



HAL
open science

The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of stress-responsive genes

Anne M. Molitor, David Latrasse, Matthias Zytnicki, Philippe P. Andrey, Nicole Houba Hérin, Mélanie Hachet, Christophe Battail, Stefania del Prete, Adriana A. Alberti, Hadi Quesneville, et al.

► To cite this version:

Anne M. Molitor, David Latrasse, Matthias Zytnicki, Philippe P. Andrey, Nicole Houba Hérin, et al.. The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of stress-responsive genes. *The Plant cell*, 2016, 28 (9), 10.1105/tpc.16.00244 . hal-02633928

HAL Id: hal-02633928

<https://hal.inrae.fr/hal-02633928>

Submitted on 27 May 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 Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in
2 concert to regulate the transcription of stress-responsive genes

3

4

5

6 Short title

7 LIF2 and LHP1 target a common set of stress genes

8

9

10

11 Anne M. Molitor^{a*}, David Latrasse^{a*}, Matthias Zytnicki^{b*}, Philippe Andrey^{a,c}, Nicole
12 Houba-Hérin^a, Mélanie Hachet^b, Christophe Battail^d, Stefania Del Prete^a, Adriana
13 Alberti^e, Hadi Quesneville^b, Valérie Gaudin^{a,1}

14

15 ^a Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay,
16 F-78000 Versailles, France

17 ^b URGI, INRA, Université Paris-Saclay, F-78000 Versailles, France

18 ^c Sorbonne Universités, UPMC Université Paris 06, UFR927, F-75005, Paris, France

19 ^d CEA-Institut de Génomique-Centre National de Séquençage, F-91057 Evry, France

20 ^e CEA-Institut de Génomique, Genoscope, Centre National de Séquençage, F-91057
21 Evry, France

22

23 * Equally contributed

24 1 Corresponding author: Valérie Gaudin

25 Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay,
26 INRA Centre de Versailles-Grignon, RD10 Route de Saint-Cyr, F-78026 Versailles,
27 France

28 E-mail: valerie.gaudin@versailles.inra.fr - Tel +33 1 30 83 35 22

29 The author responsible for distribution of materials integral to the findings presented
30 in this article in accordance with the policy described in the Instructions for Authors

31 (www.plantcell.org) is: Valérie Gaudin (valerie.gaudin@versailles.inra.fr)

32

33

34

35 **Synopsis**

36

37 CHIP-seq analyses of the RBP LIF2 and its LHP1 partner in various backgrounds and
38 stress conditions revealed target regions for the two proteins enriched in antagonistic
39 marks.

40

41

42 **Abstract**

43 LHP1-INTERACTING FACTOR2 (LIF2), a heterogeneous nuclear ribonucleoprotein
44 involved in *Arabidopsis thaliana* cell fate and stress responses, interacts with LIKE
45 HETEROCHROMATIN PROTEIN1 (LHP1), a Polycomb Repressive Complex1
46 (PRC1) subunit. To investigate LIF2-LHP1 functional interplay, we mapped their
47 genome-wide distributions in wild-type, *lif2*, and *lhp1* backgrounds, under standard
48 and stress conditions. Interestingly, LHP1-targeted regions form local clusters,
49 suggesting an underlying functional organization of the plant genome. Regions
50 targeted by both LIF2 and LHP1 were enriched in stress-responsive genes, the
51 H2A.Z histone variant, and antagonistic histone marks. We identified specific motifs
52 within the targeted regions, including a G-box-like motif, a GAGA motif, and a *telo-*
53 *box*. LIF2 and LHP1 can operate both antagonistically and synergistically. In
54 response to methyl jasmonate treatment, LIF2 was rapidly recruited to chromatin,
55 where it mediated transcriptional gene activation. Thus, LIF2 and LHP1 participate in
56 transcriptional switches in stress-response pathways.

57

58 **Introduction**

59 In eukaryotes, the control of gene expression is central to development and
60 environmental adaptation. The establishment and maintenance of specific
61 transcriptionally active and repressive chromatin states participate in this control.
62 Polycomb Repressive Complexes (PRCs) and Trithorax (Trx) Complexes shape
63 chromatin states and have general transcriptional repressor and activator activities,
64 respectively (Simon and Kingston, 2013; Del Prete et al., 2015). Over the past few
65 years, the regulatory function of PRCs has been challenged in both plants and
66 animals (Tavares et al., 2012; Simon and Kingston, 2013; Calonje, 2014; Pu and
67 Sung, 2015; Forderer et al., 2016). For instance, novel PRC1 complexes have been
68 identified; the canonical model of PRC repression, in which PCR2-dependent H3K27
69 trimethylation is followed by PRC1-dependent H2A monoubiquitination, is no longer
70 regarded as the unique mode of action (Tavares et al., 2012; Calonje, 2014); and a
71 novel transcriptional activation function has been reported for PRC1 (Gil and
72 O'Loghlen, 2014).

73 However, the mechanism underlying the transition from active to repressed
74 chromatin states remains poorly understood. Documented recruitment of Polycomb
75 group proteins (PcG) to chromatin identified thousands of target regions in eukaryotic

76 genomes. In plants, the PRC1 subunit LHP1 is distributed throughout the genome
77 and co-localizes with the H3K27me3 repressive histone mark (Turck et al., 2007;
78 Zhang et al., 2007), as observed for animal PcG proteins. The distribution of
79 FERTILIZATION INDEPENDENT ENDOSPERM (FIE), a plant PRC2 subunit,
80 somewhat overlaps with H3K27me3 regions (Deng et al., 2013). The chromatin
81 context, which is determined by the combination of specific DNA motifs, histone
82 marks, or other chromatin-associated proteins, largely determines PcG recruitment.
83 For instance, *Drosophila* PRCs contain sequence-specific DNA-binding factors and
84 are classically recruited at Polycomb/Trithorax Response Elements (PRE/TREs or
85 PREs) in the genome, which are composed of a variable combination of short DNA
86 motifs and participate in the maintenance of the transcriptional status (Bauer et al.,
87 2015). Only a few PRE-like elements have been reported in mammals (Bauer et al.,
88 2015). PRCs interact with various chromatin-associated proteins, such as histone
89 modifying enzymes or transcription factors (TFs), which may also contribute to their
90 targeting. In plants, several TFs, such as SCARECROW, ASYMMETRIC LEAVES 1
91 (AS1), and AS2, interact with PRC subunits (reviewed in (Del Prete et al., 2015)).
92 Recently, the *A. thaliana* GAGA-binding factor BPC6 was shown to recruit LHP1 to
93 GAGA motifs (Hecker et al., 2015), reminiscent of the recruitment of PcG proteins to
94 GAGA motifs present in animal PREs. Finally, whereas some long non-coding RNAs
95 (lncRNAs) are involved in the scaffolding of chromatin modifying complexes
96 associated with PRC function (Brockdorff, 2013) or mediate intrachromosomal
97 interactions (Zhang et al., 2014), they also emerged as novel interacting partners of
98 both PRC2 and PRC1 subunits (Del Prete et al., 2015) that participate in their
99 genomic recruitment. In *A. thaliana*, lncRNAs were proposed to function in the
100 transcriptional regulation of *FLOWERING LOCUS C (FLC)* mediated by PcG proteins
101 (Swiezewski et al., 2009; Heo and Sung, 2011; Csorba et al., 2014). Intriguingly,
102 LIF2, a heterogeneous nuclear ribonucleoprotein Q (hnRNP-Q) with three RNA
103 recognition motifs (RRMs), was identified as a partner of LHP1 (Latrasse et al.,
104 2011), highlighting the diversity of plant proteins associated with PRC1, and
105 suggesting that RNA-binding proteins (RBPs) mediate interactions between plant
106 PRC1 and RNA components.

107 To investigate the interplay between LIF2 and LHP1, we compared the genome-wide
108 chromatin profiles of LIF2 and LHP1 in wild-type and mutant backgrounds. This is the
109 first report of the genome-wide chromatin profile of a plant RBP. Our ChIP-seq data
110 analyses revealed that LIF2 had a more restricted distribution than LHP1, being

111 mainly present at stress-responsive genes. The spatial analysis of LHP1 distribution
112 showed that LHP1 regions tend to aggregate locally, suggesting a role for LHP1 in
113 genome topography. Specific and antagonistic histone marks were associated with
114 each protein, as well as *cis*-regulatory DNA elements. We identified the GAGA motif
115 and *tel*-box motifs in the LHP1 target genes. Also present in FIE binding sites (Deng
116 et al., 2013), these two motifs may thus be part of a PRC targeting signature. Given
117 the role of LIF2 in pathogen responses (Le Roux et al., 2014), we investigated the
118 distribution of LIF2 in response to methyl jasmonate (MeJA), a key hormone in plant
119 biotic and abiotic stress responses. We showed that LIF2 distribution was dynamic in
120 response to MeJA treatment and that LIF2 was required for transcriptional gene
121 activation. Thus, we highlighted a complex interplay between LIF2 and LHP1 in
122 stress-response pathways.

123

124

125 **Results**

126

127 **LIF2 and LHP1 target a common set of chromatin regions**

128 Prompted by the observation that hnRNP-Q LIF2 physically interacts with the
129 chromatin-associated protein LHP1 (Latrasse et al., 2011), we performed ChIP-seq
130 experiments to identify the chromatin regions enriched in LIF2 and LHP1 (enrichment
131 regions, ERs). For this purpose, we produced transgenic lines expressing 3xHA-
132 tagged LIF2 (HA:LIF2) and 3xHA-tagged LHP1 (LHP1:HA) under the control of
133 endogenous genomic regulatory sequences, in the *lif2-1* and *lhp1-4* genetic
134 backgrounds, respectively. Two independent ChIP-seq libraries were sequenced for
135 each protein (Figure 1). We observed good overlaps between replicates, as well as
136 high Pearson coefficients of the MACS peak fold-change correlations between
137 replicates (0.93 (LIF2) and 0.81 (LHP1); Supplemental Figure 1). We identified 1457
138 ERs present in both replicates for LIF2 and 4844 for LHP1 (at a false discovery rate
139 (FDR) of < 0.05), and determined the summit (i.e., position with the maximum read
140 number) in each ER. The comparison of the two genome-wide distributions allowed
141 us to identify 488 genomic regions where LIF2 and LHP1 were detected,
142 corresponding to the intersection of the two genomic distributions (named LIF2-LHP1
143 IRs, intersect regions) (Figures 1 A and 1 B). We confirmed binding to ten ERs by
144 ChIP experiments followed by quantitative PCR (ChIP-qPCR) (Supplemental Figure
145 1). We established that 52.8% of the DamID-identified LHP1 target genes (Zhang et
146 al., 2007) were present in our ChIP-seq data set (despite differences in tissue,
147 developmental stages, and growth conditions). These data suggest that LIF2 has a
148 more specialized function in the genome than does LHP1, with each protein having
149 independent and specific functions. However, in agreement with their physical
150 interactions, a subset of genomic regions was identified where the two proteins were
151 located.

152

153 **LIF2 is present in narrow chromatin regions in 5' and 3' UTRs**

154 LIF2 and LHP1 exhibited different chromatin-associated profiles. Whereas the LIF2
 155 profile had narrow, discrete peaks, LHP1 peak sizes were larger (Figure 1 C). By
 156 analyzing the distribution of LIF2 and LHP1 over annotated genomic features and
 157 comparing this distribution with a random distribution over genome regions of similar
 158 size, we found that LIF2 had a preference for 5' UTRs (2.52-fold compared to LIF2-
 159 random), exons (especially exon 1) (Supplemental Figure 2), and 3' UTRs (4-fold

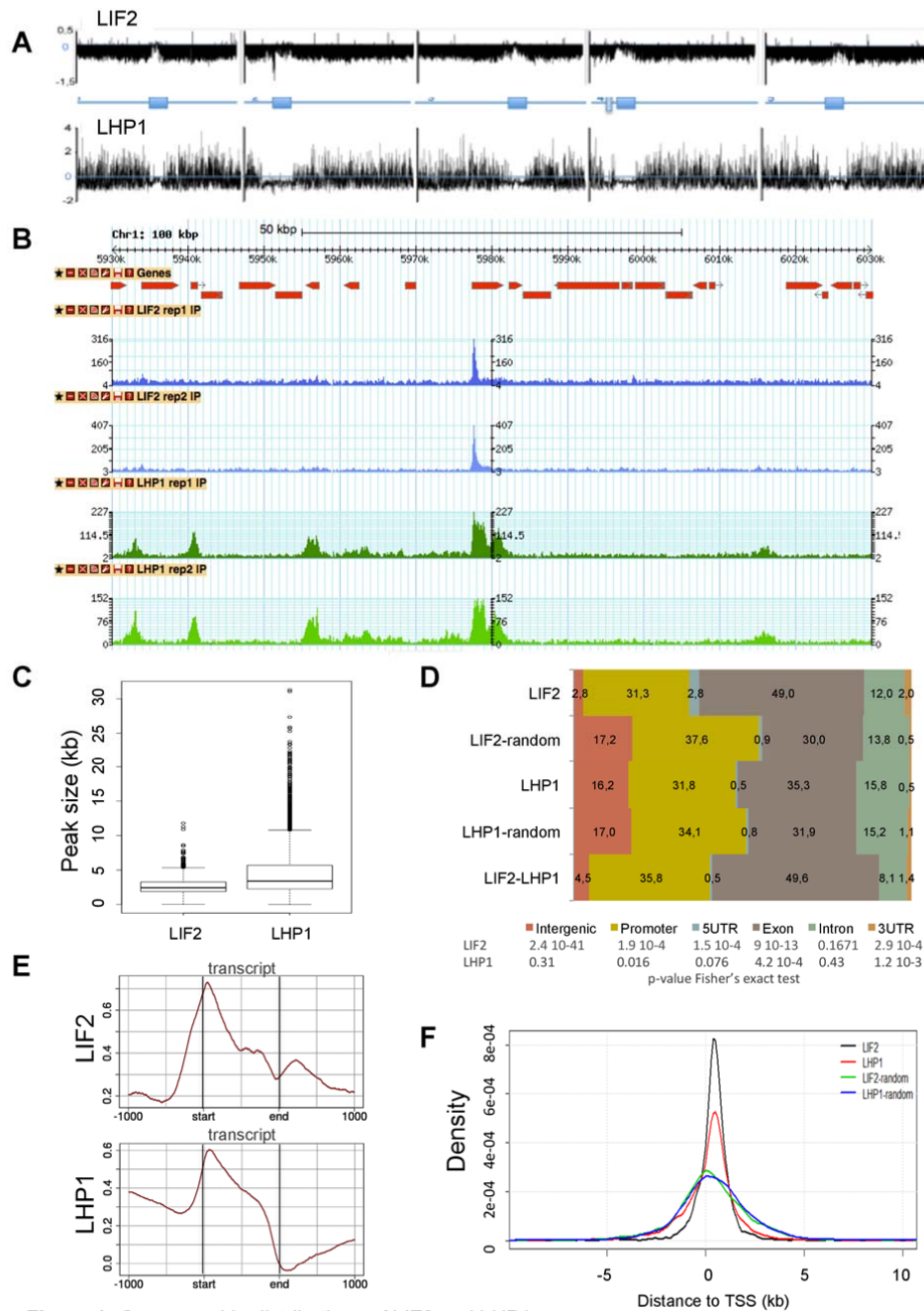


Figure 1: Genome-wide distributions of LIF2 and LHP1.

(A) Chromosomal view of the peaks using model-based analysis.

(B) Screenshot of a 100-kb window with the distributions.

(C) Size distributions of the ERs defined as intersects of MACS peaks for the biological replicates.

(D) Distributions of ER-associated annotations (percentage). Regions with identical sizes were randomly shuffled in the genome and compared with the observed ERs, using a Fisher's exact test.

(E) Distributions of IP enrichment ($\log_2(\# \text{ reads IP} / \# \text{ read input})$) over the transcript structures.

(F) Distance to closest transcriptional start sites (TSS) of LIF2 and LHP1 ERs, and the corresponding randomized regions.

160 compared to LIF2-random), whereas LHP1 had a more balanced distribution over all
161 regions (Figure 1 D). Interestingly, the distribution of the LIF2-LHP1 IRs was similar
162 to that of the LIF2 ERs. Our analysis of the peak distributions over transcripts
163 showed that LIF2 was enriched at transcription start sites (TSSs) and depleted at
164 transcription termination sites (TTSs) (Figure 1 E). The low but significant level of
165 LIF2 downstream TTSs was not due to the proximity of another TSS. LHP1 was
166 more prevalent at promoter regions and gene bodies, with a marked preference for
167 TSSs, resulting in an asymmetry between upstream and downstream genic regions
168 (Figure 1 E). The presence of LIF2 and LHP1 ERs in regions close to the TSS was
169 confirmed with the analysis of the distance of the ERs to the closest TSS compared
170 to the randomly shuffled control regions (Figure 1 F).

171

172 **The targeted regions tend to form clusters**

173 Given the role of PcG proteins in structuring the genome in animal species (Del Prete
174 et al., 2015), we analyzed the distribution of LHP1 and LIF2 along each
175 chromosome. The distribution of the number of ER summits in 1-Mb windows
176 revealed that portions of the genome were enriched for LHP1 and LIF2 (Figure 2 A,
177 Supplemental Figure 3). To analyze the distribution of LHP1 ERs further and to test
178 the existence of an underlying organization principle in this distribution, we compared
179 the summit distribution of LHP1 ERs to a random distribution model, conditioned on
180 the size of LHP1 ER regions. Observed and model-predicted distributions were
181 compared using local-scale (i.e., cumulative distribution of the distance between
182 each ER and its closest neighbor) and global-scale (i.e., cumulative distribution of the
183 inter-distances between all ERs) spatial descriptors. A significant discrepancy with
184 the completely random model was observed at the local scale, with the measured
185 distances between ERs and their closest neighbors being significantly smaller than
186 expected under a random distribution for all chromosome arms (Figure 2 B,
187 Supplemental Figure 4). Compared to the random distribution, the distance to the
188 closest ER was enriched in the range of short values of up to ~10 kb (Figure 2 C).
189 This range was constant across chromosome arms, suggesting the existence of
190 common spatial constraints despite differences in arm length and ER density.
191 Overall, no significant difference to the random distribution was observed (Figure 2
192 D, Supplemental Figure 5) when comparing the distribution of all inter-distances
193 (Figure 2 D, Supplemental Figure 4), consistent with the globally uniform distribution
194 of LHP1 ERs in 1-Mb windows (Figure 2 A, Supplemental Figure 3). LIF2 ERs

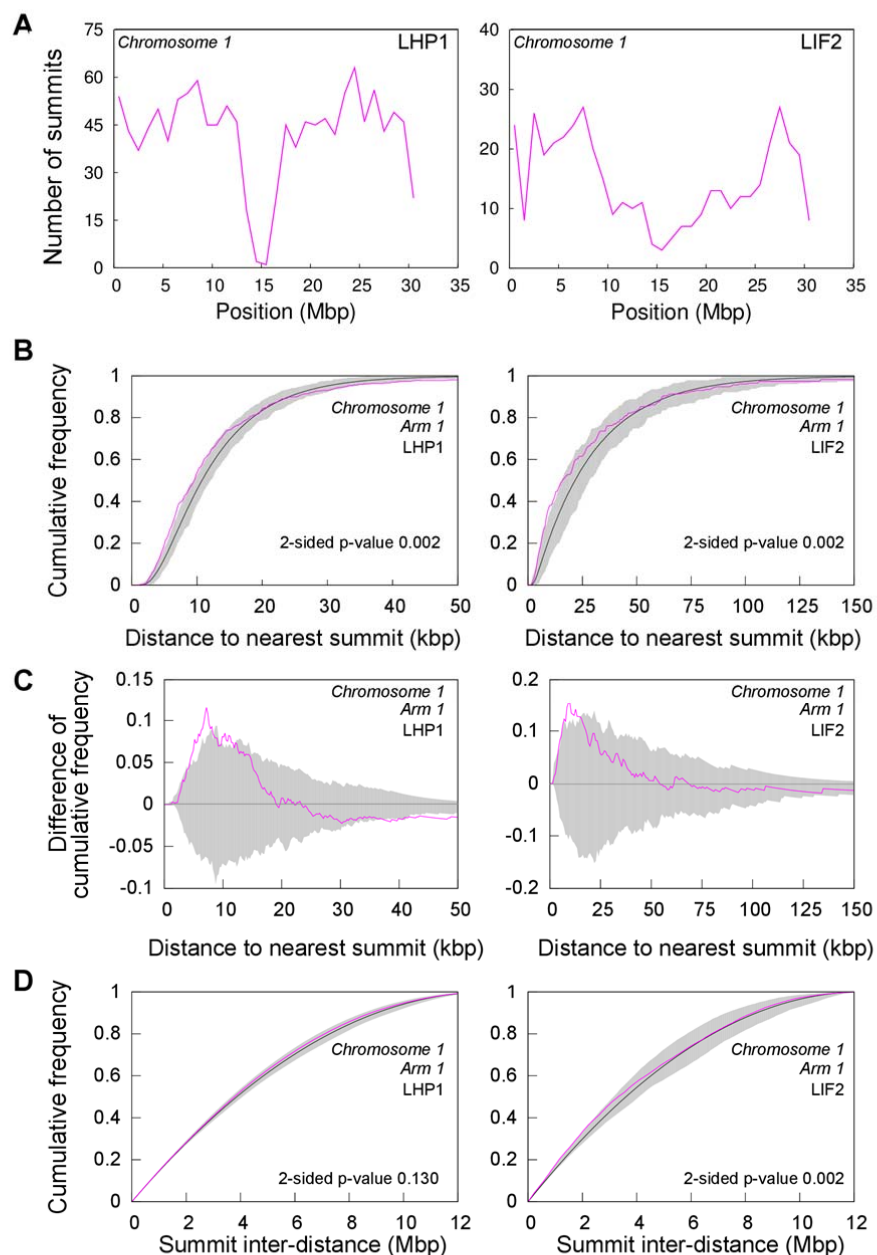


Figure 2: Non-random distributions of the LHP1 ERs and LIF2 ERs in the *A. thaliana* genome. (A) Number of summits in 1-Mb windows along Chromosome 1. (B-D) Observed (pink) and random model (black: average; grey: 95% envelope) distributions of distance to nearest ER (B-C) and of all ER inter-distances (D), on the first arm of Chromosome 1. Similar results were obtained for all chromosome arms (Supplemental Figures 2, 3 and 4).

195 exhibited spatial clustering that was similar to that of LHP1 ERs. Despite the lower
 196 density of LIF2 ERs, the range of distances between nearest neighbors was similar
 197 to that observed for LHP1 ERs (Figure 2 C), suggesting that the distributions of the
 198 two proteins' target regions were under shared constraints.

199 To further investigate the clustering of LHP1 ERs, we analyzed the relationship
200 between the LHP1 ER genome-wide distribution and the distribution of repeated
201 genes in the *A. thaliana* genome. Indeed, repeated genes may participate in this
202 clustering tendency. It was previously shown using ChIP-chip experiments that out of
203 the 679 tandemly repeated genes located on chromosome 4, 30% were targeted by
204 LHP1 (Turck et al., 2007). In the whole *A. thaliana* genome, 1564 tandem
205 duplications (T-clusters) and 1680 segmental duplications (with a 1:1 duplication
206 relation, S-clusters) were described (Haberer et al., 2004). The T-clusters contain
207 from 2 to up to 21 repeated genes, with a mean value <3 genes. Using our ChIP-seq
208 data, we observed that 20.6% of the T-cluster genes were targeted by LHP1,
209 compared to 11.7% for the S-cluster genes. However, only 23.1% of the LHP1-
210 targeted genes were located in T-clusters. On average, there was less than one
211 LHP1 target gene per T-cluster and only 11% of T-clusters had two or more LHP1
212 target genes, accounting for only 9.6% of all LHP1 targets (Supplemental Table 1).
213 These low figures suggest that LHP1 binding to T-cluster genes is not sufficient to
214 explain the clustering tendency of LHP1-targeted regions observed at the local scale
215 on the chromosome arms.

216

217 **The presence of antagonistic histone marks and H2A.Z characterize LIF2-LHP1** 218 **IRs**

219 A limited number of chromatin states, which are based on histone post-translational
220 modifications (PTMs) or histone variants, have been reported for the *A. thaliana*
221 genome (Sequeira-Mendes et al., 2014). We thus examined whether specific
222 epigenetic marks were preferentially associated with the identified ERs, using data
223 sets for nine histone marks (Luo et al., 2012) and the H2A.Z histone variant
224 (Zilberman et al., 2008; Coleman-Derr and Zilberman, 2012). We observed that
225 LHP1 ERs were enriched in the repressive mark H3K27me3, confirming our previous
226 genome-wide analysis (Zhang et al., 2007), and were depleted in active histone
227 marks, such as H3K4me3 (Figure 3, Supplemental Figure 6). By contrast, LIF2 ERs
228 were enriched in H3K4me3 and H3K9ac histone marks, which are hallmarks of
229 active/open chromatin. Interestingly, a similar enrichment in H3K4me3 and
230 H3K27me3 was observed in LIF2-LHP1 IRs, and this was associated with a
231 noticeable depletion in the active mark H3K36me3 compared to LIF2 ERs. LIF2 and
232 LHP1 ERs also had similar levels of H2A.Z, with LIF2-LHP1 IRs having slightly
233 higher levels. In *A. thaliana*, H2A.Z is enriched within the nucleosomes surrounding

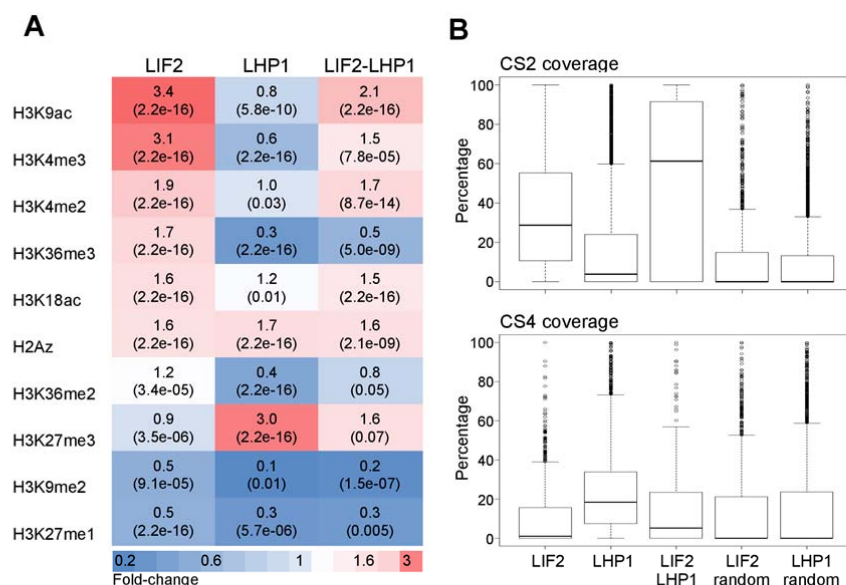


Figure 3. Post-translational histone modifications (PTMs) and the H2A.Z histone variant in the LIF2 ERs, LHP1 ERs or LIF2-LHP1 IRs .

(A) Heat map presenting the fold changes (p-value paired t-test) between targeted and randomized regions.

(B) Percentage of chromatin states 2 and 4 (CS2 and CS4; defined by Sequeira-Mendes et al., 2014) covering LHP1 ERs, LIF2 ERs, and LIF2-LHP1 IRs, and randomized control regions.

234 the TSSs of genes (Zilberman et al., 2008), but also across the bodies of genes with
 235 low transcription levels and high responsiveness (Coleman-Derr and Zilberman,
 236 2012). Our data suggest that LIF2-LHP1 IRs may correspond to subdomains of
 237 chromatin state 2 (CS2), which is characterized by relatively high levels of both active

238 H3K4me3 and inactive H3K27me3 histone marks and is mostly associated with
239 bivalent regions and highly constrained gene expression (Sequeira-Mendes et al.,
240 2014). We thus analyzed the coverage of CS2 in the distributions of LHP1 and LIF2
241 ERs and compared this coverage with CS4 coverage, CS4 having high levels of
242 H3K27me3, but reduced levels of active marks. We confirmed enrichments in CS2
243 for both LIF2 IRs and LIF2-LHP1 IRs (Figure 3 B). By comparing the lists of LIF2 and
244 LHP1 target genes with the bivalent genes identified by sequential ChIP experiments
245 (Luo et al., 2012), we observed that about 14.92% of the LIF2-LHP1 IR genes have
246 been annotated as bivalent (Luo et al., 2012), whereas only 4.97% and 5.97% of the
247 LIF2 and LHP1 target genes have been annotated as bivalent, respectively. Thus,
248 the genome-wide distributions of LIF2 and LHP1 contributed to the functional
249 topographical organization of the *A. thaliana* genome (Sequeira-Mendes et al., 2014).

250

251 **LIF2-LHP1 IRs are enriched in stress-responsive genes**

252 To predict the functions of genes of LIF2 ERs and LIF2-LHP1 IRs, we determined the
253 gene responsiveness index of the binding regions based on the expression profiles of
254 the genes located in the ERs (Aceituno et al., 2008; Coleman-Derr and Zilberman,
255 2012). We found that they were enriched in responsive genes (Figure 4 A). Our
256 analysis of the functional Gene Ontology (GO) terms revealed that LIF2 ERs and
257 LIF2-LHP1 IRs were enriched in stress-responsive genes (Figure 4 A, Supplemental
258 Figures 7 and 8). The Bio-Array Resource for Plant Functional Genomics (BAR)
259 classification Superviewer program (Provart et al., 2003) showed that both LIF2 ERs
260 and LIF2-LHP1 IRs were enriched in the GO term “response to abiotic or biotic
261 stimulus” (normed frequency (NF); LIF2 NF=3, p-value $1.399 \cdot 10^{-78}$; LIF2-LHP1
262 NF=2.7, p-value $2.662 \cdot 10^{-19}$) (Supplemental Table 2). A more detailed analysis using
263 AgriGO revealed that the first two enriched GO terms for LIF2 ERs were “aromatic
264 compound catabolic process” (NF=29.35, FDR $9.4 \cdot 10^{-5}$, p-value $3.8 \cdot 10^{-6}$) and
265 “callose deposition in cell wall during defense response” (NF=16.57, FDR $1.2 \cdot 10^{-4}$, p-
266 value $5.1 \cdot 10^{-6}$). For LIF2-LHP1 IRs, they were “response to chitin” (NF=4.98, FDR 5.9
267 $\cdot 10^{-12}$, p-value $1.7 \cdot 10^{-14}$) and “regulation of defense response” (NF=3.56, FDR $1.3 \cdot 10^{-3}$,
268 p-value $9.5 \cdot 10^{-5}$) (Supplemental Figure 7). These results were in agreement with
269 those of our previous transcriptome analysis of the *lif2* mutant (Latrasse et al., 2011)
270 and the response of *lif2* to pathogens (Le Roux et al., 2014), but also with the
271 epigenetic marks present at LIF2-LHP1 IRs. Indeed, genes present in CS2 were
272 shown to have constrained transcription profiles (Sequeira-Mendes et al., 2014).

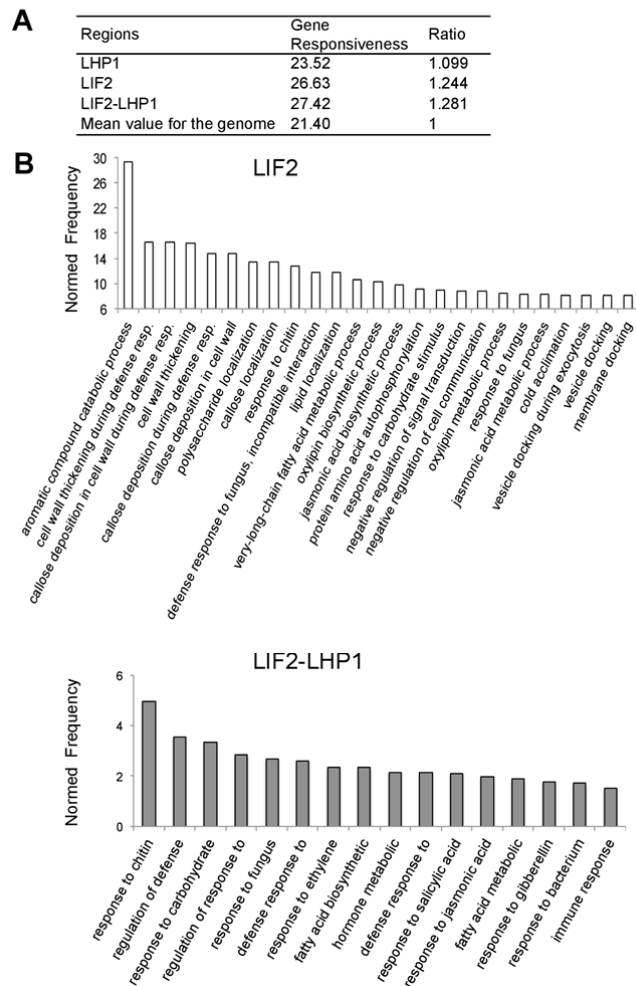


Figure 4: LIF2 binds preferentially stress-response genes.

(A) Average gene responsiveness scores were calculated based on a published data set²⁵ and normalized to the genome-wide average.

(B) GO analysis of LIF2 ERs and LIF2-LHP1 IRs using the AgriGO toolkit. The biological process GO terms, with the 25 best normed frequencies (NF) and with $NF \geq 1.5$ are presented for LIF2 ERs and LIF2-LHP1 IRs, respectively.

273 Further GO term analysis revealed an enrichment in “transcription factor activity” in
 274 the LIF2 ER and LIF2-LHP1 IR datasets ($NF=2.39$, $p\text{-value } 3.730 \cdot 10^{-18}$ and $NF=3.7$,
 275 $p\text{-value } 2.739 \cdot 10^{-17}$, respectively) (Supplemental Table 2). Using the Plant GeneSet
 276 Enrichment Analysis (PlantGSEA) (Yi et al., 2013) and Arabidopsis Gene Regulatory

277 Information Server (AGRIS) (Yilmaz et al., 2011) toolkits, we observed that the main
278 TFs targeted by LIF2 belonged to the AP2-EREBP and WRKY families, consistent
279 with a role for LIF2 in the stress response, whereas TFs in LHP1-ERs belonged to a
280 larger range of families (Supplemental Table 3). Interestingly, TFs present in the 488
281 LIF2-LHP1 IRs also belonged to the AP2-EREBP family. Some of the target genes
282 were also targeted by other TFs, such as LONG HYPOCOTYL 5 (HY5) (Lee et al.,
283 2007; Zhang et al., 2011) and PHYTOCHROME-INTERACTING FACTOR1 (PIF1)
284 (Chen et al., 2013) (Supplemental Table 4), suggesting a complex interplay between
285 LIF2, LHP1, and TFs.

286

287 **Identification of *cis*-regulatory DNA elements associated with LIF2 and LHP1** 288 **binding**

289 We next searched for putative DNA-binding motifs around the summits. Using the
290 MEME algorithm (Bailey and Elkan, 1995), two consensus motifs were discovered in
291 the 51-bp regions centered on the LIF2-binding summits: a GAGA-like motif and a
292 (C/G)ACGTG(G/T)C(A/G) consensus motif, which belongs to the ACGT-containing
293 element (ACE) family (Figure 5). The ACGTGGCA word was present at moderate
294 levels in the whole genome, mostly in the distal promoter regions of genes (region
295 from -1000 bp to -3000 bp relative to the TSS) (Supplemental Table 5). Some of the
296 ACE elements are recognized by TFs, among which HY5 and PIF1 (Song et al.,
297 2008; Chen et al., 2013), previously identified as having common targets with LIF2
298 (Supplemental Table 4) and two physically interacting TFs involved in plant growth
299 and, in particular, in the crosstalk between light and reactive oxygen species (ROS)
300 signaling. In the LHP1 datasets, we identified a GAGA-like motif as a putative
301 recognition motif (Figure 5). In addition, we identified the (A/G_T)AACCCCTA(A/G)
302 motif. Despite being less represented among the LHP1 peaks, this putative and
303 highly significant DNA motif ($-\log_{10}(\text{E-value}) > 20$) was discovered with both MEME
304 and “peak-motif” algorithms (Bailey and Elkan, 1995; Thomas-Chollier et al., 2012)
305 (Figure 5 B). This motif contains the AAACCCTA short interstitial telomere motif, also
306 named the *te/o*-box, which was originally described in the 5' regions of genes
307 encoding the translation elongation factor EF1 α and ribosomal proteins (Regad et al.,
308 1994; Gaspin et al., 2010). The AAACCCTA word/*te/o*-box is mainly present in
309 introns and 5' UTRs (Supplemental Table 5). Interestingly, the (A/G_T)AACCCCTA(A/G)
310 motif recognized by LHP1 was present in a LHP1-target subset, which was enriched
311 in the molecular function GO term “nucleic acid binding transcription factor activity”

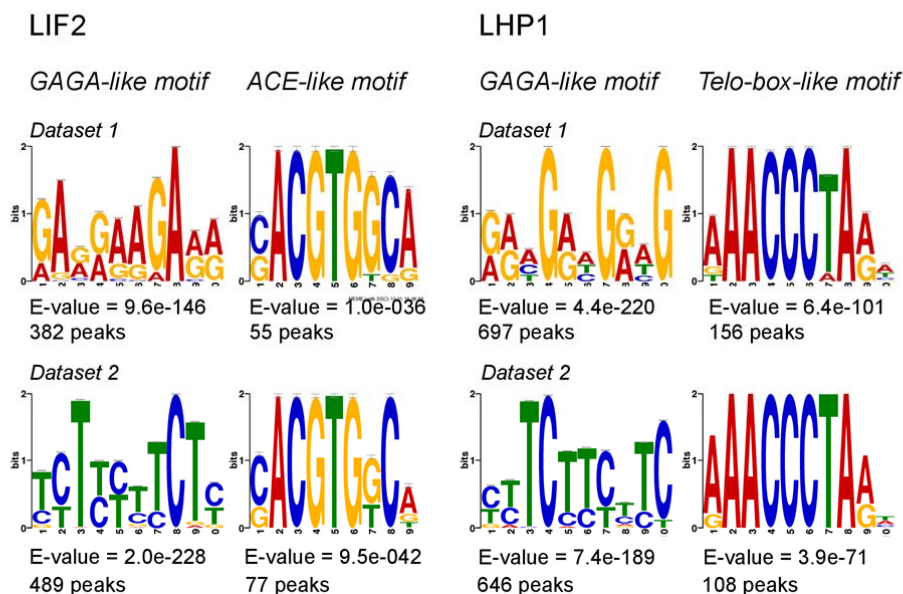


Figure 5: Identification of putative *cis*-regulatory DNA motifs in LIF2 ERs and LHP1 ERs. The regions centered on LIF2 and LHP1 summits were used to screen for putative targeting motifs. The *E*-value of MEME program is an estimate of the expected number of motifs with the given log likelihood ratio (or higher), and with the same width and number of occurrences, that one would find in a similarly sized set of random sequences.

312 (GO:0001071, fold-enrichment 3.62, p-value $2.67 \cdot 10^{-03}$) and in the biological process
 313 GO term “carpel development” (GO:0048440, fold-enrichment >5, p-value $4.04 \cdot 10^{-02}$)(Panther classification system), suggesting the existence of a small and
 314 specialized subset of LHP1 targets containing the (A/G)_TAACCCTA(A/G) *telo*-box-like
 315

316 motif. Interestingly, *TELOMERE REPEAT BINDING PROTEIN1 (TRB1)* also binds to
317 the AAACCCTA motif and it was proposed that TRB1 may act as a transcriptional
318 repressor in the absence of LHP1 (Zhou et al., 2015).

319

320 **LIF2 has a major transcriptional activation activity on its targets**

321 To better understand the mode of action of LIF2, we compared the binding profiles of
322 LIF2 with our previous transcriptome data obtained from the seedlings and rosette
323 leaves of *lif2* and *lhp1* mutants (Latrasse et al., 2011) (Figure 6). We observed a bias
324 towards down-regulated genes among LIF2 targets (23.8%), suggesting that LIF2
325 had a global transcriptional activator role on its own targets. The *lif2* mutation had no
326 significant impact on the transcription of LHP1 target genes, whereas LHP1 had a
327 general repressor activity on LIF2 targets (25.5% of the LIF2 targets were
328 deregulated in the *lhp1* mutant). A proportion of genes located in the LIF2-LHP1 IRs
329 were activated by LIF2 and repressed by LHP1, suggesting that LIF2 and LHP1 have
330 general antagonistic transcriptional roles in activation and repression, respectively.
331 Nevertheless, small sets of LIF2-LHP1 IR genes were down-regulated in the mutants
332 and enriched in stress response-associated GO terms (Figure 6 G), suggesting that
333 LIF2 and LHP1 can also act synergistically to activate specific genes.

334

335 **A complex interplay between LIF2 and LHP1 recruitments**

336 To investigate the impact of LIF2 and LHP1 on each other's binding, we crossed the
337 complemented mutant lines expressing tagged LIF2 or LHP1 with the *lif2-1 lhp1-4*
338 double mutant and selected transgenic lines in single mutant backgrounds (named
339 *lif2* LHP1 and *lhp1* LIF2). We performed ChIP-seq experiments and identified regions
340 that exhibited differences in the binding of the tagged proteins compared with the
341 binding in the original complemented mutant lines (*lif2* LIF2 and *lhp1* LHP1). The
342 analysis revealed a strong bias towards a depletion of any protein binding in the
343 double mutant background (Figure 7 A). In the absence of LHP1, LIF2 binding
344 decreased strongly in regulatory regions (UTRs and promoters) and increased
345 strongly in gene bodies (exons). Similar findings were observed for LHP1 (Figure 7
346 B). The gene set depleted in LIF2 binding in the *lhp1* background was enriched in
347 stress-related genes, and in the GO term "transcription repressor activity"
348 (GO:0016564, NF 17.1, p-value $8.9 \cdot 10^{-07}$; Supplemental Table 6). Since
349 modifications of the binding of one protein at a precise locus in the mutant
350 background could result from a direct loss of binding of the other or from indirect

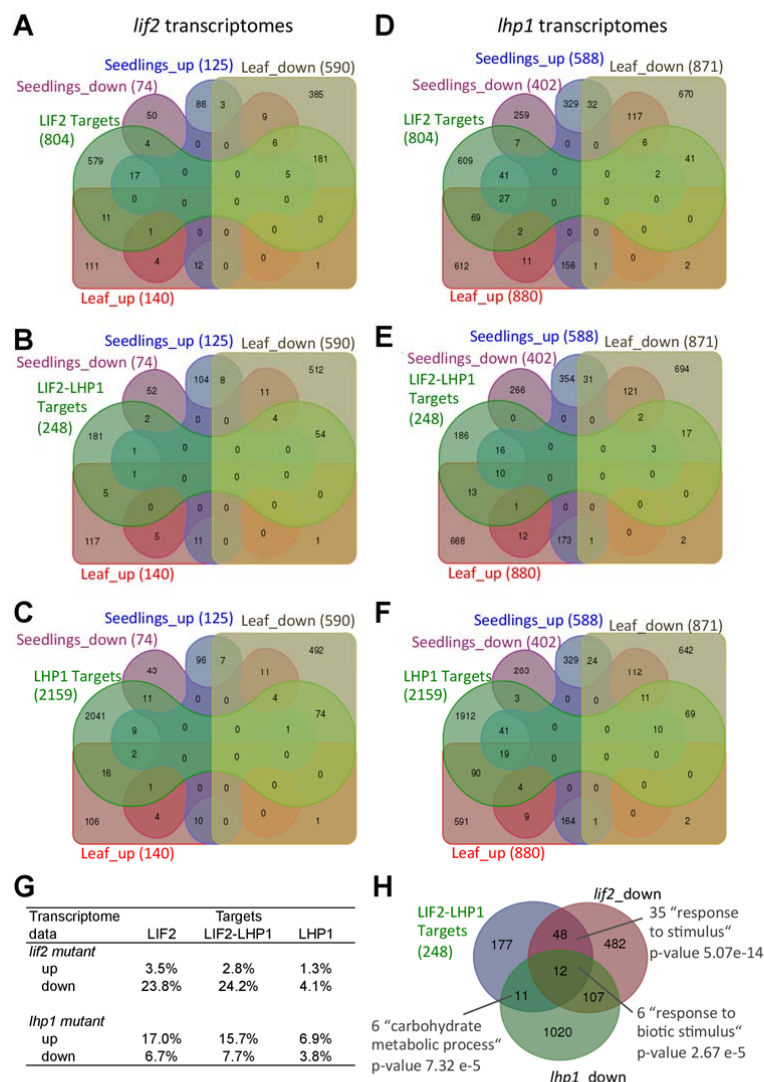


Figure 6: LIF2 functions mainly as a transcriptional activator on its targets. (A-F) Venn diagrams between genes of LIF2 ERs (A, D), LIF2-LHP1 IRs (B, E, G) and LHP1 ERs (C, F) and deregulated genes in vegetative tissues of the *lif2* (A, B, C) and *lhp1* (D, E, F) mutants. The analysis involved genes for which the binding was located in CDSs or in UTRs. (G) Comparisons between target genes and deregulated genes in *lif2* and *lhp1* mutants. (H) Venn diagram and GO annotations of LHP1-LIF2 IR genes and genes activated by LHP1 and LIF2, respectively, revealed a small set of genes that requires a synergistic and activation function of both LIF2 and LHP1.

351 effects of its loss of function, we focused our analysis on the LIF2-LHP1 IR genes
 352 identified in the first part of this study (Figure 1). Among these LIF2-LHP1 IR genes,
 353 we identified three subsets that presented an alteration in LHP1 and LIF2 binding in
 354 the *lif2* LHP1 and *lhp1* LIF2 backgrounds (Figure 7 C). The three sets were enriched

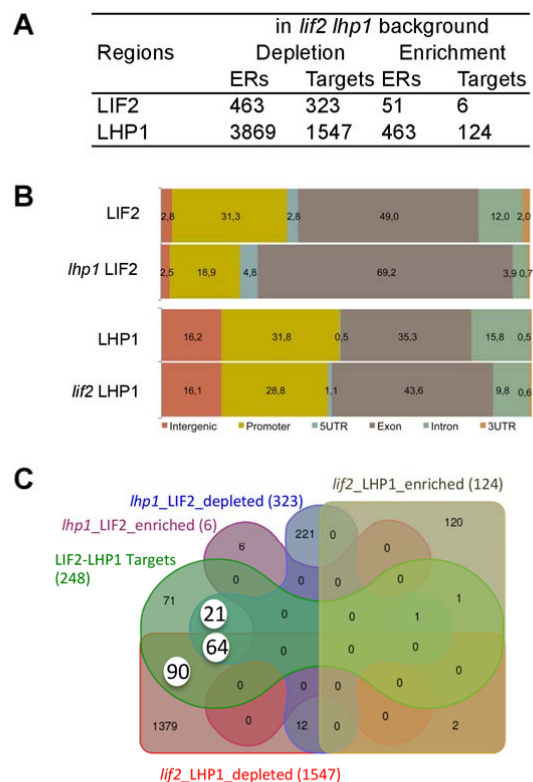


Figure 7: Complex interplay between LIF2 and LHP1 for their recruitment.

(A) LIF2 and LHP1 binding in the mutant backgrounds.

(B) Distribution of the annotations of the targeted regions.

(C) Venn diagram highlighting Set-21, Set-64 and Set-90 (white circles), which contain LIF2-LHP1 IR genes, depleted in one or the other protein, in the mutant backgrounds.

355 in stress response-associated GO terms (Supplemental Figure 8). In Set-64 (64
 356 genes), the presence of both proteins was mutually required for their binding,
 357 suggesting a synergistic mode of action, whereas for Set-90 and Set-21, LIF2 and
 358 LHP1, respectively, were necessary for the presence of the other one. Therefore,

359 these data suggest a prominent role for LIF2 in LHP1 recruitment to chromatin and
360 regulation in the LIF2-mediated stress response pathway. This role may be
361 underestimated, as we only considered locations occupied by the two proteins under
362 normal physiological conditions. Most of the genes of the three sets were not
363 deregulated in our *lif2* and *lhp1* transcriptomes (Latrasse et al., 2011). This might be
364 due to redundant mechanisms of gene regulation. Furthermore, transcriptome
365 profiles, established in mutants under normal physiological conditions, may not
366 highlight deregulation in responses to various cues. However, 45.3% of the Set-64
367 genes were down-regulated in *lif2*, in agreement with the major transcriptional activity
368 of LIF2 (Supplemental Figure 8 D).

369

370 **Rapid recruitment of LIF2 in response to methyl jasmonate**

371 Due to the enrichment in stress GO terms, such as “JA-mediated signaling pathway”,
372 in both *lif2* transcriptomes (Le Roux et al., 2014) and LIF2 ERs (Supplemental Figure
373 7), we investigated whether JA treatment affects LIF2 recruitment to chromatin by
374 comparing ChIP-seq data obtained from plants subjected or not to JA treatment. For
375 the JA treatment, we used a short-term (1 h) oxylipin-derived methyl jasmonate
376 (MeJA) treatment to avoid complex downstream regulatory events, as a 1-h
377 treatment was sufficient to transcriptionally activate JA-inducible marker genes in
378 wild-type plants (Supplemental Figure 9). For each protein, we identified a reduced
379 number of regions with binding modifications in response to MeJA (JA-ERs), and
380 observed a bias toward enrichments in LIF2 and LHP1 in response to MeJA (Figure
381 8 A). Short-term MeJA treatment promoted LIF2 binding in promoter and intergenic
382 regions and LHP1 binding at 5' UTRs (Figure 8 B). Interestingly, after MeJA
383 treatment, the ERs that exhibited the greatest enrichment in LIF2 or LHP1 were
384 enriched in “transcription factor activity” GO term, and also in the “energy pathway”
385 GO term for LIF2 ERs (Figure 8 C). When the JA ERs were compared to LIF2-LHP1
386 IRs under normal conditions, only a limited number of loci were identified, suggesting
387 that we had access to very early regulatory events, in agreement with the observed
388 enrichment in TFs, and/or that both proteins have independent functions in response
389 to MeJA (Figure 8 D). Alternatively, the use of a gene set in which both proteins
390 might already be present before the treatment introduced a bias in the analysis.

391 To further characterize LIF2 binding in response to MeJA, we examined the
392 expression of JA-inducible genes, *MYC2*, *JASMONATE-ZIM DOMAIN (JAZ1, JAZ6,*
393 *JAZ9)*, *VEGETATIVE STORAGE PROTEIN2 (VSP2)*, and *LIPOXYGENASE3*

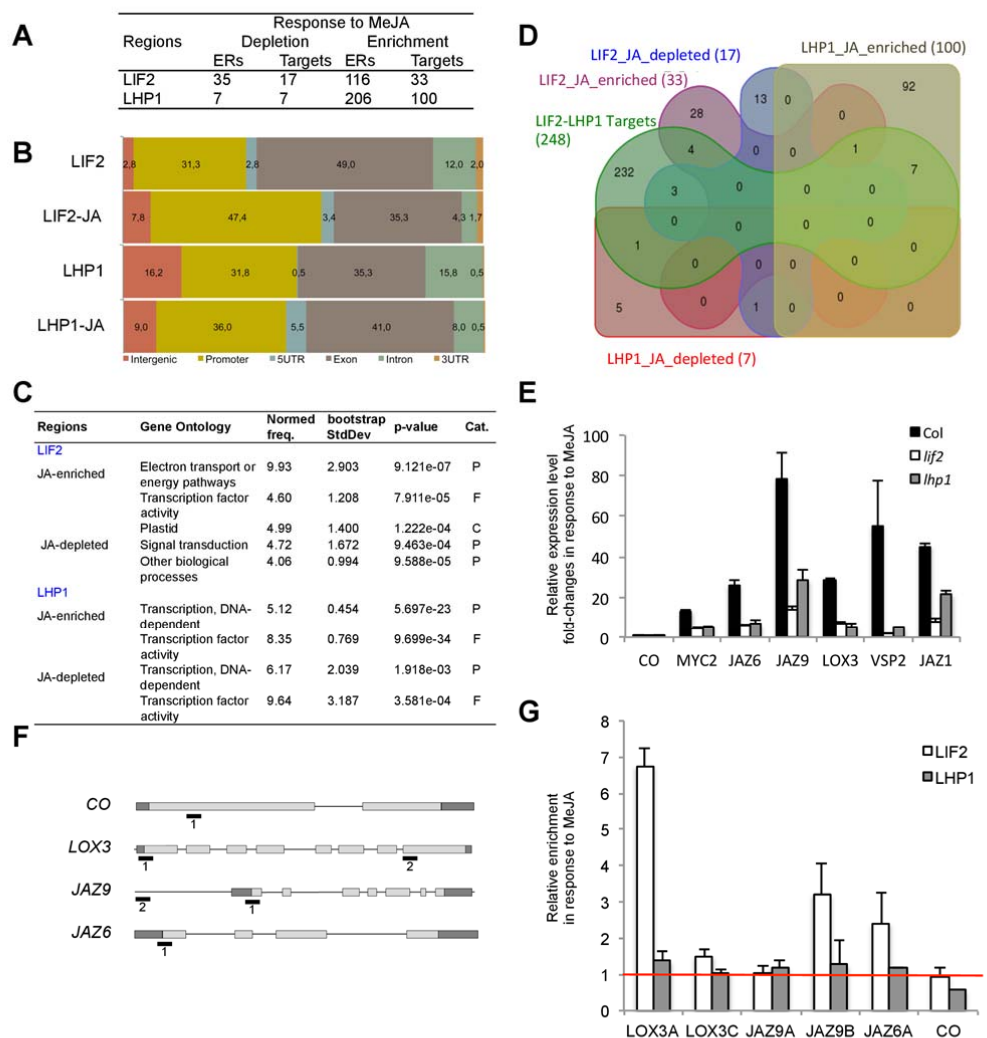


Figure 8: LIF2 and LHP1 binding in response to MeJA.

A 1-h MeJA treatment was performed on two-week-old seedlings.

(A) Dynamics of LIF2 and LHP1 binding in response to MeJA.

(B) Distribution of the annotations of the binding regions.

(C) GO terms with NF>4 (AgriGO toolkit). Cat.: category; P: process; F: function; C: cellular component.

(D) Venn diagram with the genes of the LIF2-LHP1 IRs.

(E) Fold changes of the relative expression in response to MeJA in the mutant backgrounds of stress-related genes. Mean \pm SEM. Three biological replicates were performed.

(F-G) Relative enrichments of LIF2 and LHP1 in response to MeJA. The targeted regions (i.e., 1, 2) are indicated in the schematic representation (F). ChIP-QPCR experiments (G). Three biological replicates were performed.

394 (*LOX3*). *LOX3* is among the bivalent genes identified by sequential ChIP (Luo et al.,
 395 2012). These genes were up-regulated in wild-type, *lif2-1*, and *lhp1-4* plants in
 396 response to MeJA treatment; however, the activation levels were higher in wild-type
 397 plants than in any of the mutants (Figure 8 E). Under normal growth conditions, LIF2

398 and LHP1 were present on *LOX3*, a gene present in Set-64, whereas *JAZ6* and *JAZ9*
399 were only targeted by LIF2 (our CHIP-seq data). These data suggested that the two
400 proteins were cooperatively recruited to *LOX3* (our CHIP-seq data). Upon MeJA
401 treatment, LIF2 binding increased in the TSS regions of the three loci, whereas LHP1
402 binding was not significantly affected (Figure 8 F). These data revealed that the early
403 events of the transcriptional activation of the three JA-inducible genes require LIF2
404 recruitment. The presence of LHP1 on *LOX3* seemed to be required to reach a full
405 level of activation, as suggested by *LOX3* expression in the *lhp1* mutant, but its
406 distribution on the locus was not significantly affected. Therefore, LHP1 seems to be
407 required for the early transcriptional events of JA-dependent activation of *LOX3* by
408 LIF2. Whether long-term treatments would impact LHP1 binding requires further
409 investigation.

410

411

412

413 Discussion

414 Dynamic switches that mediate the transition between active and inactive chromatin
415 states are crucial for the development and adaptation of organisms. PRC and TRX
416 complexes, with their antagonistic effects on transcriptional gene regulation, play a
417 crucial role in these chromatin-associated transitions. Chromatin may be regarded as
418 a bistable system composed of two main antagonistic chromatin states, and
419 transitory intermediate chromatin states. The mechanism by which chromatin
420 changes from one state to another remains poorly understood. To decipher this
421 mechanism, we studied two interacting partners, LHP1, a plant PRC1 subunit, and
422 the LIF2 hnRNP-Q protein. Our comparative analysis of their genome-wide binding
423 profiles in wild-type and mutant backgrounds and under normal and stress
424 conditions, and of their transcriptomes, revealed that these two proteins interact in a
425 complex manner to control gene transcription.

426 Contrasting profiles were obtained for these interacting proteins: LHP1 was
427 distributed over large genomic regions similar to histone marks, while LIF2 occurred
428 in narrow binding regions, mainly located in promoters and in proximity to TSSs,
429 which is reminiscent of TF binding at precise regulatory DNA elements. Furthermore,
430 whereas LHP1 ERs were associated with the Polycomb H3K27me3 mark, as we
431 previously reported using the DamID approach (Zhang et al., 2007), LIF2 was
432 present in chromatin states characterized by the presence of H3K9ac and H3K4me3,
433 which are usually associated with active/open chromatin.

434 The LIF2-LHP1 IRs were identified at the intersection of LIF2 and LHP1 protein
435 distributions. However, to pursue and fully demonstrate that they are simultaneously
436 binding to the exact same chromatin fiber, further analyses, such as sequential ChIP
437 experiments, would be required. The LIF2-LHP1 IRs were associated with
438 antagonistic marks, which may correspond to bivalent regions (Sequeira-Mendes et
439 al., 2014) or to intermediate heterochromatin such as telomeric heterochromatin
440 (Vrbsky et al., 2010; Vaquero-Sedas et al., 2012). Interestingly, 9.8% of the H3K9ac
441 target genes in *A. thaliana* are also marked by H3K27me3 (Zhou et al., 2010;
442 Karmodiya et al., 2012) and H3K9ac is present in bivalent chromatin regions of
443 mouse promoters of developmentally regulated genes (Karmodiya et al., 2012).
444 LHP1 interacts with MSI1 (Derkacheva et al., 2013), which associates with histone
445 deacetylase 19 (HDAC19) in the same *in vivo* complex, to maintain a low H3K9ac
446 level at genes involved in the ABA signalling pathway (Mehdi et al., 2015).
447 Furthermore, the LIF2-LHP1 IRs had a good coverage with Chromatin State 2 and

448 were enriched in stress-responsive genes, demonstrating that the LIF2/LHP1 duo
449 seems to have a specialized function in the stress response pathway and a putative
450 role in maintaining or regulating a distinctive chromatin state at a specific gene set.
451 Furthermore, the binding maps of each protein, established in the absence of its
452 partner, revealed various scenarios that were highly dependent on the genomic
453 contexts, with synergistic binding, as well as binding dependent on one or the other
454 protein.

455 The identification of GAGA motifs in LHP1 ERs confirmed a recent discovery (Hecker
456 et al., 2015). Indeed, the BASIC PENTACYSSTEINE6 (BPC6) GAGA-binding factor
457 interacts with LHP1 and recruits LHP1 at GAGA motif-containing DNA probes *in vitro*
458 (Hecker et al., 2015). Interestingly, GAGA motifs were also present in FIE ERs ¹⁰, as
459 was Motif 2, which is similar to a *telo*-like box. The presence of these two types of
460 motifs in LHP1 and FIE ERs suggests the existence of common recruitment motifs
461 between plant PRC1 and PRC2 subunits, but also between PRC1 and LIF2. Thus,
462 these different DNA motifs may correspond to modules that participate to form
463 putative plant PREs.

464 In addition to establishing the global rules governing LIF2 and LHP1 binding, we
465 observed that the two proteins exhibited different recruitment dynamics in response
466 to a short-term MeJA treatment. A rapid increase in LIF2 binding was observed,
467 especially at the TSS of *LOX3*, *JAZ6*, and *JAZ9*, with an associated increase in gene
468 expression. These data were in agreement with the global down-regulation of LIF2
469 targets in *lif2*. At *LOX3*, the presence of LHP1 was not modulated by the MeJA
470 treatment, but LHP1 was required for LIF2-mediated activation. Removal of LHP1
471 was not a prerequisite for the early transcriptional activation, suggesting that the two
472 proteins may have different kinetics of action. Thus, one hypothesis would be that the
473 RNA-binding protein LIF2 functions in transcriptional activation, especially in JA-
474 dependent activation, and may counteract gene repression via its interaction with
475 LHP1. Further investigation is needed to understand this dynamic and complex
476 interplay. For instance, it remains unclear how LIF2 specifically interacts with
477 chromatin. Perhaps this interaction is mediated by RRM. Indeed, RRMs are plastic
478 protein domains and some RRMs also have DNA-binding properties (Enokizono et
479 al., 2005; Grinstein et al., 2007; Wan et al., 2007). Alternatively, RNA molecules
480 interacting with RRMs may participate in RNA/DNA recognition, and thus help target
481 RBP via their interaction with RNA molecules. Since RNA molecules play diverse
482 functions in modulating animal PRC activities, further investigation of putative

483 interactions between LIF2 and RNA molecules will be of key importance.
484 Finally, we showed that LHP1 ERs had a significant and robust tendency to form
485 clusters (in the ~10 kb range), regardless of the chromosome arm identity. Due to the
486 large number of LHP1 ERs in the genome, the distribution of LHP1 clusters may not
487 be neutral and may influence the functional organization of the genome. Indeed,
488 proteins in the HP1 family have dimerization properties and SWI6 even has an
489 oligomerization property, which contributes to heterochromatin formation (Canzio et
490 al., 2011). Thus a clustering of the LHP1 ERs may have 3D consequences on
491 genome organization. In animals, PcG proteins contribute to the modular
492 organization of the linear epigenome, but also to the 3D genome organization
493 (Cavalli, 2014; Del Prete et al., 2015). In *A. thaliana*, recent HiC studies highlighted
494 long-range genome interactions, but the absence of large chromatin modules as
495 observed in animal genomes (Feng et al., 2014; Grob et al., 2014), possibly due to
496 resolution limitations. Although restricted to one dimension, our approach in this
497 study, in which spatial statistics are applied to genome-wide data, represents a
498 complementary tool for deciphering eukaryotic genome organization. It allowed us to
499 evaluate distribution patterns of chromatin-associated proteins at different scales and
500 highlighted the existence of short-range clusters on the linear organization of the *A.*
501 *thaliana* genome. It will be interesting to determine whether the linear proximity of
502 LHP1 ERs contributes to the formation of LHP1 foci (Gaudin et al., 2001), promotes
503 silent plant chromatin formation, or influences the 3D genome organization.
504

505 **Methods**

506 **Materials and hormonal treatment**

507 All *Arabidopsis thaliana* lines used in this study are in the Col-0 background. The *lif2-*
508 *1* and *lhp1-4* mutants were previously described (Latrasse et al., 2011). For all
509 experiments, plants were grown *in vitro* for 14 days under controlled long-day
510 conditions as previously described (Gaudin et al., 2001). For methyl jasmonate
511 (MeJA) treatments a filter paper was imbibed with 10 µl of 95% MeJA (Sigma-Aldrich,
512 Ref. 392707) and placed in a Petri dish. Plates were hermetically sealed and placed
513 for 1 h under identical growth conditions. MeJA treated and mock seedlings were
514 either directly harvested for gene expression analyses or fixed for ChIP assays after
515 the 1-h treatment. All primers are listed in Supplemental Table 7.

516

517 **Plasmid constructs**

518 For the 3xHA:LIF2 binary construct, the 3xHA tag was PCR amplified from the
519 pGWB15 vector (Invitrogen) using the 3HA-1 and 3HA-2 primers bearing *Pst*I and
520 *Xba*I restriction sites, respectively. After digestion and purification, the 3xHA fragment
521 was inserted into the pCambia1300 vector giving the pCa-HA vector. The Nos
522 terminator, amplified from plasmid pUC-SPYNE (Walter et al., 2004) using the Nost-1
523 and Nost-2 primers (bearing *Kpn*I and *Eco*R I sites, respectively) was digested, gel-
524 purified, and inserted into the *Kpn*I/*Eco*R I digested pCa-HA vector. A 3-kbp promoter
525 region of *LIF2* (including the first three codons) was amplified from the T18A10 BAC
526 plasmid (ABRC DNA stock center) using primers AD379-28 and AD379-29 (bearing
527 a *Pst*I restriction site). The *Pst*I-digested LIF2 promoter fragment was gel-purified and
528 inserted into the pCa-HA-tNos vector at the *Pst*I and blunt-made *Hind*III sites. Finally,
529 the *LIF2* genomic region was amplified from T18A10 using the primers AD379-30
530 and AD379-32, digested with *Xho*I and inserted into the *Sal*I/*Sma*I-digested pCa-
531 ProLIF2:HA-tNos vector giving the pCa-ProLIF2:HA:LIF2-tNos vector (N-terminal HA-
532 tagged gLIF2).

533 For the ProLHP1:LHP1:HA binary construct, a 3xHA fragment was PCR amplified
534 from the pGWB15 vector (Invitrogen) using the 3HA-1 and 3HA-2 primers and
535 digested with *Eco*RV and *Xho*I. The 3xHA fragment was inserted into the *Eco*RV
536 restriction site of the vector bearing a 5569-bp genomic LHP1 fragment (Latrasse et
537 al., 2011). Subsequently, the *Nco*I/*Bst*EII fragment containing the LHP1:3xHA-tagged
538 region was substituted to the wild-type genomic fragment of the pCaSSP vector

539 giving the gLHP1:HA binary plasmid (C-terminal HA tagged gLHP1). All subcloning
540 steps were confirmed by sequencing. Col-0 plants were transformed by floral dip. For
541 each construct, homozygous transgenic lines with wild-type phenotypes were
542 selected, in which the functional HA-tagged protein was detected.

543

544 **RNA extraction**

545 Total RNA was isolated from 14-day-old *in vitro*-grown seedlings, subjected or not to
546 MeJA treatment, using the RNeasy Plant Mini Kit (QIAGEN) according to supplier's
547 instructions. Total RNA (1-2 µg) was treated with RNase-free DNaseI (Invitrogen)
548 and reverse transcribed with Superscript II reverse transcriptase (Invitrogen).

549

550 **Quantitative real-time PCR**

551 Relative levels of cDNA (RT-qPCR) and immunoprecipitated DNA fragments (ChIP-qPCR)
552 were analyzed by quantitative real-time PCR on an Eppendorf Mastercycler®ep Realplex
553 using SsoAdvanced™ SYBR® (Biorad). Immunoprecipitated DNA levels were normalized to
554 input and to the internal reference gene *EF1* (AT5G60390). The cDNA levels were
555 normalized to *EF1*.

556

557 **ChIP library construction and sequencing**

558 ChIP assays were performed on five grams of 14-day-old *in vitro* seedlings from
559 transgenic lines expressing LHP1-HA or LIF2-HA in the single or double mutant
560 genetic backgrounds, using a previously published protocol (Latrasse et al., 2011),
561 with the following minor modifications. Chromatin was immunoprecipitated overnight
562 using high affinity anti-HA antibody (Roche, Ref. 11867423001). Immunoprecipitated
563 DNA enrichment was controlled by quantitative real-time PCR (qPCR). DNA quantity
564 and quality were checked using a Qubit fluorometer (ThermoFisher Scientific,
565 Waltham, MA) and Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara,
566 CA). Several independent experiences were pooled for library construction. Then,
567 10-15 ng of immunoprecipitated DNA was fragmented to a 100-500 bp range using
568 the E210 Covaris instrument (Covaris, Woburn, MA). Libraries were prepared
569 according to the Illumina standard procedure using the NEBNext DNA Sample
570 Preparation Reagent Set 1 (New England Biolabs, Ipswich, MA) and homemade
571 ligation adaptors. The ligated product was amplified by 12 cycles of PCR using
572 Platinum Pfx DNA Polymerase (ThermoFisher Scientific). Amplified material was

573 purified using Agencourt Ampure XP beads (Beckmann Coulter Genomics, Danvers,
574 MA). Libraries were then quantified by qPCR and library profiles were evaluated
575 using an Agilent 2100 Bioanalyzer. Two independent libraries for each protein were
576 sequenced using 100 base-length read chemistry in a paired-end flow cell on the
577 HiSeq2000 (Illumina, San Diego, CA).

578

579 **ChIP-seq data analyses**

580 After Illumina sequencing, Illumina read processing and quality filtering were
581 performed. An in-house quality control process was applied to reads that passed the
582 Illumina quality filters. Low quality nucleotides ($Q < 20$) were discarded from both
583 ends of the reads. Next, Illumina adapter and primer sequences were removed from
584 the reads. Then, reads shorter than 30 nucleotides after trimming were discarded.
585 These trimming and removing steps were achieved using internal software based on
586 the `FastX` package (`FASTX-Toolkit`,
587 http://hannonlab.cshl.edu/fastx_toolkit/index.html). This processing yields high quality
588 data and improves subsequent analyses. The sequencing reads were uniquely
589 mapped to the Arabidopsis genome (TAIR10; <http://www.arabidopsis.org>) using
590 Bowtie 4.1.2 mapper (Langmead et al., 2009) with default mismatch parameters, and
591 retaining only reads mapping uniquely to the genome for further analysis. The main
592 heterochromatic regions of the genome were thus excluded from our analysis.

593 To identify biologically relevant binding regions, peak prediction and normalization
594 were performed using MACS1.4.1 (Zhang et al., 2008) and peak analysis was
595 performed using S-MART (Langmead et al., 2009) or the “`annotatePeaks.pl`”
596 software from Homer (<http://homer.salk.edu/homer>; Heinz et al., 2010). High-
597 confidence target regions (i.e., enriched regions, ERs) were defined as strict overlap
598 of the MACS peaks from the corresponding biological replicates.

599 By default, a TSS region was defined from -1 kb to +100 bp from TSS and the TTS
600 region was defined from -100 bp to +1 kb from the TTS. The process of annotating
601 peaks/regions was divided into two primary parts. The first determined the distance
602 to the nearest TSS and assigned the peak to that gene. The second determined the
603 genomic annotation of the region occupied by the center of the peak/region.

604

605 **Bioinformatics analyses**

606 Motifs were predicted using the integrated online pipeline "peak-motifs"
607 (<http://plants.rsat.eu/>; Thomas-Chollier et al., 2011; Thomas-Chollier et al., 2012).
608 Briefly, 50 and 300 bp surrounding protein-binding summits were scanned for a
609 global overrepresentation of words (oligo-analyses) or spaced words (dyad-
610 analyses). Then, 5000 random, artificial 300-bp long sequences were generated by
611 the "RSAT-random sequence tool" (<http://plants.rsat.eu/>) and were used as
612 background control for motif discovery. In parallel, sequences were analyzed by the
613 motif prediction program "MEME" (Bailey and Elkan, 1995). The word occurrence
614 was determined using the word frequency program in AtcisDB from AGRIS
615 (<http://arabidopsis.med.ohio-state.edu/AtcisDB/>).

616 The functional annotation and classification of gene populations was carried out
617 using the online "AgriGO" gene Ontology tool (<http://bioinfo.cau.edu.cn/agriGO/>)
618 using pre-set parameters. Venn diagrams were generated using the online tool
619 provided by T. Hulsen (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

620 To analyze the histone mark enrichments over the ERs, ChIP-seq data presented in
621 (Luo et al., 2012) were used and available at SRA under IDs GSM701923-701931.
622 Raw data were mapped onto the TAIR10 genome with the Bowtie mapper
623 (Langmead et al., 2009) (unique hits, 1 mismatch at most). Mapped reads were
624 processed using SAMtools (Li et al., 2009) and BEDtools (Quinlan and Hall, 2010).
625 The number of reads per bp of the selected loci was counted and compared with that
626 of randomized loci (using the shuffle BEDtool). Fold enrichment/depletion was
627 calculated as the ratio between the mean read number in regions of interest versus
628 randomized regions. Statistical significance was assessed by t-tests. Boxplots
629 represent distribution of histone mark ChIP-seq reads within LIF2, LHP1, LIF2-LHP1,
630 and the corresponding randomized regions.

631

632 **Spatial distributions of the targeted regions**

633 The spatial distributions of LHP1 and LIF2 targeted regions were quantified and
634 analyzed for each individual biological replicate, using the cumulative distribution
635 functions of (1) the distance to the nearest neighbor of each targeted region and (2)
636 the inter-distance between every pair of targeted regions. Departure from
637 randomness was assessed by adapting a Monte Carlo procedure developed for 3D
638 data (Andrey et al., 2010). Observed distributions were compared to distributions
639 obtained under complete randomization of targeted regions without overlap (999

28

Comment citer ce document :

Molitor, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zytnicki, M. (Co-premier auteur) Andrey, P., Houba Héryn, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of

640 randomizations for computing averages of distance functions under randomness, 999
641 further randomizations for computing envelopes around averages). The relative
642 position of the empirical distance function within the range of variations under
643 randomness was used to estimate p-values (Andrey et al., 2010).

644

645 **Accession Numbers**

646 Sequencing data were deposited at NCBI, under the Sequence Read Archive (SRA)
647 number SRP068984.

648

649 **Supplemental data**

650

651 **Supplemental Figure 1:** ChIP-seq experiments.

652

653 **Supplemental Figure 2:** Exon distributions of LIF2 ERs.

654

655 **Supplemental Figure 3:** Distributions of the number of summits in 1-Mb windows.

656

657 **Supplemental Figure 4:** Cumulative distribution of the distance to the nearest LHP1
658 summit.

659

660 **Supplemental Figure 5:** Cumulative distribution of LHP1 summit inter-distances.

661

662 **Supplemental Figure 6:** Post-translational histone modifications and their
663 distributions in LIF2 ERs, LHP1 ERs and LIF2-LHP1 IRs.

664

665 **Supplemental Figure 7:** GO term analysis of the target loci of LIF2 and LHP1.

666

667 **Supplemental Figure 8:** Analyses of the LIF2-LHP1 IRs with binding alterations in
668 the mutant backgrounds.

669

670 **Supplemental Figure 9:** Expression kinetics of JA-induced marker genes in
671 response to MeJA treatment in wild-type plants.

672

673 **Supplemental Table 1:** Tandem duplications and LHP1 target genes.

674

675 **Supplemental Table 2:** GO term analysis of the genes present in LIF2 ERs and
676 LIF2-LHP1 IRs using the Plant Functional Genomics (BAR) classification
677 Supreviewer program.

678

679 **Supplemental Table 3:** Enrichments of LIF2 and LHP1 targets in specific
680 transcription factor families using the PlantGSEA resource.

681

682 **Supplemental Table 4:** LIF2 and LHP1 targets are also bound by specific
683 transcription factors.

684

685 **Supplemental Table 5:** Occurrences of the two identified DNA words.

686

687 **Supplemental Table 6:** GO term analysis of LIF2 or LHP1 depleted regions in the
688 mutant backgrounds (AgriGO).

689

690 **Supplemental Table 7:** List of primers.

691

692 **Acknowledgments**

693 We thank Bruno Letarnec and Hervé Ferry for plant care in the greenhouses, Dr.
694 Georg Haberer for providing the S-cluster listing and Dr. Crisanto Gutierrez for
695 helpful exchange. We are grateful to our colleagues, Dr. Franziska Turck and Dr.
696 Dierk Wanke, for critical reading of the manuscript and suggestions. D.L., A.M. and
697 M.H. were supported by fellowships from the ANR (ANR-08-BLAN-0200,
698 Polycombara) from the French Research Ministry. S.D.L. was supported by a PhD
699 fellowship provided by the European Commission Seventh Framework-People-2012-
700 ITN project EpiTRAITS (Epigenetic regulation of economically important plant traits,
701 no-316965). The Génoscope supported the sequencing in the frame of a large-scale
702 DNA sequencing project (N°11). The IJPB benefits from the support of the LabEx
703 Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS).

704

705 **Author contributions**

706 A.M., D.L., M.Z., P.A., N.H.H., M.H., C.B., S.D.P., A.A., V.G. performed the
707 experiments. A.M., M.Z., M.Z., P.A., M.H., H.Q., V.G. analyzed the data. A.M., M.Z.,

708 P.A., V.G. prepared the figures. A.M., V.G. wrote the manuscript.

709

710 **Competing financial interests**

711 The authors declare no competing financial interests.

712

713 **Figure legends**

714

715 **Figure 1:** Genome-wide distributions of LIF2 and LHP1.

716 (A) Chromosomal view of the peaks using model-based analysis.

717 (B) Screenshot of a 100-kb window with the distributions.

718 (C) Size distributions of the ERs defined as intersects of MACS peaks for the
719 biological replicates.

720 (D) Distributions of ER-associated annotations (percentage). Regions with identical
721 sizes were randomly shuffled in the genome and compared with the observed ERs,
722 using a Fisher's exact test.

723 (E) Distributions of IP enrichment ($\log_2(\# \text{ reads IP} / \# \text{ read input})$) over the transcript
724 structures.

725 (F) Distance to closest transcriptional start sites (TSS) of LIF2 and LHP1 ERs, and
726 the corresponding randomized regions.

727

728 **Figure 2:** Non-random distributions of the LHP1 ERs and LIF2 ERs in the *A. thaliana*
729 genome.

730 (A) Number of summits in 1-Mb windows along Chromosome 1.

731 (B-D) Observed (pink) and random model (black: average; grey: 95% envelope)
732 distributions of distance to nearest ER (B-C) and of all ER inter-distances (D) on the
733 first arm of Chromosome 1. Similar results were obtained for all chromosome arms
734 (Supplemental Figures 3, 4 and 5).

735

736 **Figure 3.** Post-translational histone modifications (PTMs) and the H2A.Z histone
737 variant in the LIF2-ERs, LHP1 ERs, and LIF2-LHP1 IRs.

738 (A) Heat map presenting the fold changes (p-value paired t-test) between targeted
739 and randomized regions.

740 (B) Percentage of chromatin states 2 and 4 (CS2 and CS4; defined by Sequeira-
741 Mendes et al., 2014) covering LHP1 ERs, LIF2 ERs, LIF2-LHP1 IRs, and randomized
742 control regions.

743

744 **Figure 4:** LIF2 binds preferentially stress-response genes.

745 (A) Average gene responsiveness scores were calculated based on a published data
746 set (Aceituno et al., 2008) and normalized to the genome-wide average.

747 (B) GO analysis of LIF2 ERs and LIF2-LHP1 IRs using the AgriGO toolkit. The
748 biological process GO terms, with the 25 best normed frequencies (NF) and with
749 $NF \geq 1.5$ are presented for LIF2 ERs and LIF2-LHP1 IRs, respectively.

750

751 **Figure 5:** Identification of putative *cis*-regulatory DNA motifs in LIF2 ERs and LHP1
752 ERs.

753 The regions centered on LIF2 and LHP1 summits were used to screen for putative
754 targeting motifs. The *E*-value of MEME program is an estimate of the expected
755 number of motifs with the given log likelihood ratio (or higher), and with the same
756 width and number of occurrences, that one would find in a similarly sized set of
757 random sequences.

758

759 **Figure 6:** LIF2 functions mainly as a transcriptional activator on its targets.

760 (A-F) Venn diagrams between genes of LIF2 ERs (A, D), LIF2-LHP1 IRs (B, E, G),
761 and LHP1 ERs (C, F) and deregulated genes in vegetative tissues of the *lif2* (A, B, C)
762 and *lhp1* (D, E, F) mutants. The analysis involved genes for which the binding was
763 located in CDSs or in UTRs.

764 (G) Comparisons between target genes and deregulated genes in *lif2* and *lhp1*
765 mutants.

766 (H) Venn diagram and GO annotations of LHP1-LIF2 IR genes and genes activated
767 by LHP1 and LIF2, respectively, revealed a small set of genes that requires a
768 synergistic and activation function of both LIF2 and LHP1.

769

770 **Figure 7:** Complex interplay between LIF2 and LHP1 for their recruitment.

771 (A) LIF2 and LHP1 binding in the mutant backgrounds.

772 (B) Distribution of the annotations of the targeted regions.

773 (C) Venn diagram highlighting Set-21, Set-64, and Set-90 (white circles), which
774 contain LIF2-LHP1 IR genes, depleted in one or the other protein, in the mutant
775 backgrounds.

776

777 **Figure 8:** LIF2 and LHP1 binding in response to MeJA.

778 A 1-h MeJA treatment was performed on two-week-old seedlings.

779 (A) Dynamics of LIF2 and LHP1 binding in response to MeJA.

780 (B) Distribution of the annotations of the binding regions.

781 (C) GO terms with NF>4 (AgriGO toolkit). Cat.: category; P: process; F: function; C:
782 cellular component.

783 (D) Venn diagram with the genes of the LIF2-LHP1 IRs.

784 (E) Fold changes of the relative expression in response to MeJA in the mutant
785 backgrounds of stress-related genes. Mean±SEM. Three biological replicates were
786 performed.

787 (F-G) Relative enrichments of LIF2 and LHP1 in response to MeJA. The targeted
788 regions (i.e., 1, 2) are indicated in the schematic representations (F). ChIP-QPCR
789 experiments (G). Three biological replicates were performed.

790

Parsed Citations

Aceituno, F.F., Moseyko, N., Rhee, S.Y., and Gutierrez, R.A. (2008). The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics* 9, 438.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Andrey, P., Kieu, K., Kress, C., Lehmann, G., Tirichine, L., Liu, Z., Biot, E., Adenot, P.G., Hue-Beauvais, C., Houba-Herlin, N., Duranthon, V., Devinoy, E., Beaujean, N., Gaudin, V., Maurin, Y., and Debey, P. (2010). Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei. *PLoS Computational Biology* 6, e1000853.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bailey, T.L., and Elkan, C. (1995). The value of prior knowledge in discovering motifs with MEME. *Proc Int Conf Intell Syst Mol Biol* 3, 21-29.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Bauer, M., Trupke, J., and Ringrose, L. (2015). The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma*.

Brockdorff, N. (2013). Noncoding RNA and Polycomb recruitment. *RNA* 191, 429-442.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Calonje, M. (2014). PRC1 marks the difference in plant PcG repression. *Molecular Plant* 7, 459-471.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Molecular Cell* 41, 67-81.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Cavalli, G. (2014). Chromosomes: now in 3D! *Nature Reviews Molecular Cell Biology* 15, 6.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Chen, D., Xu, G., Tang, W., Jing, Y., Ji, Q., Fei, Z., and Lin, R. (2013). Antagonistic basic helix-loop-helix/bZIP transcription factors form transcriptional modules that integrate light and reactive oxygen species signaling in *Arabidopsis*. *The Plant Cell* 25, 1657-1673.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Coleman-Derr, D., and Zilberman, D. (2012). Deposition of histone variant H2AZ within gene bodies regulates responsive genes. *PLoS Genetics* 8, e1002988.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Csorba, T., Questa, J.I., Sun, Q., and Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proceedings of the National Academy of Sciences of the United States of America* 111, 16160-16165.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Del Prete, S., Mikulski, P., Schubert, D., and Gaudin, V. (2015). One, Two, Three: Polycomb Proteins Hit All Dimensions of Gene Regulation. *Genes* 6, 520-542.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Deng, W., Buzas, D.M., Ying, H., Robertson, M., Taylor, J., Peacock, W.J., Dennis, E.S., and Helliwell, C. (2013). *Arabidopsis* Polycomb Repressive Complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes. *BMC Genomics* 14, 593.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Molitor, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zytnicki, M. (Co-premier auteur) Andrey, P., Houba Hérlin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A.,

Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The *Arabidopsis* hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgova, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W., and Hennig, L. (2013). **Arabidopsis MSI1 connects LHP1 to PRC2 complexes.** *The EMBO journal* 32, 2073-2085.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Enokizono, Y., Konishi, Y., Nagata, K., Ouhashi, K., Uesugi, S., Ishikawa, F., and Katahira, M. (2005). **Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D.** *The Journal of biological chemistry* 280, 18862-18870.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Feng, S., Cokus, S.J., Schubert, V., Zhai, J., Pellegrini, M., and Jacobsen, S.E. (2014). **Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis.** *Molecular Cell* 55, 694-707.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Forderer, A., Zhou, Y., and Turck, F. (2016). **The age of multiplexity: recruitment and interactions of Polycomb complexes in plants.** *Current Opinion in Plant Biology* 29, 169-178.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gaspin, C., Rami, J.F., and Lescure, B. (2010). **Distribution of short interstitial telomere motifs in two plant genomes: putative origin and function.** *BMC plant biology* 10, 283.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gaudin, V., Libault, M., Pouteau, S., Juul, T., Zhao, G., Lefebvre, D., and Grandjean, O. (2001). **Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis.** *Development* 128, 4847-4858.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Gil, J., and O'Loughlen, A. (2014). **PRC1 complex diversity: where is it taking us?** *Trends in Cell Biology* 24, 632-641.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Grinstein, E., Du, Y., Santourlidis, S., Christ, J., Uhrberg, M., and Wernet, P. (2007). **Nucleolin regulates gene expression in CD34-positive hematopoietic cells.** *The Journal of biological chemistry* 282, 12439-12449.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Grob, S., Schmid, M.W., and Grossniklaus, U. (2014). **Hi-C analysis in Arabidopsis identifies the KNOT, a structure with similarities to the flamenco locus of Drosophila.** *Molecular Cell* 55, 678-693.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Haberer, G., Hindemitt, T., Meyers, B.C., and Mayer, K.F. (2004). **Transcriptional similarities, dissimilarities, and conservation of cis-elements in duplicated genes of Arabidopsis.** *Plant physiology* 136, 3009-3022.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Hecker, A., Brand, L.H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D. (2015). **The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs.** *Plant physiology* 168, 1013-1024.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). **Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.** *Molecular Cell* 38, 576-589.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Heo, J.B., and Sung, S. (2011). **Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA.** *Science* 331, 76-79.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). **H3K9 and H3K14 acetylation co-occur at many**

Molitor, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zytnicki, M. (Co-premier auteur) Andrey, P., Houba Hérin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A.,

Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). **The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of**

gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13, 424.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10, R25.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Latrasse, D., Germann, S., Houba-Herlin, N., Dubois, E., Bui-Prodhomme, D., Hourcade, D., Juul-Jensen, T., Le Roux, C., Majira, A., Simoncello, N., Granier, F., Tacconat, L., Renou, J.P., and Gaudin, V. (2011). Control of flowering and cell fate by LIF2, an RNA binding partner of the polycomb complex component LHP1. *PLoS One* 6, e16592.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Le Roux, C., Del Prete, S., Boutet-Mercey, S., Perreau, F., Balague, C., Roby, D., Fagard, M., and Gaudin, V. (2014). The hnRNP-Q protein LIF2 participates in the plant immune response. *PLoS One* 9, e99343.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I., and Deng, X.W. (2007). Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *The Plant Cell* 19, 731-749.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Luo, C., Sidote, D.J., Zhang, Y., Kerstetter, R.A., Michael, T.P., and Lam, E. (2012). Integrative analysis of chromatin states in *Arabidopsis* identified potential regulatory mechanisms for natural antisense transcript production. *The Plant journal : for cell and molecular biology*.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Mehdi, S., Derkacheva, M., Ramstrom, M., Kralemann, L., Bergquist, J., and Hennig, L. (2015). MSI1 functions in a HDAC complex to fine-tune ABA signaling. *The Plant Cell*.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Provart, N.J., Gil, P., Chen, W., Han, B., Chang, H.S., Wang, X., and Zhu, T. (2003). Gene expression phenotypes of *Arabidopsis* associated with sensitivity to low temperatures. *Plant physiology* 132, 893-906.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Pu, L., and Sung, Z.R. (2015). PcG and trxG in plants - friends or foes. *Trends Genetics* 31, 252-262.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841-842.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Regad, F., Lebas, M., and Lescure, B. (1994). Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. *Journal of Molecular Biology* 239, 163-169.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the *Arabidopsis* Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *The Plant Cell* 26, 2351-2366.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

monet, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zymoni, M. (Co-premier auteur), Andrey, P., Houba Hérlin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The *Arabidopsis* hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of

Simon, J.A., and Kingston, R.E. (2013). Occupying chromatin: polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular cell* 49, 808-824.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Song, Y.H., Yoo, C.M., Hong, A.P., Kim, S.H., Jeong, H.J., Shin, S.Y., Kim, H.J., Yun, D.J., Lim, C.O., Bahk, J.D., Lee, S.Y., Nagao, R.T., Key, J.L., and Hong, J.C. (2008). DNA-binding study identifies C-box and hybrid C/G-box or C/A-box motifs as high-affinity binding sites for STF1 and LONG HYPOCOTYL5 proteins. *Plant physiology* 146, 1862-1877.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature* 462, 799-802.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., Wutz, A., Vidal, M., Elderkin, S., and Brockdorff, N. (2012). RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* 148, 664-678.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Thomas-Chollier, M., Herrmann, C., Defrance, M., Sand, O., Thieffry, D., and van Helden, J. (2011). RSAT peak-motifs: motif analysis in full-size ChIP-seq datasets. *Nucleic Acids Research* 40, e31.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Thomas-Chollier, M., Darbo, E., Herrmann, C., Defrance, M., Thieffry, D., and van Helden, J. (2012). A complete workflow for the analysis of full-size ChIP-seq (and similar) data sets using peak-motifs. *Nature protocols* 7, 1551-1568.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genetics* 3, e86.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Vaquero-Sedas, M.I., Luo, C., and Vega-Palas, M.A. (2012). Analysis of the epigenetic status of telomeres by using ChIP-seq data. *Nucleic Acids Research* 40, e163.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Vrbsky, J., Akimcheva, S., Watson, J.M., Turner, T.L., Daxinger, L., Vyskot, B., Aufsatz, W., and Riha, K. (2010). siRNA-mediated methylation of Arabidopsis telomeres. *PLoS Genet* 6, e1000986.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Walter, M., Chaban, C., Schütze, K., Batistic, O., Weckermann, K., Näke, C., Blazevic, D., Grefen, C., Schumacher, K., Oecking, C., Harter, K., and Kudla, J. (2004). Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *Plant J* 40, 428-438.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Wan, F., Anderson, D.E., Barnitz, R.A., Snow, A., Bidere, N., Zheng, L., Hegde, V., Lam, L.T., Staudt, L.M., Levens, D., Deutsch, W.A., and Lenardo, M.J. (2007). Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. *Cell* 131, 927-939.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yi, X., Du, Z., and Su, Z. (2013). PlantGSEA: a gene set enrichment analysis toolkit for plant community. *Nucleic Acids Research* 41, W98-103.

Pubmed: [Author and Title](#)

CrossRef: [Author and Title](#)

Google Scholar: [Author Only](#) [Title Only](#) [Author and Title](#)

Yilmaz, A., Mejia-Guerra, M.K., Kurz, K., Liang, X., Welch, L., and Grotewold, E. (2011). AGRIS: the Arabidopsis Gene Regulatory Information Server, an update. *Nucleic Acids Research* 39, D1118-1122.

Pubmed: [Author and Title](#)

Montoli, A. M. (Co-premier auteur), Larasse, D. (Co-premier auteur), Zymnicki, M. (Co-premier auteur), Andrey, P., Houba Hérin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of

CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhang, H., He, H., Wang, X., Yang, X., Li, L., and Deng, X.W. (2011). Genome-wide mapping of the HY5-mediated gene networks in Arabidopsis that involve both transcriptional and post-transcriptional regulation. The Plant journal 65, 346-358.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhang, H., Zeitz, M.J., Wang, H., Niu, B., Ge, S., Li, W., Cui, J., Wang, G., Qian, G., Higgins, M.J., Fan, X., Hoffman, A.R., and Hu, J.F. (2014). Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. The Journal of cell biology 204, 61-75.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhang, X., Germann, S., Blus, B.J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S.E. (2007). The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nature Structural and Molecular Biology 14, 869-871.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoutte, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biology 9, R137.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhou, J., Wang, X., He, K., Charron, J.B., Elling, A.A., and Deng, X.W. (2010). Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. Plant Molecular Biology 72, 585-595.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zhou, Y., Hartwig, B., Velikkakam James, G., Schneeberger, K., and Turck, F. (2015). Complementary activities of TELOMERE REPEAT BINDING proteins and Polycomb Group complexes in transcriptional regulation of target genes. The Plant Cell.

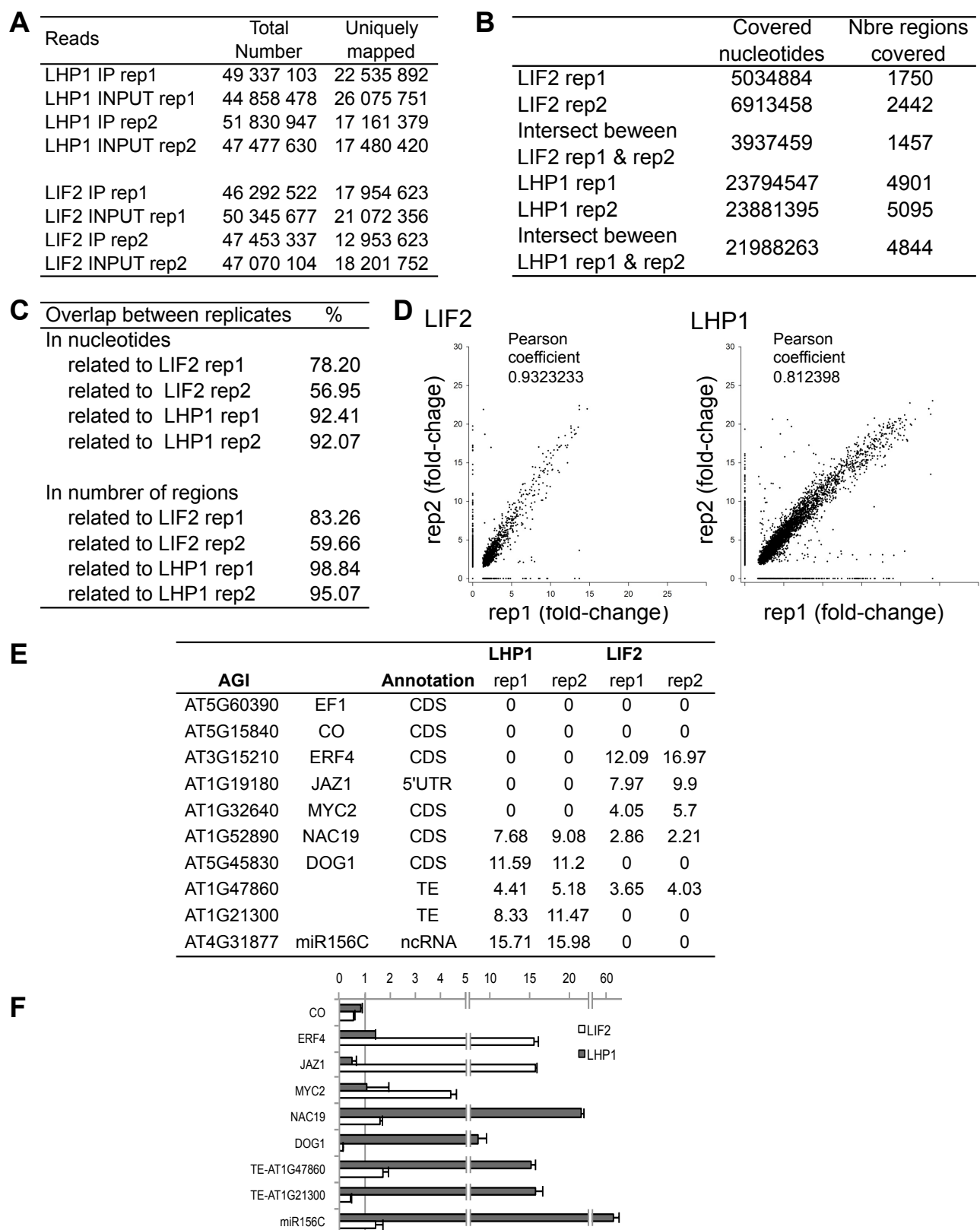
Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2AZ and DNA methylation are mutually antagonistic chromatin marks. Nature 456, 125-129.

Pubmed: [Author and Title](#)
CrossRef: [Author and Title](#)
Google Scholar: [Author Only Title Only Author and Title](#)

Comment citer ce document :

Molitor, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zytnicki, M. (Co-premier auteur), Andrey, P., Houba Héryn, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of

**Supplemental Figure 1: ChIP-seq experiments.**

(A) Read counts and mapping in the two ChIP-seq biological replicates.

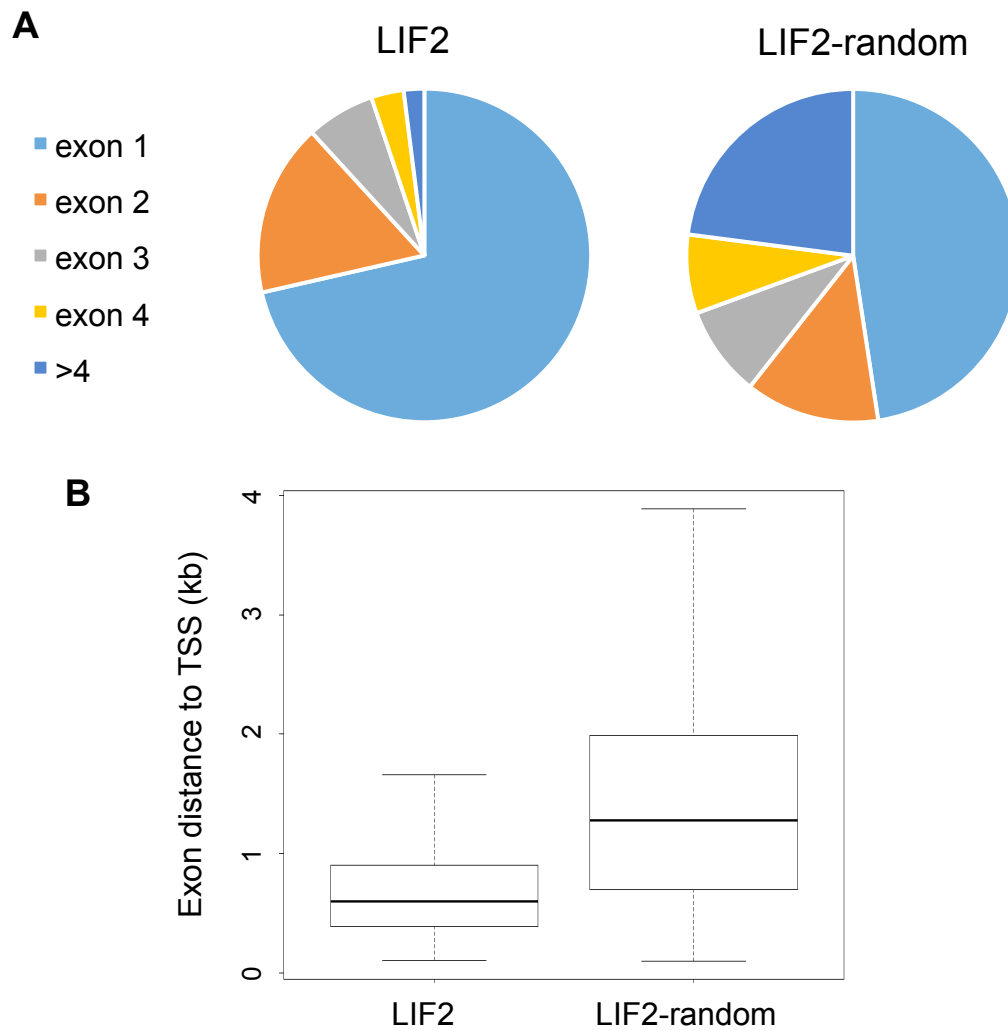
(B) Comparisons between replicates.

(C) Overlaps between replicates.

(D) MACS peak fold-change correlations between ChIP-seq replicates.

(E) Fold-changes in the two ChIP-seq biological replicates of selected target regions.

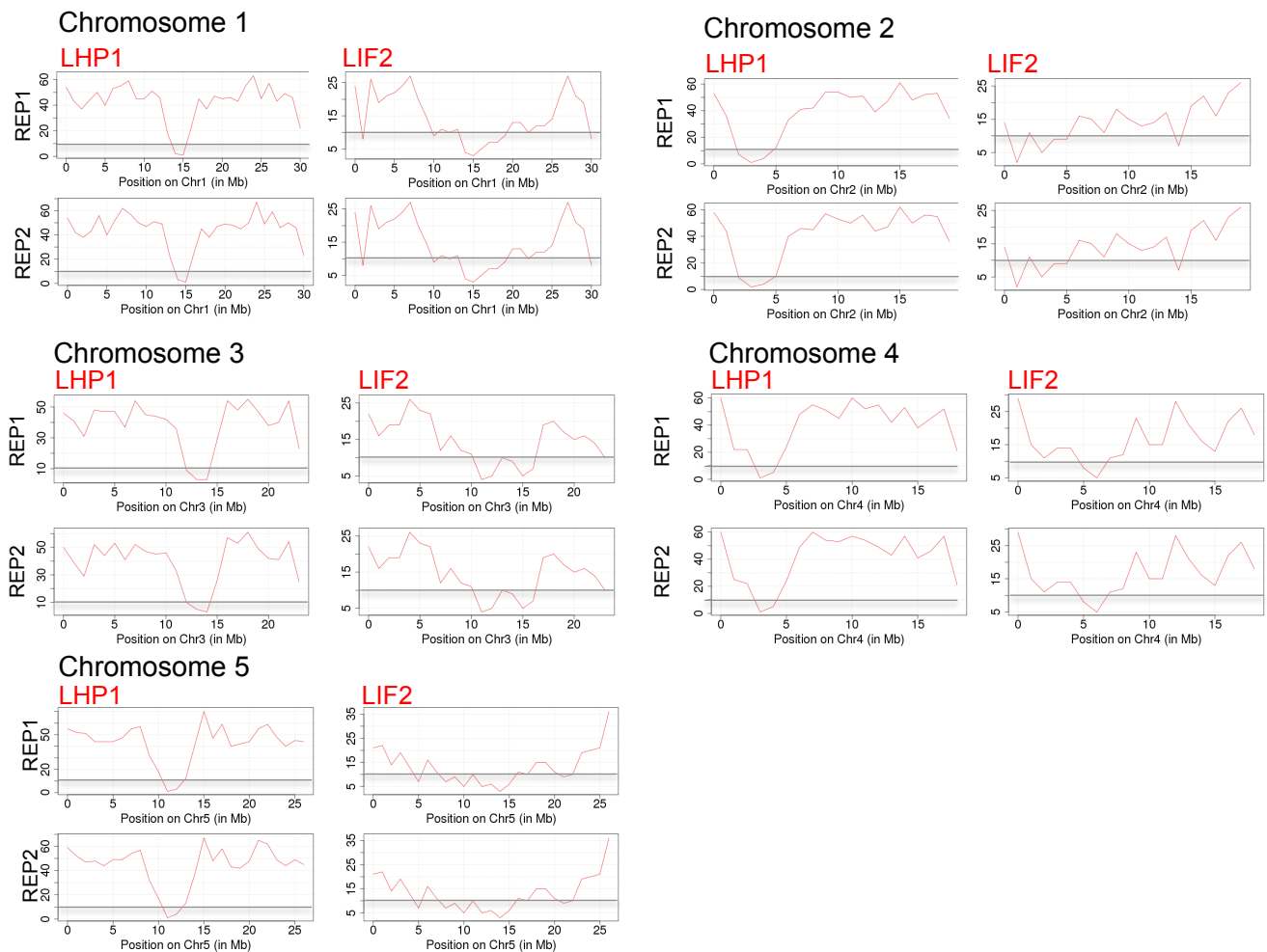
(F) Confirmation by ChIP-QPCR at targets identified by ChIP-seq. Protein enrichments were relative to input and the internal reference gene, *EF1 α* . The values correspond to the mean of two biological replicates and three technical replicates for each \pm SE.



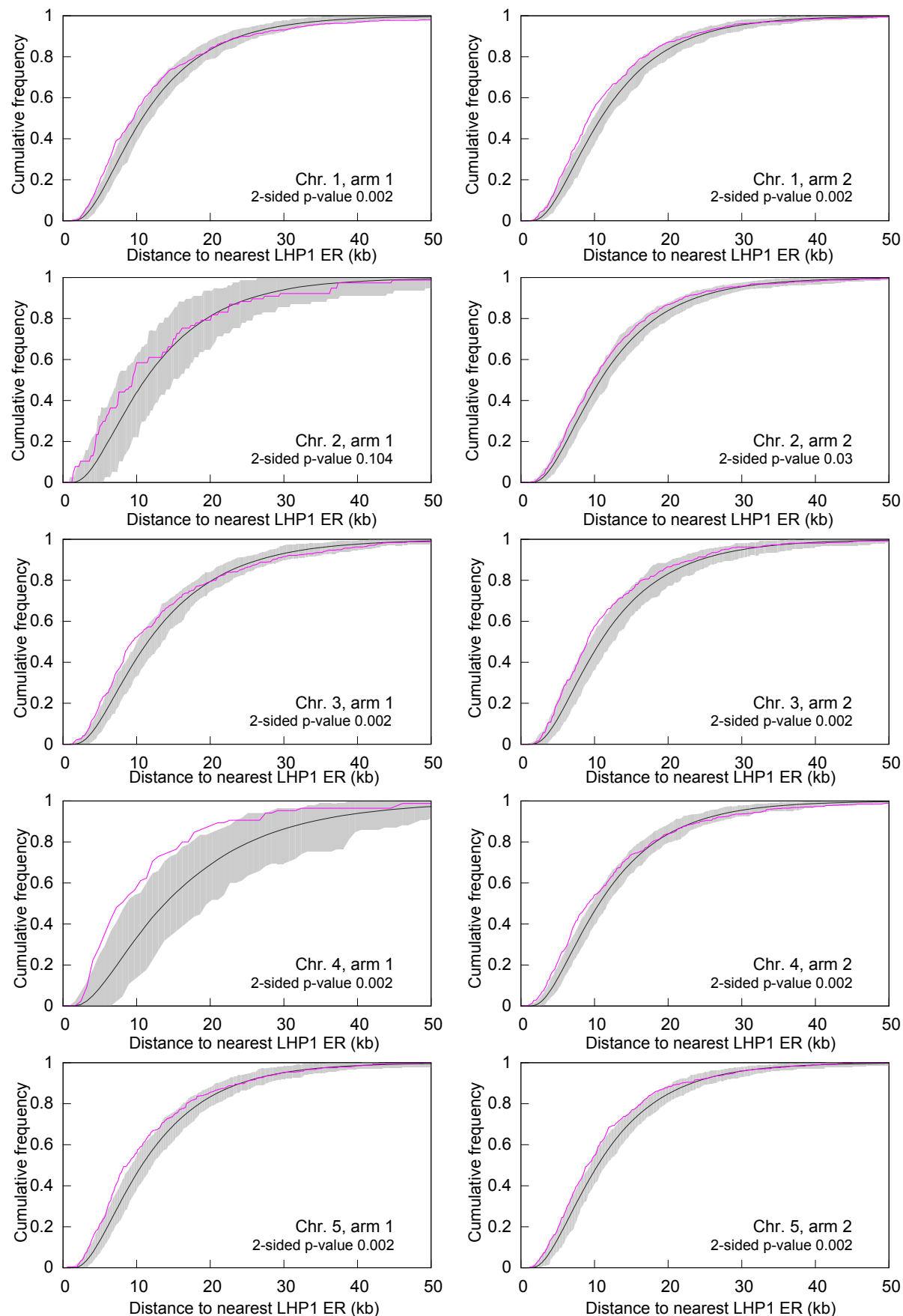
Supplemental Figure 2: Exon distributions of LIF2 ERs.

(A) Exon number distribution of LIF2 ERs and comparisons with the randomly shuffled control regions.

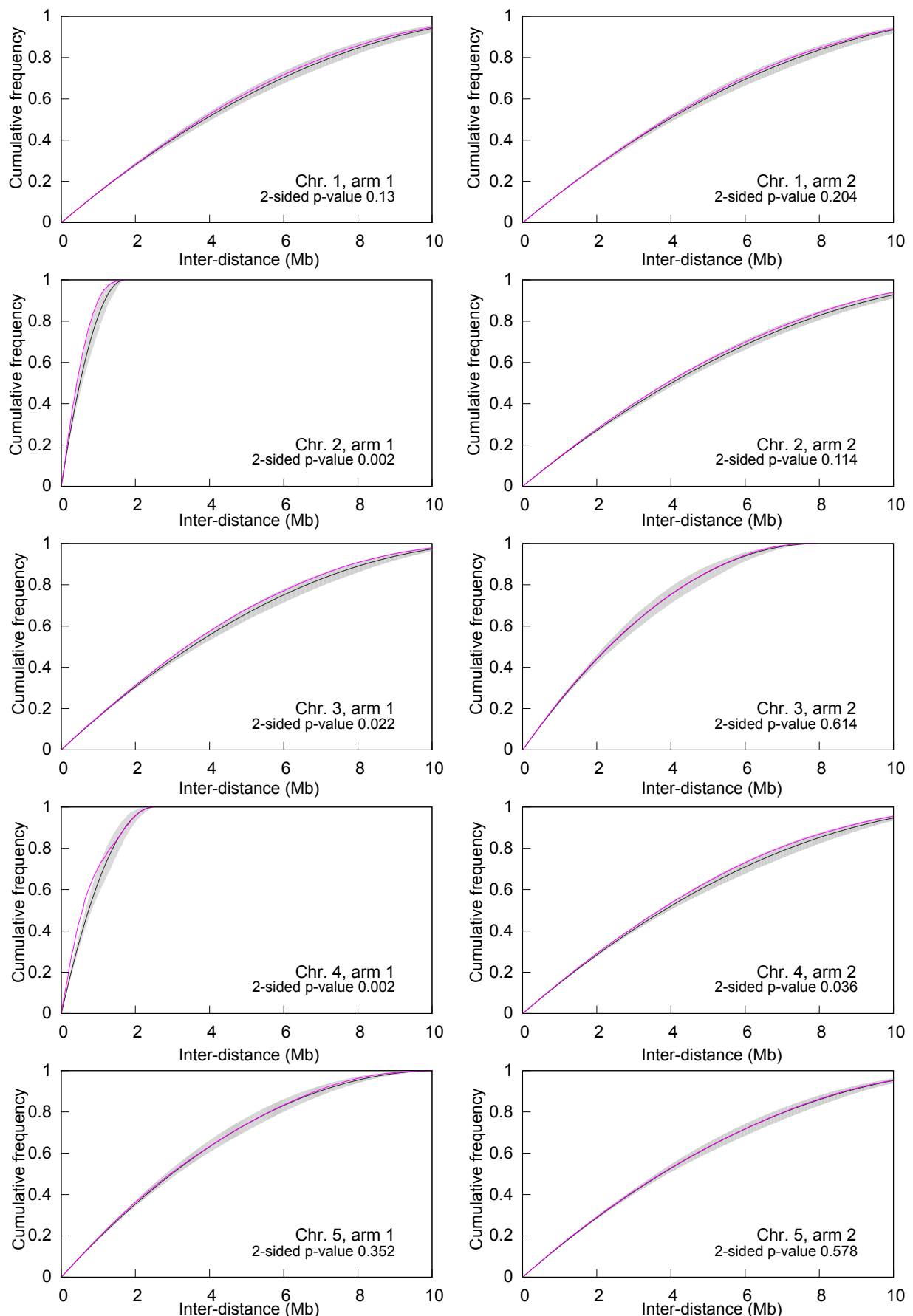
(B) Distance to TSS of the LIF2-bound exons and comparisons with the randomly shuffled control regions.



Supplemental Figure 3: Distributions of the sums of the summits in 1-Mb windows. For the two biological replicates, LHP1 and LIF2 distributions on the five chromosomes were compared by plotting the sums of their summits in 1-Mb windows.

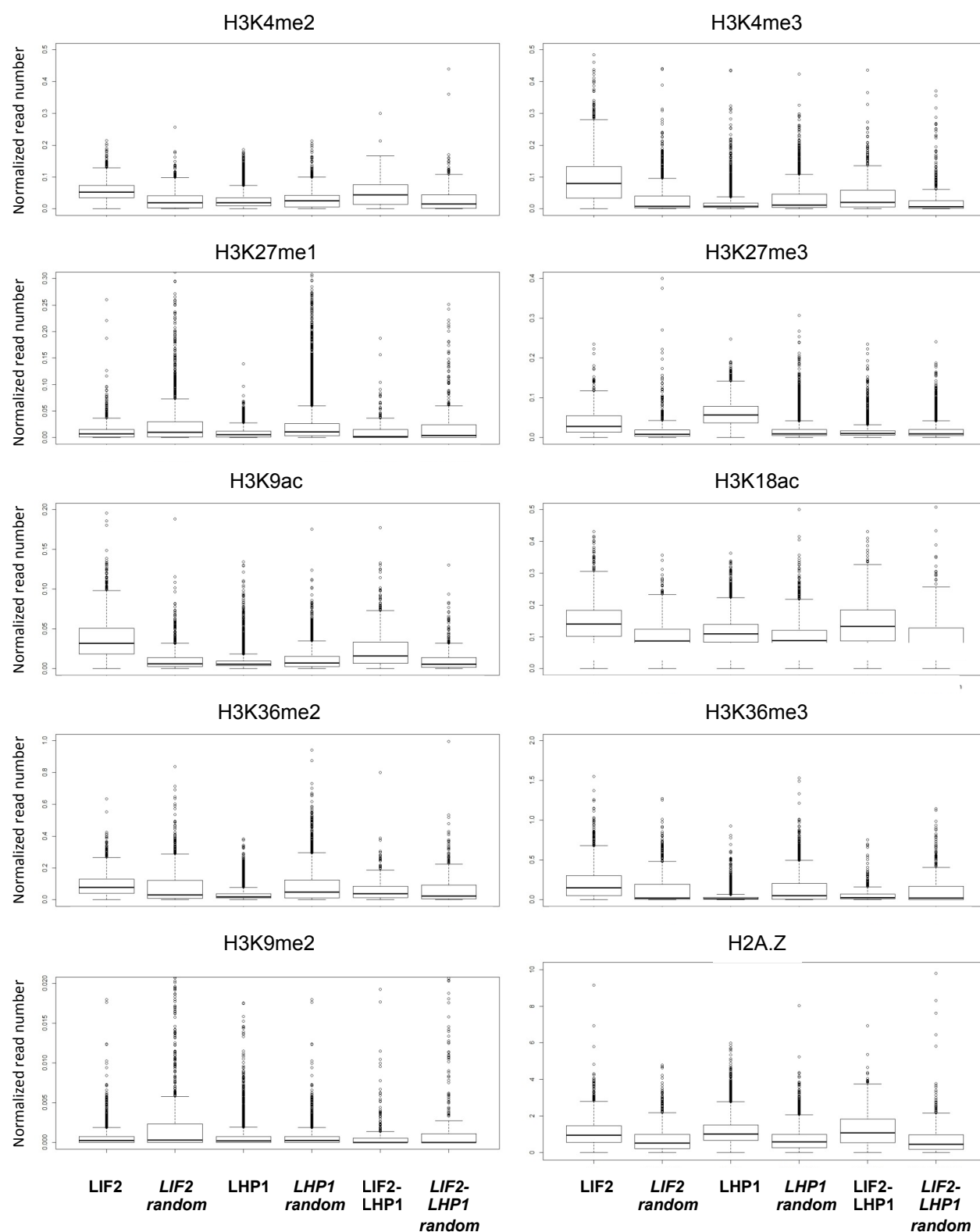


Supplemental Figure 4: Cumulative distribution of the distance to the nearest LHP1 summit. Pink: observed distribution; black: average distribution under the random model; grey: 95% envelope under the random model. Arm 1 of chromosome 2 and arm 1 of chromosome 4 bear the NOR and the knob, respectively, which introduce spatial constraints.

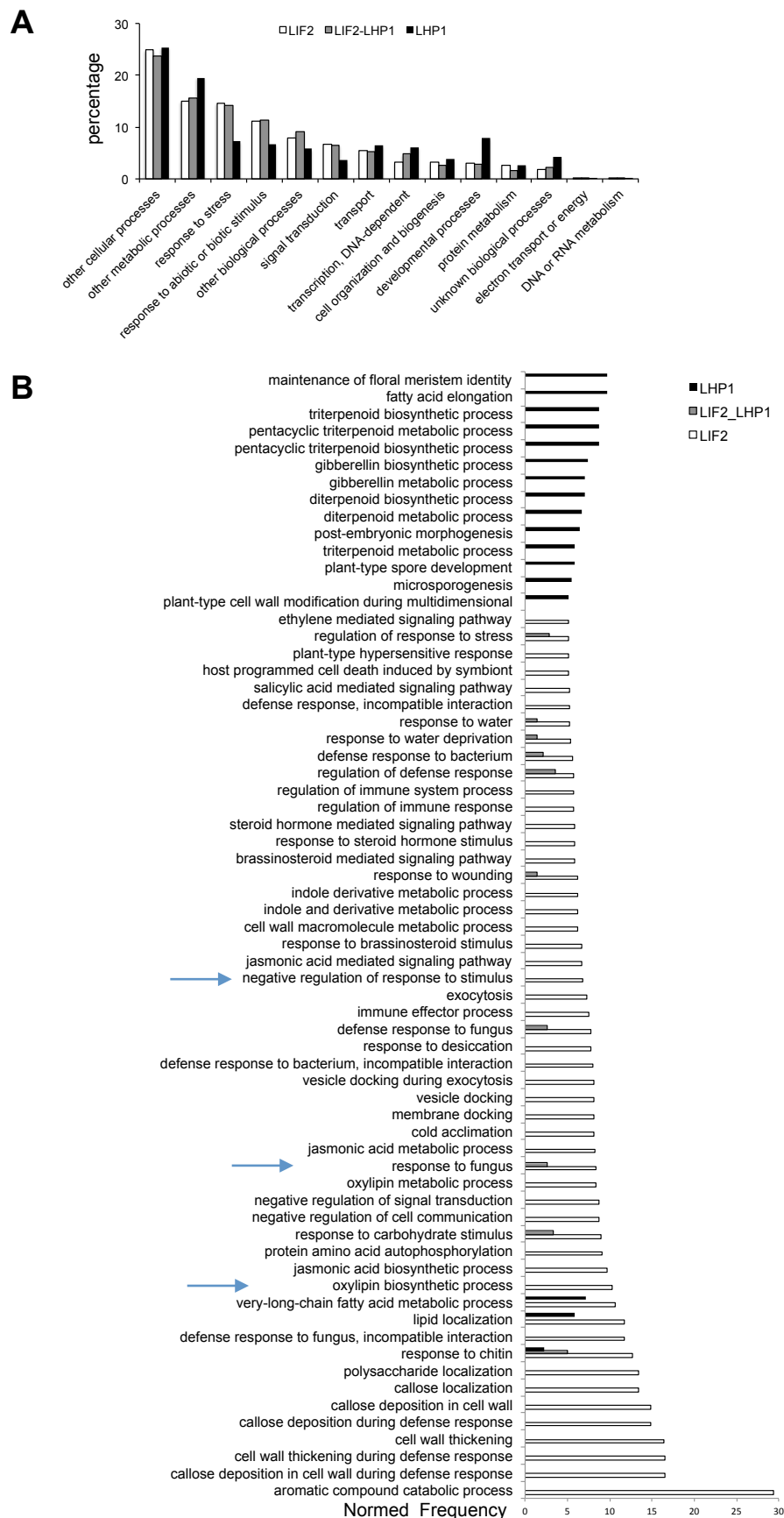


Supplemental Figure 5: Cumulative distribution of LHP1 summit inter-distances.

Pink: observed distribution; black: average distribution under the random model; grey: 95% envelope under the random model. Arm 1 of chromosome 2 and arm 1 of chromosome 4 bear the NOR and the knob, respectively, which introduce spatial constraints.



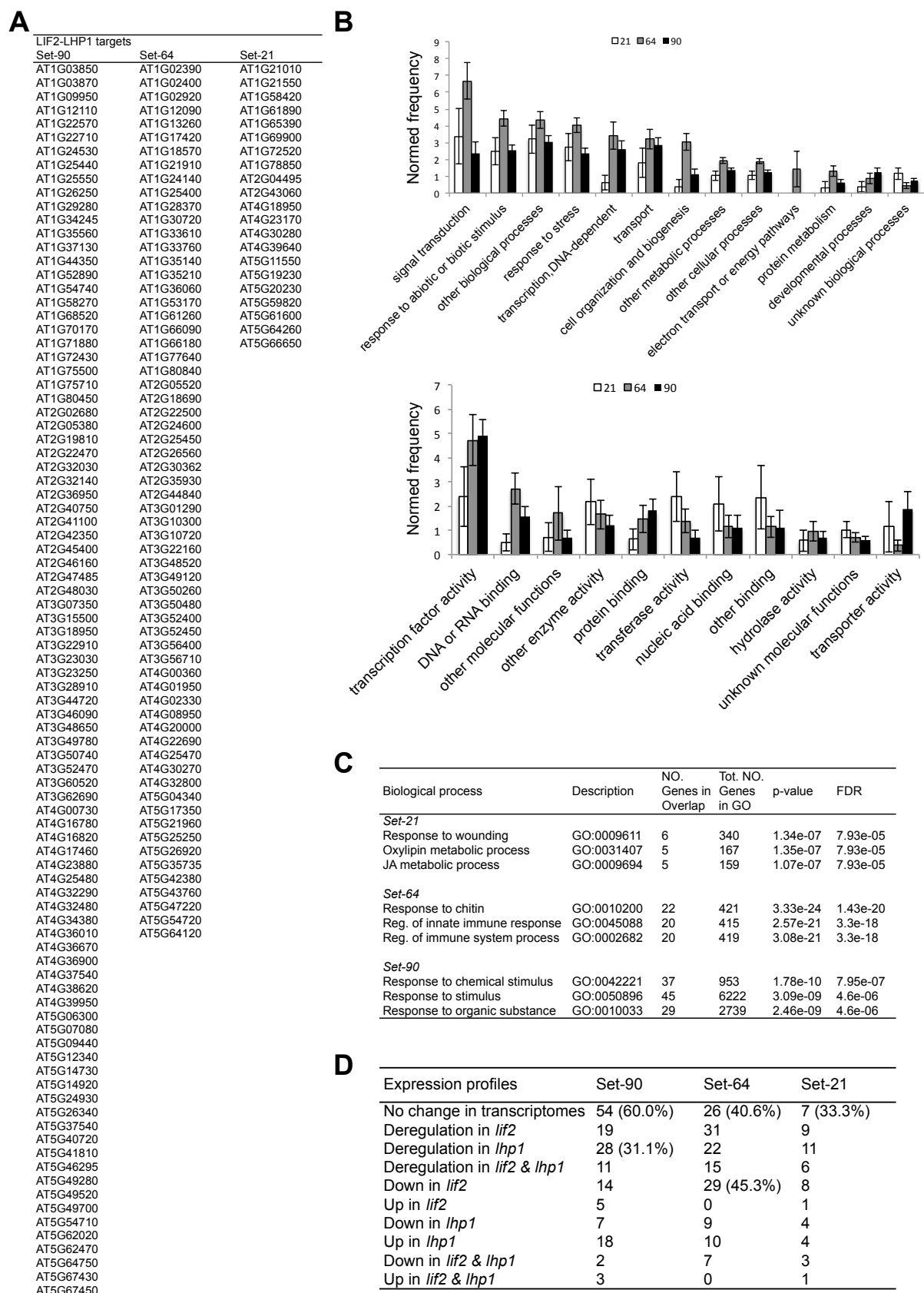
Supplemental Figure 6: Post-translational histone modifications and their distributions in LIF2 ERs, LHP1 ERs and LIF2-LHP1 IRs.



Supplemental Figure 7: GO term analysis of the target loci of LIF2 and LHP1.

(A) Functional categorization of the LIF2 ERs, LHP1 ERs and LIF2-LHP1 IRs (TAIR GO toolkit).

(B) Normed frequencies (NF) of the GO terms in the biological process categories with an over 4-fold enriched NF in at least one of the input lists (*i.e.*, LIF2 target genes, LHP1 target genes, genes of LIF2-LHP1 IRs) were compiled using AgriGO toolkit. Arrows: GO terms related to JA.

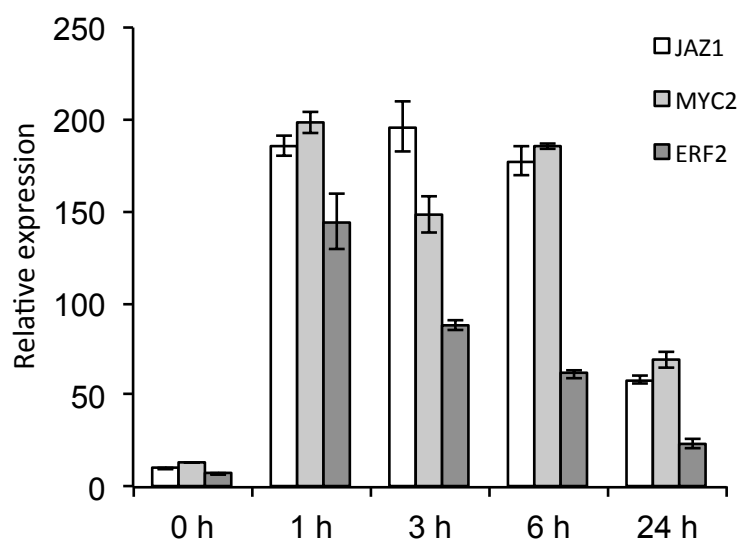


Supplemental Figure 8: Analyses of the LIF2-LHP1 IRs with binding alterations in the mutant backgrounds.

(A) Lists of the three sets of genes.

(B-C) GO analyses using BAR and Plant GSEA toolkits, respectively.

(D) Expression profiles of the gene sets. Percentages related to the number of genes in the Set were calculated for the largest classes.



Supplemental Figure 9: Expression kinetics of JA-induced marker genes in response to MeJA treatment in wild-type plants.

Two-week-old seedlings were treated with JA for 1 to 24 hours and *JAZ1*, *MYC2* and *ERF2* expression was recorded. *EF1* was used as reference gene.

Tandem duplications (T-Clusters)	Chromosome 4	Whole genome
T-Clusters	248	1564
T-Cluster genes	681	4176
Average T-Cluster size (gene number)	2.75	2.67
T-Clusters & LHP1 targets genes	132	861
% of T-Cluster genes targeted by LHP1	19.4	20.6
% of LHP1 target genes in T-Clusters	23.8	23.1
LHP1 target genes in T-Clusters (with no unannotated genes in the cluster)	n.d.	848
Clusters with 1 LHP1 target gene per cluster	n.d.	484
Clusters with 2 LHP1 target gene per cluster	n.d.	138
Clusters with 3 LHP1 target gene per cluster	n.d.	20
Clusters with 4 LHP1 target gene per cluster	n.d.	7
% of LHP1 targeted genes in T-clusters that have multiple LHP1 targets	n.d.	9.6%

Supplemental Table 1 : Tandem duplications and LHP1 target genes.
n.d. non determined.

	LIF2		LIF2-LHP1	
	Normed Frequency	p-value	Normed Frequency	p-value
<i>Biological process</i>				
signal transduction	3.34	1.753e-50	2.82	2.213e-11
other biological processes	3.22	1.652e-87	3.27	8.384e-29
response to abiotic or biotic stimulus	3	1.399e-78	2.7	2.662e-19
response to stress	2.98	2.398e-87	2.67	5.353e-21
transport	2.32	2.349e-36	2.22	5.928e-11
transcription,DNA-dependent	1.71	8.047e-10	2.15	1.356e-07
other cellular processes	1.35	1.633e-24	1.24	4.149e-05
other metabolic processes	1.34	3.208e-20	1.32	6.090e-07
cell organization and biogenesis	1.2	5.091e-03	0.97	0.079
electron transport or energy pathways	1.19	0.060	0.33	0.043
developmental processes	1.13	0.013	1.04	0.071
protein metabolism	1.05	0.026	0.63	1.059e-03
unknown biological processes	0.58	1.403e-18	0.68	1.004e-04
DNA or RNA metabolism	0.27	6.744e-06	0.12	2.346e-03
<i>Molecular function</i>				
transcription factor activity	2.39	3.730e-18	3.7	2.739e-17
DNA or RNA binding	1.16	0.011	1.39	6.066e-03
receptor binding or activity	1.43	0.100	1.31	0.255
protein binding	1.9	1.958e-13	1.29	0.031
other enzyme activity	1.2	0.033	1.25	0.082
other molecular functions	1.2	3.401e-03	1.25	0.022
transporter activity	1.12	0.050	1.16	0.097
other binding	1.17	6.249e-04	1.12	0.032
transferase activity	1.52	4.876e-07	1.05	0.078
nucleic acid binding	0.53	1.859e-04	0.72	0.068
unknown molecular functions	0.66	3.823e-12	0.71	3.973e-04
hydrolase activity	1.01	0.045	0.7	0.022
kinase activity	1.83	5.319e-07	0.68	0.070
nucleotide binding	1.15	0.011	0.64	9.169e-03
structural molecule activity	0.52	0.016	0.63	0.154

Supplemental Table 2: GO term analysis of the genes present in LIF2 ERs and LIF2-LHP1 IRs using the Plant Functional Genomics (BAR) classification Supervisor program.

TF family	Nbr Genes	Total in family	p-value	FDR
LIF2				
AP2-EREBP	24	138	1.57e-11	1.79e-08
WRKY	12	72	2.82e-06	1.07e-03
LIF2_LHP1				
AP2-EREBP	17	138	2.72e-14	1.11e-11
C2C2-CO-Like	4	30	1.97e-04	0.0202
C2H2	8	211	6.24e-04	0.0488
WRKY	6	72	5.63e-05	7.68e-03
LHP1				
C2H2	34	211	1.00E-04	0.0146
BASIC HELIX-LOOP-HELIX (bHLH)	33	162	1.94e-06	5.27e-04
BHLH	33	161	1.73e-06	5.27e-04
MADS	31	109	7.36e-09	4.66e-06
MADS-BOX	31	108	6.13e-09	4.66e-06
HOMEBOX	29	91	2.66e-09	4.66e-06
MYB	27	131	1.34e-05	2.82e-03
MYB3R- and R2R3- TYPE MYB-encoding genes	27	132	1.5e-05	2.86e-03
NAC	19	96	3.76e-04	0.0434
WRKY	18	72	4.29e-05	6.79e-03
ZINC FINGER-HOMEBOX - ZHD subfamily	7	14	3.88e-04	0.0434

Supplemental Table 3: Enrichments of LIF2 and LHP1 targets in specific transcription factor families using the PlantGSEA resource.

TF	Genes targeted	Total genomic targets	p-value	FDR	Genes targeted by TF in ERs
LIF2					
AGL15 (MADS box)	7	22	9.89e-06	1.69e-04	AT1G02400 AT1G14920 AT4G25470 AT2G45830 AT4G38680 AT1G68840 AT1G13260
AP2 (AP2/EREBP)	18	165	2.87e-06	5.73e-05	AT1G11050 AT1G13260 AT1G22530 AT1G23390 AT1G71030 AT3G55980 AT4G00730 AT4G01250 AT4G02540 AT4G08950 AT4G16490 AT4G20260 AT4G25480 AT4G26690 AT4G29190 AT5G19140 AT5G20250 AT5G49360
PIF1 (AtbHLH15)	22	189	8.38e-08	2.51e-06	AT1G25550 AT1G28330 AT1G52890 AT1G56220 AT1G60190 AT1G68670 AT2G01570 AT2G16660 AT2G27500 AT2G45820 AT2G46710 AT3G02550 AT3G04730 AT3G12920 AT3G24050 AT3G24503 AT3G25870 AT4G17460 AT5G24930 AT5G54380 AT5G64260 AT5G65430
HY5 (bZIP)	61	221	6.37e-37	7.64e-35	AT1G07135 AT1G09070 AT1G17420 AT1G18300 AT1G18570 AT1G20510 AT1G21910 AT1G24140 AT1G24530 AT1G53170 AT1G56660 AT1G60190 AT1G61340 AT1G61890 AT1G66160 AT1G68840 AT1G69760 AT1G73500 AT1G77640 AT1G78850 AT2G22500 AT2G25460 AT2G27660 AT2G30040 AT2G33580 AT2G35290 AT2G35930 AT2G39650 AT2G41100 AT2G41640 AT3G06080 AT3G14440 AT3G25250 AT3G46080 AT3G48520 AT3G52400 AT3G54810 AT3G55980 AT3G56880 AT3G62260 AT3G62720 AT4G08950 AT4G17500 AT4G25490 AT4G27310 AT4G29780 AT4G32800 AT4G33920 AT4G34150 AT4G36670 AT5G01100 AT5G43890 AT5G45110 AT5G45340 AT5G49280 AT5G49520 AT5G52020 AT5G59550 AT5G62020 AT5G62520 AT5G66650
LIF2_LHP1					
AGL15 (MADS box)	4	22	6.71e-05	7.64e-04	AT1G02400 AT1G13260 AT2G45830 AT4G25470
AP2 (AP2/EREBP)	6	165	3.61e-03	0.0274	AT4G00730 AT4G37540 AT4G25480 AT4G08950 AT5G54470 AT1G13260
PIF1 (AtbHLH15)	6	189	6.79e-03	0.0441	AT1G03850 AT1G25550 AT1G52890 AT4G17460 AT5G24930 AT5G64260
HY5 (bZIP)	30	221	1.75e-25	7.97e-24	AT1G17420 AT1G18570 AT1G21910 AT1G24140 AT1G24530 AT1G25400 AT1G53170 AT1G61890 AT1G77640 AT1G78850 AT2G22500 AT2G27660 AT2G35290 AT2G35930 AT2G37430 AT2G41100 AT3G25250 AT3G46080 AT3G48520 AT3G52400 AT4G08950 AT4G32800 AT4G36670 AT5G43890 AT5G49280 AT5G49520 AT5G52020 AT5G59820 AT5G62020 AT5G66650
LHP1					
AP2 (AP2/EREBP)	30	165	3.73e-05	1.68e-03	AT2G16760 AT2G14210 AT5G54470 AT1G71050 AT4G00730 AT4G04630 AT5G67180 AT5G07030 AT3G58780 AT4G36870 AT2G40435 AT5G15310 AT4G13210 AT2G18550 AT2G43620 AT2G45660 AT1G35910 AT1G70560 AT4G18960 AT5G64870 AT4G37540 AT5G13790 AT1G73590 AT5G67060 AT1G13260 AT1G35730 AT3G55710 AT2G42830 AT4G08950 AT4G24050
HY5 (bZIP)	42	221	4.27e-07	5.86e-05	AT1G24530 AT5G52020 AT5G25810 AT1G78990 AT1G17420 AT3G52400 AT2G47460 AT1G61890 AT5G23010 AT4G27250 AT1G02810 AT2G22500 AT5G49520 AT5G57510 AT1G16850 AT4G36670 AT3G55120 AT2G37430 AT1G53170 AT5G59820 AT5G66650 AT1G32450 AT5G44120 AT5G42800 AT2G28630 AT3G48520 AT5G62490 AT5G24140 AT5G59780 AT2G15020 AT3G21720 AT5G49280 AT1G17380 AT5G10100 AT4G08950 AT4G05100 AT2G41100 AT3G22830 AT5G62020 AT1G78850 AT1G12950 AT5G13930 AT3G54340 AT2G22540 AT3G58780 AT3G02310 AT2G45660 AT4G18960 AT5G13790 AT2G03710 AT4G24540 AT2G42830 AT5G15800

Supplemental Table 4: LIF2 and LHP1 targets are also bound by specific transcription factors.

The PlantGSEA and AGRIS toolkits were used.

Segment Name	Unique sequence occurrences	Total occurrence	Expected unique sequence	Expected occurrences	Rank	Score
<i>ACGTGGCA word</i>						
Distal Promoter	598	621	497.582	503.982	1175	109.93
Proximal Promoter	414	431	353.216	358.329	1251	65.7371
Core Promoter	172	174	141.099	152.113	539	34.061
5'UTR	26	26	15.5612	16.4062	3156	13.3462
Intron	48	48	38.8425	40.5848	21162	10.161
3'UTR	26	27	21.3065	22	14830	5.1762
Genome-wide	5	1265	5	1020.69	2410	0
<i>AAACCCCTA word</i>						
Intron	764	802	652.752	686.061	672	120.231
5'UTR	911	952	825.505	908.536	26	89.7771
Distal Promoter	3116	3674	3041.64	3238.08	2631	75.2602
Core Promoter	639	658	568.781	618.048	51	74.3858
Proximal Promoter	1606	1779	1540.73	1598.78	1209	66.6328
3'UTR	91	92	75.7094	78.332	2976	16.7401
Genome-wide	5	11340	5	9887.71	55	0

Supplemental Table 5: Occurrences of the two identified DNA words.

The word frequency calculation was performed in non-coding segments of the *A. thaliana* genome, using the Arabidopsis *cis*-regulatory element database (<http://arabidopsis.med.ohio-state.edu/AtcisDB/>).

GO term	Description	Normed frequency	p-value	FDR
<i>lhp1</i> LIF2 depleted				
<i>Biological process</i>				
GO:0052542	callose deposition during defense response	31,5	1.4e-06	2.4e-05
GO:0033037	polysaccharide localization	28,5	2.2e-06	3.3e-05
GO:0052545	callose localization	28,5	2.2e-06	3.3e-05
GO:0010200	response to chitin	24,5	2.1e-31	5.0e-29
GO:0009631	cold acclimation	20,6	8.7e-06	0.00013
GO:0002252	immune effector process	19,3	1.2e-05	0.00016
GO:0031408	oxylipin biosynthetic process	18,7	1.3e-05	0.00018
GO:0009867	jasmonic acid mediated signaling pathway	17,1	4.1e-07	7.6e-06
GO:0009743	response to carbohydrate stimulus	16,9	1.5e-29	2.4e-27
GO:0031407	oxylipin metabolic process	15,3	3.2e-05	0.00041
GO:0002682	regulation of immune system process	14,6	4.0e-05	0.00049
GO:0050776	regulation of immune response	14,6	4.0e-05	0.00049
GO:0042434	indole derivative metabolic process	13,5	9.7e-06	0.00014
GO:0042430	indole and derivative metabolic process	13,5	9.7e-06	0.00014
GO:0050832	defense response to fungus	13,3	4.1e-10	1.4e-08
GO:0009873	ethylene mediated signaling pathway	12,9	4.3e-07	7.7e-06
GO:0031347	regulation of defense response	12,7	2.6e-06	3.9e-05
GO:0080134	regulation of response to stress	11,5	9.6e-07	1.7e-05
GO:0009620	response to fungus	11,3	2.0e-11	1.1e-09
GO:0000160	two-component signal transduction system (phosphorelay)	10,4	2.0e-06	3.2e-05
GO:0009611	response to wounding	10,3	3.8e-12	2.3e-10
GO:0042742	defense response to bacterium	10,1	9.0e-11	4.3e-09
GO:0009415	response to water	10,0	8.4e-14	7.3e-12
GO:0009414	response to water deprivation	9,9	3.8e-13	3.0e-11
GO:0019760	glucosinolate metabolic process	9,6	0.00024	0.0026
GO:0016143	S-glycoside metabolic process	9,6	0.00024	0.0026
GO:0019757	glycosinolate metabolic process	9,6	0.00024	0.0026
GO:0048585	negative regulation of response to stimulus	8,7	9.5e-05	0.0011
GO:0006955	immune response	8,5	5.5e-16	6.2e-14
GO:0002376	immune system process	8,4	5.9e-16	6.2e-14
GO:0045087	innate immune response	8,3	1.2e-14	1.2e-12
GO:0009642	response to light intensity	8,0	0.00015	0.0018
GO:0009617	response to bacterium	7,7	8.6e-10	2.7e-08
GO:0009723	response to ethylene stimulus	7,2	2.4e-07	4.9e-06
GO:0006952	defense response	7,2	9.6e-25	1.3e-22
<i>Molecular function</i>				
GO:0016564	transcription repressor activity	17,1	9.0e-09	8.9e-07
<i>lif2</i> LHP1 depleted				
<i>Biological process</i>				
GO:0010876	lipid localization	16,7	5.6e-13	3.6e-11
GO:0010076	maintenance of floral meristem identity	16,7	1.2e-05	0.00037
GO:0046087	cytidine metabolic process	16,7	1.2e-05	0.00037
GO:0006216	cytidine catabolic process	16,7	1.2e-05	0.00037
GO:0009972	cytidine deamination	16,7	1.2e-05	0.00037
GO:0046135	pyrimidine nucleoside catabolic process	15,0	1.8e-05	0.00053
GO:0046133	pyrimidine ribonucleoside catabolic process	15,0	1.8e-05	0.00053
GO:0010022	meristem determinacy	13,9	0.00013	0.0027
GO:0009164	nucleoside catabolic process	13,6	2.8e-05	0.00075
GO:0042454	ribonucleoside catabolic process	13,6	2.8e-05	0.00075
GO:0034656	nucleobase, nucleoside and nucleotide catabolic process	12,5	4.0e-05	0.001
GO:0034655	nucleobase, nucleoside, nucleotide and nucleic acid catabolic process	12,5	4.0e-05	0.001
GO:0045962	positive regulation of development, heterochronic	12,5	0.00018	0.0038
GO:0046131	pyrimidine ribonucleoside metabolic process	10,7	7.8e-05	0.0019
GO:0006213	pyrimidine nucleoside metabolic process	10,0	0.00011	0.0023
GO:0009886	post-embryonic morphogenesis	9,3	2.1e-08	9.7e-07
GO:0045596	negative regulation of cell differentiation	8,1	8.1e-07	3.2e-05
GO:0065001	specification of axis polarity	7,9	9.1e-05	0.0021
GO:0016145	S-glycoside catabolic process	7,9	0.0003	0.006
GO:0019759	glycosinolate catabolic process	7,9	0.0003	0.006
GO:0019762	glucosinolate catabolic process	7,9	0.0003	0.006
GO:0010089	xylem development	7,8	0.001	0.017
GO:0016139	glycoside catabolic process	7,5	0.00038	0.0073
GO:0010074	maintenance of meristem identity	7,4	4.4e-05	0.0011
GO:0048440	carpel development	7,4	3.6e-11	2.0e-09
GO:0009943	adaxial/abaxial axis specification	7,3	0.00014	0.0031
GO:0019827	stem cell maintenance	7,3	1.7e-05	0.0005
GO:0048864	stem cell development	7,3	1.7e-05	0.0005
GO:0009944	polarity specification of adaxial/abaxial axis	7,1	0.00048	0.0088
GO:0006722	triterpenoid metabolic process	7,1	0.00048	0.0088
GO:0048467	gynoecium development	7,0	2.7e-11	1.5e-09
GO:0048863	stem cell differentiation	7,0	2.1e-05	0.00059

Supplemental Table 6: GO term analysis of LIF2 or LHP1 depleted regions in the mutant backgrounds (AgriGO). Lists of GO terms with NF \geq 7.

Primer	Sequence (5' to 3')	Application
3HA-1	ACACACACTGCAGGGGTTAATTAACATCTTTTACCC	Cloning
3HA-2	CGGAATCTAGAGTCGACGCTGCACTGAGCAGCGTAA	Cloning
3HA-2	CCGGATATCGTCGACGGGTTAATTAACATCTTTTACCC	Cloning
3HA-2	CCGGATATCTTACTCGAGGCACTGAGCAGCGTAATCTGG	Cloning
Nost-1	CCTAAGGTACCGAATTTCCCGATCGTTCA	Cloning
Nost-2	CGGGAATTCGGATCTAGTAACATA	Cloning
AD379-28	TACGCAGCTGTTGTAATCCAA	Cloning
AD379-29	CATTATCCTGCAGGTCTGACATCTGGTCATC	Cloning
AD379-30	TTCCCCTCGAGATGTCAGACGCAAGAGATAA	Cloning
AD379-32	TTGAAAGTTGAAACAAAATCAATCA	Cloning
JAZ1_F	AGCTTCACTTCACCGTTCTTGGA	qRT-PCR
JAZ1_R	TCTTGTCTTGAAGCAACGTCGTCA	qRT-PCR
JAZ9_F	TGCTGTCGAAGAACGAGGGT	qRT-PCR
JAZ9_R	CTTCCCCATTCTCTAGCTGC	qRT-PCR
LOX3_F	CGGATAGAGAAAGAGATTGAGAAAAGGAAC	qRT-PCR
LOX3_R	GGTACACCTCTACACGTAACACCAGGC	qRT-PCR
MYC2_F	GCCGAAGGAATACACGCAAT	qRT-PCR
MYC2_R	CGGGTTGTGAACGGGCTA	qRT-PCR
VSP2_F	TAGGCTTCAATATGAGATGCTTCCAGT	qRT-PCR
VSP2_R	ACCGTTGGAATTGTGGAAGAATG	qRT-PCR
JAZ6_A_F	GTCTGCTCAGCCGGTACTTG	qChIP/qRT-PCR
JAZ6_A_R	TTCGAGCCAACCCCATATTA	qChIP/qRT-PCR
CO_F	AAC AAT GAC CGA TCC AGA GAA	qChIP/qRT-PCR
CO_R	CCT CCT TGG CAT CCT TAT CA	qChIP/qRT-PCR
EF1_F	CCA AGG GTG AAA GCA AGA AGA	qChIP/qRT-PCR
EF1_R	CTG GAG GTT TTG AGG CTG GTA T	qChIP/qRT-PCR
MYC2_A_F	CTCTTCCGATATCTCAACTTTATGG	qChIP
MYC2_A_R	GGCGTCGGAGTTGTTTCA	qChIP
miR156C_F	TTGCGTGCTCACTGCTCTAT	qChIP
miR156C_R	AGAGAAAGTGAGAGATGGGAACA	qChIP
AT1G21300-F	TGTACCAACAACGCTCCACT	qChIP
AT1G21300_R	TTTCCAGATAGCGAAGTTGTCTT	qChIP
AT1G47860_F	CCGCGTTTGCACCATTAT	qChIP
AT1G47860_R	CCATTGCCTACACGTACCG	qChIP
DOG1_F	TCTCGAGTGGATGAGTTTGC	qChIP
DOG1_R	TCTTCATCACCGTGAGAT CG	qChIP
ERF4_F	GTTTTCTTGCCCGGATCTC	qChIP
ERF4_R	CGTTAGGAAGCGTCCTTGG	qChIP
NAC19_F	TCTTCATCGGTGCGGTAAAATCGG	qChIP
NAC19_R	TCCAAGAACTGACCCGTTAACGC	qChIP
JAZ1_B_F	GCAGAGAAGCAACAGCAACA	qChIP
JAZ1_B_R	TCTCGAATAGCTAAATCGATACAAAG	qChIP
JAZ9_A_F	GTGAGATAAATGGAACATATTAACC	qChIP
JAZ9_A_R	GCAATAGGACGAACACAGTTATCA	qChIP
JAZ9_B_F	TCTTCCTCTCTTTAAATTGGATGTT	qChIP
JAZ9_B_R	CAAACCTCAAATTAACGTGTTTCTC	qChIP
LOX3_A_F	CATCACAGAAAGGTCATCACTTG	qChIP
LOX3_A_R	TTGATCGAGAACTGTGTTGACTG	qChIP
LOX3_C_F	TTGGTACTCAGAATCAATCAACG	qChIP
LOX3_C_R	GGTCGTCGACGGTTGATAA	qChIP

Supplemental Table 7: List of primers.

The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of stress-responsive genes

Anne Molitor, David Latrasse, Matthias Zytnicki, Philippe Andrey, Nicole Houba-Hérin, Mélanie Hachet, Christophe Battail, Stefania Del Prete, Adriana Alberti, Hadi Quesneville and Valerie Gaudin
Plant Cell; originally published online August 5, 2016;
DOI 10.1105/tpc.16.00244

This information is current as of September 9, 2016

Supplemental Data	http://www.plantcell.org/content/suppl/2016/08/05/tpc.16.00244.DC1.html
Permissions	https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&issn=1532298X&WT.mc_id=pd_hw1532298X
eTOCs	Sign up for eTOCs at: http://www.plantcell.org/cgi/alerts/ctmain
CiteTrack Alerts	Sign up for CiteTrack Alerts at: http://www.plantcell.org/cgi/alerts/ctmain
Subscription Information	Subscription Information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: http://www.aspb.org/publications/subscriptions.cfm

© American Society of Plant Biologists

ADVANCING THE SCIENCE OF PLANT BIOLOGY

Molitor, A. M. (Co-premier auteur), Latrasse, D. (Co-premier auteur), Zytnicki, M. (Co-premier auteur), Andrey, P., Houba Hérin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., Gaudin, V. (Auteur de correspondance) (2016). The Arabidopsis hnRNP-Q Protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of