UNIVERSITY OF CALIFORNIA

Los Angeles

On Transcriptional Regulation by TOPLESS in *Arabidopsis thaliana*

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy in Molecular, Cellular, and Developmental Biology

by

Ao Liu

2018

ABSTRACT OF THE DISSERTAION

On transcriptional regulation by TOPLESS in *Arabidopsis thaliana*

by Ao Liu

Doctor of Philosophy in Molecular, Cellular and Developmental Biology

University of California, Los Angeles, 2018

Professor Jeffrey Aaron Long, Chair

How multicellular organisms modulate transcription in response to internal and external cues is a fundamental question in plant and animal development and growth. In *Arabidopsis*, TOPLESS(TPL), a Groucho/Tup1 like co-repressor, is involved in the transcriptional response to multiple plant hormones and plays a crucial role in a range of major developmental processes, such as embryonic apical-basal patterning and post-embryonic floral patterning. This dissertation describes the advances made toward the elucidation of the molecular mechanisms underlying transcriptional regulation by TPL.

As a co-repressor, TPL lacks intrinsic DNA binding activity and has been shown to be recruited by transcription factors with repressive motifs. To date, a range of transcription factors have been reported to interact with TPL, yet the functionality of most of these interactions is not known. We found that the members of the Class III Homedomain-Leucine Zipper (HD- ZIPIII) transcription factor family co-localize with TPL to form complexes at the promoters of

thousands of genes in the genome. HD-ZIPIIIs both antagonistically and cooperatively regulate entire pathways at multiple steps in a cell or tissue specific manner. The five members in the family form homo- and hetero-dimers and act both as activators and repressors. The association with TPL confers the repressive function of HD-ZIPIIIs.

Just like its counterparts in animals and fungi, TPL also associates with histone deacetylases(HDACs) to repress transcription. We characterized the binding profiles of TPL and HDA19, a class I histone deacetylase that interacts with TPL, and found they tend to bind to the promoters of dynamically expressed genes. Almost all HDA19 peaks have an overlapping TPL peak, while about half of the TPL peaks are in common with HDA19, suggesting HDA19 is involved only at a subset of TPL targets. Specifically, we found the three members of the evolutionarily conserved NAM/CUC (for NO APICAL MERISTEM and CUP-SHAPED COTYLEDON) transcription factors, *CUC1*, *CUC2* and *CUC3*, are regulated by TPL and HDA19 to control boundary specification during floral development.

Finally, we showed that TPL interacts with lysine 27-trimethylated histone H3(H3K27me3) and binds to a significant number of genomic regions with this modification, suggesting a novel layer of regulation which has not been reported in animals. A decrease in TPL binding was observed in plants that have a mutation in the Polycomb Repressive Complexes 2 (PRC2), which mediates the deposition of H3K27me3. TPL itself affects H3K27me3 levels and interacts with a component of the PRC1 complex, which recognizes H3K27me3 and further contributes to chromatin compaction. Genes potentially regulated by both TPL and PRCs are involved in multiple developmental processes and stress responses. These data suggest that upon

the recruitment of TPL, binding affinity is enhanced by H3K27me3 marked nucleosomes and TPL may further contributes to the repressive transcriptional state via the association with Polycomb group (PcG) proteins. Overall, we propose a model of transcriptional regulation by TPL that consists of three steps: TPL first is recruited by specific transcription factors; the association of TPL at some target genes is enhanced by the presence of H3K27me3; and TPL forms co-repressor complexes with HDA19 and possibly PcGs to repress gene expression.

The dissertation of Ao Liu is Approved.

Amander Therese Clark

Steven Erik Jacobsen

Siavash K. Kurdistani

Jeffrey Aaron Long, Committee Chair

University of California, Los Angeles

2018

DEDICATION

I would like to dedicate my thesis to my parents for their love and unconditional support.

Table of Contents

List of Figures and Tables

Acknowledgements

I first would like to thank my thesis committee: Dr. Amander Clark, Dr. Steve Jacobsen, Dr. Siavash Kurdistani and Dr. Jeff Long. They have been an incredible group and have provided me insightful scientific advice. My committee has been incredibly supportive and guided me to become a mature scientist. I would specially thank Dr. Jeff Long, my thesis advisor, for providing a research environment that permits the search for answers to bold questions at the frontier of science, and for his guidance, expertise and support for my graduate work.

I would like to thank previous and current members in the Long lab. Dr. Naden Krogan launched the work in Chapter 4 and taught me the techniques of protein purification, with which I established the foundation of the work in Chapter 5. Dr. Jared Sewell, who is the primary author in the manuscript in Chapter 3, offered invaluable advice and scientific discussions throughout the four years we overlapped. My undergraduates Nik, Vishal, Chandler and Angela assisted me with greenhouse work, genetic screen and basic molecular experiments. I would like to thank members of Dr. Steve Jacobsen's lab as well, Dr. Sylvain Bischof for scientific discussion and collaboration in performing long and tedious experiments, Linda Yen for advice and assistance in biochemistry. I also would like to thank Chengyang Wang from Dr. Yi Xing's lab for advice and assistance in bioinformatics, Dr. Anna Vitlin as well as Andrew Rashoff from our lab and Dr. Siobhan Braybrook for scientific discussions.

I would also like to thank my family and friends. I would like to thank my parents, my girlfriend Tan Nguyen and our two caring Chihuahuas for emotional support during the most stressful

times. I would like to thank my lifetime friends, Jing Yan and Fengdan Shen, with whose companion and support I survived the most difficult times in my life. And last, I would like to thank all my friends at UCLA who made life in LA fun.

Chapter 3 contains material from a manuscript submitted to *Nature*. The dissertation author will be the co-first author of this manuscript. Chapter 4 and 5 contain unpublished materials.

I was supported by Chinese Scholarship Council fellowship, Whitcome Fellowship and the Dissertation Year Fellowship at UCLA, in chronological order. And I would like to thank the Academic Advisor of MBI, Ashley TerHorst, for all the assistance she provided.

VITA

2008-2012 Bachelor of Science, Zhejiang University

2012-2018 Doctor of Philosophy, University of California, Los Angeles

PUBLICATIONS

Jared A. Sewell, Ao Liu, Zachary R. Smith, Ajay A. Vashisht, Brian Nadel, Sylvain Bischof, James A. Wohlschlegel, Steven E. Jacobsen and Jeff A. Long. Combinatorial and Differential Pathway Regulation by a Family of Transcription Factors. (submitted)

Ao Liu and Jeff A. Long. TPL mediated transcriptional repression: Interplay with Histone Modifications. 2018. Abstract. Plant Molecular Biology Gordon Research Conference - Dynamic Plant Systems.

Ao Liu and Jeff A. Long. Read the Code: The Involvement of Histone Modifications in Transcriptional Repression by TOPLESS. 2016. Abstract. Keystone Symposia-Plant Epigenetics: From Genotype to Phenotype.

Hale CJ, Potok ME, Lopes J, Do T, Liu A, Gallego-Bartolome J, Michaels SD, Jacobsen SE. Identification of Multiple Proteins Coupling Transcriptional Gene Silencing to Genome Stability in Arabidopsis thaliana. *PLoS Genet*. 2016 Jun 2;12(6): e1006092

Chapter 1: Introduction

1.1 TPL, a Groucho/Tup1 like corepressor in *Arabidopsis*

In multicellular organisms, precise transcriptional regulation is crucial for proper growth and development. The activation and repression of gene expression are achieved by the actions of diverse factors including co-repressors. Transcription co-repressors lack the ability to bind directly to DNA and are recruited by transcription factors via intrinsic repression domains. The co-repressor Groucho(Gro) was initially identified in *Drosophila*, where mutations in the gene would result in increased number of bristles over the eyes, resembling the bushy eyebrows of Groucho Marx[1]. Homologs of Gro exist in almost all metazoan and play a vital role in development throughout the life span of animals[2]. Co-repressors share homology to Gro are also present in fungi and plants, for example, Tup1 in Saccharomyces cerevisiae[3,4]. Gro/Tup1 co-repressors have similar structural organization, with an amino-terminal(N-terminal) domain important for self-association as well as binding to specific transcription factors, and a set of peptide repeats each with approximately 40 amino acid terminating with tryptophan (W) and aspartate (D) (WD40 repeats) at the carboxyl-terminus(C-terminus), which mediates the physically association with most of the interacting transcription factors[4,5].

Gro/Tup1 co-repressors are also characterized by a conserved mechanism for transcriptional repression, with histone de-acetylation being an important component. Gro/Tup1 co-repressors are able to interact with the class I histone deacetylase Rpd3/HDAC1, which leads to the deacetylation of lysine residues in histones H3 and H4 and hence the repression of transcription[6,7]. Additionally, the N-terminal domain of Gro/Tup1 has been reported to bind to all

four core histones, with highest affinity when the histone tails are hypo-acetylated[4,8-10]. It was long believed that the self-association of Gro/Tup1 and the ability to bind to histone tails may result in the recruitment of more Gro/Tup1 protein and contribute to the spreading of the repressive state of chromatin. Indeed, it has been shown that Gro/Tup1 localize to genomic regions beyond some of the loci where Gro/Tup1 are initially recruited by the repressors[7,8,11,12]. However, recent study of Gro/Tup1 corepressors using chromatin immunoprecipitation followed by high throughput sequencing analysis (ChIP-seq) suggests the "spreading" of the protein along the chromosome happens only at a limited number of loci[13-15]. With this evidence, a possible multi-step model emerges: Gro/Tup1 co-repressors are first recruited by DNA-bound transcription factors; then they recruit histone deacetylases to catalyze the de-acetylation of local histones, leading to transcriptional repression; at some targets, the ability of Gro/Tup1 to interact with itself and histones may enable the recruitment of additional co-repressors, which could further contribute to the establishment of a repressive chromatin environment. However, transcriptional repression by Gro/Tup1 is only partially dependent on histone deacetylases. It has been shown that Gro/Tup1 can also interact with the core transcriptional machinery[16-19], and that Gro can mask the activation domains of transcriptional activators to reduce the transactivation potential[20].

In plants, Gro/Tup1 co-repressors constitute a small family that can be divided into 2 subclasses represented by TPL/TPL-RELATED 1- 4 (TPR1- 4) and LEUNIG(LUG)/LEUNIG-HOMOLOG (LUH). Both LUG/LUH and TPL/TPRs have a dimerization motif at the N terminus, the Lissencephaly Homology (LisH) domain (Fig 1.1). TPL/TPRs have an additional protein-protein interaction domain C-terminal to LisH domain (CTLH), while LUG and LUH do not have a

CTLH domain but contain a conserved LUFS (LUG, LUH, Flo8, SSDP) domain that is required for transcription repression and for direct interaction with SEUSS (SEU), an adaptor protein that physically links LUG/LUH to a range of transcription factors[21-25]. The N-terminal domains (LisH and CTLH) of TPL/TPRs are able to interact with transcription factors harboring the Ethylene-responsive element binding factor-associated Amphiphilic Repression (EAR) motif. TPL/TPRs also differ from LUG/LUHs in that the former encode for two sets of predicted WD40 domains: one set of five repeats at the C terminus, as well as an additional set of six repeats centrally located in the protein, whereas the latter have only one set of seven repeats at their C termini (Fig 1.1). In *Arabidopsis*, both TPL/TPRs and LUG/LUH can form repressor complexes with histone deacetylase HDA19[26,27]. LUG and TPL have also been shown to interact with Mediator, a multiprotein complex that passes on integrated information from transcriptional factors to RNA polymerase II (RNAPII)[26,28].

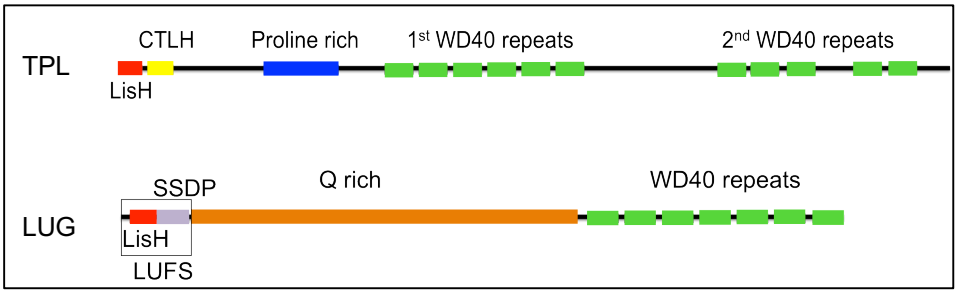


Fig 1.1 Domain organization of TPL and LUG proteins

Both TPL/TPRs and LUG/LUH are important for multiple major developmental processes. LUG was the first Gro/Tup1-like co-repressor identified in plants. In *lug* mutant, there is a sepal-to-carpel transformation, as well as a reduced number of petals and stamens due to ectopic expression of the flower homeotic gene *AGAMOUS* (*AG*)[29]. In addition to flower development, it has been shown that LUG and LUH play a role in leaf development, embryonic shoot apical meristem (SAM) initiation, postembryonic SAM maintenance, gynoecium development, seed

germination, as well as stress response[25,29-32]. The function of TPL/TPRs was first characterized in a temperature-sensitive mutant *tpl-1*. The *tpl-1* mutation acts as dominant negative against TPL and all TPRs. At the restrict temperature of 29˚C, *tpl-1* mutants display ectopic expression of the root-promoting *PLETHORA* genes and exhibit a homeotic transformation of the apical pole into a second basal pole[33,34]. Besides apical-basal patterning, TPL/TPRs have also been shown to be important for a range of biological processes including meristem organization[35-38], floral development and gametophyte growth[27,39], circadian rhythm[40], leaf development[41,42] and immune responses[43].

Plant hormones are small compounds that orchestrate plant growth, development and environmental adaptation. Extensive studies have been made on the role TPL/TPRs play in integrating hormone signaling to regulate transcription. The first line of evidence was that the founding member, TPL, forms a co-repressor complex with AUXIN RESPONSE FACTOR 5 [ARF5, also known as MONOPTEROS (MP)] in the presence of EAR-motif-containing INDOLE-3-ACETIC ACID INDUCIBLE 12 [IAA12, also known as BODENLOS (BDL)], which blocks the activation of ARF5 targets[44]. Auxin, one of the major hormones involved in practically all cellular processes, induces the degradation of IAA12, freeing ARF5 and hence activating the transcription of auxin response genes. It also has been shown that TPL bridges IAA14 and Mediator to regulate transcription of the targets of ARFs in an auxin-dependent way[28]. Jasmonic Acid(JA) signaling pathway utilizes TPL in the same manner, where TPL physically interacts with the jasmonate ZIM-domain (JAZ) repressor-interacting protein, NINJA, to negatively regulate jasmonate responses[45]. Similarly, TPR2 has been shown to interact with SUPPRESSOR OF MORE AXILLARY GROWTH2-LIKE proteins (SMXLs), which repress

Strigolactone (SL) response genes and are subject to degradation upon SL induction[46]. Brassinosteroids (BRs) are known to counteract Abscisic Acid (ABA) in seed germination and young seedling development. BR induces Brassinazole Resistant 1 (BZR1) and BRI1-EMS-SUPRESSOR 1 (BES1), which recruit TPL to repress the expression of ABA regulators in early seedling development[47,48]. BES1 also interacts with TPL to regulate boundary formation in shoots and meristem development in roots[36]. TPR4, another member in the TPL/TPR family, is involved in the signaling pathway of a canonical hormone named Gibberellic Acid(GA)[49].

The ability of TPL/TPRs to interact with the above regulators resides in its N terminal domains. In two recently published paper on the crystal structure of the N terminus of TPL, it was shown that the N termini form two distinct regions proving surface for tetramerization and for interacting with EAR-motif-containing repressors[50,51]. While the function of the N terminal domains and the structural basis have been revealed, little is known about the C terminal WD40 repeats in TPL/TPRs. Moreover, TPL is ubiquitously expressed in every cell throughout the life of *Arabidopsis*, while the phenotypes of the dominant negative allele *tpl-1* and the *tpl/tpr* quadruple loss-of-function mutant indicate cell or tissue specific functions of TPL. Therefore, what provides the functional specificity of TPL remains to be answered.

1.2 Homeodomain-leucine Zipper III (HD-ZIP III) transcription factors

In a search for suppressors of *tpl-1* double root phenotype, dominant mutations in the HD-ZIP Class III transcription factor family were identified[34]. The HD-ZIP class of proteins are transcription factors unique to plants. HD-ZIPs share a conserved N terminal DNA binding and

dimerization domain characterized by a homeodomain (HD) closely followed by a leucine zipper motif (ZIP)[52]. In *Arabidopsis*, there are 48 HD-ZIP proteins, which can be divided into four families, HD-ZIP I to IV, based on conservation within the HD-ZIP domain and the presence of additional motifs[53-57]. Among the four families, the HD-ZIP III protein family is the smallest yet plays a pivotal role in development. The HD-ZIP III family consists of only five members: ATHB8, CORONA (CNA) [also known as INCURVATA4 (ICU4)], PHABULOSA (PHB), PHAVOLUTA (PHV) and REVOLUTA (REV). These HD-ZIP III proteins also contain a steroidogenic acute regulatory protein-related lipid transfer domain (START), followed by MEHKLA/PAS domain[58,59]. While in animals, START-domain proteins have been shown to bind to lipid ligands, the function of START domain in plants is still under investigation. At the post-transcriptional level, HD-ZIPIIIs are under the control of microRNAs (miR) 165/166 by recognizing sequences in the START domain[60-63]. MEKHLA/PAS domains act as sensors that are able to perceive redox potentials[64]. It has been shown that the MEKHLA domain of REV acts as a negative regulatory domain by inhibiting dimerization of REV[65]. MEKHLA domains are reported to be required for HD-ZIP IIIs to interact with the APETALA2-like (AP2-like) transcription factors DORNROSCHEN (DRN) and DORNROSCHEN-like (DRNL)[66]. Interestingly, a missense mutation in the MEKHLA domain of CNA enables the plant to auto-regenerate shoots without the application of exogenous hormones[67].

HD-ZIP IIIs are important regulators of both embryonic and post-embryonic patterning. The five of them play partially overlapping roles in plant development, with REV being the only one that displays a phenotype in the single loss-of-function mutant. In a *rev* mutant, the axilliary meristems and floral meristems constantly fail to complete normal development and the cells in

the leaves undergo extra divisions[68]. Double mutants of *phb/phv* show no obvious phenotypes, while the triple mutant of *phb/phv/rev* or *phb/rev/icu4* displayed a single ventralized radial cotyledon that lacks an apical meristem[69,70]. Interestingly, the function of the HD-ZIP III proteins is not always redundant but also antagonistic. Mutations in *ICU4* and *ATHB8* (*icu4/athb8/rev*) can partially suppress the *rev* mutant phenotype and *phb/phv/icu4* mutants display an increase in meristem size[70]. In contrast to the loss-of-function alleles, all dominant gain-of-function (gof) alleles contianing mutations in the miRNA binding sequences, exhibit developmental defects: Dominant *phb* and *phv* mutants displayed adaxialized cotyledons and leaves[61], and in some alleles also severely adaxialized petals[34]; dominant *rev* mutants lose adaxial-abaxial polarity in vascular bundles, although the leaves appear normal[69]; *icu* mutants exhibit dorsalized leaves and all the gof alleles identified in the suppressor screen can rescue the double-root phenotype when in combination with *tpl-1*[34].

In *Arabidopsis* embryos, HD-ZIP III proteins are master regulators of apical fate[34]. During embryogenesis, *HD-ZIP III* mRNA is restricted to the apical half, and later accumulates in the SAM, the adaxial region of the developing cotyledons and the provascular cells[61,69-71]. During the post-embryonic growth phase, the expression of *HD-ZIP III*s continues to be tightly regulated. We observed similar protein accumulation patterns for HD-ZIP IIIs in the SAM and in the adaxial side of leaf primordia. The genetic interaction between *tpl-1* and gof *HD-ZIP III* allels suggests a possible link in the function of TPL and HD-ZIP IIIs. In that case, the cell or tissue specific expression of *HD-ZIP III*s might result in functional specificity of TPL with a corresponding pattern.

1.3 Histone acetylation and deacetylation

In eukaryotes, the genome is packed into a well-organized structure called chromatin, which allows the genomic DNA of meters long to fit into the microscopic space of the nucleus. The fundamental unit, the nucleosome, consists of a core with roughly 146 base pair(bp) of DNA wrapped around a histone octamer (2 copies of histones H2A, H2B, H3, and H4), and a linker with up to 80 bp DNA bound by the linker histone H1. The N terminal tails of histones protrude from the core and these highly basic regions are able to interact with the acidic regions of its neighboring nucleosomes[72] and the linker DNA[73]. The amino acids within the N-terminal tails can be covalently modified. These modifications can either change the interaction of histone tails with DNA and with adjacent nucleosomes or can result in the recruitment of chromatin remodelers and chromatin readers, and therefore affect gene expression.

Histone acetylation is one of the most well studied modifications. Acetylation takes place on the $NH3^+$ group of Lysine residues at multiple sites of the histone tails. It is a reversible process involving the antagonistic activity of histone acetyltransferases(HAT) and deacetylases(HDAC). Based on sequence characterization, these modifiers can be divided into 4 families of HATs: the GNAT, MYST, CBP and $TAF_{II}250$ superfamilies; and 3 families of HDACs: the RPD3, SIR2 and the plant-specific HD-tuins superfamilies[74]. In *Arabidopsis*, 12 HATs and 18 HDACs have been identified (summarized in table 1.1). Acetylation decreases the interaction between histone tails and the negatively charged DNA by removing the positive charge on the histones. As a consequence, acetylation decondenses the chromatin, thereby positively regulating transcription.

The reversible action of histone acetylation and de-acetylation plays an important role in transcriptional regulation in response to developmental and environmental cues (table 1.1).

Family	Gene name	Function
RPD3/HDA1 Class I	HDA19	Seed development and dormancy, flower development, leaf morphogenesis, light signaling and hypocotyl growth, ethylene, JA, ABA, SA pathway and defense response[27,33,47,74-83]
	HDA7	Seed germination, plant growth, female gametophyte and embryo development[84]
	HDA6	Seed development and germination, ABA and salt stress response, JA signaling, cold stress response and ethylene pathway, light signaling, leaf morphogenesis, Circadian regulation, flowering control and DNA methylation[78,79 45,85-95]
	HDA9	Flowering time[96]
Class II	HDA5	Flowering time[97]
	HDA15	Phytochrome pathway and hypocotyl growth[98]
	HDA18	Root hair development[99]
Class III unclassified	HDA2	–
	HDA8	–
	HDA10	–
	HDA14	–
	HDA17	–
HD-tuins	HDT1	Leaf morphogenesis[100]
	HDT2	Leaf morphogenesis and seed dormancy[100,101]
	HDT3	ABA pathway, salt stress response and seed germination[85,102]
	HDT4	–
Sirtuin	SRT1	Ethylene-induced transcriptional repression[103]
	SRT2	Defense to pathogen, mitochondrial function and Ethylene-induced transcriptional repression[103-105]
GNAT	HAG1	Root, shoot and flower development, light signaling, low temperature response [33,76,106,107]
	HAG2	–

	HAG3	Cell proliferation[108]
MYST	HAM1	Flowering time and gamete formation[109,110]
	HAM2	Flowering time and gamete formation[109,110]
CBP	HAC1	Sugar response and flowering time[111,112]
	HAC2	–
	HAC4	–
	HAC5	–
	HAC12	–
$TAF_{II}250$	HAF1	–
	HAF2	Light signaling[113]

Table 1. 1 a summary of HATs and HDACs in Arabidopsis

Histone acetylation is involved in many developmental processes with TPL in *Arabidopsis*. Mutations in *HAG1,* a member of the GNAT family, partially suppress the phenotype of *tpl-1* while mutations in *HDA19,* a member of the RPD3 family, enhance the *tpl-1* phenotype, suggesting histone acetylation and deacetylation affect the function of TPL during embryogenesis[33]. It was further confirmed that TPL can form a repressor complex with the transcription factor AP2 and with HDA19 and this co-repressor complex is involved in repressing the expression of a floral homeotic gene *AG*[27]. It also has been shown that during early seedling development the TPL-HDA19 complex directly facilitates the histone deacetylation of *ABA INSENSITIVE 3* (*ABI3*) chromatin, leading to the suppress of ABA signaling output[47]. While histone acetylation has been extensively studied, limited information is known on the direct targets of histone acetylases and deacetylases in plants. The development of next-generation sequencing based tools enable the precise mapping of genomic sites bound by

proteins. Although in *Arabidopsis*, the first binding profile of Rpd3 histone deacetylases was recently published on HDA9[114], high quality data are still needed to decipher the process of deacetylation and subsequent transcriptional repression.

1.4 H3K27me3 and Polycomb Repressive Complex in *Arabidopsis*

While acetylation usually correlates with active transcription, histone methylation can be associated with either transcriptional repression or activation, depending on the location of the residue and the degree of methylation. Histone methylation in general occurs on the lysine (K) and arginine (R) residues. Among the various forms of histone methylation, histone H3 lysine 27 trimethylation (H3K27me3) is considered to turn off the expression of genes and maintain a silenced state in the cells where the genes are supposed to be repressed. In both animals and plants, H3K27me3 plays a pivotal role in the establishment and maintenance of cell identity and is crucial for multiple biological processes including imprinting[115-117].

H3K27me3 is dynamically deposited by SET domain-containing Polycomb group (PcG) proteins and removed by histone demethylases throughout development. PcG proteins were first identified in *Drosophila* as negative regulators of homeotic *Hox* genes, which are required for segmentation during embryo development[118]. Polycomb Repressive Complex 1 (PRC1) and PRC2 are two of the most well-characterized PcG protein complexes. PRC2 catalyzes the trimethylation of K27 while PRC1 recognizes H3K27me3 and subsequently ubiquitylates K119 of Histone H2A to further condense the chromatin. In *Drosophila*, PRC2 consists of 4 main components: the SET domain-containing Enhancer of zeste (E(z)) with methyltransferase activity, Extra sex comb (Esc) and Suppressor of zeste 12 (Su (z)12) that are both required for

E(z) function, and a non-essential nucleosome remodeling factor Nurf55 or p55[119]. PRC1 is also comprised of four components in *Drosophila*: the chromodomain-containing Polycomb (Pc) with the ability to recognize H3K27me3, the RING domain-containing dRING/Sex combs extra (Sce) and the zinc finger-containing Posterior sex combs (Psc) that monoubiquitylate histone H2A, and another zinc finger-containing protein Polyhomeotic (Ph)[120]. PcG proteins are evolutionarily conserved throughout eukaryotes. In *Arabidopsis*, there is one Esc homologue [FERTILIZATION INDEPENDENT ENDOSPERM (FIE)], three homologues of E(z) [the partially redundant CURLY LEAF(CLF) and SWINGER(SWN) in the sporophyte, and MEADEA (MEA) in the gametophyte], three Su(z)12 homologues [EMBRYONIC FLOWER2 (EMF2), VERNALIZATION2 (VRN2), and FERTILIZATION INDEPENDENT SEED2 (FIS2)], and five homologues of p55 [MULTICOPY SUPRESSOR OF IRA 1–5 (MSI1–5)]. Both FIE and MSI1 are indispensable components of *Arabidopsis* PRC2 complexes. As for the interchangeable components, with the presence of CLF or SWN, VRN2 and EMF2 are required for vernalization and vegetative development, respectively, whereas with MEA, FIS regulates female gametophyte and seed development[121]. PRC1 components in *Arabidopsis* are less conserved, with no clear homologues of Pc or Ph. There are two homologs of *Drosophila* dRING - AtRING1A and AtRING1B[122], and three homologs of Psc - AtBMI1A, AtBMI1B and AtBMI1C[122,123]. The five of them have been shown to monoubiquitylate K119 of H2A[122,124]. Like Heterochromatin Protein 1/Terminal Flower 2 (LHP1/TFL2), which recognizes and binds to H3K27me3, is a proposed Pc-like component of PRC1[125,126]. In addition, *Arabidopsis* PRC1 also has some plant-specific components - EMF1, VRN1 and VP1/ABI3-Like 1/2/3 (VAL1/2/3) – that may contribute to the recruitment of PcG complexes[127-129].

To ensure precise regulation of genes at different times in different cells, the recruitment of PRC1 and PRC2 needs to be tightly controlled. The paradigm for targeting PcG complexes originated from research in *Drosophila*, where PRCs are recruited to Polycomb response elements (PREs). PREs are sequences in promoters that harbor consensus binding motifs for multiple transcription factors, including the GAGA factor (GAF) and Pho[130]. In mammals, although the core PcG complexes are highly conserved from *Drosophila*, the TFs that bind *Drosophila* PREs are mostly absent, with one exception being the zinc finger protein PHO/YY1. This suggests PRE-directed recruitment might not be common among all eukaryotes. Indeed, only two elements that have similarities to *Drosophila* PREs, D11.12 from the human *HOX* cluster and PRE-kr, have been identified[131,132]. It appears that unmethylated CpG islands and Long non-coding RNAs (lncRNAs) play a more widespread role than PREs for genome-wide targeting of PcG in mammals[133-137]. In *Arabidopsis*, the GAGA repeat, a cis motif also found in *Drosophila* PREs, has been identified as a possible component of plant PREs due to its enrichment in the FIE binding sites[138]. However, these elements are located within the gene body, leading to the question whether they may serve as recruiting elements or to strengthen the interaction of PRC2 at target sites. It also has been shown that genes with conserved noncoding sequences (CNSs), which are enriched in transcription factor binding sites, are significantly overrepresented among the genes with H3K27me3, suggesting a possible role of CNSs in recruiting PcGs[139]. lncRNAs are involved in the recruitment of PcGs in plant as well. During vernalization, CLF binds to the lncRNA COLD ASSISTED INTRONIC NONCODING RNA (COLDAIR) to trigger the tri-methylated of H3K27 at the floral repressor gene, FLOWERING LOCUS C (FLC), and repress its expression[140]. LHP1 binds to the lncRNA AUXIN REGULATED PROMOTER LOOP RNA (APOLO) to regulate the expression of PINOID

(PID)[141]. Although more and more studies on the recruitment of PcGs and on the nucleation as well as spreading of H3K27me3 have been carried out in *Arabidopsis*, it is still unclear how PRC1s and PRC2s crosstalk and integrate information from the above-mentioned machineries to shape the epigenetic landscape during development.

Abbreviations in this chapter:

GNAT: GCN5-related N-terminal acetyltransferases
MYST: MYST is from its four founding members: human MOZ (monocytic leukemia zinc finger protein), yeast Ybf2 (renamed Sas3, for something about silencing 3), yeast Sas2 (84) and mammalian TIP60 (HIV Tat-interacting 60 kDa protein)
CBP: CREB-binding protein
$TAF_{II}250$: TATA-binding protein-associated factor
RPD3: Reduced Potassium Deficiency 3
Sir2: Silent Information Regulator 2

Chapter 2: Transcriptional regulation by TPL and HD-ZIP III transcription factors in *Arabidopsis* development

Summary

Chapter 2 contains the contents of the manuscript titled "Combinatorial and Differential Pathway Regulation by a Family of Transcription Factors". For this manuscript, I performed bioinformatics analysis, RNA-seq experiments with the transgenic lines to confirm the rescue by the transgenes at molecular level, co-immunoprecipitations (co-IP) and sequential ChIPs to show the physical association in planta, and repeated REV and ATHB8 ChIP-seq experiments. In this work, TPL was shown to form complexes with the members of the HD-ZIP III transcription factor family. The five members of the family form heterodimers and bind in different combinations to the promoters of thousands of genes, acting both as activators and repressors. HD-ZIPIIIs proteins accumulate with a highly tissue-specific pattern and are involved in regulating multiple steps of different biological pathways both antagonistically and cooperatively. TPL and HD-ZIP IIIs simultaneously localize at the genomic regions that have overlapping peaks. The genes with both TPL and HD-ZIP IIIs peaks in the promoters showed significantly lower expression levels. There results suggest HD-ZIP IIIs forms complexes with TPL to carry out repressor function.

Reproduction of the Manuscript

Combinatorial and differential pathway regulation by a family of transcription factors

Abstract: **Genome duplication is a common occurrence in eukaryotes and can lead to large gene families that have both redundant and divergent functions[142,143]. In *Arabidopsis,* the five-member Class III Homedomain-Leucine Zipper (HD-ZIPIII) TF family controls key developmental processes, such as shoot stem cell formation, and displays complex genetic interactions[70]. The specific molecular mechanisms behind these genetic observations remain unresolved. Here, we show that HD-ZIPIIIs collaboratively and antagonistically regulate entire pathways at multiple, sometimes opposing steps, which can occur at a cell type, tissue or stage specific level. We found that HD-ZIPIIIs form preferential heterodimers *in planta*, co-occupy hundreds of genomic locations and collaboratively regulate target gene expression both positively and negatively. Transcriptional repression by HD-ZIPIIIs is correlated to co-occupancy and association with the co-repressor TOPLESS. Our results demonstrate how a highly-related family of TFs can collectively regulate important signaling networks to stabilize or fine-tune gene expression through cooperative target binding.**

Genetic studies have shown that HD-ZIPIII family members (REVOLUTA/REV, PHABULOSA/PHB, PHAVOLUTA/PHV, INCURVATA4/ICU4 and ARABIDOPSIS THALIANA HOMEOBOX 8/ATHB8) display both redundant and antagonistic functions, and play pivotal roles in specifying and maintaining the shoot apical meristem (SAM) as well as patterning the adaxial side of newly formed organs[61,69]. Despite the critical function of these TFs during development, there is limited information regarding their native binding sites[144,145]. To identify genome-wide targets of all five HD-ZIPIIIs, we used recombineering[146] to generate endogenously expressed fusion proteins (e.g. REVg:2xYPet-3xHA) (Fig. 1a and Extended Data

Fig. 1a-o). These constructs were capable of rescuing developmental and molecular phenotypes in their respective mutant backgrounds (Extended Data Fig. 1p-q and Extended Data Fig. 2a-e). Therefore, we proceeded to use these lines to perform chromatin immunoprecipitation (ChIP) followed by deep sequencing (ChIP-seq). Thousands of binding sites were detected for each family member with as few as 2,790 for PHB and as many as 10,383 for ICU4 (Fig. 1B, Table 1, and Extended Data tables 1 and 2).

We found that known REV direct targets, *LITTLE ZIPPER 1* and *4 (ZPR1, ZPR4*) and the Class II HD-ZIP members *HAT2* and *ARABIDOPSIS THALIANA HOMEOBOX-LEUCINE ZIPPER PROTEIN 4* (*ATHB4*) (Fig.1c) were bound by all five HD-ZIPIII members at identical locations, validating our approach (Extended Data Fig. 3b)[144,147]. When we compared the overlap of peaks between 4 family members (*REV*, *PHB*, *PHV*, *ICU4*) or all five family members we found it to be relatively low (683 and 243 respectively) (Fig. 1b, Extended Data Fig. 3c, Extended Data Fig. 4, and Extended Data table 3). Analysis of HD-ZIPIII binding profiles with publicly available ChIP-seq data[148-159] revealed that the five members significantly overlap with each other and the co-occupancy is specific to the HD-ZIPIII family (Extended Data Fig. 4). Moreover, we found that ICU4 and ATHB8 displayed a significant number of unique targets (Fig. 1b, and Extended data Fig. 5a,b). These results are in agreement with previous genetic studies[70] and expression profiles for their respective loss of function mutants (Extended Data Fig. 5f), which indicate diverging functional roles for these family members. Together, these data suggest that a majority of HD-ZIPIII binding sites are unique to the family and not a consequence of binding highly occupied target regions. Further these results indicate that there is a small core group of shared candidate target genes, as well as target genes either unique to individual family member or

specific to different HD-ZIPIII combinations. We next compared ChIP-seq read abundance among HD-ZIPIIIs and found distinct combinations of enrichment between family members, suggesting preferential co-regulation of specific target genes (Fig. 1d, e, and Extended Data Fig. 3e-l). Strikingly, when we compared the basal expression levels of genes bound by one or multiple family members, we found a cooperative effect on gene expression levels based on the number of factors bound (Fig. 1f, and Extended Data Fig. 5g-k). However, when we performed generalized linear regression on the binding effects of each HD-ZIPIII member on basal expression levels, the individual binding effects on transcriptional levels differed as REV and ATHB8 displayed negative effects on basal expression levels (Fig. 1f, and Extended Data Fig. 5g-k). These results suggest that HD-ZIPIII family members can collaboratively fine-tune transcription levels contingent on the number of factors bound and are consistent with previous genetic studies, which indicate both shared and unique functions within this gene family[70,160].

Among all targets, we observed enriched Gene Ontology (GO) categories for genes involved in post-embryonic development, meristem maintenance, response to far red light and hormone-mediated signaling pathways (Extended Data Fig. 3n), which fit well with the characterized developmental roles for these TF's[69,144,161-163]. We then performed a *de novo* cis-regulatory motif discovery analysis on bound regulatory regions and found enrichments of the canonical HD-ZIP binding motif AAT(C/G)ATT[144,164]. We also identified enrichment of the G-box motif, CACGTG, which is typically associated with basic Helix-Loop-Helix (bHLH) proteins[165] (Extended Data Fig. 1m). Our data indicate that HD-ZIPIII family members co-occupy many regulatory regions (Fig. 1b-e). Although studies have shown that HD-ZIPIIIs are capable of forming dimers *in vitro*[65,164], there are no reports of these interactions *in vivo*. To determine

whether dimerization could explain co-occupancy near target genes, we immunoprecipitated REV, PHB and ICU4 followed by mass spectrometry (IP-MS). Within our IP-MS data, we found peptides unique to every HD-ZIPIII member (Table 2). We validated the IP-MS results by performing co-immunoprecipitation assays *in planta* with REV (Fig. 1g). Our IP-MS data also revealed enriched interaction ratios between certain family members (Table 2). For example, ICU4 consistently pulled down more PHB specific peptides than those of other members. This is in agreement with our ChIP-seq analyses that showed PHB is more highly enriched at ICU4 bound regions than other HD-ZIPIIIs (Fig. 1d and e). This may also suggest the formation of higher order protein complexes necessary for collaborative transcriptional regulation. Taken together, our results suggest this TF family works together at target loci in specific, preferential combinations.

To understand if DNA binding events lead to a transcriptional response, we correlated our ChIP-seq data with a transcriptional over-expression profile for each HD-ZIPIII family member. We created inducible, microRNA (miR) resistant versions of each HD-ZIPIII member fused to the glucocorticoid receptor (ex. *p35s:PHB*miR-GR*) (Extended Data Fig. 6f-n)[166]. To capture dynamic changes in gene expression we induced each fusion protein for 15 minutes, 1 hour or 4 hours. Following induction, RNA-seq analyses revealed a unique molecular signature for each HD-ZIPIII (Fig. 1h, and Extended Data Fig. 6o). Comparisons of our RNA-seq and ChIP-seq data sets, yielded approximately 400-800 direct target genes for each HD-ZIPIII member (Extended Data Fig. 6p, Table 1 and Extended Data table 5). Notably, genes neighboring regulatory regions co-occupied by multiple family members were more sensitive to HD-ZIPIII induction at earlier time points (Fig. 1i) and diminished over the time course treatments. These

results support the notion that HD-ZIPIIIs cooperatively control gene expression by co-occupying regulatory regions. This mechanism may be important for the immediate induction of genes and for fine-tuning expression levels to a steady state. We found these putative direct targets to be both up and down regulated in response to HD-ZIPIII induction and sometimes in opposing manners among individual family members (Fig.1h). These data indicate that all HD-ZIPIII members act as both activators and repressors of gene expression. To investigate the biological relevance of HD-ZIPIIIs acting as both activators and repressors, we fused a tetramer of the transactivation domain from the herpes simplex virus (VP64) to endogenously expressed REV, PHV and ICU4. These fusion proteins should now act as strong activators[167,168]. Compared to wild-type (WT), *REV/PHV/ICU4-VP64* transgenic lines displayed severe morphological defects in the SAM (Fig. 2a-d) indicating that a balance between HD-ZIPIII-mediated transcriptional repression and activation is necessary for their role in maintaining the SAM.

HD-ZIPIIIs lack known domains associated with transcriptional repression[169], however our IP-MS showed that REV, PHB and ICU4 associate with transcriptional co-repressors *TOPLESS (TPL*) and *TOPLESS RELATED* proteins *(TPR1, TPR3, TPR4)*[27,33,44], which is confirmed by co-IP (Fig 2e). This suggests these transcriptional co-repressors may be recruited to some HD-ZIPIII target sites in order to provide repressive activity (Table 2, Fig 2e, h and i). TPL has previously been shown to act in concert with the HD-ZIPIIIs during embryonic patterning[34]. To test whether TPL associates with genomic regions bound by HD-ZIPIIIs, we performed ChIP-seq on TPLg:2xYPet-3xHA lines (Extended Data table 6) and compared these data to our HD-ZIPIII ChIP-seq data sets. Among the genomic regions co-occupied by TPL and HD-ZIPIIIs, we observed a modest proportion of overlapping peaks with REV and ATHB8 (29% and 9%,

respectively) while greater than 60% of TPL peaks overlap with PHB, PHV and ICU4 peaks (Extended Data table 7). Consistent with previous genetic studies[34], we observed binding by TPL and HD-ZIPIII members upstream of *PLETHORA1* and *2* (*PLT1, PLT2)* (Extended Data Fig. 7a, b). Furthermore, all pairwise combinations between TPL and HD-ZIPIII bound regions displayed a positive Pearson correlation coefficient (Fig. 2g). When we compared TPL targets with those shared by the HD-ZIPIII family (Fig. 1b), >80% of those genomic regions were co-occupied with TPL (Fig. 2f and Extended Data Fig. 5c). To determine if TPL is simultaneously bound at these loci, we performed sequential ChIP[170] with TPL and HD-ZIPIII members. In all cases tested, fold enrichments were higher in the sequential ChIP assays compared to HD-ZIPIII ChIPs alone (Fig. 2j and Extended data Fig. 7d,e). Notably, genes that contained both upstream HD-ZIPIII binding coupled with TPL occupancy displayed lower levels of steady state RNA expression compared to regions lacking TPL (Fig. 2k). Taken together, these results strongly suggest HD-ZIPIIIs and TPL form protein complexes to repress gene expression.

We were further interested in the pathways that display co-regulation by multiple family members. We found that HD-ZIPIIIs directly bound and/or regulated entire families of genes and entire hormone signaling pathways, including cytokinin, gibberellic acid and auxin, at multiple sometimes opposing, steps (Fig. 3a-c, Extended Data Fig. 8a-d and Extended Data table 8). This observation suggests HD-ZIPIIIs function to integrate expression of multiple genes at several steps within a signaling network. To understand the influence that HD-ZIPIII family members exert on genes important for development, we focused on cytokinin biosynthetic enzymes as they have known roles in patterning tissues[171-174], are regulated by the HD-ZIPIIIs in the root[161,174], and play a critical function in preserving stem cell activity within the SAM[173,175]. Moreover, we

concentrated on REV and ICU4 targets within the SAM as genetic studies have suggested these two family members display both genetic antagonism and redundancy[70]. Our genomics data indicate that *CYTOKININ OXIDASE 3* (*CKX3*) is a negatively regulated target of ICU4, while other family members did not affect its expression (Fig. 3a, c, and Extended Data Fig. 6a). CKX's are enzymes responsible for catalyzing cytokinin degradation and have been shown to control postembryonic meristem size by degrading cytokinin in the Organizing Center (OC) of the SAM[171]. We generated a transcriptional reporter for *CKX3* (*CKX3p::NLS-2xYPet)* and examined accumulation in WT and HD-ZIPIII dominant gain-of-function (g-o-f) mutant meristems. In WT plants, *CKX3p::NLS-2xYPet* strongly accumulates in the OC of the inflorescence meristem yet is completely absent from the L1/epidermis of *icu4-1D* meristems (Fig. 3d, e). In a *rev-10D* g-o-f background, there were no observable changes in reporter expression (Extended Data Fig. 9b and c). These experiments uncovered an additional cell type-specific effect, as repression of *CKX3* was only observed in the L1 layer of the meristem, and not in the underlying cell layers where *ICU4* is also expressed.

Reduced *CKX3* expression in the L1/epidermis may increase cytokinin levels and activity. To better understand the outcome of this cell type-specific regulation, we took advantage of a Two Component Signaling (TCSn::GFP) sensor, which provides a readout of the transcriptional response downstream of cytokinin signaling[176]. We observed TCSn expression within the inflorescence meristem, but it was excluded from the first two cell layers in WT (Fig. 3f). TCSn::GFP signal was expanded into the L1/epidermis of *icu4-1D* meristems (Fig. 3g) suggesting that ICU4 plays a role in increasing cytokinin levels, specifically in the epidermis of meristems. Furthermore, this result is in agreement with recent reports suggesting that

epidermally-derived cytokinin can act as a positional cue to establish patterning domains within the SAM[172,177] and thus, HD-ZIPIII activity can be placed upstream of this regulation.

We next examined the expression of *LONELY GUY 8* (*LOG8*) using a transcriptional reporter (*LOG8p::NLS-2xYPet*) as LOG enzymes positively mediate cytokinin biosynthesis and have been shown to play a pivotal role in maintaining the SAM[173]. Our genomics data showed that REV, but not ICU4, negatively regulates *LOG8* expression (Fig. 3b, c). Similar to *CKX3*, the *LOG8p* reporter accumulates to high levels within the OC of a WT SAM (Fig. 3h). We found that the *LOG8p* reporter was greatly diminished in *rev-10d* (Fig. 3i, and Extended Data Fig. 9e), but not visibly affected in *icu4-1D* mutants. To test the downstream effect of this regulation, we examined TCSn::GFP activity in *rev-10D* and found it was barely detectable within these meristems (Fig. 3j). Taken together, these results are consistent with our genomics data and suggest that REV and ICU4 regulate a hormone response pathway in an opposing manner. This functional opposition at the molecular level uncovers one mechanism for the antagonism between REV and ICU4 observed in genetic studies.

Given the critical roles HD-ZIPIIIs play throughout development[34,61,69,160,163], we next asked whether the genomic data obtained from seedlings were applicable during embryogenesis. We first used publicly available data[178] to assess whether cytokinin-related targets were expressed during embryogenesis. While *LOG8* did not appear to be expressed during this developmental stage, *CYTOKININ OXIDASE 5* (*CKX5*) and *CKX3* were. *CKX5* is positively regulated by REV (Fig. 3a and Extended Data Fig. 9a), and as expected we observed a strong increase in *CKX5* mRNA expression in provascular cells of *rev-10D* embryos compared to WT (Fig. 4a and b).

Consistent with our results from post-embryonic tissue, we observed a decrease in TCSn activity in *rev-10D* g-o-f mutant embryos compared to WT in the same cells that *CKX5* was up-regulated (Fig. 4c and d). These results demonstrate that REV plays an inhibitory role on cytokinin signaling during the course of development through direct transcriptional regulation of enzymes that produce and degrade cytokinin.

We next examined the expression of the *CKX3* reporter. We found that in WT embryos, the *CKX3* reporter is expressed in cells underlying the embryonic SAM (Fig. 4e). Unexpectedly, *CKX3p:NLS-2xYPet* expands down into the provascular cells in *icu4-1D* embryos (Fig. 4f). This result is in contrast to our observation of negative *CKX3* regulation in post-embryonic tissue (Fig. 3e), which suggests that target gene regulation can change during the course of development. We hypothesized that an increase and expansion of *CKX3* would lead to a reduction in TCS activity in these cells. In agreement with this, *icu4-1D* embryos displayed much lower expression of TCSn compared to WT (Fig. 4g). Together, these results highlight the unique and specific regulatory roles that individual HD-ZIPIII members have on gene expression in both space and developmental time. Moreover, these data demonstrate how these TFs can redundantly regulate hormone response by acting on different genes with similar functions.

Cooperativity between TFs is an emerging trend in gene regulation and has been observed for both plants[149] and animals[179]. Our findings provide new insight into the molecular basis of previous genetic observations[3] which appear to be the result of both cooperative and antagonistic regulation of individual genes, as well as higher-level network regulation. Further, these results portray the importance of studying interfamily functional relationships by showing that

cooperativity coupled with the ability to dimerize with co-repressors allow for gene expression to be fine tuned (Fig. 4h). Our data also highlight the importance of TFs families in controlling the gene expression of enzyme encoding genes. Recent studies have suggested that this type of regulation is necessary for integrating environmental and developmental cues in order to precisely stabilize metabolic networks in both plants and animals[35,36]. In summary, these results will help to unravel the inherent complexities of TF families in the regulation of developmental networks at a cell type, tissue and organ system level.

Acknowledgements: We thank the Müller laboratory for the TCSn::GFP plasmid and the Kay laboratory for the VP64 construct; Dr. DR Kelley, Dr. JA Osterhout and Dr. JM Van Norman for comments on the manuscript; J. DaDude and J. Meister for the support. S.B. is supported by a postdoctoral fellowship of the Swiss National Science Foundation. S. E. J. is an Investigator of the Howard Hughes Medical Institute. A.A.V. and J.A.W. were supported by NIH grants GM112763 and GM089778. A. L is supported by the Chinese Scholarship Council and Whitcome Fellowship. J.A.S., A.L. and J.A.L. were supported by NIH grant R01GM072764.

Author Contributions: J.A.S. and J.A.L. designed the experiments. J.A.S. generated and analyzed reporters. J.A.S and A.L. performed ChIP-seq, RNA-seq and co-IP experiments. A.L. and B.N. performed computational analyses. S.B., A.A.V., J.A.W. performed IP-MS experiments and analyses. Z.R.S. generated 35S::GR constructs. J.A.S. and J.A.L. wrote the manuscript and S.B and A.L. revised it. All authors read and approved the manuscript.

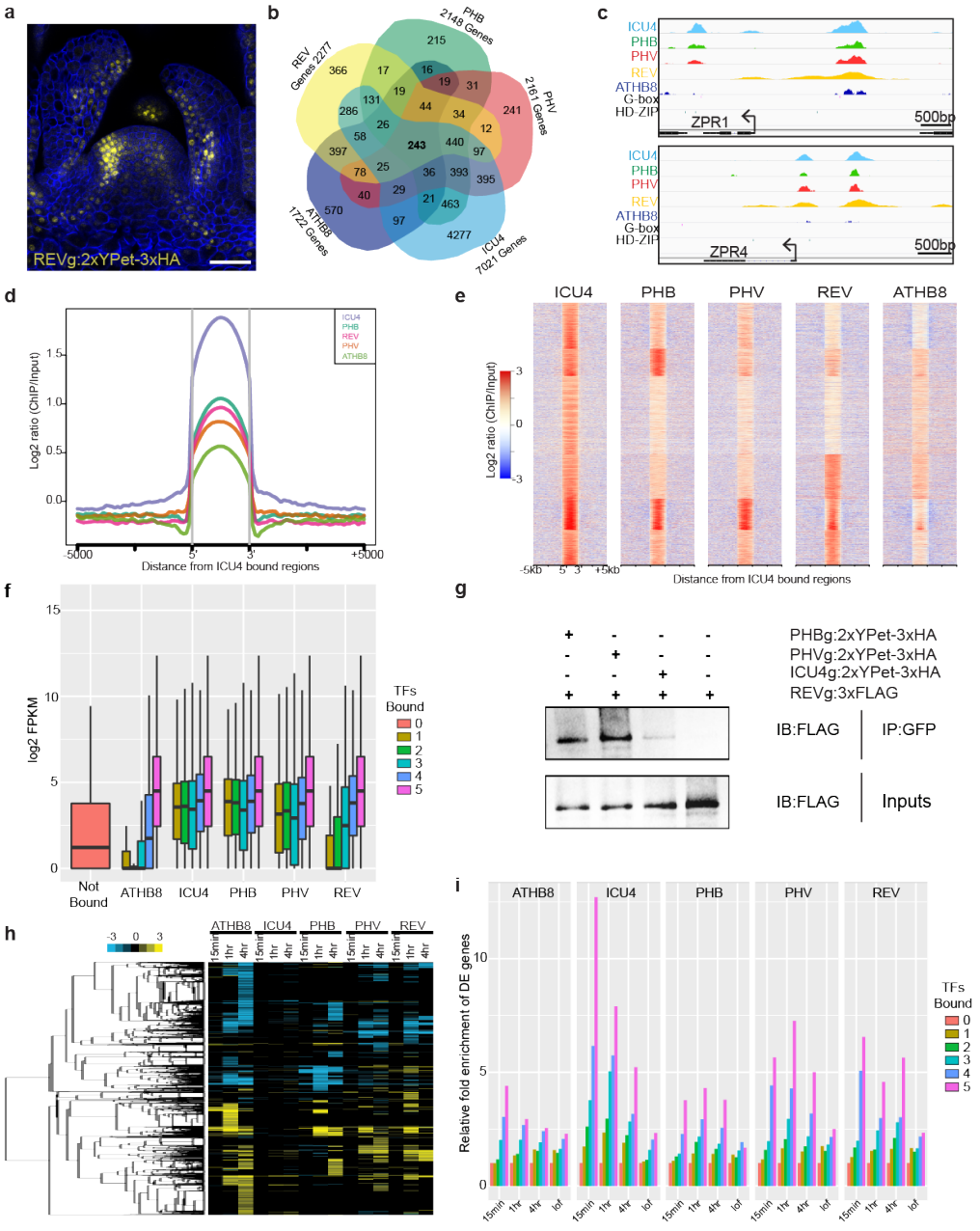


Figure 1 | Genome-wide analysis of HD-ZIPIII targets. a, Protein accumulation of HD-ZIPIII member (REVg:2xYPet-3xHA). **b,** Venn diagram comparing target genes identified by ChIP-seq. **c,** ChIP-seq profiles of HD-ZIPIIIs show co-occupancy on genomic regions for ZPR1 and ZPR4. **d, e,** Metaplot (**d**) and heat map (**e**) showing relative ChIP-seq read enrichment by ICU4, PHB, PHV, REV and ATHB8 (± 5000 bp from ICU4 bound regions). **f,** Box plots representing basal expression levels of genes neighboring HD-ZIPIII occupied and co-occupied genomic regions. Colored boxes represent the number of HD-ZIP family members co-occupied and their effect on gene expression. **g,** FLAG-tagged REV co-immunoprecipitates with YFP-tagged PHB, PHV and ICU4. IB, immunoblot; IP, immunoprecipitation. **h,** Global RNA-seq gene expression changes after 15 minutes, 1 hour and 4 hours of induction HD-ZIPIIIs. **i,** Bar graphs representing the fold enrichment of differentially expression genes after dex inductions for co-

occupied genes. Color bar represents $\log_2$ values from +3 (yellow) to -3 (blue) for (h). Scale bar represent 20μm in (**a**).

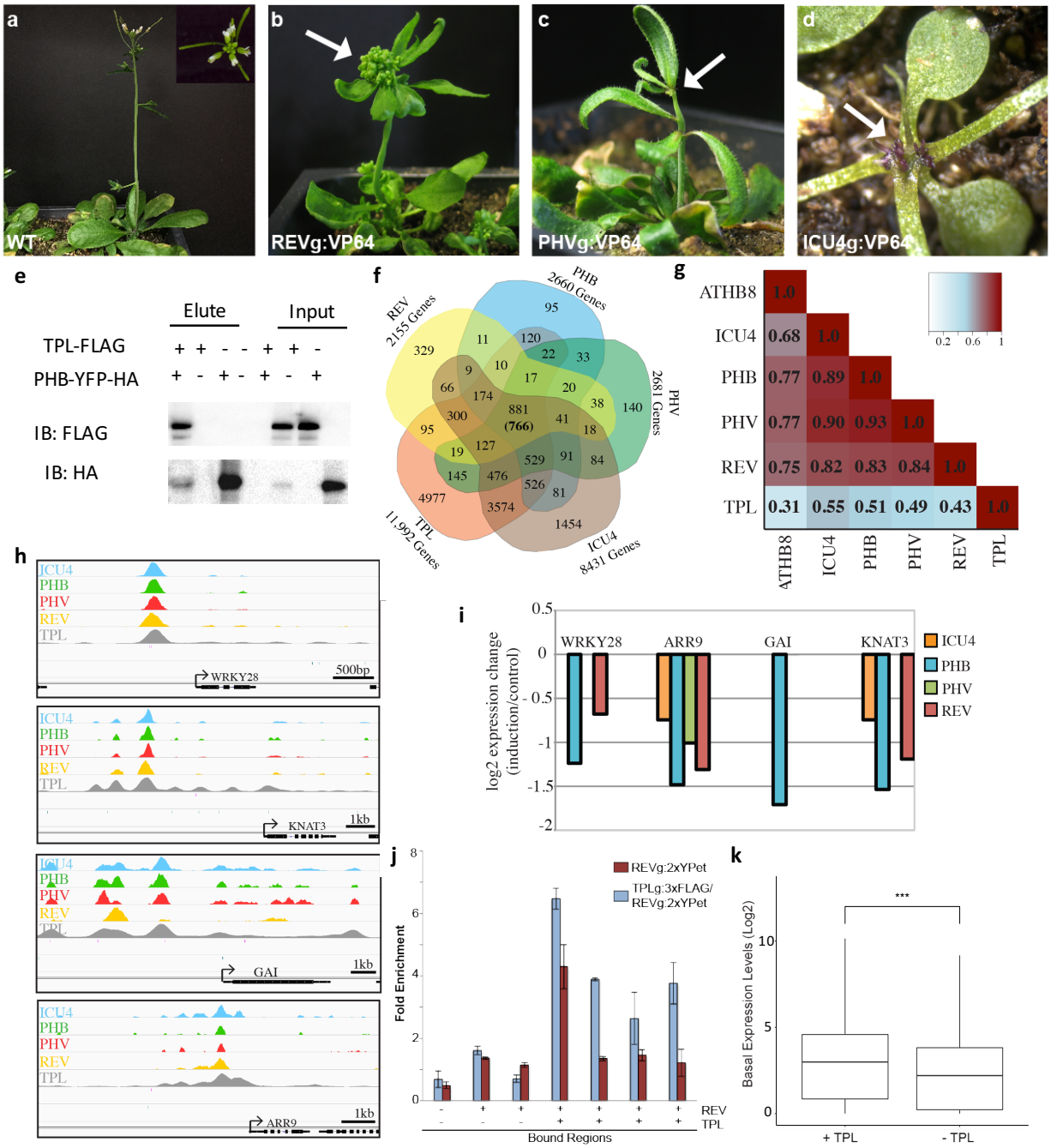


Figure 2 | HD-ZIPIIIs can act as repressors and recruit TPL to repress targets. a, Wild-type (WT) plant with inlet showing aerial view of flower. **b,** Endogenously expressed REV fused to VP64 (REVg:VP64) displays an altered meristem compared to WT. **c, d,** PHVg:VP64 (**c**) and ICU4g:VP64 (**d**) transgenic plants display terminated meristems. **e,** YFP-HA-tagged PHB co-immunoprecipitates with FLAG-tagged TPL. IB, immunoblot; **f,** Venn diagram comparing target genes identified by ChIP-seq between TPL and HD-ZIPIII members REV, PHB, PHV and ICU4. **g,** Pearson correlation coefficients for pairwise comparisons of peaks identified by ChIP-seq between HD-ZIPIII members and TPL. **h,** ChIP-seq profiles for HD-ZIPIIIs and TPL on select common targets. **i,** Transcriptional responses of TPL and

HD-ZIPIII common targets after HD-ZIPIII induction. **j,** Bar graph representing sequential ChIP assays performed on REVg:2xYPet and REVg:2xYPet/TPLg:3xFLAG. **k,** Boxplot of steady state expression levels (log 2 FPKM) of HD-ZIPIII bound regions lacking TPL (-TPL) or co-occupied with TPL (+TPL) in the promoter (defined as the 3kb region upstream of TSS).

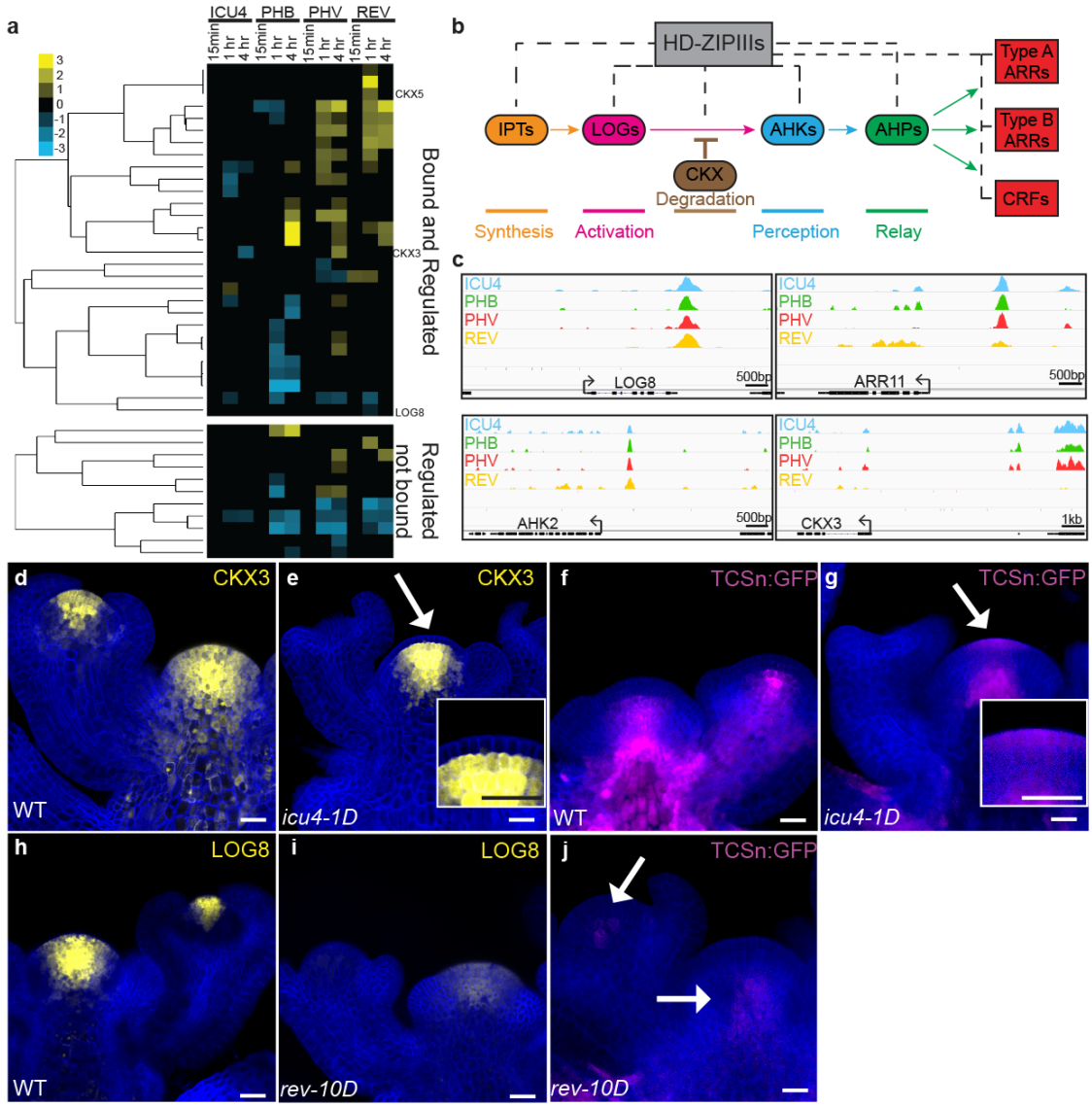


Figure 3 | HD-ZIPIIIs antagonistically control the cytokinin-signaling pathway in the inflorescence meristem. a, Heatmap depicting hierarchal clustering analysis of gene expression changes after HD-ZIPIII induction within the cytokinin-signaling pathway. **b,** Model of HD-ZIPIII-cytokinin pathway interactions. **c,** ChIP-seq profiles of select cytokinin pathway genes. **d, e,** A transcriptional reporter for CKX3 (CKX3p::NLS-2xYPet) is expressed throughout (**d**) WT inflorescence meristems and is absent in the L1 of an (**e**) *icu4-1D* meristem (arrow and inlet). **f, g,** TCSn::GFP (magenta) is expressed within (**f**) WT post-embryonic meristems but excluded from L1 and L2 cell layers yet expands into the L1 of (**g**) *icu4-1D* meristems (arrow and inlet). **h, i,** A transcriptional reporter for LOG8 (LOG8p::NLS-2xYPet) is expressed throughout a (**h**) WT post-embryonic meristem (yellow) and displays lower expression in (**i**) *rev-10D* mutants. **j,** TCSn::GFP is reduced in *rev-10D*. Scale bars represent 20µm in (**d-j**). Cells stained with SR2200 (blue).

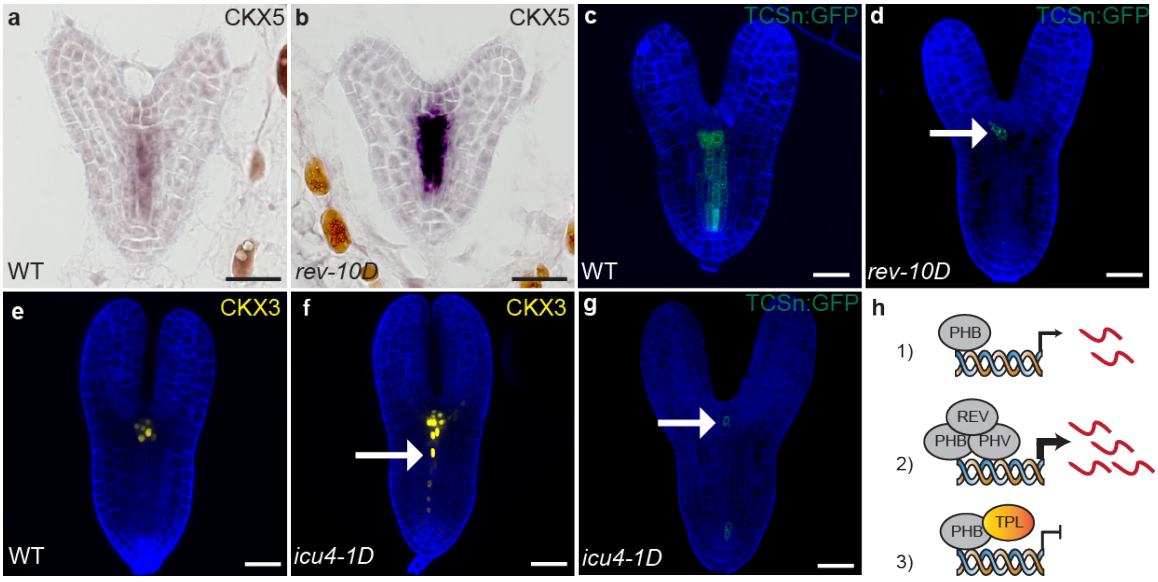


Figure 4 | HD-ZIPIIIs redundantly control cytokinin activity during embryogenesis and a model for cooperative gene regulation. a, b, *CKX5* mRNA is expressed in provascular cells of (**a**) WT embryos and is expressed at higher levels in (**b**) *rev-10D* embryos. **c, d,** TCSn::GFP (green) is expressed in provascular cells of (**c**) WT embryos and is greatly reduced in (**d**) *rev-10D* embryos. **e, f,** A transcriptional reporter for a cytokinin degradation enzyme (CKX3p::NLS-2xYPet), is expressed in cells underlying the SAM in a (**e**) WT embryo (yellow) and expands down into provascular cells (arrow) of (**f**) *icu4-1D* embryos. **g,** TCSn::GFP accumulates at lower levels in *icu4-1D* embryos. **h,** Model of HD-ZIPIII mediated transcriptional regulation. Width of arrow represents strength of transcriptional output and red lines represent mRNA transcripts. Scale bars represent 20μm in (**a-g**). Cells stained with SR2200 (blue).

Table 1. The number of HD-ZIPIII and TPL targets identified by ChIP-seq and RNA-seq

Factor Analyzed	# of peaks (ChIP-seq)	Genes associated with peaks	Bound and regulated targets (ChIP-seq + RNA-seq)	Peaks co-occupying WUS bound sites
REV	5,815	2,277	**15 min**: 108 (7D, 101U) **1 hr**: 608 (337D, 271U) **4 hr**: 292 (165D, 127U)	14/159
PHB	2,790	2,148	**15 min:** 132 (82D, 50U) **1 hr:** 526 (320D, 206U) **4 hr:** 593 (294D, 299U)	26/159
PHV	2,872	2,161	**15 min:** 74 (23D, 51U) **1 hr:** 572 (368D, 204U) **4 hr:** 796 (419D, 377U)	23/159
ICU4	10,383	7,021	**15 min:** 139 (59D, 80U) **1 hr:** 355 (162D, 193U) **4 hr:** 621 (306D, 315U)	52/159
ATHB8	4,813	1,722	**15 min:** 105 (33D, 72U) **1 hr:** 336 (181D, 155U)	4/159

			4 hr: 433 (222D, 211U)	
TPL	12,785	9,797	n/a	74/159

The total number of identified binding sites (peaks) and target genes associated with the peaks (within +/- 5kb) identified for REV, PHB, PHV, ICU4, ATHB8 and TPL after ChIP-seq analysis. These target genes were compared to transcriptional changes identified through RNA-seq analysis after 15 minutes, 1 hour or 4 hour of dexamethasone inductions. The overlap between the two experiments were considered to be bound and regulated targets for both up (U) and down (D) differentially expressed transcripts. HD-ZIPIIIs peaks were also compared to WUS ChIP-ChIP data [16].

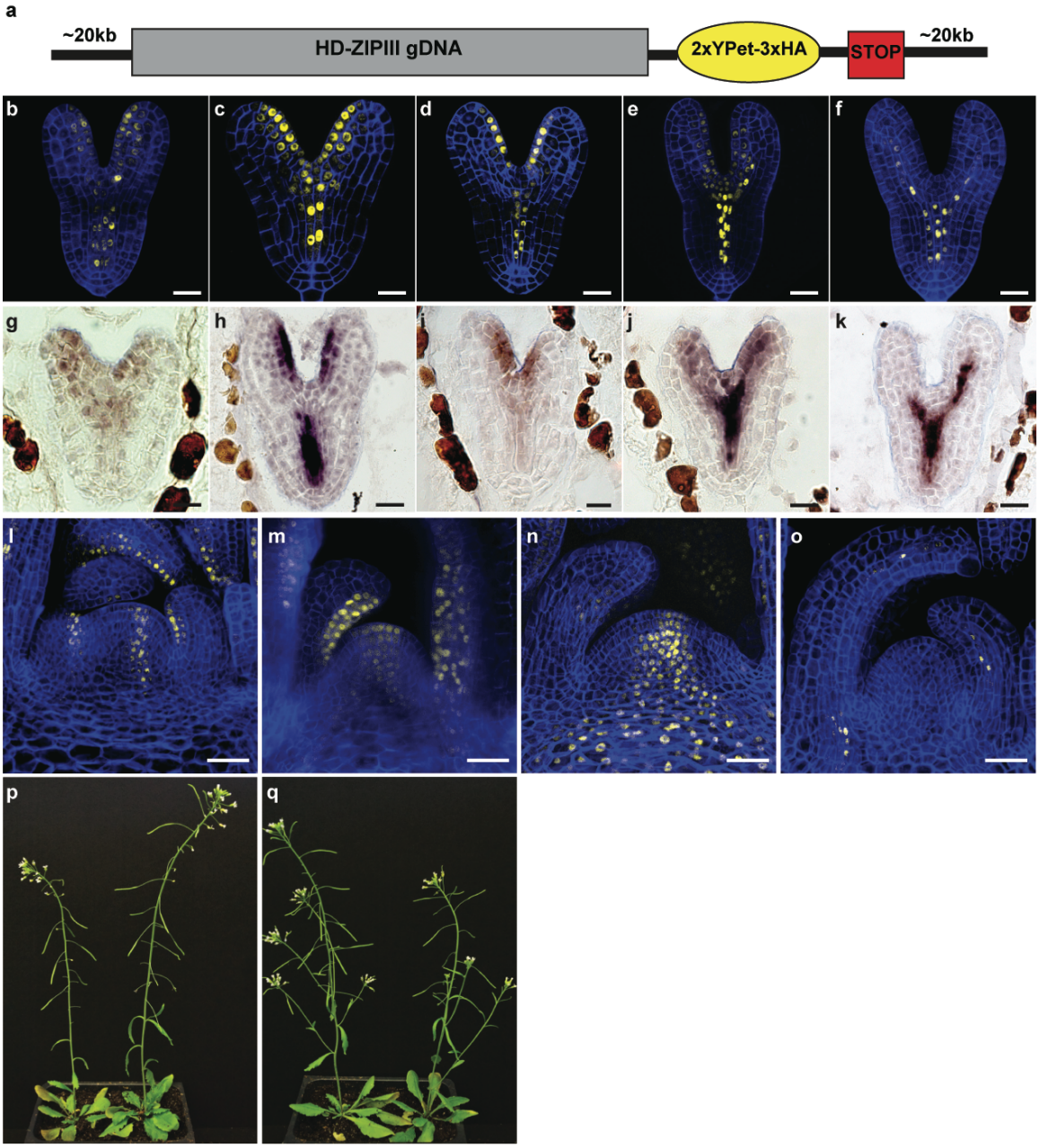
Table 2. HD-ZIPIIIs form heterodimers and associate with TPL and TPR proteins.

	LocusID	Description	Spectra		NSAFe5		% of REV	
REV-Flag	AT5G60690	REV	30	88	448	294	100	100
	AT1G30490	PHV	14	44	209	147	47	50
	AT2G34710	PHB	13	40	192	132	43	45
	AT1G52150	ICU4	7	22	105	74	23	25
	AT4G32880	ATHB-8	10	17	151	57	34	20
	AT1G15750	TPL	2	12	22	30	5	10
	AT1g80490	TPR1	-	12	-	30	-	10
	AT5G27030	TPR3	-	4	-	10	-	3
	AT3G15880	TPR4	-	6	-	15	-	5

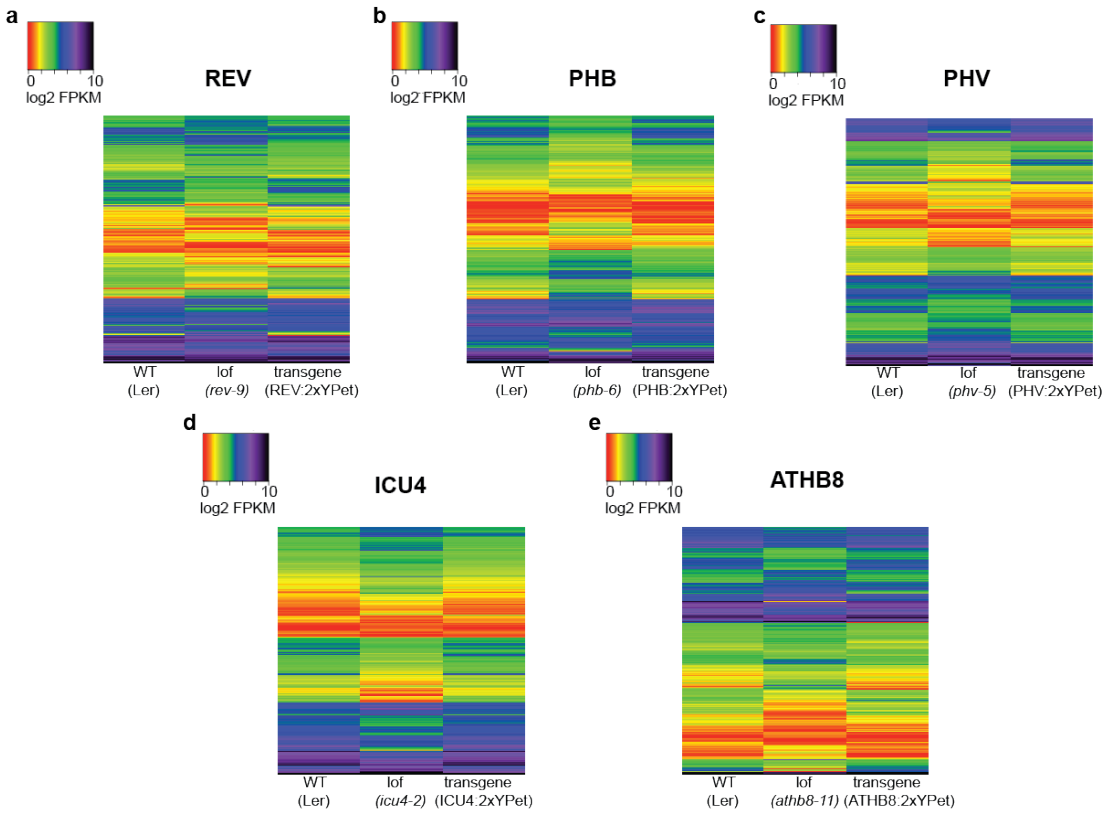
	LocusID	Description	Spectra		NSAFe5		% of PHB	
PHB-Flag	AT2G34710	PHB	24	50	479	97	100	100
	AT5G60690	REV	14	17	283	33	59	34
	AT1G30490	PHV	8	14	162	27	34	28
	AT1G52150	ICU4	-	5	-	10	-	10
	AT4G32880	ATHB8	8	-	163	-	34	-
	AT1G15750	TPL	-	18	-	25	-	26
	AT1g80490	TPR1	-	16	-	25	-	26
	AT3G15880	TPR4	-	10	-	14	-	15

	LocusID	Description	Spectra		NSAFe5		% of ICU4	
ICU4-Flag	AT1G52150	ICU4	20	45	262	139	100	100
	AT5G60690	REV	7	10	91	31	35	22
	AT1G30490	PHV	-	18	-	55	-	40
	AT2G34710	PHB	14	20	180	60	69	44
	AT4G32880	ATHB-8	9	19	118	59	45	42
	AT1G15750	TPL	2	9	19	21	7	15
	AT1g80490	TPR1	-	9	-	21	-	15
	AT5G27030	TPR3	-	4	-	9	-	7

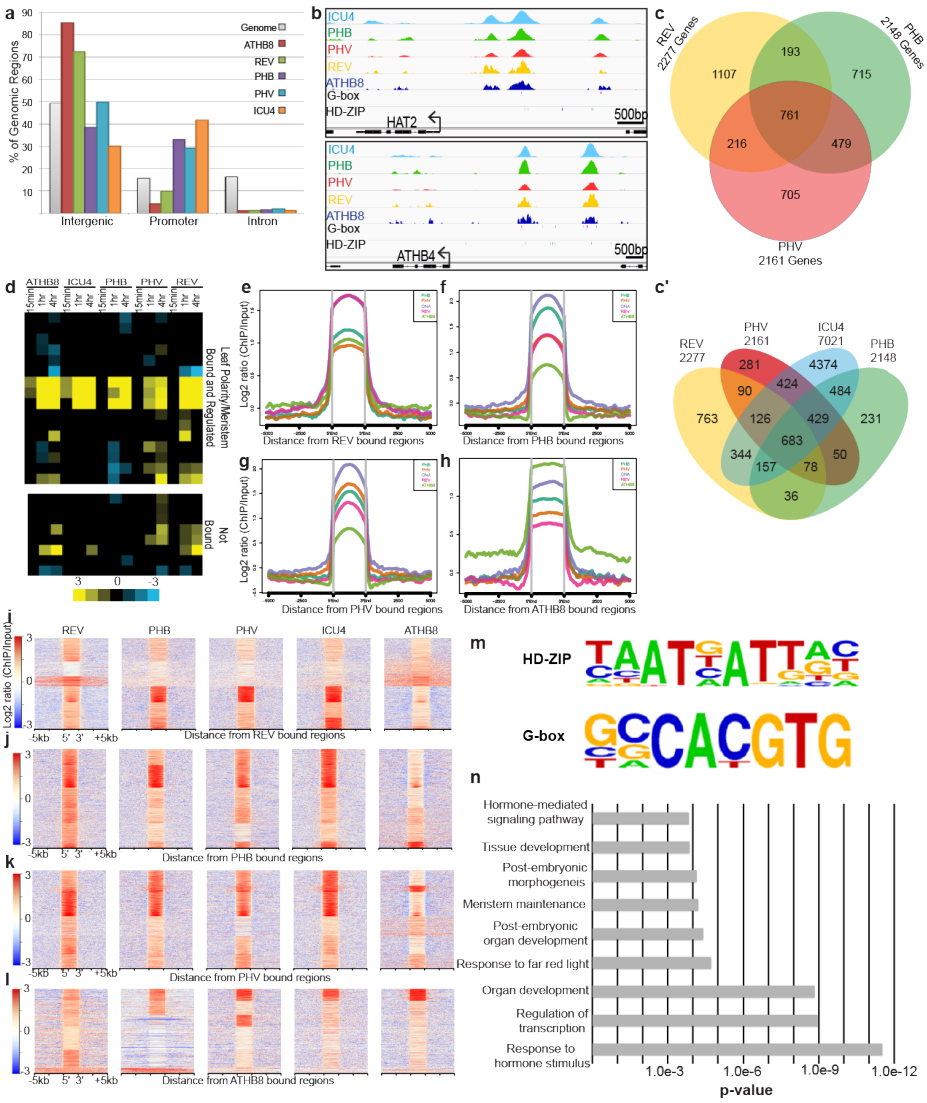
The total number of identified spectra, the normalized spectral abundance factor (NSAF) and the percentage relative to REVg:3xFLAG, PHBg:3xFLAG or ICU4g:3xFLAG are given for two biological replicates.



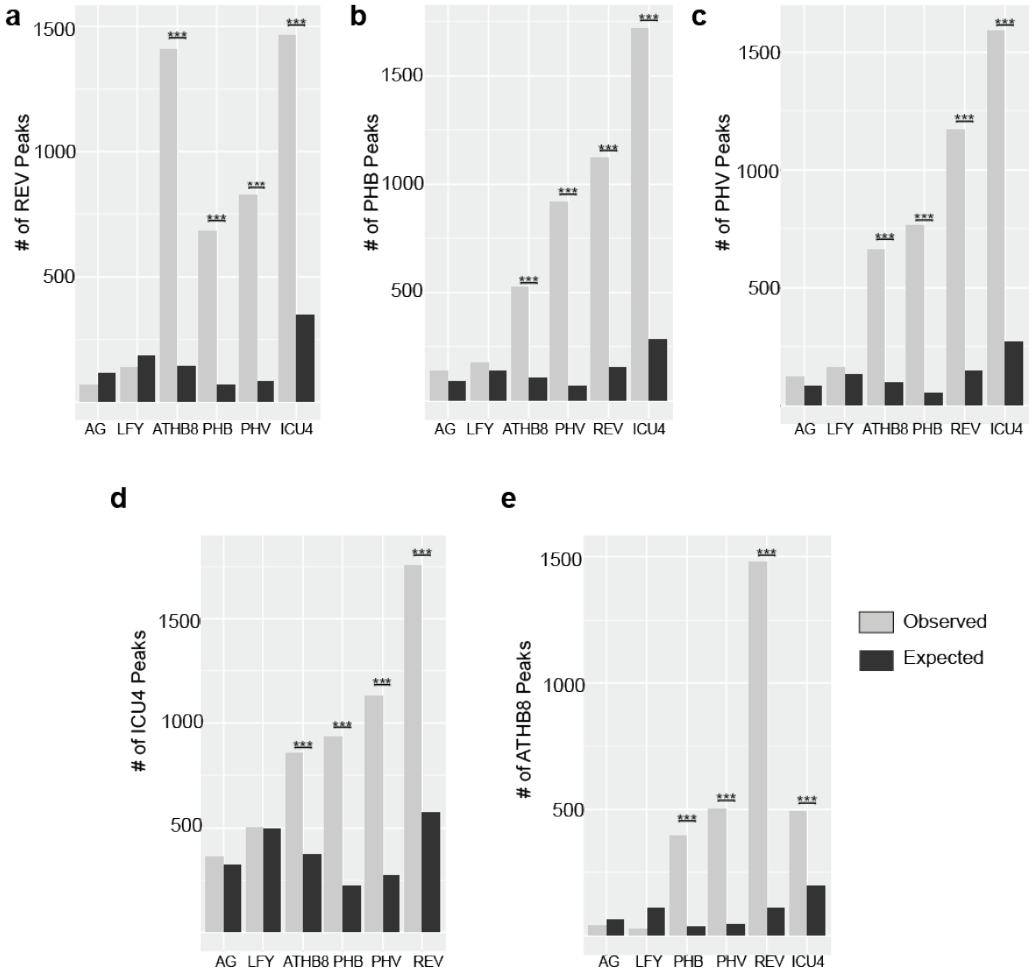
Extended Data Figure 1 | Protein accumulation and rescue by HD-ZIPIII recombineered constructs. **a,** Schematic of C-terminal fusion proteins made with recombineering. **b-k,** Protein accumulation of REV (**b**), PHB (**c**), PHV (**d**), ICU4 (**e**) and ATHB8 (**f**) during embryogenesis matches their respective mRNA expression patterns (**g-k**). **l-n,** Recombineered reporters accumulate in the meristem and adaxial/dorsal side of leaves for PHB (**l**), PHV (**m**) and ICU4 (**n**). **o,** ATHB8 is expressed in vascular cells only and not the meristem (F). **p, q,** Recombineered REVg:2xYPet-3xHA (**q**) rescues the axillary meristem phenotype seen in *rev-1* plants (**p**).



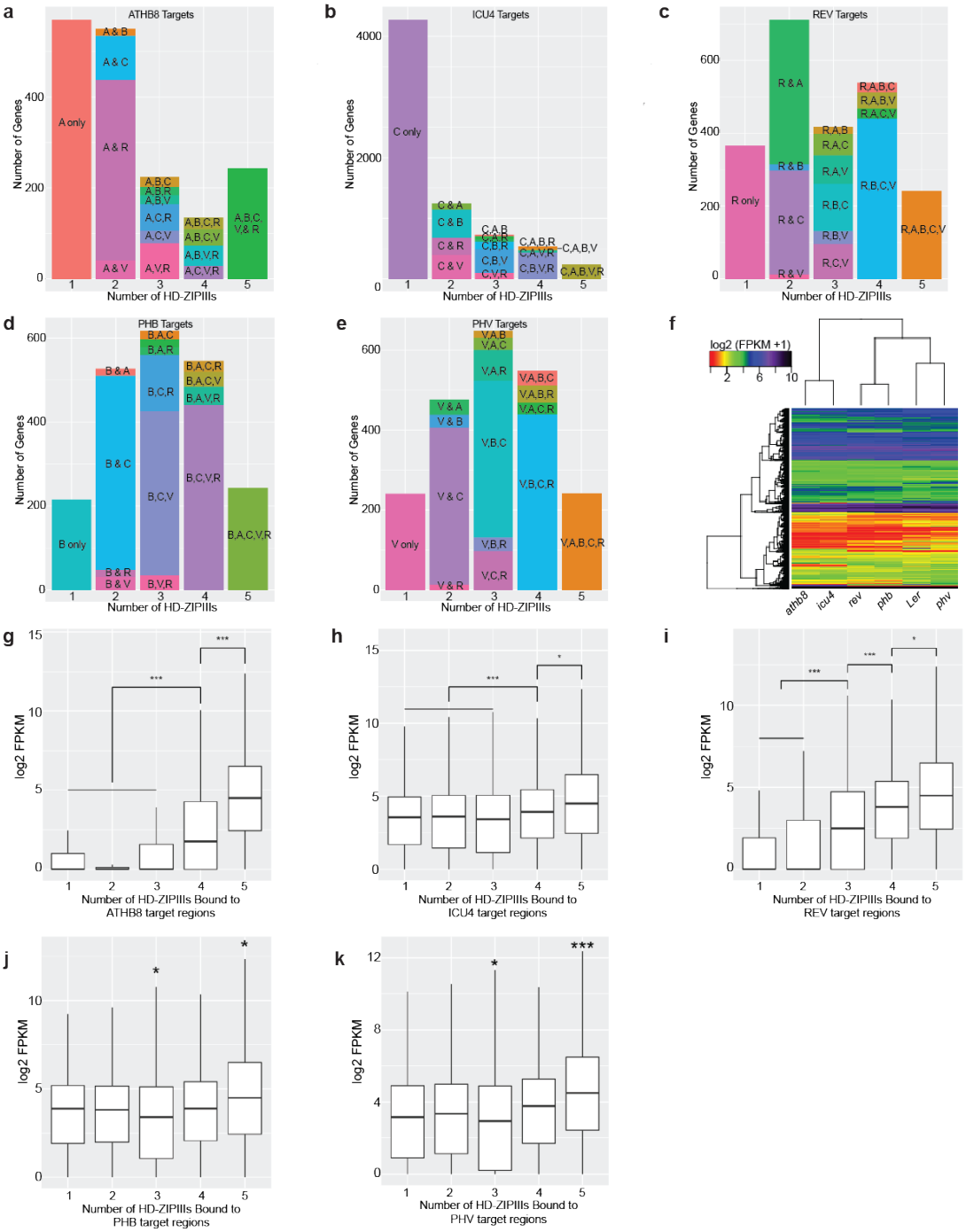
Extended Data Figure 2 | Molecular rescue by HD-ZIPIII recombineered constructs. a-e, Heatmaps representing RNA-seq gene expression analyses of recombineered lines for REV (**a**), PHB (**b**), PHV (**c**), ICU4 (**d**), and ATHB8 (**e**). Genes that showed a significant change in expression ($q < 0.05$) in loss of function (lof) mutants vs. wild type (WT) were plotted.



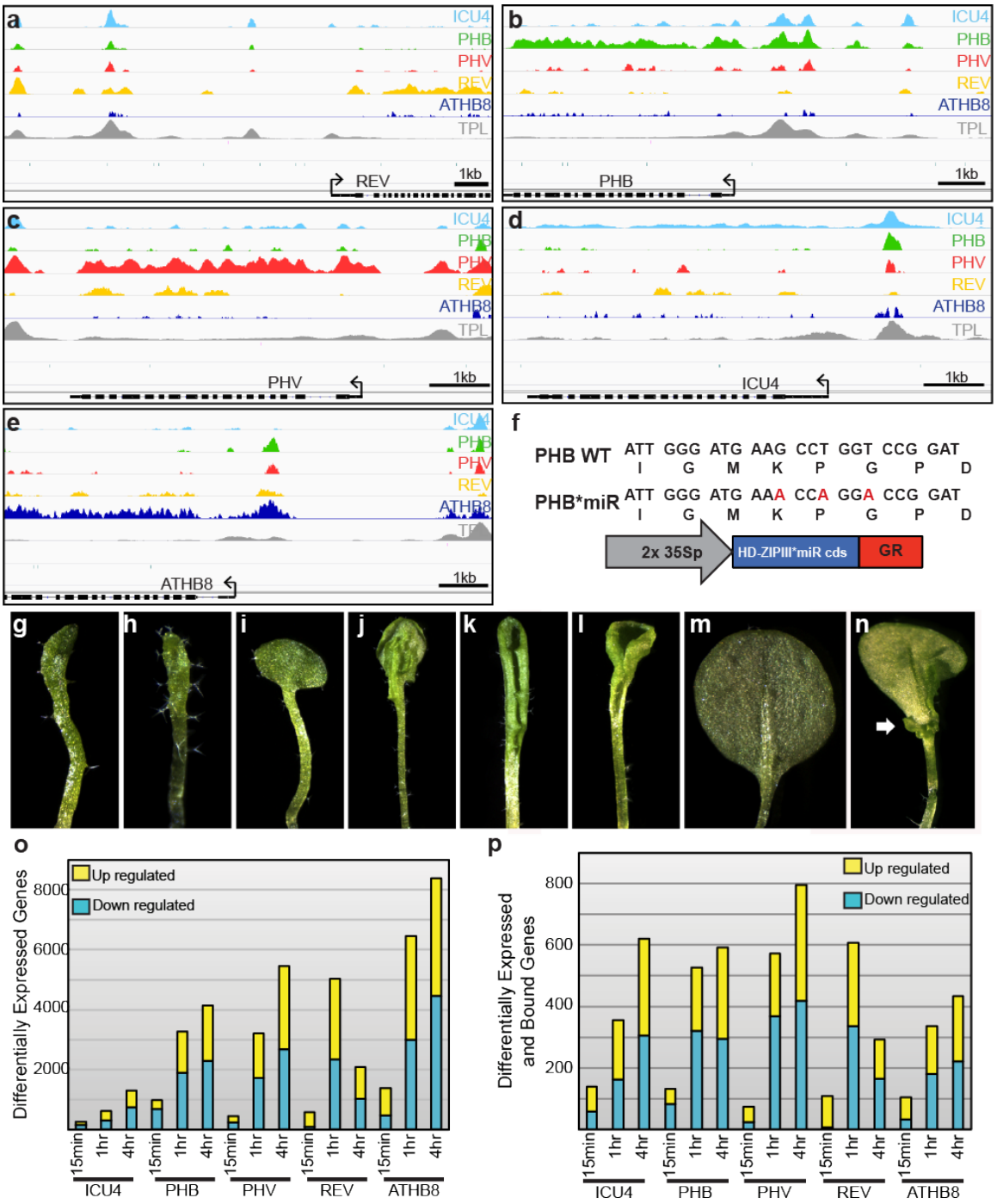
Extended Data Figure 3 | ChIP-seq analyses of genome-wide HD-ZIPIII targets. a, Distribution of HD-ZIPIII occupancy across genomic regions. Intergenic regions were defined as being greater than 2kb away from transcriptional start site (TSS) but not within coding regions. Promoter regions were defined as being with 2kb of the TSS. **b,** ChIP-seq profiles of HD-ZIPIIIs show co-occupancy of HAT2 and ATHB4. **c,** Venn diagrams comparing target genes between (**c**) PHB, PHV, REV and (**c'**) PHB, PHV, REV and ICU4. **d,** Hierarchal cluster analysis of differentially expressed target genes that are key regulators of the leafy polarity and meristem patterning pathways. **e-l,** Metaplots (**e-h**) and heat maps (**i-l**) showing relative ChIP-seq read enrichment by HD-ZIPIIIs to bound regions for REV (**e, i**), PHB (**f, j**), PHV (**g, k**) and ATHB8 (**h, l**) (± 5000 bp from peak). **m.** *De-novo* motif analysis identifies two dominant motifs enriched among HD-ZIPIII peaks, defined as HD-ZIP motif (AAT(G/C)ATT) and G-box motif (CACGTG). **n.** Select enriched GO terms among overlapping targets



Extended Data Figure 4 | Overlap of HD-ZIPIII binding sites with publicly available ChIP-seq datasets. a-e, Bar graphs representing the observed and expected number of overlap of binding sites between HD-ZIPIII members for REV (**a**), PHB (**b**), PHV (**c**), ICU4 (**d**) and ATHB8 (**e**). The transcription factors LEAFY (LFY) and AGAMOUS (AG) were plotted as representatives of the combined data . *** represents p value < 0.0001.

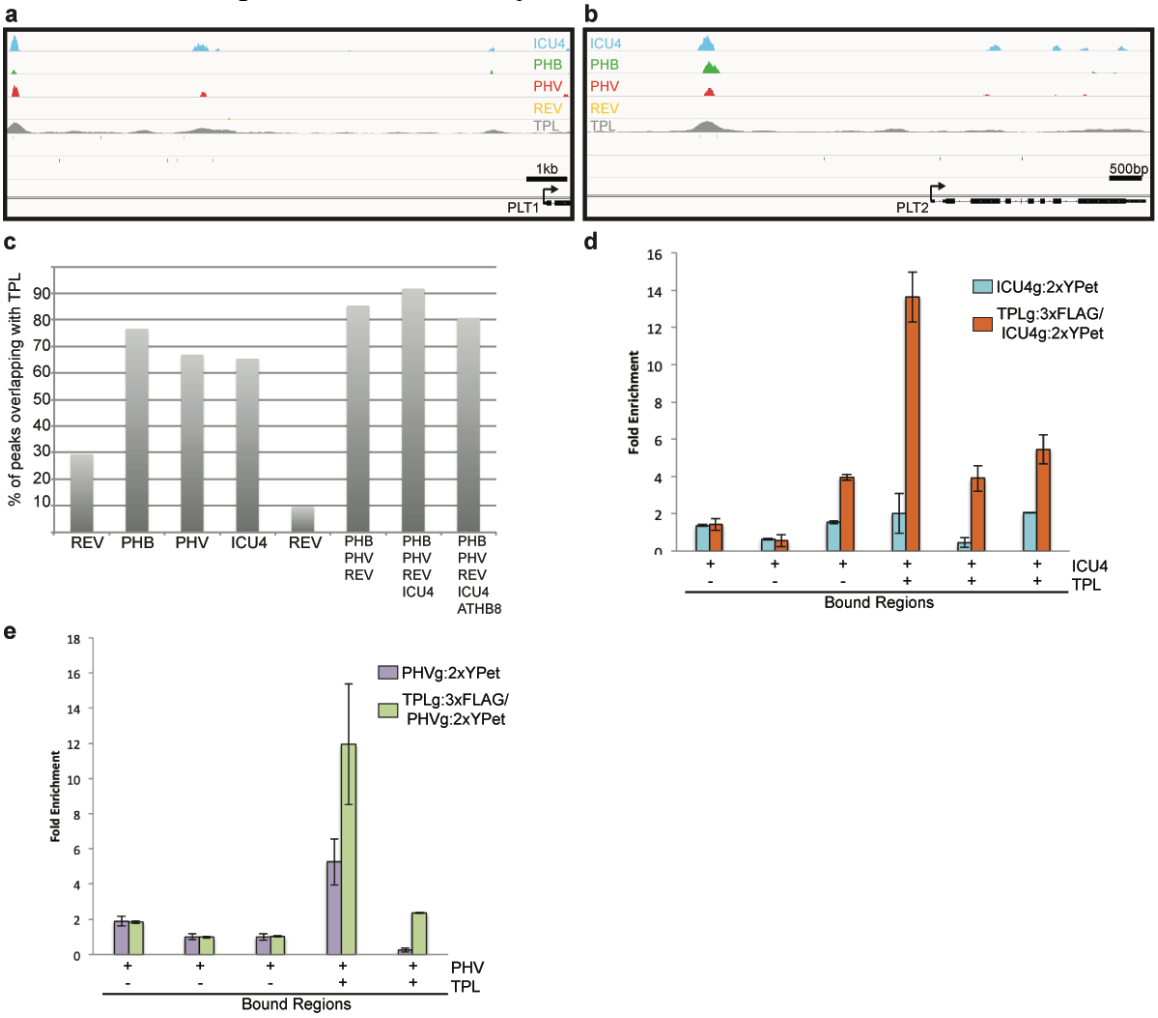


Extended Data Figure 5 | ChIP-seq and RNA-seq design and analysis. a-e, Clustered bar graphs representing significant unique and shared targets between HD-ZIPIII members for ATHB8 (**a**), ICU4 (**b**), REV (**c**), PHB (**d**) and PHV (**e**). Letter representation for HD-ZIPIII members are A (ATHB8), B (PHB), V (PHV), C (ICU4) and R (REV). Statistical significance determined using hypergeometric tests (Extended Data Figure 4). **f,** Heatmap representing gene expression changes in loss of function mutants vs. wild type. **g-k,** Box plots representing basal expression levels of genes neighboring HD-ZIPIII occupied and co-occupied genomic regions for ATHB8 (**g**), ICU4 (**h**), REV (**i**), PHB (**j**), and PHV (**k**). *** represents p value < 0.0001, * represents p value < 0.01

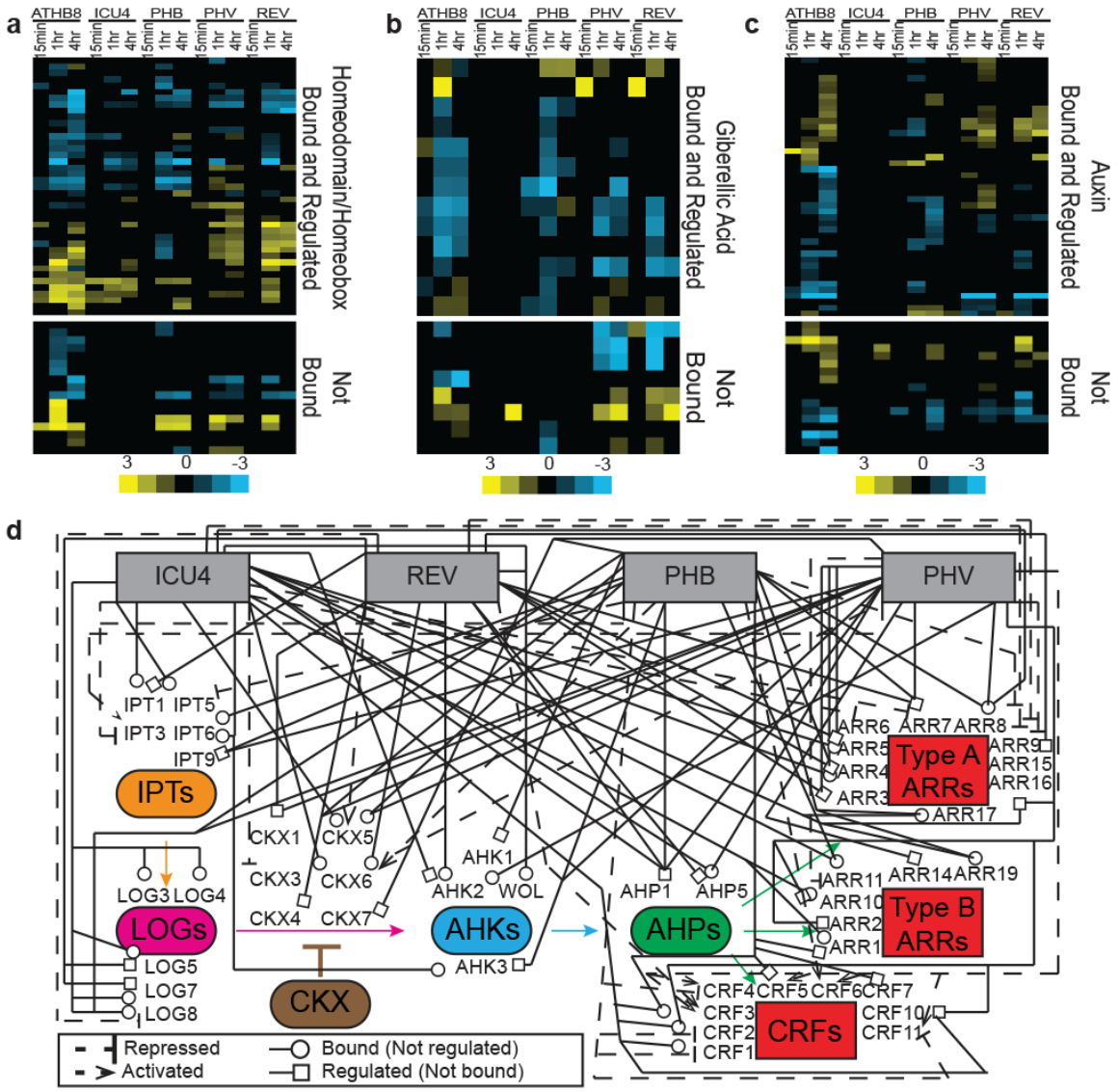


Extended Data Figure 6 | ChIP-seq and RNA-seq design and analysis. a-e, ChIP-seq profiles for HD-ZIPIIIs show co-occupancy at upstream regions for REV (**a**), PHB (**b**), PHV (**c**), ICU4 (**d**) and ATHB8 (**e**). In each panel, the topmost trace is ICU4 (blue), followed by PHB (green), PHV (red), REV (yellow), ATHB8 (dark blue) and TPL (grey). **f,** To identify differently expressed targets, microRNA resistant (*miR) versions of each HD-ZIPIII member were created by making silent mutations (red) in the miR site and fused to the inducible glucocorticoid receptor (ex. p35S:PHB*miR-GR). **g-n,** Induced lines used for transcriptional profiling experiments phenocopy gain-of-function mutants for *phb-1D* (**g**), PHB-GR (**h**), *phv-1D* (**i**), PHV-GR (**j**), *icu4-1D* (**k**), ICU4-GR (**l**), *rev-10D* (**m**) and REV-GR (**n**). **o,** The number of

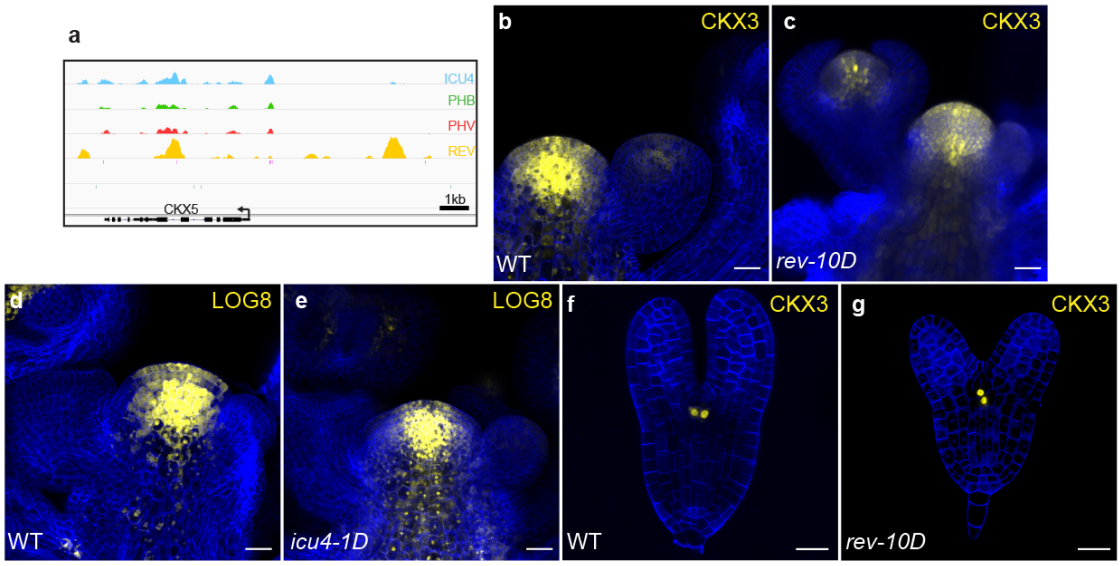
differentially expressed genes following dexamethasone induction from RNA-seq analyses. **p,** The number of direct targets with differential expression after inductions.



Extended Data Figure 7 | Overlap of targets and sequential ChIP assay between HD-ZIPIIIs/TPL. **a,b,** ChIP-seq profiles for HD-ZIPIIIs and TPL show co-occupancy at upstream regions for (**a**) PLT1 and (**b**) PLT2. **c,** Bar graph representing the percentage of peak overlap between TPL, individual HD-ZIPIII members and common peaks between combinations of HD-ZIPIIIS as observed by ChIP-seq analyses. **d-e,** Bar graphs representing sequential ChIP assays performed on ICU4g:2xYPet and TPLg:3xFLAG (**d**) and PHVg:2xYPet and TPLg:3xFLAG (**e**).



Extended Data Figure 8 | Regulation of select pathways by HD-ZIPIIIs. a-c, Hierarchal cluster analysis of differentially expressed genes including (**a**) homeodomain/homeobox proteins, (**b**) Gibberellic Acid pathway, and the (**c**) Auxin pathway. Color bar represents $\log_2$ values from +3 (yellow) to -3 (blue). **d,** Model of a cytokinin gene-regulatory network as controlled by ICU4, REV, PHB and PHV. Dashed arrows represent gene activation, dashed blunt lines represent gene repression, solid lines ending with a circle represent binding but no observed transcriptional changes and solid lines ending with a square represent genes with transcriptional changes but not bound by HD-ZIPIII members.



Extended Data Figure 9 | CKX3 and LOG8 reporter analysis in HD-ZIPIII gain-of-function mutants. a, ChIP-seq profiles for HD-ZIPIIIs at the CKX5 locus. **b, c,** A transcriptional reporter for CKX3 (CKX3p::NLS-2xYPet) is expressed in cells throughout the inflorescence meristems of (**b**) WT (yellow) and (**c**) *rev-10D*. **d, e,** A transcriptional reporter for LOG8 (LOG8p::NLS-2xYPet) in WT (**d**) and *icu4-1D* (**e**) inflorescence. **f, g,** CKX3 reporter (CKX3p:NLS-2xYPet) expression in a (**f**) WT (**g**) *rev-10d* embryo (yellow). Scale bars represent 20µm in (**c-j**). Cells stained with SR2200 (blue).

Materials and Methods

Growth Conditions and mutant alleles

Plants were grown on either soil or Petri dishes containing Linsmaier and Skoog (LS) salts medium. The wild-type plants used in this study were all in the Landsberg (Ler) background. Arabidopsis thaliana seeds were stratified for 3 to 5 days at 4°C prior to sowing. For ChIP-seq studies, plants were grown for 10 days after germination on 1.5% agar supplemented with LS plates and bulk harvested. For RNA-seq studies, growth and inductions were performed as previously described [9]. All plants were grown at 22 to 24°C under 16 hour light/8 hour dark cycle conditions. Creation of transgenic lines and reporter constructs utilized previously described mutants for *phb-6*, *phv-5*, *rev-10d*, *phb-14d*, *icu4-1d*, *tpl-1* [34] and *rev-1*, *athb8-11*, *icu4-2* [70].

Recombineering

Recombineering experiments were performed as previously described[180] with the following modifications. ChIP constructs utilized 2xYPet-3xHA and were fused in frame with the C terminus of REV (JAtY78C12 (57878 bp)), PHB (JAtY77D23 (54213 bp)), PHV (JAtY77A11 (79129 bp)), ICU4 (JAtY53N08 (69067 bp)), ATHB8 (JAtY59J22 (48651 bp)) and TPL (JatY67H04 (79026 bp)) in a transformation-competent artificial chromosome clone using a bacterial recombineering approach[146,181,182]. Each construct was transformed via the *Agrobacterium* floral dip method[183] into their respective loss of function background. For IP-MS experiments, constructs utilized a birA-3xFLAG and were fused in frame with the C terminus of REV, PHB and ICU4. TCS transcriptional reporters utilized an NLS-2xYPet with an ATG and were fused before the start codon of CKX3 (JAtY72D11 (59880 bp)) and LOG8 (JAtY71C10 (70378 bp)). Primers used for recombineering are described in Table S9.

35s:GR Constructs

To generate 2x35Sp::HD-ZIPIII*miR-GR fusions, full-length cDNA for each HD-ZIPIII was first amplified by PCR (as a *BamHI-Bam*HI fragment). Silent mutations were then made in the microRNA recognition sequence of each HD-ZIPIII using site directed mutagenesis and cloned downstream of a 2x35S promoter and upstream of the rat glucocorticoid receptor (GR) in plasmid pBJ36. Each transgene was transferred into the *Not*I site of binary vector pMLBART and transformed into their respective loss of function mutant background with the exception of ICU4 and ATHB8, which were transformed into Ler. Western blot analysis using an anti-GR

antibody (Pierce PA1-516) was then performed on transgenic lines to isolate those with similar protein levels.

Histology and Microscopy

For GFP and YPET analysis, ovules, seedlings and inflorescence meristems (IMs) were fixed in 4% paraformaldehyde, vacuum infiltrated, rinsed with phosphate buffer saline (PBS), vacuum infiltrated with 2% SCRI Renaissance 2200 (Renaissance Chemicals) and 4% DMSO, then washed twice in PBS. Seedlings and IM's were embedded in 3% agarose and sectioned using a Thermo Scientific HM 650 V Vibration microtome. 50 uM sections were mounted in 20% glycerol while ovules were directly mounted after fixing [25]. Embryos, seedlings and IM's were imaged using a Zeiss LSM 710 laser scanning confocal microscope. For each reporter, >10 transgenic lines were analyzed and the most over represented patterns of expression were selected. SR2200 was excited with the ultraviolet diode 405 nm line, and emission was measured between 420 and 470 nm. GFP was excited with a 488nm argon laser line and emission was measured at 500–535 nm. YPET was excited with a 517nm argon laser line and emission was measured at 500–535 nm.

In-situs

In situ hybridizations were detected with digoxigenin-labelled riboprobes using the methods found at http://www.its.caltech.edu/~plantlab/protocols/insitu.htm. *PHB*, *PHV*, and *REV* probes were generated as described[34]. *ICU4* and *ATHB8* probes were generated using full-length cDNAs. The *CKX5* probe was generated as described[171]. For each probe, >50 embryos were screened and the most over represented expression pattern was shown.

ChIP

ChIP experiments were performed as previously described[34] in duplicates with the following modifications. Transgenic homozygous lines were used for each construct. Starting material consisted of 10 grams of 10 day-old seedlings grown on 1.5% agar supplemented with 1x LS. Nuclear extracts were first isolated as previously described[184] prior to immunoprecipitations. Protein A/G magnetic beads (Pierce 88803) were used with a polyclonal anti-GFP antibody (Life Technologies A6455) in immunoprecipitations.

Sequential ChIP

For each ChIP, ~50g of 10 day old seedlings were crosslinked as previously descirbed[34]. First, constructs harboring 2xYPet where ChIP'd with an anti-GFP antibody, protein complexes were then eluted in 30ul of elution buffer supplemented with 2% SDS, 15mM DTT and protease inhibitors at room temperature (21˚C) for 30 minutes. Eluates were pooled and diluted to at least 1:30 with ChIP dilution buffer. Protein complexes were concentrated with Millipore centrifugal filter units with a cut-off of 100 kDa. Sequential ChIP was performed on concentrated DNA/protein complex with 100 ul Sigma anti-FLAG M2 beads and incubated overnight and then washed 4 times with ChIP buffer and chromatin was eluted similar to ChIP protocol above.

ChIP-seq library preparation and sequencing

Immunoprecipitated DNA was first purified using AmpureXP beads and libraries were constructed using the NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina (NEB) using Illumina adapters according to the manufacturers protocol. Input DNA was used as a control. Suitable size distribution of the libraries was validated by agarose gel (between 200-

500bp). Libraries were multiplexed with a total of 10 libraries/lane and sequenced on an Illumina HiSeq 2000 (single-end 50-bp run).

ChIP-seq Analysis

Reads were mapped using bowtie2, against the TAIR10 *Arabidopsis* genome. Only uniquely mapped reads were used to call peaks using the MACS2 peak calling algorithm[185]. Replicates were used as the experiment and the input file was used as a control. MACS2 was run using default settings with a q value cutoff of 0.05. A pool of genomic intervals was generated by combining the publicly available ChIPseq data and our ChIPseq data. The significance of overlapping was checked by random sampling from the pool for 5000 times and p value is an estimation to obtain an observed or larger overlap. De-novo motif discovery was conducted using the HOMER tool (findMotifsGenome)[186]. The inputs to HOMER were the sequences of 200bp around the summits of peaks, as called by MACS2. In some cases, only the top 10% or 25% quality peaks were used, where quality was defined by the q-value output of macs2. Pearson correlation between two factors was calculated by using the reads at the union of the peak files of the respective factors. Each peak from the union is considered a separate observation. Heat maps and metaplot profiles of relative enrichment (log2 ChIP/Input) at defined genomic regions and were generated by ngs.plot[187] followed by k means clustering. Mapped reads were visualized using the Integrative Genomics Viewer (IGV) using bigwig files for each TF. Genes regulated by the HD-ZIPs are defined by BETA[188] with a distance cut-off of 5kb.

Inductions and RNA extraction

Transgenic homozygous lines for each construct were used. ~200 seedlings were grown in liquid culture as previously described[147] with the following modifications. Each genotype was grown in liquid culture for 10 days in duplicates. Plants were induced for 15min, 1 hour or 4 hours using 50 uM dexamethasone and flash frozen using liquid N2. Un-induced samples were treated as controls as well as Ler mock treated plants. Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) followed by DNase digestion (Promega) following the manufacturers protocol.

RNA-seq

RNA quality was first assessed by agarose gel analyses. Libraries were constructed using the TruSeq RNA Library Preparation Kit v2 (Illumina) following the manufactures protocol. Samples were sequenced on an Illumina HiSeq 2000 (single-end 50-bp run) with the 8 samples multiplexed per lane.

RNA-seq Analysis

Raw reads were mapped to the TAIR10 *Arabidopsis* genome using TopHat aligner (standard parameters, except for intron length of 40). Uniquely mapped reads were then used to assemble and quantify transcripts using Cufflinks using default parameters. Cuffdiff was then used to call differentially expressed transcripts between induced samples and un-induced samples (15min induction vs. un-induced) using a q value < 0.05[189]. Box-plot distribution of FPKM values was generated using R and statistical significance was calculated using a Mann-Whitney test.

Heat Maps

Heatmaps were generated using Cluster 3.0 using hierarchical clustering with uncentered correlation and average linkage. Clusters were then visualized using Java TreeView.

IP-MS

Immunoprecipitation and mass spectrometry analysis were done accordingly to[190]. Briefly, ten grams of 2-week-old seedling tissue were ground in liquid nitrogen using a RETSCH homogenizer and resuspended in 25 mL ice-cold IP buffer [50 mM Tris·HCl pH 8.0, 150 mM NaCl, 5 mM MgCl2, 0.1% Nonidet P-40, 10% (vol/vol) glycerol, 1x Protease Inhibitor Mixture (Roche)]. Resuspended tissue was dounced twice using a glass douncer, filtered once through Miracloth and centrifuged 10 min at 16,000g at 4°C. 200 µL M2 magnetic FLAG-beads (SIGMA, M8823) were added to the supernatant and incubated 120 min rotating at 4 °C. M2 magnetic FLAG-beads were washed five times in ice-cold IP buffer for 5 min rotating at 4 °C. Immunoprecipitated proteins were eluted two times with 300 µL 250ug/mL 3xFLAG peptides (SIGMA, F4799) in TBS [50mM Tris-HCl pH 7.4, 150mM NaCl] for 15 min shaking at 37 °C. Colombia (Col) wild-type tissue was used as a negative control. IPs were then digested by the sequential addition of lys-C and trypsin proteases, fractionated using reversed phase chromatography, and then analyzed by tandem mass spectrometry on a Q-Exactive mass spectrometer (Thermofisher). Data was analyzed using the IP2 suite of software tools.

Co-IP

For co-immunoprecipitation experiments, one gram of floral tissue was isolated from F1 crosses between recombineered 2xYPet-3xHA and birA-3xFLAG or tdTomato-3xFALG constructs. Immunoprecipitations from detergent extracts were carried out using the micro-MACS GFP-

tagged protein isolation kit (Miltenyi Biotec), according to the manufacturer's instructions. Immunoblot analysis was performed with an antibody against FLAG (SIGMA, A8592).

Extended Data table.

Recomb_REV F	ACAACAATCTGCATTGTCTTGCCTTCTCCTTTGTAAACTGGTCTTTTGTGGGTTCGAAATCGATAAGC
Recomb_REV R	TGAGTTCTCTAACGTAAATTTAAATTAGTCTTTTTCTGTCAATCGAATCAATCTCATTAAAGCAGGACTCT
Recomb_PHB F	ATGAATCAAACCACTGCTTAGCTTTCATGTTCGTGAATTGGTCGTTCGTTGGTTCGAAATCGATAAGC
Recomb_PHB R	AAGGGAAAAATGGAGTAAAATCTGGTCTTCTTATTTCTTATTATTCTTCAATCTCATTAAAGCAGGACTCT
Recomb_PHV F	ACGAATCAAACCACTGCTTGGCTTTTACCCTCGTTAGTTGGTCGTTTGTTGGTTCGAAATCGATAAGC
Recomb_PHB R	AACATCCCCTAACTTTTAATATCTAACATAACCAACCTTTTCCTTCATCAATCTCATTAAAGCAGGACTCT
Recomb_ICU4 F	AAGAAAATGCTCATTGCATCTGCTTTGTGTTCATCAATTGGTCCTTTGTGGGTTCGAAATCGATAAGC
Recomb_ICU4 R	AGCCAAAAGGCAAAAGCTTAGTCTGAAAATACAAAATACAATAAATCTCAATCTCATTAAAGCAGGACTCT
Recomb_ATHB8 F	ACGAAGATCCTCATTGTATCTGCTTCATGTTCCTCAACTGGTCTTTTATAGGTTCGAAATCGATAAGC
Recomb_ATHB8 R	AAAAAAGGTCTGTGTTTTGATAATTAAGGAAAAACATGGAAAAAAGGTCAATCTCATTAAAGCAGGACTCT
TPL Recomb F	CACCAACTGCACCTTCCGTTGGAGCCTCTGCATCTGATCAGCCTCAGAGAGGTTCGAAATCGATAAGC
TPL Recomb R	TAAATAGAAGAAAGGGAATCAGAGAGGTAGGTGGCTCTCGGATCAGATCAATCTCATTAAAGCAGGACTCT
CKX3 Recomb P F	GATAAGAATCAAGCTATTCATAAAATAAAAAGACTTATTTCTCAAAAAAAGGTTCGAAATCGATAAGC
CKX3 Recomb P R	ACTATTGTTATTGCTATAAGACGAACTTGTGAACGAAGATTATAACTCGCAGCTTGCATGCCGGTCCTGCT
LOG8 Recomb P F	TTGTTTGACGACCTGAGGGAATACCTTGATTTGTAGATTTTTATCCATTAGGTTCGAAATCGATAAGC
LOG8 Recomb P R	CTTCCGCAAAAGACACAGATTTTTCTGAATCTGCTTCGCTGATTATCTTCAGCTTGCATGCCGGTCCTGCT
REV_5'_BAMHI-F	GGATCCATGGAGATGGCGGTGGCTAACCACC
REV_3' BamHI Fus-R	GGATCCACAAAAGACCAGTTTACAAAGGAG
PHB_5'_BamHI-F	GGATCCTATGATGATGGTCCATTCGATGAGC
PHB_3'_Fus_BamHI-R	GGATCCACGAACGACCAATTCACGAAC
PHV_BamHI-F	GGATCCATGATGGCTCATCACTCCATGGACC
PHV_fus_BamHI-R	GGATCCACAAACGACCAACTAACGAGGG

ICU4_BamHI-F	GGATCCATGGCAATGTCTTGCAAGGATGG
ICU4_fus_BamHI-R	GGATCCACAAAGGACCAATTGATGAACAC
ATHB-8_BamHI-F	GGATCCATGGGAGGAGGAAGCAATAATAG
ATHB-8_fus_Bam-R	GGATCCAATATAAAAGACCAGTTGAGGAACATG

Chapter 3: TPL recruits HDA19 to regulate gene expression

Introduction

In multicellular organisms, the formation of complex body patterns often requires precise spatial and temporal regulation of fate-specifying genes. How the transcription regulators carry out this function to establish the identities of different types of cells and tissues is a critical question in developmental biology. While transcriptional activation has been extensively studied and appreciated, transcriptional repression is emerging to be as important as activation[191]. In eukaryotes, transcriptional repression is achieved by a variety of strategies that can be divided into two categories: active and passive repression. Passive repression involves mechanisms such as competing with transcriptional activators for DNA binding and inactivating activators by physical interaction. Active repressors have intrinsic repression domains and the repression function is independent of activators. In both animals and plants, many patterning processes require active repression of transcription by DNA-binding transcription factors (TFs) that recruit co-repressors. Co-repressors further recruit factors such as chromatin-modifying enzymes to contribute to transcriptional repression[192]. In animal, the Groucho (Gro)/Transducin-like enhancer of split (TLE) family of corepressors recruit class I histone deacetylase Rpd3/HDAC1 and have been implicated in multiple developmental processes[2]. In *Arabidopsis*, the TOPLESS(TPL)/ TPL-RELATED(TPR) family proteins resemble Groucho(Gro)/TLE transcriptional co-repressors and the founding member, TPL, plays a pivotal role in a wide range of patterning processes during both embryonic and post-embryonic development[27,33,37].

The mechanisms of transcriptional repression by Gro/TLE co-repressors have been extensively studied. Gro/TLE proteins form oligomers through the interaction of the amino (N) termini. It has been proposed that the oligomerization over broad regions of chromatin leads to long-range repression on target genes[193]. However, recent studies using chromatin immunoprecipitation followed by high throughput sequencing analysis (ChIP-seq) showed spreading of the protein along the chromosome happens only at a few loci and a loss of the ability to oligomerize does not significantly affect the binding pattern[13-15]. TPL also has an N-terminal dimerization domain, which has been shown to mediate the formation of tetramers[50,51]. Whether and how oligomerization contribute to the association of TPL in the genome is still largely unknown.

In *Arabidopsis*, the flower consists of four types of lateral organs organized in concentric whorls, which from the peripheral to the central are: sepals, petals, stamens and carpels. The fate of each whorl is tightly regulated by floral organ identity genes, the misregulation of which can result in homeotic transformation of one type of floral organ to another[194,195]. In *Arabidopsis*, TPL-mediated transcription repression is crucial for floral development. It has been shown that TPL is recruited by the floral organ identity TF APETALA2(AP2), forming a co-repressor complex with histone deacetylase HDA19 to repress the expression of AP2 targets[27]. The *tpl/tpr1-4* quintuple loss-of-function mutants and late-arising flowers of *hda19-1* mutants exhibit carpeloid sepals, phenocopying the strong *ap2-2* allele[27]. Therefore, the flower is a useful model for studying the mechanisms of transcriptional repression mediated by TPL in the context of development.

Although Gro/TLE type of transcription co-repressors in *Arabidopsis* have been shown to recruit HDACs to their target sites, the detailed function of these co-repressor complexes during development remains to be fully characterized. In addition, high quality ChIP-seq data of histone deacetylases are needed to understand the detailed function of histone deacetylation. Here we assayed the genome-wide binding profiles of TPL and HDA19 in inflorescence meristems and young flowers. We found that HDA19 largely co-localize with TPL at discrete regions in the genome, suggesting a non-spreading model similar to that in animals. Specifically, we found that TPL fine-tunes the expression of three evolutionarily conserved NAM/CUC (for NO APICAL MERISTEM and CUP-SHAPED COTYLEDON) transcription factors, *CUC1*, *CUC2* and *CUC3*, in different ways either dependent or independent of HDA19.

Results

Genome-wide characterization of TPL and HDA19 recruitment

To identify genome-wide targets of TPL and HDA19, we performed ChIP-seq with inflorescence meristems. DNA binding events by chromatin-associating proteins such as TFs are characterized by a bimodal pattern of reads distribution[196]. Both our TPL and HDA19 ChIP-seq data displayed perfect bimodality with high signal-to-noise ratio, which is an indication of good quality (Fig 3.1a). TPL has been shown to form tetramers through the interaction of the N-terminal domains[50,51]. Oligomerization of Gro/TLE protein was proposed to contribute to the spreading of the protein over the chromatin, which will result in broad peaks with the size of several kilobases (kb). However, recent ChIP-seq data showed that most Gro peaks spanned less than 1kb, suggesting that the model of spreading does not hold true[13,14]. Just like Gro in *Drosophila,* most of the TPL peaks were smaller than 1kb, a sign that TPL also typically occupies narrow regions

in the genome (Fig 3.1b). TPL tended to bind to clusters of regions (Fig 3.1d), with over 60% of the peaks located within 5kb of the neighboring peaks (Fig 3.1c). These data suggest that TPL localizes at clusters of discrete loci in *Arabidopsis* genome to regulate transcription.

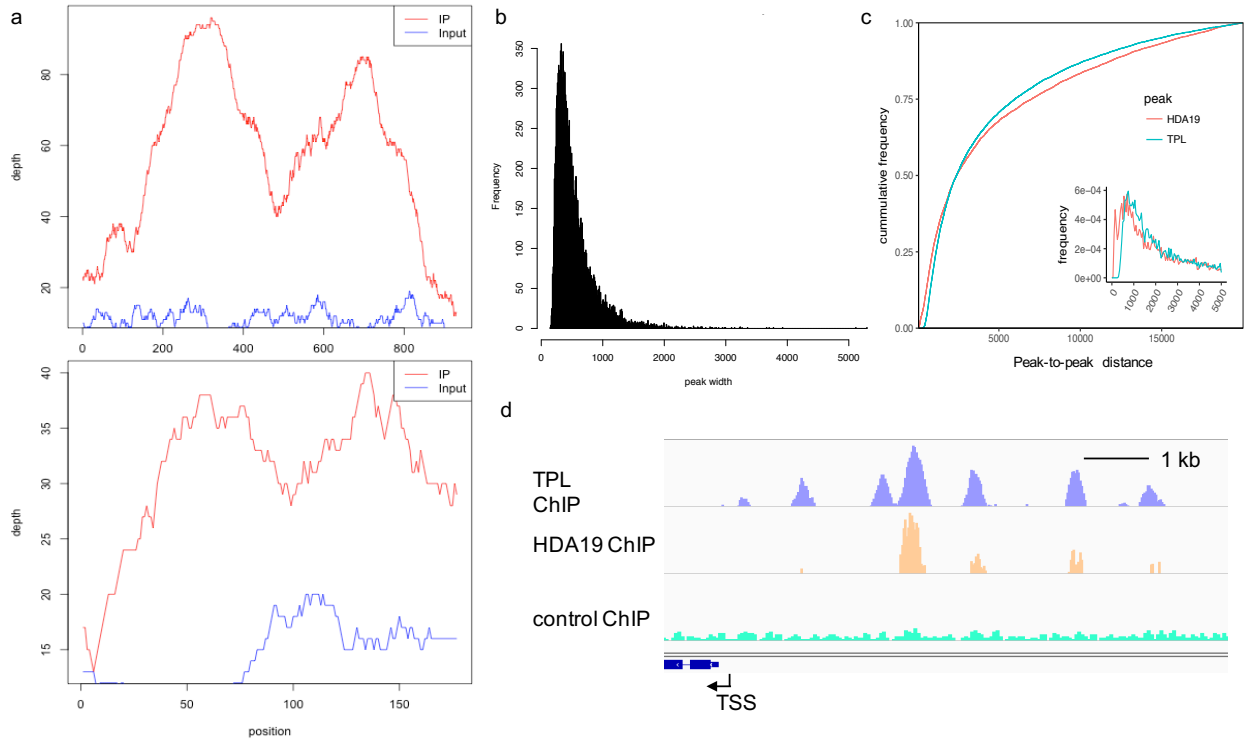


Fig 3.1: characterization of genome-wide TPL and HDA19 binding. **a,** ChIP-seq reads pile up TPL (upper) and HDA19 (lower) peaks, red-ChIP signal, blue-input signal. **b**, Length distribution of TPL peaks. **c,** Cumulative frequency and frequency of TPL (green) and HDA19 (red) peak-to-peak distances. **d,** Screen shot of TPL (top), HDA19 (middle) peaks. ChIP with Ler control is shown as the bottom track; Blue boxes indicate exons while blue lines indicate introns; arrow indicates transcription direction.

HDA19 has been shown to form a co-repressor complex with TPL to repress the expression of floral homeotic genes during flower development[27]. Consistent with the model, the majority of the HDA19 peaks co-localized with the TPL peaks (Fig 3.1d, 3. 2a and 2e). Both TPL and HDA19 peaks were highly enriched at the promoters in the genome, with over 60% of peaks located within 3kb upstream of transcription start sites (TSS) (Fig 3.2b, 2c and 2d). Peaks that were unique to HDA19, numbering in the low hundreds, showed an enrichment at TSS sites, significantly distinct from the pattern of overlapping peaks or TPL unique peaks (Kolmogorov–

Smirnov test p<2.2e-16), which points to a possible unique function of HDA19 at TSS's (Fig 3.2b and 2c). Additionally, HDA19 only occupied about 50% of the regions that were bound by TPL (Fig 3.2a). The significant number of unique TPL peaks suggests the existence of other TPL co-repressor complexes in addition to the TPL/HDA19 complex.

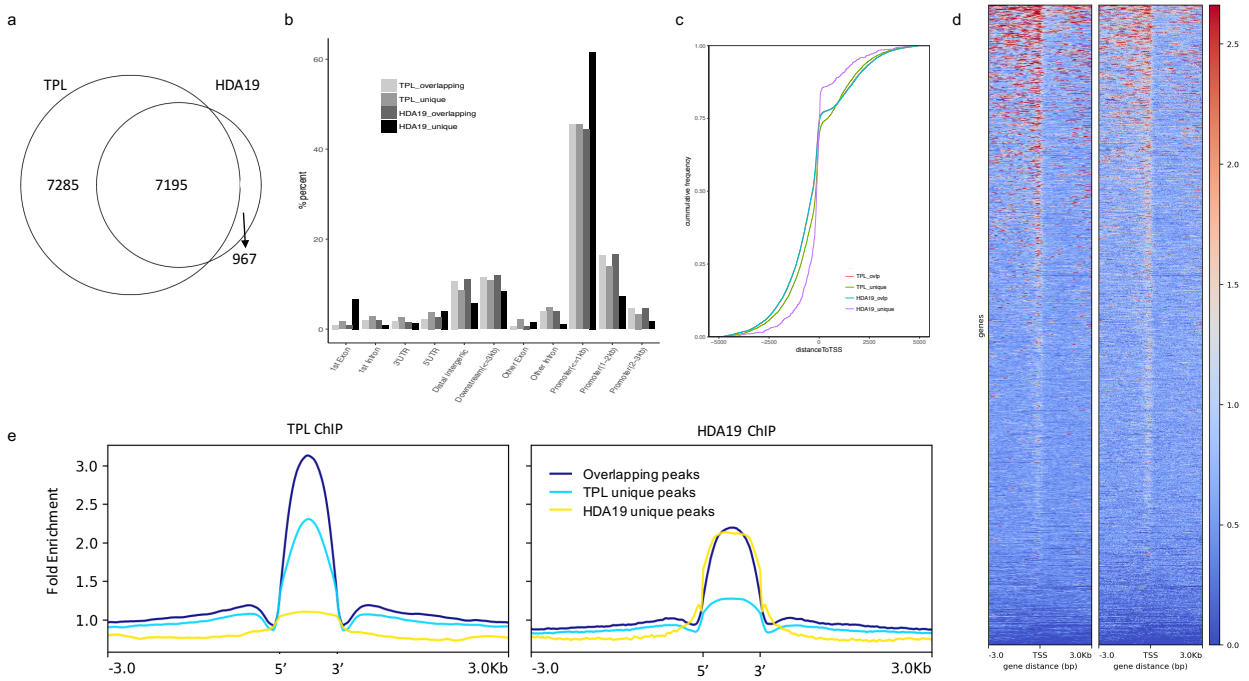


Fig 3.2: HDA19 overlaps with TPL. **a**, Venn diagram illustrating the overlap between peaks of TPL and HDA19. **b**, Annotations of TPL and HDA19 unique or overlapping peaks. **c,** cumulative distribution of TPL and HDA19 common and unique peaks around TSS. **d**, fold enrichment of TPL (left) and HDA19 (right) ChIP over input around TSS in the genome. **e,** ChIP fold enrichment of TPL (left) and HDA19 (right) around the indicated peaks.

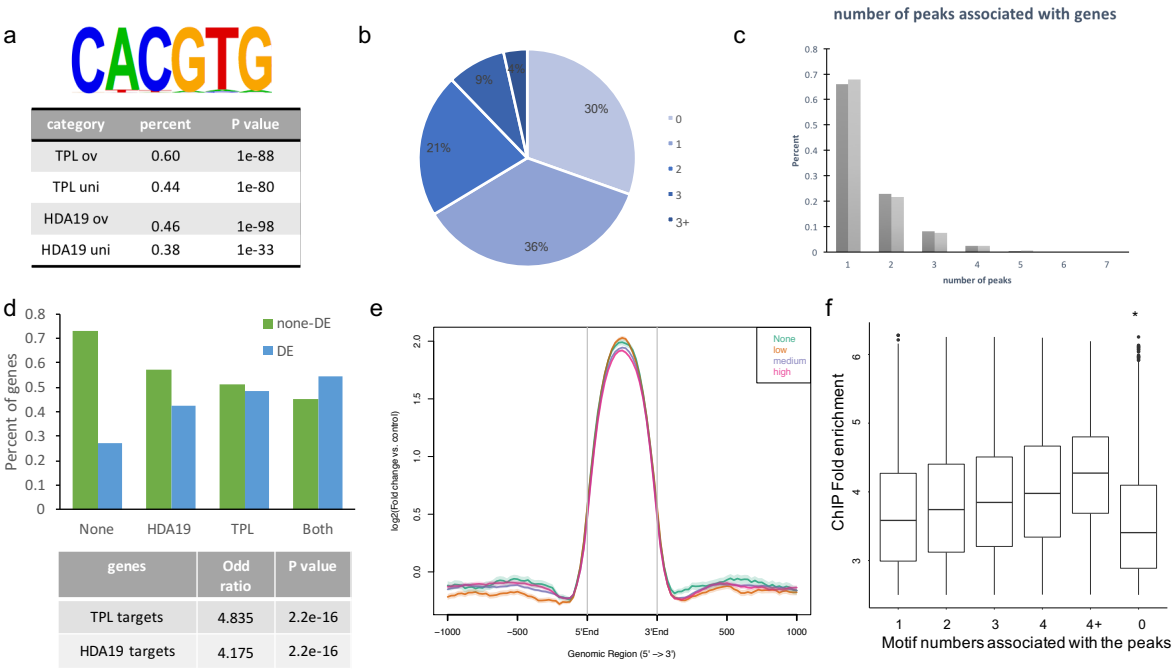
TPL and HDA19 bind to known motifs of varieties of TFs

TPL has been shown to interact with diverse TFs via their repressive EAR motifs (protein sequence: LXLXL) or B3 motifs (R/KLFGV)[197]. Therefore no consensus DNA sequence was found associated with TPL peaks, but a variety of motifs specific to different families of TFs were significantly enriched. Both TPL and HDA19 peaks are enriched for the G-Box motif (CACGTG, Fig 3.3a), a cis-element typically bound by basic helix-loop-helix (bHLH) proteins such as MYC2 and MYC3, which have been shown to form a repressor complex with JAZ,

NINJA and TPL in Jasmonic Acid signaling[45,198,199]. We have previously showed that the G-Box motif is also enriched within HD-ZIP III binding sites (chapter2 extended Fig 3m). WUS, the first TPL-interacting TF identified, interacts with the G-box as well and this motif confers the repressive function of WUS[35]. The fact that both TPL and HDA19 peaks display G-Box enrichment suggests the TPL/HDA19 co-repressor complex potentially endows WUS, possibly as well as HD-ZIP III and JAZ13-MYC2/3, the ability to repress transcription. TPL and HDA19 overlapping peaks were also enriched in the CG rich motif GGWCC (W: A or T). This motif was found in the binding sites of a non-canonical bHLH TF family named Teosinte branched1/Cycloidea/ Proliferating cell factors (TCPs), which have profound effects on the growth patterns of meristems and lateral organs through the crosstalk with hormone signaling[200]. It has been shown that TCPs interact with TPL via yeast two-hybrid (Y2H) and with EAR-motif containing adaptors by GST pull down[41,197,201], suggesting that TPL and HDA19 might also be recruited by TCPs to regulate hormone mediated patterning events.

TPL was also shown to interact with members from protein families such as APETALA2/ethylene response factor (AP2/ERF)-type transcription factors, C2H2 zinc finger proteins, MADS (short for MCM1, AGAMOUS, DEFICIENS, SRF) domain proteins, SQUAMOSA promoter binding (SBP) proteins, basic leucine zippers (bZIPs), MYB DNA-binding domain-containing proteins and WRKY domain transcription factor in a high throughput Y2H screen[197]. Correspondingly, we found that TPL peaks were also enriched in motifs overrepresented by transcription factors from those families, although to a less extent than the G-box and TCP motifs (supplemental Excel file 3.1). Those transcription factors with motifs enriched only at a small percentage of TPL peaks might contribute to the specificity of TPL.

Enrichment of multiple motifs was common at TPL peaks (Fig 3.3b), which suggests a possible role of TPL as a protein interacting hub for multiple TFs. Additionally, the fact that 70% of TPL peaks were enriched in TF binding motifs also indicates TPL binding is mostly restricted to where it is recruited, as spreading of TPL would deplete sequence specificity of most of the binding sites.



category	percent	P value
TPL ov	0.60	1e-88
TPL uni	0.44	1e-80
HDA19 ov	0.46	1e-98
HDA19 uni	0.38	1e-33

genes	Odd ratio	P value
TPL targets	4.835	2.2e-16
HDA19 targets	4.175	2.2e-16

Fig 3.3: Cis elements contribute to the enrichment of TPL. **a,** enrichment of G-box motif within both TPL and HDA19 peaks. **b,** number of motifs identified within each TPL peak. **c,** number of peaks associated with defined target genes. **d**, enrichment of differentially expressed (DE) genes in the HDA19 unique targets, TPL unique targets and shared targets (p value from Fisher's test). **e,** log2 (fold enrichment) of ChIP at TPL peaks grouped into categories based on the expression level of associated genes (None: not expressed; low/medium/high: lowly/medially/highly expressed). **f,** fold enrichment of TPL at peaks harboring different numbers of motifs (Wilcoxon test, peaks with no motifs enriched had significantly lower ChIP enrichment).

TPL and HDA19 tend to associate with dynamically expressed genes

To examine the effect of TPL and HDA19 binding on gene expression, we assigned the peaks to the nearest gene within 3kb. Almost 70% of the target genes had only one peak assigned to them while the rest had multiple peaks, usually as clusters of peaks at the promoter of the genes (Fig 3.3c and 3.1d). We found that TPL and HDA19 tended to bind to the promoters of expressed

genes. When we compared the expression level of those targets in different tissues, we found that TPL and HDA19 targets were more dynamically expressed in different tissue types, with much more percent of genes showing differential expression in different tissues, for example in seedlings versus in flowers (Fig 3.3d). The broad spectrum of expression levels of these targets may mask the effect of suppression by TPL or HDA19 in a specific subset of cells. Therefore, while it's possible just like *Gro* in Drosophila, TPL attenuates the transcription of those expressed genes[14], it is also possible that TPL and HDA19 regulate transcription by silencing the targets in specific cells, which is not distinguishable in data obtained from whole tissue.

In yeast, unrelated proteins with no biological relevance were shown to localize at the most highly expressed genes by ChIP methods[202]. Those spurious binding events are highly suggestive of a technical issue due to hyper-ChIPable open chromatin, raising the question whether ChIP profiles represent bona fide binding events. Among the TPL targets, binding events at relatively highly expressed genes are not likely due to the amenability of local DNA to ChIP, since the read enrichment of TPL at peaks that associated with highly expressed genes are not significantly higher than that at lowly or not expressed genes (Fig 3.3e). As a matter of fact, it appears that the existence of motifs but not open chromatin itself contributes to TPL binding, with motif-enriched peaks showing significantly more fold enrichment of TPL (Fig 3.3f). What's more, the overrepresentation of motifs in TPL peaks is also suggestive of functional binding.

TPL and HDA19 regulates a common group of genes

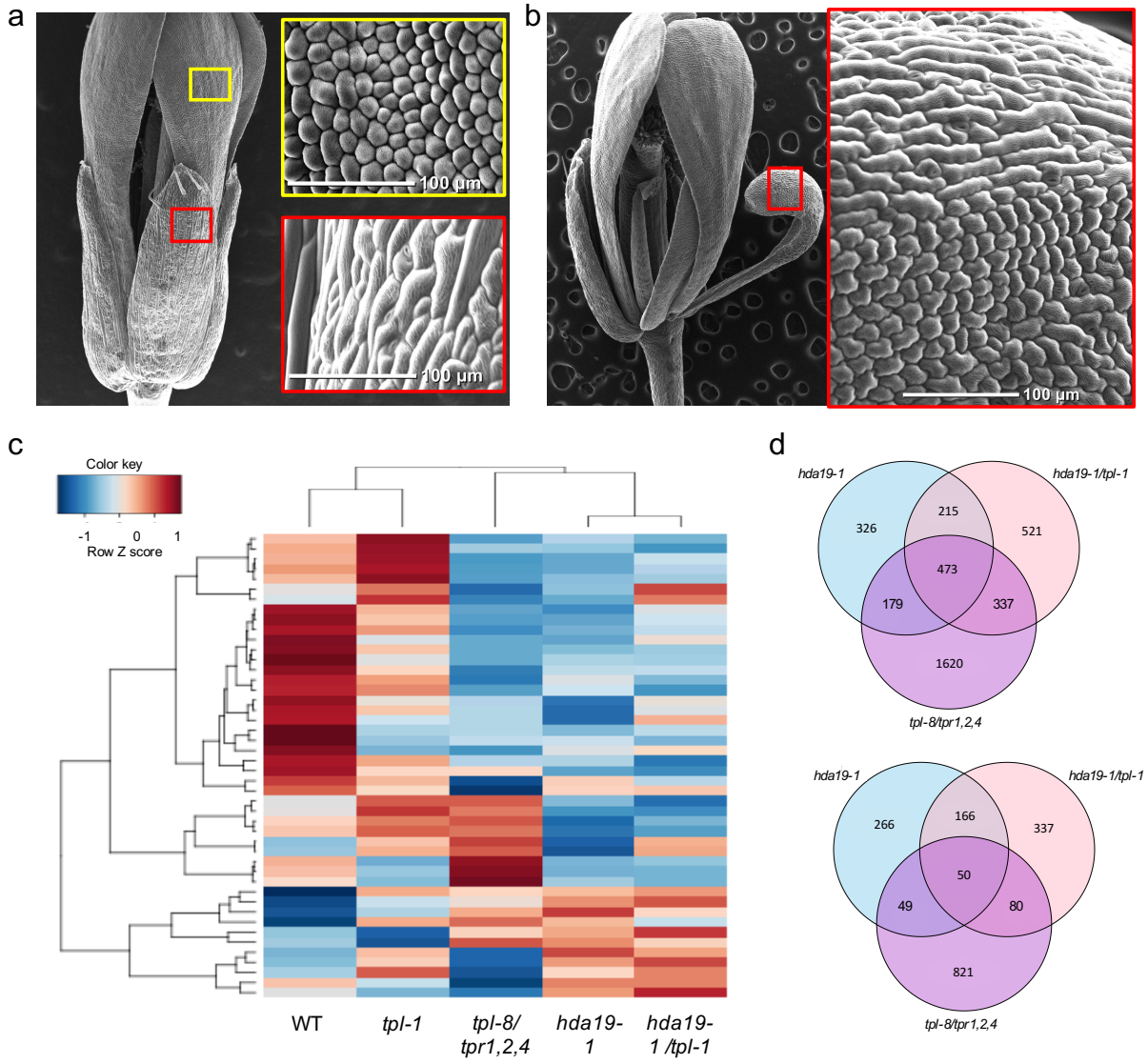


Fig 3.4: TPL and HDA19 regulate a common group of genes. **a**, a WT type flower with long rod-like sepal cells (in red) and round, uniform petal cells (in yellow). **b,** a flower of the *tpl-8/tpr1,2,4* quadruple mutant with sepal-to-petal homeotic transformation. **c,** heatmap of relative gene expression at genes that showed at least 1.5-fold change in all four mutants. Blue: relatively low expression. Red: relatively high expression. **d,** overlap of genes that significantly down(upper)- or up(lower)-regulated regulated by 1.5-fold in the mutants.

To further investigate the function of TPL and HDA19 in transcriptional regulation, we conducted RNA-seq with inflorescence meristems at the same stage as those in ChIP-seq experiments. *tpl-1* doesn't show much of a change in global gene expression (Fig 3.4c), which might explain the relatively weak floral defect in *tpl-1* mutants. We therefore performed RNA-

seq with *tpl-8/tpr1,2,4* quadruple mutants, which showed more severe phenotypes in the flower[27] (Fig 3.4b). A common group of genes were ectopically expressed in *tpl-8/tpr1,2,4* mutants and *hda19* mutants. Unexpectedly, we saw more genes downregulated than upregulated and more overlap in the genes downregulated in different mutant backgrounds (Fig 3.4c and 4d). However, in the group of genes that are upregulated in two or three mutants, significantly more of them have TPL peaks. For genes that are upregulated in all three mutants, almost all of them are associated with TPL peaks, while the commonly downregulated genes showed less than 30% association with TPL peaks, which is expected by chance in the genome. These results suggest a model where the genetic perturbation of TPL-HDA19 complex leads to the de-repression of only a core group of genes that are directly regulated by the complex, leading to the down-regulation of more downstream genes.

TPL and HDA19 regulates boundary specification genes in floral development

Combining ChIP-seq and RNA-seq data, we found the CUP-SHAPED COTYLEDON (CUC) transcription factors were directly regulated by TPL in different modes with HDA19 (Fig 3.5). The three members in this family, CUC1, 2 and 3, are important regulators of organ boundary formation both during embryogenesis and in post-embryonic development[203-209]. TPL has been shown to suppress CUC3 expression with BES1, a master regulator of the brassinosteroid (BR) signaling pathway[47]. Consistent with this, we saw clear peaks of TPL both at the promoter and the second intron of *CUC3*, and an increase in RNA level in *tpl-8/tpr1,2,4* quadruple mutants (Fig 3.5c). *CUC1* and *CUC2* were both repressed by TPL too, with TPL bound at the promoters and an elevation in the expression level in quadruple mutants (Fig 3.5a and b). In *tpl-1* mutants,

only *CUC2* showed a significant increase in expression, which suggests the *tpl-1* allele is only partially dominant to TPRs in repressing *CUC*s in the inflorescence.

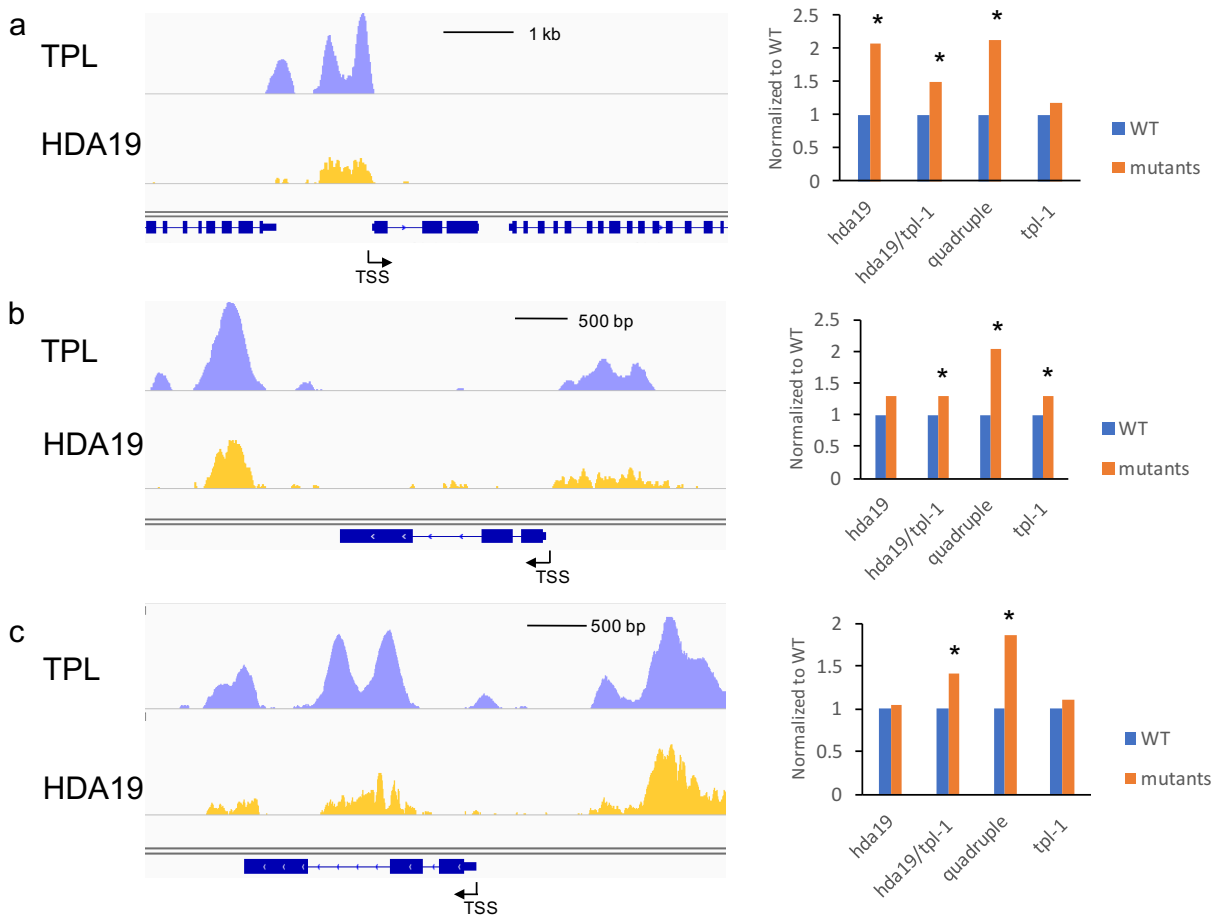


Fig 3.5: TPL regulates the expression of CUCs with HDA19..Snapshots of TPL and HDA19 ChIP signal enrichment (left) and bar charts of expression changes in the indicated mutants (right) at: **a**, *CUC1*, **b**, *CUC2* and **c**, *CUC3*.

Interestingly, the effect of a mutation in *hda19* on the expression of different *CUC*s was divergent. *CUC1* had HDA19 bound to the region where TPL peaks were located, and was activated in *hda19* and *hda19/tpl-1* mutants (Fig 3.5a). The fact that mutations in either *HDA19* or *TPL/TPRs* led to an increase in transcription indicates that both TPL and HDA19 are indispensable for the repression of *CUC1*. *CUC2* had only negligible enrichment of HDA19 at the promoter, suggesting a unique repressive role of TPL. Indeed, only in *hda19* was *CUC2* not

significantly upregulated (Fig 3.5b). With a second HDA19 and TPL overlapping peak in the intron, *CUC3* seemed to be regulated at an additional level. *CUC3* was only upregulated in *hda19/tpl-1* double mutants and *tpl-8/tpr1,2,4* quadruple mutants (Fig 3.5c). It's a possible *CUC3* is repressed by two mechanisms, one unique to TPL and the TPRs that *tpl-1* is dominant negative to, and the other dependent on the synergistic action of HDA19 and the TPRs that are resistant to *tpl-1*. *CUC1* and *CUC2* are under post-transcriptional regulation by microRNA miR164[205]. It has been reported that plants expressing a miR164-resistant *CUC2* gene displayed increased sepal spacing due to an increase size of boundary regions[205]. This was also seen in the *tpr* quadruple mutants, which had a dramatic increase in *CUC2* mRNA (Fig 3.5b and 6f).

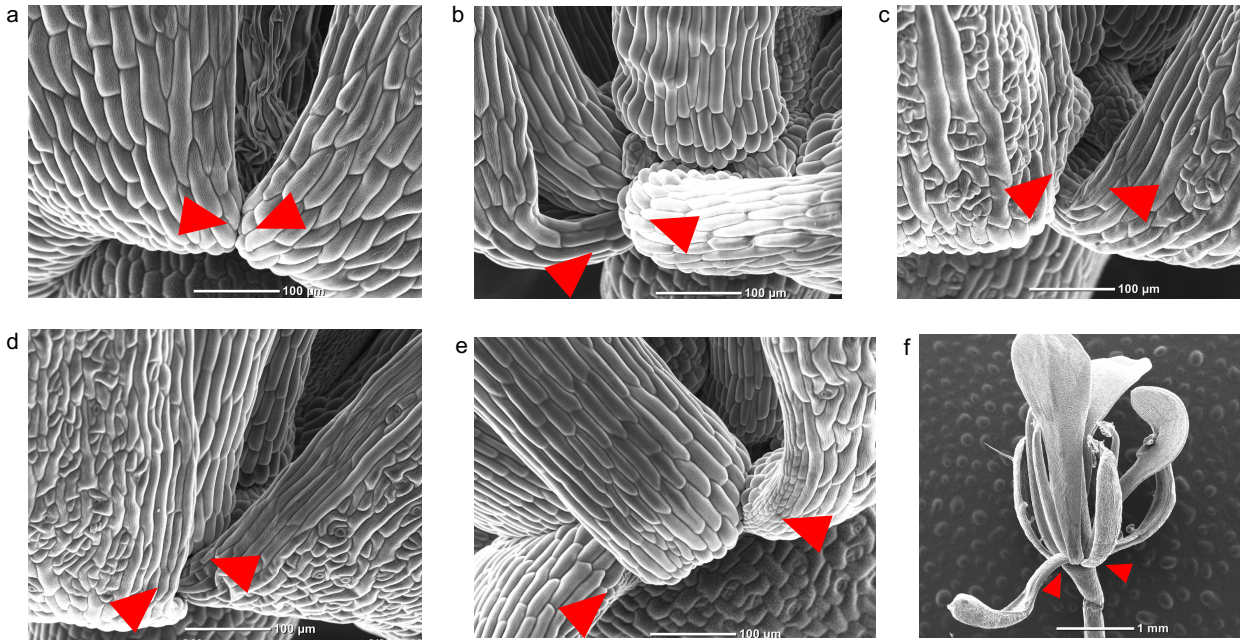


Fig 3.6: Floral defects in *tpl* and *hda19* mutants. Scanning electron microscope (SEM) images of sepal-to-sepal arrangements of flowers in **a,** WT, **b**, *tpl-1*, **c,** *hda19-1*, **d**, *tpl-1/hda19-1*, **e and f**, *tpl-8/tpr1,2,4.* Red arrows point to the connection between sepals.

In addition to enlarged sepal boundary in the quadruple mutants and the sepal-to-petal homeotic conversion, we observed a distinct pattern of sepal arrangement in all the mutants (Fig 3.6a-f). Sepals in WT plants were oriented in the first girdling whorl, with two adjacent sepals in contact

at their bases (Fig 3.6a). Sepals in the mutants, however, did not appear to be spaced alongside each other. Instead, the bases of sepals protruded towards the inner whorl, disrupting the whorled phyllotaxy (Fig 3.6b-e).

Plants undergo changes in phyllotaxy pattern during development. During leaf primordia and floral meristem formation, *Arabidopsis* displays a spiral phyllotaxy with successive primordia emerging approximately 137 degrees radially from each other. *Arabidopsis* transitions from spiral to whorled phyllotaxy in floral organ development, when the initiation of floral organs within the same whorl takes place synchronously. Upon the establishment of dome-shaped floral meristems, sepal primordia arise in a continuous spiral sequence, with the emergence of abaxial primordium first, followed by the adaxial and then the two lateral primordia. The initiation of sepal primordia takes place so fast in *Arabidopsis* (within 18 hours) that the four sepals appear in one whorl of the flower[210]. *tpl* and *hda19* mutants displayed delayed growth compared to WT plants, which might allow the spiral pattern of sepals to be observed. The overexpression of *CUC*s in the mutants might account for this phenomenon, for CUCs have been shown to inhibit cell division and elongation in organ boundaries[205,206,209]. *tpl-8/tpr1,2,4* quadruple mutants often displayed homeotic transformation of sepals into carpelloids and decreased number of petals, suggestive of ectopic expression of multiple floral homeotic genes, which in combination with delayed cell proliferation and elongation, may contribute to the rare cases in which two neighboring sepals were separated by a petal in the middle in a spiral pattern (Fig 3.6f).

Discussion

In this chapter, we characterize the function of TPL as well as HDA19 in the context of development. We show that TPL binds to clusters of discrete narrow regions in the genome without spreading into multi-kilo base domains, resembling its counterparts in animals[13,14]. The proposed "long range" repression model was based on the observations that Gro/TLE type co-repressors located beyond the regions where the protein is recruited[7,8,11,12]. In *Arabidopsis*, we seldom see this phenomenon. Instead, we find that TPL binding is highly centered around the core sequence bound by TFs that recruit them, which suggests TPL mainly functions locally without extending to distal regions.

TPL forms co-repressor complexes HDA19 at half of the TPL targets. The two proteins co-localize at promoters of thousands of genes in *Arabidopsis*. Those genes are highly enriched in genes that have a broad spectrum of expression levels. Therefore, we reason that the observation of co-repressor binding at expressed genes is a result of the high expression level in a subset of cells masking the repressive function of the co-repressor in other cells. In that case, the broad expression pattern of target genes will serve as an indication that transcriptional regulation by TPL is a highly dynamic process. Nevertheless, TPL may play a similar role as Gro in attenuating transcription rather than silencing genes[14]. In addition, only a subset of genes that are associated with TPL or HDA19 peaks showed an increase in the mutants. While redundancy and extra factors may account for this, the fact that the data are in multiple types of organs might also largely lead to a poor correlation, with transcriptional regulation in different directions in different cells canceling each other out. Tissue- specific ChIP-seq and RNA-seq need to be performed to clarify these questions and concerns.

We also show different genetic interactions of *TPL* and *HDA19* in regulating the expression of *CUC1,2* and *3*, three closely related NAC domain-containing transcription factors. We see a sepal separation phenotype in the quadruple mutant, resembling that in the miR164-resistant *CUC2* mutant. The quadruple mutant of *tpl-8/tpr1,2,4* was the only mutant that displayed increased mRNA level at all three *CUC*s and had the highest fold change in transcription levels, which suggests that TPL/TPR is the fundamental component of the repressive machinery in regulating this family of transcription factors and that the mutant phenotype may be dosage-dependent. Another possible explanation for the quadruple mutant being the only one displayed boundary defect phenotype is that *CUC*s are under the regulation of multiple layers including miR164 and SHOOT MERISTEMLESS (STM) and consequently have highly specific pattern[203,205,211]. To address that, in situ hybridization will be performed.

We find that all *tpl* and *hda19* mutants exhibited spiral orientation instead of whorled organization of sepals. We reason that the ectopic expression of *CUC*s can possibly inhibit cell division and cell growth, leading to the emergence of spiral phyllotaxy of the sepals, which is insignificant in WT flowers where the initiation of sepal primordia takes place transiently. To confirm this, cell division marker H4 will be used for in situ to label cell division event and cell sizes will be quantified.

ChIP has been an effective tool to study the molecular mechanism of transcriptional regulation. Plants have rigid cell walls, high levels of secondary compounds, and large vacuoles in cells, adding difficulty to performing ChIP with transcription factors and co-factors, whose association with DNA might be weak or indirect. Here we successfully profiled genome-wide TPL and

HDA19 binding in *Arabidopsis* inflorescence. Our high-quality data are a good resource for the study of transcriptional regulation in plants.

Materials and Methods

Growth Conditions and mutant alleles

Seeds were stratified for at least 3 days at 4°C in the dark prior to sowing on soil. The controls used were all wild-type plants in the same ecotypes to the mutants or the transgenic plants for each experiment. Plants were grown at 21°C under 16-hour light/8-hour dark cycle in highly controlled growth chambers. Inflorescence tissue with unopened buds were collected from those that already had 6 open flowers.

Recombineering

Transgenic lines with fluorescence protein fusions were constructed using Recombineering method as previously described[212]. For ChIP, 2xYPet-3xHA were fused to the C terminus of the genes. Each construct was transformed into the respective loss of function allele via the *Agrobacterium* floral dip method[213].

SEM

Plants were grown side by side. Flowers with sepal-to-petal homeotic transformation were imaged directly by Nikon JEOL neoscope without fixation.

ChIP-seq

ChIP experiments were performed as previously described[27] with the following modifications. 2 grams of tissue from Transgenic homozygous lines were vacuum crosslinked with 1% Formaldehyde in PBS for 2 hours at 4 °C and vacuum quenched with 0.125M glycine in PBS for 20 minutes. Protein A/G magnetic beads (Pierce 88803) were used with anti-GFP antibody (abcam ab290) in immunoprecipitations. Immunoprecipitated DNA was purified using AmpureXP beads and libraries were constructed using the Nugen Ovation Ultralow Library System according to the manufacturers protocol. Input DNA or Ler ChIP DNA was used as a control. Libraries were sequenced 50bp with single end on an Illumina HiSeq 2000. Reads were mapped to the TAIR10 *Arabidopsis* genome using bowtie2. Uniquely mapped reads were retained for peak calling using the MACS2[214] and for metaplots. Each ChIP contained at least 2 biological replicates. MACS2 was run using default settings with a q value cutoff of 0.01 and peaks with at least two-fold enrichment versus input were selected. De-novo motif discovery was conducted using HOMER (http://homer.ucsd.edu/homer/) with the findMotifsGenome function. Sequences 200bp surrounding the summits of peaks were used. In some cases, only the top 1000 high quality peaks with the lowest q-value were used. Heat maps and metaplot profiles of relative enrichment (log2 ChIP/Input) at defined genomic regions and were generated by deeptools[215]. BigWig files of the mapped reads were visualized by the Integrative Genomics Viewer (IGV)[216,217]. Genes were annotated by ChIPseeker with a 3kb cutoff[218].

RNA-seq

100mg of seedlings or inflorescence from homozygous mutants as well as WT were harvested and ground using liquid N2. Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) followed by DNase digestion (Promega) following the manufacturers' protocols. Libraries were

constructed with the TruSeq RNA Library Preparation Kit v2 (Illumina) as per manufacturer's instruction. Samples were sequenced on an Illumina HiSeq 2000 (single-end 50-bp run). Raw reads were mapped to the TAIR10 *Arabidopsis* genome using TopHat with standard parameters except for intron length of 40). Uniquely mapped reads were then used for Cufflinks for differential expression analysis[219].

Chapter 4: The involvement of H3K27me3 and PRC complexes in TPL mediated transcriptional regulation

Introduction

Active repression of genes by the Gro/Tup1 family of co-repressors has been shown to regulate multiple cellular and developmental processes[2 3]. Gro/Tup1 co-repressors in animals and fungi are characterized by same structural organization, with an N-terminal domain important for self-association as well as binding to specific transcription factors, and a set of WD40 repeats at the C-terminus, which mediates the physically association with most of the interacting transcription factors. The canonical model of transcriptional repression by Gro/Tup1 type of co-repressors is that they are recruited by transcription factors via intrinsic repression domains and in turn recruit class I histone deacetylase Rpd3/HDAC1 to repress transcription[220]. TPL is a Gro/Tup1 like co-repressor in *Arabidopsis.* It is a master regulator involved in a number of processes including apical-basal embryonic patterning, floral development, hormone response, and circadian clock regulation, to name a few[27,33,40,44,45]. TPL also has been shown to interact with a wide range of repressors and repress transcription by forming a complex with histone deacetylase[27,41,43,44,197,201,221,222]. However, unlike in animals and fungi, all the protein-protein interactions characterized so far are conferred by the N terminus of TPL, while the function of its C-terminal WD 40 repeats still remains to be uncovered.

Active repression by the epigenetic machinery also controls multiple aspects of developmental processes. Among them, the evolutionarily conserved Polycomb group (PcG) proteins have been shown to regulate tissue-specific gene expression and are essential for the normal development

of multicellular organisms[115-117,119]. Polycomb Repressive Complex 1 (PRC1) and PRC2 are two of the most well-characterized PcG protein complexes. Histone H3 lysine 27 tri-methylation (H3K27me3) play a crucial role in transcriptional repression mediated by PcG proteins, with PRC2 catalyzing the tri-methylation of K27 while PRC1 recognizing H3K27me3 and subsequently ubiquitinating Histone H2A to further condense the chromatin. To ensure precise regulation of genes at different times in different cells, the recruitment of PRC1 and PRC2 needs to be tightly controlled. The mechanism of recruitment was first characterized in *Drosophila*, where PRCs are recruited to Polycomb response elements (PREs)[130]. In mammals, un-methylated CpG islands and Long non-coding RNAs (lncRNAs) play a major role in genome-wide targeting of PcG[133-137]. In *Arabidopsis*, extensive efforts have been made to search for the recruitment mechanism of PcGs. While PREs and lncRNAs both were reported to play a role[138-141], it is still unclear how PRC1s and PRC2s crosstalk and integrate information from the above-mentioned machineries to shape the epigenetic landscape during development.

Here we show that K27 tri-methylated histone H3 can be recognized by the C terminal WD40 repeats of TPL and therefore can enhance TPL's association with the chromatin after the initial recruitment by transcription factors. TPL interacts with a proposed PRC1 component - Like Heterochromatin Protein 1/Terminal Flower 2 (LHP1/TFL2), co-localizes with a PRC2 component - CURLY LEAF (CLF) and affects H3K27me3 level in the genome. Our data reveal a novel layer of transcriptional repression mediated by the interplay between histone modification and Gro/Tup1 type co-repressors and suggest a possible role of co-repressors in the function of PcGs.

Results

TPL interact with histone H3 and H3K27me3 through its WD40 repeats

TPL contains two sets of WD40 repeats at the C terminus (Fig 4.1c). The WD40 repeat is a short structural motif typically composed of 40 amino acids (aa), which preferentially terminates with a tryptophan (W) and an aspartic acid (D). Most WD40 repeats serve as a scaffold for protein interactions[223] and it has been shown that chromatin readers such as WDR5 can interact with histones and modified histones via its WD40 domain[224,225]. To investigate if the WD40 repeats of TPL recognize chromatin modifications, we purified recombinant protein from *E.coli* that comprises the C terminus of TPL fused to Glutathione-S-Transferase (GST). The recombinant protein was hybridized to a histone peptide array that contains multiple combinations of modifications. We found that the C terminus of TPL bound to unmodified histone H3 from aa 16 to 35, and that it showed a higher affinity to K27 tri-methylated H3, a hallmark of a silent chromatin state (Fig 4.1a). We also found that the recombinant proteins with only the first set (Fig 4.1c, 1d) or the second set (data not shown) of WD40 repeats lost the ability to bind to K27 tri-methylated H3, suggesting the interaction with H3K27me3 is dependent on the cooperation of both sets of WD40's.

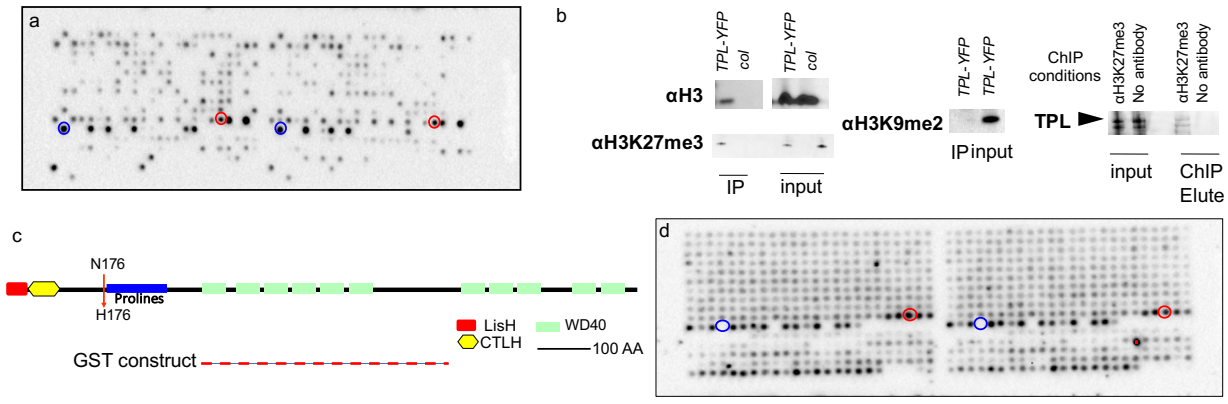


Fig 4.1: TPL interact with histone H3 and H3K27me3 through its WD40 repeats. a, The WD40 repeats of TPL bind to histone H3 and K27 tri-methylated histone H3 peptides in vitro. Red: unmodified

histone H3(aa16-35), blue: K27 tri-methylated H3(aa16-35). **b,** TPL pulls down histone H3 and K27 tri-methylated histone H3(left), but not K9 di-methylated H3(middle) *in vivo,* and TPL is associated with the DNA-protein complexes pulled down in the H3K27me3 ChIP (right). Co-IP was conducted with anti-GFP beads. *col*: WT plants without TPL-YFP transgene. **c,** Diagram of TPL structure and illustration of the GST construct for the 1st set of WD40 repeats. **d,** the 1st set of WD40 repeats lost binding specificity to H3K27me3.

To examine the interaction *in vivo*, we performed Co-immunoprecipitations (co-IP). We found that TPL was able to precipitate histone H3, and that H3K27me3 but not H3K9me2, another repressive mark in *Arabidopsis*, were effectively pulled down by TPL, suggesting the interaction we saw is specific to H3K27me3 and is not the mere result of general association with repressive chromatin (Fig 4.1 b). We also conducted Chromatin immunoprecipitations (ChIP) and found TPL presented in the chromatin marked by H3K27me3, which further confirmed the interaction between TPL and H3K27me3 (Fig 4.1 b).

TPL co-localizes with H3K27me3 genome-wide

To characterize the association of TPL to H3K27me3 marked chromatin, we performed ChIP followed by deep sequencing (ChIP-seq) on TPL and H3K27me3. Our data revealed a significant overlap between TPL peaks and H3K27me3 regions in the genome (Fig 4.2a, 2b and 2c). Over 40% of H3K27me3 regions showed TPL occupancy (Fig 4.2c). At those regions, we found that TPL tended to co-localize with H3K27me3 regions at promoters but was absent from gene bodies that were heavily marked by H3K27me3 (Fig 4.2a and 2d). This finding is in consistence with the fact that TPL is recruited by transcription factors to promoters but not gene bodies.

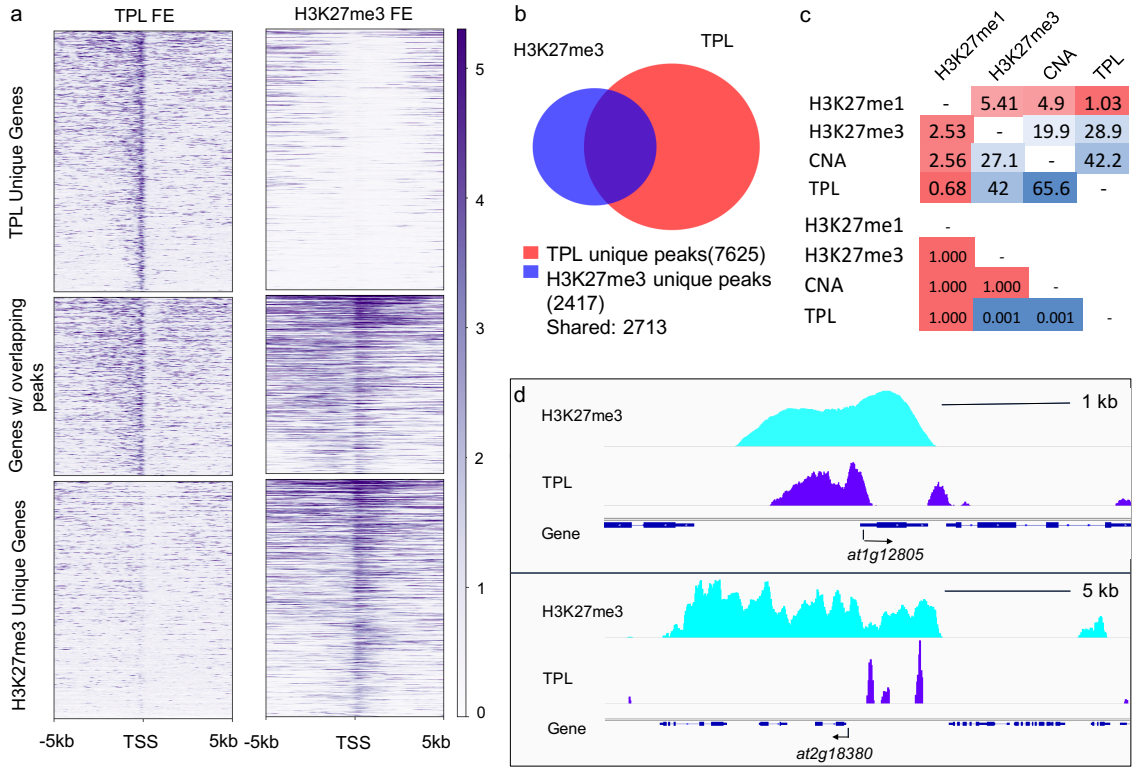


Fig 4.2: genome-wide localization of TPL and H3K27me3. **a,** ChIP/Control fold enrichment of TPL (left) and H3K27me3 (right) along the Transcription Start Site (TSS) **b,** Venn diagram of TPL (right) and H3K27me3 (left) peaks. **c,** Heatmap of the percent (upper) and p-value (lower) of overlap between TPL and different factors. H3K27me1: a control histone modification. CORONA (CNA): a control transcription factor known to interact with TPL. **d,** TPL (magenta) and H3K27me3 (Cyan) fold enrichment at co-localizing sites.

Mutations in PRC1/2 complexes enhance *tpl-1* phenotypes

H3K27me3 is catalyzed by the PRC2 complex. The *Arabidopsis* genome encodes three homologous genes for the catalytic core of PRC2, *MEDEA (MEA)* in the gametophyte, and the partially redundant *CLF* and *SWINGER (SWN)* in the sporophyte. *clf* mutants displayed an early transition to reproductive growth and reduced H3K27me3 level whereas *swn* mutants showed no obvious phenotypes[226]. It is therefore proposed that CLF plays a dominant role in sporophyte growth. H3K27me3 is recognized by PRC1, which further contributes to the compaction of the chromatin by H2A ubiquitination. In *Arabidopsis*, the PRC1 component LHP1 reads the H3K27me3 mark and interacts with PRC2 components[125,126]. Both CLF and LHP1 have been

shown to be important for the spreading of H3K27me3 at *FLOWERING LOCUS C (FLC)*[227]. We found that CLF and LHP1 proteins were widely accumulated in the embryos (Fig 4.3a and 3b), where TPL is highly expressed in every cell. We therefore explored the biological relevance between H3K27me3 and TPL using the *clf* and *lhp1* mutants.

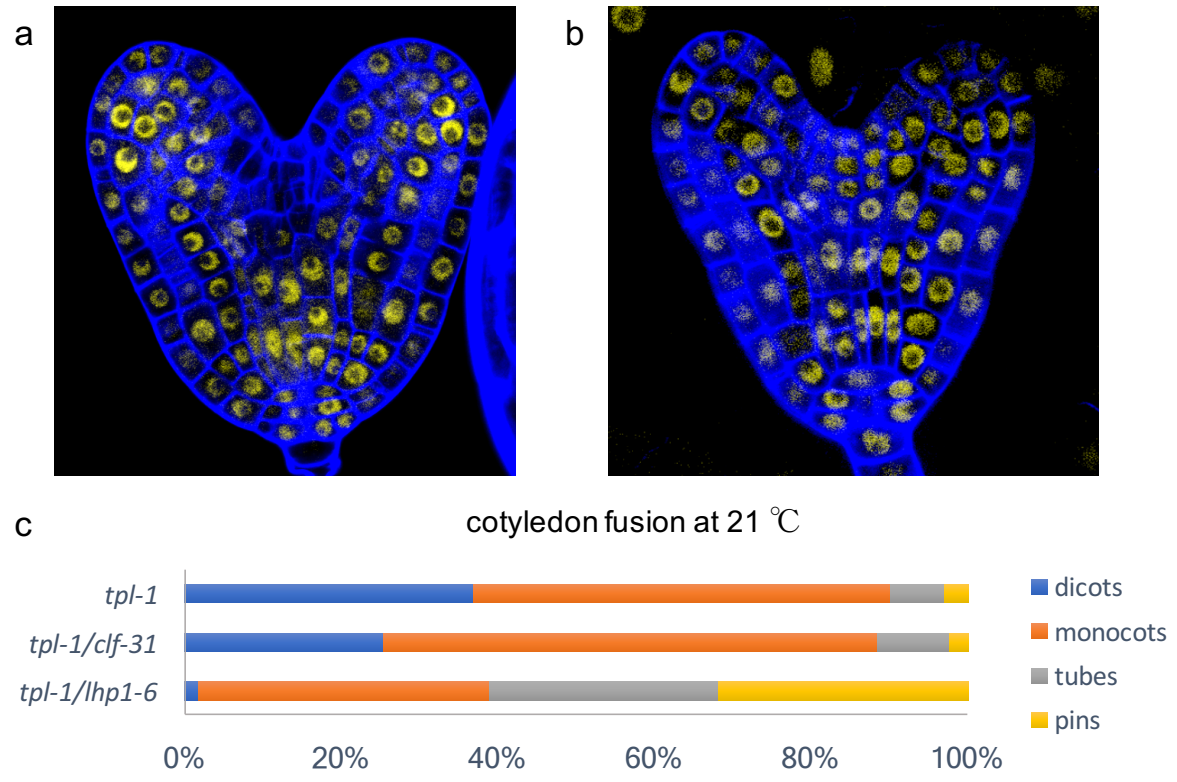


Fig 4.3. ***clf-31* or *lhp1-6* loss of function mutation enhances the *tpl-1* phenotype**. Confocal images of **a,** *CLFg-2xYPet-3xFLAG* and **b,** *LHP1g-2xYPet-3xHA* at the heart stage of the embryo (In blue: cell wall; in yellow: accumulation of the YPet labeled protein) **c,** Total number of seedlings assayed*: tpl-1* 274, *clf-31* 200, *lhp1-6* 200, *tpl-1/clf-31* 209, *tpl-1/lhp1-6* 226. All 200 *clf-31 or lhp1-6* plants assayed have dicots.

TPL plays a pivotal role in apical-basal patterning during embryogenesis. Plants carrying the dominant negative allele of TPL, *tpl-1*, display defects in shoot apical meristem formation, with a transformation of the shoot into a range of structures including, from less severe to more severe, a monocot, a tube, a pin and a second root[33]. The penetrance of the phenotype is temperature sensitive, with plants showing the most severe phenotypes at the restrictive temperature of

29˚C[33]. Mutations in *CLF* and *LHP1*, had no effect on the establishment of apical-basal axis alone. However, they enhanced the phenotype when in combination with *tpl-1* (Fig 4.3c), suggestive of a functional connection between TPL and H3K27me3. Strikingly, *tpl-1/lhp1-6* double mutant displayed a much more severe phenotype, with more than 95% of the seedlings displayed a mutant phenotype even at a permissive temperature of 21˚C, which suggests LHP1 is actively involved in the function of TPL during embryogenesis.

Reduced TPL binding and H3K27me3 level in *clf-28* mutant

The WD40 repeats of TPL showed higher affinity towards K27 tri-methylated H3 peptides (supplemental Excel sheet 4.1), leading to the question whether H3K27me3 affects the function of TPL by recruiting TPL or by enhancing TPL binding at its targets. To address this, we performed TPL ChIP-seq in *clf-28* mutant. We rarely saw a dramatic decrease or complete loss of H3K27me3 in the mutant. Instead, we observed an overall decrease in the width of H3K27me3 peaks in *clf-28*, suggesting a role of CLF in the spreading of H3K27me3 genome-wide (Fig 4.4a and 4b). In *clf-28*, TPL peaks that overlapped with H3K27me3 regions showed a significant decrease in ChIP signal enrichment compared to TPL unique peaks (Fig 4.4b and 4c). While TPL fold enrichment decreased accompanying the decrease in H3K27me3 level, TPL still showed mild enrichment even at the overlapping peaks that lost H3K27me3 (Fig 4.4a). *TPL* displayed a slight increase in expression level in *clf-28* (Fig4.4d). Therefore, the decrease of TPL ChIP signal was not likely due to reduced TPL. In fact, we found that TPL was less tightly associated with chromatin in the mutant (Fig 4.4e), which may reflect the reduced occupancy of TPL. These findings support the hypothesis that H3K27me3 strengthens the association of TPL with its targets after the initial recruitment.

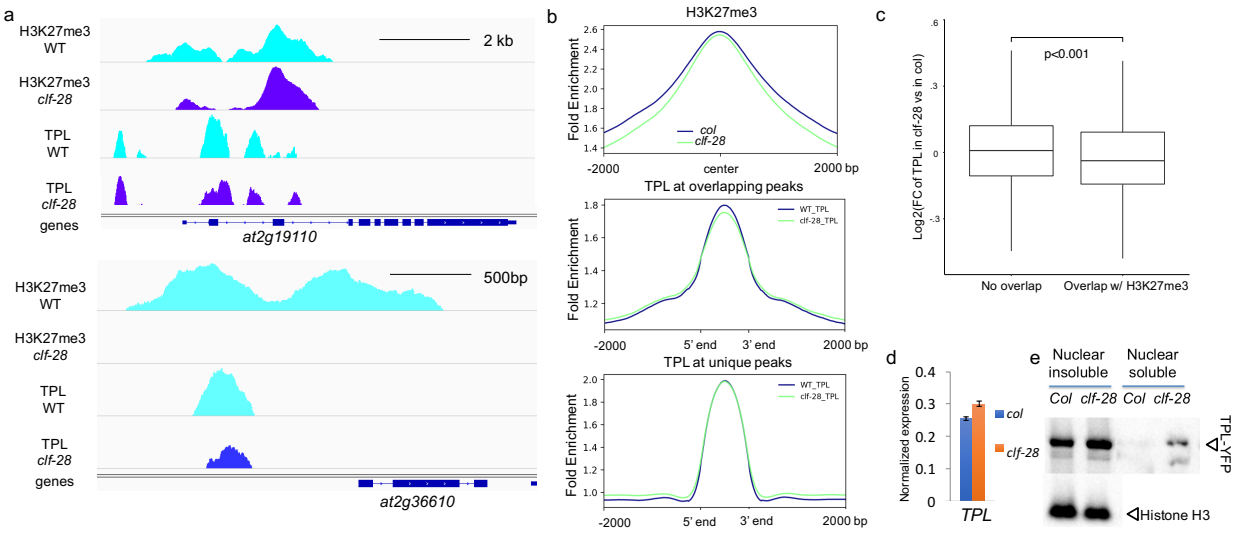


Fig 4.4: **Reduced TPL binding and H3K27me3 level in *clf-28* mutant**. **a**, Snapshots of H3K27me3 and TPL read enrichment change in WT(cyan) and *clf-28*(magenta). **b,** Average plots of H3K27me3 fold enrichment (top), and TPL fold enrichment at overlapping region(middle) and TPL unique peaks(bottom) in WT(blue) and *clf-28*(green). **c,** boxplot of log2 fold change of TPL at reduced(Y) or none-reduced(N) H3K27me3 peaks. **d,** *TPL* expression is not reduced in the *clf-28* mutant. **e,** TPL is less tightly associated with chromatin in *clf-28*.

TPL cooperatively regulates transcription with PcGs and H3K27me3

To further characterize the functional interaction between TPL and H3K27me3, we performed ChIP-seq with LHP1 and CLF. Consistent with LHP1's role in recognizing H3K27me3, the majority of LHP1 peaks overlapped with H3K27me3 regions (Fig 4.5a and 5b). CLF, however, showed only partial overlap, with about 50% of peaks unique to H3K27me3 regions (Fig 4.5a and 5b). Surprisingly, we found that both LHP1 and CLF peaks largely overlapped with TPL binding sites, which suggests TPL may form repressive complexes with PcGs. We performed Co-IPs and found TPL could be pulled down by LHP1 (Fig 4.5c). The physical interaction between TPL and LHP1 might explain the strong enhancement of *tpl-1* phenotype by *lhp1-6*. We failed to detect an interaction between TPL and CLF in our Co-IP, which was possibly due to the relatively low protein level of CLF in the plants.

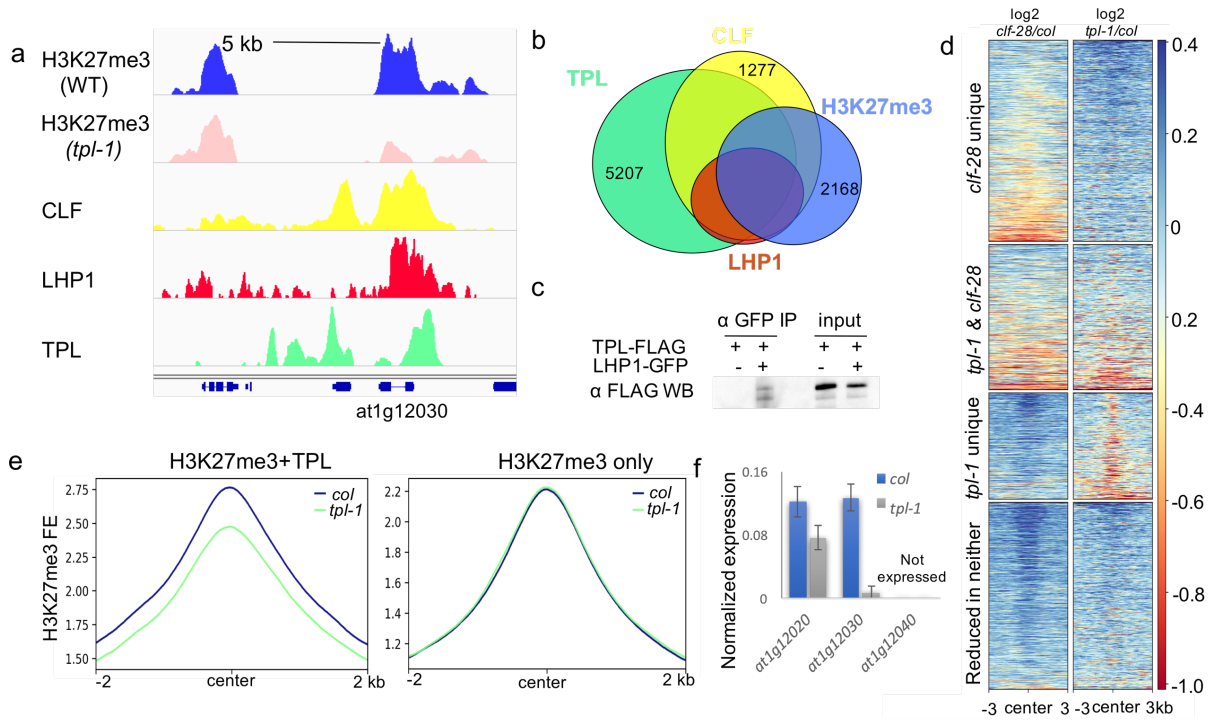


Fig 4.5: **TPL co-localizes with PRC components and affects H3K27me3 level**. **a,** A snapshot of H3K27me3 (WT: blue, *tpl-1*: pink), CLF (yellow), LHP1 (red) and TPL (green) fold enrichment at a co-localizing site. **b,** Venn diagram of peak overlapping. **c,** LHP1 pulls down TPL in vivo. **d,** Heatmap of log2 fold change of H3K27me3 in *clf-28* (left) or *tpl-1* (right) vs in WT at peaks that has reduced H3K27me3 level in the indicated genetic background. **e,** Average plot of H3K27me3 enrichment at TPL overlapping (left) or H3K27me3 unique peaks (right) in WT (blue) and *tpl-1* (green). **f,** Expression of the gene indicated in A and its neighboring genes.

The connection between TPL and PcG proteins raised the question whether TPL itself played a role in H3K27me3 deposition. We therefore examined H3K27me3 levels in the *tpl-1* mutant. Intriguingly, the H3K27me3 level in *tpl-1* mutants showed similar change to that in *clf-28*, with narrower H3K27me3 peaks at regions where TPL was bound in wild-type (Fig 4.5a and 5e). In *tpl-1* as many as 880 genes were significantly upregulated by at least 1.5 fold (Fig 4.6a). Although some genes with reduced H3K27me3 level were activated, many genes associated with decreased H3K27me3 showed no difference at RNA level and in rare cases, the transcription was downregulated (Fig 4.5f), which suggests a reduction of H3K27me3 in *tpl-1* is not a passive result of transcriptional activation but is due to a disruption of the function of TPL protein itself.

More peaks showed a reduction in H3K27me3 signal and more than 50% of those peaks that showed reduced H3K27me3 in *tpl-1* also displayed decreased enrichment in *clf-28* (Fig 4.5d), which indicates CLF may play a more important role than TPL in the spreading of H3K27me3.

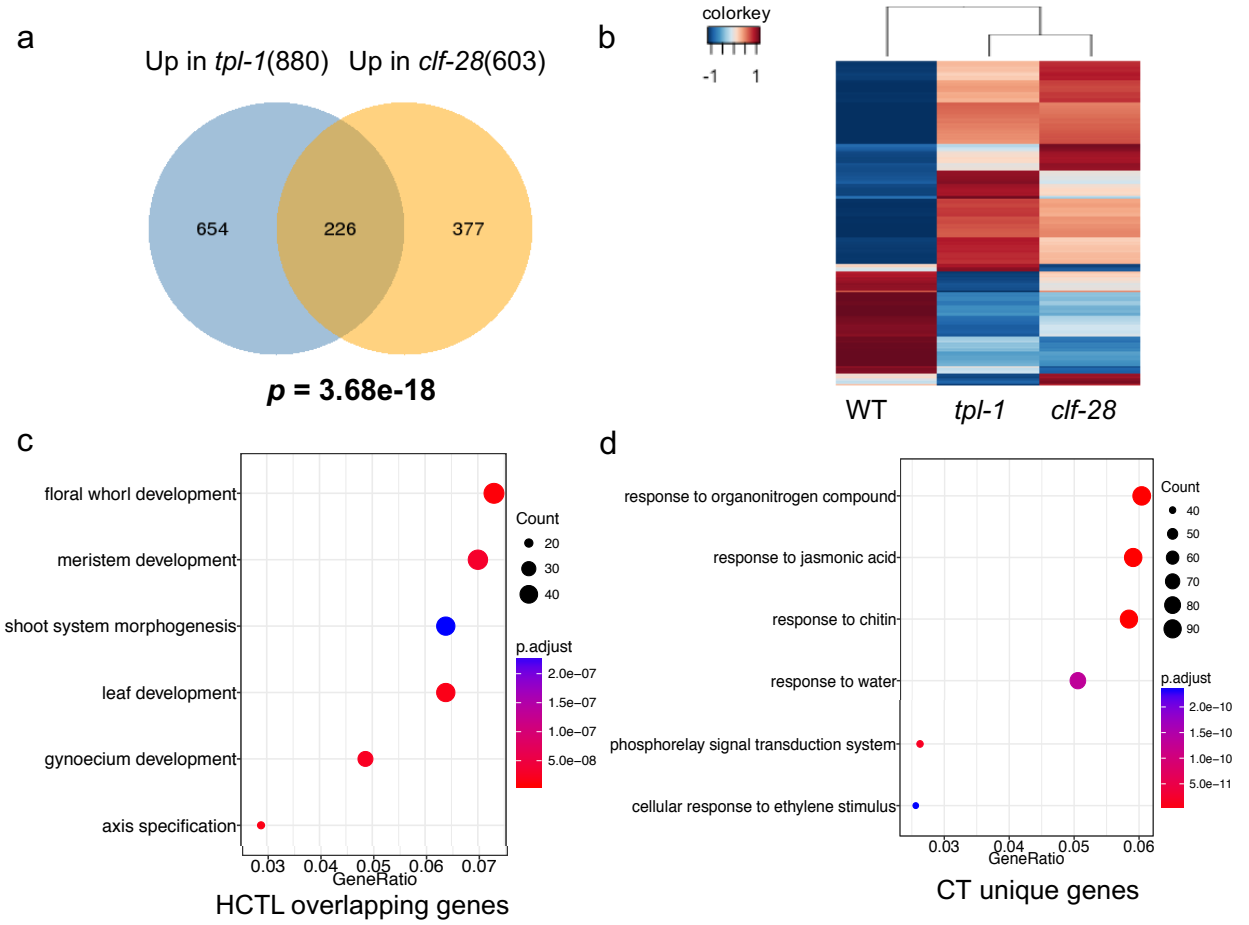


Fig 4.6. TPL cooperatively regulates transcription with PcGs and H3K27me3. a, Venn diagram of genes that are up-regulated in *clf-28* and/or in *tpl-1*. **b**, Heatmap of z-score transformed expression level of genes that showed significant changes (q_value<0.05) in *tpl-1* and *clf-28*. Blue: relatively low, red: relatively high. TPL and CLF mutants showed similar trend in expression change. GO enrichment of **c,** H3K27me3, CLF, TPL and LHP1 genes, and **d,** CLF and TPL only genes: suggest TPL + CLF with H3K27me3 play a role in developmental processes while TPL + CLF without H3K27me3 in response to stimulus (genes with H3K27me3 but no CLF didn't show an enrichment in developmental processes). H: H3K27me3, C: CLF, T: TPL, L: LHP1.

We next performed RNA-seq with *clf* mutant and found a significant number of genes are up-regulated in both *tpl-1* and *clf-28* mutants (Fig 4.6a). When we compared the genes that were differential expressed in *tpl-1* and in *clf-28*, we found a similar trend in the direction of

differential expression (Fig 4.6b). We also performed Gene Ontology (GO) analysis and found that TPL together with CLF and LHP1 contributed to the repression of developmentally important genes that were marked by H3K27me3, such as WUS, a Homeobox transcription factor that controls the stem cell pool[228], MP, an Auxin response factor important for root formation[229], and many floral homeotic genes (Fig 4.6c). CLF, together with TPL, may play a role in stress response independent of H3K27me3 (Fig 4.6d).

Discussion

Understanding how plants use transcriptional repression methods to properly regulate gene expression in response to developmental and environmental signals is key to deciphering the gene regulation codes in plants. New findings in plants may shed light on regulatory mechanisms of conserved counterparts in animals and fungi. Here we showed a novel layer of transcriptional regulation by Gro/Tup1 like co-repressors that has never been reported in animals and fungi. We found that TPL, a Gro/Tup1 type co-repressor in *Arabidopsis*, is able to interact with K27me3 marked histone H3. The repressive mark enhances the binding of TPL to the targets and TPL in turn affects local H3K27me3 level.

We established a link between co-repressors and PcG proteins by showing that TPL physically interacts with LHP1 and co-regulating a common group of genes with PRC complexes. Both CLF and TPL are involved in the spreading of H3K27me3 mark and the ability of TPL to recognize H3K27me3 indicates TPL may perform an LHP1-like role in *Arabidopsis*. A mutation in *LHP1* greatly enhances the *tpl-1* phenotype, which is another indication of functional convergence. What's more, *tpl-1/lhp1-6* displayed earlier termination of flowering than the *lhp1-*

6 single mutant, and showed new phenotypes such as disrupted silique phyllotaxy (data not shown), suggesting a synergistic interaction between *TPL* and *LHP1*.

Although CLF peaks greatly overlap with TPL peaks, we were not able to detect a physical interaction between the two proteins. *CLF* is relatively lowly expressed in *Arabidopsis*, which might explain why regular co-IP with CLF failed. It's also possible the association between TPL and CLF is weak or indirect, which would partially explain the weak enhancement of *tpl-1* phenotype by *clf-31*. The slight enhancement of phenotype might also be explained by the fact that CLF protein is absent from the apical meristems during embryogenesis (Fig 4.3a and data not shown). Many developmentally important genes, such as floral homeotic genes, are strictly silenced and H3K27 tri-methylated in the apical meristems, suggests a non-redundant function of MEA or SWN in mediating H3K27me3 deposition. Interestingly, we found CLF peaks are not always localized at H3K27me3 marked regions and those peaks correlate with genes that are enriched in response genes. Those genes need to be readily activated upon biotic or abiotic stimuli. The plant PRC2 catalytic subunit CLF may have adopted this novel function independent of H3K27me3 to ensure immediate response.

How PRC complexes are recruited as well as how H3K27me3 is established and spreads are still unclear in plants. H3K27me3 regions in *Arabidopsis* are enriched in conserved noncoding sequences (CNSs) that are enriched with transcription factor binding sites[139]. It was reported that the PRC2 complex in plants might be recruited by transcription factors[230]. However, in our IP-MS we failed to detect transcription factors that interact with LHP1 or CLF. TPL has been

shown to interact with a wide range of transcription factors, which makes it a good candidate for connecting the TFs with PcGs.

Materials and Methods

Growth Conditions and mutant alleles

Plants were grown either on soil or on Petri dishes containing Linsmaier and Skoog (LS) salts medium and 0.9% of agar. The controls used were all wild-type plants in the same ecotypes to the mutants or the transgenic plants for each experiment. For petri dish growth, *Arabidopsis thaliana* seeds were sterilized by ethanol and stratified for at least 3 days at 4°C in the dark prior to sowing. Seedlings were harvested 10 days after germination on 1.5% agar supplemented with LS plates. For floral studies, the unopened buds were collected from inflorescences that had 6 open flowers. Plants were grown at 17, 21 or 24°C under 16-hour light/8-hour dark cycle stable conditions in highly controlled growth chambers. Plants for meristem imaging were grown under short day conditions with an 8-hour light/16-hour dark cycle.

Recombineering

Transgenic lines with fluorescence protein fusions were constructed using Recombineering method as previously described[212]. For ChIP and imaging, 2xYPet-3xFLAG or 2xYPet-3xHA were fused to the C terminus of all genes except for *CLF* to the N terminus. tdTomato-3xFLAG was also used to introduce another tag compatible for co-IP. Each construct was transformed into the respective loss of function allele via the *Agrobacterium* floral dip method[213].

Histology and Microscopy

For YPET analysis, ovules and inflorescence meristems (IMs) were fixed in 4% paraformaldehyde with 4% DMSO, vacuum infiltrated for three rounds for a total of 45 minutes, washed with phosphate buffer saline (PBS, PH 7.0), vacuum infiltrated with 2% SCRI Renaissance 2200 (Renaissance Chemicals) and 4% DMSO, followed by two washes with PBS. IM's were embedded in 3% agarose and sectioned using a Vibration microtome (Thermo Scientific HM 650 V). Ovules and sections of IM were mounted in 20% glycerol. Slides were imaged using a laser scanning confocal microscope (Zeiss LSM 710) with an exciting wave length of 405 nm and emission of 420-470nm for SR2200, and an exciting wave length of 517 nm and emission of 500-535nm for YPET.

Histone peptide array

GST fusion proteins were expressed in BL21 cells. Bacteria cells from 1L culture were collected and lysed in 50ml PBS supplemented with 0.05g Lysozyme, 125 ul of 0.4mg/ml Leupeptin, 50 ul of 1mg/ml Pepstatin A and 500ul 0.1M PMSF. The lysate was then sonicated until it became clear. Triton X-100 was added to a final concentration of 1%. The lysate was incubated at 4°C for 30min with rotating and then centrifuged at top speed. The supernatant was incubated with 1050ul anti-GST beads for 30min to collect the fusion protein. Beads were washed 5 times with PBS and eluted with elution buffer containing 10mM reduced glutathione. The elute was dialyzed in a Float-A-Lyzer dialysis device at 4°C overnight for a total of 14 hours with interaction buffer (100mM KCl, 20mM HEPES with pH 7.5, 1mM EDTA, 10% glycerol and 0.1mM DTT). Buffer was changed at 2 hours, 4 hours, 6 hours and 12 hours. The purified

protein was then used for histone peptide array (Active Motif 13005) as per manufacturer's instruction with the following modifications: 3% Bovine serum albumin (BSA) instead of 5% non-fat milk was used for blocking buffer. 50nM GST fusion protein was incubated on the array for 1 hour at RT and anti-GST-HRP in 3ml blocking buffer was used for detection purpose.

ChIP-seq

ChIP experiments were performed as previously described[27,212] with the following modifications. 10 grams of 10-day old seedlings from Transgenic homozygous lines were vacuum crosslinked with 1% Formaldehyde in PBS for 15 minutes and vacuum quenched with 0.125M glycine in PBS for 5 minutes. Protein A/G magnetic beads (Pierce 88803) were used with anti-GFP antibody (abcam ab290) in immunoprecipitations. Immunoprecipitated DNA was purified using AmpureXP beads and libraries were constructed using the Nugen Ovation Ultralow Library System according to the manufacturers protocol. Input DNA or Ler ChIP DNA was used as a control. Libraries were sequenced 50bp with single end on an Illumina HiSeq 2000. Reads were mapped to the TAIR10 *Arabidopsis* genome using bowtie2. Uniquely mapped reads were retained for peak calling using the MACS2[214] and for metaplots. Each ChIP contained at least 2 biological replicates. MACS2 was run using default settings with a q value cutoff of 0.01 and peaks with at least two-fold enrichment versus input were selected. De-novo motif discovery was conducted using HOMER (http://homer.ucsd.edu/homer/) with the findMotifsGenome function. Sequences 200bp surrounding the summits of peaks were used. In some cases, only the top 1000 high quality peaks with the lowest q-value were used. Heat maps and metaplot profiles of relative enrichment (log2 ChIP/Input) at defined genomic regions and were generated by

deeptools[215]. BigWig files of the mapped reads were visualized by the Integrative Genomics Viewer (IGV)[216,217]. Genes were annotated by ChIPseeker with a 3kb cutoff[218].

RNA-seq

100mg of 10-day old seedlings from homozygous lines were harvested from the petri and ground using liquid N2. Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) followed by DNase digestion (Promega) following the manufacturers' protocols. Libraries were constructed with the TruSeq RNA Library Preparation Kit v2 (Illumina) as per manufacturer's instruction. Samples were sequenced on an Illumina HiSeq 2000 (single-end 50-bp run). Raw reads were mapped to the TAIR10 *Arabidopsis* genome using TopHat with standard parameters except for intron length of 40). Uniquely mapped reads were then used for Cufflinks for differential expression analysis[219].

Co-IP

For co-immunoprecipitation experiments, one gram of floral tissue was isolated from F1 crosses between recombineered 2xYPet-3xHA and tdTomato-3xFLAG constructs, ground in liquid nitrogen and resuspended in 4 mL ice-cold IP buffer [50 mM Tris·HCl pH 8.0, 150 mM NaCl, 5 mM MgCl2, 0.1% Tergitol NP-40, 10% (vol/vol) glycerol, 1x Protease Inhibitor Mixture (Roche)]. Resuspended tissue was filtered once through two layers of Miracloth and centrifuged for 10 min at 20,000g at 4°C. Supernatant was incubated with 50 µL of anti-GFP beads from the micro-MACS GFP-tagged protein isolation kit (Miltenyi Biotec) for 30 min on ice. Colombia (Col) wild-type tissue was used as a negative control for TPL-histone co-IP and tissue from *TPL-tdTomato-3xFLAG* were used as a negative control for other co-IPs. Beads were washed and

eluted as per manufacturer's instructions. Immunoblot analysis was performed with an antibody against histones (H3-abcam 1791, H3K27me3-millipore 07-449 and H3K9me2-abcam 1220) or FLAG (SIGMA, A8592).

Chapter 5: Conclusions and Future Perspectives

In this thesis, we characterized the mechanism of transcriptional repression by TPL from three different levels and came up with a model that consists of the following steps: Firstly, TPL is recruited by tissue-specific transcription factors; at some targets, the association of TPL with local chromatin is enhanced by the presence of H3K27me3 after the initial recruitment; TPL in turn recruit HDA19 to form co-repressor complexes and may recruit PcGs to contribute the spreading of H3K27me3 to repress gene expression (Fig 5.1).

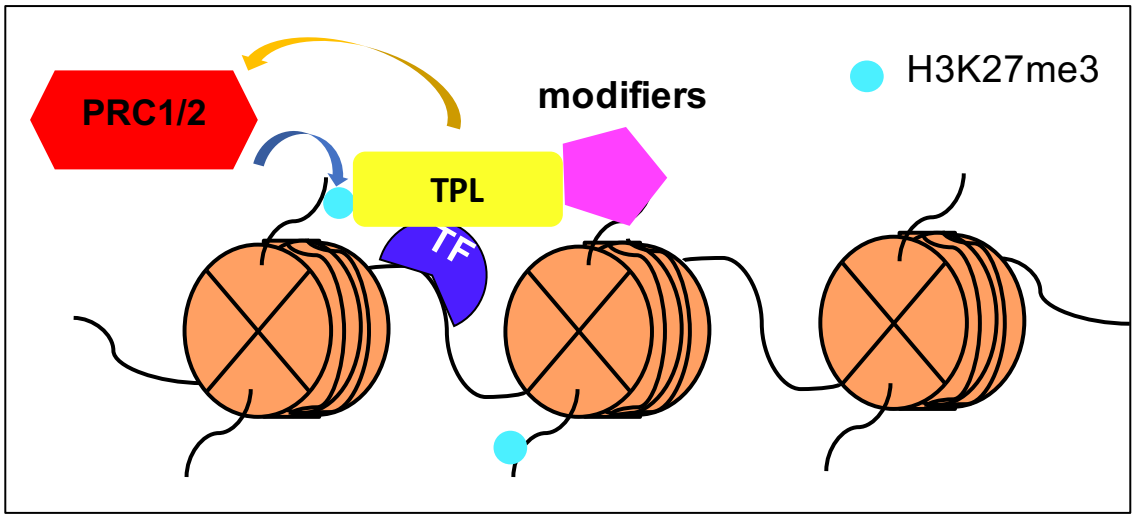


Fig 5.1: a model of transcriptional repression mediated by TPL

5.1 Tissue specificity of TPL function

The phenotypes of the dominant negative allele *tpl-1*, quadruple and quintuple mutants of *tpl* and *tprs* showed a tissue specific pattern. The predominant phenotype of *tpl-1* and the quintuple mutant is the loss of apical fate during embryogenesis, leading to a transformation of the apical pole into a second basal pole. We also saw a clear phenotype of floral homeotic transformation in the *tpl-8/tpr1,2,4* quadruple mutants. These observations suggest the function of TPL and

TPRs is highly tissue specific. However, TPL is ubiquitously expressed in every cell at every stage of *Arabidopsis* development and growth, leading to the question what contributes to this specificity.

Co-repressors lack the ability to bind directly to DNA and need to be recruited to their targets by transcription factors. TPL has been shown to interact with a wide range of transcription factors[197]. In this thesis, we showed that TPL interacts with the HD-ZIP III family of transcription factors. HD-ZIP III proteins specifically accumulate in the SAM and in the adaxial side of leaf primordia. Transcription factors with restricted expression patterns such as HD-ZIP IIIs likely confer tissue specificity of TPL function.

Histone modifications are other candidates for providing specificity. Gro/Tup1 like co-repressors have been shown to repress gene expression by recruiting histone deacetylases[220]. We found that in *Arabidopsis*, HDA19 colocalizes with only about 50% of the regions bound by TPL, suggestive of the existence of different TPL co-repressor complexes. We showed that *HDA19* is involved in transcriptional repression with TPL by different genetic interactions. The different modes of action may partially explain functional diversity. We also showed that TPL interacts with H3K27me3 and LHP1, a PRC1 component in plants. The direct interaction of TPL with histones offers another layer of regulation which may result in functional specificity.

5.2 Dynamic transcriptional regulation

Precise temporal and spatial regulation of gene expression is essential for normal plant development. We found a number of important regulators for development are the direct targets

of TPL (supplemental Excel sheet 4.2). Interestingly, these factors may be dynamically self-regulated, with the proteins themselves also interacting with TPL. For example, EMF1, a plant-specific component of PRC1, which may contribute to the recruitment of PRC1 complexes[127], physically interacts with TPL[197] and is a direct target of TPL and PcGs.

As sessile organisms, plants are required to quickly respond to biotic and abiotic signals. Here we showed that CLF, a catalytic core of PRC2 complexes in *Arabidopsis*, may play a role in regulating response genes with TPL but without H3K27 tri-methylation. Consistent with CLF being involved in responses to stimuli, we constantly observed pleiotropic phenotypes in *clf* mutants grown in the less steady environment in greenhouse.

5.3 tissue specific ChIP-seq and RNA-seq

While we identified transcription factors that are of great importance for plant development, we failed to capture all the regulatory events due to technical restrictions. With all the experiments done with whole plants or inflorescence which consist of different types of cells, we were not able to establish a clear connection between the binding of transcriptional regulators and change of expression of target genes in the mutants. The fact that the results are the "average" of different types of tissue at a steady state in response to genetic perturbation makes the identification of direct regulation difficult, especially when TPL is a master regulator which is expressed ubiquitously yet has diverse functions or targets in different types of cells. Therefore, cell-type specific approaches, such as either single cell sequencing[231] or INTACT[232], are required to explore transcriptional regulation at a better resolution.

5.4 Connection between histone acetylation and H3K27me3

In *Arabidopsis*, the mechanisms of initiation and maintenance of Polycomb silencing are largely unknown. Transcription factor binding sites are overrepresented in H3K27me3 regions[139,230], suggesting a possible role of transcription factors in recruiting PcGs. We showed that TPL associates with PcGs in the genome and affects H3K27me3 level, which indicates that co-repressors may also involve in the deposition of H3K27me3 by PRC complexes. TPL also interacts with transcription factors and showed an enrichment of TF binding motifs identified within H3K27me3 peaks (supplemental Excel 3.1), which makes it a possible bridge between TFs and PcGs.

The function of the cooperation between TPL and PcG proteins is unclear. H3K27me3 is in general considered a relatively stable modification that turn off gene expression faithfully. By incorporating TPL in this machinery the plant may acquire an opportunity for less strict silencing of expression. It has been reported that linc RNAs are involved in the recruitment of PcGs at some targets[140,141]. TPL may prevent the promoters from heavy H3K27me3 methylation and faithful silencing, thus providing an opportunity window for lincRNA synthesis. It has also been shown that HDA6, another HDAC that interacts with TPL, is not only required for histone deacetylation but also for heterochromatin condensation[233-235]. Another promising model will be that TPL is involved in the transition of the targets from active transcription to stable silencing by recruiting HDACs to clear off acetylation for the subsequent K27 tri-methylation of H3. To test these hypotheses, results from tracking the dynamic change in acetylation and H3K27me3 deposition in an inducible system will be informative.

Overall, we described a model of transcriptional regulation by TPL that involved transcription factors and crosstalk with histone modifications. Our data suggest a novel layer of transcription repression by Gro/Tup1 like co-repressors in *Arabidopsis* and the possible involvement of co-repressors in the function of PcG proteins. This work in plants may shed light on similar studies in animals and fungi.

References

1 Lindsley, D. L. & Grell, E. H. Genetic Variations of Drosophila melanogaster. *Carnegie Institution of Washington*, 627 (1968).

2 Agarwal, M., Kumar, P. & Mathew, S. J. The Groucho/Transducin-like enhancer of split protein family in animal development. *IUBMB Life* **67**, 472-481, doi:10.1002/iub.1395 (2015).

3 Liu, Z. & Karmarkar, V. Groucho/Tup1 family co-repressors in plant development. *Trends Plant Sci* **13**, 137-144, doi:10.1016/j.tplants.2007.12.005 (2008).

4 Flores-Saaib, R. D. & Courey, A. J. Analysis of Groucho-histone interactions suggests mechanistic similarities between Groucho- and Tup1-mediated repression. *Nucleic Acids Res* **28**, 4189-4196, doi:DOI 10.1093/nar/28.21.4189 (2000).

5 Jennings, B. H. *et al.* Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol Cell* **22**, 645-655, doi:10.1016/j.molcel.2006.04.024 (2006).

6 Davie, J. K., Trumbly, R. J. & Dent, S. Y. R. Histone-dependent association of Tup1-Ssn6 with repressed genes in vivo. *Mol Cell Biol* **22**, 693-703, doi:Doi 10.1128/Mcb.22.3.693-703.2002 (2002).

7 Winkler, C. J., Ponce, A. & Courey, A. J. Groucho-Mediated Repression May Result from a Histone Deacetylase-Dependent Increase in Nucleosome Density. *Plos One* **5**, doi:ARTN e10166 10.1371/journal.pone.0010166 (2010).

8 Edmondson, D. G. & Roth, S. Y. Interactions of transcriptional regulators with histones. *Methods* **15**, 355-364, doi:DOI 10.1006/meth.1998.0639 (1998).

9 Palaparti, A., Baratz, A. & Stifani, S. The groucho/transducin-like enhancer of split transcriptional repressors interact with the genetically defined a amino-terminal silencing domain of histone H3. *J Biol Chem* **272**, 26604-26610, doi:DOI 10.1074/jbc.272.42.26604 (1997).

10 Watson, A. D. *et al.* Ssn6-Tup1 interacts with class I histone deacetylases required for repression. *Gene Dev* **14**, 2737-2744, doi:DOI 10.1101/gad.829100 (2000).

11 Martinez, C. A. & Arnosti, D. N. Spreading of a corepressor linked to action of long-range repressor hairy. *Mol Cell Biol* **28**, 2792-2802, doi:10.1128/Mcb.01203-07 (2008).

12 Sekiya, T. & Zaret, K. S. Repression by Groucho/TLE/Grg proteins: Genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. *Mol Cell* **28**, 291-303, doi:10.1016/j.molcel.2007.10.002 (2007).

13 Chambers, M. *et al.* Mechanisms of Groucho-mediated repression revealed by genome-wide analysis of Groucho binding and activity. *Bmc Genomics* **18**, doi:ARTN 215 10.1186/s12864-017-3589-6 (2017).

14 Kaul, A., Schuster, E. & Jennings, B. H. The Groucho Co-repressor Is Primarily Recruited to Local Target Sites in Active Chromatin to Attenuate Transcription. *Plos Genet* **10**, doi:ARTN e1004595 10.1371/journal.pgen.1004595 (2014).

15 Yasuoka, Y. *et al.* Occupancy of tissue-specific cis-regulatory modules by Otx2 and TLE/Groucho for embryonic head specification. *Nat Commun* **5**, doi:ARTN 4322 10.1038/ncomms5322 (2014).

16 Gromoller, A. & Lehming, N. Srb7p is a physical and physiological target of Tup1p. *Embo J* **19**, 6845-6852, doi:DOI 10.1093/emboj/19.24.6845 (2000).

17 Kuchin, S. & Carlson, M. Functional relationships of Srb10-Srb11 kinase, carboxy-terminal domain kinase CTDK-I, and transcriptional corepressor Ssn6-Tup1. *Mol Cell Biol* **18**, 1163-1171, doi:Doi 10.1128/Mcb.18.3.1163 (1998).

18 Papamichos-Chronakis, M., Conlan, R. S., Gounalaki, N., Copf, T. & Tzamarias, D. Hrs1/Med3 is a Cyc8-Tup1 corepressor target in the RNA polymerase II holoenzyme. *J Biol Chem* **275**, 8397-8403, doi:DOI 10.1074/jbc.275.12.8397 (2000).

19 Zhang, H. & Emmons, S. W. Caenorhabditis elegans unc-37/groucho interacts genetically with components of the transcriptional mediator complex. *Genetics* **160**, 799-803 (2002).

20 Cai, Y., Brophy, P. D., Levitan, I., Stifani, S. & Dressler, G. R. Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation. *Embo J* **22**, 5522-5529, doi:DOI 10.1093/emboj/cdg536 (2003).

21 Gregis, V., Sessa, A., Colombo, L. & Kater, M. M. AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *Plant Cell* **18**, 1373-1382, doi:10.1105/tpc.106.041798 (2006).
22 Pfluger, J. & Zambryski, P. The role of SEUSS in auxin response and floral organ patterning. *Development* **131**, 4697-4707, doi:10.1242/dev.01306 (2004).
23 Simonini, S. *et al.* Basic pentacysteine proteins mediate MADS domain complex binding to the DNA for tissue-specific expression of target genes in Arabidopsis. *Plant Cell* **24**, 4163-4172, doi:10.1105/tpc.112.103952 (2012).
24 Sridhar, V. V., Surendrarao, A. & Liu, Z. C. APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development (vol 133, pg 3159, 2001). *Development* **133**, 3496-3496, doi:10.1242/dev.02562 (2006).
25 Stahle, M. I., Kuehlich, J., Staron, L., von Arnim, A. G. & Golz, J. F. YABBYs and the transcriptional corepressors LEUNIG and LEUNIG_HOMOLOG maintain leaf polarity and meristem activity in Arabidopsis. *Plant Cell* **21**, 3105-3118, doi:10.1105/tpc.109.070458 (2009).
26 Gonzalez, D., Bowen, A. J., Carroll, T. S. & Conlan, R. S. The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. *Mol Cell Biol* **27**, 5306-5315, doi:10.1128/Mcb.01912-06 (2007).
27 Krogan, N. T., Hogan, K. & Long, J. A. APETALA2 negatively regulates multiple floral organ identity genes in Arabidopsis by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. *Development* **139**, 4180-4190, doi:10.1242/dev.085407 (2012).
28 Ito, J. *et al.* Auxin-dependent compositional change in Mediator in ARF7-and ARF19-mediated transcription. *P Natl Acad Sci USA* **113**, 6562-6567, doi:10.1073/pnas.1600739113 (2016).
29 Conner, J. & Liu, Z. C. LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. *P Natl Acad Sci USA* **97**, 12902-12907, doi:DOI 10.1073/pnas.230352397 (2000).
30 Cnops, G. *et al.* The TORNADO1 and TORNADO2 genes function in several patterning processes during early leaf development in Arabidopsis thaliana. *Plant Cell* **18**, 852-866, doi:10.1105/tpc.105.040568 (2006).
31 Geng, X. Y. *et al.* LEUNIG_HOMOLOG transcriptional co-repressor mediates aluminium sensitivity through PECTIN METHYLESTERASE46-modulated root cell wall pectin methylesterification in Arabidopsis. *Plant J* **90**, 491-504, doi:10.1111/tpj.13506 (2017).
32 Simon, M. K., Skinner, D. J., Gallagher, T. L. & Gasser, C. S. Integument Development in Arabidopsis Depends on Interaction of YABBY Protein INNER NO OUTER with Coactivators and Corepressors. *Genetics* **207**, 1489-1500, doi:10.1534/genetics.117.300140 (2017).
33 Long, J. A., Ohno, C., Smith, Z. R. & Meyerowitz, E. M. TOPLESS regulates apical embryonic fate in Arabidopsis. *Science* **312**, 1520-1523, doi:10.1126/science.1123841 (2006).
34 Smith, Z. R. & Long, J. A. Control of Arabidopsis apical-basal embryo polarity by antagonistic transcription factors. *Nature* **464**, 423-U121, doi:10.1038/nature08843 (2010).
35 Busch, W. *et al.* Transcriptional Control of a Plant Stem Cell Niche. *Dev Cell* **18**, 849-861, doi:10.1016/j.devcel.2010.03.012 (2010).
36 Espinosa-Ruiz, A. *et al.* TOPLESS mediates brassinosteroid control of shoot boundaries and root meristem development in Arabidopsis thaliana. *Development* **144**, 1619-1628, doi:10.1242/dev.143214 (2017).
37 Pi, L. M. *et al.* Organizer-Derived WOX5 Signal Maintains Root Columella Stem Cells through Chromatin-Mediated Repression of CDF4 Expression. *Dev Cell* **33**, 576-588, doi:10.1016/j.devcel.2015.04.024 (2015).
38 Xiang, D. Q. *et al.* POPCORN Functions in the Auxin Pathway to Regulate Embryonic Body Plan and Meristem Organization in Arabidopsis. *Plant Cell* **23**, 4348-4367, doi:10.1105/tpc.111.091777 (2011).
39 Borg, M. *et al.* An EAR-Dependent Regulatory Module Promotes Male Germ Cell Division and Sperm Fertility in Arabidopsis. *Plant Cell* **26**, 2098-2113, doi:10.1105/tpc.114.124743 (2014).
40 Wang, L., Kim, J. & Somers, D. E. Transcriptional corepressor TOPLESS complexes with pseudoresponse regulator proteins and histone deacetylases to regulate circadian transcription. *P Natl Acad Sci USA* **110**, 761-766, doi:10.1073/pnas.1215010110 (2013).
41 Tao, Q. *et al.* The TIE1 Transcriptional Repressor Links TCP Transcription Factors with TOPLESS/TOPLESS-RELATED Corepressors and Modulates Leaf Development in Arabidopsis. *Plant Cell* **25**, 421-437, doi:10.1105/tpc.113.109223 (2013).

42 Gonzalez, N. *et al.* A repressor protein complex regulates leaf growth in Arabidopsis (vol 27, pg 2273, 2015). *Plant Cell* **28**, 824-824, doi:10.1105/tpc.16.00202 (2016).
43 Zhu, Z. H. *et al.* Arabidopsis resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *P Natl Acad Sci USA* **107**, 13960-13965, doi:10.1073/pnas.1002828107 (2010).
44 Szemenyei, H., Hannon, M. & Long, J. A. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science* **319**, 1384-1386, doi:10.1126/science.1151461 (2008).
45 Pauwels, L. *et al.* NINJA connects the co-repressor TOPLESS to jasmonate signalling. *Nature* **464**, 788-U169, doi:10.1038/nature08854 (2010).
46 Wang, L. *et al.* Strigolactone Signaling in Arabidopsis Regulates Shoot Development by Targeting D53-Like SMXL Repressor Proteins for Ubiquitination and Degradation. *Plant Cell* **27**, 3128-3142, doi:10.1105/tpc.15.00605 (2015).
47 Ryu, H., Cho, H., Bae, W. & Hwang, I. Control of early seedling development by BES1/TPL/HDA19-mediated epigenetic regulation of ABI3. *Nat Commun* **5**, doi:ARTN 4138
10.1038/ncomms5138 (2014).
48 Oh, E., Zhu, J. Y., Ryu, H., Hwang, I. & Wang, Z. Y. TOPLESS mediates brassinosteroid-induced transcriptional repression through interaction with BZR1. *Nat Commun* **5**, doi:ARTN 4140
10.1038/ncomms5140 (2014).
49 Fukazawa, J. *et al.* DELLAs Function as Coactivators of GAI-ASSOCIATED FACTOR1 in Regulation of Gibberellin Homeostasis and Signaling in Arabidopsis. *Plant Cell* **26**, 2920-2938, doi:10.1105/tpc.114.125690 (2014).
50 Ke, J. Y. *et al.* Structural basis for recognition of diverse transcriptional repressors by the TOPLESS family of corepressors. *Sci Adv* **1**, doi:ARTN e1500107
10.1126/sciadv.1500107 (2015).
51 Martin-Arevalillo, R. *et al.* Structure of the Arabidopsis TOPLESS corepressor provides insight into the evolution of transcriptional repression. *P Natl Acad Sci USA* **114**, 8107-8112, doi:10.1073/pnas.1703054114 (2017).
52 Ruberti, I., Sessa, G., Lucchetti, S. & Morelli, G. A Novel Class of Plant-Proteins Containing a Homeodomain with a Closely Linked Leucine Zipper Motif. *Embo J* **10**, 1787-1791, doi:DOI 10.1002/j.1460-2075.1991.tb07703.x (1991).
53 Sessa, G., Steindler, C., Morelli, G. & Ruberti, I. The Arabidopsis Athb-8, -9 and -14 genes are members of a small gene family coding for highly related HD-ZIP proteins. *Plant Mol Biol* **38**, 609-622, doi:Doi 10.1023/A:1006016319613 (1998).
54 Baima, S. *et al.* The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol* **126**, 643-655, doi:DOI 10.1104/pp.126.2.643 (2001).
55 Henriksson, E. *et al.* Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. *Plant Physiol* **139**, 509-518, doi:10.1104/pp.105.063461 (2005).
56 Ohto, M. A., Hayashi, S., Sawa, S., Hashimoto-Ohta, A. & Nakamura, K. Involvement of HLS1 in sugar and auxin signaling in Arabidopsis leaves. *Plant Cell Physiol* **47**, 1603-1611, doi:10.1093/pcp/pcl027 (2006).
57 Ciarbelli, A. R. *et al.* The Arabidopsis Homeodomain-leucine Zipper II gene family: diversity and redundancy. *Plant Mol Biol* **68**, 465-478, doi:10.1007/s11103-008-9383-8 (2008).
58 Schrick, K., Nguyen, D., Karlowski, W. M. & Mayer, K. F. X. START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. *Genome Biol* **5**, doi:ARTN R41
DOI 10.1186/gb-2004-5-6-r41 (2004).
59 Mukherjee, K., Brocchieri, L. & Burglin, T. R. A Comprehensive Classification and Evolutionary Analysis of Plant Homeobox Genes. *Mol Biol Evol* **26**, 2775-2794, doi:10.1093/molbev/msp201 (2009).
60 Floyd, S. K. & Bowman, J. L. Gene regulation: Ancient microRNA target sequences in plants. *Nature* **428**, 485-486, doi:10.1038/428485a (2004).
61 McConnell, J. R. *et al.* Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* **411**, 709-713, doi:Doi 10.1038/35079635 (2001).
62 McHale, N. A. & Koning, R. E. MicroRNA-directed cleavage of Nicotiana sylvestris PHAVOLUTA mRNA regulates the vascular cambium and structure of apical meristems. *Plant Cell* **16**, 1730-1740, doi:10.1105/tpc.021816 (2004).

63 Mallory, A. C. *et al.* MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5 ' region. *Embo J* **23**, 3356-3364, doi:10.1038/sj.emboj.7600340 (2004).
64 Moglich, A., Ayers, R. A. & Moffat, K. Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure* **17**, 1282-1294, doi:10.1016/j.str.2009.08.011 (2009).
65 Magnani, E. & Barton, M. K. A Per-ARNT-Sim-Like Sensor Domain Uniquely Regulates the Activity of the Homeodomain Leucine Zipper Transcription Factor REVOLUTA in Arabidopsis. *Plant Cell* **23**, 567-582, doi:10.1105/tpc.110.080754 (2011).
66 Chandler, J. W., Cole, M., Flier, A., Grewe, B. & Werr, W. The AP2 transcription factors DORNROSCHEN and DORNROSCHEN-LIKE redundantly control Arabidopsis embryo patterning via interaction with PHAVOLUTA. *Development* **134**, 1653-1662, doi:10.1242/dev.001016 (2007).
67 Catterou, M. *et al.* hoc: an Arabidopsis mutant overproducing cytokinins and expressing high in vitro organogenic capacity. *Plant J* **30**, 273-287, doi:DOI 10.1046/j.1365-313X.2002.01286.x (2002).
68 Talbert, P. B., Adler, H. T., Parks, D. W. & Comai, L. The Revoluta Gene Is Necessary for Apical Meristem Development and for Limiting Cell Divisions in the Leaves and Stems of Arabidopsis-Thaliana. *Development* **121**, 2723-2735 (1995).
69 Emery, J. F. *et al.* Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr Biol* **13**, 1768-1774 (2003).
70 Prigge, M. J. *et al.* Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. *Plant Cell* **17**, 61-76, doi:10.1105/tpc.104.026161 (2005).
71 Otsuga, D., DeGuzman, B., Prigge, M. J., Drews, G. N. & Clark, S. E. REVOLUTA regulates meristem initiation at lateral positions. *Plant J* **25**, 223-236, doi:DOI 10.1046/j.1365-313x.2001.00959.x (2001).
72 Richmond, T. J. Hot papers - Crystal structure - Crystal structure of the nucleosome core particle at 2.8 angstrom resolution by K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond - Comments. *Scientist* **13**, 15-15 (1999).
73 Angelov, D., Vitolo, J. M., Mutskov, V., Dimitrov, S. & Hayes, J. J. Preferential interaction of the core histone tail domains with linker DNA. *P Natl Acad Sci USA* **98**, 6599-6604, doi:DOI 10.1073/pnas.121171498 (2001).
74 Ma, X. J., Lv, S. B., Zhang, C. & Yang, C. P. Histone deacetylases and their functions in plants. *Plant Cell Rep* **32**, 465-478, doi:10.1007/s00299-013-1393-6 (2013).
75 Baena-Gonzalez, E., Rolland, F., Thevelein, J. M. & Sheen, J. A central integrator of transcription networks in plant stress and energy signalling. *Nature* **448**, 938-U910, doi:10.1038/nature06069 (2007).
76 Benhamed, M., Bertrand, C., Servet, C. & Zhou, D. X. Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell* **18**, 2893-2903, doi:10.1105/tpc.106.043489 (2006).
77 Kim, K. C., Lai, Z. B., Fan, B. F. & Chen, Z. X. Arabidopsis WRKY38 and WRKY62 Transcription Factors Interact with Histone Deacetylase 19 in Basal Defense. *Plant Cell* **20**, 2357-2371, doi:10.1105/tpc.107.055566 (2008).
78 Perrella, G. *et al.* Histone Deacetylase Complex1 Expression Level Titrates Plant Growth and Abscisic Acid Sensitivity in Arabidopsis. *Plant Cell* **25**, 3491-3505, doi:10.1105/tpc.113.114835 (2013).
79 Tanaka, M., Kikuchi, A. & Kamada, H. The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol* **146**, 149-161, doi:10.1104/pp.107.111674 (2008).
80 Tian, L. & Chen, Z. J. Blocking histone deacetylation in Arabidopsis induces pleiotropic effects on plant gene regulation and development (vol 98, pg 200, 2001). *P Natl Acad Sci USA* **98**, 7647-7647 (2001).
81 Zhou, C. H., Zhang, L., Duan, J., Miki, B. & Wu, K. Q. HISTONE DEACETYLASE19 is involved in jasmonic acid and ethylene signaling of pathogen response in Arabidopsis. *Plant Cell* **17**, 1196-1204, doi:10.1105/tpc.104.028514 (2005).
82 Zhou, Y. *et al.* HISTONE DEACETYLASE19 Interacts with HSL1 and Participates in the Repression of Seed Maturation Genes in Arabidopsis Seedlings. *Plant Cell* **25**, 134-148, doi:10.1105/tpc.112.096313 (2013).
83 Choi, S. M. *et al.* HDA19 is required for the repression of salicylic acid biosynthesis and salicylic acid-mediated defense responses in Arabidopsis. *Plant J* **71**, 135-146, doi:10.1111/j.1365-313X.2012.04977.x (2012).
84 Cigliano, R. A. *et al.* Histone Deacetylase AtHDA7 Is Required for Female Gametophyte and Embryo Development in Arabidopsis. *Plant Physiol* **163**, 431-440, doi:10.1104/pp.113.221713 (2013).

85 Luo, M. *et al.* HD2C interacts with HDA6 and is involved in ABA and salt stress response in Arabidopsis. *J Exp Bot* **63**, 3297-3306, doi:10.1093/jxb/ers059 (2012).
86 Luo, M. *et al.* Histone Deacetylase HDA6 Is Functionally Associated with AS1 in Repression of KNOX Genes in Arabidopsis. *Plos Genet* **8**, doi:ARTN e1003114 10.1371/journal.pgen.1003114 (2012).
87 Zhu, Z. Q. *et al.* Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in Arabidopsis. *P Natl Acad Sci USA* **108**, 12539-12544, doi:10.1073/pnas.1103959108 (2011).
88 To, T. K. *et al.* Arabidopsis HDA6 is required for freezing tolerance. *Biochem Bioph Res Co* **406**, 414-419, doi:10.1016/j.bbrc.2011.02.058 (2011).
89 Tessadori, F. *et al.* PHYTOCHROME B and HISTONE DEACETYLASE 6 Control Light-Induced Chromatin Compaction in Arabidopsis thaliana. *Plos Genet* **5**, doi:ARTN e1000638 10.1371/journal.pgen.1000638 (2009).
90 Scofield, S. & Murray, J. A. H. KNOX gene function in plant stem cell niches. *Plant Mol Biol* **60**, 929-946, doi:10.1007/s11103-005-4478-y (2006).
91 Wang, Y. Z. *et al.* HISTONE DEACETYLASE 6 represses pathogen defence responses in Arabidopsis thaliana. *Plant Cell Environ* **40**, 2972-2986, doi:10.1111/pce.13047 (2017).
92 Gu, X. F. *et al.* Arabidopsis Homologs of Retinoblastoma-Associated Protein 46/48 Associate with a Histone Deacetylase to Act Redundantly in Chromatin Silencing. *Plos Genet* **7**, doi:ARTN e1002366 10.1371/journal.pgen.1002366 (2011).
93 Wu, K., Zhang, L., Zhou, C., Yu, C. W. & Chaikam, V. HDA6 is required for jasmonate response, senescence and flowering in Arabidopsis. *J Exp Bot* **59**, 225-234, doi:10.1093/jxb/erm300 (2008).
94 Yu, C. W. *et al.* HISTONE DEACETYLASE6 Interacts with FLOWERING LOCUS D and Regulates Flowering in Arabidopsis. *Plant Physiol* **156**, 173-184, doi:10.1104/pp.111.174417 (2011).
95 Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, M. & Matzke, A. J. M. HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *Embo J* **21**, 6832-6841, doi:DOI 10.1093/emboj/cdf663 (2002).
96 Kim, W., Latrasse, D., Servet, C. & Zhou, D. X. Arabidopsis histone deacetylase HDA9 regulates flowering time through repression of AGL19. *Biochem Bioph Res Co* **432**, 394-398, doi:10.1016/j.bbrc.2012.11.102 (2013).
97 Luo, M. *et al.* Regulation of flowering time by the histone deacetylase HDA5 in Arabidopsis. *Plant J* **82**, 925-936, doi:10.1111/tpj.12868 (2015).
98 Liu, X. C. *et al.* PHYTOCHROME INTERACTING FACTOR3 Associates with the Histone Deacetylase HDA15 in Repression of Chlorophyll Biosynthesis and Photosynthesis in Etiolated Arabidopsis Seedlings. *Plant Cell* **25**, 1258-1273, doi:10.1105/tpc.113.109710 (2013).
99 Liu, C. *et al.* HDA18 Affects Cell Fate in Arabidopsis Root Epidermis via Histone Acetylation at Four Kinase Genes. *Plant Cell* **25**, 257-269, doi:10.1105/tpc.112.107045 (2013).
100 Ueno, Y. *et al.* Histone deacetylases and ASYMMETRIC LEAVES2 are involved in the establishment of polarity in leaves of Arabidopsis. *Plant Cell* **19**, 445-457, doi:10.1105/tpc.106.042325 (2007).
101 Yano, R., Takebayashi, Y., Nambara, E., Kamiya, Y. & Seo, M. Combining association mapping and transcriptomics identify HD2B histone deacetylase as a genetic factor associated with seed dormancy in Arabidopsis thaliana. *Plant J* **74**, 815-828, doi:10.1111/tpj.12167 (2013).
102 Colville, A. *et al.* Role of HD2 genes in seed germination and early seedling growth in Arabidopsis. *Plant Cell Rep* **30**, 1969-1979, doi:10.1007/s00299-011-1105-z (2011).
103 Zhang, F., Wang, L. K., Ko, E. E., Shao, K. & Qiao, H. Histone Deacetylases SRT1 and SRT2 Interact with ENAP1 to Mediate Ethylene-Induced Transcriptional Repression. *Plant Cell* **30**, 153-166, doi:10.1105/tpc.17.00671 (2018).
104 Wang, C. Z. *et al.* Arabidopsis Putative Deacetylase AtSRT2 Regulates Basal Defense by Suppressing PAD4, EDS5 and SID2 Expression (vol 51, pg 1291, 2010). *Plant Cell Physiol* **51**, 1820-1820, doi:10.1093/pcp/pcq141 (2010).
105 Konig, A. C. *et al.* The Arabidopsis Class II Sirtuin Is a Lysine Deacetylase and Interacts with Mitochondrial Energy Metabolism. *Plant Physiol* **164**, 1401-1414, doi:10.1104/pp.113.232496 (2014).
106 Kornet, N. & Scheres, B. Members of the GCN5 Histone Acetyltransferase Complex Regulate PLETHORA-Mediated Root Stem Cell Niche Maintenance and Transit Amplifying Cell Proliferation in Arabidopsis. *Plant Cell* **21**, 1070-1079, doi:10.1105/tpc.108.065300 (2009).

107 Vlachonasios, K. E., Thomashow, M. F. & Triezenberg, S. J. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell* **15**, 626-638, doi:10.1105/tpc.007922 (2003).
108 Nelissen, H. *et al.* The elongata mutants identify a functional Elongator complex in plants with a role in cell proliferation during organ growth. *P Natl Acad Sci USA* **102**, 7754-7759, doi:10.1073/pnas.0502600102 (2005).
109 Latrasse, D. *et al.* The MYST histone acetyltransferases are essential for gametophyte development in Arabidopsis. *Bmc Plant Biol* **8**, doi:Artn 121
10.1186/1471-2229-8-121 (2008).
110 Xiao, J. *et al.* Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of FLC in regulating flowering in Arabidopsis. *J Plant Physiol* **170**, 444-451, doi:10.1016/j.jplph.2012.11.007 (2013).
111 Deng, W. W. *et al.* Involvement of the histone acetyltransferase AtHAC1 in the regulation of flowering time via repression of FLOWERING LOCUS C in Arabidopsis. *Plant Physiol* **143**, 1660-1668, doi:10.1104/pp.106.095521 (2007).
112 Heisel, T. J., Li, C. Y., Grey, K. M. & Gibson, S. I. Mutations in HISTONE ACETYLTRANSFERASE1 affect sugar response and gene expression in Arabidopsis. *Front Plant Sci* **4**, doi:ARTN 245
10.3389/fpls.2013.00245 (2013).
113 Bertrand, C. *et al.* Arabidopsis HAF2 gene encoding TATA-binding protein (TBP)-associated factor TAF1, is required to integrate light signals to regulate gene expression and growth. *J Biol Chem* **280**, 1465-1473, doi:10.1074/jbc.M409000200 (2005).
114 Chen, X. S. *et al.* POWERDRESS interacts with HISTONE DEACETYLASE 9 to promote aging in Arabidopsis. *Elife* **5**, doi:ARTN e17214
10.7554/eLife.17214 (2016).
115 Sawarkar, R. & Paro, R. Interpretation of Developmental Signaling at Chromatin: The Polycomb Perspective. *Dev Cell* **19**, 651-661, doi:10.1016/j.devcel.2010.10.012 (2010).
116 Jiang, H. & Kohler, C. Evolution, function, and regulation of genomic imprinting in plant seed development. *J Exp Bot* **63**, 4713-4722, doi:10.1093/jxb/ers145 (2012).
117 Brockdorff, N. Chromosome silencing mechanisms in X-chromosome inactivation: unknown unknowns. *Development* **138**, 5057-5065, doi:10.1242/dev.065276 (2011).
118 Lewis, E. B. A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565-570 (1978).
119 Schuettengruber, B. *et al.* Functional Anatomy of Polycomb and Trithorax Chromatin Landscapes in Drosophila Embryos. *Plos Biol* **7**, 146-163, doi:ARTN e1000013
10.1371/journal.pbio.1000013 (2009).
120 Mozgova, I. & Hennig, L. The Polycomb Group Protein Regulatory Network. *Annu Rev Plant Biol* **66**, 269-296, doi:10.1146/annurev-arplant-043014-115627 (2015).
121 Derkacheva, M. & Hennig, L. Variations on a theme: Polycomb group proteins in plants. *J Exp Bot* **65**, 2769-2784, doi:10.1093/jxb/ert410 (2014).
122 Xu, L. & Shen, W. H. Polycomb Silencing of KNOX Genes Confines Shoot Stem Cell Niches in Arabidopsis. *Current Biology* **18**, 1966-1971, doi:10.1016/j.cub.2008.11.019 (2008).
123 Sanchez-Pulido, L., Devos, D., Sung, Z. R. & Calonje, M. RAWUL: A new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs. *Bmc Genomics* **9**, doi:Artn 308
10.1186/1471-2164-9-308 (2008).
124 Bratzel, F., Lopez-Torrejon, G., Koch, M., Del Pozo, J. C. & Calonje, M. Keeping Cell Identity in Arabidopsis Requires PRC1 RING-Finger Homologs that Catalyze H2A Monoubiquitination. *Current Biology* **20**, 1853-1859, doi:10.1016/j.cub.2010.09.046 (2010).
125 Turck, F. *et al.* Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *Plos Genet* **3**, 855-866, doi:ARTN e86
10.1371/journal.pgen.0030086 (2007).
126 Zhang, X. *et al.* The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat Struct Mol Biol* **14**, 869-871, doi:10.1038/nsmb1283 (2007).
127 Calonje, M., Sanchez, R., Chen, L. & Sung, Z. R. EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in Arabidopsis. *Plant Cell* **20**, 277-291, doi:10.1105/tpc.106.049957 (2008).
128 Yang, C. *et al.* VAL- and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in Arabidopsis. *Curr Biol* **23**, 1324-1329, doi:10.1016/j.cub.2013.05.050 (2013).

129 Beh, L. Y., Colwell, L. J. & Francis, N. J. A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. *P Natl Acad Sci USA* **109**, E1063-E1071, doi:10.1073/pnas.1118678109 (2012).
130 Kassis, J. A. & Brown, J. L. Polycomb Group Response Elements in Drosophila and Vertebrates. *Adv Genet* **81**, 83-118, doi:10.1016/B978-0-12-407677-8.00003-8 (2013).
131 Sing, A. *et al.* A Vertebrate Polycomb Response Element Governs Segmentation of the Posterior Hindbrain. *Cell* **138**, 885-897, doi:10.1016/j.cell.2009.08.020 (2009).
132 Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J. & Kingston, R. E. A Region of the Human HOXD Cluster that Confers Polycomb-Group Responsiveness. *Cell* **140**, 99-110, doi:10.1016/j.cell.2009.12.022 (2010).
133 Tsai, M. C. *et al.* Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* **329**, 689-693, doi:10.1126/science.1192002 (2010).
134 Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-U1148, doi:10.1038/nature08975 (2010).
135 Plath, K. *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**, 131-135, doi:10.1126/science.1084274 (2003).
136 Hunkapiller, J. *et al.* Polycomb-Like 3 Promotes Polycomb Repressive Complex 2 Binding to CpG Islands and Embryonic Stem Cell Self-Renewal. *Plos Genet* **8**, doi:ARTN e1002576
10.1371/journal.pgen.1002576 (2012).
137 Ku, M. *et al.* Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. *Plos Genet* **4**, doi:ARTN e1000242
10.1371/journal.pgen.1000242 (2008).
138 Deng, W. W. *et al.* Arabidopsis Polycomb Repressive Complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes. *Bmc Genomics* **14**, doi:Artn 593
10.1186/1471-2164-14-593 (2013).
139 Berke, L. & Snel, B. The Histone Modification H3K27me3 Is Retained after Gene Duplication and Correlates with Conserved Noncoding Sequences in Arabidopsis. *Genome Biol Evol* **6**, 572-579, doi:10.1093/gbe/evu040 (2014).
140 Heo, J. B. & Sung, S. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science* **331**, 76-79, doi:10.1126/science.1197349 (2011).
141 Ariel, F. *et al.* Noncoding Transcription by Alternative RNA Polymerases Dynamically Regulates an Auxin-Driven Chromatin Loop. *Mol Cell* **55**, 383-396, doi:10.1016/j.molcel.2014.06.011 (2014).
142 Jaillon, O., Aury, J. M. & Wincker, P. "Changing by doubling", the impact of Whole Genome Duplications in the evolution of eukaryotes. *C R Biol* **332**, 241-253, doi:10.1016/j.crvi.2008.07.007 (2009).
143 Semon, M. & Wolfe, K. H. Consequences of genome duplication. *Curr Opin Genet Dev* **17**, 505-512, doi:10.1016/j.gde.2007.09.007 (2007).
144 Brandt, R. *et al.* Genome-wide binding-site analysis of REVOLUTA reveals a link between leaf patterning and light-mediated growth responses. *Plant J* **72**, 31-42, doi:10.1111/j.1365-313X.2012.05049.x (2012).
145 O'Malley, R. C. *et al.* Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**, 1280-1292, doi:10.1016/j.cell.2016.04.038 (2016).
146 Alonso, J. M. & Stepanova, A. N. A recombineering-based gene tagging system for Arabidopsis. *Methods Mol Biol* **1227**, 233-243, doi:10.1007/978-1-4939-1652-8_11 (2015).
147 Wenkel, S., Emery, J., Hou, B. H., Evans, M. M. & Barton, M. K. A feedback regulatory module formed by LITTLE ZIPPER and HD-ZIPIII genes. *Plant Cell* **19**, 3379-3390, doi:10.1105/tpc.107.055772 (2007).
148 DS, O. M. *et al.* Control of reproductive floral organ identity specification in Arabidopsis by the C function regulator AGAMOUS. *Plant Cell* **25**, 2482-2503, doi:10.1105/tpc.113.113209 (2013).
149 Heyndrickx, K. S., Van de Velde, J., Wang, C., Weigel, D. & Vandepoele, K. A functional and evolutionary perspective on transcription factor binding in Arabidopsis thaliana. *Plant Cell* **26**, 3894-3910, doi:10.1105/tpc.114.130591 (2014).
150 Huang, W. *et al.* Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. *Science* **336**, 75-79, doi:10.1126/science.1219075 (2012).
151 Immink, R. G. *et al.* Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. *Plant Physiol* **160**, 433-449, doi:10.1104/pp.112.202614 (2012).
152 Kaufmann, K. *et al.* Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower. *PLoS Biol* **7**, e1000090, doi:10.1371/journal.pbio.1000090 (2009).

153 Liu, T., Carlsson, J., Takeuchi, T., Newton, L. & Farre, E. M. Direct regulation of abiotic responses by the Arabidopsis circadian clock component PRR7. *Plant J* **76**, 101-114, doi:10.1111/tpj.12276 (2013).
154 Moyroud, E. *et al.* Prediction of regulatory interactions from genome sequences using a biophysical model for the Arabidopsis LEAFY transcription factor. *Plant Cell* **23**, 1293-1306, doi:10.1105/tpc.111.083329 (2011).
155 Nakamichi, N. *et al.* Transcriptional repressor PRR5 directly regulates clock-output pathways. *Proc Natl Acad Sci U S A* **109**, 17123-17128, doi:10.1073/pnas.1205156109 (2012).
156 Pedmale, U. V. *et al.* Cryptochromes Interact Directly with PIFs to Control Plant Growth in Limiting Blue Light. *Cell* **164**, 233-245, doi:10.1016/j.cell.2015.12.018 (2016).
157 Wuest, S. E. *et al.* Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proc Natl Acad Sci U S A* **109**, 13452-13457, doi:10.1073/pnas.1207075109 (2012).
158 Yant, L. *et al.* Orchestration of the floral transition and floral development in Arabidopsis by the bifunctional transcription factor APETALA2. *Plant Cell* **22**, 2156-2170, doi:10.1105/tpc.110.075606 (2010).
159 Zhiponova, M. K. *et al.* Helix-loop-helix/basic helix-loop-helix transcription factor network represses cell elongation in Arabidopsis through an apparent incoherent feed-forward loop. *Proc Natl Acad Sci U S A* **111**, 2824-2829, doi:10.1073/pnas.1400203111 (2014).
160 Kelley, D. R., Skinner, D. J. & Gasser, C. S. Roles of polarity determinants in ovule development. *Plant J* **57**, 1054-1064, doi:10.1111/j.1365-313X.2008.03752.x (2009).
161 Dello Ioio, R. *et al.* A PHABULOSA/cytokinin feedback loop controls root growth in Arabidopsis. *Curr Biol* **22**, 1699-1704, doi:10.1016/j.cub.2012.07.005 (2012).
162 Lee, C. & Clark, S. E. A WUSCHEL-Independent Stem Cell Specification Pathway Is Repressed by PHB, PHV and CNA in Arabidopsis. *PLoS One* **10**, e0126006, doi:10.1371/journal.pone.0126006 (2015).
163 Mandel, T. *et al.* The ERECTA receptor kinase regulates Arabidopsis shoot apical meristem size, phyllotaxy and floral meristem identity. *Development* **141**, 830-841, doi:10.1242/dev.104687 (2014).
164 Sessa, G., Morelli, G. & Ruberti, I. DNA-binding specificity of the homeodomain-leucine zipper domain. *J Mol Biol* **274**, 303-309, doi:10.1006/jmbi.1997.1408 (1997).
165 Lau, O. S. *et al.* Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. *Science* **345**, 1605-1609, doi:10.1126/science.1256888 (2014).
166 Aoyama, T. & Chua, N. H. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J* **11**, 605-612 (1997).
167 Gendron, J. M. *et al.* Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proc Natl Acad Sci U S A* **109**, 3167-3172, doi:10.1073/pnas.1200355109 (2012).
168 Wigge, P. A. *et al.* Integration of spatial and temporal information during floral induction in Arabidopsis. *Science* **309**, 1056-1059, doi:10.1126/science.1114358 (2005).
169 Kagale, S. & Rozwadowski, K. EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. *Epigenetics* **6**, 141-146 (2011).
170 Geisberg, J. V. & Struhl, K. Quantitative sequential chromatin immunoprecipitation, a method for analyzing co-occupancy of proteins at genomic regions in vivo. *Nucleic Acids Res* **32**, e151, doi:10.1093/nar/gnh148 (2004).
171 Bartrina, I., Otto, E., Strnad, M., Werner, T. & Schmulling, T. Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in Arabidopsis thaliana. *Plant Cell* **23**, 69-80, doi:10.1105/tpc.110.079079 (2011).
172 Chickarmane, V. S., Gordon, S. P., Tarr, P. T., Heisler, M. G. & Meyerowitz, E. M. Cytokinin signaling as a positional cue for patterning the apical-basal axis of the growing Arabidopsis shoot meristem. *Proc Natl Acad Sci U S A* **109**, 4002-4007, doi:10.1073/pnas.1200636109 (2012).
173 Kurakawa, T. *et al.* Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature* **445**, 652-655, doi:10.1038/nature05504 (2007).
174 Sebastian, J. *et al.* PHABULOSA controls the quiescent center-independent root meristem activities in Arabidopsis thaliana. *PLoS Genet* **11**, e1004973, doi:10.1371/journal.pgen.1004973 (2015).
175 Yanai, O. *et al.* Arabidopsis KNOXI proteins activate cytokinin biosynthesis. *Curr Biol* **15**, 1566-1571, doi:10.1016/j.cub.2005.07.060 (2005).
176 Zurcher, E. *et al.* A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant Physiol* **161**, 1066-1075, doi:10.1104/pp.112.211763 (2013).
177 Gruel, J. *et al.* An epidermis-driven mechanism positions and scales stem cell niches in plants. *Sci Adv* **2**, e1500989, doi:10.1126/sciadv.1500989 (2016).

178 Schmid, M. *et al.* A gene expression map of Arabidopsis thaliana development. *Nat Genet* **37**, 501-506, doi:10.1038/ng1543 (2005).
179 Jolma, A. *et al.* DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384-388, doi:10.1038/nature15518 (2015).
180 Crawford, B. C. *et al.* Plant development. Genetic control of distal stem cell fate within root and embryonic meristems. *Science* **347**, 655-659, doi:10.1126/science.aaa0196 (2015).
181 Liu, Y. G. *et al.* Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning. *Proc Natl Acad Sci U S A* **96**, 6535-6540 (1999).
182 Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. & Copeland, N. G. Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res* **33**, e36, doi:10.1093/nar/gni035 (2005).
183 Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**, 735-743 (1998).
184 Honda, S. I., Honglada.T & Laties, G. G. A New Isolation Medium for Plant Organelles. *J Exp Bot* **17**, 460-&, doi:DOI 10.1093/jxb/17.3.460 (1966).
185 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
186 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
187 Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284, doi:10.1186/1471-2164-15-284 (2014).
188 Wang, S. *et al.* Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat Protoc* **8**, 2502-2515, doi:10.1038/nprot.2013.150 (2013).
189 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578, doi:10.1038/nprot.2012.016 (2012).
190 Moissiard, G. *et al.* Transcriptional gene silencing by Arabidopsis microrchidia homologues involves the formation of heteromers. *Proc Natl Acad Sci U S A* **111**, 7474-7479, doi:10.1073/pnas.1406611111 (2014).
191 Krogan, N. T. & Long, J. A. Why so repressed? Turning off transcription during plant growth and development. *Curr Opin Plant Biol* **12**, 628-636, doi:10.1016/j.pbi.2009.07.011 (2009).
192 Courey, A. J. & Jia, S. T. Transcriptional repression: the long and the short of it. *Gene Dev* **15**, 2786-2796 (2001).
193 Gray, S. & Levine, M. Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. *Gene Dev* **10**, 700-710, doi:DOI 10.1101/gad.10.6.700 (1996).
194 Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. The ABC model of flower development: then and now. *Development* **139**, 4095-4098, doi:10.1242/dev.083972 (2012).
195 Coen, E. S. & Meyerowitz, E. M. The War of the Whorls - Genetic Interactions Controlling Flower Development. *Nature* **353**, 31-37, doi:DOI 10.1038/353031a0 (1991).
196 Wilbanks, E. G. & Facciotti, M. T. Evaluation of Algorithm Performance in ChIP-Seq Peak Detection. *Plos One* **5**, doi:ARTN e11471
10.1371/journal.pone.0011471 (2010).
197 Causier, B., Ashworth, M., Guo, W. J. & Davies, B. The TOPLESS Interactome: A Framework for Gene Repression in Arabidopsis. *Plant Physiol* **158**, 423-438, doi:10.1104/pp.111.186999 (2012).
198 Thireault, C. *et al.* Repression of jasmonate signaling by a non-TIFY JAZ protein in Arabidopsis. *Plant J* **82**, 669-679, doi:10.1111/tpj.12841 (2015).
199 Geerinck, J., Pauwels, L., De Jaeger, G. & Goossens, A. Dissection of the one-MegaDalton JAZ1 protein complex. *Plant Signal Behav* **5**, 1039-1041 (2010).
200 Nicolas, M. & Cubas, P. TCP factors: new kids on the signaling block. *Curr Opin Plant Biol* **33**, 33-41, doi:10.1016/j.pbi.2016.05.006 (2016).
201 Chen, G. H., Sun, J. Y., Liu, M., Liu, J. & Yang, W. C. SPOROCYTELESS Is a Novel Embryophyte-Specific Transcription Repressor that Interacts with TPL and TCP Proteins in Arabidopsis. *J Genet Genomics* **41**, 617-625, doi:10.1016/j.jgg.2014.08.009 (2014).
202 Leonid, T. B., Thurtl, D. M., Rine, J. & van Oudenaarden, A. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *P Natl Acad Sci USA* **110**, 18602-18607, doi:10.1073/pnas.1316064110 (2013).

203 Aida, M., Ishida, T. & Tasaka, M. Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. *Development* **126**, 1563-1570 (1999).
204 Hibara, K., Takada, S. & Tasaka, M. CUC1 gene activates the expression of SAM-related genes to induce adventitious shoot formation. *Plant J* **36**, 687-696, doi:10.1046/j.1365-313X.2003.01911.x (2003).
205 Laufs, P., Peaucelle, A., Morin, H. & Traas, J. MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development* **131**, 4311-4322, doi:10.1242/dev.01320 (2004).
206 Peaucelle, A., Morin, H., Traas, J. & Laufs, P. Plants expressing a miR164-resistant CUC2 gene reveal the importance of post-meristematic maintenance of phyllotaxy in Arabidopsis. *Development* **134**, 1045-1050, doi:10.1242/dev.02774 (2007).
207 Takada, S., Hibara, K., Ishida, T. & Tasaka, M. The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. *Development* **128**, 1127-1135 (2001).
208 Vroemen, C. W., Mordhorst, A. P., Albrecht, C., Kwaaitaal, M. A. C. J. & de Vries, S. C. The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in Arabidopsis. *Plant Cell* **15**, 1563-1577, doi:10.1105/tpc.012203 (2003).
209 Xu, B. *et al.* Arabidopsis genes AS1, AS2, and JAG negatively regulate boundary-specifying genes to promote sepal and petal development (vol 146, pg 566, 2008). *Plant Physiol* **146**, 2054-2054, doi:10.1104/pp.104.900253 (2008).
210 Smyth, D. R., Bowman, J. L. & Meyerowitz, E. M. Early Flower Development in Arabidopsis. *Plant Cell* **2**, 755-767 (1990).
211 Scofield, S. *et al.* Coordination of meristem and boundary functions by transcription factors in the SHOOT MERISTEMLESS regulatory network. *Development* **145**, doi:UNSP dev157081 10.1242/dev.157081 (2018).
212 Crawford, B. C. W. *et al.* Genetic control of distal stem cell fate within root and embryonic meristems. *Science* **347**, 655-659, doi:10.1126/science.aaa0196 (2015).
213 Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**, 735-743, doi:DOI 10.1046/j.1365-313x.1998.00343.x (1998).
214 Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9**, doi:ARTN R137 10.1186/gb-2008-9-9-r137 (2008).
215 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-W165, doi:10.1093/nar/gkw257 (2016).
216 Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178-192, doi:10.1093/bib/bbs017 (2013).
217 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).
218 Yu, G. C., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383, doi:10.1093/bioinformatics/btv145 (2015).
219 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks (vol 7, pg 562, 2012). *Nat Protoc* **9**, 2513-2513, doi:10.1038/nprot1014-2513a (2014).
220 Turki-Judeh, W. & Courey, A. J. Groucho: A Corepressor with Instructive Roles in Development. *Curr Top Dev Biol* **98**, 65-96, doi:10.1016/B978-0-12-386499-4.00003-3 (2012).
221 Kieffer, M. *et al.* Analysis of the transcription factor WUSCHEL and its functional homologue in Antirrhinum reveals a potential mechanism for their roles in meristem maintenance. *Plant Cell* **18**, 560-573, doi:10.1105/tpc.105.039107 (2006).
222 Gallavotti, A. *et al.* The control of axillary meristem fate in the maize ramosa pathway. *Development* **137**, 2849-2856, doi:10.1242/dev.051748 (2010).
223 Migliori, V., Mapelli, M. & Guccione, E. On WD40 proteins Propelling our knowledge of transcriptional control? *Epigenetics-Us* **7**, 815-822, doi:10.4161/epi.21140 (2012).
224 Han, Z. F. *et al.* Structural basis for the specific recognition of methylated histone H3 lysine 4 by the WD-40 protein WDR5. *Mol Cell* **22**, 137-144, doi:10.1016/j.molcel.2006.03.018 (2006).
225 Couture, J. F., Collazo, E. & Trievel, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat Struct Mol Biol* **13**, 698-703, doi:10.1038/nsmb1116 (2006).
226 Chanvivattana, Y. *et al.* Interaction of polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**, 5263-5276, doi:DOI 10.1242/dev.01400 (2004).

227 Yang, H. C. *et al.* Distinct phases of Polycomb silencing to hold epigenetic memory of cold in Arabidopsis. *Science* **357**, 1142-1145, doi:10.1126/science.aan1121 (2017).
228 Mayer, K. F. X. *et al.* Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* **95**, 805-815, doi:Doi 10.1016/S0092-8674(00)81703-1 (1998).
229 Hardtke, C. S. & Berleth, T. The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *Embo J* **17**, 1405-1411, doi:DOI 10.1093/emboj/17.5.1405 (1998).
230 Xiao, J. *et al.* Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat Genet* **49**, 1546-+, doi:10.1038/ng.3937 (2017).
231 Efroni, I. & Birnbaum, K. D. The potential of single-cell profiling in plants. *Genome Biol* **17**, doi:ARTN 65 10.1186/s13059-016-0931-2 (2016).
232 Deal, R. B. & Henikoff, S. A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue. *Dev Cell* **18**, 1030-1040, doi:10.1016/j.devcel.2010.05.013 (2010).
233 Tan, L. M. *et al.* The PEAT protein complexes are required for histone deacetylation and heterochromatin silencing. *Embo J*, doi:10.15252/embj.201798770 (2018).
234 Blevins, T. *et al.* Identification of Pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in Arabidopsis. *Elife* **4**, doi:ARTN e09591 10.7554/eLife.09591 (2015).
235 Earley, K. W. *et al.* Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. *Gene Dev* **24**, 1119-1132, doi:10.1101/gad.1914110 (2010).