2014, *65*, 3875-3887.
567 12. Xiao, W.; Custard, K.D.; Brown, R.C.; Lemmon, B.E.; Harada, J.J.; Goldberg, R.B.; Fischer, R.L. DNA
568 methylation is critical for *Arabidopsis* embryogenesis and seed viability. *The Plant Cell* 2006, *18*, 805-814.
569 13. Cocciolone, S.M.; Chopra, S.; Flint-Garcia, S.A.; McMullen, M.D.; Peterson, T. Tissue-specific patterns of a
570 maize Myb transcription factor are epigenetically regulated. *The Plant Journal* 2001, *27*, 467-478.
571 14. Song, Q.-X.; Lu, X.; Li, Q.-T.; Chen, H.; Hu, X.-Y.; Ma, B.; Zhang, W.-K.; Chen, S.-Y.; Zhang, J.-S. Genome-
572 wide analysis of DNA methylation in soybean. *Molecular Plant* 2013, *6*, 1961-1974.
573 15. Lafon-Placette, C.; Faivre-Rampant, P.; Delaunay, A.; Street, N.; Brignolas, F.; Maury, S. Methylome of
574 DNase I sensitive chromatin in *Populus trichocarpa* shoot apical meristematic cells: a simplified approach
575 revealing characteristics of gene-body DNA methylation in open chromatin state. *New Phytologist* 2013, *197*,
576 416-430.
577 16. Batut, P.; Dobin, A.; Plessy, C.; Carninci, P.; Gingeras, T.R. High-fidelity promoter profiling reveals
578 widespread alternative promoter usage and transposon-driven developmental gene expression. *Genome*
579 *Research* 2013, *23*, 169-180.
580 17. Maunakea, A.K.; Nagarajan, R.P.; Bilenky, M.; Ballinger, T.J.; D/'Souza, C.; Fouse, S.D.; Johnson, B.E.; Hong,
581 C.; Nielsen, C.; Zhao, Y., *et al.* Conserved role of intragenic DNA methylation in regulating alternative
582 promoters. *Nature* 2010, *466*, 253-257.
583 18. Sørensen, M.B.; Müller, M.; Skerritt, J.; Simpson, D. Hordein promoter methylation and transcriptional
584 activity in wild-type and mutant barley endosperm. *Molecular and General Genetics* 1996, *250*, 750-760.
585 19. Zhang, X.; Yazaki, J.; Sundaresan, A.; Cokus, S.; Chan, S.W.L.; Chen, H.; Henderson, I.R.; Shinn, P.;
586 Pellegrini, M.; Jacobsen, S.E., *et al.* Genome-wide high-resolution mapping and functional analysis of DNA
587 methylation in *Arabidopsis*. *Cell* 2006, *126*, 1189-1201.
588 20. Illingworth, R.; Kerr, A.; DeSousa, D.; Jørgensen, H.; Ellis, P.; Stalker, J.; Jackson, D.; Clee, C.; Plumb, R.;
589 Rogers, J., *et al.* A novel CpG island set identifies tissue-specific methylation at developmental gene loci.
590 *PLoS Biology* 2008, *6*, e22.
591 21. Hsieh, T.-F.; Ibarra, C.A.; Silva, P.; Zemach, A.; Eshed-Williams, L.; Fischer, R.L.; Zilberman, D. Genome-
592 wide demethylation of *Arabidopsis* endosperm. *Science* 2009, *324*, 1451-1454.
593 22. Ibarra, C.A.; Feng, X.; Schoft, V.K.; Hsieh, T.-F.; Uzawa, R.; Rodrigues, J.A.; Zemach, A.; Chumak, N.;
594 Machlicova, A.; Nishimura, T., *et al.* Active DNA demethylation in plant companion cells reinforces
595 transposon methylation in gametes. *Science* 2012, *337*, 1360-1364.

596 23. Zemach, A.; Kim, M.Y.; Silva, P.; Rodrigues, J.A.; Dotson, B.; Brooks, M.D.; Zilberman, D. Local DNA
597 hypomethylation activates genes in rice endosperm. *Proceedings of the National Academy of Sciences* 2010, *107*,
598 18729-18734.
599 24. Eichten, S.R.; Vaughn, M.W.; Hermanson, P.J.; Springer, N.M. Variation in DNA methylation patterns is
600 more common among maize inbreds than among tissues. *The Plant Genome* 2013, *6*, 1-10.
601 25. Mayer, K.F.X.; Waugh, R.; Brown, J.W.S.; Schulman, A.; Langridge, P.; Platzer, M.; Fincher, G.B.;
602 Muehlbauer, G.J.; Sato, K.; Close, T.J., *et al.* A physical, genetic and functional sequence assembly of the
603 barley genome. *Nature* 2012, *491*, 711-716.
604 26. Brockman, W.; Alvarez, P.; Young, S.; Garber, M.; Giannoukos, G.; Lee, W.L.; Russ, C.; Lander, E.S.;
605 Nusbaum, C.; Jaffe, D.B. Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome*
606 *Research* 2008, *18*, 763-770.
607 27. Mahalanobis, P.C. In *On the generalised distance in statistics*, Proceedings National Institute of Science, India,
608 1936; pp 49-55.
609 28. Enot, D.P.; Lin, W.; Beckmann, M.; Parker, D.; Overy, D.P.; Draper, J. Preprocessing, classification modeling
610 and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data.
611 *Nature Protocols* 2008, *3*, 446-470.
612 29. Osabe, K.; Clement, J.D.; Bedon, F.; Pettolino, F.A.; Ziolkowski, L.; Llewellyn, D.J.; Finnegan, E.J.; Wilson,
613 I.W. Genetic and DNA methylation changes in cotton (*Gossypium*) genotypes and tissues. *PLoS One* 2014,
614 *9*, e86049.
615 30. Kitimu, S.R.; Taylor, J.; March, T.J.; Tairo, F.; Wilkinson, M.J.; Rodriguez Lopez, C.M. Meristem
616 micropropagation of cassava (*Manihot esculenta*) evokes genome-wide changes in DNA methylation.
617 *Frontiers in Plant Science* 2015, *6*, 1-12.
618 31. Zhong, S.; Fei, Z.; Chen, Y.-R.; Zheng, Y.; Huang, M.; Vrebalov, J.; McQuinn, R.; Gapper, N.; Liu, B.; Xiang,
619 J., *et al.* Single-base resolution methylomes of tomato fruit development reveal epigenome modifications
620 associated with ripening. *Nature Biotechnology* 2013, *31*, 154-159.
621 32. Candaele, J.; Demuynck, K.; Mosoti, D.; Beemster, G.T.S.; Inze, D.; Nelissen, H. Differential methylation
622 during maize leaf growth targets developmentally regulated genes. *Plant Physiology* 2014, *164*, 1350-1364.
623 33. Brandeis, M.; Ariel, M.; Cedar, H. Dynamics of DNA methylation during development. *BioEssays* 1993, *15*,
624 709-713.
625 34. Zhang, X. The epigenetic landscape of plants. *Science* 2008, *320*, 489-492.
626 35. Bewick, A.J.; Ji, L.; Niederhuth, C.E.; Willing, E.-M.; Hofmeister, B.T.; Shi, X.; Wang, L.; Lu, Z.; Rohr, N.A.;
627 Hartwig, B., *et al.* On the origin and evolutionary consequences of gene body DNA methylation. *Proceedings*
628 *of the National Academy of Sciences* 2016, *113*, 9111-9116.
629 36. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes & Development* 2011, *25*, 1010-
630 1022.
631 37. Xie, H.; Konate, M.; Sai, N.; Tesfamicael, K.G.; Cavagnaro, T.; Gilliham, M.; Breen, J.; Metcalfe, A.; Stephen,
632 J.; DeBei, R., *et al.* Global DNA Methylation Patterns Can Play a Role in Defining Terroir in Grapevine (*Vitis*
633 *vinifera* cv. Shiraz). *Frontiers in Plant Science* **2017**, *8*.
634 38. Li-Byarlay, H.; Li, Y.; Stroud, H.; Feng, S.; Newman, T.C.; Kaneda, M.; Hou, K.K.; Worley, K.C.; Elsik, C.G.;
635 Wickline, S.A., *et al.* RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative
636 splicing in the honey bee. *Proceedings of the National Academy of Sciences* 2013, *110*, 12750-12755.
637 39. Jones, P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews*
638 *Genetics* 2012, *13*, 484-492.
639 40. Hollister, J.D.; Gaut, B.S. Epigenetic silencing of transposable elements: A trade-off between reduced
640 transposition and deleterious effects on neighboring gene expression. *Genome Research* 2009, *19*, 1419-1428.
641 41. 40. Druka, A.; Muehlbauer, G.; Druka, I.; Caldo, R.; Baumann, U.; Rostoks, N.; Schreiber, A.; Wise, R.;
642 Close, T.; Kleinhofs, A., *et al.* An atlas of gene expression from seed to seed through barley development.
643 *Functional & Integrative Genomics* 2006, *6*, 202-211.
644 42. Orgel, L.E.; Crick, F.H.C. Selfish DNA: the ultimate parasite. *Nature* 1980, *284*, 604-607.
645 43. Lister, R.; O'Malley, R.C.; Tonti-Filippini, J.; Gregory, B.D.; Berry, C.C.; Millar, A.H.; Ecker, J.R. Highly
646 integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 2008, *133*, 523-536.
647 44. Hirsch, C.D.; Springer, N.M. Transposable element influences on gene expression in plants. *Biochimica et*
648 *Biophysica Acta - Gene Regulatory Mechanisms* 2016, *1860*, 157–165.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 April 2018 doi:10.20944/preprints201804.0328.v1

649 45. Marquez, Y.; Brown, J.W.S.; Simpson, C.; Barta, A.; Kalyna, M. Transcriptome survey reveals increased
650 complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Research* 2012, *22*, 1184-1195.
651 46. Simpson, Craig G.; Lewandowska, D.; Fuller, J.; Maronova, M.; Kalyna, M.; Davidson, D.; McNicol, J.;
652 Raczynska, D.; Jarmolowski, A.; Barta, A., *et al.* Alternative splicing in plants. *Biochemical Society*
653 *Transactions* 2008, *36*, 508-510.
654 47. Barbazuk, W.B.; Fu, Y.; McGinnis, K.M. Genome-wide analyses of alternative splicing in plants:
655 Opportunities and challenges. *Genome Research* 2008, *18*, 1381-1392.
656 48. Warf, M.B.; Berglund, J.A. The role of RNA structure in regulating pre-mRNA splicing. *Trends in biochemical*
657 *sciences* 2010, *35*, 169-178.
658 49. Pan, Q.; Shai, O.; Lee, L.J.; Frey, B.J.; Blencowe, B.J. Deep surveying of alternative splicing complexity in
659 the human transcriptome by high-throughput sequencing. *Nature Genetics* 2008, *40*, 1413-1415.
660 50. Lorincz, M.C.; Dickerson, D.R.; Schmitt, M.; Groudine, M. Intragenic DNA methylation alters chromatin
661 structure and elongation efficiency in mammalian cells. *Nature Structural Molecular Biology* 2004, *11*, 1068-
662 1075.
663 51. Johnston, R.; Wang, M.; Sun, Q.; Sylvester, A.W.; Hake, S.; Scanlon, M.J. Transcriptomic analyses indicate
664 that maize ligule development recapitulates gene expression patterns that occur during lateral organ
665 initiation. *The Plant Cell* 2014, *26*, 4718-4732.
666 52. Glass, A.D.M.; Shaff, J.E.; Kochian, L.V. Studies of the uptake of nitrate in barley: IV. Electrophysiology.
667 *Plant Physiology* 1992, *99*, 456-463.
668 53. Tischner, R. Nitrate uptake and reduction in higher and lower plants. *Plant, Cell & Environment* 2000, *23*,
669 1005-1024.
670 54. Wise, R.R. The diversity of plastid form and function. In *The Structure and Function of Plastids*, Wise, R.R.;
671 Hoober, J.K., Eds. Springer, the Netherlands: Dordrecht, 2006; pp 3-26.
672 55. Berger, B.; Regt, B.; Tester, M. Trait dissection of salinity tolerance with plant phenomics. In *Plant Salt*
673 *Tolerance*, Shabala, S.; Cuin, T.A., Eds. Humana Press: 2012; Vol. 913, pp 399-413.
674 56. Zadoks, J.C.; Chang, T.T.; Konzak, C.F. A decimal code for the growth stages of cereals. *Weed Research* 1974,
675 *14*, 415-421.
676 57. Xia, Z.; Zou, M.; Zhang, S.; Feng, B.; Wang, W. AFSM sequencing approach: a simple and rapid method for
677 genome-wide SNP and methylation site discovery and genetic mapping. *Scientific Reports* 2014, *4*, 7300-
678 7308.
679 58. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.-L. Development of high-density genetic maps for barley
680 and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 2012, *7*, e32253.
681 59. Mayne, B.T.; Leemaqz, S.Y.; Buckberry, S.; Rodriguez Lopez, C.M.; Roberts, C.T.; Bianco-Miotto, T.; Breen,
682 J. msgbsR: An R package for analysing methylation-sensitive restriction enzyme sequencing data. *Scientific*
683 *Reports* 2018, *8*, 2190.
684 60. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to
685 multiple testing. *Journal of the royal statistical society. Series B (Methodological)* 1995, 289-300.
686 61. Dunn, O.J. Multiple comparisons among means. *Journal of the American Statistical Association* 1961, *56*, 52-
687 64.
688 62. Rodríguez López, C.M.; Morán, P.; Lago, F.; Espiñeira, M.; Beckmann, M.; Consuegra, S. Detection and
689 quantification of tissue of origin in salmon and veal products using methylation sensitive AFLPs. *Food*
690 *Chemistry* 2012, *131*, 1493-1498.
691 63. Quinlan, A.R.; Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features.
692 *Bioinformatics* 2010, *26*, 841-842.
693 64. Gentleman, R. annotate: Annotation for microarrays. R package version 1.52.0.
694 https://www.bioconductor.org/packages/release/bioc/html/annotate.html (4 March 2017),
695 65. Carlson, M. GO.db: A set of annotation maps describing the entire gene ontology. R package version 3.4.0.
696 http://bioconductor.org/packages/release/data/annotation/html/GO.db.html (4 March 2017),
697 66. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology
698 terms. *PLoS One* 2011, *6*, e21800.