Functional specialization of Arabidopsis VEL polymerization domains in the switch to Polycomb silencing

Anna Schulten[1], Geng-Jen Jang[1], Alex Payne-Dwyer[2,3], Mathias L. Nielsen[1],†, Mariann Bienz[4], Mark C. Leake[2,3], Caroline Dean[1,4]*

[1]Department of Cell and Developmental Biology, John Innes Centre; Norwich, United Kingdom.

[2]School of Physics, Engineering and Technology, University of York; York, United Kingdom.

[3]Department of Biology, University of York; York, United Kingdom.

[4] Medical Research Council Laboratory of Molecular Biology, Cambridge, United Kingdom.

†Present address: Department of Immunology and Regenerative Biology, Weizmann Institute of Science; Rehovot, Israel.

*Corresponding author. Email: caroline.dean@jic.ac.uk

Abstract:

Epigenetic gene silencing by Polycomb Repressive Complex 2 (PRC2) is essential for development in eukaryotes, yet what initiates silencing is still unclear. Polycomb silencing at Arabidopsis *FLOWERING LOCUS C* (*FLC*) requires PRC2 and accessory proteins VIN3 and VRN5, both containing a structurally conserved polymerization module. Here, we show that polymerization of the VIN3 VEL domain increases the size of nuclear assemblies, with enhanced local concentration promoting chromatin association. This increased avidity enables effective H3K27me3 nucleation, underlying the digital switch to an epigenetically silenced state. However, VRN5 VEL is not essential for this PRC2 nucleation and VRN5 VEL is unable to replace VIN3 VEL function. This work thus defines a specific role for polymerization-mediated multivalency in initiating and maintaining PRC2 nucleation and extends our understanding of the principles that initiate epigenetic silencing.

Main Text:

The activity of the Polycomb Repressive Complex 2 (PRC2) is essential for the deposition of H3K27me3 as one of the hallmarks of mitotically inheritable gene silencing. Since the first identification of *cis*-regulatory Polycomb Response Elements (PREs) in Drosophila, a complex picture has emerged in which non-sequential multifactorial protein interactions with the local chromatin environment, also shaped by transcriptional activity, give rise to genome-wide Polycomb silencing patterns (*1-4*). Accessory proteins that interact with the highly conserved PRC2 core complex crucially define PRC2 subcomplexes with different activities and genomic enrichment sites (*5-10*). These mediate nucleation of Polycomb complexes and are required to maintain the Polycomb silenced state through cell division (*10-12*). Nucleation generally involves local deposition of H3K27me3 over a small number of nucleosomes (*10, 13*), so it is unclear how accessory protein function facilitates the digital switch to an epigenetically stable state and ensures inheritance of the silenced state to daughter cells during DNA replication.

In the plant *Arabidopsis thaliana*, proteins of the VEL family (VIN3, VRN5 and VEL1) have been identified as accessory proteins to a PRC2 complex with the core subunit VRN2 (one of

three homologs of mammalian SUZ12) (*14, 15*). Their genome-wide localization suggests a
widespread role in PRC2 silencing of many loci in the Arabidopsis genome (*15*). Except for
VEL1, their function is essential for the epigenetic silencing of the floral repressor gene
FLOWERING LOCUS C (*FLC)* during the winter cold, a process called vernalization (*16-19*).
At *FLC*, Polycomb silencing initiates via cold-induced nucleation of H3K27me3 over the first
exon-intron boundary of *FLC*. This confers metastable epigenetic silencing that holds the silent
state for ~10 cell cycles. The metastable state is converted into a long-term epigenetically
silenced state through the spreading of H3K27me3 across the *FLC* gene body, which occurs
when the plants start to grow more rapidly following the return to ambient temperatures (*13*).
Nucleation, which is strongly reduced in *vin3*, *vrn5* and *vrn2* single mutants, can be separated
from the stable spread Polycomb state in mutants defective in one of the PRC2
methyltransferases genes *CURLY LEAF* (*CLF*), or *LIKE HETEROCHROMATIN PROTEIN 1*
required to maintain long-term silencing (*20*). Studies in these mutants have shown that VEL
proteins promote the stochastic PRC2-mediated nucleation of individual *FLC* alleles, whereby
the fraction of nucleated *FLC* loci increases over prolonged cold of winter (*21*). In other words,
the slow accumulation of H3K27me3 at the 3-4 nucleosome nucleation region is the result of
a low probability digital ON/OFF switch (i.e. PRC2-VEL bound or not), which occurs
independently at each allele (*13, 22*).

A key question was which features of the VEL proteins confer the ability to aid PRC2
engagement with these local sites, and to maintain epigenetic silencing through cell division.
The VEL proteins share a domain architecture consisting of a tripartite plant homeodomain
(PHD) superdomain, a fibronectin type III (FNIII) domain and a C-terminal VEL domain (*23,
24*) (Figure 1A). Early data assumed they may be functionally similar, with one protein being
cold-induced (VIN3) and the others (VRN5, VEL1) constitutively expressed. However, recent
structural analysis has revealed that specifically in VRN5 there is a close packing of the central
PHD superdomain and FNIII domain and this mediates its interaction with the PRC2 core
complex. In contrast, VIN3 has a more open conformation of these domains and depends on
VRN5 to interact with PRC2 (*15*).

The VEL domain mediates homo- and heterotypic interactions between the VEL proteins.
Structural and functional analysis of the VEL domain revealed it to be a head-to-tail
polymerization domain, conserved through the green lineage (*25*). It is only the third head-to-
tail polymerization domain described in biology to date, their discovery often hindered by the
insolubility of protein polymers in solution (*26*). We found that the spontaneous polymerization
of VIN3 and VEL1 via their VEL domains drives their assembly into dynamic biomolecular
condensates detected by confocal microscopy following protein expression in heterologous
systems (*25*). Mathematical modelling of epigenetic state switching and memory at *FLC* had
predicted the requirement of additional protein memory storage elements that positively
feedback to reinforce themselves; for example, molecular assemblies that are maintained in
sufficiently high numbers to persist at the locus even through the nucleosome perturbations
that occur during DNA replication (*27*). VEL-dependent protein polymerization and condensate
formation could provide such a mechanism. Here, we therefore experimentally tested the
importance of VEL polymerization domain in the switch to the epigenetically stable silenced
state.

Distinct PRC2-related phenotypes in VIN3 and VRN5 VEL domain mutants

We had previously shown that stable Arabidopsis lines expressing VIN3-GFP with single amino acid polymerization blocking point mutations located in the head and tail interface of the VEL domain, respectively (R556D/I575D [RI>DD], Fig. 1A), fail to rescue impaired *FLC* silencing in the *vin3* mutant (*25*). This pointed to the importance of polymerization in *FLC* silencing. However, these assays used multi-copy transgenes; since polymerization is concentration-dependent we needed to directly compare the effect of VIN3-GFP wild-type (WT) and RI>DD transgenes when expressed at endogenous levels. Two homozygous single insertion lines with protein levels equal to a single-insertion VIN3-GFP WT line (Fig. 1B, E) were selected from a larger T2 generation, all displaying the same non-complementation phenotype (Fig. S1). Stable lines with a full deletion of the VEL domain in VIN3-GFP mirrored the impaired *FLC* shutdown during cold exposure and post-cold reactivation observed for VIN3-GFP RI>DD (Fig. 1A and B, and Fig. S1). Levels of H3K27me3 at the *FLC* locus in VIN3-GFP WT and VIN3-GFP RI>DD lines were measured using chromatin immunoprecipitation (ChIP). As in the *vin3* mutant, the cold-induced H3K27me3 peak at the nucleation region was disrupted in VIN3-GFP RI>DD (Fig. 1C). We thus conclude that the polymerization feature of the VIN3 VEL domain is required to promote PRC2-mediated deposition of H3K27me3 in the *FLC* nucleation region.

Our prediction from the heterotypic interactions observed between VIN3 and VRN5 was that the VEL domain would be required for silencing by mediating the VRN5-dependent interaction between VIN3 and PRC2 (*15, 16*). However, independent homozygous stably transformed Arabidopsis VRN5-SYFP2 ΔVEL lines with single transgene insertions in a *vrn5* mutant background (Fig. 1A and Fig. S2) showed an unexpected result; *FLC* was fully silenced comparable to silencing in *vrn5* lines expressing VRN5-SYFP2 WT (Fig. 1B). The deposition of H3K27me3 at the *FLC* nucleation region was not impaired in VRN5 ΔVEL plants (Fig. 1C).

To further understand the different phenotypes observed for the VEL domain deletions of VIN3 and VRN5 *in vivo,* we determined the interaction partners of VIN3 and VRN5 dependent on the VEL domain. Native GFP co-immunoprecipitation (coIP) followed by mass spectrometry was undertaken with respective WT and ΔVEL seedlings vernalized for six weeks. Both VIN3 and VRN5 WT proteins co-precipitated all components of one of the Arabidopsis PRC2 core complexes (VRN2, FIE, MSI1, SWN and to a lesser extent CLF) as well as each other and VEL1 as expected (Fig. 1D). Consistent with our prediction, the deletion of the VIN3 VEL domain resulted in the loss of enrichment of the PRC2 core subunits and VRN5/VEL1. Probing for the PRC2 subunit FIE with immunoblot analysis after coIP in extracts of vernalized plants confirmed the loss of FIE interaction in both VIN3 ΔVEL and VIN3 RI>DD lines (Fig. 1E). In contrast, the interaction with PRC2 was maintained in VRN5 ΔVEL plants (Fig. 1D). This is in accordance with our previous results from heterologous coIPs in mammalian cells which mapped the PHD superdomain and the FNIII of VRN5, but not the VEL domain to the interface for PRC2 interaction (*15*). Surprisingly, the interaction of VRN5 ΔVEL with VEL1 was also maintained, whereas the co-enrichment of VIN3 was strongly reduced. Although VEL1 has the same open conformation of the PHD superdomain and the FNIII domain as VIN3, its ability to interact specifically with the PRC2 core subunit MSI1 may explain the association with VRN5 (*15*). It remains unclear whether VEL1 function might also contribute to the rescue of *FLC* silencing observed in VRN5 ΔVEL lines.

Role of VEL domain in nuclear assembly formation of VIN3 and VRN5

The molecular assemblies predicted to be involved in Polycomb epigenetic switching could be realised by VEL polymerization-based assemblies or by multimeric assemblies mediated by other interactors such as PRC2 (or combinations thereof). We therefore used the stable Arabidopsis transgenic VIN3-GFP and VRN5-SYFP2 lines to investigate the effects of VEL domain disruptions on the stoichiometry of both VIN3 and VRN5 in molecular assemblies *in vivo*. We performed single particle tracking of the fluorescently tagged proteins by using Slimfield microscopy (*28, 29*) and employing stepwise photobleaching to count the number of tagged molecules in detected fluorescent foci (*30*) but using an oblique illumination angle of 60 ± 5°degrees to enhance imaging contrast to enable single-molecule detection sensitivity.

In roots of seedlings vernalized for 6 weeks, we observed that the VIN3-GFP RI>DD mutations resulted in VIN3 assemblies with lower number of molecules, only 29% of tracks corresponded to assemblies larger than 10 molecules compared to 46% in VIN3-GFP WT (Fig. 2A). No reduction in molecular assembly size was found for VRN5-SYFP2 ΔVEL in seedlings vernalized for 6 weeks (Fig. 2B). Based on the track stoichiometry distribution, we calculated the periodicity, i. e. the intervals between nearest-neighbour peaks, of VIN3 and VRN5 within the assemblies as a proxy for their internal order. VIN3-GFP WT and VRN5-SYFP2 WT consistently showed dimeric units, which were disrupted in VIN3-GFP RI>DD but generally maintained for VRN5-SYFP2 ΔVEL (Fig. 2C, D). Notably, we have previously observed VIN3 VEL domain crystals composed of protofilaments with a dimeric unit; the result of a mutual domain swapping of VEL domain helices between two VIN3 monomers (*25*). It remains to be determined whether PRC2 might also contribute to VIN3 assembly formation during vernalization. Assembly size and order appear to be VEL-independent for VRN5, concurrent with VRN5-SYFP2 ΔVEL complementing the *vrn5* mutant. Given the direct interaction between VRN5 and PRC2, we speculate that VEL-independent dimeric periodicity of VRN5 may be a result of the reported ability of PRC2 to undergo dimer formation (*31-33*).

The lack of effect on VRN5 stoichiometry and *FLC* silencing in lines carrying VRN5 ΔVEL led us to question the properties of the VRN5 VEL domain. Because VIN3 polymerization results in the formation of biomolecular condensates, we thus investigated foci formation of VRN5 after transient overexpression in leaf epidermal cells of *Nicotiana benthamiana*. In comparison to the VEL- and polymerization dependent discrete foci formed by GFP-VIN3 (*25*), GFP-VRN5 formed much smaller foci, with more VRN5 protein diffusely distributed throughout the nucleus (Fig. 2E and Fig. S3A, B). Upon expression in HeLa cells, GFP-VRN5 previously also appeared diffuse (*25*). Tagging VRN5 with mScarlet gave the same result as observed for GFP-VRN5 in the *N. benthamiana* system, where the deletion of the VRN5 VEL domain abolished the formation of any discrete foci (Fig. S3C-D). These observations are consistent with VRN5 having a lower propensity to concentrate into foci in a VEL-dependent manner. In confirmation of our in vivo coIP results, we observed the recruitment of mScarletI-VRN5 into GFP-VIN3 foci in the *N. benthamiana* system, a co-localization that was dependent on the VRN5 VEL domain (Fig. 2F-H). Thus, the VRN5 VEL domain has the ability to homo- and heteropolymerize but its properties appear different to the VIN3 VEL domain.

VRN5 VEL domain is not functionally equivalent to VIN3 VEL

To further examine the properties of the VRN5 VEL domain, we tested whether VRN5 VEL can functionally replace VIN3 VEL in *FLC* silencing. A VIN3-GFP construct, in which the VIN3 VEL domain was replaced with the VRN5 VEL domain (Fig. 3A), was transformed into the *vin3* mutant. *FLC* silencing during vernalization remained impaired in two independent homozygous lines as well as in multiple other lines tested in the T2 generation (Fig. 3B and Fig. S4). This would suggest that VRN5 VEL is not functionally equivalent to VIN3. With equal VIN3 protein expression, co-immunoprecipitation of the PRC2 subunit FIE was barely detected in the VIN3-GFP VRN5VEL line compared to VIN3-GFP WT (Fig. 3C), presumably due to inefficient interaction with endogenous VRN5. This result is consistent with reduced foci formation of VRN5 in *Nicotiana benthamiana* cells (Fig. 2E). The experimentally determined structures of VIN3 and VEL1 VEL domains align closely with the Alphafold (AF) structure predictions (*25*) and comparing the latter to the AF prediction for the VRN5 VEL domain again shows a close superimposition (Fig. 3D). Thus, no major structural differences are predicted to underpin the functional differences between the VEL domains. Specific amino acid differences (either individual or combinations thereof), possibly associated with different post-translational modifications *in vivo*, may contribute to the functional differences between VIN3 and VRN5 VEL and deserve further investigation in the future (Fig. 3E). The residue $I575_{VIN3}/I664_{VEL1}$ in the VEL head interface engages in hydrophobic or electrostatic interactions with two basic residues in the complementary tail surface to mediate polymerization, blocked by the mutation I575D. A threonine in the corresponding position of Arabidopsis VRN5 (T565) and other VRN5 paralogs may be functionally significant as this is a neutral rather than a hydrophobic amino acid (*25*).

Multivalency emerging from polymerization promotes VIN3 chromatin association

To understand the specific contribution of VIN3 polymerization to *FLC* silencing, we determined whether the loss of PRC2 in VIN3-GFP RI>DD, VIN3-GFP ΔVEL as well as VIN3-GFP *vrn5* would affect the association of VIN3 with the *FLC* locus. ChIP-qPCR experiments with vernalized seedlings revealed a strongly reduced association of VIN3-GFP RI>DD and VIN3-GFP ΔVEL with the *FLC* nucleation region (Fig. 4A and B). We also tested other VIN3 targets, previously identified by ChIP-seq experiments (*15*), and found reduced VIN3-GFP RI>DD association at several of these loci (Fig. S5). In contrast, VIN3-GFP *vrn5* showed the same binding as VIN3-GFP WT at *FLC* (Fig. 4C), suggesting that the interaction between VIN3 and PRC2 is not required to maintain VIN3 chromatin binding if the VIN3 protein is intact. This further implicates VEL-mediated polymerization in promoting and maintaining VIN3 chromatin association with *FLC* and other loci in a PRC2-independent manner. The chromatin association of VRN5 ΔVEL at the *FLC* locus was WT-like in comparison (Fig. S6).

An emerging paradigm for the function of head-to-tail protein polymerization is the increase in local concentration of the polymerizing protein and its ligand binding sites, whereby combinatorial multivalent interactions with the ligand result in enhanced binding avidity (functional affinity) (*26*). To investigate which VIN3 domains might contribute to a polymerization-based mechanism for chromatin association, we turned to the tripartite PHD superdomain of VIN3. Although this domain is an atypical PHD domain in that it exhibits no histone H3 tail binding activity, it has a weak affinity for negatively charged DNA or RNA polymers in vitro (*24*). We generated stable transgenic Arabidopsis plants carrying VIN3-GFP with a deletion of the entire PHD superdomain (ΔPHDsuper) in the *vin3* mutant background.

The construct did not complement the *vin3* mutant in *FLC* vernalization time course
experiments in two homozygous lines or in multiple other lines tested in the T2 generation
(Fig. 4D and Fig. S7). While VIN3 ΔPHDsuper maintained its interaction with PRC2 based on
FIE immunoprecipitation in vernalized seedlings (Fig. 4E), chromatin association with the *FLC*
nucleation region was abolished (Fig. 4F). This demonstrates that the VIN3 PHD superdomain
is necessary for chromatin association and has chromatin (or chromatin-associated) ligands
other than histone H3 tails. Interestingly, H2A was among the significant VIN3 interactors
identified by native coIP-MS (Fig. 1D), more specifically the histone variant H2A.W which is
known to associate with H3K9me2-marked heterochromatin (*34*). For VIN3-GFP RI>DD, the
binding peak at the *FLC* nucleation region was much smaller than for VIN3-GFP WT but not
entirely abolished (Fig. 4A). This may reflect the weak chromatin affinity of VIN3 monomers,
mediated by the PHD superdomain in the absence of the VIN3 polymer. These findings
suggest that VEL-mediated polymerization of VIN3 promotes a high local concentration of the
PHD superdomain, enhancing avidity to mediate initiation and maintenance of VIN3 chromatin
association. In a wild-type scenario, such prolonged chromatin association of VIN3 would also
increase PRC2 residence time, bridged by VRN5, to facilitate H3K27me3 nucleation.
238
That the VRN5 VEL domain is not essential for H3K27me3 nucleation suggests that the
association of the VIN3 polymer promotes H3K27me3 nucleation by other means in addition
to the recruitment of VRN5 and PRC2, eg. by modulating chromatin properties such as
nucleosome dynamics or by binding other protein factors. We previously found an interaction
between VIN3 and the transcriptional repressor VAL1, which serves as an assembly platform
for co-transcriptional repressors and chromatin regulators (*15, 35*). VAL1 binds to two RY-
motifs in the first *FLC* intron and would thus also provide a sequence-specific link to the *FLC*
locus (*36*). However, while transgenic Arabidopsis plants carrying a point mutation in the first
RY site fail to nucleate H3K27me3, they still show an accumulation of H3K27me2 specifically
in the nucleation region, indicating that VAL1 is only one of the multifactorial mechanisms that
ensure targeting of VEL-PRC2 to this region (*15, 36*). Our findings for the VEL proteins have
striking resemblance to a polymerization network involving multiple Polycomb Repressive
Complex 1 (PRC1) subunits that engage in a combination of heterotypic and homotypic
interactions to promote transcriptional silencing in Drosophila. These interactions are
mediated by the head-to-tail co-polymerization of SAM domains of Polyhomeotic (Ph) and Sex
combs of midleg (Scm) and are linked to the Pho-repressive complex (PhoRC) subunit Sfmbt,
which undergoes heterotypic SAM interactions with Scm but is unable to homopolymerize (*37-*
39). The DNA-binding activity of the PhoRC complex contributes to this polymerization-
mediated hub to promote the nucleation of PRC1 complexes at target loci. SAM-containing
Polycomb proteins have not been found in plants (*40*), where VEL proteins may function with
PRC2 in mechanistic parallels.
260
Conclusions
In summary, we show that the PRC2 accessory protein VIN3 forms VEL polymerization-
dependent assemblies necessary to promote PRC2-mediated deposition of H3K27me3 in the
nucleation region of *FLC*. Although VIN3 and VRN5 interact via their VEL domains, which also
connects VIN3 with the PRC2 complex, the VEL domain of VRN5 is not essential for PRC2
nucleation activity and is also unable to replace VIN3 VEL function. We propose that the
polymerization of the VIN3 VEL domain results in a high local concentration of VIN3 to promote
its chromatin association, mediated by emergent multivalent interactions between VIN3 PHD
superdomains and chromatin ligands. This would both facilitate recruitment of VRN5-PRC2

and potentially influence other local chromatin properties to overcome the nucleation threshold
and promote H3K27me3 nucleation. Taken together, our analysis defines a distinct role for
VIN3 polymerization-mediated multivalency in the digital switch to the Polycomb silenced
state.

274
275
276

Materials and methods

Plant materials and growth conditions
Seeds were surface sterilized with chlorine gas, sown on MS media plates without glucose
(pH 5.7) and stratified at 4 °C in the dark for 2 days. For non-vernalized (NV, before cold)
conditions, seedlings were grown for 10 days in long-day conditions (16 h light, 8 h darkness,
20 °C). After this pre-growth, seedlings were vernalized for 6 weeks (8 h light, 16 h darkness,
5 °C) (6WT0). Vernalized seedlings were returned to NV conditions described above for 10
days (6WT10) for post-cold samples. For seed generation, seedlings were transferred to soil
after vernalization and cultivated in a glasshouse with controlled 22°C 16 h day and 20°C 8 h
night conditions.
Arabidopsis Col-FRI is Col-0 with an introgressed active Sf2 FRIGIDA allele (FRI) and was
previously described, as were the *vin3-1 FRI (vin3) (17)* and the *vrn5-8* FRI (*vrn5*) mutants
(*15*). The transgenic lines pVIN3:VIN3-eGFP/*vin3-4* FRI (VIN3-GFP #22)(*36*), pVIN3:VIN3-
eGFP/*vin3-4 vrn5-8* FRI (VIN3-GFP #22 *vrn5-8)(15)*, pVIN3:VIN3-eGFP/*vin3-1* FRI (VIN3-
GFP WT #3-8) and pVIN3:VIN3-eGFP R556D/I575D *vin3-1* FRI (VIN3-GFP RI>DD) (*25*) were
described previously.

Generation of transgenic Arabidopsis lines
All primers used for cloning newly generated constructs are listed in Supplemental Table S1.
The genomic pENTR pVIN3::VIN3-GFP construct (*36*) was modified to make pENTR
pVIN3::VIN3-GFPΔPHDsuper and pENTR pVIN3::VIN3-GFP ΔVEL by Quickchange, using
Phusion DNA Polymerase (Thermo Scientific). To generate pENTR pVIN3::VIN3-GFP
VRN5VEL, the VRN5 VEL domain was amplified from pENTR pVRN5-SYFP2 and swapped
into pENTR pVIN3-VIN3-GFP by restriction-free cloning.
To generate pENTR pVRN5::VRN5-SYFP2, the VRN5 genomic region including endogenous
promoter and terminator were amplified from genomic Col-0 DNA and cloned into pENTR
using In-Fusion cloning (Takara). SYFP2 was then inserted with In-Fusion cloning, followed
by the insertion of a 10 amino acid linker between VRN5 and SYFP2 with Quickchange PCR
using Phusion DNA Polymerase (Thermo Scientific). This wild-type construct was modified by
Quickchange, using Phsion DNA polymerase, to generate pENTR pVRN5::VRN5 -SYFP2
ΔVEL. All plasmids were verified by sequencing.

The pENTR-constructs were transferred to the binary vector pSLJ-DEST (based on
pSLJ755I6) with L/R reaction (Invitrogen) and then transformed into *vin3-1 FRI* or *vrn5-8* FRI
mutants mediated by Agrobacterium C58 using the floral dip method. Transgene copy number
was determined in T1 transformants by IDna Genetics (Norwich Research Park).

314
315
316

RNA extraction and qRT-PCR

RNA was extracted as described (*36*), using acidic phenol followed by lithium chloride precipitation. RNA was DNase treated with Turbo DNA Free DNase, then transcribed into cDNA with SuperScript Reverse Transcriptase IV (both Life Technologies) with gene-specific reverse primers (Supplemental Table S1). qPCR was performed using SYBRGreen Master Mix II on a LightCycler 480 II (both Roche) with primer pairs listed in Supplemental Table S1.

Co-immunoprecipitation and immunoblotting

For co-immunoprecipitation (coIP) analysis followed by immunoblot analyses, total proteins were extracted from 2-3 aliquots of 3 g frozen ground Arabidopsis seedling tissue with IP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% NP-40, 1% Triton-X, EDTA-free protease inhibitor cocktail [Roche]). Lysates were cleared by filtering through miracloth followed by centrifugation (6,000 g, 30 min, 4°C) and then incubated with GFP-Trap (Chromotek) for 4h. Immunoprecipitates were washed four times with IP buffer and eluted by boiling in 4x NuPAGE LDS sample buffer for 10 min. Input and coIP fractions were separated by polyacrylamide gel electrophoresis (SDS-PAGE) and blotted onto polyvinylidine difluoride (PVDF) membranes. Primary antibodies anti-GFP (11814460001, Roche) and anti-FIE (AS12 2616, Agrisera) were diluted 1:1000, anti-Actin was diluted 1:5000 (AS132640, Agrisera). Secondary antibodies were HRP-coupled. Blots were washed with TBS containing 0.05% Tween-20 and developed with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific).

Co-immunoprecipitation and mass spectrometry

For coIP followed by mass spectrometry, total proteins were extracted from 3 g of frozen ground Arabidopsis seedling tissue with the IP buffer described above, with the addition of PhosSTOP according to manufacturer instructions (4906845001, Roche). IP was performed as described above. Immunoprecipitates were eluted by boiling for 15 min in 20 mM Tris-HCl pH 8, 2% SDS. Proteins in the eluate were precipitated with chloroform/methanol (1:4) on ice for 30 min, the pellet was then washed twice with methanol and once with acetone before drying. Protein pellets were resuspended in 50 µl of 1.5% sodium deoxycholate (SDC; Merck) in 0.2 M EPPS-buffer (Merck), pH 8.5 and vortexed under heating. Cysteine residues were reduced with dithiothreitol, alkylated with iodoacetamide, and the proteins digested with trypsin in the SDC buffer according to standard procedures. After the digest, the SDC was precipitated by adjusting to 0.2% trifluoroacetic acid (TFA), and the clear supernatant subjected to C18 SPE using home-made stage tips with C18 Reprosil_pur 120, 5 µm (Dr Maisch, Germany). Aliquots were analysed by nanoLC-MS/MS on an Orbitrap Eclipse™ Tribrid™ mass spectrometer coupled to an UltiMate® 3000 RSLCnano LC system (Thermo Fisher Scientific, Hemel Hempstead, UK). The samples were loaded onto a trap cartridge (PepMap™ Neo Trap Cartridge, C18, 5um, 0.3x5mm, Thermo) with 0.1% TFA at 15 µl min^{-1} for 3 min. The trap column was then switched in-line with the analytical column (Aurora Frontier TS, 60 cm nanoflow UHPLC column, ID 75 µm, reversed phase C18, 1.7 µm, 120 Å; IonOpticks, Fitzroy, Australia) for separation at 55°C using the following gradient of solvents A (water, 0.1% formic acid) and B (80% acetonitrile, 0.1% formic acid) at a flow rate of 0.26 µl min^{-1} : 0-3 min 1% B (parallel to trapping); 3-10 min increase B (curve 4) to 8%; 10-102 min linear increase B to 48; followed by a ramp to 99% B and re-equilibration to 0% B, for a total of 140 min runtime. Mass spectrometry data were acquired with the FAIMS device set to three compensation voltages (-35V, -50V, -65V) at standard resolution for 1.0 s each with the following MS settings in positive ion mode: OT resolution 120K, profile mode, mass range m/z 300-1600, normalized AGC target 100%, max inject time 50 ms; MS2 in IT Turbo mode: quadrupole isolation window

1 Da, charge states 2-5, threshold 1e^4, HCD CE = 30, AGC target standard, max. injection
time dynamic, dynamic exclusion 1 count for 15 s with mass tolerance of ±10 ppm, one charge
state per precursor only.
The mass spectrometry raw data were processed and quantified in Proteome Discoverer 3.1
(Thermo); all mentioned tools of the following workflow are nodes of the proprietary Proteome
Discoverer (PD) software. The *A. thaliana* TAIR10 protein database (arabidopsis.org; 32785
entries) was modified by removing accessions AT4G30200.1, AT4G30200.3, and
AT4G30200.4 corresponding to 3 versions of the VEL1 protein. Only AT4G30200.2
corresponding to the canonical version of VEL1 was left in the database for clearer search
and quantification results. The database search including a decoy search was performed with
Mascot Server 2.8.3 (Matrixscience, London; in house server) with a fragment tolerance of 0.5
Da, enzyme trypsin with 2 missed cleavages, variable modifications were oxidation (M), Acetyl
(Protein N-term), phosphorylation (STY), methylation/dimethylation/trimethylation (K); fixed
modification carbamidomethyl (C).
Validation in PD was then performed using Percolator based on q-values and FDR targets
0.01 (strict) and 0.05 (relaxed). The workflow included the Minora Feature Detector with min.
trace length 7, S/N 3, PSM confidence high. The consensus workflow in the PD software was
used to evaluate the peptide identifications and to measure the abundances of the peptides
based on the LC-peak intensities. For identification, an FDR of 0.01 was used as strict
threshold, and 0.05 as relaxed threshold.
For quantification, three replicates per condition were measured. In PD3.1, the following
parameters were used for ratio calculation: normalisation on total peptide abundances, protein
abundance-based ratio calculation using the top3 most abundant peptides, missing values
imputation by low abundance resampling, hypothesis testing by t-test (background based),
adjusted p-value calculation by BH-method. The results were exported into a Microsoft Excel
table including data for protein abundances, ratios, p-values, number of peptides, protein
coverage, the search identification score and other important values.
Protein lists obtained for VIN3 and VRN5 wild-type proteins were filtered for interactors that
were positively enriched in comparison to a Col-FRI non-transgenic control sample with an
adjusted p-value ≤ 0.05. Enrichment ratios of interactors predicted to have a nuclear
localization were log2-transformed for both WT and mutant samples to generate the heatmaps
visualising the IP-MS results. The full mass spectrometry proteomics data have been
deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the
dataset identifier PXD048844 and 10.6019/PXD048844.
399
Chromatin immunoprecipitation (ChIP)
Histone ChIP was performed with 2 g of formaldehyde-crosslinked material as described
previously (*15*) with the following modifications: after nuclei extraction with Honda buffer, nuclei
were layered on a Percoll density gradient (75%/40% Percoll in Honda) and extracted from
the interface between these layers after centrifugation (7.000 x *g* in a swing bucket rotor for
30 min at 4 °C) prior to lysis of nuclei. Immunoprecipitation was performed with antibodies α-
H3K27me3 (Abcam, ab192985) and α-H3 antibody (Abcam, ab1791), using 3 µg per IP
reaction.
Non-histone ChIP (VIN3/VRN5) was performed as described for GFP/YFP-tagged proteins
(*15*). For lines with endogenous level VIN3 expression, each ChIP replicate was generated by
pooling chromatin from three aliquots of 3 g of formaldehyde-crosslinked material for IP.
Immunoprecipitation was performed with α-GFP (Abcam, ab290) using 3 µg per IP reaction.
412

Heterologous Nicotiana benthamiana transfections

The generation of *p35S:Ω-GFP-VIN3* was described previously (*25*). This plasmid was modified with seamless megaprimer cloning to generate *p35S:Ω-GFP-VRN5* and *p35S:Ω-mScarlet-VRN5* with the coding sequence of VRN5. Plasmids were transformed into *Agrobacterium tumefaciens* GV3101 using electroporation. Agrobacteria containing the desired construct at OD_{600} 0.05 were equally co-infiltrated with the silencing suppressor P19 into three-week-old *Nicotiana benthamiana* leaves. Confocal imaging of infiltrated epidermal leaf cells of *N. benthamiana* was performed on a Leica confocal Stellaris 8 microscope using a 63x/1.2 water objective and 4x zoom, excitation at 488 nm, detection at 507-542 nm for GFP and excitation at 561 nm, detection at 575-625 nm. Images were acquired 24 hr after infiltration with a laser speed of 600 Hz, a Z-step size of 4.7 µm and a pinhole size of 1 AU. The same settings were used at all imaging sets to allow direct comparison between constructs. The analysis was performed in Arivis Vision4D ver. 4.1.0. (Zeiss). Firstly, the blob finder algorithm was applied to the GFP channel using a diameter value of 0.8 mm, a probability threshold of 50%, and a split sensitivity of 65%. Then, the blob finder algorithm was applied to the mScarlet channel using the same settings for diameter value, probability, and split sensitivity. Afterwards, the intersection between the output of the two blob finder operations was calculated. Finally, metrics such as volume were computed for the objects generated by each of the blob finder operations, as well as for their intersection.

Acknowledgements

We thank Shuqin Chen and Caroline Smith for their excellent technical help. We are grateful to Gerhard Saalbach and Carlo Martin from the JIC proteomics facility for performing mass spectrometry analysis. We also thank all members of the Dean and Howard groups as well as Elsa Franco-Echevarría and Marc Fiedler for discussions. This work was funded by the European Research Council Advanced Grant (EPISWITCH-833254), Wellcome Trust (210654/Z/18/Z), Biotechnology and Biological Sciences Research Council Institute Strategic Programmes (BB/J004588/1 and BB/P013511/1), EPSRC (EP/T00214X/1, EP/T002166/1 and EP/W024063/1) and a Royal Society Professorship (RP\R1\180002) to C.D..

Author contributions

M.B. and C.D. conceived the study. A.S., G-J.Y., A.P.-D. and M.L.N. performed the investigation and analysed the data. A.S. and C.D. wrote the manuscript; all authors reviewed and edited the manuscript. M.C.L. and C.D. supervised the study, acquired the funding, and were the project administrators.

Data and materials availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD048844 and 10.6019/PXD048844.

References

1. A. Laugesen, J. W. Højfeldt, K. Helin, Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Molecular Cell* **74**, 8-18 (2019).
2. D. Holoch *et al.*, A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nature Genetics* **53**, 1686-1697 (2021).
3. J. Déjardin *et al.*, Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. *Nature* **434**, 533-538 (2005).
4. J. Müller, J. A. Kassis, Polycomb response elements and targeting of Polycomb group proteins in Drosophila. *Current Opinion in Genetics & Development* **16**, 476-484 (2006).
5. G. van Mierlo, G. J. C. Veenstra, M. Vermeulen, H. Marks, The Complexity of PRC2 Subcomplexes. *Trends in Cell Biology* **29**, 660-671 (2019).
6. M. Perino *et al.*, Two Functional Axes of Feedback-Enforced PRC2 Recruitment in Mouse Embryonic Stem Cells. *Stem Cell Reports* **15**, 1287-1300 (2020).
7. D. T. Youmans, A. R. Gooding, R. D. Dowell, T. R. Cech, Competition between PRC2.1 and 2.2 subcomplexes regulates PRC2 chromatin occupancy in human stem cells. *Molecular Cell* **81**, 488-501.e489 (2021).
8. A. Petracovici, R. Bonasio, Distinct PRC2 subunits regulate maintenance and establishment of Polycomb repression during differentiation. *Molecular Cell* **81**, 2625-2639.e2625 (2021).
9. S. Chen, L. Jiao, M. Shubbar, X. Yang, X. Liu, Unique Structural Platforms of Suz12 Dictate Distinct Classes of PRC2 for Chromatin Binding. *Molecular Cell* **69**, 840-852.e845 (2018).
10. O. Oksuz *et al.*, Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Molecular Cell* **70**, 1149-1162.e1145 (2018).
11. R. T. Coleman, G. Struhl, Causal role for inheritance of H3K27me3 in maintaining the OFF state of a Drosophila HOX gene. *Science* **356**, eaai8236 (2017).
12. F. Laprell, K. Finkl, J. Müller, Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science* **356**, 85-88 (2017).
13. H. Yang *et al.*, Distinct phases of Polycomb silencing to hold epigenetic memory of cold in Arabidopsis. *Science* **357**, 1142-1145 (2017).
14. F. De Lucia, P. Crevillen, A. M. Jones, T. Greb, C. Dean, A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc Natl Acad Sci U S A* **105**, 16831-16836 (2008).
15. E. Franco-Echevarría *et al.*, Distinct accessory roles of Arabidopsis VEL proteins in Polycomb silencing. *Genes & Development* **37**, 801-817 (2023).
16. T. Greb *et al.*, The PHD Finger Protein VRN5 Functions in the Epigenetic Silencing of Arabidopsis FLC. *Current Biology* **17**, 73-78 (2007).
17. S. Sung, R. M. Amasino, Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* **427**, 159-164 (2004).
18. A. R. Gendall, Y. Y. Levy, A. Wilson, C. Dean, The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis. *Cell* **107**, 525-535 (2001).
19. D.-H. Kim, S. Sung, The Plant Homeo Domain finger protein, VIN3-LIKE 2, is necessary for photoperiod-mediated epigenetic regulation of the floral repressor, MAF5. *Proceedings of the National Academy of Sciences* **107**, 17029-17034 (2010).

20. J. S. Mylne *et al.*, LHP1, the Arabidopsis homologue of HETEROCHROMATIN
PROTEIN1, is required for epigenetic silencing of *FLC*. *Proceedings of the National*
Academy of Sciences **103**, 5012-5017 (2006).
21. A. Angel *et al.*, Vernalizing cold is registered digitally at *FLC*. *Proceedings of the*
National Academy of Sciences **112**, 4146-4151 (2015).
22. S. Berry, C. Dean, M. Howard, Slow Chromatin Dynamics Allow Polycomb Target
Genes to Filter Fluctuations in Transcription Factor Activity. *Cell Systems* **4**, 445-
457.e448 (2017).
23. J. Mylne, T. Greb, C. Lister, C. Dean, Epigenetic Regulation in the Control of Flowering.
Cold Spring Harbor Symposia on Quantitative Biology **69**, 457-464 (2004).
24. E. Franco-Echevarría, T. J. Rutherford, M. Fiedler, C. Dean, M. Bienz, Plant
vernalization proteins contain unusual PHD superdomains without histone H3 binding
activity. *J Biol Chem* **298**, 102540 (2022).
25. M. Fiedler *et al.*, Head-to-tail polymerization by VEL proteins underpins cold-induced
Polycomb silencing in flowering control. *Cell Reports* **41**, (2022).
26. M. Bienz, Head-to-Tail Polymerization in the Assembly of Biomolecular Condensates.
Cell **182**, 799-811 (2020).
27. C. Lövkvist *et al.*, Hybrid protein assembly-histone modification mechanism for PRC2-
based epigenetic switching and memory. *eLife* **10**, e66454 (2021).
28. M. Plank, G. H. Wadhams, M. C. Leake, Millisecond timescale slimfield imaging and
automated quantification of single fluorescent protein molecules for use in probing
complex biological processes. *Integrative Biology* **1**, 602-612 (2009).
29. A. L. Payne-Dwyer, M. C. Leake, in *Chromosome Architecture: Methods and Protocols,*
M. C. Leake, Ed. (Springer US, New York, NY, 2022), pp. 311-328.
30. M. C. Leake *et al.*, Stoichiometry and turnover in single, functioning membrane protein
complexes. *Nature* **443**, 355-358 (2006).
31. S. Chen, L. Jiao, X. Liu, X. Yang, X. Liu, A Dimeric Structural Scaffold for PRC2-PCL
Targeting to CpG Island Chromatin. *Molecular Cell* **77**, 1265-1278.e1267 (2020).
32. D. Grau *et al.*, Structures of monomeric and dimeric PRC2:EZH1 reveal flexible
modules involved in chromatin compaction. *Nature Communications* **12**, 714 (2021).
33. P. V. Sauer *et al.*, Activation of automethylated PRC2 by dimerization on chromatin.
bioRxiv, 2023.2010.2012.562141 (2023).
34. Z. J. Lorković *et al.*, Compartmentalization of DNA Damage Response between
Heterochromatin and Euchromatin Is Mediated by Distinct H2A Histone Variants.
Current Biology **27**, 1192-1199 (2017).
35. P. Mikulski *et al.*, VAL1 acts as an assembly platform co-ordinating co-transcriptional
repression and chromatin regulation at Arabidopsis FLC. *Nature Communications* **13**,
5542 (2022).
36. J. I. Qüesta, J. Song, N. Geraldo, H. An, C. Dean, Arabidopsis transcriptional repressor
VAL1 triggers Polycomb silencing at *FLC* during vernalization. *Science* **353**, 485-488
(2016).
37. F. Frey *et al.*, Molecular basis of PRC1 targeting to Polycomb response elements by
PhoRC. *Genes Dev* **30**, 1116-1127 (2016).
38. H. Kang *et al.*, Sex comb on midleg (Scm) is a functional link between PcG-repressive
complexes in Drosophila. *Genes & Development* **29**, 1136-1150 (2015).
39. C. Alfieri *et al.*, Structural basis for targeting the chromatin repressor Sfmbt to
Polycomb response elements. *Genes Dev* **27**, 2367-2379 (2013).

40.   Y. Huang *et al.*, Evolution and conservation of polycomb repressive complex 1 core
555   components and putative associated factors in the green lineage. *BMC Genomics* **20**,
556   533 (2019).

Figure 1: VIN3 and VRN5 VEL domain mutant lines have different phenotypes

(A) Domain architecture of Arabidopsis VEL proteins VIN3 and VRN5. The tripartite PHD superdomain consists of a zinc finger, an atypical PHD domain and a four-helix bundle. Above: R556D and I575D indicate polymerization-blocking point mutations in the VIN3 VEL domain; below: amino acids deleted in stable transgenic lines VIN3-GFP ΔVEL and VRN5-SYFP2 ΔVEL, respectively. (B) qRT-PCR assays of *FLC* transcript levels during a vernalization timecourse. RNA was extracted before vernalization (before cold), at the end of a 6-week cold exposure (6WT0), or 10 days post-cold (6WT10). Data presented are relative to the geometric mean of *UBC* and *PP2A*, with different genotypes normalized to *FLC* levels before the cold. Error bars represent standard deviations (*n* = 3 biological replicates). (C) H3K27me3 ChIP across the *FLC* locus (position 0 corresponds to the transcriptional start site [TSS]) before cold and after a 6-week cold treatment. Data are expressed relative to input. Error bars represent SEM (*n* = 3 biological replicates). (D) Heatmap from IP-MS samples showing nuclear proteins co-precipitated with the corresponding VIN3 and VRN5 bait proteins in transgenic seedlings vernalized for 6 weeks. log2 fold-change (FC) was calculated in comparison to non-transgenic Col-FRI (adj. p-value ≤ 0.05 for wild-type proteins, *n* = 3 biological replicates). (E) Immunoblots of α-GFP immunoprecipitates from extracts of vernalized plants bearing the indicated VIN3-GFP transgenes, probed with α-FIE (PRC2 core subunit) antibody. Non-transgenic Col-FRI was used as a negative control (NC). Blots shown are a representative of three replicates.

Figure 2: VEL-dependent VIN3 and VRN5 assemblies in vivo and after transient heterologous expression

(A, B) Distributions of stoichiometry (molecule number) of tracked assemblies of (A) VIN3-GFP WT and VIN3-GFP RI>DD and (B) VRN5-SYFP2 WT and VRN5-SYFP2 ΔVEL in root nuclei of seedlings vernalized for 6 weeks, determined with SlimField microscopy. Individual replicates (*n* = 3-8 experiments, with 9-18 nuclei each) are shown in grey, the coloured line indicates the mean distribution. Insets show the frequency of tracks per cell. (C, D) Distributions of intervals between nearest neighbour stoichiometry peaks (periodicity) of tracked assemblies in (A, B). (E, F) Representative confocal images of epidermal leaf cell nuclei in *N. benthamiana,* (E) transiently expressing GFP-VIN3 or GFP-VRN5, and (F) transiently co-expressing GFP-VIN3 (green) and mScarlet-VRN5 (magenta) wild-type or mutant as indicated; scale bars 5 μm. (G, H) Quantification of number per nucleus (G) and volume of foci (H) observed in (F). Error bars represent standard deviation (*n* = 20-25 nuclei for (G) and *n* = 266 (VIN3 + VRN5) or 83 (VIN3 + VRN5 ΔVEL) foci for (H)).

Figure 3: VRN5 VEL cannot replace VIN3 VEL function in planta

(A) Schematic showing the VIN3-GFP VRN5 VEL chimera construct transformed into the *vin3* mutant. (B) qRT-PCR assays of *FLC* transcript levels during a vernalization timecourse. RNA was extracted before vernalization (before cold), at the end of a 6-week cold exposure (6WT0), or 10 days post-cold (6WT10). Data presented are relative to the geometric mean of *UBC* and *PP2A*, with different genotypes normalized to *FLC* levels before the cold. Error bars represent standard deviations (n = 3 biological replicates). (C) Immunoblots of α-GFP immunoprecipitates from extracts of vernalized plants bearing the indicated VIN3-GFP transgenes, probed with α-FIE (PRC2 core subunit) antibody. Non-transgenic Col-FRI was used as a negative control (NC). Blots shown are a representative of three replicates. (D) Superpositions of VEL domains of Arabidopsis VRN5$_{515-592}$ (teal), VIN3$_{529-601}$ (orange) and VEL1$_{618-690}$ (yellow) as predicted by Alphafold. VIN3 and VEL1 AF predictions superpose very close onto experimentally determined structures (*25*). (E) Sequence alignment of Arabidopsis VEL domains from *VIN3, VRN5 and VEL1*. α-helices are indicated underneath; above, VIN3 mutations that block self-association. Coloured residues are conserved in all angiosperm VELs (yellow), in VRN5 paralog VELs (grey) or in VIN3/VEL1 paralog VELs (pink).

Figure 4: VIN3 chromatin association is promoted by VEL-mediated polymerization and chromatin binding of the PHD superdomain

(A-C, F) ChIP-qPCR showing enrichment of VIN3-GFP (wildtype or mutant as indicated) across the *FLC* locus in seedlings vernalized for 6 weeks. Non-transgenic Col-*FRI* plants were used as a negative control sample (NC). Data are shown relative to an input control, error bars represent SEM (*n* = 2-4). (D) qRT-PCR assays of *FLC* transcript levels during a vernalization timecourse. RNA was extracted before vernalization (before cold), at the end of a 6-week cold exposure (6WT0), or 10 days post-cold (6WT10). Data presented are relative to the geometric mean of *UBC* and *PP2A*, with different genotypes normalized to *FLC* levels before the cold. Error bars represent standard deviations (*n* = 3 biological replicates). (E) Immunoblots of α-GFP immunoprecipitates from extracts of vernalized plants bearing the indicated VIN3-GFP transgenes, probed with α-FIE (PRC2 core subunit) antibody. Non-transgenic Col-FRI was used as a negative control (NC). Blots shown are a representative of three replicates.