

The lncRNA MARS modulates the epigenetic reprogramming of the marneral cluster in response to ABA

Thomas Roulé, Federico Ariel, Caroline Hartmann, Jose Gutierrez-Marcos, Nosheen Hussain, Martin Crespi, Thomas Blein

▶ To cite this version:

Thomas Roulé, Federico Ariel, Caroline Hartmann, Jose Gutierrez-Marcos, Nosheen Hussain, et al.. The lncRNA MARS modulates the epigenetic reprogramming of the marneral cluster in response to ABA. 2020. hal-03016338

HAL Id: hal-03016338 https://hal.science/hal-03016338

Preprint submitted on 20 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 The IncRNA *MARS* modulates the epigenetic reprogramming of the marneral 2 cluster in response to ABA

Thomas Roulé¹², Federico Ariel³, Caroline Hartmann¹², Jose Gutierrez-Marcos⁴, Nosheen
Hussain⁴, Martin Crespi^{12*} and Thomas Blein¹²

¹Institute of Plant Sciences Paris-Saclay, Centre Nationale de la Recherche, Institut National
de la Recherche Agronomique, Université Evry, Université Paris-Saclay, 91405 Orsay,
France

- 8 ²Institute of Plant Sciences Paris-Saclay, Université de Paris, 91405 Orsay, France
- 9 ³ Instituto de Agrobiotecnología del Litoral, CONICET, FBCB, Universidad Nacional del
- 10 Litoral, Colectora Ruta Nacional 168 km 0, 3000 Santa Fe, Argentina
- 11 4 School of Life Sciences, University of Warwick, Coventry, UK
- 12 *Correspondence to: MC (martin.crespi@universite-paris-saclay.fr)

13 ABSTRACT

Clustered organization of biosynthetic non-homologous genes is emerging as a characteristic 14 15 feature of plant genomes. The co-regulation of clustered genes seems to largely depend on epigenetic reprogramming and three-dimensional chromatin conformation. Here we identified 16 the long noncoding RNA (IncRNA) MARneral Silencing (MARS), localized inside the 17 Arabidopsis marneral cluster, and which controls the local epigenetic activation of its 18 19 surrounding region in response to ABA. MARS modulates the POLYCOMB REPRESSIVE COMPLEX 1 (PRC1) component LIKE-HETEROCHROMATIN PROTEIN 1 (LHP1) binding 20 throughout the cluster in a dose-dependent manner, determining H3K27me3 deposition and 21 22 chromatin condensation. In response to ABA, MARS decoys LHP1 away from the cluster and 23 promotes the formation of a chromatin loop bringing together the MARNERAL SYNTHASE 1 (MRN1) locus and a distal ABA-responsive enhancer. The enrichment of co-regulated 24 25 IncRNAs in clustered metabolic genes suggests that the acquisition of noncoding 26 transcriptional units constitute an additional regulatory layer driving the evolution of biosynthetic pathways. 27

KEYWORDS: IncRNA, enhancer, cluster, chromatin conformation, LHP1, ABA, seed
 germination, epigenetics, marneral

30 INTRODUCTION

In eukaryotes, functionally related genes are usually scattered across the genome. However, a growing number of operon-like clustered organization of non-homologous genes participating in common metabolic pathways point at an emerging feature of animal, fungi and plant genomes¹.

In plants, synthesis of numerous secondary metabolic compounds is important for the 35 dynamic interaction with their environment, affecting their life and survival². Terpenoids are 36 bioactive molecules of diverse chemical structure³. In Arabidopsis thaliana, the biosynthesis 37 of four triterpenes, namely thalianol⁴, tirucalla-7,24-dien-3b-ol⁵, arabidiol⁶ and marneral⁷, is 38 governed by enzymes encoded by genes organized in clusters⁸. The thalianol and marneral 39 related genes are located in the smallest metabolic clusters identified in plants to date, each 40 41 being less than 40kb in size⁸. Both compounds are derived from 2,3-oxidosqualene and the 42 corresponding gene clusters contain the oxidosqualene cyclases (OSCs), thalianol synthase (THAS) and marneral synthase (MRN1), respectively. The marneral cluster includes two 43 additional protein-coding genes, CYP705A12 and CYP71A16, participating in marneral 44 $oxidation^7$. 45

Growing evidence indicates that the co-regulation of clustered genes relies on 46 epigenetic mechanisms. It has been shown that the deposition of the histone variant H2A.Z 47 positively correlates with transcriptionally active clusters. Accordingly, nucleosome stability 48 precluding gene expression is dependent on ARP6, a component of the SWR1 chromatin 49 remodeling complex required for the deposition of H2A.Z into nucleosomes⁹. Additionally, it 50 51 was shown that the thalianol and marneral clusters exhibit increased expression in the 52 Polycomb mutant curly leaf (clf) with compromised H3K27me3 deposition, and reduced expression in the trithorax-group protein mutant pickle (pkl), a positive regulator that 53 counteracts H3K27me3 silencing¹⁰. Strikingly, it has been recently shown that biosynthetic 54 gene clusters are embedded in local hot spots of three-dimensional (3D) contacts that 55 segregate cluster regions from the surrounding chromosome environment in a tissue-56 dependent manner. Notably, H3K27me3 appeared as a central feature of the 3D domains at 57 silenced clusters¹¹. 58

Long noncoding RNAs (IncRNAs) have emerged as important regulators of eukaryotic
 gene expression at different levels¹². In plants, several IncRNAs have been shown to interact
 with the Polycomb Repressive Complex 1 and 2 components LIKE HETEROCHROMATIN
 PROTEIN 1 (LHP1) and CLF, respectively, which are related to H3K27me3 distribution^{13,14}.
 Furthermore, it has been proposed that IncRNAs can modulate the transcriptional activity of

neighboring genes by shaping local 3D chromatin conformation^{15–17}. Here we show that the 64 marneral cluster in Arabidopsis includes three noncoding transcriptional units. Among them, 65 the IncRNA MARS influences the expression of marneral cluster genes in response to ABA 66 through modification of the epigenetic landscape. MARS deregulation affects H3K27me3 67 distribution, LHP1 deposition and chromatin condensation throughout the cluster. 68 Furthermore, an ABA responsive chromatin loop dynamically regulates MRN1 transcriptional 69 activation by bringing together the MRN1 proximal promoter and an enhancer element 70 enriched in ABA-related transcription factors (TF) binding sites. MARS-mediated control of 71 72 the marneral cluster affects seed germination in response to ABA. The general co-regulation 73 of genes located within IncRNA-containing clusters in Arabidopsis points to noncoding 74 transcription as an important feature in coordinated transcriptional activity of clustered loci.

75 **RESULTS**

The marneral gene cluster includes three noncoding transcriptional units

The small marneral cluster includes three genes: marneral synthase (*MRN1*), CYP705A12 and CYP71A16 that *are* two P450 cytochrome-encoding genes (**Fig. 1a**), all participating in the biosynthesis and metabolism of the triterpene marneral⁷.

The advent of novel sequencing technologies has allowed the identification of an 81 increasing number of IncRNAs throughout the Arabidopsis genome. According to the latest 82 annotation (Araport 11¹⁸), three additional transcriptional units are located within the marneral 83 cluster, between the CYP71A16 and the MRN1 loci. The AT5G00580 and the pair of 84 antisense genes AT5G06325 and AT5G06335 are located upstream of the MRN1 gene at 85 6kpb and 3kbp, respectively (Fig. 1a). The 1,941bp-long AT5G00580 locus generates four 86 transcript isoforms ranging from 636 nt to 1,877 nt in length (Fig. 1b and 1c). In contrast, 87 88 each of the antisense genes AT5G06325 and AT5G06335 are transcribed into only one RNA molecule of 509 nt and 367 nt, respectively (Fig. 1a). All these transcripts were classified as 89 IncRNAs when using two coding prediction tools, CPC¹⁹ and CPC2²⁰, because of their low 90 91 coding potential and their length (over 200 nt), similarly to previously characterized IncRNAs $(COLDAIR^{21}; APOLO^{15}; and ASCO^{22})$ (Fig. 1d). 92

According to available transcriptomic datasets (Araport11), *AT5G00580* transcriptional accumulation positively correlates with that of marneral genes, whereas *AT5G06325* and *AT5G06335* RNAs do not (**Supplementary Fig. 1**). Notably, our analysis of

the transcriptional behavior of the noncoding gene AT5G00580 and the marneral cluster 96 protein-coding genes revealed a correlated expression in response to phosphate and nitrate 97 starvation, heat stress, as well as to exogenous auxin and ABA (Fig. 1e). Interestingly, the 98 99 AT5G00580 IncRNA exhibited the strongest induction in response to heat stress and exogenous ABA, in comparison with *MRN1* and the two *CYP* genes (Fig. 1e). Altogether, our 100 observations uncovered that the marneral cluster includes three noncoding transcriptional 101 units, one of which is actively transcribed and co-regulated with its neighboring protein-102 coding genes. 103

104 The IncRNA *MARS* shapes the transcriptional response of the 105 marneral gene cluster to ABA

106 It has been shown that IncRNAs can regulate the expression of their neighboring genes through epigenetic mechanisms^{14,23}. Thus, we wondered if the IncRNA derived from 107 108 the AT5G00580 locus may regulate the transcriptional activity of the protein-coding genes included in the marneral cluster. To this end, we modified the IncRNA expression without 109 affecting the cluster DNA region using an RNAi construct targeting the first exon of 110 AT5G00580, and isolated two independent lines. The RNAi line 1 impaired the transcriptional 111 112 accumulation of AT5G00580 without affecting the rest of the cluster. Interestingly, the RNAi line 2 exhibited a strong down-regulation of AT5G00580 together with a significant basal 113 repression of the two CYP genes (Fig. 2a). Strikingly, the response of the three protein-114 115 coding genes of the marneral cluster to exogenous ABA was significantly deregulated in both independent RNAi lines. Transcriptional levels of MRN1 and the two CYP genes increased 116 earlier in RNAi seedlings (15 min) than in the wild-type (Col-0, 30 min) (Fig. 2b bottom 117 panel). In addition, the transcriptional accumulation of these genes later reached three-fold 118 119 higher levels in the RNAi lines compared to Col-0 (Fig. 2b top panel). To further support our 120 observations, we isolated the insertional mutant SALK_133089 located 200 bp upstream the 121 transcription start site (TSS) of AT5G00580 gene. Although this insertion did not affect 122 AT5G00580 basal levels (Fig. 2a), it partially impaired its induction in response to ABA. In 123 agreement with both RNAi lines. MRN1 and CYP705A12 genes were strongly induced by exogenous ABA (Supplementary Fig. 2), in contrast to CYP71A16, whose promoter region 124 may be locally affected by the T-DNA insertion. Notably, *MRN1*, which encodes the marneral 125 synthase, is the gene in the cluster most strongly affected by AT5G00580 down-regulation. 126 Collectively, our results indicate that the noncoding transcriptional activity derived from the 127 AT5G00580 locus represses the dynamic expression of the marneral cluster genes, mainly 128 MRN1, in response to ABA. Therefore, we named the AT5G00580-derived noncoding 129 130 transcript MARneral Silencing (MARS) IncRNA.

131

MARS affects seed germination likely through its impact on MRN1 expression

134 The phytohormone ABA has been implicated in the perception and transduction of environmental signals participating also in a wide range of growth and developmental events 135 such as seed maturation, dormancy and germination²⁴. Considering that the marneral cluster 136 exhibited a strong MARS-dependent response to ABA, we wondered what the physiological 137 impact of MARS deregulation was during seed germination. To this end, we assessed seed 138 germination in Col-0 and MARS down-regulated lines (RNAi lines 1/2 and SALK 133089) 139 with or without exogenous ABA. Notably, MARS silencing resulted in a delayed germination 140 compared to Col-0 (Supplementary Fig. 3a), as revealed by an increase in T50 (time for 141 50% of germination) of nearly 5 hours (Supplementary Fig. 3b). Interestingly, in response to 142 0.5µM ABA, the germination of RNAi-MARS was further delayed than Col-0 with an increase 143 144 of T50 of nearly 10 hours (Supplementary Fig. 3c and 3d). Accordingly, transgenic plants over-expressing MRN1 also exhibited a delayed germination phenotype regardless of the 145 treatment with ABA (**Supplementary Fig. 3**). The behavior of 35S:MRN1 seedlings suggests 146 that *MRN1* up-regulation in RNAi-*MARS* lines could be linked to the increased sensitivity to 147 ABA (Fig. 2b). Altogether, our results indicate that MARS can modulate seed germination 148 149 through the regulation of the expression of MRN1.

150 MARS controls the epigenetic status of the marneral locus

It has been shown that gene clusters in plants are tightly regulated by epigenetic 151 modifications, including the repressive mark H3K27me3¹⁰. According to publicly available 152 ChIP-Seq datasets²⁵, the marneral cluster region is highly enriched in H3K27me3 deposition 153 in shoots, extensively coinciding with LHP1 recognition (Supplementary Fig. 4). 154 Consistently, an ATAC-Seq available dataset²⁶ revealed that the marneral cluster exhibits a 155 high chromatin condensation in shoots (Supplementary Fig. 4). Thus, the marneral cluster 156 157 is characterized by an epigenetically silent status in aerial organs coinciding with its low expression level¹⁰. 158

We wondered if the transcriptional activation of the marneral cluster in response to exogenous ABA was associated with a dynamic epigenetic reprogramming. We first assessed H3K27me3 deposition across the marneral cluster, including the gene body of *MRN1*, *MARS* and the two *CYP* loci (**Fig. 3a and Supplementary Fig. 5**). Interestingly,

exogenous ABA triggered a strong reduction of H3K27me3 deposition throughout the 163 marneral cluster (Fig. 3a and Supplementary Fig. 5). Strikingly, H3K27me3 basal levels 164 were also significantly lower in RNAi-MARS seedlings. Remarkably, H3K27me3 deposition 165 was even lower across the body of all genes of the cluster in response to ABA in the RNAi-166 MARS lines when compared with Col-0, in agreement with the stronger induction by ABA of 167 this subset of genes upon MARS silencing (Fig. 3a, Supplementary Fig. 5 and Fig. 2b). 168 Furthermore, we assessed LHP1 recognition of the marneral cluster. In agreement with 169 previous observations (Supplementary Fig. 4²⁵), LHP1 distribution was high in MRN1 170 171 promoter and more weakly across MARS gene body and the intergenic region between 172 CYP71A16 and MARS (Fig. 3b and Supplementary Fig. 6). Remarkably, LHP1 recognition 173 was strongly impaired in response to ABA as well as in RNAi-MARS seedlings (Fig. 3b and Supplementary Fig. 6). Therefore, our results indicate that ABA triggers an epigenetic 174 reprogramming of the marneral cluster, likely in a process involving the IncRNA MARS. 175

176 *MARS* is directly recognized by LHP1 and modulates local 177 chromatin condensation

It has been shown that the deposition of the repressive mark H3K27me3 and the 178 concomitant recognition of the plant PRC1 component LHP1 are correlated with high 179 chromatin condensation²⁷. Therefore, we determined by Formaldehyde-Assisted Isolation of 180 Regulatory Elements (FAIRE) the chromatin condensation of the whole marneral cluster. In 181 contrast to Col-0 showing a highly condensed chromatin, RNAi-MARS seedlings exhibit a 182 183 lower chromatin condensation in control conditions. Notably, the global chromatin status of the cluster was even less condensed in RNAi-MARS seedlings in response to ABA (Fig. 4a 184 and Supplementary Fig. 7), in agreement with a decrease of both H3K27me3 deposition 185 and LHP1 binding (Fig. 3b and Supplementary Fig. 6) and the concomitant transcriptional 186 187 activation of the clustered genes (Fig. 2b). Consistently with our observations, *lhp1* mutant 188 seedlings also showed a global chromatin decondensation in control conditions, comparable 189 to Col-0 in response to ABA. Notably, further chromatin decondensation triggered by ABA 190 was completely impaired (Fig. 4b and Supplementary Fig. 8), supporting the role of LHP1 191 in the dynamic epigenetic silencing of the marneral cluster.

192 It has been shown that LHP1 can recognize RNAs *in vitro*¹³ and the IncRNA *APOLO* 193 *in vivo*¹⁵. Moreover, it has been proposed that *APOLO* over-accumulation can decoy LHP1 194 away from target chromatin²⁸. Therefore, we wondered whether *MARS* IncRNA was able to 195 interact with the chromatin-related protein LHP1 participating in the modulation of the local 196 epigenetic environment. Thus, we first determined that *MARS* was enriched in the nucleus, compared with total RNA, as the previously characterized IncRNAs *APOLO* and *ASCO* that
 interact respectively with nuclear epigenetic and splicing machineries, and the spliceosome
 structural ncRNA *U6* (**Supplementary Fig. 9a**). Then, we confirmed by RNA
 immunoprecipitation (RIP) that LHP1 can interact with *MARS in vivo*, in contrast to the *MRN1* mRNA or housekeeping RNAs taken as negative controls (**Fig. 4c**).

202 LHP1 binding to the marneral cluster was impaired both in response to exogenous 203 ABA (inducing MARS, Fig. 1e and Supplementary Fig. 3b) and in RNAi-MARS seedlings, hinting at a stoichiometry-dependent action of MARS on LHP1 recognition. Therefore, we 204 used chromatin extracts from RNAi-MARS line 1 seedling, with very low MARS transcript 205 levels (Fig. 2a) to assess LHP1 recognition of the marneral cluster upon the addition of 206 increasing concentrations of in vitro-transcribed MARS RNA. Strikingly, we found that low 207 208 MARS RNA concentrations (between 0.01 and 0.1 µg of RNA; Fig. 4d and Supplementary Fig. 9b) successfully promoted LHP1 binding to the cluster, in contrast to higher 209 concentrations (between 1 and 10 µg of RNA), supporting the relevance of MARS-LHP1 210 stoichiometry for LHP1-target recognition (Fig. 4d and Supplementary Fig. 9b). Altogether, 211 our results suggest that the physical interaction of the nuclear-enriched lncRNA MARS to 212 LHP1 modulates its binding to proximal chromatin likely participating in the modulation of the 213 214 dynamic chromatin condensation of the marneral cluster.

MARS expression modulates an LHP1-dependent chromatin loop bringing together the *MRN1* locus and an ABA enhancer element

It was recently observed that the spatial conformation of cluster-associated domains 217 218 differs between transcriptionally active and silenced clusters. In Arabidopsis, segregating 3D 219 contacts are distinguished among organs, in agreement with the corresponding transcriptional activity of clustered genes¹¹. Therefore, we explored whether MARS could 220 participate in the dynamic regulation of the local 3D chromatin conformation modulating the 221 transcription of the marneral cluster. According to available HiC datasets^{25,29} there is a 222 significant interaction linking the intergenic region between CYP71A16 and MARS and the 223 MRN1 locus (indicated as "Chromatin loop" in Fig. 5a). By using Chromatin Conformation 224 225 Capture (3C), we determined that the formation of this chromatin loop drastically increased over 30 min with exogenous ABA and remained formed for at least 4 hours (Fig 5b). Thus, 226 the formation of this chromatin loop positively correlates with the transcriptional accumulation 227 228 of the marneral cluster genes in response to ABA (Fig. 2b).

229 The MARS locus is encompassed in the ABA-dependent chromatin loop (Fig. 5a). In order to determine the role of MARS in the modulation of local 3D chromatin conformation, 230 we assessed the formation of the chromatin loop in RNAi-MARS lines. Notably, RNAi-MARS 231 seedlings exhibit enhanced chromatin loop formation, which remained unchanged in 232 response to exogenous ABA (Fig. 5b). Interestingly, LHP1 has been implicated in shaping 233 local 3D conformation of target regions²⁵, suggesting that the LHP1-MARS module may 234 dynamically switch the epigenetic status of the marneral cluster from a condensed-linear to a 235 decondensed-3D structured chromatin conformation. Supporting this hypothesis, *lhp1* mutant 236 237 seedlings exhibited enhanced chromatin loop formation compared to Col-0 (Fig. 5c). Overall, 238 our results demonstrate that a chromatin loop within the marneral cluster is regulated by 239 LHP1 and the interacting IncRNA MARS, encoded in the region encompassed by the loop.

240 To better understand the role of the MARS-dependent chromatin loop in response to ABA we looked for ABA-related *cis* regulatory sequences throughout the marneral cluster. 241 We extracted from ChIP-seq the binding distribution of 13 ABA-related transcription factors 242 (TFs)³⁰. Interestingly, a high enrichment in ABA TF binding sites was found at the MARS 243 locus, as well as in the intergenic region between the CYP71A16 and MARS loci, notably 244 surrounding the contact point brought into close spatial proximity with the MRN1 locus by the 245 246 ABA-dependent 3D chromatin loop (Fig. 5a). We thus assessed the activating capacity of 247 this region, potentially acting as a distant enhancer element of the *MRN1* proximal promoter. To this end, we made use of a *GUS*-based reporter system (as described in ³¹), fusing two 248 249 regions of interest to a minimal 35S promoter. Two additional DNA regions nearby the 250 putative enhancers were used as negative controls: one between CYP705A12 and CYP71A16 and the other at the 3' end of AT5G42620 locus (Fig. 5a indicated in red). Among 251 the two putative enhancers regions tested, one was able to activate GUS expression 252 (Intergenic region 2, Fig. 5a and 5d), coinciding with the region showing a high enrichment 253 254 of ABA-related TF binding sites close to the chromatin loop anchor point (Fig. 5a). Collectively, our results indicate that an ABA-driven chromatin loop brings into close spatial 255 proximity the MRN1 locus and a transcriptional activation site likely acting as an ABA 256 enhancer element. Notably, this chromatin reorganization process depends on the LHP1-257 258 MARS module.

259 Long noncoding RNAs as emerging regulators of gene clusters

260 Physically linked genes organized in clusters are generally coregulated⁸. Considering 261 that the IncRNA *MARS* is implicated in the regulation of the marneral cluster, we wondered 262 whether the presence of noncoding transcriptional units may constitute a relevant feature of

gene cluster organization. Therefore, we collected two lists of Arabidopsis clustered genes, 263 i.e. one of co-expressed neighboring genes¹⁰ and one of predicted metabolic gene clusters 264 (by PlantiSMASH³²). According to the latest Arabidopsis genome annotation (Araport11), 265 among the 390 clusters of co-expressed neighboring genes), 189 (48%) contain at least one 266 267 embedded IncRNA inside the cluster. Most importantly, among the 45 metabolic clusters, 28 (62%) include IncRNAs (Fig. 6a). Furthermore, among the clusters containing a IncRNA, a 268 269 correlation analysis based on the maximum strength of co-expression between a IncRNA 270 and any clustered gene revealed that the metabolic clusters exhibit a significantly higher 271 correlation than co-expressed clusters (Fig. 6b). Altogether, our analyses suggest that 272 IncRNA-mediated local epigenetic remodeling may constitute an emerging feature of non-273 homologous genes metabolic clusters in plants.

274 **DISCUSSION**

The cell nucleus is a dynamic arrangement of DNA, RNAs and proteins^{33,34}. Genome topology has emerged as an important feature in the complex network of mechanisms regulating gene activity and genome connectivity, leading to regionalized chromosomal spatial distribution and the clustering of diverse genomic regions with similar expression patterns³⁵.

In the last few years, noncoding transcription has been implicated in shaping 3D 280 nuclear organization³⁶. Notably, RNase-A micro-injection into the nucleus revealed that long 281 282 nuclear-retained RNAs maintained euchromatin in a biologically active decondensed state, whereas heterochromatin domains exhibited an RNA-independent structure^{37,38}. More 283 recently, HiC analyses were performed in mammalian cells exposed or not to RNase, before 284 and after crosslinking, or upon transcriptional inhibition³⁹. As a result, it was observed that 285 topologically associated domains (TAD) boundaries remained mostly unaffected by RNase 286 287 treatment, whereas compartmental interactions suffered a subtle disruption. In contrast, transcriptional inhibition led to weaker TAD boundaries, hinting at different roles of steady-288 state RNA vs. active transcription in nuclear organization³⁹. 289

In plants, several lncRNAs have been implicated in local chromatin conformation dynamics affecting the transcriptional activity of neighboring genes^{14,40}. Notably, the lncRNA *COLDWRAP* participates in the formation of an intragenic chromatin loop blocking the transcription of the flowering time regulator *FLOWERING LOCUS C* (*FLC*¹⁶) in response to cold, in a process involving the recruitment of PRC2 by direct interaction with the component CLF. The lncRNA *APOLO* also controls the transcriptional activity of its neighboring gene *PINOID* (*PID*) by dynamically modulating the formation of an intergenic chromatin loop

encompassing the divergent promoter of *PID* and *APOLO*¹⁵, in a process involving the PRC1 297 component LHP1. More recently, it was proposed that high levels of APOLO can decoy 298 LHP1 away from multiple loci in *trans*, modulating the 3D conformation of target genes²⁸. In 299 rice, the expression of the leucine-rich repeat receptor kinase clustered genes RLKs is 300 modulated by the locally-encoded IncRNA LRK ANTISENSE INTERGENIC RNA (LAIR). It 301 was proposed that LAIR may directly recruit OsMOF (MALES ABSENT ON THE FIRST) and 302 OsWDR5 (WD REPEAT DOMAIN 5), involved in H4K16 acetylation and chromatin 303 304 remodeling⁴¹. Here we showed that the IncRNA *MARS* contributes to the co-regulation of a 305 set of physically linked genes in *cis* in Arabidopsis. We demonstrated that the relative 306 abundance of in vitro-transcribed MARS fine-tunes LHP1 binding to the cluster region in a 307 stoichiometry-dependent manner, thus explaining how MARS affects H3K27me3 deposition and chromatin condensation. It has been shown in yeast that histone depletion boosts 308 chromatin flexibility and facilitates chromatin loop formation on the kilobase pair scale⁴². In 309 310 agreement thereof, we uncovered here the dynamic role of the LHP1-MARS module affecting nucleosome distribution across the marneral cluster in response to ABA, thus 311 promoting the formation of an intra-cluster chromatin loop. 312

It has been recently observed that biosynthetic gene clusters are embedded in local 313 three-dimensionally organized hot spots that segregate the region from the surrounding 314 chromosome environment¹¹. Here we showed that active noncoding transcriptional units 315 316 within the cluster may contribute to 3D conformation dynamics switching from silent to active 317 states. Our results indicated that the MARS-dependent chromatin loop may bring the MRN1 318 locus and a distal ABA-responsive element into close spatial proximity, likely acting as an enhancer. Notably, MARS-dependent LHP1 and H3K27me3 removal in Col-0, RNAi-MARS 319 and the *lhp1* mutant correlated with chromatin decondensation, loop formation and increased 320 marneral genes transcriptional activity in response to ABA. According to this model, 321 chromatin loop conformation is related to LHP1 binding and is modulated by MARS. LHP1 322 recognition at basal MARS levels maintains a possibly linear conformation of the region, 323 precluding the enhancer-MRN1 locus interaction, whereas the positively activating chromatin 324 loop is formed in the absence of LHP1. MARS transcriptional accumulation directly 325 modulates LHP1 binding to the marneral cluster and high levels of MARS then titrate LHP1 326 away from the cluster (Fig. 6c; in response to ABA). Interestingly, when MARS levels are too 327 low compared to basal levels (as in the RNAi lines), recruitment of LHP1 to the cluster is also 328 329 impaired (Fig. 6c; MARS repression).

In mammals, growing evidence supports the role of lncRNAs in chromatin conformation determination⁴³ and enhancer activity (e.g. $PVT1^{44}$ and $CCAT1-L^{45}$). Here, we

showed that the nuclear-enriched IncRNA MARS brings together the MRN1 proximal 332 promoter and a putative enhancer element enriched in ABA-responsive TF binding sites. 333 Interestingly, it has been shown that human IncRNAs can modulate the binding of TFs to 334 their target chromatin (DHFR⁴⁶) and PANDA⁴⁷, whereas TFs have been implicated in 335 chromatin loop formation in plants³⁵. Furthermore, it was shown that in addition to the TF NF-336 YA, the IncRNA PANDA interacts with the scaffold-attachment-factor A (SAFA) as well as 337 with PRC1 and PRC2 to modulate cell senescence⁴⁸. Therefore, further research will be 338 needed to determine what ABA-responsive TFs are in control of the marneral cluster and to 339 340 elucidate how they participate in chromatin loop formation along the area, in relation with the 341 PRC1-interacting IncRNA MARS.

Plants are a tremendous source of diverse chemicals which are important for their life 342 and survival¹⁰. Marneral biosynthesis has been linked to root and leaf development, flowering 343 time and embryogenesis². Here we found that the Arabidopsis marneral cluster is activated 344 by the phytohormone ABA, in a IncRNA-dependent epigenetic reprogramming. MARS 345 deregulation affects the cluster response to ABA, impacting seed germination. Interestingly, 346 noncoding transcription had already been associated with seed germination. The DELAY OF 347 GERMINATION 1 (DOG1) locus is a major actor regulating seed dormancy strength in 348 Arabidopsis. Indeed, an antisense IncRNA (asDOG1) is able to repress DOG1 transcription 349 in mature plants. Notably, it was observed that DOG1 suppression is released by shutting 350 down antisense transcription, which is induced by ABA and drought⁴⁹. 351

It was proposed that the marneral cluster was founded by the duplication of ancestral 352 genes, independent events of gene rearrangement and the recruitment of additional genes⁷. 353 The exploration of the noncoding transcriptome in Arabidopsis recently served to identify 354 ecotype-specific IncRNA-mediated responses to the environment⁵⁰. It was suggested that the 355 noncoding genome may participate in multiple mechanisms involved in ecotype adaptation. 356 Collectively, our results indicate that the acquisition of novel noncoding transcriptional units 357 358 within biosynthetic gene clusters may constitute an additional regulatory layer behind their natural variation in plant evolution. 359

360 **METHODS**

361 Lead contact and materials availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Martin Crespi (martin.crespi@ips2.universite-parissaclay.fr).

365 Plant lines generated in this study are available from the Lead Contact with a 366 completed Materials Transfer Agreement.

367 Lines selection and generation

All plants used in this study are in Columbia-0 background. RNAi-*MARS* were obtained using the pFRN binary vector⁵¹ bearing 250bp of the first exon of *MARS* gene (see primers in **Supplementary Table 1**), previously sub-cloned into the pENTR/D-TOPO vector. Arabidopsis plants were transformed using *Agrobacterium tumefaciens* Agl-0⁵². The T-DNA inserted line *SALK_133089* was ordered to NASC (N633089).

373 **Growth conditions**

Seeds were sown in plates vertically disposed in a growing chamber in long day 374 conditions (16 h in light 150uE; 8 h in dark; 21°C) for all the experiments. Plants were grown 375 on solid half-strength MS medium (MS/2) supplemented with 0.7% sucrose, and without 376 377 sucrose for the germination assay. For nitrate starvation assay, KNO₃ and Ca(NO₃) $_2$ were 378 replaced from MS/2 by a corresponding amount of KCI and CaCl₂ respectively, 2.25 mM 379 NH₄HCO₃ was added for nitrate-containing medium. For the phosphate starvation assay, growth medium contained 0.15 mM MgSO₄, 2.1 mM NH₄NO₃, 1.9 mM KNO₃, 0.34 mM CaCl₂, 380 0.5 µM KI, 10 µM FeCl₂, 10 µM H₃BO₃, 10 µM MnSO₄, 3 µM ZnSO₄, 0.1 µM CuSO₄, 0.1 µM 381 CoCl, 0.1 µM Na,MoO₄, 0.5 g.L⁻¹ sucrose supplemented with 500uM Pi for Pi containing 382 medium versus 10uM for Pi free medium. All media were supplemented with 0.8g/L agar 383 (Sigma-Aldrich, A1296 #BCBL6182V) and buffered at pH 5.6 with 3.4mM 2-(N-morpholino) 384 ethane sulfonic acid. For the treatment with exogenous ABA or auxin, seedlings were 385 sprayed with 10uM ABA and 10uM 1-Naphthaleneacetic acid (NAA), respectively. For heat 386 387 stress, plates were transferred to a growth chamber at 37°C under the same lightning conditions. For nitrate and phosphate starvation assays, seedlings have been transferred at 388 day 12 after sowing (DAS) from respectively nitrate and phosphate containing medium to 389 nitrate and phosphate free medium. Finally, for seed germination assay, 0.5uM ABA was 390 391 supplemented or not to the medium. Germination rate was evaluated twice a day. Seeds were considered germinated when the seed coats were perforated by elongating radicle. For 392 all the experiments, samples were taken from 12 DAS starting two hours after light 393 394 illumination, at different time-points, after cross-linking or not, depending on the experiment.

395 **RT-qPCR**

Total RNA was extracted from whole seedlings using TRI Reagent (Sigam-Aldrich) 396 397 and treated with DNase (Fermentas) as indicated by the manufacturers. Reverse transcription was realized on 1ug total RNA using the Maxima Reverse Transcriptase 398 (Thermo Scientific). gPCR was performed on a Light Cycler 480 with SYBR Green master I 399 (Roche) in standard protocol (40 cycles, 60°C annealing). Primers used in this study are 400 401 listed in **Supplementary Table 1**. Data were analyzed using the delta delta Ct method using PROTEIN PHOSPHATASE 2A SUBUNIT A3 (AT1G13320) for gene normalization⁵³ and time 402 0 for time-course experiment. 403

404 Chromatin Immunoprecipitation (ChIP)

405 ChIP was performed using anti-IgG (Millipore,Cat#12-370), anti-H3K27me3 (Millipore, 406 Cat#07-449) and anti-LHP1 (Covalab, Pab0923-P), as previously described¹⁵, starting from two grams of seedlings crosslinked in 1% (v/v) formaldehyde. Chromatin was sonicated in a 407 water bath Bioruptor Plus (Diagenode; 60 cycles of 30s ON and 30s OFF pulses at high 408 intensity). ChIP was performed in an SX-8G IP-Star Compact Automated System 409 (Diagenode). Antibody coated into Protein A Dynabeads (Invitrogen) were incubated 12 410 hours at 4 °C with the samples. Recovering of immunoprecipitated DNA was realized using 411 Phenol:Chloroform:Isoamilic Acid (25:24:1, Sigma) followed by ethanol precipitation and 412 analyzed by gPCR. For input samples, non-immunoprecipitated sonicated chromatin was 413 processed in parallel. 414

415 In-vitro transcribed MARS RNA was obtained from a PCR product amplified from wild-416 type genomic DNA using the T7 promoter on the 5' PCR primer (Supplementary Table 1). 417 PCR products were verified using agarose electrophoresis, and purified using NucleoSpin kit (Macherey-Nagel). 1µg of purified DNA was used for in-vitro transcription following the 418 manufacturer instructions (HiScribe T7 High Yield RNA Synthesis Kit, NEB). Purified non-419 crosslinked chromatin obtained from five grams of MARS RNAi line 1 seedlings were 420 resuspending in 1 mL of nuclei lysis buffer and split into five tubes. An increasing quantity of 421 422 MARS RNA was added to each tube from 0 to 10 µg RNA and incubated under soft rotation during 3 h at 4 °C. Chromatin samples were then cross-linked using 1% (v/v) of 423 formaldehyde for five minutes. Sonication and the following ChIP steps were performed as 424 above. 425

426 Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)

427 FAIRE was performed as described by ⁵⁴. After chromatin purification as for ChIP, 428 only 50 μ I from the 500 μ I of purified chromatin were used (diluted to 500 μ I with 10 mM Tris-429 HCI pH 8). For the qPCR, the same set of primers as for ChIP were used.

430 Nuclear purification

Non-cross-linked seedlings were used to assess the sub-cellular localization of RNAs. To obtain the nuclear fraction, chromatin was purified as for ChIP and resuspended, after the sucrose gradient, into 1mL of TRI Reagent (Sigam-Aldrich). For the total fraction, 200 μ L of cell suspension in cell resuspension solution, were collected and completed with 800 μ L of TRI Reagent to follow with the RNA extraction. RNA's samples were treated by DNase, and RT was performed using random hexamers prior to qPCR analysis.

437 **RNA immunoprecipitation (RIP)**

For RIP, the *lhp1* mutant complemented with the *ProLHP1:LHP1:GFP*⁵⁵ were used after 4 hours of treatment with ABA. After crosslinking and chromatin extraction as for ChIP, ten percent of resuspended chromatin was conserved at -20 °C as the input. Chromatin was sonicated in a water bath Bioruptor Plus (Diagenode; 5 cycles of 30 s ON and 30 s OFF pulses at high intensity). Anti-LHP1 RIP was performed using the anti-GFP antibody (Abcam ab290), as previously described¹⁵. Results were expressed as the percentage of cDNA detected after IP taking the input value as 100%.

445 **Chromosome conformation capture (3C)**

3C was performed as described by ⁵⁶. Briefly, chromatin was extracted from two grams of cross-linked seedlings as for ChIP. Overnight digestion at 37 °C were performed using 400U of Hind III enzyme (NEB). Digested DNA was ligated during 5 h incubation at 16 °C with 100 U of T4 DNA ligase (NEB). DNA was recovered after reverse crosslinking and Proteinase K treatment (Invitrogen) by Phenol:Chloroform:Isoamyl Acid (25:24:1; Sigma) extraction and ethanol precipitation. Interaction frequency was calculated by qPCR using a DNA region uncut by Hind III to normalize the amount of DNA.

453 **Transcriptional activation assay in tobacco leaves**

The *GUS* reporter system for validating the activity of the putative enhancer element was adapted from ³¹. Different DNA fragments were cloned in the GreenGate system⁵⁷ fused to a minimal 35S promoter element from CAMV (synthesized by Eurofins Genomics). The 457 sub-unit B3 from 35S promoter element from CAMV⁵⁸ was synthesized and used as a 458 positive control. All primers used for cloning are indicated in **Supplementary Table 1**.

A. tumefaciens-mediated transient transformation was performed on 5-week-old tobacco plants using a needle-less syringe. Together with enhancer constructs, another vector containing mCherry driven by 35S promoter was co-transfected to control the transformation efficiency. Two leaf discs were collected near the infiltration site. One, to determine the transfection efficiency by mCherry fluorescence observation under epifluorescent microscope. The second was used for GUS staining, as previously described⁵⁹. Samples were incubated 4 h in the dark at 37 °C before observation.

466 Identification of IncRNA loci in Arabidopsis gene clusters

The genes of co-expressed clusters were retrieved from ¹⁰. The boundaries of the gene clusters were extracted using Araport11 annotations. The boundaries of the metabolic clusters were extracted from the plantiSMASH predicted clusters on Arabidopsis³². Using Araport11 GFF, the lncRNAs (genes with a locus type annotated as "long_noncoding_rna", "novel_transcribed_region" or "other_rna") present within the boundaries of the cluster were retrieved.

473 Gene expression correlation analysis

To compute the correlation of expression in different organ of Arabidopsis we used 474 the 113 RNA-seq datasets that were used for the Araport11 annotations (10.1111/tpj.13415). 475 These datasets were generated from untreated or mock-treated wild-type Col-0 plants. After 476 removing the adaptors with Trim Galore with default parameters, the reads were mapped on 477 TAIR10 with STAR v2.7.2a (10.1093/bioinformatics/bts635) and the parameters '--478 alignIntronMin 20 --alignIntronMax 3000'. Gene expression was then quantified with 479 featureCounts v2.0.0 (10.1093/bioinformatics/btt656) with the parameters "-B -C -p -s 0" 480 using the GFF of Araport11. Raw counts were then normalized by median of ratios using the 481 DESeq2 R package⁶⁰. 482

For the correlation of expression inside the marneral cluster, the expression of the genes of the clusters and 25kb around it (four genes upstream and two downstream) were collected for the correlation analysis. Pearson's correlations for each pair of genes were computed after log 2 transformation of the normalized counts. The correlation value and associated p-value were plotted with the corrplot R package⁶¹. Inside each co-expressed and metabolic clusters of genes, Pearson's correlation was computed between every possible pairs of IncRNA and coding gene as for the genes inside the marneral cluster. The maximum correlation value was kept as an indication of IncRNAs correlation with the genes of the cluster.

492 **Quantification and statistical analyses**

For all the experiments, at least two independent biological samples were considered. 493 For RT-gPCR, each sample was prepared from a pool of 5 to 10 individual seedlings. For 494 biochemistry assays (ChIP, FAIRE, nuclear purification, RIP and 3C) two to five grams of 495 seedlings were prepared for each independent biological sample. For validation of enhancer 496 497 function, the four leaf discs were taken from four independent tobacco plants. The tests used 498 for statistical analyses are indicated in the respective figure legends. Statistical test and associated plots have been generated using R software (v3.6.3⁶²) with the help of the 499 500 tidyverse package⁶³.

501 ACKNOWLEDGMENTS

502 IPS2 benefits from the support of Saclay Plant Sciences-SPS (ANR-17-EUR-0007). We 503 thank Jeremie Bazin and Aurélie Christ from IPS2 for the helpful discussion about results 504 interpretation and design of the experiments. We also thank Moussa Benhamed (IPS2) for 505 helpful advice on epigenetic regulation. We thank Olivier Martin for critical reading of the 506 manuscript.

507 AUTHORS' CONTRIBUTIONS

508 TR, FA, TB and MC conceived and designed the experiments. TR performed the 509 experiments. TR and TB analyzed the data. All authors discussed the results and edited the 510 manuscript.

511 COMPETING FINANCIAL INTERESTS

512 The authors declare no competing financial interests.

513 **REFERENCES**

1. Nützmann, H.-W., Scazzocchio, C. & Osbourn, A. Metabolic Gene Clusters in Eukaryotes. *Annu.*

515	Rev.	Genet.	52,	159-	-183	(2018).
-----	------	--------	-----	------	------	---------

- 516 2. Go, Y. S. *et al.* Identification of marneral synthase, which is critical for growth and 517 development in Arabidopsis. *Plant J.* **72**, 791–804 (2012).
- Yasumoto, S., Fukushima, E. O., Seki, H. & Muranaka, T. Novel triterpene oxidizing activity of
 Arabidopsis thaliana CYP716A subfamily enzymes. *FEBS Lett.* 590, 533–540 (2016).
- 520 4. Field, B. & Osbourn, A. E. Clusters in Different Plants. *Science (80-.).* **194**, 543–547 (2008).
- 5. Boutanaev, A. M. *et al.* Investigation of terpene diversification across multiple sequenced
 plant genomes. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E81–E88 (2015).
- 6. Castillo, D. A., Kolesnikova, M. D. & Matsuda, S. P. T. An effective strategy for exploring
 unknown metabolic pathways by genome mining. *J. Am. Chem. Soc.* 135, 5885–5894 (2013).
- Field, B. *et al.* Formation of plant metabolic gene clusters within dynamic chromosomal
 regions. *Proc. Natl. Acad. Sci.* **108**, 16116–16121 (2011).
- 527 8. Nützmann, H. W., Huang, A. & Osbourn, A. Plant metabolic clusters from genetics to
 528 genomics. *New Phytol.* 211, 771–789 (2016).
- 529 9. Nützmann, H. W. & Osbourn, A. Regulation of metabolic gene clusters in Arabidopsis thaliana.
 530 *New Phytol.* 205, 503–510 (2015).
- 531 10. Yu, N. *et al.* Delineation of metabolic gene clusters in plant genomes by chromatin signatures.
 532 *Nucleic Acids Res.* 44, 2255–2265 (2016).
- Nützmann, H. *et al.* Active and repressed biosynthetic gene clusters have spatially distinct
 chromosome states. 1–10 (2020) doi:10.1073/pnas.1920474117.
- 535 12. Rinn, J. L. & Chang, H. Y. Long Noncoding RNAs: Molecular Modalities to Organismal
 536 Functions. *Annu. Rev. Biochem.* 89, 283–308 (2020).
- Berry, S., Rosa, S., Howard, M., Bühler, M. & Dean, C. Disruption of an RNA-binding hinge
 region abolishes LHP1-mediated epigenetic repression. *Genes Dev.* **31**, 2115–2120 (2017).
- Lucero, L., Fonouni-Farde, C., Crespi, M. & Ariel, F. Long noncoding RNAs shape transcription
 in plants. *Transcription* **00**, 1–12 (2020).
- 541 15. Ariel, F. *et al.* Noncoding transcription by alternative rna polymerases dynamically regulates
 542 an auxin-driven chromatin loop. *Mol. Cell* 55, 383–396 (2014).
- 543 16. Kim, D.-H. & Sung, S. Vernalization-triggered intragenic chromatin-loop formation by long

544 noncoding RNAs. *Dev. Cell* **176**, 100–106 (2017).

- 545 17. Gagliardi, D. *et al.* Dynamic regulation of chromatin topology and transcription by inverted
 546 repeat-derived small RNAs in sunflower. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 17578–17583
 547 (2019).
- 548 18. Cheng, C. Y. *et al.* Araport11: a complete reannotation of the Arabidopsis thaliana reference
 549 genome. *Plant J.* **89**, 789–804 (2017).
- Kong, L. *et al.* CPC: Assess the protein-coding potential of transcripts using sequence features
 and support vector machine. *Nucleic Acids Res.* **35**, 345–349 (2007).
- 552 20. Kang, Y. J. *et al.* CPC2: A fast and accurate coding potential calculator based on sequence
 553 intrinsic features. *Nucleic Acids Res.* 45, W12–W16 (2017).
- Heo, J. B. & Sung, S. Vernalization-mediated epigenetic silencing by a long intronic noncoding
 RNA. *Science (80-.).* 331, 76–79 (2011).
- 556 22. Bardou, F. *et al.* Long Noncoding RNA Modulates Alternative Splicing Regulators in
 557 Arabidopsis. *Dev. Cell* **30**, 166–176 (2014).
- 558 23. Jarroux, J., Morillon, A. & Pinskaya, M. Long Non Coding RNA Biology. *Adv. Exp. Med. Biol.*559 **1008**, 1–46 (2017).
- Vishwakarma, K. *et al.* Abscisic acid signaling and abiotic stress tolerance in plants: A review
 on current knowledge and future prospects. *Front. Plant Sci.* 8, 1–12 (2017).
- 562 25. Veluchamy, A. *et al.* LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional
 563 Conformation of the Arabidopsis Genome. *PLoS One* **11**, 1–25 (2016).
- Sijacic, P., Bajic, M., McKinney, E. C., Meagher, R. B. & Deal, R. B. Changes in chromatin
 accessibility between Arabidopsis stem cells and mesophyll cells illuminate cell type-specific
 transcription factor networks. *Plant J.* 94, 215–231 (2018).
- Yang, X., Tong, A., Yan, B. & Wang, X. Governing the silencing state of chromatin: The roles of
 polycomb repressive complex 1 in arabidopsis. *Plant Cell Physiol.* 58, 198–206 (2017).
- 569 28. Ariel, F. *et al.* R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Mol. Cell* 77,
 570 1–11 (2020).
- 571 29. Liu, C. *et al.* Genome-wide analysis of chromatin packing in Arabidopsis thaliana at single-gene
 572 resolution. *Genome Res.* 26, 1057–1068 (2016).

- Song, L. *et al.* A transcription factor hierarchy defines an environmental stress response
 network. *Science (80-.).* 354, 598 (2016).
- 575 31. Yan, W. *et al.* Dynamic control of enhancer activity drives stage-specific gene expression
 576 during flower morphogenesis. *Nat. Commun.* **10**, 1–16 (2019).
- 577 32. Kautsar, S. A., Suarez Duran, H. G., Blin, K., Osbourn, A. & Medema, M. H. PlantiSMASH:
- 578 Automated identification, annotation and expression analysis of plant biosynthetic gene 579 clusters. *Nucleic Acids Res.* **45**, W55–W63 (2017).
- 580 33. Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat. Struct. Mol. Biol.* 20,
 581 290–299 (2013).
- 582 34. Gibcus, J. H. & Dekker, J. The Hierarchy of the 3D Genome. *Mol. Cell* **49**, 773–782 (2013).
- 583 35. Rodriguez-Granados, N. Y. *et al.* Put your 3D glasses on: Plant chromatin is on show. *J. Exp.* 584 *Bot.* 67, 3205–3221 (2016).
- 36. Quinodoz, S. & Guttman, M. Long non-coding RNAs: An emerging link between gene
 regulation and nuclear organization. *Trends Cell Biol.* 24, 651–663 (2014).
- 587 37. Caudron-Herger, M. *et al.* Coding RNAs with a non-coding function: Maintenance of open
 588 chromatin structure. *Nucleus* 2, (2011).
- 589 38. Caudron-Herger, M. & Rippe, K. Nuclear architecture by RNA. *Curr. Opin. Genet. Dev.* 22, 179–
 590 187 (2012).
- 39. Barutcu, A. R., Blencowe, B. J. & Rinn, J. L. Differential contribution of steady-state RNA and
 active transcription in chromatin organization . *EMBO Rep.* 20, 1–13 (2019).
- 593 40. Gagliardi, D. & Manavella, P. A. Short-range regulatory chromatin loops in plants. *New Phytol.*594 1–6 (2020) doi:10.1111/nph.16632.
- Wang, Y. *et al.* Overexpressing lncRNA LAIR increases grain yield and regulates neighbouring
 gene cluster expression in rice. *Nat. Commun.* 9, 1–9 (2018).
- 597 42. Diesinger, P. M., Kunkel, S., Langowski, J. & Heermann, D. W. Histone depletion facilitates
 598 chromatin loops on the kilobasepair scale. *Biophys. J.* 99, 2995–3001 (2010).
- Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* 21, 102–117 (2020).
- 601 44. Cho, S. W. et al. Promoter of IncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary

602 Element. *Cell* **173**, 1398-1412.e22 (2018).

- 45. Xiang, J. F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range
 chromatin interactions at the MYC locus. *Cell Res.* 24, 513–531 (2014).
- 46. Martianov, I., Ramadass, A., Serra Barros, A., Chow, N. & Akoulitchev, A. Repression of the
 human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445, 666–
 670 (2007).
- Hung, T. *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle
 promoters. *Nat. Genet.* 43, 621–629 (2011).
- 48. Puvvula, P. K. *et al.* Long noncoding RNA PANDA and scaffold-attachment-factor SAFA control
 senescence entry and exit. *Nat. Commun.* 5, (2014).
- 49. Yatusevich, R. *et al.* Antisense transcription represses Arabidopsis seed dormancy QTL DOG 1
 to regulate drought tolerance . *EMBO Rep.* 18, 2186–2196 (2017).
- 614 50. Blein, T. *et al.* Landscape of the non-coding transcriptome response of two Arabidopsis
 615 ecotypes to phosphate starvation. *Plant Physiol.* **183**, pp.00446.2020 (2020).
- Ariel, F. *et al.* Two direct targets of cytokinin signaling regulate symbiotic nodulation in
 medicago truncatula. *Plant Cell* 24, 3838–3852 (2012).
- 618 52. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for Agrobacterium-mediated
 619 transformation of Arabidopsis thaliana. *Plant J.* 16, 735–743 (1998).
- 53. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. Genome-Wide
 Identification and Testing of Superior Reference Genes for Transcript Normalization in
 Arabidopsis. *Plant Physiol.* 139, 5–17 (2005).
- 54. Simon, J. M., Giresi, P. G., Davis, I. J. & Lieb, J. D. Using formaldehyde-assisted isolation of
 regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* 7, 256–267 (2012).
- 55. Nakahigashi, K., Jasencakova, Z., Schubert, I. & Goto, K. The Arabidopsis HETEROCHROMATIN
 PROTEIN1 homolog (TERMINAL FLOWER2) silences genes within the euchromatic region but
 not genes positioned in heterochromatin. *Plant Cell Physiol.* 46, 1747–1756 (2005).
- 56. Louwers, M., Splinter, E., van Driel, R., de Laat, W. & Stam, M. Studying physical chromatin
 interactions in plants using Chromosome Conformation Capture (3C). *Nat. Protoc.* 4, 1216–
 1229 (2009).

- 57. Lampropoulos, A. *et al.* GreenGate A novel, versatile, and efficient cloning system for plant
 transgenesis. *PLoS One* 8, (2013).
- 633 58. Moreno-risueno, M. A. *et al.* NIH Public Access. **329**, 1306–1311 (2010).
- 59. Jefferson, R. A., Kavanagh, T. A. & Bevan, M. W. GUS fusions: ,B-glucuronidase as a sensitive
 and versatile gene fusion marker in higher plants. *EMBO J.* 6, 3901–3907 (1987).
- 636 60. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
 637 RNA-seq data with DESeq2. *Genome Biol.* 15, 1–21 (2014).
- 638 61. Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation
- 639 Matrix (Version 0.84). Available online at <u>https://github.com/taiyun/corrplot</u>
- 640 62. R Core Team (2018). R: A language and environment for statistical computing. R Foundation
- 641 for Statistical Computing, Vienna, Austria. Available online at <u>https://www.R-project.org/</u>.
- 642 63. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).



Fig. 1, *AT5G00580* is a IncRNA transcribed from the marneral cluster locus and its expression correlates with its neighboring genes

a, Schematic illustration of the marneral cluster. Genes are indicated with plain rectangles and white arrows indicate the sense of transcription. The square indicates the region displayed in (B).

b, Schematic illustration of the different isoforms of *AT5G00580* transcripts. First line corresponds to *AT5G00580* genomic region whereas the other lines present the various isoforms. For each isoform, exons are indicated with rectangles and introns with solid lines. Black arrows indicate the primers used for the cDNA amplification in (C).

c, cDNA amplification of *AT5G00580* isoforms. The position of the primers used for the amplification are displayed on B. MWM stands for Molecular Weight Marker (GeneRuler 1 kb Plus DNA Ladder, Thermo Scientific).

d, Coding potential of the transcripts located in the marneral cluster genomic region. Scores were calculated using CPC1 (left) and CPC2 (right) programs^{19,20}. For each algorithm, the threshold between coding and non-coding genes is displayed with a horizontal solid black line. Coding genes are situated above the threshold, whereas non coding genes are situated below. *COLDAIR*, *APOLO* and *ASCO* are used as positive controls of non-coding transcripts.

e, Dynamic transcriptional levels of co-regulated genes of the marneral cluster in response to phosphate and nitrate starvation, heat stress, and exogenous ABA and auxin. Gene expression data are represented as the mean \pm standard error (n \geq 3) of the log2 fold change compared to time 0h.

a Transcript abundance





Fig. 2, *MARS* transcriptional activity modulates the response to ABA of the marneral cluster

a, Transcript levels of the marneral cluster genes in control conditions in RNAi lines targeting *AT5G00580/MARS* and *SALK_133089* line. Transcriptional levels are represented as the mean \pm standard error (n = 3) of the log2 fold change compared to Col-0. Letters indicate statistic group determined by one-way analysis of variance (ANOVA) followed by Tukey's post-hoc test. For each gene, each letter indicates statistical difference between genotypes (p ≤ 0.05).

b, Transcript levels of the genes of the marneral cluster in response to ABA treatment in RNAi lines targeting *AT5G00580/MARS*. Gene expression data are represented as the mean \pm standard error (n = 3) of the log2 fold change compared to time 0h.





10

0



Fig. 3, MARS modulates the epigenetic landscape of MRN1 locus

a, H3K27me3 deposition over the MRN1 gene in Col-0 and RNAi-MARS seedlings under control conditions and in response to ABA. Higher values of ChIP-gPCR indicate more H3K27me3.

b, LHP1 binding to the MRN1 gene in Col-0 and RNAi-MARS seedlings in the same conditions as in (A). Higher values of ChIP-qPCR indicate more LHP1 deposition. In (A) and (B), values under the dotted line are considered as not enriched. Results are represented as the mean ± standard error (n = 2) of the H3K27me3/IgG or LHP1/IgG ratio. Numbers are pvalue of the difference between the two genotypes determined by Student t-test.













Fig. 4, MARS influences chromatin condensation of MRN1 gene through its interaction with LHP1 protein

a, Chromatin condensation in MRN1 gene of Col-0 and RNAi-MARS seedlings in control conditions and in response to ABA, determined by Formaldehyde Assisted Isolation of Regulatory Element (FAIRE)-qPCR.

b, Evolution of the chromatin condensation in MRN1 gene of Col-0 and Ihp1 mutant subjected to ABA treatment determined by Formaldehyde Assisted Isolation of Regulatory Element (FAIRE) qPCR.

In a and b, results are expressed as the mean ± standard error (n = 3) of the percentage of input (signal measured before isolation of decondensed region of chromatin, free of nucleosomes). Lower value indicates more condensed chromatin. Numbers are p-value of the difference between the two genotypes determined by Student t-test.

c, LHP1-MARS interaction was assessed by RNA immunoprecipitation (RIP) using LHP1-GFP seedlings. Negative controls include a housekeeping gene (PP2A) and MRN1 mRNA. The interaction between APOLO and LHP1 was taken as a positive control¹⁵. Results are expressed as the mean \pm standard error (n = 4) of the percentage of input (signal measured before immunoprecipitation).

d, LHP1 binding to the MRN1 promoter region in chromatin from RNAi-MARS seedlings upon increasing amounts of *in-vitro* transcribed MARS RNA. After incubation (see Methods), the samples were crosslinked for LHP1 ChIP-qPCR. Higher values indicate more LHP1-DNA interaction. Results are expressed as the mean \pm standard error (n = 2) of the LHP1/Igg ratio.

In **c** and **d** numbers are p-value of the difference between the different corresponding genes determined by Student t-test.



Fig. 5, An LHP1-dependent chromatin loop brings together the *MRN1* locus and a putative enhancer element in response to ABA

a, Schematic illustration of the loop linking the *MRN1* locus with the intergenic region between *CYP71A16* and *MARS*. Forward (F) and Reverse (R) oligonucleotides used for 3C-qPCR (in B–C) are indicated with arrows. The orange track shows the number of different ABA-related transcription factor binding sites (HB6, HB7, GBF2, GBF3, MYB3, MYB44, NF-YC2, NF-YB2, ANAC102, ANAC032, ABF1, ABF3, ABF4, RD26, ZAT6, FBH3, DREBA2A, AT5G04760, HAT22 and HSFA6A) found along the marneral cluster³⁰. Green and red rectangles indicate the putative enhancer region and the negative controls, respectively, tested for the GUS-based reporter system in **d**.

b, Relative chromatin loop formation in response to ABA in Col-0 and RNAi-*MARS* seedlings. Results are expressed as the mean \pm standard error (n = 2) from 3C-qPCR using primer F and R shown on **a**, compared to time 0h.

c, Relative chromatin loop formation in response to ABA treatment in Col-0 and *lhp1* mutant. Data are represented as the mean \pm standard error (n = 3) from 3C-qPCR using primer F and R shown on **a**, compared to time 0h.

d, Constructs used for GUS-based reporter system are illustrated on the left. Corresponding transformed tobacco leaf discs are on the right (n = 4). First line represents the positive control in which 35S sub-unit controls *GUS* expression. The second and third lines show two independent negative controls in which *GUS* gene is driven by a genomic region that does not contain ABA-related binding sites as indicated in **a**. In the remaining lines, the transcriptional activity is assessed for the two intergenic regions indicated in **a**.



High MARS RNA level

Low MARS RNA level

Fig. 6, Regulation of metabolic clusters in plants by IncRNA

a, The proportion of metabolic clusters including lncRNA loci is higher than for other clusters. The co-expressed clusters were predicted in ¹⁰ and correspond to co-expressed neighboring genes. The metabolic clusters are co-expressed neighboring genes involved in the biosynthesis of a particular secondary metabolite predicted by plantiSMASH³².

b, Maximum level of correlation between a IncRNA and any clustered gene calculated in each cluster. The number shown above indicates the p-value of the difference between the two type of clusters determined by Student t-test.

c, The IncRNA *MARS* regulates the expression of the marneral cluster genes through epigenetic reprogramming and chromatin conformation. In control conditions (upper panel) the chromatin of the marneral cluster is enriched in H3K27me3 and LHP1, which results in a condensed and possibly linear chromatin conformation. In response to ABA (bottom left panel) *MARS* over-accumulated transcripts titrate LHP1 away from the cluster. The decrease of LHP1 deposition diminishes H3K27me3 distribution, relaxes the chromatin and as a consequence allows the formation of a chromatin loop that brings together the enhancer element and *MRN1* proximal promoter, leading to transcriptional activation. When *MARS* is repressed, LHP1 recruitment to the cluster is impaired, thus leading to a similar chromatin loop conformation. Under this chromatin state, the clustered genes become highly responsive to the ABA treatment.