The LEAFY floral regulator displays pioneer transcription factor properties

Xuelei Lai, Romain Blanc-Mathieu, Loïc Grandvuillemin, Ying Huang, Arnaud Stigliani, Jérémy Lucas, Emmanuel Thévenon, Jeanne Loue-Manifel, Laura Turchi, Hussein Daher, et al.

The LEAFY floral regulator displays pioneer transcription factor properties

Xuelei Lai[1,*], Romain Blanc-Mathieu[1,*], Loïc GrandVuillemin[1,^], Ying Huang[2,^], Arnaud Stigliani[1,3], Jérémy Lucas[1], Emmanuel Thévenon[1], Jeanne Loue-Manifel[1,4], Laura Turchi[1], Hussein Daher[1,5], Eugenia Brun-Hernandez[1], Gilles Vachon[1], David Latrasse[2], Moussa Benhamed[2,6], Renaud Dumas[1], Chloe Zubieta[1] and François Parcy[1,**]

[1]Laboratoire Physiologie Cellulaire et Végétale, Univ. Grenoble Alpes, CNRS, CEA, INRAE, IRIG-DBSCI-LPCV, 17 avenue des martyrs, F-38054, Grenoble, France

[2]Université Paris-Saclay, CNRS, INRAE, Univ. Evry, Institute of Plant Sciences Paris-Saclay (IPS2), 91405, Orsay, France.

[3]The Bioinformatics Centre, Department of Biology and Biotech and Research Innovation Centre, University of Copenhagen, Ole Maaløes Vej 5, DK2200 Copenhagen N, Denmark

[4]Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, F-69342 Lyon, France

[5]Institut de Biologie Structurale, Université Grenoble Alpes, CEA, CNRS, Grenoble, France

[6]Université de Paris, Institute of Plant Sciences Paris-Saclay (IPS2), F-75006 Paris, France

* contributed equally

^ contributed equally to the work

** Correspondence: François Parcy (francois.parcy@cea.fr)

Short summary: LFY is a master floral regulator in Arabidopsis and has been suggested to act as a pioneer TF, a special class of TFs that are able to access closed chromatin regions and trigger gene expression. Here, we show that LFY fulfills several pioneer TF properties that may contribute to launch the floral gene expression program.

Abstract

Pioneer transcription factors (TFs) are a special category of TFs with the capacity to bind to closed chromatin regions in which DNA is wrapped around histones and may be highly methylated. Subsequently, pioneer TFs are able to modify the chromatin state to initiate gene expression. In plants, LEAFY (LFY) is a master floral regulator and has been suggested to act as a pioneer TF in Arabidopsis. Here, we demonstrate that LFY is able to bind both methylated and non-methylated DNA using a combination of *in vitro* genome-wide binding experiments and structural modeling. Comparisons between regions bound by LFY *in vivo* and chromatin accessibility data suggest that a subset of LFY bound regions is occupied by nucleosomes. We confirm that LFY is able to bind nucleosomal DNA *in vitro* using reconstituted nucleosomes. Finally, we show that constitutive LFY expression in seedling tissues is sufficient to induce chromatin accessibility in the LFY direct target genes, *APETALA1* and *AGAMOUS*. Taken together, our study suggests that LFY possesses key pioneer TF features that contribute to launch the floral gene expression program.

Introduction

Proper gene regulation is essential to all living organisms, controlling processes from basic development to environmental response. Gene regulation requires the finely orchestrated activity of transcription factors (TFs) that recognize specific DNA sequences in gene regulatory regions and activate or repress transcription of their target genes. While the binding of most TFs to DNA is restricted to accessible regions of the genome, a specific type of TF, called a "pioneer", is able to access its cognate binding site even in closed, nucleosome-rich chromatin regions (Magnani et al., 2011; Iwafuchi-Doi and Zaret, 2014; Iwafuchi-Doi and Zaret, 2016; Zaret, 2020). The ability to bind nucleosomal DNA *in vivo* and *in vitro* is a defining characteristic of pioneer TFs and has been well-established for diverse mammalian pioneer TFs (Fernandez Garcia et al., 2019). As DNA in closed chromatin regions is often highly methylated, another emerging feature of pioneer TFs is their capability to bind DNA in a methylation insensitive manner (Zhu et al., 2016; Mayran and Drouin, 2018). Some pioneer TFs are even able to directly recruit DNA demethylases at methylated sites, thereby facilitating the remodeling of closed regions (Iwafuchi-Doi, 2018).

Pioneer TFs are often master regulators controlling developmental transitions, with the mammalian pluripotency factors Octamer binding TF (OCT4), SRY (sex determining region Y)-box 2 (SOX2), and Kruppel-like factor 4 (KLF4) representing some of the most well-studied (Soufi et al., 2015). These factors bind to closed chromatin regions and induce their opening or remodeling, so that genes they contain can be activated by the pioneer TFs themselves or by other TFs called settlers (Sherwood et al., 2014; Slattery et al., 2014). The modification of the chromatin landscape by pioneer TF can be accomplished either directly by triggering DNA detachment from nucleosomes (Dodonova et al., 2020; Michael et al., 2020), or indirectly by the recruitment of ATP-dependent cellular machineries, such as chromatin remodelers that remove or modify adjacent nucleosomes in order to prime downstream regulatory events (Hu et al., 2011; King and Klose, 2017). Such capacity to modify DNA accessibility is another defining feature of pioneer TFs (Iwafuchi-Doi and Zaret, 2014).

In plants, the only TF reported as pioneer TF so far is LEAFY COTYLEDON1 (LEC1), a seed specific TF involved in embryonic epigenetic reprogramming (Tao et al., 2017). LEC1 was shown to promote the initial establishment of an active chromatin state of its target gene in silenced chromatin and activate its expression *de novo*. Pioneer TF activity was also suggested for two types of factors controlling flower development, the MADS homeotic TFs (Pajoro et

al., 2014; Denay et al., 2017) and the master floral regulator, LEAFY (LFY) (Sayou et al., 2016). The MADS TFs, including APETALA1 (AP1) and SEPALLATA3, were shown to be able to access closed chromatin regions to specify floral organs, and were thus postulated to act as pioneer TFs (Pajoro et al., 2014). However, mammalian MADS TFs do not seem to act as pioneer factors and thus the identification of AP1 and SEP3 as potential pioneers remains speculative (Sherwood et al., 2014). In contrast to the MADS TFs, one previous study suggest that LFY may have pioneer activity (Sayou et al., 2016). LFY is a master regulator specifying the floral identity of meristems. It directly induces the floral homeotic genes *AP1*, *APETALA3 (AP3)* and *AGAMOUS (AG)* (Parcy et al., 1998; Wagner et al., 1999; Lohmann et al., 2001; Chae et al., 2008; Yamaguchi et al., 2013; Chahtane et al., 2013). *AG* and *AP3* are known to be under the repression of Polycomb repressive complexes in seedlings (Goodrich et al., 1997; Turck et al., 2007; Calonje et al., 2008). This suggests that their activation during flower development requires modifications of their chromatin landscape and that the direct binding of LFY to their regulatory regions might trigger. Consistent with this, LFY was suggested to be able to access closed chromatin regions *in vivo* (Sayou et al., 2016). Moreover, LFY's role is not confined to conferring a flower fate to meristems. It can also contribute to meristem emergence (Moyroud et al., 2010; Chahtane et al., 2013; Yamaguchi et al., 2013), and together with its co-regulators such as the homeodomain TF WUSCHEL or the F-Box protein UNUSUAL FLORAL ORGANS, it can even induce floral meristem formation from root or leaf tissue, respectively (Levin and Meyerowitz, 1995; Gallois et al., 2004; Risseeuw et al., 2013). Taken together, these data indicate that LFY has the full capability of reprogramming cell fate, a property often requiring pioneer activity. However, whether LFY is truly able to directly bind closed chromatin regions and change their status has yet to be demonstrated.

Here, we address the pioneer activity of LFY *in vitro* and *in vivo*. Firstly, we determined whether LFY binding was sensitive to DNA methylation. For this, we combined *in vitro* LFY genome-wide binding data using methylated and unmethylated genomic DNA and structural analysis. These experiments demonstrated that LFY binding is only mildly sensitive to DNA methylation. In order to test whether LFY binding was compatible with the presence of nucleosomes, we compared LFY binding data from chromatin immunoprecipitation sequencing (ChIP-seq) and chromatin accessibility data. Based on these comparisons, we found that LFY could access a number of closed chromatin regions and that LFY colocalizes with nucleosomes in some regions *in vivo*. Using electrophoretic mobility shift assays (EMSA), we further showed that LFY was able to directly bind nucleosomes *in vitro*. Finally, chromatin accessibility assays

demonstrated that LFY constitutive expression was sufficient to increase chromatin accessibility in genomic regions including its known target genes *AP1* and *AG*. Taken together, these data establish that LFY is able to act as a pioneer TF in the regulation of important target genes critical for the establishment of floral fate.

Results and discussion

LFY is weakly sensitive to DNA methylation

Both the presence of nucleosomes and DNA methylation usually reduce TFs access and binding to their target DNA (Yin et al., 2017; Klemm et al., 2019). DNA methylation in promoter regions (including in euchromatin) is often associated with transcriptional silencing (Zhang et al., 2006). This is also the case in the process of flowering and flower development (Yang et al., 2015), suggesting a crosstalk between the DNA methylation landscape and TF action during this process.

In order to assess the effect of DNA methylation on LFY binding, we applied DNA Affinity Purification sequencing (DAP-seq) (O'Malley et al., 2016). Similar to ChIP-seq, this technique allows the identification of the genomic regions bound by a TF but uses naked DNA and a recombinant TF. We used Arabidopsis genomic DNA extracted from seedlings that was either PCR amplified (ampDAP, DNA cleared of methylation) or not amplified (DAP, DNA retaining methylation). Both experiments were performed in triplicates with high reproducibility (Supplemental Figure 1; Supplemental Table 1). As controls, we used two TFs described as methylation sensitive based on available ampDAP and DAP datasets (O'Malley et al., 2016) (Supplemental Figure 2). For each genomic region bound by a given TF, we plotted the DAP/ampDAP signal ratio as a function of i) the methylation density in the whole bound region (based on Arabidopsis seedling methylation maps (Zhang et al., 2016) (Figure 1A-C).), and ii) the number of methylated cytosines within the best TF binding site (TFBS), identified using position weight matrices in each bound region (Figure 1D-F). Whereas an increased number of methylated cytosines in the whole bound region or in the TFBS itself strongly decreases the binding for the two methylation sensitive TFs in DAP relative to ampDAP, LFY binding was only mildly affected (Figure 1A-F). Finally, we designed a specific procedure to compute the effect of methylation on each individual cytosine possibly present in the best TFBS (Supplemental Figure 3-5). In the case of LFY, we identified two positions where the binding is increased by cytosine methylation (positions 4 on the forward DNA strand and 5 on the reverse), and other positions (2,3,7,8 on the forward strand and 1,3,4,9 on the reverse) where

the binding is only mildly inhibited (Figure 1G). In contrast, methylation is inhibitory for the two methylation sensitive TFs in most positions where a cytosine could possibly be present (Figure 1H-I). We then examined this result in the context of the protein-DNA structural data. It is known that DNA methylation inhibits the DNA binding of most TFs because the 5-methyl group of methylcytosine often clashes with protein residues that are involved in specific base readout (Yin et al., 2017). Some TFs, however, are not sensitive or even favor methylated DNA because direct hydrophobic interactions form between the methyl group and the TF, as it is the case for homeodomain TFs (Yin et al., 2017) or for some basic leucine zipper TFs (Weber et al., 2019). In the case of LFY, the structural analysis of its DNA binding domain in complex with DNA (Hamès et al., 2008) (PDB 2VY1 and 2VY2) provided a biochemical explanation for the observed positive and negative effects (Supplemental Figure 6). In particular, the hydrophobic contacts between LFY and DNA are likely to be enhanced by the presence of a methyl group in positions 4 and 5 of the LFY binding site (LFYBS), consistent with the DAP versus ampDAP analysis (Figure 1G). How much this weak sensitivity to DNA methylation might help LFY to perform its master function during flowering remains to be determined.

Such computational analysis has the potential to be generalized to all TFs for which DAP and ampDAP data are available. It represents a powerful complement to methylation-sensitive SELEX (systematic evolution of ligands by exponential enrichment) analysis which was used to detect the effect of methylation to TF-DNA binding using randomized DNA sequences (Yin et al., 2017).

A subset of LFY binding occurs at closed chromatin regions

Next, we analyzed how *in vivo* factors (including the chromatin state) affect LFY DNA binding. For this, we compared LFY binding *in vitro* and *in vivo* by plotting the coverage of LFY DAP-seq peaks versus that of LFY ChIP-seq peaks. LFY ChIP-seq was obtained from *35S::LFY* seedlings or floral meristems (Sayou et al., 2016; Goslin et al., 2017). This analysis identified genome regions well bound in both experiments (Figure 2A; Supplemental Figure 7A; colored in light purple to red). However, it also highlighted the existence of regions much better bound *in vivo* (ChIP-specific regions colored in deep purple) or *in vitro* (DAP-specific regions colored in orange). The existence of ChIP-specific regions indicated that LFY DNA binding might increase due to interactions with *in vivo* factors. The presence of DAP-specific regions indicated that the *in vivo* context inhibits LFY from binding to some genomic regions despite their high affinity for LFY binding observed in DAP-seq.

To understand whether chromatin conformation could play a role in this inhibition, we analyzed the chromatin state of each region using DNaseI-seq data obtained in two-week-old seedlings (Zhang et al., 2012), a high DNaseI-seq signal being indicative of an open region (Figure 2B; Supplemental Figure 7B). We found that many of the DAP-specific regions have a low DNaseI-seq signal, typical of closed chromatin regions. This suggests that a closed chromatin state inhibits LFY binding. However, as previously observed (Sayou et al., 2016), a number of regions are bound in ChIP-seq despite low DNaseI-seq signal (right panels on Figure 2B and Supplemental Figure 7B). Overall, this analysis suggests that while the closed chromatin context is generally inhibitory for LFY binding, some closed chromatin regions can still be bound. To analyze what type of closed regions are most likely to be bound, we analyzed the upper and lower deciles of regions ranked based on their ChIP-seq signal, the upper decile contains regions well bound in ChIP-seq whereas the lower has regions poorly bound in ChIP-seq (but bound in DAP-seq). The distribution of nine chromatin states (as defined in the literature (Sequeira-Mendes et al., 2014)) changes drastically between the two deciles (Figure 2C; Supplemental Figure 7C). Chromatin states 7, 8, and 9 (the most compacted states that includes heterochromatin) are highly represented among regions better bound in DAP or in a control set of unbound regions but they are less present in the ChIP-specific regions. Conversely, states 1-5, which are closer to gene units or targets of Polycomb repression (state 5) are more frequently found in regions better bound in ChIP-seq than in DAP-seq. This analysis underlines the fact that LFY can bind closed chromatin regions but not those with the highest degree of compactness.

As closed chromatin regions are often occupied by nucleosomes, and since *in vivo* data suggests that LFY might be able to bind some of these regions, we wondered whether LFY binding was compatible with the presence of nucleosomes. To test this, we compared the position of LFY ChIP-seq peaks with that of nucleosomes (based on MNase-seq data (Zhang et al., 2015)). We found that nucleosomes were indeed enriched at the center of LFY ChIP-seq peaks in closed regions (Figure 2D; Supplemental Figure 7D), but not in open ones (Figure 2E; Supplemental Figure 7), suggesting that LFY might be able to directly bind nucleosomal DNA *in vivo*. We mapped the LFYBS in nucleosome-occupied LFY ChIP-seq peaks that are also found in DAP-seq experiments to ensure they contain a *bona fide* LFY dimer binding site. We found a slight enrichment at the center of the nucleosome, around the dyad position which is a site commonly bound by pioneer TFs (Figure 3A; Supplemental Figure 8) (Zaret, 2020). However, since these genomic data are established on mixtures of tissues, they are not sufficient to firmly establish that LFY is indeed able to bind nucleosomal DNA.

LFY binds nucleosomal DNA at specific sites *in vitro*

Next, we tested whether LFY has the capacity to bind nucleosomal DNA *in vitro*. We first assembled nucleosomes using the Widom 601 strong nucleosome positioning sequence (Lowary and Widom, 1998; McGinty and Tan, 2015), in which a LFYBS was inserted at different positions (C1-C7 around the dyad and E1-E7 farther away) (Figure 3B; Supplemental Table 2). With nucleosomes assembled with a LFYBS at position C2 and C7, we observed a gel shift upon addition of LFY that is absent with LFYBS at positions C1, C3-C6, E1-E7 or with no LFYBS, demonstrating that LFY binds nucleosomal DNA in a sequence specific manner and only with a LFYBS present at specific positions (C2, located around the dyad, and C7, located one helix turn apart from C2, with the LFYBS exposed to the outer nucleosome surface (Figure 3B and C; Supplemental Figure 9)). This property is consistent with structural data showing that LFY binds a single side of the DNA (Hamès et al., 2008) that needs to be exposed to the outer surface (like C2 or C7) and not hidden by histones (C1, C3 to C6). It is worth noting that LFY DNA binding presents structural similarities with that of the pioneer TF FoxA in animals, whose helix-turn-helix DNA binding domain (DBD) mainly contacts the major groove, and with a flexible loop contacting the minor groove, both of which are located in the same side of the DNA (Zaret and Carroll, 2011; Fernandez Garcia et al., 2019). Using the same methodology, as a negative control, we tested nucleosomal DNA binding of the TF REGULATOR OF AXILLARY MERISTEMS 1 (RAX1), a direct downstream target of LFY (Chahtane et al., 2013). We found that RAX1 cannot associate with nucleosomes even when its binding site is exposed to the outer nucleosome surface and at the dyad (Supplemental Figure 10), suggesting that RAX1 is unlikely a pioneer TF. We also assembled nucleosomes with two regions of the *AP1* gene, a known early activated LFY target (Parcy et al., 1998; Wagner et al., 1999; Benlloch et al., 2011). These regions were taken from *AP1* first intron and *AP1* promoter (annotated as *AP1* intron and *AP1* pro, respectively, in Figure 3D). They are both bound by LFY *in vivo* (ChIP-seq (Moyroud et al., 2011; Winter et al., 2011; Sayou et al., 2016; Goslin et al., 2017)) and *in vitro* (DAP-seq in Figure 3D), and with well-defined nucleosome signals from MNase-seq in both seedlings and flower tissues (Zhang et al., 2015) (Figure 3D). We observed that LFY was able to bind to these nucleosomes (Figure 3E and Supplemental Figure 11), showing that LFY nucleosomal DNA binding also occurs within Arabidopsis genomic regions.

LFY constitutive expression induce changes in chromatin accessibility and nucleosome positioning

One key characteristic feature of pioneer TFs is their ability to modify the status of closed chromatin regions (Iwafuchi-Doi and Zaret, 2014). To test whether LFY is able to do so, we examined whether it could alter chromatin accessibility when ectopically expressed in seedlings. We selected regions bound by LFY in ChIP-seq (Figure 4A) (Sayou et al., 2016) that are mapped with a nucleosome in wild-type seedlings but not in closed flower buds (Zhang et al., 2015). These regions are adjacent to 496 genes including 54 that are upregulated by LFY (induced by LFY overexpression or downregulated in *lfy* mutants (William et al., 2004; Schmid et al., 2005; Winter et al., 2011)). Among these genes are present those encoding for early floral regulators AP1, AG and ULTRAPETALA 1 (ULT1) (Moreau et al., 2016), as well as other TFs and cell wall remodeling enzymes (Supplemental Table 4). We focused on these three well-established floral regulators. Using Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE)-qPCR that identifies accessible chromatin regions (including those depleted of nucleosomes) and MNase-qPCR that locates the nucleosome (Figure 4 and Supplemental Figure 12), we tested whether ectopic LFY expression (*35S::LFY*) could alter the local chromatin as compared to two-week-old Col-0 seedlings where endogenous LFY is not yet highly induced. Focusing on different types of regions (promoters or introns), we found that LFY ectopic expression increased accessibility of these regions (Figure 4A) but not for three control regions (*Actin2, AT2G38220* and *AT4G22285*) with poor accessibility in seedlings and where LFY does neither bind *in vivo* nor *in vitro* (Figure 4B). The analysis of the *AP1* locus is particularly interesting. This gene is a direct and early target of LFY and contains two LFY binding peaks, one in its promoter (P2-P4 in *AP1*; Figure 4A) and one in its first intron (P1 in *AP1*; Figure 4A). *AP1* promoter can be induced by LFY in seedling leaves, independently of flower formation (Parcy et al., 1998). According to DNaseI-seq signal, *AP1* promoter is already open in seedlings and with two nucleosomes detected by MNase-seq whereas the first intron is closed (Figure 4A). We found that LFY expression triggers a strong accessibility increase in the intron (3-fold, P1 in *AP1,* Figure 4A) and a more moderate increase in the promoter (P2 and P3 in *AP1*; Figure 4A). We also found increased accessibility in *ULT1* promoter and *AG* regulatory intron (Figure 4A). Consistent with FAIRE-qPCR, MNase-qPCR experiments on the same regions gave comparable results, indicating that the increased accessibility is most likely due to nucleosome depletion (Supplemental Figure 12). Our results are consistent with a recent report using LFY induction in root explants, a system where *AP1* promoter appears in a closed chromatin state and opens after LFY induction (Jin et al., 2020).

To conclude, we have obtained evidence that LFY shares some properties with pioneer TFs. However, the pioneer function is likely a spectrum of activities: TFs that play central roles in developmental transitions, such as LFY, are able to fulfill a pioneer role under certain chromatin conditions or cellular contexts, for example in the presence of specific cofactors (Zaret, 2020) and/or for a few distinct loci (Li et al., 2019). Taken together, our *in vitro* and *in vivo* results demonstrate the essential properties of pioneer TFs- the competence to bind closed chromatin and the ability to trigger subsequent opening of these closed regions- are properties of LFY in the context of at least a few key floral regulatory targets.

Methods

Detailed methods are available in Supplemental Information (SI).

DAP-seq and AmpDAP-seq

The input library of ampDAP-seq was PCR amplified from Col-0 genomic DNA constructed according to published protocol (O'Malley et al., 2016; Bartlett et al., 2017; Lai et al., 2020). For the DAP-seq input library, genomic DNA was extracted from two-weeks-old seedlings of a *35S::LFY* line (pCA26 #15) (Sayou et al., 2016) grown on 0.5 x Murashige and Skoog medium in long-day conditions. The LFY protein was produced using an *in vitro* transcription/translation system (Promega L3260). DAP-seq was carried out according to published protocol with minor modifications (O'Malley et al., 2016; Bartlett et al., 2017). The immunoprecipitated DNA fragments were PCR amplified for 20 cycles, and purified using AMPure XP magnetic beads (Beckman). Individual libraries were pooled with equal molarity, and sequenced on Illumina HiSeq (Genewiz). Both DAP-seq and ampDAP-seq were performed in triplicates.

Bioinformatic analyses

Regions bound by LFY *in vitro*, to methylated and non-methylated genomic DNA, were detected from DAP-seq data generated in this study as described previously (Lai et al., 2020). *In vivo* LFY bound regions were obtained from ChIP-seq data from two weeks-old seedling *35S::LFY* tissue (Sayou et al., 2016) and inflorescence tissue of *35S:LFY-GR ap1 cal* (Goslin et al., 2017). DNA accessibility, nucleosomes position and probability of methylcytosines were obtained from processed DNaseI-seq, MNase-seq and bisulfite sequencing data (Zhang et al., 2012, Zhang et al., 2015, Zhang et al., 2016), respectively. Microarray analyses were performed

on published data of LFY overexpression lines (GEO: GSE911) (William et al., 2004), and (GEO: GSE28062) (Winter et al., 2011) and data from *lfy* lines (GEO: SE576) (Schmid et al., 2003) and available on AtGenExpress (Schmid et al., 2005). *Ad hoc* Python, R and Shell scripts developed in our laboratory are available at https://github.com/Bioinfo-LPCV-RDF/TF_genomic_analysis.

Nucleosome reconstruction and Electrophoretic mobility shift assay (EMSA)

The recombinant AtLFYΔ40 was produced in *E. coli* Rosetta2 (DE3) strain (Novagen) and purified by nickel affinity purification (GE Healthcare). The production and purification of histones was carried out according to published protocols (Shim et al., 2012). The nucleosome assembly was performed by salt dilution method (Okuwaki et al., 2005) using the corresponding DNA probes (Supplemental Tables 2 and 3). Nucleosomes of interest were incubated with 500 µM AtLFYΔ40 for 1 hour at room temperature. The EMSAs were run on 5 % non-denaturing polyacrylamide gels for one hour at 4 °C at 120 V, and visualized by Cy5 signal on ChemiDoc MP Imager (BIO-RAD).

FAIRE-qPCR

FAIRE-qPCR was performed on two-week-old seedlings of Col-0 and *35S::LFY* (Sayou et al., 2016). 1 g of plant material was crosslinked by formaldehyde for 15 min using vacuum infiltration. Crosslinking was quenched by adding glycine solution to 0.125 M. Nuclei were isolated using NI buffer (see SI) and then resuspended in 1 mL of FAIRE Lysis Buffer (see SI). The crosslinked DNA was sheared to an average size of 200 - 300 bp using Covaris S220. An aliquot was used as control DNA and directly treated with RNAse A+T1 cocktail enzyme mix (Thermo Fisher Scientific) followed by proteinase K treatment and reverse crosslinked. The non-de-crosslinked samples were treated as for control DNA and subject to a phenol/chloroform extraction. DNA was then quantified using Qubit dsDNA HS kit (Thermo Fisher Scientific) and the ratio between nucleosome-free DNA versus total DNA was determined by qPCR analysis using 20 ng of template DNA for each reaction.

Accession Numbers

LFY DAP-seq sequencing data from this article can be found in the NCBI GEO data libraries under accession numbers GSE160013.

Acknowledgements

We thank K. Kaufmann for discussions. This project was supported by the ANR-DFG project Flopinet (ANR-16-CE92-0023-01) to C.Z. and F.P., and GRAL, a program from the Chemistry Biology Health (CBH) Graduate School of University Grenoble Alpes (ANR-17-EURE-0003). to C.Z., F.P. and A.S.

Authors contributions

F.P., C.Z., and R.D. designed and supervised the project, R.B.M., J.L., A.S., and L.T. performed bioinformatics analyses, L.G., G.V., J.L.M., H.D., E.T. and E.B.H. performed biochemical analyses, X.L. performed DAP-seq, Y.H. performed FAIRE-qPCR and MNase-qPCR supervised by M.B. and D.L., F.P., C.Z., and X.L. wrote the paper with the help of all authors.

References

Benlloch, R., Kim, M. C., Sayou, C., Thévenon, E., Parcy, F., and Nilsson, O. (2011). Integrating long-day flowering signals: A LEAFY binding site is essential for proper photoperiodic activation of APETALA1. *Plant J.* **67**:1094–1102.

Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., and Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**:1659–1672.

Calonje, M., Sanchez, R., Chen, L., and Sung, Z. R. (2008). EMBRYONIC FLOWER1 Participates in Polycomb Group–Mediated *AG* Gene Silencing in *Arabidopsis*. *Plant Cell* **20**:277–291.

Chae, E., Tan, Q. K.-G., Hill, T. A., and Irish, V. F. (2008). An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* **135**:1235–1245.

Chahtane, H., Vachon, G., Le Masson, M., Thévenon, E., Périgon, S., Mihajlovic, N., Kalinina, A., Michard, R., Moyroud, E., Monniaux, M., et al. (2013). A variant of LEAFY reveals its capacity to stimulate meristem development by inducing RAX1. *Plant J.* **74**:678–689.

Chua, E. Y. D., Vasudevan, D., Davey, G. E., Wu, B., and Davey, C. A. (2012). The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.* **40**:6338–6352.

Denay, G., Chahtane, H., Tichtinsky, G., and Parcy, F. (2017). A flower is born: an update on Arabidopsis floral meristem formation. *Curr. Opin. Plant Biol.* **35**:15–22.

Dodonova, S. O., Zhu, F., Dienemann, C., Taipale, J., and Cramer, P. (2020). Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature* **580**:669–672.

Fernandez Garcia, M., Moore, C. D., Schulz, K. N., Alberto, O., Donague, G., Harrison, M. M., Zhu, H., and Zaret, K. S. (2019). Structural Features of Transcription Factors Associating with Nucleosome Binding. *Mol. Cell* **75**:921–932.

Gallois, J.-L., Nora, F. R., Mizukami, Y., and Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes Dev.* **18**:375–380.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M., and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**:44–51.

Goslin, K., Zheng, B., Serrano-Mislata, A., Rae, L., Ryan, P. T., Kwaśniewska, K., Thomson, B., Ó'Maoiléidigh, D. S., Madueño, F., Wellmer, F., et al. (2017). Transcription Factor Interplay between LEAFY and APETALA1/CAULIFLOWER during Floral Initiation. *Plant Physiol.* **174**:1097–1109.

Hamès, C., Ptchelkine, D., Grimm, C., Thevenon, E., Moyroud, E., Gérard, F., Martiel, J.-L., Benlloch, R., Parcy, F., and Müller, C. W. (2008). Structural basis for LEAFY floral switch function and similarity with helix-turn-helix proteins. *EMBO J.* **27**:2628–2637.

Hu, G., Schones, D. E., Cui, K., Ybarra, R., Northrup, D., Tang, Q., Gattinoni, L., Restifo, N. P., Huang, S., and Zhao, K. (2011). Regulation of nucleosome landscape and transcription factor targeting at tissue-specific enhancers by BRG1. *Genome Res.* **21**:1650–1658.

Iwafuchi-Doi, M. (2018). The mechanistic basis for chromatin regulation by pioneer transcription factors. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11**:e1427.

Iwafuchi-Doi, M., and Zaret, K. S. (2014). Pioneer transcription factors in cell reprogramming. *Genes Dev.* **28**:989–998.

Iwafuchi-Doi, M., and Zaret, K. S. (2016). Cell fate control by pioneer transcription factors. *Development* **143**:1833–1837.

Jin, R., Klasfeld, S., Zhu, Y., Fernandez Garcia, M., Xiao, J., Han, S.-K., Konkol, A., and Wagner, D. (2021). LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nat. Commun.* **12**:626.

King, H. W., and Klose, R. J. (2017). The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife* **6**:e22631.

Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **20**:207–220.

Lai, X., Stigliani, A., Lucas, J., Hugouvieux, V., Parcy, F., and Zubieta, C. (2020). Genome-wide binding of SEPALLATA3 and AGAMOUS complexes determined by sequential DNA-affinity purification sequencing. *Nucleic Acids Res.* **48**:9637–9648.

Levin, J. Z., and Meyerowitz, E. M. (1995). UFO: An Arabidopsis Gene lnvolved in Both Floral Meristem and Floral Organ Development. *Plant Cell* **7**:529–548.

Li, S., Bo Zheng, E., Zhao, L., and Liu, S. (2019). Nonreciprocal and Conditional Cooperativity Directs the Pioneer Activity of Pluripotency Transcription Factors. *Cell Rep.* **20**:2689–2703.

Lohmann, J. U., Hong, R. L., Hobe, M., Busch, M. A., Parcy, F., Simon, R., and Weigel, D. (2001). A Molecular Link between Stem Cell Regulation and Floral Patterning in Arabidopsis. *Cell* **105**:793–803.

Lowary, P. T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**:19–42.

Magnani, L., Eeckhoute, J., and Lupien, M. (2011). Pioneer factors: Directing transcriptional regulators within the chromatin environment. *Trends Genet.* **27**:465–474.

Mayran, A., and Drouin, J. (2018). Pioneer transcription factors shape the epigenetic landscape. *J. Biol. Chem.* **293**:13795–13804.

McGinty, R. K., and Tan, S. (2015). Nucleosome Structure and Function. *Chem. Rev.* **115**:2255–2273.

Michael, A. K., Grand, R. S., Isbel, L., Cavadini, S., Kozicka, Z., Kempf, G., Bunker, R. D., Schenk, A. D., Graff-Meyer, A., Pathare, G. R., et al. (2020). Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science* **368**:1460–1465.

Moreau, F., Thévenon, E., Blanvillain, R., Lopez-Vidriero, I., Franco-Zorrilla, J. M., Dumas, R., Parcy, F., Morel, P., Trehin, C., and Carles, C. C. (2016). The Myb-domain protein ULTRAPETALA1 INTERACTING FACTOR 1 controls floral meristem activities in *Arabidopsis*. *Development* **143**:1108–1119.

Moyroud, E., Kusters, E., Monniaux, M., Koes, R., and Parcy, F. (2010). LEAFY blossoms. *Trends in Plant Sci.* **15**:346–352.

Moyroud, E., Minguet, E. G., Ott, F., Yant, L., Posé, D., Monniaux, M., Blanchet, S., Bastien, O., Thévenon, E., Weigel, D., et al. (2011). Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor. *Plant Cell* **23**:1293–1306.

Okuwaki, M., Kato, K., Shimahara, H., Tate, S., and Nagata, K. (2005). Assembly and Disassembly of Nucleosome Core Particles Containing Histone Variants by Human Nucleosome Assembly Protein I. *Mol. Cell. Biol.* **25**:10639–10651.

O'Malley, R. C., Huang, S. shan C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A., and Ecker, J. R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.

Pajoro, A., Madrigal, P., Muiño, J. M., Matus, J. T., Jin, J., Mecchia, M. A., Debernardi, J. M., Palatnik, J. F., Balazadeh, S., Arif, M., et al. (2014). Dynamics of chromatin

accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome. Biol.* **15**:R41.

Parcy, F., Nilsson, O., Busch, M. A., Lee, I., and Weigel, D. (1998). A genetic framework for floral patterning. *Nature* **395**:561–566.

Risseeuw, E., Venglat, P., Xiang, D., Komendant, K., Daskalchuk, T., Babic, V., Crosby, W., and Datla, R. (2013). An activated form of UFO alters leaf development and produces ectopic floral and inflorescence meristems. *PLoS ONE* **8**:e83807.

Sayou, C., Nanao, M. H., Jamin, M., Pose, D., Thevenon, E., Gregoire, L., Tichtinsky, G., Denay, G., Ott, F., Llobet, M. P., et al. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat. Commun* **7:11222**.

Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J. U. (2003). Dissection of floral induction pathways using global expression analysis. *Development* **130**:6001–6012.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. *Nat. Genet.* **37**:501–506.

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. *Plant Cell* **26**:2351–2366.

Sherwood, R. I., Hashimoto, T., O'Donnell, C. W., Lewis, S., Barkal, A. A., Van Hoff, J. P., Karun, V., Jaakkola, T., and Gifford, D. K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**:171–178.

Shim, Y., Duan, M.-R., Chen, X., Smerdon, M. J., and Min, J.-H. (2012). Polycistronic coexpression and nondenaturing purification of histone octamers. *Anal. Biochem* **427**:190–192.

Slattery, M., Zhou, T., Yang, L., Dantas Machado, A. C., Gordân, R., and Rohs, R. (2014). Absence of a simple code: How transcription factors read the genome. *Trends Biochem. Sci.* **39**:381–399.

Soufi, A., Garcia, M. F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K. S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**:555–568.

Tao, Z., Shen, L., Gu, X., Wang, Y., Yu, H., and He, Y. (2017). Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature* **551**:124–128.

Turck, F., Roudier, F., Farrona, Sara., 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 Genet.* **3(6):e86**.

Wagner, D., Sablowski, R. W. M., and Meyerowitz, E. M. (1999). Transcriptional Activation of APETALA1 by LEAFY. *Science* **285**:582–584.

Weber, E., Buzovetsky, O., Heston, L., Yu, K.-P., Knecht, K. M., El-Guindy, A., Miller, G., and Xiong, Y. (2019). A Noncanonical Basic Motif of Epstein-Barr Virus ZEBRA Protein Facilitates Recognition of Methylated DNA, High-Affinity DNA Binding, and Lytic Activation. *J. Virol.* **93**:e00724-19.

William, D. a, Su, Y., Smith, M. R., Lu, M., Baldwin, D. a, and Wagner, D. (2004). Genomic identification of direct target genes of LEAFY. *Proc. Natl. Acad. Sci. USA* **101**:1775–1780.

Winter, C. M., Austin, R. S., Blanvillain-Baufumé, S., Reback, M. A., Monniaux, M., Wu, M. F., Sang, Y., Yamaguchi, A., Yamaguchi, N., Parker, J. E., et al. (2011). LEAFY Target Genes Reveal Floral Regulatory Logic, cis Motifs, and a Link to Biotic Stimulus Response. *Dev. Cell* **20**:430–443.

Yamaguchi, N., Wu, M. F., Winter, C. M., Berns, M. C., Nole-Wilson, S., Yamaguchi, A., Coupland, G., Krizek, B. A., and Wagner, D. (2013). A Molecular Framework for Auxin-Mediated Initiation of Flower Primordia. *Developmental Cell* **24**:271–282.

Yang, H., Chang, F., You, C., Cui, J., Zhu, G., Wang, L., Zheng, Y., Qi, J., and Ma, H. (2015). Whole-genome DNA methylation patterns and complex associations with gene structure and expression during flower development in Arabidopsis. *Plant J.* **81**:268–281.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **356**:eaaj2239.

Zaret, K. S. (2020). Pioneer Transcription Factors Initiating Gene Network Changes. *Annu. Rev. Genet.* **54**:367–385.

Zaret, K. S., and Carroll, J. S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25**:2227–2241.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W.-L., Chen, H., Henderson, I. R., Shinn, P., Pellegrini, M., Jacobsen, S. E., et al. (2006). Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. *Cell* **126**:1189–1201.

Zhang, W., Zhang, T., Wu, Y., and Jiang, J. (2012). Genome-Wide Identification of Regulatory DNA Elements and Protein-Binding Footprints Using Signatures of Open Chromatin in Arabidopsis. *Plant Cell* **24**:2719–2731.

Zhang, T., Zhang, W., and Jiang, J. (2015). Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol.* **168**:1406–1416.

Zhang, Q., Wang, D., Lang, Z., He, L., Yang, L., Zeng, L., Li, Y., Zhao, C., Huang, H., Zhang, H., et al. (2016). Methylation interactions in Arabidopsis hybrids require

RNA-directed DNA methylation and are influenced by genetic variation. *Proc. Natl. Acad. Sci. USA* **113**:E4248–E4256.

Zhu, H., Wang, G., and Qian, J. (2016). Transcription factors as readers and effectors of DNA methylation. *Nat Rev. Genet.* **17**:551–565.

Figure Legends

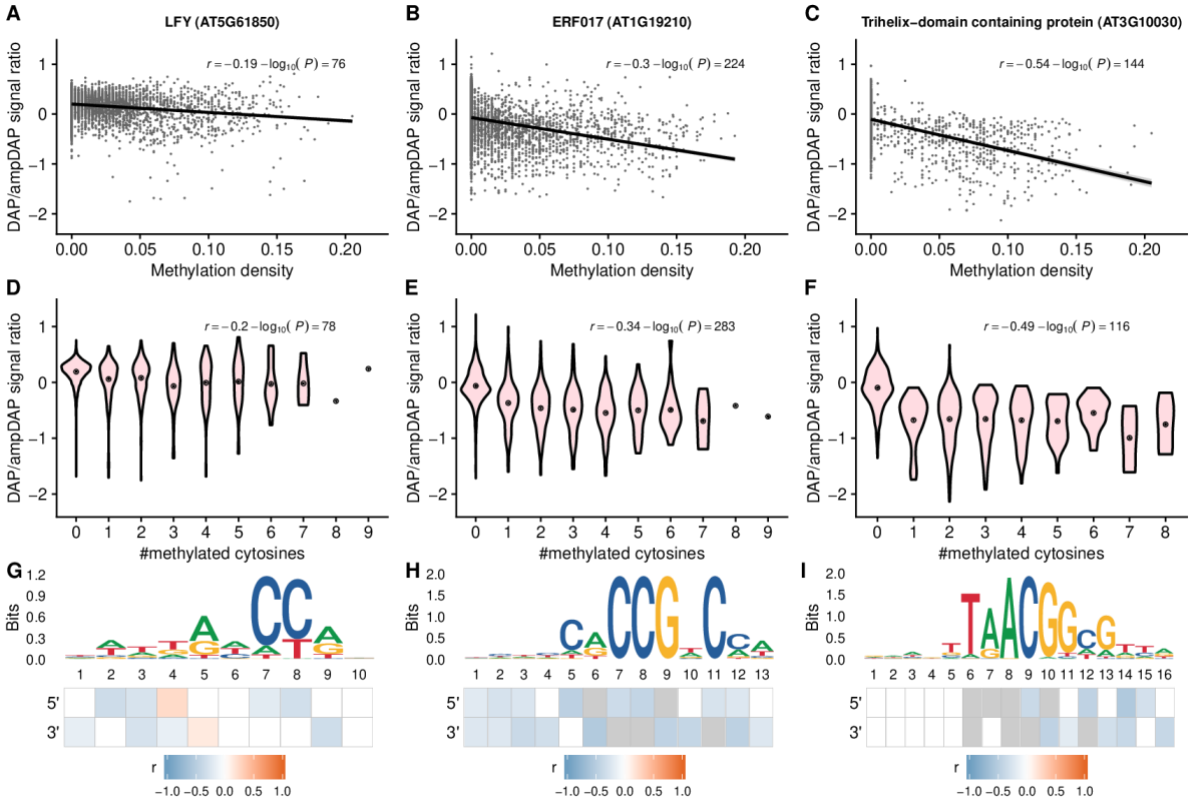


Figure 1: Cytosine methylation has a mild effect on LFY DNA-binding intensity.

Effect of cytosine methylation on DNA binding for three transcription factors: LFY (left), ERF017 (middle) and a trihelix-domain containing protein (right). (A-C) Biplots between the DAP/ampDAP signal ratio (peak normalized read coverage in the DAP experiment divided to that in the ampDAP experiment) in a $\log_{10}$ scale and methylation density (proportion of cytosines with a probability of methylation greater than 0.5) within transcription factor bound regions. The increasing methylation density has weaker effect on LFY than on the two other TFs. (D-F) Violin plots of DAP/ampDAP signal ratio in a $\log_{10}$ scale as a function of the number of methylated cytosines in the best TF binding site (TFBS) of each bound region. LFY binding is barely affected by the increased number of methylated cytosines. (G-I) Binding site sequence

motif for each TF and the methylation effect on each individual position. For LFY, a single half of the symmetric motif is shown. Heatmaps show the Pearson's correlation coefficient (*r*) between the DAP/ampDAP signal ratio in a $\log_{10}$ scale and the probability of methylation at each position of the best TFBSs. Blank positions have a high false discovery rate (> 5%) and grey indicates positions with less than ten cytosines in the dataset. Correlation are tested on both sides of a symmetric motif (G) or on both strands for non-symmetric motifs (H-I).

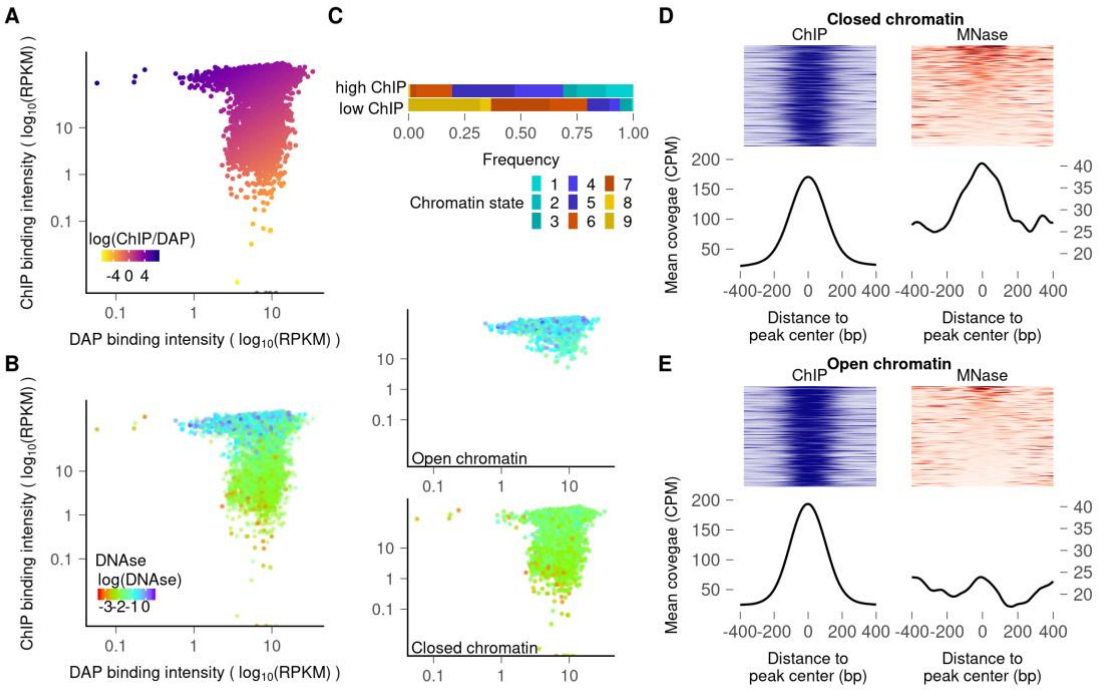
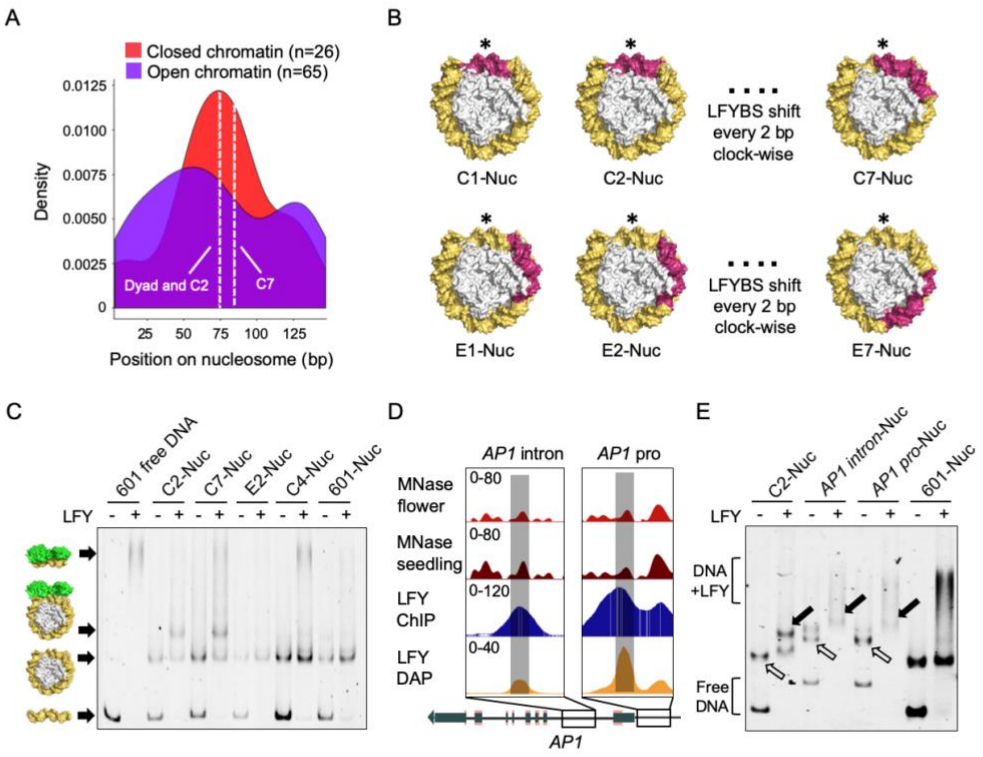


Figure 2: LFY is able to bind nucleosomes in closed chromatin regions.

(A) Plots comparing the LFY binding intensities (peak coverages) in ChIP-seq (Y-axis) vs DAP-seq (X-axis) experiments. Heat map is based on the ChIP-seq/DAP-seq intensity ratio. (B) Overlay of DNaseI signal (heat map) on LFY bound regions. The two panels on the right show the same regions split into open (upper panel) and closed (lower panel) chromatin states. (C) Distribution of chromatin states 1 to 9 according to (Sequeira-Mendes et al., 2014) for the first and last decile of LFY bound regions based on ChIP-seq signal and for a comparable set of regions not bound by LFY. (D-E) MNase signal around ChIP-seq peak centers in closed (D) or open (E) chromatin regions. Upper panels show ChIP-seq and MNase-seq coverage for each peak ordered based on MNase-seq signal. Lower panels represent the mean coverage. Regions above the dotted lane show a MNase signal in their center indicative of the presence of a nucleosome (see methods).

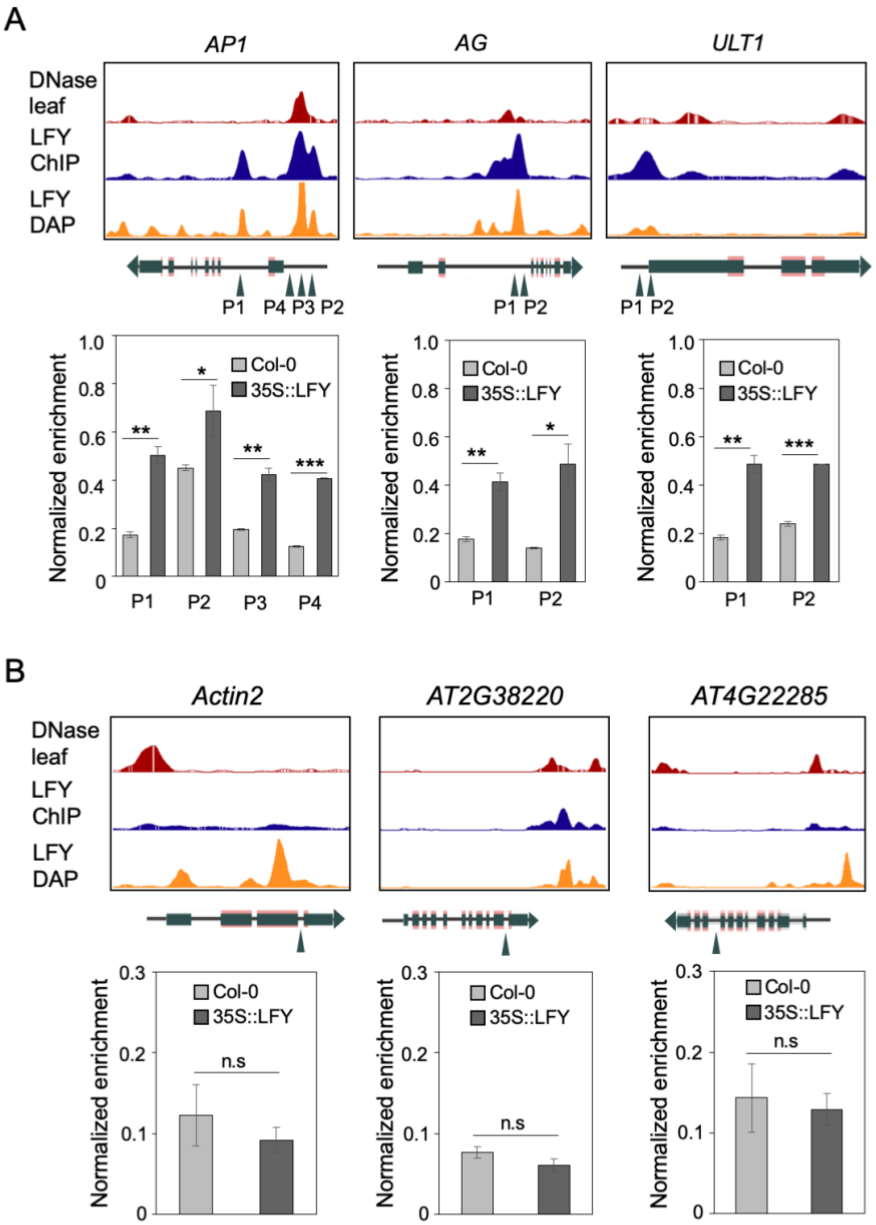


559

560 **Figure 3: LFY binds nucleosomes *in vitro*.**

561 (A) Density plot of the LFY best binding site present in ChIP-seq peaks along a canonical 147-
562 bp nucleosomal sequence in open and closed chromatin contexts for flower tissues. An
563 enrichment for LFY binding sites (LFYBS) around the dyad position (the center of the

nucleosomal DNA) is observed in closed chromatin regions. C2 (at dyad) and C7 positions are indicated. Alternative plots for different datasets and thresholds for binding sites selection are reported in Supplemental Figure 8. (B) Design of Widom 601 sequences (yellow orange) with a LFYBS (warm pink) inserted at different positions (central C1-C7 (top) and external E1-E7 (bottom)) on nucleosome (PDB: 3UT9 (Chua et al., 2012). * indicates the dyad. (C) Representative EMSA showing LFY binding to 601 nucleosomes with a LFYBS at positions C2 (labelled C2-Nuc) and C7, but not at E2, C4 or 601 nucleosome without a LFYBS (refer to Supplemental Figure 9 for the screening of LFY nucleosomal DNA binding at all other positions). Free DNA (C2 in the first 2 lanes, or present in the nucleosomal preparations) is shifted at the very top of the gels. 601-Nuc is made with wild-type 601 sequence (without a LFYBS): only free DNA is shifted due to non-sequence specific interactions with LFY. Cartoon on the side from bottom to top are free DNA, nucleosome alone, LFY-nucleosome complex and free DNA-LFY complex. (D) Genomic snapshots of LFY DAP-seq, ChIP-seq (seedlings tissue), and MNase-seq (seedlings and closed flower buds) at the *AP1* loci. *AP1* intron and *AP1* pro sequences used to assemble nucleosomes in (E) are highlighted in grey. Both regions are bound in DAP and ChIP, and with well-defined nucleosome signals, lower in floral tissue than in seedlings. (E) EMSA showing LFY binding to nucleosomes assembled with *AP1* intron and *AP1* pro sequences. AP1-intron-Nuc and AP1-pro-Nuc are longer than 601 due to the presence of amplification primers. Note some free 601 DNA is shifted despite the absence of LFYBS in the last lane. The hollow and solid arrows indicate the position of reconstituted nucleosomes and the shifted nucleosomes signals, respectively.



585

586 **Figure 4: LFY constitutive expression increases chromatin accessibility.**

587 (A) (Top) Genomic snapshots of chromatin accessibility (DNaseI-seq from 2-week-old Col-0

588 seedlings (Zhang et al., 2012)), LFY binding *in vitro* (DAP-seq using genomic DNA from 2-

589 week-old *35::LFY* seedlings), *in vivo* (ChIP-seq of 2-week-old *35::LFY* seedlings (Sayou et al.,

2016)) at *AP1*, *AG* and *ULT1* loci. The regions tested in FAIRE-qPCR are indicated by triangle arrows (P1-4 labels). (Bottom) FAIRE-qPCR of the indicated regions are performed in 2-week-old seedlings of Col-0 (pale gray) and *35S::LFY* (dark gray), respectively. Error bars represent means ± standard deviation. Significance test is performed by one-tailed students' t-test, *p<0.05, **p<0.01, ***p<0.001, n.s, not significant. (B) (Top) genome browser snapshots of three genomic regions devoid of LFY binding and poorly accessible in 2-week-old seedlings. (Bottom) FAIRE-qPCR on the indicated regions. Significance test is performed as per (A). The FAIRE-qPCR is performed by two biological replicates, with three technical replicates for each. The enrichment is normalized by input DNA in each experiment.