Spatio-temporal regulation of Arabidopsis *Flowering Locus C*

Svenja Marie Reeck

September 2023

John Innes Centre, Norwich, United Kingdom

A thesis submitted to the University of East Anglia

for the degree of Doctor of Philosophy

Acknowledgements

Firstly, I want to thank my examiners Eirini Kaiserli and Anthony Dodd for the stimulating discussion that could have lasted forever. It was the perfect finish of my exciting PhD experience.

I want to thank my two brilliant supervisors Caroline Dean and Martin Howard for giving me the opportunity and support to explore the scientific world and learn how to become an independent researcher. Your passion for science will never fail to amaze and inspire me.

Second of all, I want to thank my mentor Rea L. Antoniou-Kourounioti: You are the hero of my PhD. Your guidance helped me to be confident in my results and your curiosity kept me motivated. Our meetings made the Monday morning a brilliant start of every week.

I also want to thank every single member of the Howard and Dean groups. The team spirit in each group meeting is unique to anything I have seen. It has been a blessing working, discussing, and brainstorming with you. From the Dean side of the family: Anna, Rob, Yusheng, Pawel, Philip, Congyao, Xiaofeng, Pan, Mathias, Tricia, Elsa, Marc, Colette, Silvia, Eduardo, Minglei, Geng-Jen, Miguel, Tina, Chichi, Caroline. And from the Howard side: Amy, Cecilia, Govind, Ander, John and Chris. Thank you all so much!

I want to sincerely thank Stefanie and Anis, who not only have been our collaborators but also my scientific family and support team since the start.

The John Innes Centre is a wonderful place that brings together people who are fascinated about plant research from all over the world. I am very grateful I could be part of this stimulating and exciting environment. I met many people without their help, this work would not be possible. Many thanks to Susan Duncan for the helpful discussions on imaging and feedback on Chapter 4. Thank you to the whole bioimaging team, and especially to Eva and Sergio or their help, and to Tim from the horticultural service for taking great care of my plants. I also want to phrase my appreciation to everyone from the media kitchen for making our work as carefree as possible.

My PhD cohort, the PEP-Net team of Gemma, Karen, Paniz, Jenny, Nancy, Alex, Jan, Jana, Ana, Andrea, Oz, Meredith and Brigitte. It was a pleasure to be on this journey with you, can't wait to see more of the amazing science you will be creating. I am especially grateful for the support of Jane Mellor, and all the other PEP-Net PIs especially Leonie, Marc, Davide, Kim and Nacho.

Last but not least the biggest love and appreciation to my family and my friends that became family. Norwich family: Franziska, Anna, Marina, Aura, Anna, Emma, Josh, Magda, Scott and Oli. Berlin family: Sophie, Amelia, Anton, Felix, Doris and Michael. Thank you for the support and for always believing in me. I am so lucky to have you.

Good Luck, with love,
Svenja
December 2023, Norwich.

This work was funded by the European Union's HORIZON 2020 research and innovation programme Marie Sklodowska-Curie Actions: PEP-NET 813282.

Abstract

The regulation of gene expression of developmental genes is pivotal for an organism's lifecycle. At the single-cell level, gene expression can occur through digital or analogue mechanisms. In the digital mode, genes exhibit binary ON/OFF states, different fractions of which generate the cell population average, while the analogue mode features graded gene expression with varying transcriptional output at each gene copy. Both modes are known to exist but how they are integrated to regulate one locus is unknown.

The floral repressor *FLOWERING LOCUS C (FLC)* in *Arabidopsis* exhibits digital Polycomb silencing in response to prolonged cold, creating quantitative epigenetic memory. However, how quantitatively different *FLC* levels are achieved independent of cold exposure remains elusive. To investigate this, an allelic series of mutations in *FCA*, a key regulatory player in the cold-independent autonomous pathway of *FLC* regulation, was used to perform quantitative fluorescent imaging. A dual regulation mechanism for *FLC* is observed, with analogue preceding digital regulation. Intermediate analogue *FLC* levels facilitated gradual digital silencing at the population level, as more cells switched OFF over time. Live imaging during embryo development showed that cells are initially ON, followed by an analogue and digital response post-germination. The digital response depends on the previously set-up analogue transcription level and is mediated by the Polycomb Repressive Complex 2.

Fluctuating winter-like temperatures, including a below-freezing period, induced the upregulation of *FLC* antisense transcripts, collectively named *COOLAIR*. Using Fluorescence in situ hybridization, I show that the upregulation can be accounted for by an increased number of cells expressing *COOLAIR* and more *COOLAIR* transcripts accumulating in a defined focus at active transcription sites. In a spatial context, specific meristematic cells were identified as enriched with *COOLAIR*.

These findings show the interplay between analogue transcriptional and digital epigenetic pathways to set up transcriptional levels. The importance of elucidating the mechanism at play to achieve the correct timing of developmentally important genes is highlighted, which could include the role of antisense transcription in fine-tuning the response in specific cells when exposed to freezing temperatures.

Table of contents

List of Abbreviations

°C – degrees Celsius

% – Percent

3D – Three-dimensional

ATS – active transcription site

BLAST – Basic Local Alignment Search Tool

bp – base pair

BSA – Bovine serum albumin

CAS – cryptic antisense transcripts

CaVIAR – *COOLAIR:* visualization of Antisense RNA in the embryo

CC – Constant Cold (5°C)

CDK – Cyclin dependent kinase

ChIP – Chromatin immunoprecipitation

CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats

CZ – Central Zone

Da – Dalton

DAP – Days after Pollination

DAPI – 4'-6-diamidino-2-phenylindole

DMSO – Dimethylsulfoxide

DNA – Deoxyribonucleic acid

dNTP – Deoxyribonucleotide triphosphate

DZ – Differentiation zone

EDTA – Ethylenediaminetetraacetic acid

EZ – Elongation zone

Fig –Figure

FISH – fluorescence in situ hybridization

FM – Fluctuating mild

FS – Fluctuating strong

H3K27me3 – trimethylation on the Lysine (position 27) of the histone tail of H3

H3K4me1/H3K4me3 – mono-/trimethylation of the Lysine (position 4) at histone tail of H3

H3K36me3 – Trimethylation of the Lysine (position 36) of the histone tail of H3

hr(s) – Hour(s)

IF – Immunofluorescence

kDa – kiloDalton

Ler – Landsberg erecta

lncRNA – long noncoding RNA

M – Molar

MES – 2-N-morpholino-ethane-sulfonic acid

ml – millilitre

mm – millimetre

mM – millimolar
μM – micromolar
mRNA – messenger RNA
MS – Murashige Skoog
MW – Molecular weight
nm – Nanometre
nt – nucleotides
NV – Non-vernalized
PDF – Probability Density Function
PBS – Phosphate-buffered saline
PcG – Polycomb group
PER – Primer Exchange Reaction
PFA – Paraformaldehyde
PHD – Plant homeo domain
PI – Propidium iodide
PRC2 – Polycomb Repressive Complex 2
PZ – Peripheral zone
qRT-PCR – quantitative reverse transcriptase polymerase chain reaction
QC – Quiecent centre
RAM – Root apical meristem
RdDM – RNA-directed DNA Methylation
RM – Rib meristem
RNA – Ribonucleic acid
RNA Pol II – RNA Polymerase 2
ROI – Region of interest
SABER – single-stranded DNA amplification by exchange reaction
SAM – Shoot apical meristem
SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAIR – The Arabidopsis Information Resource
TE – Transposable Elements
TSS – Transcription start site
UTR – Untranslated region
V – Vernalized
w/v – Weight/Volume
v/v –Volume/Volume
WB – Western blot
WT – Wild type

Name of genes/long non-coding RNAs

ABI3 – ABSCISIC ACID-INSENSITIVE 3
ABI5 – ABSCISIC ACID-INSENSITIVE 5
ATXR7 – ARABIDOPSIS TRITHORAX-RELATED 7
CBF – C-repeat containing Binding Factors
CLF – CURLY LEAF
CstF – cleavage stimulation factors
COOLAIR –*FLC* antisense lncRNA
ELF6 – EARLY FLOWERING 6
FCA – FLOWERING CONTROL LOCUS
FES1 – FRIGIDA-ESSENTIAL 1
FLC –FLOWERING LOCUS C
FLK – FLOWERING LOCUS K HOMOLOGY DOMAIN
FLX – FLC EXPRESSOR
FRI – FRIGIDA
FRL1 – FRIGIDA-LIKE 1
FT – FLOWERING LOCUS T
JMJ13 – JUMONJI DOMAIN-CONTANING PROTEIN
LD – LUMINIDEPENDENS
LHP1 – LIKE HETEROCHROMATIN PROTEIN 1
KPL – KOKOPELLI
MEA – MADEA
NTL8-D3 – NAC with TM motif1-like 8
PRP8 – PRE-RNA PROCESSING 8
REF6– RELATIVE OF EARLY FLOWERING 6
SDG26 – SET Domain Group 26
SOC1 – SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1
SUF4 – SUPPRESSOR OF FRIGIDA 4
SPCH – SPEECHLESS
SWN – SWINGER
UBC – UBC UBIQUITIN-CONJUGATING ENZYME 21
VIN3 – VERNALIZATION INSENSITIVE 3
VRN2 – VERNALIZATION 2 (VRN2)
VRN5 – VERNALIZATION 5
WUS – WUSCHEL

List of Figures

List of Tables

CHAPTER 1 : INTRODUCTION

1.1 Transcription regulation

Most of the genetic information needed for all biological processes is stored in the cell nucleus in the sequence of deoxyribonucleic acid (DNA), which is inherited through cell division. Functional units of the DNA, genes, are transcribed to mRNA and further translated and modified to become functional proteins that guide cell homeostasis. In complex organisms consisting of many cells with specialised functions, the set of genes expressed varies between cells (e.g., some genes are actively transcribed in one cell and repressed in another), yet the information in the DNA is identical (Alberts B, Johnson A, Lewis J, 2010).

The process of transcription needs to be coordinated tightly to control the right timing of activation for genes involved in the cell cycle, cell specification and developmental transitions (Fenderson, 2008). At the same time, transcription needs to be flexible to respond to external cues from the environment. Furthermore, the cell has the capacity to 'remember' expression states of genes and can inherit those states through replication. This 'memory' is not fixed but needs to be maintained by feedback cycles, which must also allow these states to be switched when appropriate. Overall, one of the most fundamental questions in molecular biology is how quantitative gene expression differences between cells is achieved.

Generically, two types of cell memory systems exist: trans and cis memory (Fig. 1-1). Trans memory can be maintained by circuits of diffusible factors that reinforce their regulatory state through a feedback network. The diffusible factor regulates transcription in a concentration-dependent manner (reviewed in (Bonasio et al., 2010)). One well-understood example of trans memory is the Sxl-lethal feedback loop that determines sex identity in Drosophila (Bell et al., 1991). During cell division, the diffusible factor will be divided roughly equally between daughter cells feeding back into reinforcing the transcriptional state (see Fig. 1-1, trans regulation). Traditionally, such regulation is ascribed to sequence-specific transcription factors A that bind to regulatory DNA elements (e.g. the promoter) of a gene A. According to the concentration of the transcription factor A, gene expression can then be quantitatively up or downregulated.

Additionally, how the local chromatin is organized can affect transcription and this can encode cis memory (Summarised in Fig1-1, cis regulation). Several known molecular processes modify the accessibility of a promoter/gene by altering the structure or dynamics of chromatin (reviewed in (Dillon, 2006)). Epigenetic modifications, the addition of chemical groups to DNA/histones or different histone variants, can structurally modify chromatin interactions (Moazed, 2011). The degree of compaction brought about by nucleosomes, the smallest unit of the chromatin, and other associated proteins (e.g. linker

histones, chromatin remodellers, cohesin) can restrict access to the genomic DNA on different levels and during various nuclear processes (Bannister and Kouzarides, 2011).

By modulating the local chromatin environment of a gene, quantitative transcriptional regulation can also arise. For example, by varying the type and level of histone modifications, DNA accessibility can be radically altered (Ahmad et al., 2022). In one scenario, nucleosome positioning affects the ability of transcription factors to bind. Another possibility is that alteration of the chromatin environment directly affects the kinetics of transcription (Coulon et al., 2014) by altering how fast the RNA polymerase binds and elongates.

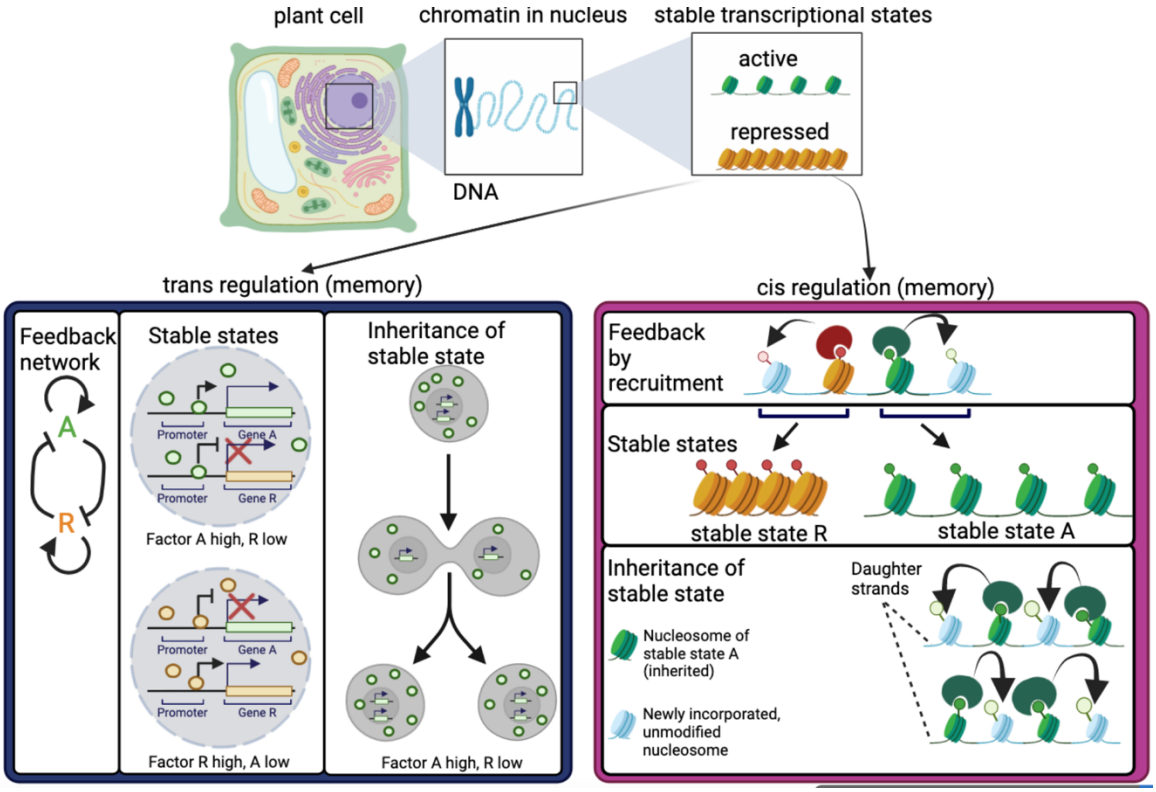


Figure 1–1 Maintenance of gene expression states.

Schematic presentation of the mechanism of epigenetic memory by stable transcriptional states in the cell nucleus. **Trans regulation** involves trans factors that determine the transcriptional states based on their availability (concentration) resulting in a feedback network reinforcing that state. Here, two stable states are depicted depending on the concentration of a diffusible factor A (green) or R (orange) activating or repressing the respective other factor. If factor A is high, factor R is low which can be inherited by an equal distribution of factor A during cell division. **Cis regulation** involves modification on the DNA as heritable information carriers, here shown as histone modifications. Modifications recruit more modifications of their type (e.g., state A being active/open chromatin, green). During S-phase, modified histones are distributed between the two daughter strands. The newly incorporated histones get modified by a read-and-write feedback recruitment (newly recruited histone modifications symbolized in light green). Figure created with BioRender.com

Another example of cis memory is by the conserved histone methyltransferase complex, Polycomb Repressive Complex 2 (PRC2), targeting many key developmental genes in multicellular organisms to epigenetically repress transcription by positioning H3K27me3 at its target sites. In mammals, a process required for proper female development is the inactivation of the second copy of the X chromosome. Here, the long noncoding RNA (lncRNA) *Xist* is associated with recruiting PRC2 to the *Xist*-expressing X chromosome, thus positioning the repressive histone mark H3K27me3 along the chromosome. X chromosome inactivation is maintained by a read-and-write-feedback mechanism of PRC2 throughout the rest of development indicating that the modified histones are somatically inherited through cell division. Parental histones carrying epigenetic information are recycled during DNA replication, with the replisome coupling new histone deposition with fork progression (Groth et al., 2007). The distribution of the parental modified histones to the two daughter strands is mediated by two different mechanisms for each strand and is thought to be at random, causing an approximate two-fold dilution of parental marks (Foltman et al., 2013). Read-write feedback by cooperative recruitment of PRC2 by neighbouring H3K27me3-marked histones, and then allosteric activation to deposit more H3K27me3 nearby, is thought to re-establish the proper chromatin state.

lncRNAs can impact gene expression in several ways. They can act as a scaffold by connecting DNA and proteins or by binding to DNA and recruiting other factors to the site (Statello et al.). Important to mention is that while some lncRNA function in trans, being able to diffuse in the cell, others affect only the site of synthesis before RNA polymerase release (in cis).

Another key mechanism of transcriptional regulation is modifications directly on the DNA. In the case of DNA methylation (addition of a methyl group, CH_3, on the cytosine (C) nucleotide it can be inherited to the progeny cell (Jones and Taylor, 1980). Importantly, the methylation has no effect on base-pairing, however, it can influence the chromatin compaction and the accessibility of proteins involved in gene expression (Buitrago et al., 2021). Methylation occurs most commonly at the dinucleotide sequence CG, where CG is base-paired to a GC, which is also methylated. Inheritance through semi-conservative replication is achieved by the methylated parental strand acting as a template for the unmethylated newly synthesized daughter strand. The consequent recruitment of a maintenance methyltransferases to the hemi-methylated site adds a methyl group to the unmethylated cytosine to maintain the specific CG patterns (Probst et al., 2009). It has been shown in Arabidopsis, among other species, that the DNA trinucleotide sequences CHG and CHH can be methylated, in which H could be the nucleotide A, C, or T. There are several known pathways involved in the maintenance of asymmetric methylation sites (CHH methylation) in plants e.g. the RNA-directed DNA Methylation (RdDM) pathway (Stroud et al., 2014; Zhang et al., 2006).

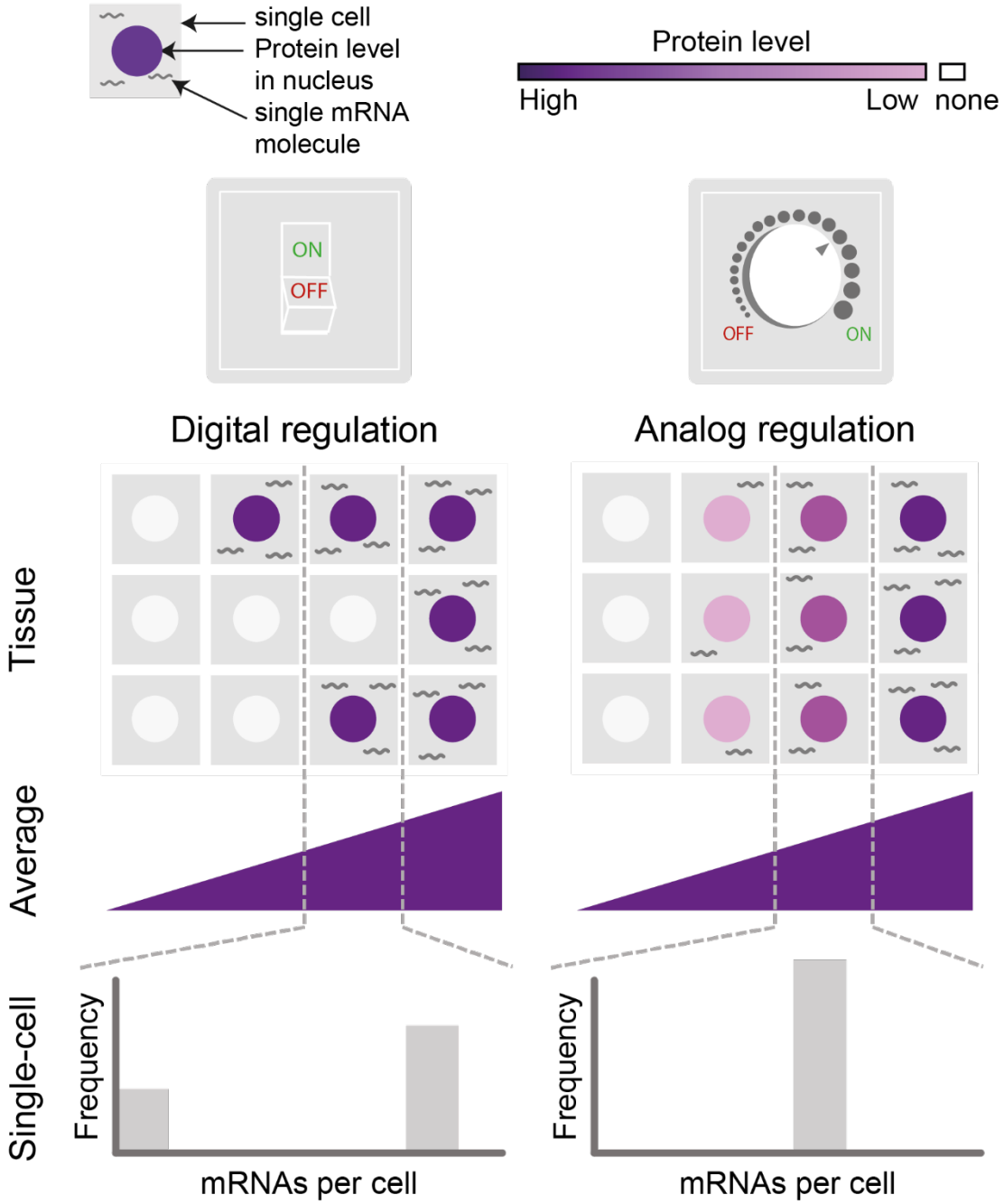


Figure 1–2 Digital and analogue modes of gene regulation.

Schematic illustration of digital and analogue gene regulation. Digital regulation (left) corresponds to loci being in an 'ON' state (purple) or 'OFF' state (white), where we assume for simplicity that there is only one gene copy per cell. At the tissue level, moving from low to high average expression (columns left to right) is achieved by a change in the fraction of cells in each of the two states. This mode is distinct from analogue regulation (right), where each cell has a graded expression level that roughly corresponds to the overall population average.

How quantitative gene regulation is stably maintained on a single-cell level has been of great interest. Whole tissue intermediate transcriptional levels could be explained by two single-cell scenarios (or a combination of both). Either digital, with a fraction of loci being ON and a fraction of loci being OFF similar to a light switch, or alternatively analogue (Munsky and Neuert, 2015), with the expression level at all gene copies being uniformly tuneable resulting in a graded range of transcriptional intensities, similar to a dimmer switch as highlighted Figure 1-2.

1.2 *FLC* as a model system

The timing of flowering is important for plant reproduction. Plants must be able to sense when environmental and internal cues are favourable before they can switch to reproductive development. To do this, plants have evolved a complex regulatory network that consists of multiple pathways that regulate a set of genes called floral integrators. The activity of these genes triggers the transition of the meristem to reproductive development (Taiz and Zeiger, 2014).

The timing of flowering is predominantly influenced by two external cues: prolonged low temperature (vernalization), and photoperiod integrating cues on day lengths and light composition (Taiz and Zeiger, 2014). These external cues work in conjunction with internal factors such as plant age, phytohormones and carbohydrate levels, as well as the autonomous pathway, consisting of RNA-binding and processing factors. Collectively, those factors determine the reproductive strategy.

The vernalization pathway and autonomous pathway converge on the regulation of *FLOWERING LOCUS C* (*FLC*). *FLC* encodes a MADS-box DNA-binding protein functioning as a flowering repressor by inhibiting the expression of floral integrators, such as *FLOWERING LOCUS T* (*FT*), FD, and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (SOC1) (Michaels and Amasino, 1999). FLC functions by repressing the expression of the mobile flowering signal, FT. If the repression is lifted because the repressor, FLC, is not expressed anymore, FT travels to the shoot to activate downstream flowering events (Corbesier et al., 2007; Li et al., 2011). FLC is crucially involved in determining the timing of the transition from vegetative to reproductive growth to align flowering with spring or summer and is thereby essential for reproductive success critical to determine lifestyle (Figure 1-3A).

FLC has been shown to bind to many targets beyond the floral integrators , suggesting its role in regulating other developmental transitions (Deng et al., 2011). It has an effect on temperature-dependent germination by impacting seed dormancy (Blair et al., 2017; Chiang et al., 2009). Changes in shoot branching patterns of late flowering mutants indicate the effect of FLC on stem branching (Huang et al., 2013) most likely via weak binding of FLC to BRANCHED 1 (BRC 1) (Aguilar-Martínez et al., 2007), a key transcription factor integrating several signals to determine if axillary buds stay dormant or grow to branches inside the axillary meristem. Additionally, FLC levels have also been reported to influence the circadian rhythm (Edwards et al., 2006), via the clock integrator *LATE ELONGATED HYPOCOTYL* (*LHY*) (Spensley et al., 2009).

Quantitative variation (different levels) of *FLC* expression influences if there is a vernalization requirement (Col*FRI*) or if the plant has a rapid cycling lifestyle (like the accessions Col-0 or Ler). *FLC*'s constitutive levels are set by two opposing pathways that are competing in a "tug-of-war" at early development. Depending on those levels, the reproductive strategy is determined. In rapid-cycling, summer-annual accessions, *FLC*

expression is repressed by the autonomous pathway resulting in early flowering independent from cold. The **autonomous pathway** consists of RNA binding and processing factors influencing the chromatin environment that represses *FLC* transcription (Wu et al., 2019b). It has been shown to affect the transcriptional state of *FLC* by alternative processing of *FLC* and its antisense long non-coding RNAs, collectively called *COOLAIR* (Schon et al., 2021; Swiezewski et al., 2009). These actions are antagonised by the **FRIGIDA (FRI) pathway** which promotes *FLC* transcription thus creating an active chromatin environment (Koornneef et al., 1998).

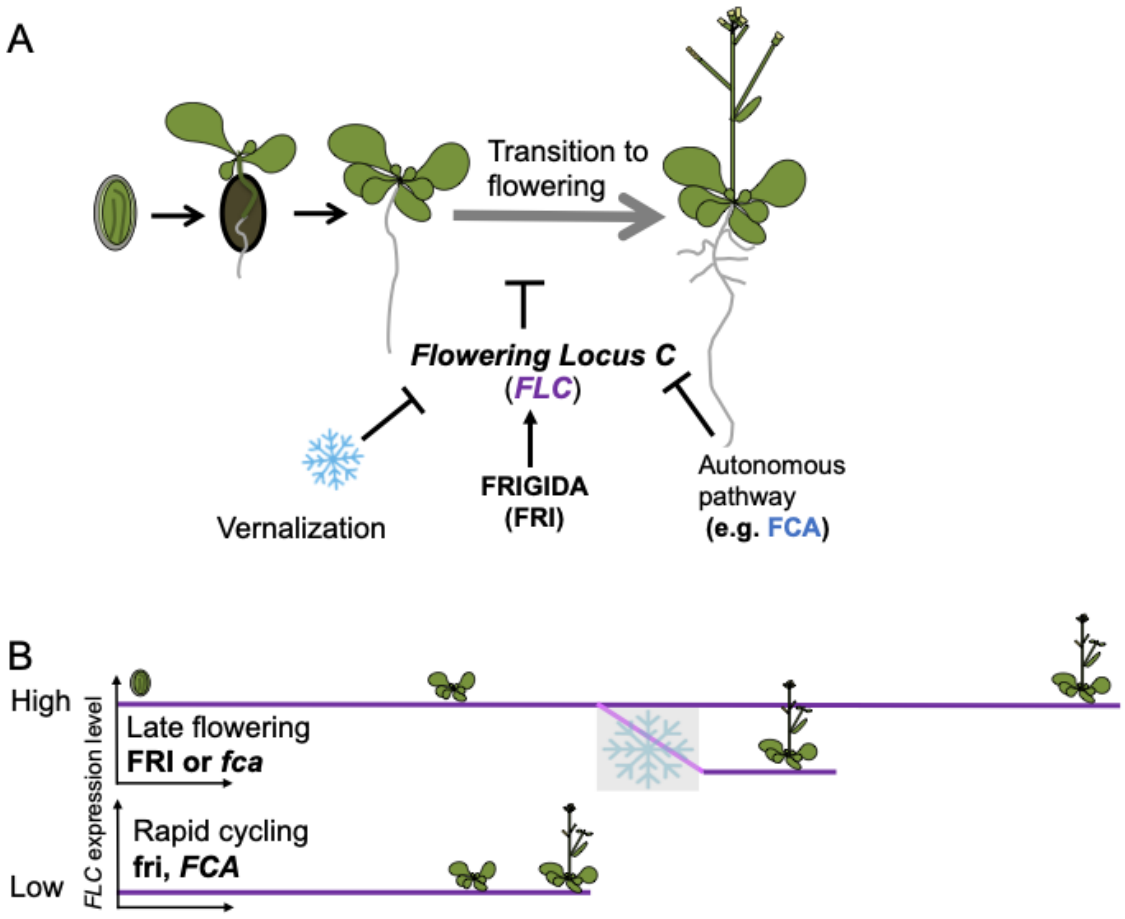


Figure 1–3 *FLC* regulation.

(A) Schematic representation of the regulation acting on *FLC* levels. The environmental response to cold is called vernalization and leads to *FLC* levels being switched OFF by the Polycomb Repressive Complex 2 (PRC2). **(B)** The *FLC* expression level determines reproductive strategy. High *FLC* expression levels lead to a winter-annual, late flowering phenotype as observed in Col*FRI* accessions or *fca* mutants. Low *FLC* expression levels contribute to a summer-annual, rapid cycling lifestyle.

The coiled-coil protein FRI acts as a scaffolding protein, interacting with the components FRIGIDA-LIKE 1 (FRL1), FRIGIDA-ESSENTIAL 1 (FES1), SUPPRESSOR OF FRIGIDA 4 (SUF4) and FLC EXPRESSOR (FLX) as an transcriptional activator of *FLC* (Choi et al., 2011). What is known about the molecular mechanism of how FRI functions in enhancing *FLC* expression is that the FRI complex acts as a transcription factor recruiting transcription

machinery and chromatin modification factors to *FLC* (Choi et al., 2011). Direct interaction with the nuclear Cap-Binding Complex leads to coupled effects on transcription and splicing elucidating a co-transcriptional mechanism (Geraldo et al., 2009). The FRI complex also interacts with chromatin modifiers such as the H3K36 methyltransferase SET DOMAIN GROUP8 (SDG8) setting an transcription promoting open chromatin environment (Hyun et al., 2017).

In the environmentally induced, temperature dependent **Vernalization pathway**: *FLC* is progressively silenced by prolonged exposure to cold temperatures (vernalization), aligning flowering with spring. This is achieved through individual *FLC* gene copies making a cis-mediated, digital switch from an "ON" (expressing) state to an "OFF" (silenced) state (Angel et al., 2011). This switch is asynchronous between loci, even in the same cell, with the number of switched OFF gene copies increasing over time in the cold (Berry et al., 2015). The overall change in the proportion of cells being transcriptionally active is an output of how long the plant has experienced winter-like cold conditions (Yang et al., 2017).

There are several loss-of-function mutations in *FRI* in the natural accessions of Arabidopsis resulting in an early flowering phenotype (Zhang and Jiménez-Gómez, 2020). This suggests that summer-annual varieties evolved from winter-annual ancestors through multiple independent losses of FRI function. The commonly used laboratory accession Columbia-0 (Col-0) carry a non-functional allele of FRI and Landsberg erecta (Ler) carries a functionally impaired allele. Ler carries a deletion of 376 bp combined with an insertion of 31 bp in the promoter region, resulting in the loss of the translational start site (Le Corre et al., 2002; Johanson et al., 2000). An alternative start codon produces a lower expressed mRNA with reduced activity (Schmalenbach et al., 2014). The Columbia allele contains a 16 bp deletion that introduces a premature stop codon. This truncation results in a non-functional protein lacking a portion of its C-terminus. (Johanson et al., 2000).

Extensive focus on the *FLC* locus over the last decades has unravelled components associated with the high and low expression states that quantitatively regulate *FLC*. A detailed description of the main components and mechanisms identified so far follows in the next sections.

1.2.1 Autonomous pathway vs. FRIGIDA

When flowering occurs independent of external, environmental cues such as photoperiod and vernalization and instead in response of internal developmental factors, it is autonomously regulated (Taiz and Zeiger, 2014). In Arabidopsis, although flowering is induced by environmental cues, it can also occur in its absence. The internal mechanisms of the autonomous pathway orchestrate the establishment of a repressive chromatin environment to repress *FLC* expression through RNA processing. These regulatory actions are counteracted by the FRIGIDA (FRI) pathway, which operates by actively promoting

FLC transcription, thereby contributing to creating an active chromatin environment (Koornneef et al., 1998) (see Fig. 1-4).

How exactly the different chromatin environments are established involves many players and complex interactions discussed in the following section with the exact order of events still remaining elusive. The *FLC* gene undergoes transcription in both the sense (*FLC*) and antisense (*COOLAIR*) directions, and these two transcriptional events are mutually exclusive (Rosa et al., 2016).

One key protein of the autonomous pathway is FLOWERING CONTROL LOCUS *(*FCA*)*, an RNA binding protein with a WW protein interaction domain as well as two prion-like domains (Macknight et al., 1997). FCA, the RNA-binding proteins FPA and FLOWERING LOCUS K HOMOLOGY DOMAIN (FLK) (Lim et al., 2004; Schomburg et al., 2001), as well as the 3' processing factors FY (homolog of Polyadenylation factor 1 subunit 2 (Pfs2p)) (Simpson et al., 2003) and the cleavage stimulation factors (CstF) CstF77/64 (Liu et al., 2010), promote proximal polyadenylation of *COOLAIR*. Also, the positive splicing factor PRE-RNA PROCESSING8 (PRP8), a core spliceosome component essential for the usage of the proximal *COOLAIR* splicing site, represses *FLC* transcription (Marquardt et al., 2014). Ultimately, splicing and polyadenylation separate *COOLAIR* isoforms into two main classes: Class I, which is proximally polyadenylated, and class II, distally polyadenylated (here two isoforms, class II.i and class II.ii are shown in Fig. 1-4A). Proximal polyadenylation is associated with a functional autonomous pathway and low transcription at *FLC*, and distal polyadenylation is associated with increased expression levels (Fig. 1-4B). While it is unclear whether the utilization of the distal site initiates or

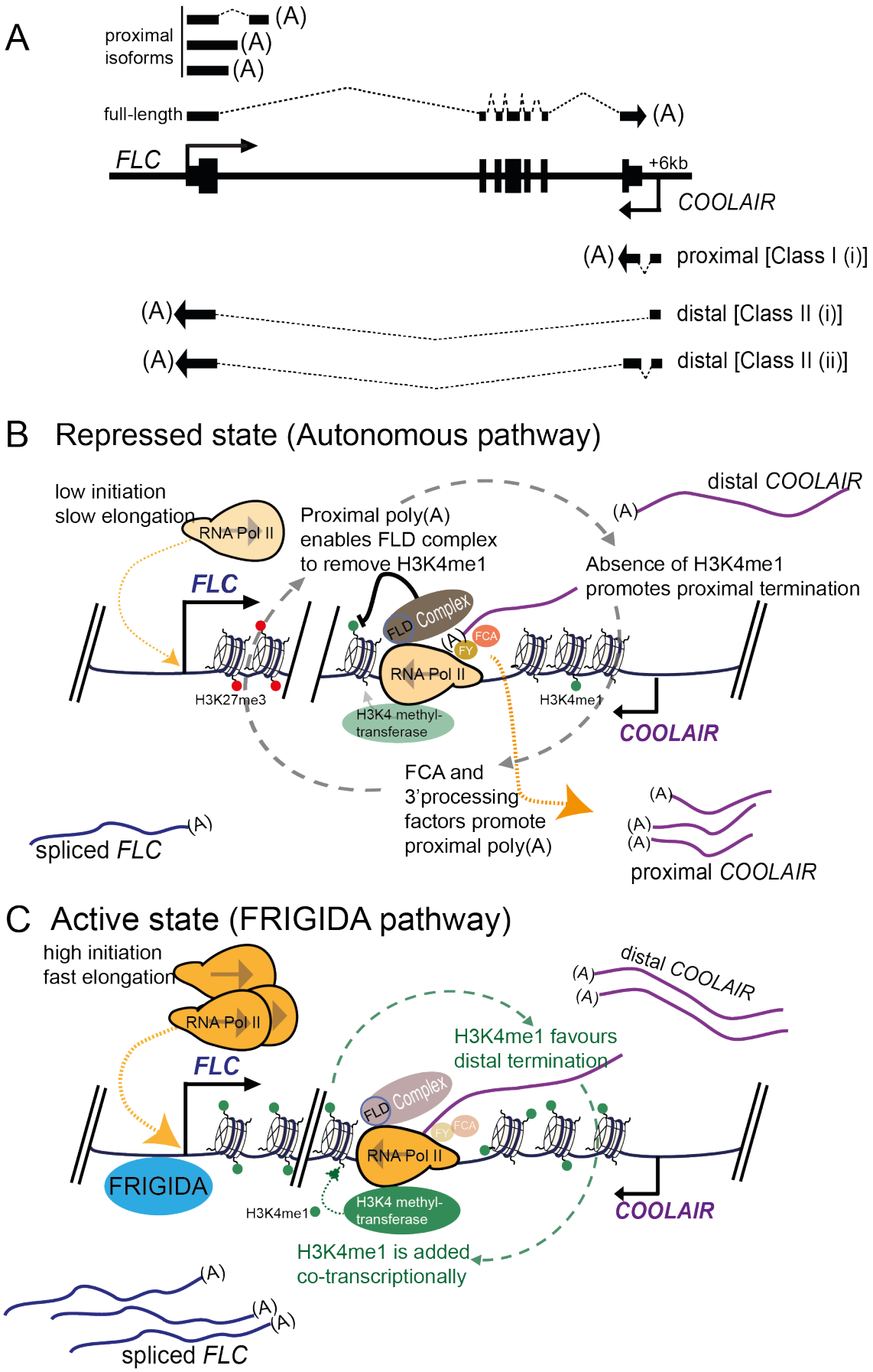

A
proximal isoforms
(A)
(A)
(A)
full-length
(A)
FLC
+6kb
COOLAIR
(A)
proximal [Class I (i)]
(A)
distal [Class II (i)]
(A)
distal [Class II (ii)]
B Repressed state (Autonomous pathway)
low initiation
slow elongation
RNA Pol II
FLC
H3K27me3
Proximal poly(A)
enables FLD complex
to remove H3K4me1
FLD Complex
FCA
FY
RNA Pol II
H3K4 methyl-
transferase
distal COOLAIR
(A)
Absence of H3K4me1
promotes proximal termination
H3K4me1
COOLAIR
FCA and
3'processing
factors promote
proximal poly(A)
(A)
(A)
(A)
proximal COOLAIR
(A)
spliced FLC
C Active state (FRIGIDA pathway)
high initiation
fast elongation
RNA Pol II
FLC
FRIGIDA
H3K4me1
distal COOLAIR
(A)
(A)
H3K4me1 favours
distal termination
FLD Complex
FY
FCA
RNA Pol II
H3K4 methyl-
transferase
COOLAIR
H3K4me1 is added
co-transcriptionally
(A)
(A)
(A)
spliced FLC

Figure 1–4 The autonomous pathway acts via transcription-coupled repression at *FLC*.

(previous page)

(A) Representation of the *FLC* genomic locus with the most abundant sense and antisense transcripts. Black boxes are exons, splicing patterns are represented by dashed lines. (A) is indicative of the polyadenylation site. Both sense and antisense transcripts are proximally polyadenylated, mediated by *FCA*. **(B)** Schematic model of interactions of selective autonomous pathway players at *FLC*. Activating H3K4me1 can be added during transcription. RNA processing-related factors (*FCA*, FY) promote proximal polyadenylation and termination. The repressive chromatin environment is further influenced by FLD demethylating H3K4me1 creating reduced transcriptional initiation and elongation of *FLC* and *COOLAIR*. The removal of H3K4me1 plays a crucial role as it can impede co-transcriptional processing by preventing the addition of another activating histone modification H3K36me3. Consequently, this impediment promotes additional proximal termination events. This intricate process creates a feedback loop, as indicated by the black dashed arrows, where the initial occurrence of proximal termination triggers subsequent proximal termination events. **(C)** Active transcription involves FRIDIGA and high levels of RNA Pol II initiation and fast elongation. Histone Methyltransferases add active histone marks such as H3K4me1, a histone mark associated with distal poly(A) of *COOLAIR*. Image modified from Menon et al. (submitted).

results from the enhanced transcription of *FLC*, the autonomous pathway players associated with proximal *COOLAIR* mediated by FLOWERING LOCUS D (FLD, homolog of the human H3K4me demethylase LSD1) are associated with removing H3K4 methylation at the *FLC* gene body (Liu et al., 2007; Wu et al., 2016). As H3K4 methylation is a hallmark of actively transcribed chromatin, its removal is correlated with transcriptional repression. FLD itself forms a complex with LUMINIDEPENDENS (LD, homeodomain-containing protein) and SET Domain Group 26 (SDG26, histone demethylase domain) and transiently interacts with FY (Fang et al., 2020). This process involves the 3′ processing Cleavage and polyadenylation specificity factor (CPSF) complex mediating the cleavage and polyadenylation necessary for subsequent alteration of chromatin modifications, which in turn results in *FLC* transcriptional silencing. High levels of H3K27me3 histone marks are present thus resulting in low transcription initiation, as well as slow elongation rates (He et al., 2003; Liu et al., 2010; Yang et al., 2014). In this way, transcription termination is linked with the delivery of chromatin modifications to influence subsequent transcription (Mateo-Bonmatí et al., 2023).

COOLAIR transcription has been associated with the formation of an R-loop (DNA-RNA hybrid) at the *FLC* 3′ end, approximately spanning from the *COOLAIR* promoter to its proximal polyadenylation site (Sun et al., 2013). The autonomous pathway player *FCA* is required for resolving the R-loops (Fang et al., 2020; Xu et al., 2021), which are associated with genome destabilization and DNA damage (Gan et al., 2011; Stirling et al., 2012). The resolution involves a transcription termination process that subsequently instigates chromatin modifications within the local genomic environment (Mateo-Bonmati, 2023). Additionally, the resolution is mediated by m^6A modifications at the *COOLAIR* polyadenylation site. The m^6A modifications may affect *COOLAIR* secondary structure leading to increased binding affinities to *FCA* and other RNA-binding proteins (Xu et al. 2021). This suggests that m^6A may be an important mark for effective R-loop removal by *FCA* being part of a feedback mechanism of the repressive chromatin state. Interestingly,

most of the RNA-binding proteins associated with *FLC* carry intrinsically disordered domains required for compartmentalization into biomolecular condensates indicating they phase-separate. FLL2, a protein carrying prion-like domains, has been shown to promote liquid-liquid phase separation (LLPS) with *FCA* (Fang et al., 2019).

NDX, a homeodomain-containing DNA binding protein, binds to the antisense strand of *FLC*, at the *COOLAIR* promoter site and stabilized a R-loop. Loss of R-loop stability in *ndx1* mutants increased *COOLAIR* transcription and released *FLC* repression (Sun et al., 2013), suggesting that R-loops suppress transcription.

Changes in the ratio of distal to proximal *COOLAIR* isoforms, as well as in the chromatin environment, correlate with respective fast and slow transcription rates of *FLC* as shown by RNA Pol II occupancy, a 25-fold decrease in initiation and 10-fold slower elongation in a repressed compared to an active state (Wu et al., 2016).

In summary, *FLC* regulation by the autonomous pathway involves an interplay of co-transcriptional processing, transcription termination, and RNA Pol II activity, resulting in the establishment of a chromatin landscape conducive to limited gene expression.

The active state (Fig. 1-4C) requires the FRI complex to implement a high *FLC* expression state. It is associated with high RNA Pol II occupancy, and high activating epigenetic marks such as acetylation, H3K4me3/H3K36me3 around the transcriptional start site. Accumulation of H3K4me1 in the *FLC* gene body (He et al., 2003; Yang et al., 2014) is added co-transcriptionally by H3K4 methyltransferases e.g. *ARABIDOPSIS* TRITHORAX-RELATED7 (ATXR7) (Oya et al., 2022; Tamada et al.). The active state can oppose FLDs function and favours distal termination of the *COOLAIR*. Those factors jointly contribute to high transcription initiation and fast elongation (Wu et al., 2016). The repressive and active transcriptional states may feedback into reinforcing themselves (Fig. 1-4 BC, dashed arrows).

1.2.2 The vernalization pathway

During the cold, the chromatin environment changes from active to repressive at *FLC*. By combining experimental data of chromatin immunoprecipitation (ChIP), live imaging and mathematical modelling of the *FLC* locus, a picture of a digital, bistable model for epigenetic memory arose depending on local repressive histone marks at the *FLC* mediated by the Polycomb Repressive Complex 2 (PRC2) (Angel et al., 2011). Interestingly, live imaging of two differently labelled reporter-protein-fusions of the two *FLC* copies showed that *FLC* shut-down is locus-specific and heritable (Berry et al., 2015; Yang et al., 2017) suggesting that memory of gene expression is stored in cis, in the local chromatin environment and in a digital "all-or-nothing"-manner for each locus (Fig.1-1- cis-regulation, Fig.1-2 -digital regulation). Over the period of vernalization, the number of loci being silenced increases quantitatively (Fig.1-6B). How the switch-off is initiated is still actively investigated and involves the PRC2 component VERNALIZATION 2 (VRN2), as well as its accessory

proteins, the Plant Homeodomain (PHD)- proteins VERNALIZATION INSENSITIVE 3 (VIN3) and VERNALIZATION 5 (VRN5) that have been shown to be able to form polymerisation-based assemblies potentially contributing to the switch off (Fiedler et al., 2022).

Early cold (0–4 weeks) boosts the expression of total *COOLAIR* and proximal *COOLAIR* (Rosa et al., 2016; Swiezewski et al., 2009). In line with the function of the autonomous pathway in the warm, the proximal termination of *COOLAIR* correlates with changes in the chromatin environment. This co-occurrence of proximal splicing and polyadenylation of *COOLAIR* with early cold might suggest its importance for *FLC* repression.
It has been shown that *COOLAIR* is not a requirement for *FLC* silencing as shown by plants lacking the *COOLAIR* transcript still sufficiently vernalizing in the cold (by either *COOLAIR* promoter-terminator exchange systems or Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) deletions of the *COOLAIR* promoter (Csorba et al., 2014; Luo et al., 2019). However, the presence of cryptic antisense transcripts (CAS) in place of the canonical *COOLAIR* isoforms suggests that antisense transcription happening (Zhao et al., 2021).

Plant growth is impacted in the cold, leading to reduced plant size and an upregulation of genes involved in cold acclimation, e.g. the C-repeat containing Binding Factors (CBF), shown to be involved in freezing protection by synthesis of membrane-stabilizing proteins (Lee et al., 2009). CBFs have also been shown to bind to the promoter region of *FLC*'s antisense transcripts, *COOLAIR* (Jeon et al., 2023). In addition, NTL8 encodes a NAC domain transcription factor that directly binds to the *COOLAIR* promoter (Zhao et al., 2020a). *ntl8-D* alleles cause a truncation in the protein deleting the C-terminal transmembrane domain which results in an elevated nuclear localization (Zhao et al., 2020a) promoting all endogenous *COOLAIR* transcripts.

Furthermore, natural temperature fluctuations with night temperatures below 0°C, as observed during the autumn-to-winter seasonal transition, have been shown to strongly upregulate *COOLAIR* while accelerating *FLC* transcriptional shutdown (Zhao et al., 2021).

In the cold, the most abundant distal poly(A) isoform has been shown to change conformation (Yang et al., 2022). In contrast, the most abundantly upregulated proximally polyadenylated *COOLAIR* showed no structural changes upon cold.

Our current knowledge about *FLC* shutdown can be divided into two distinct memory phases, separated in time, with both necessary to gain stable inheritance. First, upon transition to cold *FLC* is repressed followed by H3K27me3 deposition at a specific genomic location downstream of the *FLC* transcription start site, the *FLC* nucleation region. The first phase is considered metastable as only in the second phase upon return to warm, H3K27me3 spreads over the whole locus resulting in heterochromatinization and stable silencing (Yang et al., 2017). Once the gene is fully silenced, this state is maintained through

cell divisions and only reset during the gametogenesis of the next generation (Tao et al., 2017). The temporal uncoupling of these two phases showed the involvement of the H3K27 Methyltransferases CURLY LEAF (CLF) and SWINGER (SWN, SET-domain proteins), as well as the H3K27me3 binding Protein LHP1, in the spreading of H3K27me3 over the whole gene body that occurs upon return to the warm and which is needed for the long-term memory of the silenced state. Mathematical modelling of *FLC*-reactivation in single cells of mutants deficient in H3K27me3 spreading (LIKE HETEROCHROMATIN PROTEIN, *lhp1*) upon return to warm, revealed the potential existence of additional cis-memory factors needed to keep the metastable memory (Lövkvist et al., 2021).

Modelling of the digital memory system confirmed that it is able to integrate information about the duration of the previously experienced period of cold (measured in the fraction of quantitatively silenced loci), as well as withstand noise in the input signal caused, e.g., during cell division or from fluctuating temperatures (Angel et al., 2015). (Figure 1-2, digital regulation). A similar digital response is assumed for other Polycomb targets in Arabidopsis, the apical meristem maintenance regulator WUSCHEL (Lee et al., 2019) and the stomata identity regulator SPEECHLESS (Sun et al., 2019), summarized in (Menon et al., 2021).

1.2.3 *FLC* expression in the embryo

1.2.3.1 *FLC* resetting

After the transition to flowering, *FLC* levels need to 'reset' so the progeny generations still have a vernalization requirement. In annual Arabidopsis accessions, *FLC* expression is reset during embryogenesis and increases from the heart embryo stage (Hofmann et al., 2019; Roeder and Yanofsky, 2006)(Fig-1-5A). The reset requires the deletion of the repressive H3K27me3 epigenetic marks at some developmental stage. The sperm and egg cells have a reduced number of H3 variants that are rapidly turned over after fertilization in a DNA replication-independent manner (Ingouff et al., 2010) making it difficult for epigenetic information in the form of histone modifications to be transmitted from the parent to offspring.

FLC re-setting requires the transcription factor LEC1 to bind to *FLC* approximately 24 h after fertilization. This changes the chromatin environment to be accessible for LEC2 and FUS3 at the RY motifs, located in intron 1, allowing transcription (Tian et al., 2019).

It has been shown that *FLC* is a maternally imprinted gene, as the epigenetic information of the male germline copy is wiped whereas the maternal copy remains silenced (Choi et al., 2009; Sheldon et al., 2008). The two alleles are reset independently, with the maternal copy remaining H3K27me3-silenced by CLF-PRC2 for at least two days after pollination (2 DAP) (Luo et al., 2020). The same study found that at DAP6, CLF was absent from the *FLC* locus, correlating with the resetting dynamics. Note, however, that ChIP here was performed on whole seed material and therefore might not reflect the embryo state.

In addition, the activity of Jumonji domain H3K27me3 demethylases such as EARLY FLOWERING 6 (ELF6), RELATIVE OF EARLY FLOWERING 6 (REF6) and JUMONJI DOMAIN-CONTANING PROTEIN (JMJ13) are required for resetting of Polycomb targets (Crevillén et al., 2014; Lu et al., 2011; Zheng et al., 2019). For *FLC* in embryo, *elf6* mutants showed a heritable partially vernalized state highlighting its role in *FLC* resetting by preventing the propagation of epigenetic marks (Crevillén et al., 2014).

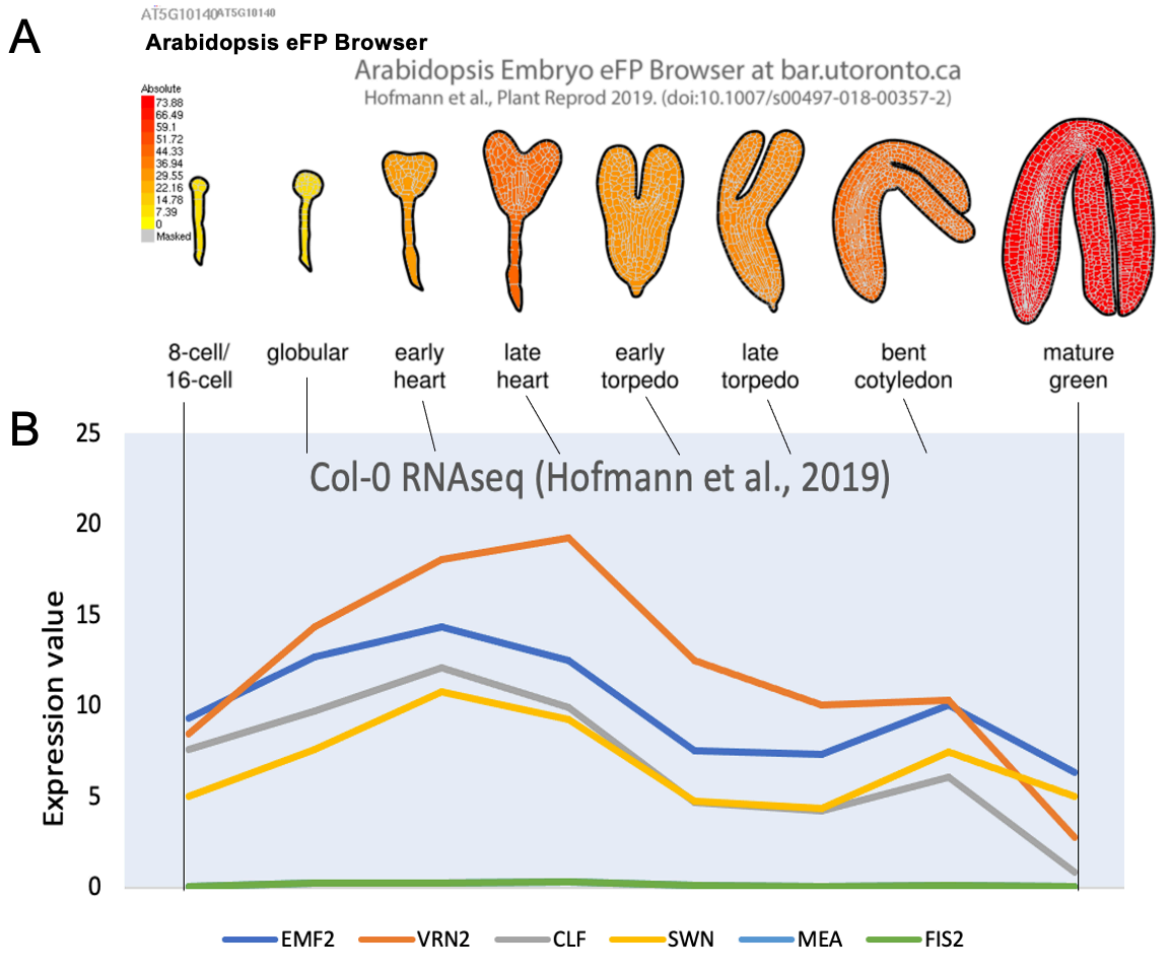


Figure 1–5 *FLC* and Polycomb component expression throughout embryo developmental stages

(A) *FLC* transcription in the different embryo developmental stages in Col-0. Graphics obtained from the Arabidopsis eFP Browser (bar.utoronto.ca). **(B)** RNA sequencing data from Hofmann et al., 2019 of PRC2 components. Time points are aligned with the developmental stages of embryos in (A).

Once reset, *FLC* mRNA expression levels in the embryo are determined by a delicate antagonism of the autonomous pathway and FRI. It has been shown that the main autonomous pathway component *FCA* plays a role in promoting the early termination of the *FLC* sense transcript, leading to the downregulation of *FLC* transcription. On the other hand, FRI promotes transcriptional read-through to a distal site (Schon et al., 2021). *FLC* levels are further elevated by embryo-specific factors, such as the B3 transcription factor

(TF) ABSCISIC ACID-INSENSITIVE 3 (ABI3) together with the basic leucine zipper TF ABSCISIC ACID-INSENSITIVE 5 (ABI5), which binds to an ABA-responsive cis-element in the *FLC* promoter region (Xu et al., 2022).

In summary, after fertilization *FLC* resetting occurs allele-specifically. The vernalized paternal copy is stripped prior to fertilization. The maternal vernalization state is imprinted and transmitted through the egg cell and maintained PRC2-silenced until 2DAP. At 6DAP *FLC* is reset (via removal of H3K27me3) and activated (proximal and distal transcription) according to trans factors. Not much more is known about PRC2 acting on *FLC* after the 6DAP stage.

1.2.3.2 PRC2 components during embryogenesis

During development, cells differentiate into specialized functional tissue. Since the genetic code in each of these cells is identical, cellular differentiation is hugely dependent on the correct setting up and maintenance of epigenetic states. PRC2 complexes regulate many genes controlling developmental transitions by adding and keeping the repressive histone mark H3K27me3. This can be mediated by multiple PRC2 complexes or "flavours" that consist of different homologs of subunits with cell-type specific or developmental roles.

Whereas it is well known that PRC2 is essential for normal development during mammalian embryogenesis (Meng et al., 2020; O'Carroll et al., 2001) it is not that clear during embryo development in *Arabidopsis.* The embryo-specific FIS(MEA)-PRC2 is required for endosperm development (nutritive tissue that supports embryo growth) by coordinating the expression of imprinted genes (genes that are expressed predominantly from one parental allele)(Grossniklaus et al., 1998).

Interestingly, PRC2 components CLF and MADEA (MEA) have been observed in embryo tissue (Luo et al., 2020; Simonini et al., 2021). Additionally, gene expression data in rapid-cycling genotypes (no functional FRI) showed that all core PRC2 units are expressed in the embryo, peaking at heart stage (Fig-1-5B) suggesting a potential function at that time (RNA-sequencing data from Hofmann et al., 2019).

Gene expression data show a very low abundance of FIS-PRC2 (containing FIS2 and MEA) in the embryo. This complex is known to coordinate imprinted gene expression in the endosperm. Embryo abortion in FIS-PRC2 is considered to be a result of failure of the endosperm to provide nutrients to the embryo. When separating embryo and endosperm development by using single-spermed pollen (*kokopelli, kpl*) in combination with *mea*, an effect on early embryo patterning and cell lineage determination was observed by impacting normal cell cycle progression (Simonini et al., 2021), showing that MEA has indeed a role in normal embryo patterning. In other cases, mutations in major PRC2 components still produced viable embryos and established a normal cellularization/ body plan.

1.3 Transcriptional modes at *FLC*

1.3.1 Noisy transcriptional levels

There is often strong variability within and between samples which could be due to several influences. While experiments are carefully replicated under controlled environmental conditions to reduce abiotic and biotic factors, there can still be internal influences.

An essential question in plant biology revolves around understanding the impact of inter-cell variability in gene expression or inter-plant variability. A recent RNA-sequencing investigation conducted on 14 individual Arabidopsis seedlings revealed a notable finding: roughly 9% of expressed genes displayed significant levels of inter-plant variability in their expression patterns (Cortijo et al., 2019). Those classified Highly Variable Genes include *FLC*. Furthermore, it became evident that *FCA* exhibited differential expression patterns over a daily (24-hour) cycle with slightly higher expression levels during the night.

Many genes show cell-to-cell heterogeneity of transcriptional levels which is only partly attributed to differences in cell size and DNA content (Melaragno et al., 1993). This additional level of variation can come from differences in the expression of the two alleles of a gene in the same cell (intrinsic noise) or from the level of variation between cells (extrinsic noise).

While many causes of variability exist, the necessity of sufficient sample size, replicates and single-cell observations to interpret the results becomes clear.

1.3.2 Digital switching during vernalization

Environmental signals are also noisy, as they naturally fluctuate. How noisy environmental signals are perceived and integrated to achieve *FLC* silencing has been of great interest in several studies in recent years. These studies have clarified a picture of distributed temperature sensing jointly contributing to a multi-layered regulation network with inputs ranging in time scale from hours to days to several weeks (Hepworth et al., 2018, Qüesta et al., 2020).

When *FLC* levels are measured during different times of vernalization on the whole plant (tissue) level, it is quantitatively graded. As mentioned in the section above (1.2.2 the vernalization pathway), *FLC* silencing occurs in cis for each *FLC* copy independently with a random probability for the switch to occur at each gene copy in each cell in the cold. This ON or OFF manner with an "all-or-nothing" switch between two bistable states (Figure 1-2A) is similar to a digital computer reading information in 'bits' (0 or 1).

Analogue nucleation would lose its sensitivity to a short period of cold, as a small peak is not sufficient in switching chromatin state at the locus (Angel 2015). Instead, digital silencing enables robust switching that is resistant to temperature fluctuations, and can also help buffer noisy transcription (Berry et al., 2017).

While *FLC* loci are known to have digital behaviour during and after cold treatment, allowing them to robustly hold epigenetic memory of cold exposure, it remains unknown how the starting expression levels (prior to cold) are regulated quantitatively in terms of analogue versus digital control.

1.3.3 *fca* alleles

The requirement of vernalization is dependent on the expression level of *FLC* in the seed/germinating seedling (Fig. 1-3B) and can be affected by variations in DNA sequence as observed in natural accessions fine-tuning flowering time by their local environment (Bloomer and Dean, 2017). After plants germinate, they grow as seedlings that have not experienced any cold exposure, or digital Polycomb switching, so quantitative variation in *FLC* expression seen in young seedlings could represent different cellular proportions of digitally regulated *FLC*, as well as graded transcriptional changes. More generally, elucidating the interplay between analogue and digital control is essential for a more in-depth understanding of quantitative gene regulation. Although digital and analogue modes of repression have been separately studied in the past (see e.g., Munsky and Neuert 2015) how these two fundamentally different modes of regulation might be combined has not been considered. *FLC* is an ideal system to study this question, due to its digital control in the cold, as well as the wealth of knowledge about its regulation at all stages (Berry and Dean, 2015; Wu et al., 2019b).

When understanding the differences between transcriptional states, a picture is formed in which different transcriptional states are first initiated by the antagonism of the autonomous pathway and FRIGIDA. In both pathways, the players are present in the cells at all times, without the need for an active memory component to remember the transcriptional state, as is the case for the vernalization pathway in the cold. Furthermore, it has been shown that transcriptional states can switch by gradually changing over time in individual cells and is maintained during cell divisions as observed during the establishment of silencing at the non-Polycomb targets *HML* and *HMR* loci in *Saccharomyces cerevisiae* (Goodnight and Rine, 2020).

Although the mechanisms initiating *FLC* silencing are fundamentally different during vernalization, the components required for successful repression still involve H3K27me3 maintained by PRC2 (Fang et al., 2020), suggesting that the mechanism of maintenance in a digital state is similar. Overall, *FLC* regulation by the autonomous pathway may therefore have both an analogue and a digital component.

To better understand the regulation of *FLC* by the autonomous pathway, and to decipher possible digital and analogue components of *FLC* regulation, we have used a series of quantitative knock-down mutants of one of the major autonomous pathway components,

FCA. The allelic series of *FCA* carries mutations in the gene that result in reduced functionality of the protein. We are particularly interested in the transcriptional modes of the intermediate cases, where the difference between analogue and digital regulation should be clearest. In addition, varying analogue regulation could make the system more sensitive to noise and less effective for long-term memory storage (Angel et al., 2015).

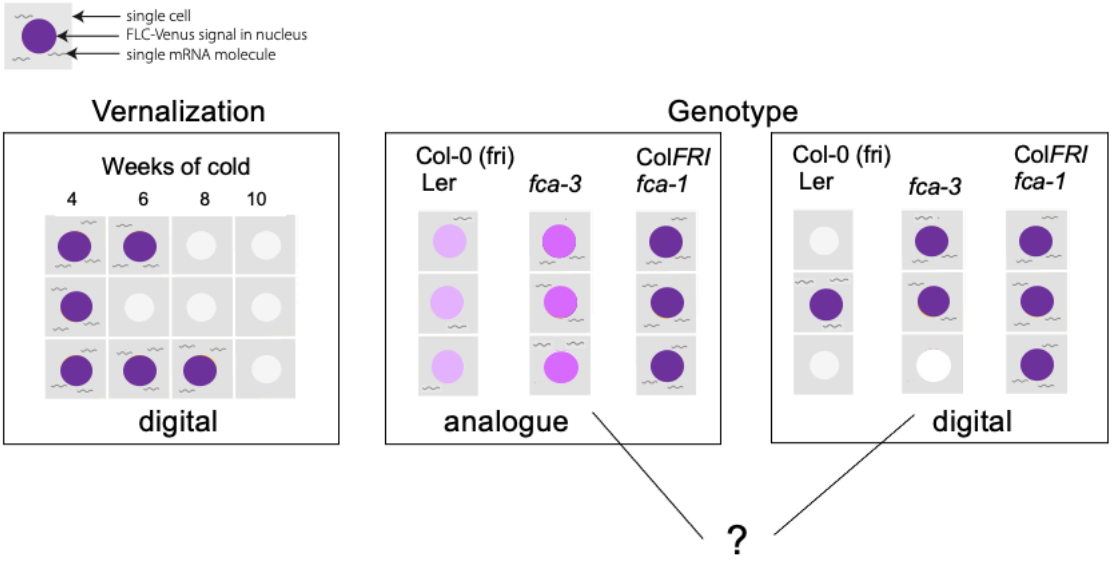


Figure 1–6 Potential modes of gene regulation at *FLC* at the single cell level.

Schematic illustration of possible modes of transcription regulation. During vernalization, *FLC* is digitally silenced. Digital regulation (left) corresponds to loci being in an 'ON' state (purple 4 weeks vernalization) or 'OFF' state (white, 10 weeks of vernalization), where we assume for simplicity that there is only one gene copy per cell. A change in the fraction of cells in each of the two states at the tissue level, shows switching off cells to from high to low average expression (columns left to right) while in analogue regulation each cell has a graded expression level that roughly corresponds to the overall population average. Intermediate average transcription levels as observed in *fca-3* could be observed by analogue and digital modes of gene regulation (middle and right box).

Additionally, our plant system does not require environmental (vernalization) stimulus to flower because it carries a natural mutation of FRI. Hence, those plants are rapidly cycling. This system allows us to investigate quantitative differences in *FCA* and their impact on *FLC* regulation. Using this methodology allows us to get a better understanding of co-transcriptional regulation and potentially different modes thereof at Polycomb target genes.

In the context of climate change, where current temperature extremes become normal, we cannot rely on existing vernalization for plants to gain the competence to flower. Therefore, it is of clear importance to understand the layers of gene regulation at the *FLC* locus. In the future, this could have large implications for the agricultural sector, as well as for food security.

1.3.4 Potential switching of states?

An open question that remains is how the switching from an ON to an OFF state occurs: does the switching occur in one direction (ON→ OFF) or two (ON →OFF → ON …)? So far, little is known about the timing of switching in the intermediate transcriptional state. It is tempting to hypothesize that the most destabilizing process is during replication when half of the histones containing epigenetic information are lost as information carriers.

1.4 Growth and cell division patterns

In Arabidopsis, post-embryonic cell growth and division occur in defined tissues, the pluripotent meristems. To potentially understand the heritability and switching of *FLC* transcriptional states in single cells, I will introduce the patterns of cell growth and division in the vegetative seedlings. The shoot apical meristem (SAM) gives rise to aerial organs like leaves and stems, while the root apical meristem (RAM) gives rise to the root.

In plants, cell division occurs within the meristems and cells maintain their position with respect to surrounding cells. Additionally, newly dividing cells push older cells from the dividing pluripotent part to the differentiating zone allowing us to determine the age of a cell based on its location in relation to the meristem.

1.4.1 Shoot apical meristem organisation

The SAM exhibits a zonal organisation consisting of several distinct layers or zones. These are the central zone (CZ), peripheral zone (PZ), and rib meristem (RM) (Medford, 1992). Cells in the CZ are responsible for maintaining the meristem and do not differentiate into specific tissues. The PZ gives rise to various aerial organs like leaves and stems. Cell division predominantly occurs in the CZ, where cells continually divide to replenish the meristem and provide cells for differentiation. As cells move away from the CZ, they start differentiating into specific cell types, leading to the development of leaves, stems, and other shoot structures resulting in the continuous production of new tissues and organs for the above-ground growth of the plant. Within the CZ, there are three different meristematic layers (L1-3). While L1 and L2 divide anticlinal (perpendicular to the surface), L3 divides both anticlinal and periclinal (Reddy et al., 2004).

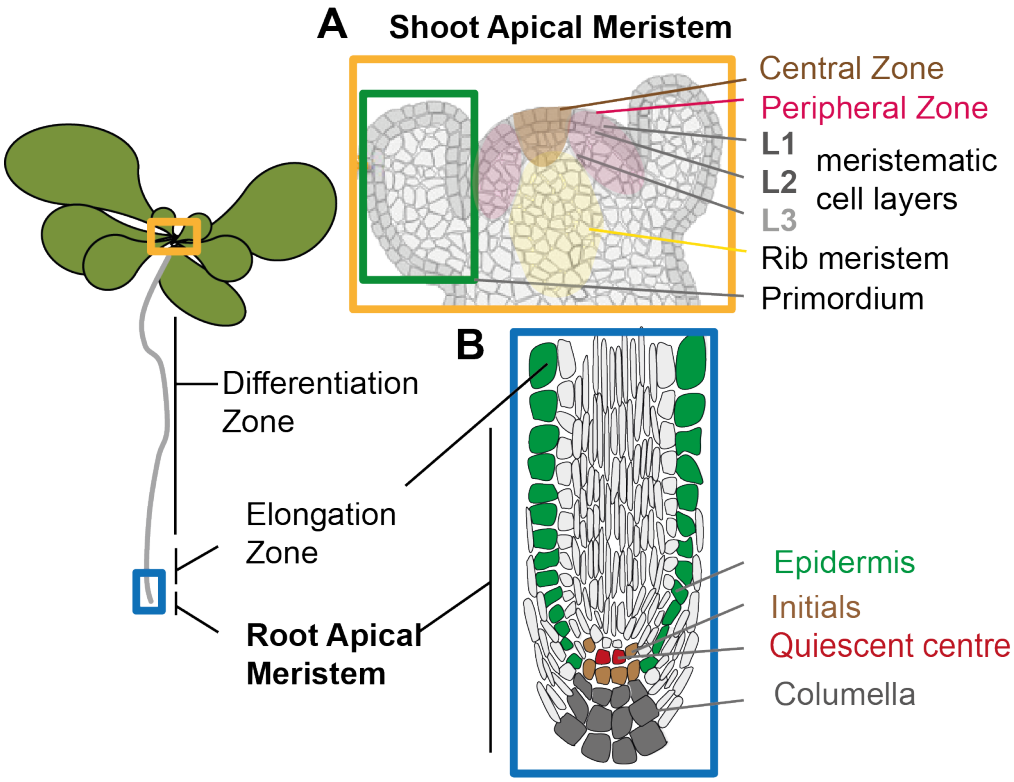


Figure 1–7 Apical Meristems of *Arabidopsis*

Schematic representation of the location of the shoot apical meristem (SAM) (orange box, A) and Root Apical Meristem (RAM) (blue box, B) in the vegetative seedling.

1.4.2 Root apical meristem organisation

The RAM also exhibits zonal organization, typically with three primary zones: the quiescent centre (QC), the proximal meristem (PM), and the distal meristem (DM). The QC contains slow-dividing or quiescent cells that help maintain the stem cell population. The PM contains actively dividing cells that give rise to the root cap and the vascular tissue. The DM consists of rapidly dividing cells that contribute to root elongation. Cell division occurs primarily in the PM and DM (Rahni and Birnbaum, 2019), with the QC playing a critical role in regulating stem cell fate. Cells undergo differentiation to various root tissues, including the root cap, epidermis, cortex, and vascular tissues. The further cells get away from the QC, the more they elongate, subsequently forming the elongation zone in which endoreduplication begins (cell cycle progression without cell division) (Hayashi et al., 2013; Joubès and Chevalier, 2000). The RAM is responsible for root growth and development, facilitating both primary root elongation and lateral root formation. In the DM, both the epidermal and underlying cortical tissues align their division plane in an anticlinal and transverse orientation. This orientation results in the formation of a distinct pattern characterized by parallel and adjacent cell files that are arranged along the longitudinal axis of the root (Dolan et al., 1993). Cells in one file are daughter cells of the cell-type defining initial cell and are clonally related (Dolan et al., 1994).

1.5 Imaging gene expression states of *FLC*

Central to understanding the regulation of *FLC* by the autonomous pathway, is to gain a better comprehension of *COOLAIR* expression. So far, most of our knowledge regarding gene expression is based on population averages of cells using methods such as RNA-sequencing or qPCR. To truly understand differences in cell types and developmental stages it is necessary to look at individual single cells or even individual loci.

To better understand the circuitry/feedback/interdependence of *FLC* and *COOLAIR* transcription it is of high interest to observe the *COOLAIR* transcripts at a single cell level. Using a single-cell approach unravelled the mutual exclusivity of *FLC* and *COOLAIR* transcription at the *FLC* locus (Rosa et al., 2016). Given the central role of *COOLAIR* in the regulation of *FLC*, it is also essential to discriminate the different isoforms at a single-cell level. As we will see, the development of an amplification-based single-molecule detection method will elucidate proximal, distal and nascent transcripts at a single-cell level.

Recent advances in imaging approaches have given rise to new possibilities to understand gene expression at a single molecule resolution. Single-molecule RNA fluorescence in situ hybridization (smFISH) is a method to visualize single RNA molecules in fixed cells using 18-22 nt long oligonucleotides labelled with a fluorophore that are each unique and complementary to the RNA (Fig. 1-8, left). Using this technique for *FLC* and nascent *COOLAIR* greatly improved our understanding of cell-to-cell variability (Ietswaart et al., 2017; Rosa et al., 2016). The limitation of the smFISH probe set is the relatively large size of the RNA that the probes need to hybridize with to be detected as diffraction-limited spots. Here, we are particularly interested in visualizing the 0.5 kb proximal *COOLAIR*. ClampFISH and SABER FISH are sensitive RNA detection techniques that take different approaches to overcome this limitation.

ClampFISH is a padlock system that uses a circular DNA backbone that covalently locks when bound to its target RNA, forming a DNA:RNA double helix (Fig. 1-8, middle). Several recursive rounds of the circular DNA padlock allow signal amplification. One concern of this method regarding the *FLC* 3′ end and *COOLAIR* initiation site (R-loop), is the accessibility of proximal *COOLAIR* to binding of the circular system. It is difficult to predict how the circular conformation of the padlock system might interfere with the secondary structure of the RNA and this approach therefore needs to be tested experimentally (Rouhanifard et al., 2019).

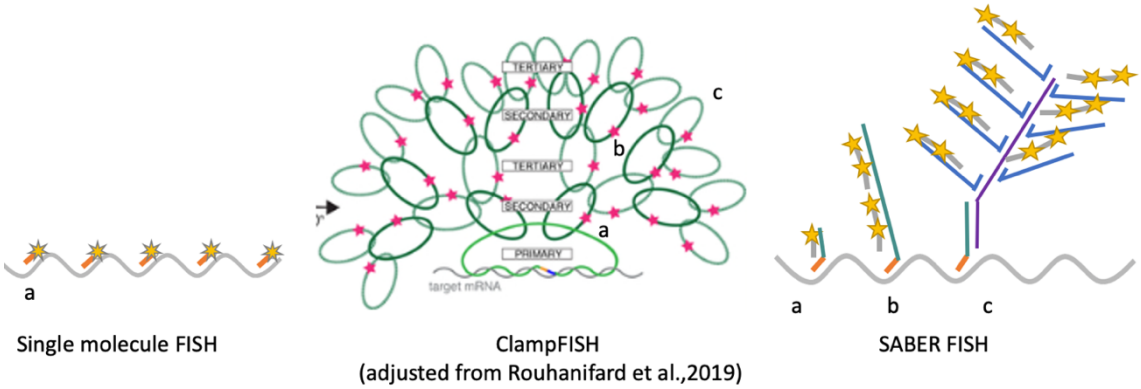


Figure 1–8 Comparison of amplification-based RNA detection systems.

ClampFISH and SABER FISH both use amplification cycles to increase signal. (a-c) represent the amplification cycles with (a) being no additional amplification, (b) being amplification to some degree and (c) amplification to a higher degree.

SABER FISH uses branched DNA probes that consecutively hybridize to the target RNA (Fig. 1-8, right). The unlabelled primary probes have an adjustable tail that can be extended by rolling cycle reactions according to the desired signal amplification. Additionally, a bridge sequence can be added that hybridizes to the tail sequence. The bridge sequence binds several fluorophore-tagged imager probes (Kishi et al., 2019).

Using amplification-based hybridization methods could overcome detection limitations due to fewer probe binding sites but could compromise sequence specificity. Moreover, the detection of proximal *COOLAIR* can become problematic if the RNA-probe binding sites are not accessible, due to, e.g., RNA binding proteins, secondary structures or R-loops. This complication means controls and optimisations are necessary. Overall, these methods are a powerful tool for understanding transcriptional differences at the single-cell level.

1.6 Outline of thesis

This thesis aims to understand the different modes of regulation of *FLC*, specifically how the autonomous pathway controls switching between *FLC* expression states and when such switching might occur.

In Chapter 1, our current knowledge of the regulation of *FLOWERING LOCUS C* (*FLC*) and its role in developmental transitions is introduced.

Chapter 2 investigates how analogue and digital modes of gene regulation both act on *FLC*. Quantitative fluorescent imaging is used to investigate quantitative differences in *FLC* expression revealing temporal separation between analogue and digital regulation. Most findings have been recently published in Antoniou-Kourounioti et al (2023).

In **Chapter 3** the modes of transcription regulation at *FLC* in earlier developmental time points are investigated by performing live-imaging at different embryo stages. The developmental time window in which cells gain the competence to digitally switch off is

identified. This switch-off involves the polycomb repressive complex as live imaging in a *clf-2* mutant (subunit of PRC2) revealed.

In **Chapter 4** the upregulation of *FLC's* antisense transcript (*COOLAIR)* in response to fluctuating cold conditions is investigated using FISH.

The results of Chapter 2-4 are discussed and put in a broader context in **Chapter 5** with implications of the thesis' findings for understanding *FLC* regulation and developmental transitions.

A detailed description of the experimental methods and materials including primer sequences and antibodies used in the thesis is described in **Chapter 6**. Abbreviations and acronyms are defined when they are first used and are summarised in the list of abbreviations.

Chapter 7 consists of the bibliography.

This thesis includes research material from the following published work that is attached in the Appendix.

1.7 Statement of Collaborations and Published work

This PhD project was part of a broader research endeavour aimed at unravelling modes of transcriptional regulation at *FLC*. Throughout this project, several collaborations were established, both internally and externally, and their collective efforts have been integrated into this thesis.

Chapter 2 showcases work that was performed by Anis Meschichi (Swedish University of Agricultural Sciences [SLU], Uppsala, now Eidgenössiche Technische Hochschule [ETH], Zurich), Rea Antoniou-Kourounioti (John Innes Centre [JIC], now University of Glasgow), Scott Berry (JIC, now University of New South Wales, Sydney), Lihua Zhao (SLU, Uppsala) and Yusheng Zhao (JIC, now Chinese Academy of Science, Beijing). In Chapter 3, specified data presented is attributed to the work of Eduardo Mateo-Bonmati (JIC, now Universidad Politécnica de Madrid [UPM], Madrid).

To acknowledge and ensure transparency regarding each contributor's work, every chapter begins with an abstract of the chapter including a statement summarizing the specific results they have provided. Moreover, within the Figures that incorporate their contributions, a footnote is included to highlight the corresponding experiment. All research findings presented and discussed in this thesis have been conducted with the full consent and collaboration of the individuals involved.

Published work integrated in this thesis:

Rea L Antoniou-Kourounioti, Anis Meschichi, Svenja Reeck, Scott Berry, Govind Menon, Yusheng Zhao, John Fozard, Terri Holmes, Lihua Zhao, Huamei Wang, Matthew Hartley, Caroline Dean, Stefanie Rosa, Martin Howard (2023) **Integrating analogue and digital modes of gene expression at Arabidopsis FLC** eLife 12:e79743 https://doi.org/10.7554/eLife.79743

Menon, Govind and Mateo-Bonmati, Eduardo and Reeck, Svenja and Maple, Robert and Wu, Zhe and Ietswaart, Robert and Dean, Caroline and Howard, Martin (2024) **Proximal Termination Generates a Transcriptional State that Determines the Rate of Establishment of Polycomb Silencing**. Molecular Cell https://doi.org/10.1016/j.molcel.2024.05.014

Published work not integrated in this work:

Lövkvist, Cecilia & Mikulski, Pawel & Reeck, Svenja & Hartley, Matthew & Dean, Caroline & Howard, Martin. (2021). **Hybrid protein assembly-histone modification mechanism for PRC2-based epigenetic switching and memory**. eLife. 10. 10.7554/eLife.66454.

Meschichi, Anis & Zhao, Lihua & Reeck, Svenja & White, Charles & Da Ines, Olivier & Sicard, Adrien & Pontvianne, Frédéric & Rosa, Stefanie. (2022). **The plant-specific DDR factor SOG1 increases chromatin mobility in response to DNA damage**. EMBO reports. 23. 10.15252/embr.202254736.

CHAPTER 2 : FROM ON TO OFF: UNDERSTANDING THE MODES OF GENE REGULATION AT FLOWERING LOCUS C

Abstract (with statement of main **collaborations**[1])

Proper regulation of developmental genes is vital for the successful lifecycle of organisms, and changes in transcriptional states play a crucial role in this process. Quantitative gene expression on a whole organism level can be achieved in two distinct ways at the single-cell level: through digital or analogue mechanisms. In the digital mode, gene copies exhibit binary ON/OFF expression states, and it is the proportion of copies in each state that changes at the cell population level. Conversely, in the analogue mode, gene expression is graded, with varying levels of transcription per gene copy. In this context, it is already known that the floral repressor *FLOWERING LOCUS C* (*FLC*) in Arabidopsis undergoes digital Polycomb silencing in response to cold, enabling quantitative epigenetic memory.

In this chapter, it is investigated whether analogue or digital regulation is involved in *FLC* repression prior to cold exposure. The chapter begins with a characterisation of an allelic series of mutations in one of the key players in the autonomous pathway, *FCA*, resulting in quantitative differences in *FLC* expression. A combination of quantitative fluorescent imaging (partly performed by **Dr. Anis Meschichi** [AM]) of *FLC* mRNA and protein, as well as mathematical modelling (performed by **Dr. Rea L. Antoniou-Kourounioti** [RAK]) reveal both analogue and digital modes of *FLC* regulation with a temporal separation between the two, with analogue regulation preceding digital regulation. Notably, intermediate analogue *FLC* expression levels facilitate a gradual population-level digital silencing, as individual gene copies stochastically and heritably switch OFF without cold stimulus. These findings present a novel perspective on gradual repression and shed light on the integration of analogue transcriptional and digital epigenetic memory pathways.

2.1 *Fca* alleles as tools for investigating quantitative expression of *FLC* before cold

FLC levels, which play a crucial role in the development and flowering of *Arabidopsis thaliana* plants, are determined during the early stages of development (embryogenesis) through a competition between an activator called FRIGIDA (FRI) and a repressive pathway referred to as the autonomous pathway. Quantitative variation in *FLC* expression determines the timing of the developmental transition to flowering (Bloomer and Dean, 2017). In the commonly used Arabidopsis accessions Ler and Col-0, mutations in the *FRI* gene allow the autonomous pathway to dominate, leading to repression of *FLC* during vegetative development and resulting in a fast-cycling summer annual lifestyle (Le Corre et al., 2002; Johanson et al., 2000). However, when an active FRI allele (Col*FRI*) is

[1] In this Chapter, as it was a collaborative effort of many contributors, major contributions are acknowledged under each Figure using footnotes with special emphasis on my contribution. Detailed contributions are mentione in the Methods .

introduced, it results in an initially high expression of *FLC*, which can be repressed by cold temperatures, leading to flowering. These extreme cases, with either high (Col*FRI*) or low (Ler, Col-0) *FLC* expression, do not provide the necessary range of *FLC* expression levels to understand the interplay between digital (all-or-nothing) and analogue (graded) control.

A genotype with intermediate *FLC* expression at the cell population level during embryogenesis is required to address this limitation. At a single gene copy level, such a genotype could exhibit either graded (analogue) or all-or-nothing (digital) *FLC* expression before exposure to cold, allowing the investigation of whether either or both forms of regulation are involved.

Interestingly, most of the identified mutants in the autonomous pathway have been discovered in the Landsberg *erecta* (Ler) background, in which *FLC* expression levels are lower due to the presence of a transposable element (TE) in *FLC*'s intron1, leading to reduced *FLC* RNA levels and early flowering (Liu et al., 2004).

In this work, we use a series of mutant alleles in *FLOWERING CONTROL LOCUS A* (*FCA*), a key player of the autonomous pathway consisting of several RNA binding factors that jointly and cotranscriptionally decrease *FLC* expression by impacting RNA processing of *FLCs* antisense transcript *COOLAIR* (Wang et al., 2014). This is accompanied by changes to a more repressive chromatin environment (Liu et al., 2007). The allelic series in the Ler background, without a dominant FRI, and with different mutations in *FCA,* results in quantitative differences in overall transcription levels of *FLC*. Moreover, we gain a system that expands the strictly functional and non-functional classes of repressing *FLC.* Instead, mutants with intermediate effect on *FLC* expression were chosen, which, in combination with single-cell assays, will systematically dissect the interplay of different modes of transcription. It is important to highlight that this study focuses on the effect of *FCA* mutations on *FLC*'s transcriptional regulation, rather than characterising how mutants in *FCA* may function. The genotypes used in this chapter will be characterised below.

FCA gene expression is regulated via alternative processing of the pre-mRNA producing four different mRNA isoforms (α, β, δ and γ) of which only the fully spliced *FCA*-γ transcript affects flowering time regulation (Macknight et al., 2002). The mRNA expression patterns show that *FCA* is strongest expressed in floral meristematic tissue (Schomburg et al., 2001), with its protein predominantly localised in the nucleus (Quesada et al., 2003). The *FCA* protein contains two RNA recognition motives (RRMs) that bind single-stranded RNA (Macknight et al., 1997), and its closest identified homolog, ETR-1, is an RNA-binding protein in *C. elegans* involved in muscle development by binding to specific mRNAs and regulating their translation (Macknight et al., 1997). *FCA* protein negatively regulates its own expression through alternative polyadenylation (poly(A)) of an early (proximal) poly(A) site. This interaction requires an intact WW domain (protein interaction domain) located at the C-terminal end between two prion-like domains involved in phase separation with other autonomous pathway components (Fang et al., 2019; Quesada et al., 2003).

Here, the transcript levels of *FCA*-γ for Ler, *fca-3* and *fca-1* as estimated by *FCA*-γ-specific primers biding to a 3′ region of the gamma isoform (Fig.2-1AB, binding site: orange) show significantly different expression levels between *fca-3* – Ler and *fca-3* – *fca-1* (p-values *fca-3* – Ler: 0.0030, *fca-3* – *fca-1: 0.0230*). Although they are significantly different, *FCA* transcript levels are not reduced. One possibility in *fca-3* and *fca-1* is that the lack of functional *FCA* is interfering with its autoregulation inducing compensating feedback loops onto *FCA*'s transcription. We conclude that the observed effect on *FLC* transcription is not due to straightforwardly varying levels of *FCA* mRNA. The wildtype, Ler, has a fully functional autonomous pathway, and exhibits low *FLC* levels and early flowering (Fig. 2-1A-C,E).

Fca-1 arose from an ethyl methanesulfonate (EMS) mutagenesis screen for late flowering mutants (Koornneef et al., 1991). The single nucleotide substitution (C to T) in exon 13 led to an amino acid change creating an early in-frame stop codon resulting in a truncated *FCA* protein lacking the WW domain (Macknight et al., 1997) (Fig.2-1 A). Even without the autoregulatory protein interaction domain, *FCA*-γ transcript levels as detected by the primer pair, are not significantly different from Ler. This indicates that the lack of auto-feedback from the resulting impaired or non-functional protein still leads to consistent levels of transcription. No further characterisation of the transcript was performed as in this work I am interested in the function of FCA on *FLC* rather than *FCAs* transcriptional regulation. Downstream targets of FCA are affected as shown by high *FLC* expression levels, and late flowering similar to Col*FRI* (Fig. 2-1C,E).

Fca-3 and *fca-4* are the result of X-ray-mutagenesis (Koornneef et al., 1991). They display intermediate cell population-level *FLC* expression (Fig. 2-1C,E, Fig. 2-2B) and flowering time (Fig. 2-1E). *fca-3* carries a mutation from G to A of the 3′ splice site in intron 6 (Chr4:9,210,672). An Alternative Splice Site Predictor (ASSP) was used to predict a new 3′ splice site in exon 7 (Fig. 2-1 D: see grey box in Ler (top) and *fca-3* (bottom) for constitutive splice site). The closest predicted weak alternative splice site (Chr4:9,210,582; 90 bp downstream, Fig. 2-1D turquoise box) is within intron 7, with the next constitutive acceptor splice site in intron 7 (Chr4:9,210,518; 154 bp downstream, Fig. 2-1D, green box). This suggests that either intron 6 is not spliced out or exon 7 is skipped with implications on protein folding and/or RRM2 functionality. Using a C-terminal polyclonal serum antibody (Quesada et al., 2003), no protein was detected as previously shown in both *fca-3* or *fca-1* (Macknight et al., 2002) and Fig. 2.1F. However, more characterization is needed to determine the exact protein that is made in *fca-3*.

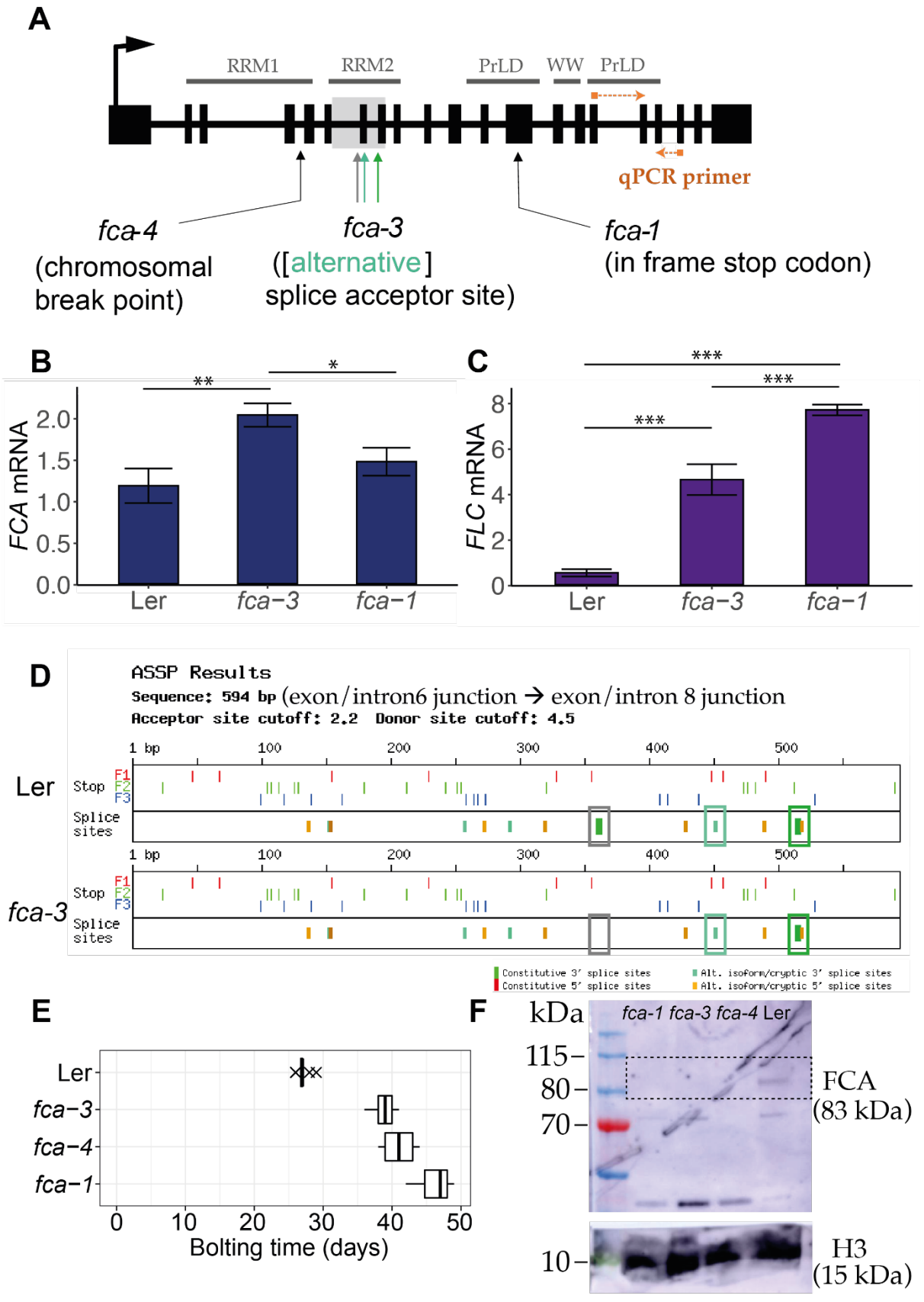

A
RRM1
RRM2
PrLD
WW
PrLD
qPCR primer
fca-4
(chromosomal break point)
fca-3
([alternative] splice acceptor site)
fca-1
(in frame stop codon)
B
FCA mRNA
Ler
fca−3
fca−1
C
FLC mRNA
Ler
fca−3
fca−1
D
ASSP Results
Sequence: 594 bp (exon/intron6 junction → exon/intron 8 junction
Acceptor site cutoff: 2.2 Donor site cutoff: 4.5
Ler
fca-3
Stop
Splice sites
Constitutive 3′ splice sites
Constitutive 5′ splice sites
Alt. isoform/cryptic 3′ splice sites
Alt. isoform/cryptic 5′ splice sites
E
Ler
fca−3
fca−4
fca−1
Bolting time (days)
F
kDa
fca-1 fca-3 fca-4 Ler
FCA
(83 kDa)
H3
(15 kDa)

Figure 2–1 Characterisation of *fca* mutation series.[2] *(previous page)*

(A) Schematic summary of mutations in *FCA* used in this study. Exons are represented as black boxes and introns by lines. The primer pair used (primer binding sites across exon junctions indicated by orange lines) is specific to the *FCA*-γ isoform in a wildtype context. **(B)** *FCA* -γ mRNA **(C)** *FLC* mRNA. Statistical tests: ANOVA suggests that the effect of genotype on *FCA* expression is statistically significant and large (p-value = 0.00277) with a 95% Confidence Interval. An additional pairwise comparison of estimated marginal means between genotypes showed that there is a significant difference between *fca-3* and Ler (*fca-1* - *fca-3* p-value: 0.0230, *fca-1* - Ler p-value: 0.2683, *fca-3* - Ler p-value: 0.0030). ANOVA suggests that the effect of genotype on *FLC* expression is statistically significant and large (p-value = $2.62e^{-06}$) with a 95% Confidence Interval. An additional pairwise comparison of (Bonferroni-adjusted) estimated marginal means between genotypes showed that all genotypes are significantly different from each other ((*fca-1*) - (*fca-3*) p-value: 0.0004; (*fca-1*) - Ler p-value: <.0001; (*fca-3*) - Ler p-value: 0.0001. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. (B-C) n=2 **(D)** Predicted missing splice site due to mutation in *fca-3*. Splice site predictions in wildtype Ler (top) and *fca-3* (bottom) starting from exon6/intron6 junction to exon8/intron8 junction (F1 - frame 1, F2 - frame 2, F3 - frame 3). **(E)** Flowering time as an indicator for transition to flowering is measured in days from sowing until the first appearance of a floral meristem (bolting). 12 plants were measured for each genotype. Boxplots show median and interquartile range (IQR), with data points outside median ± 1.5 IQR shown as crosses. **(F)** Immunoblot of *FCA* in *fca* mutants using a serum antibody (K4) against C-terminal *FCA* (Quesada et al., 2003) with H3 loading control. Unspecific bands can also be used as loading controls. Dashed black box highlights the 82 kDa *FCA* protein observed in Ler.

Fca-4 carries a breakpoint and chromosomal rearrangement within intron 4 of *FCA* (Page et al., 1999) (Fig.2-1 A) resulting in *FCA* intron 4 being fused to a predicted gene, *At4g1550* (GenBank accession No. Z97335), of unknown function. Protein detection using a serum antibody specific to the C-terminal half of *FCA* did not yield any detectable protein (Fig. 2-1F). Another strong allele (*fca-6*) that terminates the open reading frame very close to the breaking point in *fca-4*, has a similarly strong effect on *FLC* as *fca-1*, whereas *fca-4*'s intermediate effect (similar to *fca-3*) suggests that the fused protein containing the C-terminal *FCA* region may be partially functional and could contribute to the flowering time phenotype (Macknight et al., 1997).

Fca-3 and *fca-4* were empirically selected as intermediate mutants for this study based on their effects on *FLC* levels and on flowering time. To further dissect how the intermediate level can be achieved, the expression of *FLC* in single cells in the four genotypes was investigated by genetic transformation of an *FLC* reporter system, FLC-Venus into Ler.

Since regulation of the *FLC* locus is complex, the insertion site of the reporter system within the *FLC* locus is relevant to avoid disruptions, e.g., by inserting a reporter tag within a regulatory element in the sequence. Previous efforts in the lab in the Col*FRI* ecotype identified *FLC* exon 6 to be a suitable location as null mutants transformed with those constructs not only complemented Col*FRI FLC* mRNA levels but also restored the wild-type phenotype. This is indicated by a late flowering time that can be circumvented by vernalizing the plants, suggesting that FLC protein generated from the transgene is, in fact, functional (Berry et al., 2015).

[2] (E) was measured by Scott Berry (SB), I performed experiments of (B, C) and (F) and assessed splice site using ASSP (D).

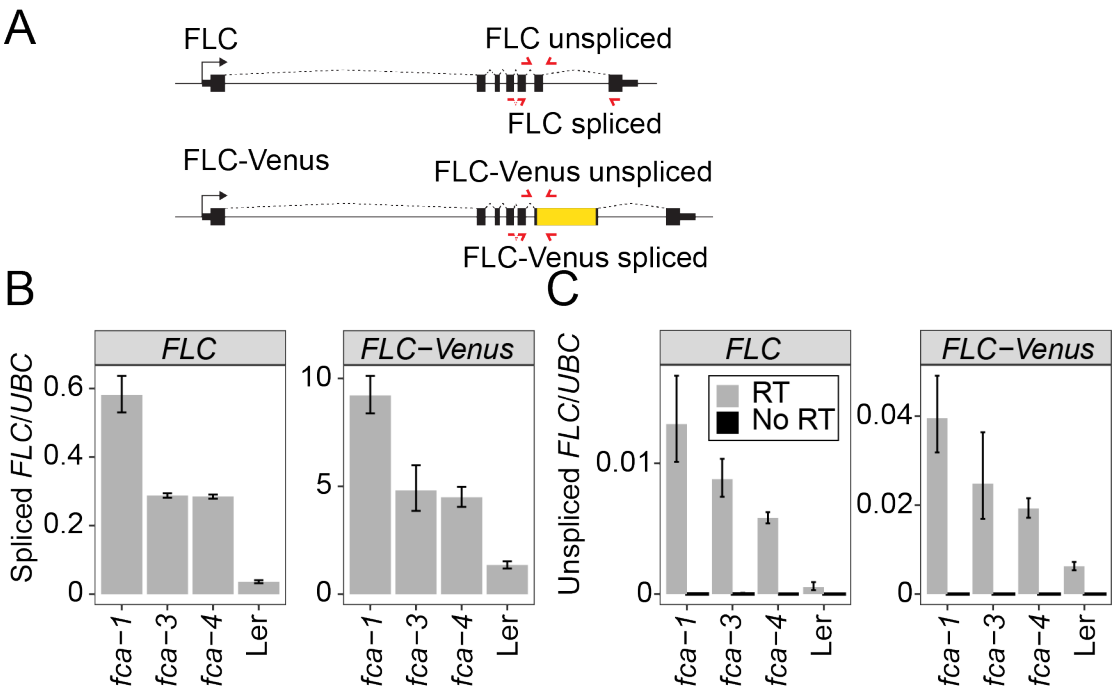


Figure 2–2 Characterisation of *fca* alleles and FLC-Venus transgene.[3]

(A) Diagram showing positions of qPCR primers (red) to quantify spliced and unspliced endogenous *FLC* and transgenic *FLC-Venus* RNA. The *Venus* sequence was inserted into the FLC sequence within exon 6. Transgene spans a region from 3.4 kb upstream to 8.6 kb downstream of the *FLC* transcription start site as shown in (Berry et al., 2015). **(B-C)** *FLC* and *FLC-Venus* RNA measured by qPCR from whole seedlings, relative to *UBC* RNA. No RT (Reverse Transcription) negative control in C shows no expression with same unspliced primers. Error bars show SEM of n=2 biological replicates for *fca-3* and n=3 biological replicates for other genotypes, harvested 14 days after sowing.

For that reason, and to avoid additional complications associated with the TE in Ler intron 1, the Col-0 *FLC-Venus* construct was transformed into the Ler background (Fig. 2-2A, Transformations performed by Huamei Wang). A single copy transgenic line was identified before crossing it into the mutant genotypes (*fca-1*, *fca-3*, *fca-4*). Comparing endogenous and *FLC-Venus* unspliced and spliced expression levels between the four genotypes, we observe >15-fold higher *FLC-Venus* expression compared to endogenous *FLC* (Fig. 2-2B) possibly due to differences in basal *FLC* transcription levels associated with the Ler *FLC* containing the TE or due to the position of the transgene. Importantly, the relationship between transcriptional levels and genotype is unaffected as indicated by a similar expression pattern for spliced and unspliced *FLC* specific for endogenous or -*Venus FLC*. In line with the RNA levels, the FLC-Venus protein signal was observed in nuclei of *fca-3*, *fca-4*, *fca-1* plants by confocal microscopy (see below, Figs. 2-4, 2-5, 2-6) reflecting mutant behaviours with Ler FLC-Venus protein being almost non-detectable due to low expression.

[3] All experiments in Figure 2-2 performed by SB

Taken together, the Col-0 *FLC-Venus* transgene insertion increases transcript levels compared to endogenous levels, while maintaining regulation by *FCA* making it a suitable system to investigate *FLC* transcriptional regulation on the single cell level.

2.2 Analysis of *fca* allelic series reveal both analogue and digital regulation at *FLC*

To gain a better understanding of transcriptional states in single cells of the intermediate genotype *fca-3*, individual mRNA molecules in cells of the meristematic root area were visualised using single molecule Fluorescence *in situ* Hybridization (smFISH) (Fig. 2-3 A,B) and quantified per cell (Fig. 2-4A). Although *FLC* is expressed in the root meristem, its biological function in the root tissue is not clear. Its impact on repressing flowering in the shoot meristem is has not been shown. Nevertheless, FLC-Venus Protein has been shown to respond in the same manner in primary leaves compared to roots as shown later in this Chapter (Fig.2-7) suggesting that we can insinuate the same regulatory mechanism. Performing smFISH in *fca-3* and *fca-1* targeting endogenous *FLC* mRNA (indicated by the purple-coloured probes binding to *FLC* mRNA sequence) in root squashes of the root tip, revealed no single molecule foci detection but rather *FLC* mRNA accumulated in the nucleolus (Fig. 2-3A).. As it cannot be ruled out that the accumulations are a side-effect of the TE, smFISH probes specific to the Venus sequence (30 individual oligonucleotides) and with a few *FLC*-specific probes (10 individual oligonucleotides, see Fig. 2-3 B) were used. These probes generated a signal specific to the transgenic *FLC* copy in these plants since no signal was detected in lines without the transgene. A comparison of whole-plant gene expression showed similar relative behaviour in the mutants for both the endogenous and transformed *FLC* (Fig. 2-2*)*. Targeting the *Venus* sequence gave us the additional advantage of being able to use the same lines for both mRNA and protein level quantifications. In this chapter, when referring to *FLC* mRNA in an smFISH context, *FLC-Venus* mRNA is meant. In the root squashes, exact cells types cannot be attributed.

Quantifying *FLC* mRNA transcripts per cell, we found that the average numbers per genotype were in line with previous gene expression results, with Ler having the lowest (mean of 2 (1.9)), *fca*-1 having the highest (mean of 37) and *fca*-3 showing intermediate (mean of 3 (2.9)) mRNA molecules per cell (Fig. 2-4A). The distributions of each genotype were significantly different from each other (as shown by performing a 3-way comparison with Kruskal-Wallis ($\boldsymbol{\chi^2(2) = 1611.21}$, p-value = 0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni adjustment $\boldsymbol{p \leq 1.2 \cdot 10^{-9}}$). For Ler, manual observation of the images revealed that the automated detection method for low mRNA numbers generated distributions that were indistinguishable from the background, as manual random sampling of cells did not detect any cells with mRNA. In contrast, when only focusing on the mRNA counts between 0 and 10 (Fig. 2-4A Insets, where 10 is the 95th percentile of the *fca*-3 data), there seems to be a difference in the tail in *fca*-3 suggesting a small population of cells with higher mRNA numbers in *fca*-3 compared to Ler. In *fca*-1, 89% of the cells have higher expression than the 95th percentile of *fca*-3 cells.

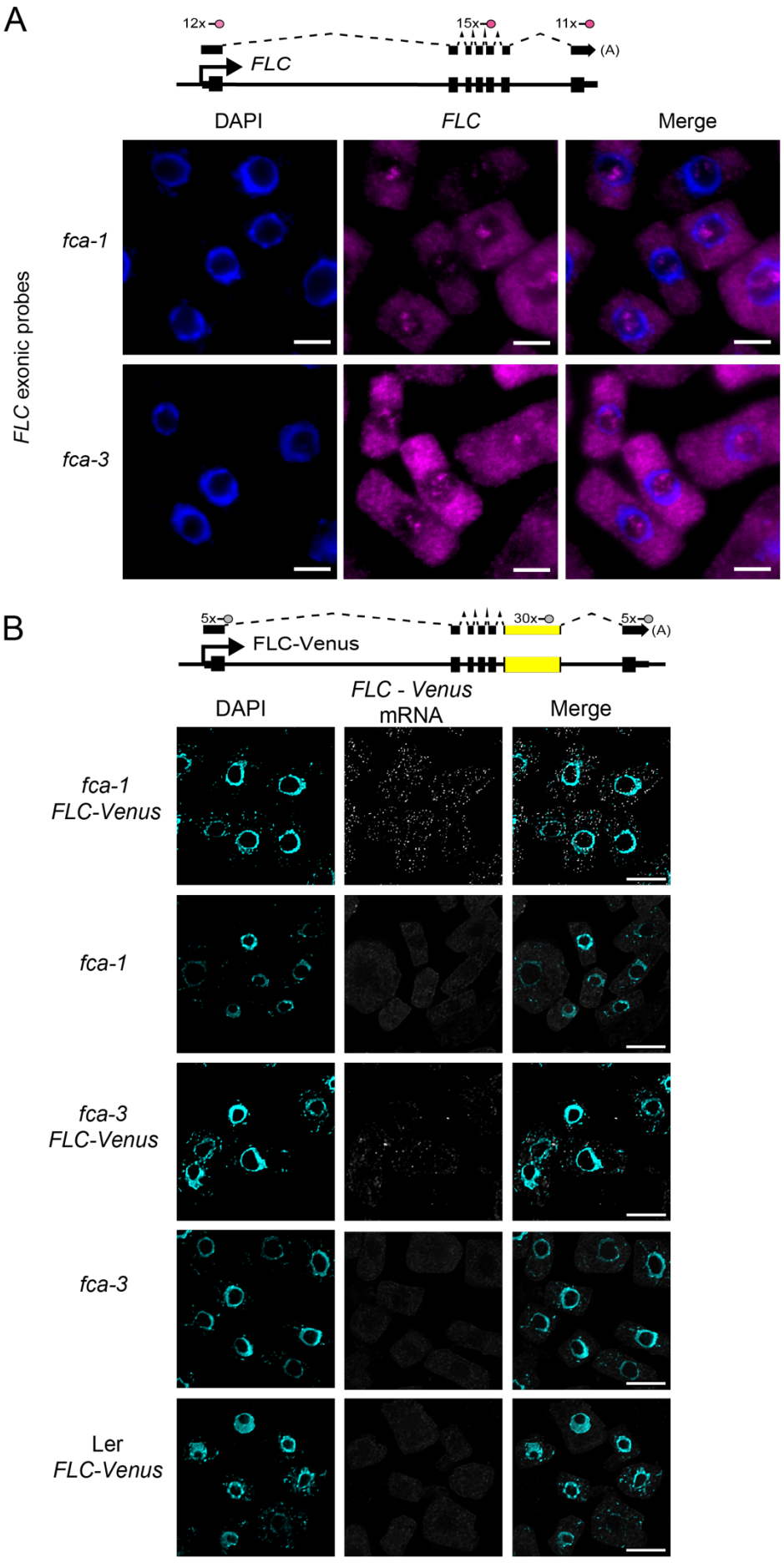

A
12x
15x
11x
(A)
FLC
DAPI
FLC
Merge
fca-1
fca-3
FLC exonic probes
B
5x
30x
5x
(A)
FLC-Venus
DAPI
FLC - Venus
mRNA
Merge
fca-1
FLC-Venus
fca-1
fca-3
FLC-Venus
fca-3
Ler
FLC-Venus

Figure 2–3 smFISH method for FLC-Venus imaging.[4] *(previous page)*

(A) Detection of ambiguous nucleolar signal using *FLC* mRNA -specific probes (magenta) in Ler background in *fca-1* and *fca-3 Arabidopsis thaliana* root cell squashes. Binding sites of 38 probes are depicted in the schematic on the top. The nucleus is stained with DAPI (blue). Scale bar, 5 μm. **(B)** Representative deconvolved images of FLC-Venus probes (grey) in the whole cell (nucleus and cytoplasm) which bind specifically to the transgenic *FLC-Venus* RNA and not to the endogenous *FLC* transcript (10 probes binding to *FLC* mRNA without Venus do not create a focus that can be detected with our imaging method). Binding sites of 40 probes are depicted in the schematic on top. All data from seedlings 7 days after sowing. The nucleus is stained with DAPI (cyan). Scale bar, 10 μm.

To gain a robust measurement for transcriptionally ON cells (expressing mRNA) for each genotype, different ON cell thresholds were tested ranging from >0 – >5 molecules per cell using an automated pipeline (Fig. 2–4B-F). We found that for all ON threshold levels, *fca-3* has considerably more ON cells than in Ler (Fig. 2-4), indicating that counts for *fca-3* are not simply background. As the threshold was increased away from zero, so that the background signal was progressively removed, the fraction of *fca-3* ON cells approached twice that in Ler. Hence, it appears that *fca-3* does have a proportion of ON cells, though less than in *fca-1*, and the levels of expression in these ON cells are much lower than in *fca-1*. Visual inspection and manual counting of images further confirmed these conclusions. Strikingly, when choosing a rather strict threshold of >3 molecules for the ON cells (Fig. 2-4E), the mRNA numbers in *fca-3* were only around 1/5 of those in *fca--1*, much less than one-half. This observation rules out the possibility that reduced levels in *fca-3* were solely due to the silencing of one of the two gene copies.

Nevertheless, mRNA numbers in many *fca-3* cells were close to zero, suggesting that digital silencing may also be relevant in this case. Hence, the differences in overall expression between these three genotypes appeared to have two components.

Firstly, a substantial proportion of cells exhibited a lack of detectable expression (designated as OFF), indicated by the presence of three or fewer mRNA counts. Specifically, in Ler, approximately 83% of cells displayed this OFF state, while in the *fca-3*, the percentage was around 72%. *fca-1* cells exhibited a significantly lower fraction of OFF cells, with only 1.5% falling into this category. Note, the fraction of OFF cells in Ler is almost certainly higher, as our automated analysis pipeline often mistakes a noisy background signal for mRNAs and therefore overestimates the mRNA number (see methods for details). Secondly, among the cells that did express *FLC* mRNA (designated as ON), the presence of *FCA* mutations resulted in an analogue (gradual) difference in their expression levels. Notably, the *fca-1* ON cells exhibited higher levels of *FLC* mRNA compared to the *fca-3* ON cells (as shown in Fig. 2–4). This observed behaviour clearly diverged from the digital regulatory pattern observed in *FLC* following exposure to cold treatment (as schematically depicted in Chapter 1, Fig. 1A) (Angel et al., 2011; Rosa et al., 2016).

[4] Imaging of smFISH performed by AM; Image preparation performed by SvR (Svenja Reeck)

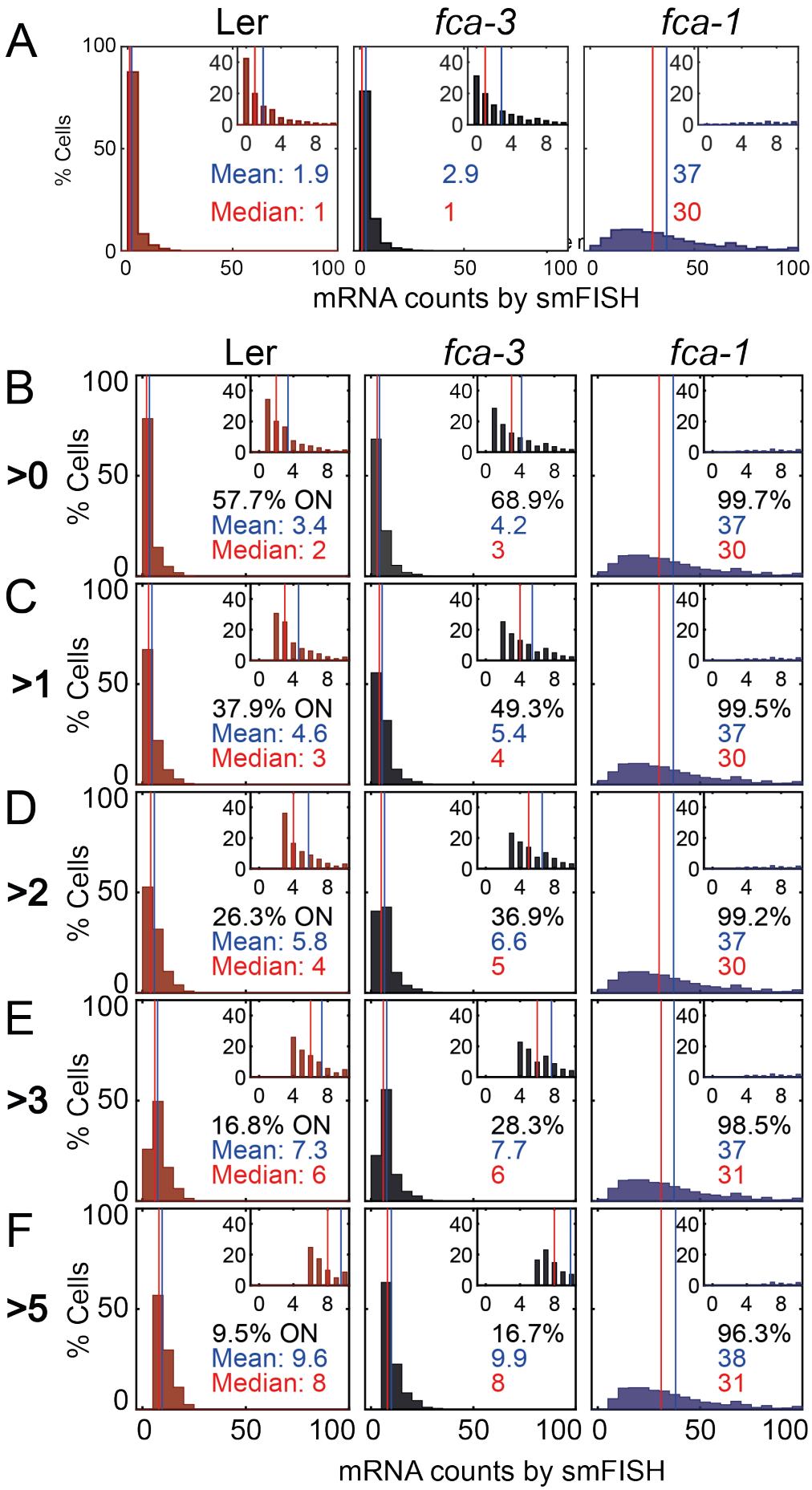

A
Ler
fca-3
fca-1
% Cells
Mean: 1.9
Median: 1
2.9
1
37
30
mRNA counts by smFISH
B
>0
57.7% ON
Mean: 3.4
Median: 2
68.9%
4.2
3
99.7%
37
30
C
>1
37.9% ON
Mean: 4.6
Median: 3
49.3%
5.4
4
99.5%
37
30
D
>2
26.3% ON
Mean: 5.8
Median: 4
36.9%
6.6
5
99.2%
37
30
E
>3
16.8% ON
Mean: 7.3
Median: 6
28.3%
7.7
6
98.5%
37
31
F
>5
9.5% ON
Mean: 9.6
Median: 8
16.7%
9.9
8
96.3%
38
31
mRNA counts by smFISH

Figure 2–4 Single molecule RNA quantification.[5] *(previous page)*

(A) All measurement histograms of smFISH results for each genotype (Ler, *fca-3*, *fca-1*) showing the number of single-molecule *FLC-Venus* RNAs detected per cell. The means and medians of the distributions are indicated in each panel. Insets show the same data for mRNA counts between 0 and 10 (where 10 is the 95th percentile of the *fca-3* data). For Ler n=853, *fca-3* n=1088, *fca-1* n=792; data from 7 days after sowing (two independent experiments). Statistical tests: 3-way comparison with Kruskal-Wallis ($\chi^2(2) = 1611.21$, p-value = 1.5e-285 (comparing all 3)) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni adjusted p-values for *fca-3* – *fca-1*: adj. p-value = $5.5 \cdot 10^{-275}$; *fca-3* – Ler: adj. p-value = $1.2 \cdot 10^{-9}$; *fca-1* – Ler: adj. p-value = $2.9 \cdot 10^{-257}$.
(B-F) Threshold for ON/OFF state of cells in smFISH experiments to determine ON fraction. The data plotted for the three genotypes but showing only measurements in "ON cells" with different thresholds defined as cells with more than the specified *FLC-Venus* mRNA molecules per cell, indicated at the beginning of each row. The Red number indicates the ON cells' median, the blue number indicates the mean, and the black number is the percentage of cells that satisfy the threshold condition. Note, the higher the threshold, the lower the number of ON cells, e.g., for (**E**), Ler n=143, *fca-3* n=308, *fca-1* n=780, from 7 days post sowing. Statistical test: Wilcoxon rank sum test, p-value= $4.8 \cdot 10^{-114}$. Kruskal-Wallis p-value = 5.9e-147. Comparison using Dunn-Sidak of *fca-3* and Ler: p-value = 0.97. Comparison using Wilcoxon rank sum test of *fca-3* and Ler: p-value = 0.49135, B-adjusted p-value = 1 (1.474)

We additionally used confocal live imaging to investigate FLC protein levels in individual cells in epidermal cells of the root meristem (Fig. 2-5, 2-6). After cell segmentation, FLC-Venus intensity within each nucleus was quantified, comparing the different genotypes (Fig. 2-5, 2-6, all differences are significant with $p \leq 1.7 \cdot 10^{-32}$).

By examining the mean intensity levels of FLC-Venus per cell and the overall histogram distribution, we again observed intermediate levels of FLC protein in *fca-3* compared to Ler and *fca-1 (*mean fluorescence intensity values ranging from Ler: 0.29, *fca-3*: 0.71 and *fca-1*: 3.8). Again, several indications pointed to the presence of a digital component in *FLC* regulation: In cells with the lowest protein levels, we observed similar levels in both *fca-3*- and Ler, which reinforced the contribution of a digital OFF state. Enhancing *fca-3* images to increase the Venus intensity to a similar level as in *fca-1*, a mixture of distinct ON and OFF cells was observed. In contrast, all cells in *fca-1* were in the ON state, while in Ler, cells appeared to be OFF even with an equivalent adjustment. These observations were depicted in Figure 2-5 (iv-v, enhanced) and Figure 2-6 (enhanced). Due to the growth-specific pattern and cell division in the root, cells within one 'file' derive from the same initial and hence are clonally related (ref). When observing cells with high FLC-Venus intensities in *fca-3* in a spatial context, we noticed the presence of occasional short adjacent ON cell sequences within a file suggesting a potential heritability of the ON or/and OFF states (indicated by yellow boxes in Figure 2-5 (v), Fig. 2-6). To gain confidence that the presence of these short ON sequences was a general feature rather than being specific to *fca-3*, an additional mutant allele, *fca-4*, was imaged also displaying intermediate *FLC* levels with similar short ON sequences (yellow boxes in Fig. 2-6).

[5] Manual cell segmentation performed by RAK, SvR, manual foci counting AM, SvR; Graphs: RAK, statistical analysis: RAK and SvR

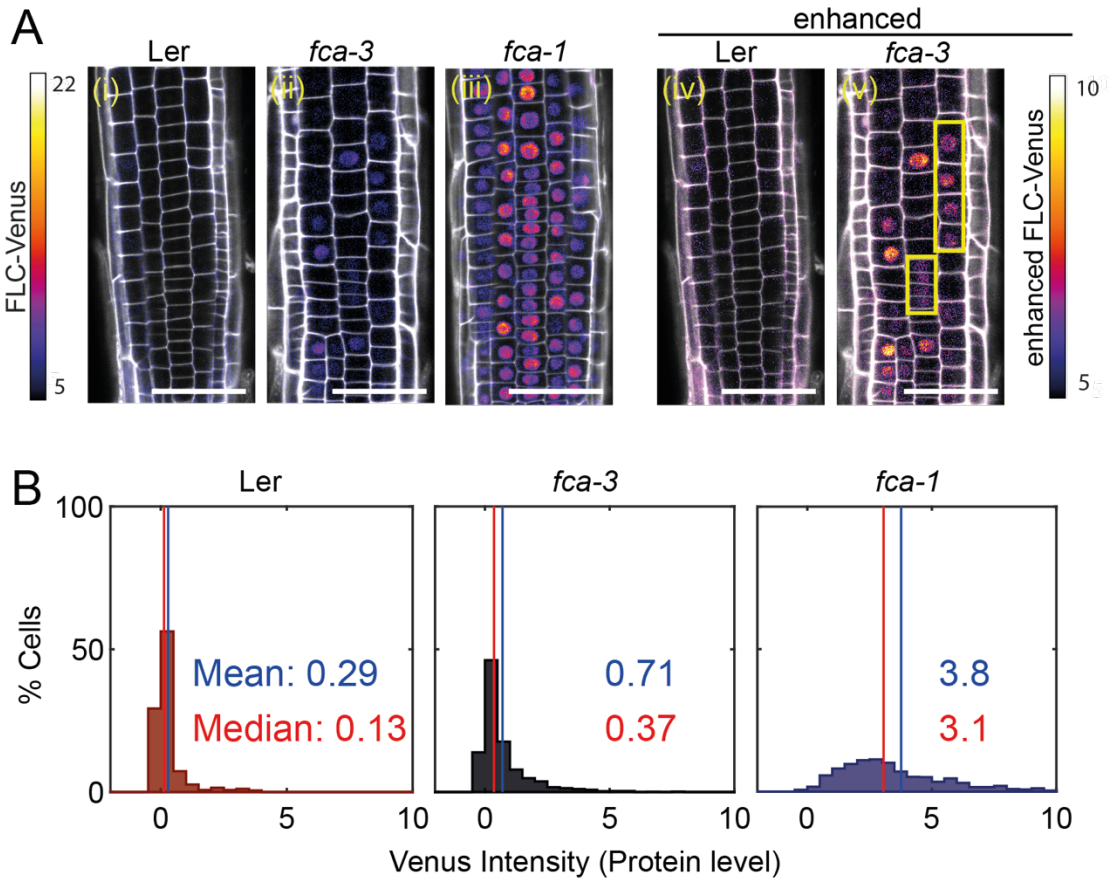


Figure 2–5 FLC-Venus per cell in *fca* mutants. [6]

(A) Representative confocal images of roots for each genotype. FLC-Venus intensity is indicated by colour maps; Grey shows the propidium iodide (PI) channel. The same settings were used for imaging and image presentation in (i)-(iii). Images in (iv) and (v) are the same as (i) and (ii), respectively, but adjusted to enhance the Venus signal by changing brightness and contrast (please note the different scale of colour map). Yellow boxes in (v) show short files of ON cells. Scale bar, 50µm. **(B)** Histograms of FLC-Venus intensity per cell in each genotype. The means and medians of the distributions are indicated in each panel. For Ler n=537 cells from 6 roots, *fca-3* n=1515 cells from 14 roots, *fca-1* n=1031 cells from 11 roots; data from 7 days after sowing (two independent experiments). Statistical tests: 3-way comparison with Kruskal-Wallis ($\chi^2(2) = 1607.56$, p-value = 0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni adjusted p-values for *fca-3* – *fca-1*: adj. p-value = $8.7 \cdot 10^{-281}$; *fca-3* – Ler: adj. p-value = $1.7 \cdot 10^{-32}$; *fca-1* – Ler: adj. p-value = $1.2 \cdot 10^{-201}$.

[6] Imaging in (A) performed by AM, Image preparations SvR

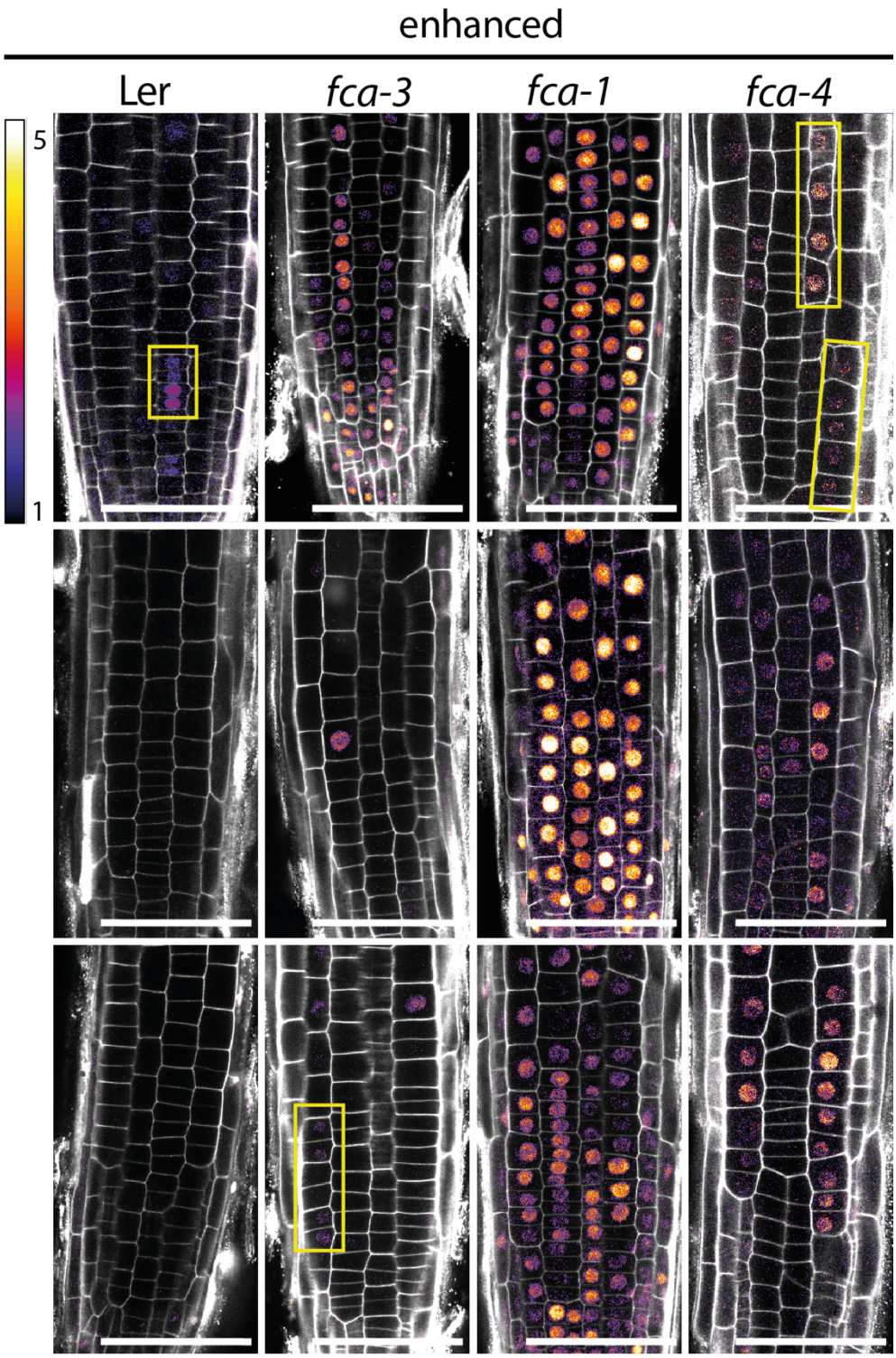


Figure 2–6 FLC-Venus imaging in *fca* alleles – root replicates and *fca-4*[7].

Representative images of FLC-Venus signal in four genotypes (Ler, *fca-3*, *fca-1*, *fca-4*) from 7 days post sowing. FLC-Venus intensity indicated by colour map; Grey shows the propidium iodide (PI) channel. Same settings were used for imaging and image presentation. Yellow boxes show files of ON cells in Ler, *fca-3* and *fca-4*. Scale bar, 50 μm. All imaging is enhanced (see text).

[7] Experiments and Image preperation performed by SvR

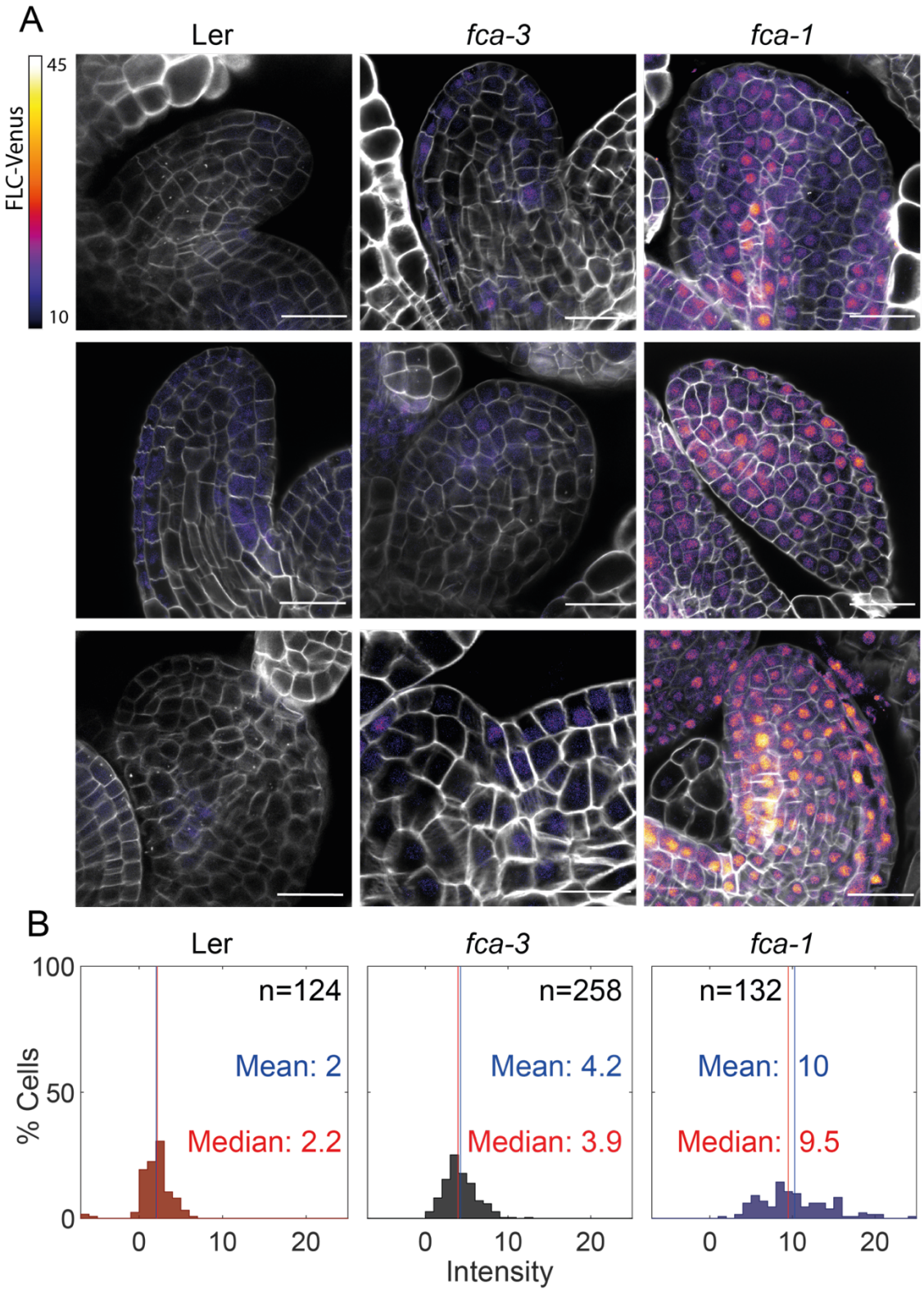


Figure 2–7 FLC-Venus imaging in *fca* mutants and wildtype in young leaf tissue. [8]

(A) Three representative images of FLC-Venus signal in the three genotypes (*Ler*, *fca-3*, *fca-1*) from 9 days post sowing. FLC-Venus intensity is indicated by colour map; Grey shows the cell wall dye (SR2200). Scale bar, 20 µm. **(B)** Quantification of FLC-Venus intensity per nucleus.

[8] Experiments performed by AM, LZ, quantification: SvR, Image projection: SvR, Statistical analysis: SvR

While using root imaging has the advantage of being able to integrate information on relative protein levels with the spatial distribution of cells in the root, it is not the main tissue contributing to the transition to flowering. For that reason, FLC-Venus imaging in fixed and cleared young leaves (primordia) (Fig. 2-7) was performed. Quantification of FLC-Venus fluorescence intensity per nucleus showed clear differences in the distributions between genotypes with mean values for Ler: 2, *fca-3*: 4.2 and *fca-1*: 10, consistent with what has been observed in the root, suggesting a global effect of each respective mutation in different tissues. Although division patterns in the primary leaves are more difficult to trace compared to the root, the presence of clustered cells with visible nuclear-localised protein in *fca-3*, and the absence of such in Ler, suggest an interplay of heritable analogue and digital components in *FLC* regulation in the leaf. Therefore, the conclusions we draw from root experiments can be extended to *FLC* regulation in primary leaves and help to extrapolate to all dividing tissues. Since imaging roots have the additional advantage of easily tracing heritable transcriptional states, further microscopy experiments were performed in root tissue only.

Overall, our results provide support for a combination of analogue and digital regulation in the case of *FLC*. In Ler, most cells were digitally OFF, while in *fca-1*, all cells were digitally ON. However, in *fca-3*, a fraction of cells was in the OFF state, and for those cells that were in the ON state, the level of *FLC* expression was reduced in an analogue manner relative to *fca-1*.

2.3 *FLC* RNA and protein are degraded quickly relative to the cell cycle length

Our findings strongly suggest that *FLC* regulation in mutants of the autonomous pathway involves digital switching. Yet, there is also an analogue component as indicated by the differences in expression (protein and RNA) of the loci in the ON state between genotypes. A previous study in yeast examined the expected RNA distributions for analogue and digital regulation scenarios (Goodnight and Rine, 2020). Here, we considered half-lives of the mRNA and protein as an important component when interpreting the analogue/digital nature of the regulation.

In a digital scenario, where switching between states occurs infrequently, long half-lives (similar to or longer than the cell cycle duration, which is approximately one day (Rahni and Birnbaum, 2019)) could lead to broader histograms of mRNA/protein levels due to the extended time required for levels to increase or decrease after switching states. This could potentially lead to misinterpretation of digital for analogue regulation, for example, if intermediate levels of mRNA/protein persist in OFF cells that descended from ON cells. Conversely, short half-lives would result in clearer bimodality. Additionally, what may appear as a heritable transcriptional state could also be due to the slow dilution of a stable

protein, as observed in other cases (Kueh et al., 2013; Zhao et al., 2020b). Therefore, it was necessary to measure the half-lives of mRNA and protein to interpret our observations.

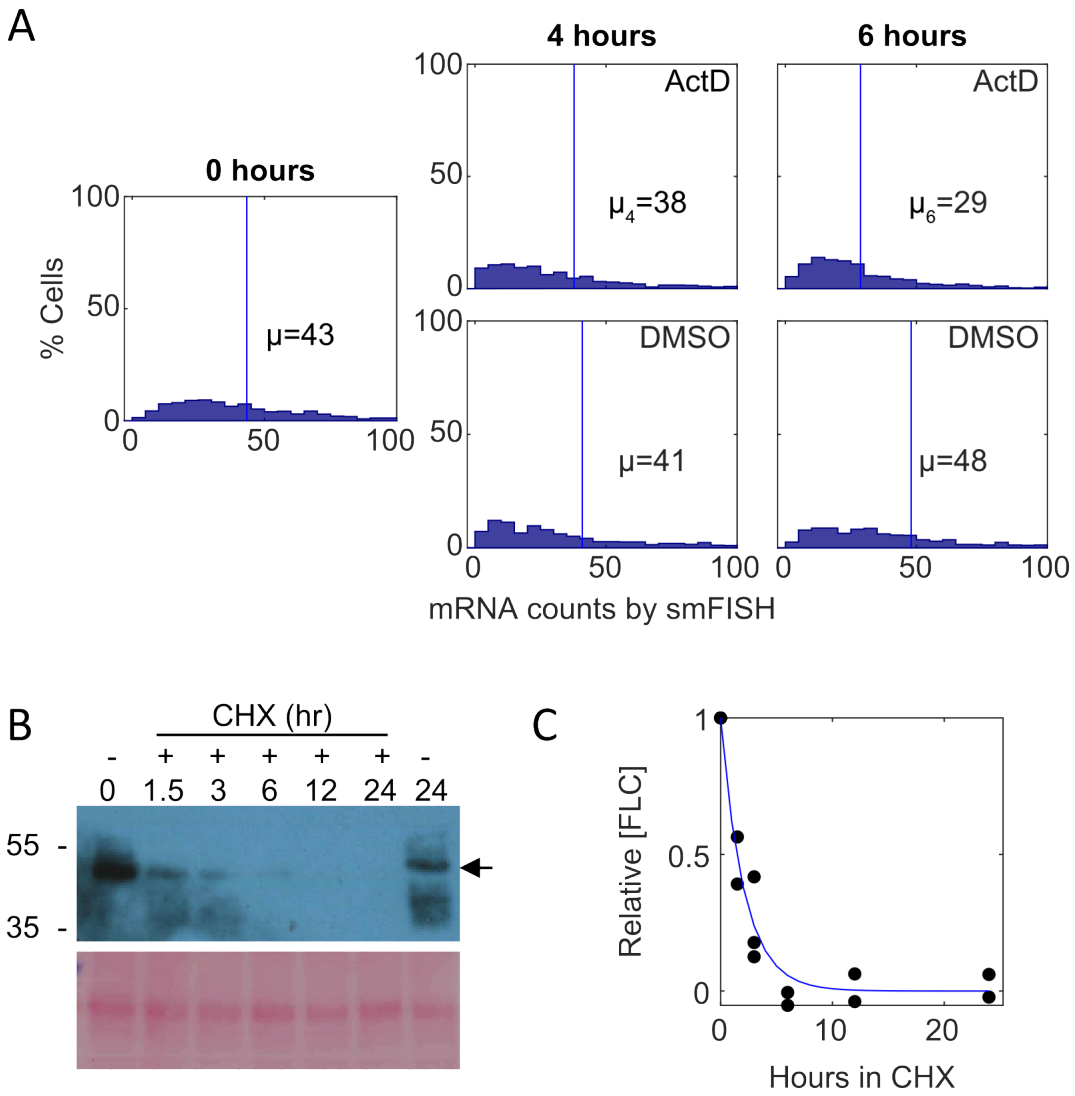


Figure 2–8 Degradation rate of *FLC* mRNA and protein.[9]

(A) Histogram of *FLC-Venus* mRNAs per cell measured by smFISH after 4 hours or 6 hours of ActD treatment (top) or mock DMSO treatment (bottom). Intronic PP2A probes were used to confirm the effect of ActD on transcription. Mean mRNA number per cell, μ, is indicated in each panel. Half-life $t_{1/2}$ of 5 hours calculated as in (Ietswaart et al., 2017) using the formula $t_{1/2} = ln(2)/d$, with $d = \ln(\mu_4/\mu_6)/2$, where d is the degradation rate and μ_4, μ_6 are the mean RNA counts after 4 and 6 hours respectively. For 0 h, n=1554, for ActD 4 h, n=820, for ActD 6 h, n=784, for DMSO 4 h, n=759, for DMSO 6 h, n=776, in seedlings 7 days after sowing. Three independent experiments were performed and combined. **(B)** Analysis of FLC-Venus protein stability with cycloheximide (CHX) treatment. 7-day old seedlings of *fca-1* plants carrying FLC-Venus reporter were subjected to liquid MS solution supplemented with 100μM CHX for 0, 1.5, 3, 6, 12, or 24 hours. Blot shows representative replicate. Non treatment control was also included (right-most well). Arrow indicates FLC-Venus detected by anti-GFP antibody. Ponceau staining was used as processing controls. **(C)** Quantification of the FLC-Venus band from blot shown in (B) and additional replicates. Combined data from n=3 biological replicates are shown with exponential decay profile fitted to first 3 timepoints (half-life=1.45hr).

[9] Experiments performed by (A) AM (B) YZ, quantification by (A) AM, SvR, RAK (B) RAK

In the Col*FRI* genotype, *FLC* mRNA has previously been shown to have a half-life of approximately 6 hours (Ietswaart et al., 2017). To exclude the possibility of genotype-specific differences, the half-life in the highly expressing *FLC* line, *fca-1*, as well as the protein half-life of FLC-Venus, were measured (Fig. 2-8) by inhibiting transcription using Actinomycin D (ActD) or translation using cycloheximide (CHX), respectively. The obtained mRNA/protein levels acquired by smFISH or Immuno-detection using the Venus antibody after different lengths of incubation with the respective drug gave a decay curve that was used to calculate the half-life. Our results revealed rapid degradation rates relative to the cell cycle duration, with mRNA and protein half-lives of approximately 5 (5.13) and 1.5 (~1.45 based on fit of first 3 time points) hours, respectively, whereas the DMSO mock control for mRNA as well as the 24hr untreated control for FLC-Venus protein showed no obvious decay.

Therefore, the apparent analogue regulation and observed heritability in our root images are unlikely to be attributed to slow degradation. Furthermore, the short protein half-life suggests that any potential effects from growth-induced dilution, which would result in reduced protein concentrations, are expected to be minimal.

2.4 Switching of *FLC* loci to an OFF state over time in *fca-3*

To understand the context of potential digital switching, it is important to determine the directionality of the switch, meaning whether switching occurs from ON to OFF, OFF to ON, or in both directions. If most loci are switching one-way only, in either direction, this would lead to a gradual change of overall *FLC* expression levels over time. Alternatively, two-way switching or non-switching in at least a few cells would be necessary to have a constant concentration of *FLC* mRNA/protein over time. These considerations, therefore, raised the related question of whether cell population-level silencing is at steady state at the time of observation, or, whether we are capturing a snapshot of a transient behaviour, with cells continuing to switch over developmental time.

By performing time course *FLC* expression analysis at 7, 15 and 21 days after sowing in the intermediate *fca-3* mutant, as well as *fca-1* and Ler, overall changes in expression could be observed. Indeed, a decreasing trend in *fca-3* and Ler was detected (Fig. 2-9A), which did not seem to be due to a change in *FCA* expression over the same timescale (Fig. 2-9B). This was particularly clear when comparing *FLC* and *FCA* in *fca--3* between 7 and 15 days: since *FCA* is a repressor of *FLC* expression, a decrease in both suggests that *FCA* is not the cause of the *FLC* expression decrease. We therefore concluded that the most likely explanation was primarily one-way switching from an otherwise heritable *FLC* ON state, to the heritable silenced OFF state occurring digitally and independently at individual loci in *fca-3* and Ler. We note that Ler was already mostly OFF at the starting time of 7 days and so the significant

but very slow rate of decrease (slope: -0.038, p-value: $4.8 \cdot 10^{-6}$, Fig. 2-9A; slope: -0.023, p-value: 0.0030, Fig. 2-10B) is interpreted as a small number of remaining ON cells continuing to switch OFF.

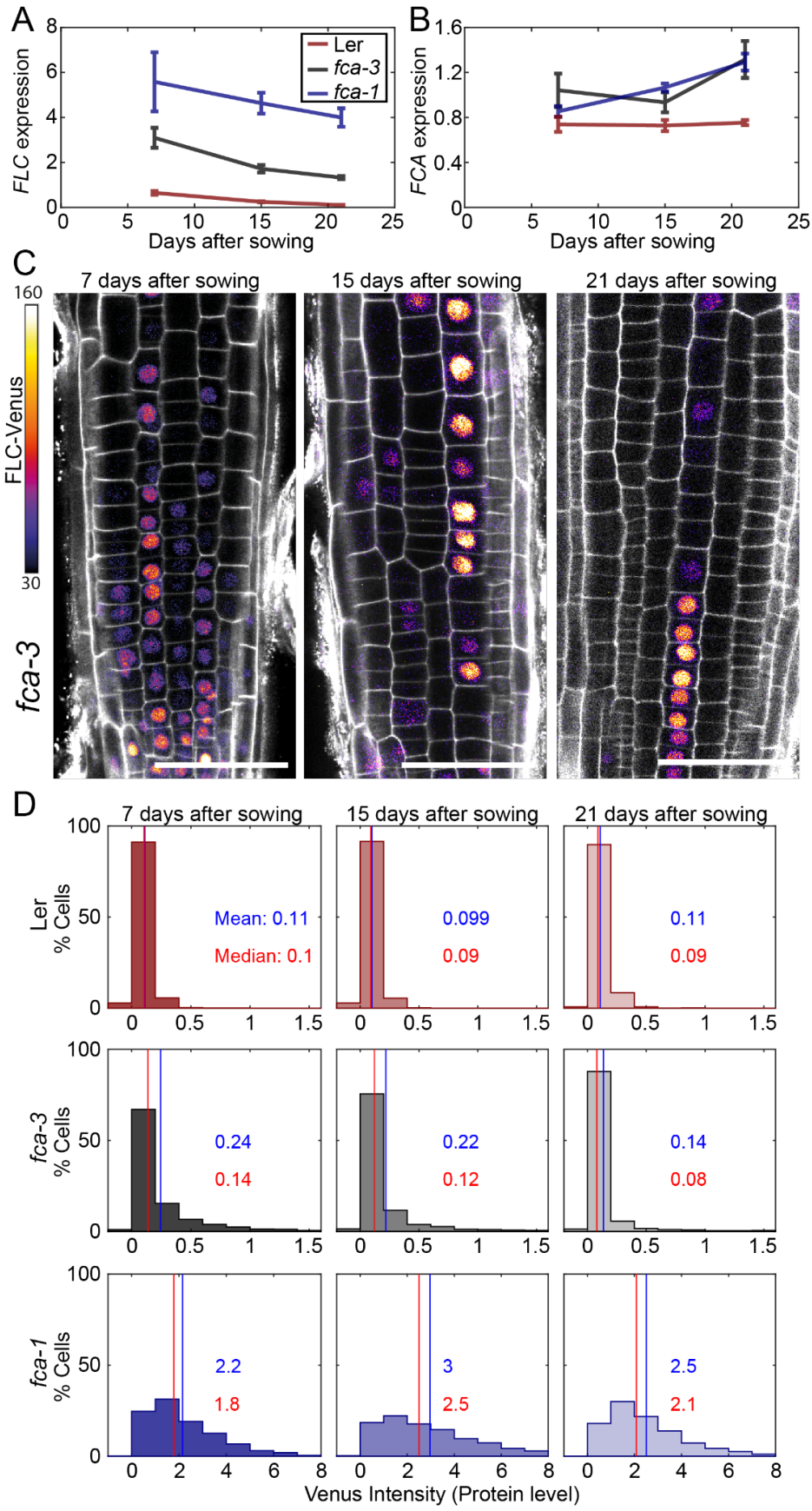


Figure 2–9 *FLC* loci accumulate in OFF state over time in *fca-3* experimental observation of gradual *FLC* silencing.[10] *(previous page)*

(A) Timeseries of *FLC* expression in *fca* mutant alleles transformed with the FLC-Venus construct. Expression is measured by qPCR in whole seedlings relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC)*. Error bars show SEM of n=3 biological replicates. Statistical tests: Samples were excluded as outliers based on Grubbs' test with alpha= 0.05. Linear regression on timeseries for each genotype. Slope for *fca-3* = -0.13, p-value = $2.0 \cdot 10^{-4}$; Slope for *fca-1* = -0.11, p-value = 0.12; Slope for Ler = -0.038, p-value = $4.8 \cdot 10^{-6}$. **(B)** Timeseries of *FCA* expression, otherwise as in (A). Statistical tests: Samples were excluded as outliers based on Grubbs' test with alpha= 0.05. Linear regression on timeseries for each genotype. Slope for *fca-3* = 0.021, p-value = 0.22; Slope for *fca-1* = 0.032, p-value = $3.8 \cdot 10^{-4}$; Slope for Ler = $1.1 \cdot 10^{-3}$, p-value = 0.84. **(C)** Representative images of *fca-3* roots by confocal microscopy. FLC-Venus intensity indicated by colour map; Grey shows the PI channel. Same settings were used for imaging and image presentation. Scale bar, 50 μm. **(D)** Histograms of FLC-Venus intensity per cell at each timepoint. The means and medians of the distributions are indicated in each panel. For Ler: 7 days, n=1121 cells from 10 roots (three independent experiments); 15 days, n=1311 cells from 9 roots (two independent experiments); 21 days, n=1679 cells from 12 roots (three independent experiments). For *fca-3*: 7 days, n=2875 cells from 24 roots (three independent experiments); 15 days, n=3553 cells from 23 roots (three independent experiments); 21 days, n=3663 cells from 21 roots (three independent experiments). For *fca-1*: 7 days, n=1022 cells from 9 roots (three independent experiments); 15 days, n=1770 cells from 12 roots (three independent experiments); 21 days, n=2124 cells from 12 roots (three independent experiments). Statistical tests: Linear regression on timeseries for each genotype. Slope for Ler = -0.00015, p-value = 0.44; Slope for *fca-3* = -0.0077, p-value = $4.0 \cdot 10^{-46}$; Slope for *fca-1* = 0.018, p-value = 0.00064.

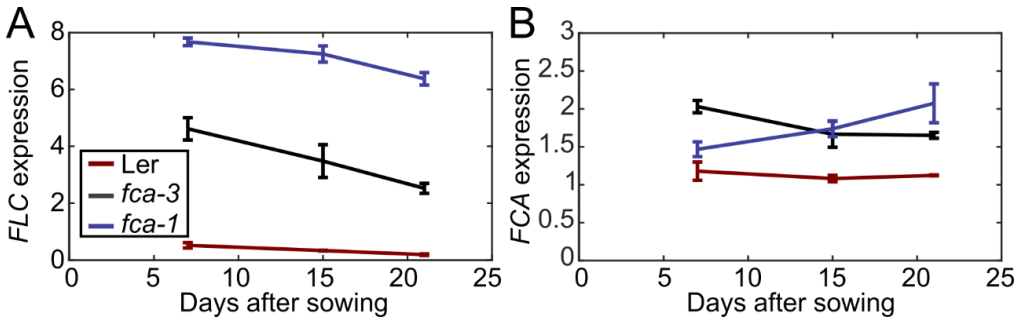


Figure 2–10 *FLC* and *FCA* expression in whole seedlings over time.[11]

(A) Timeseries of *FLC* expression in *fca* mutant alleles. Expression is measured by qPCR relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC)*. Error bars show SEM of n=3 biological replicates. *FLC* 7d timepoint is repeated from Fig. 2-1C. Statistical tests: Linear regression on timeseries for each genotype. Slope for *fca-3* = -0.15, p-value = 0.0063; Slope for *fca-1* = -0.091, p-value = 0.0054; Slope for Ler = -0.023, p-value = 0.0030. **(B)** Timeseries of *FCA* expression, otherwise as in (A). Statistical tests: Linear regression on timeseries for each genotype. Slope for *fca-3* = -0.028, p-value = 0.043; Slope for *fca-1* = 0.043, p-value = 0.031; Slope for Ler = -0.0044, p-value = 0.56.

[10] (A-C)_ Experiments performed by SvR (D) graphs: RAM
[11] (A-B) Experiments performed by SvR

However, at the later timepoints Ler had transitioned to flowering and so no biological conclusions were drawn from the *FLC* dynamics at those times in this genotype. Rather, this data was used as a negative control for the root imaging experiments below. In *fca-1* a downward trend was not statistically significant ($p = 0.12$) in Fig. 2–9A, likely due to the wide error bars at the first timepoint. In another experiment, however, (Fig. 2-10A) there was slow but significant decrease also in *fca-1*, suggesting there might be some cells switching OFF also in that case, but with a slower rate compared to *fca-3*.

By imaging FLC-Venus over time, we observed that the fraction of ON cells was indeed decreasing in *fca-3* over the time course (Fig. 2–9CD, Fig. 2–11), with mean fluorescence intensity values reducing by nearly a half between day 7 and day 21. In fact, after 21 days the pattern of ON/OFF cells in the *fca-3* roots was very similar to that of plants that had experienced cold leading to partial cell population-level *FLC* shutdown, with the majority of cell files being stably repressed, but still with occasional long sequences of cells in which *FLC* was ON (compare timepoint 21 in Fig. 2-9C and figures in (Berry et al., 2015). Similar decreasing trends of an ON fraction with prolonged developmental progression were observed in the other intermediate genotype *fca-4* (Fig. 2-12, note the slight difference in age at imaging). Contrary to that, other genotypes did not show a reduction in FLC-Venus levels over time in our imaging experiments. In *fca-1*, the fraction of ON cells did not show a consistent trend but the shape of the data stayed very similar, while in Ler, the fraction of ON cells remained –although at low levels– essentially constant. These results differed from the qPCR data (see above), which might indicate that different tissues could have different switching dynamics, e.g., epidermis cells in the root tip could switch off early, while a small number of cells remain ON in other tissues of the plant that continue to switch off, thereby explaining the slow decrease in the qPCR experiment (Fig. 2-9A, Fig 2-10A). Additional to that, there is the possibility of switching kinetics being different depending on the developmental stage of the root, e.g., when primary root growth is superseded by lateral root growth (Shishkova et al., 2008).

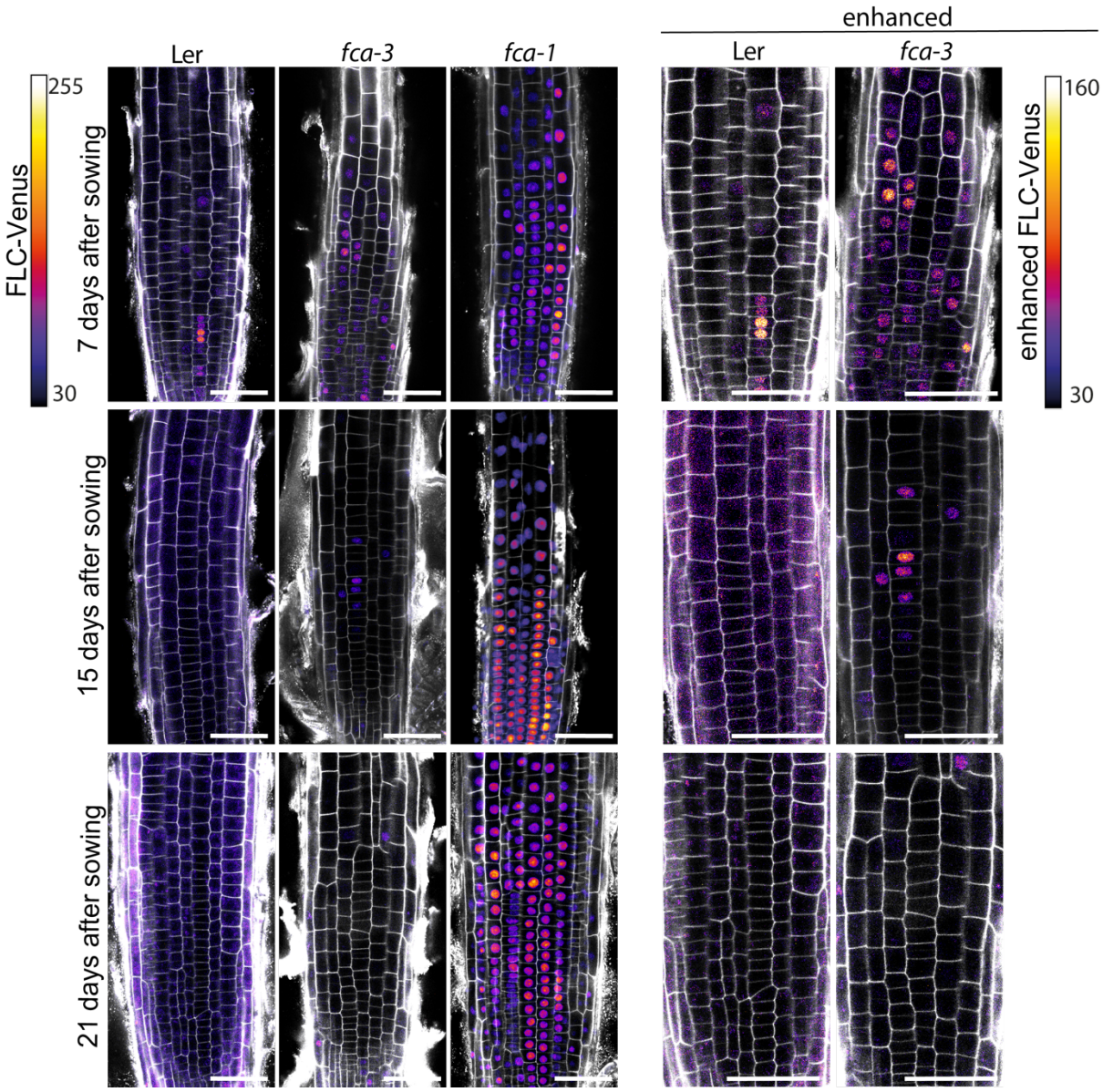


Figure 2–11 FLC-Venus timecourse replicates in *fca* alleles.[12]

Additional representative images of FLC-Venus signal in the three genotypes (Ler, *fca-3*, *fca-1*) at the three timepoints. FLC-Venus intensity is indicated by colour maps; Grey shows the PI channel. Same settings were used for imaging and image presentation in left three columns. "Enhanced" images (two columns on right) are a zoomed-in area of the same respective images, adjusted to enhance FLC-Venus signal to similar level as *fca-1* by changing brightness and contrast (please note different scale of colour map). Scale bar, 50 μm.

[12] Experiments performed by SvR

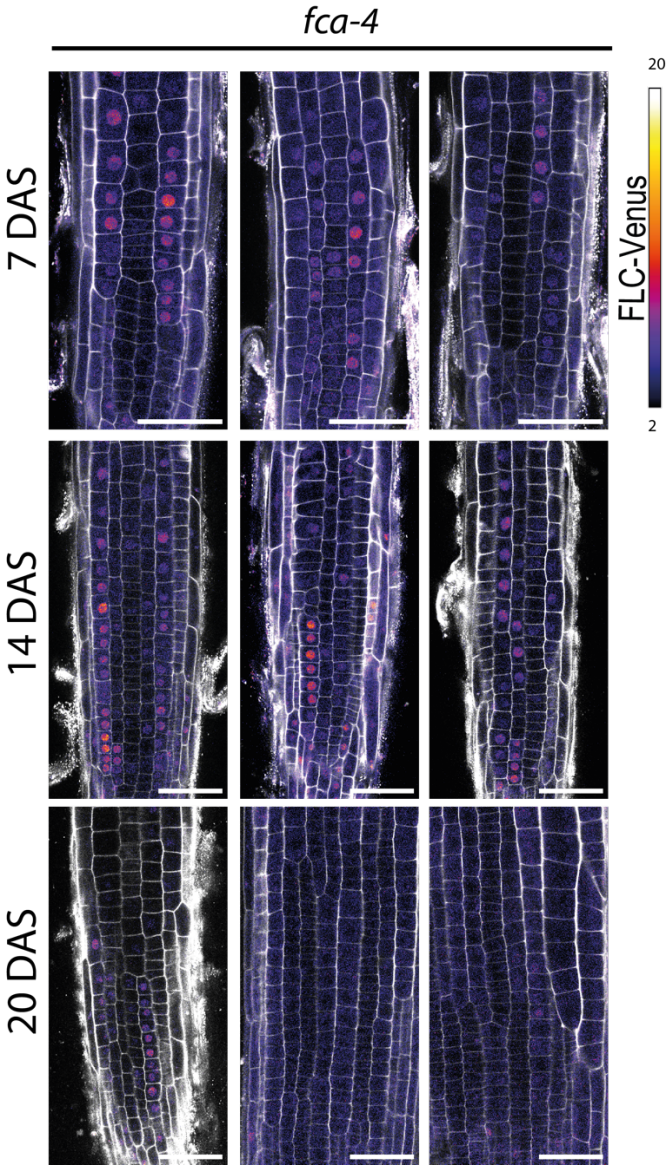


Figure 2–12 *fca-4* time series.[13]

Representative images of *fca-4* roots by confocal microscopy. FLC-Venus intensity is indicated by colour map; Grey shows the PI channel. The same settings were used for imaging and image presentation. Scale bar, 50 μm. *fca-4* 7 days after sowing (DAS) images are the same representative images as shown in Fig. 2-6.

We also examined whether the change observed in *fca-3* could be due to an analogue change in the fluorescence intensity of the ON cells rather than a decreasing number of ON cells over time. By setting a threshold at an intensity of 1, we looked at only the tail of the ON cell distribution (normalised by the total number of these cells in each condition) (Fig. 2-13). We see that high intensity cells are still present at later times and do not show a reduction in intensity relative to earlier timepoints. This finding is consistent with digital regulation: a decrease in the number of ON cells, but with all remaining ON cells expressing FLC-Venus at similarly high levels over time. Overall, our results in *fca-3* are consistent with progressive digital switching of *FLC* loci, primarily in the ON-to-OFF direction. Still, it

[13] Experiment performed by SvR

remains possible that there is a low level of switching in the opposite direction, from OFF to ON, in *fca-3* (see below).

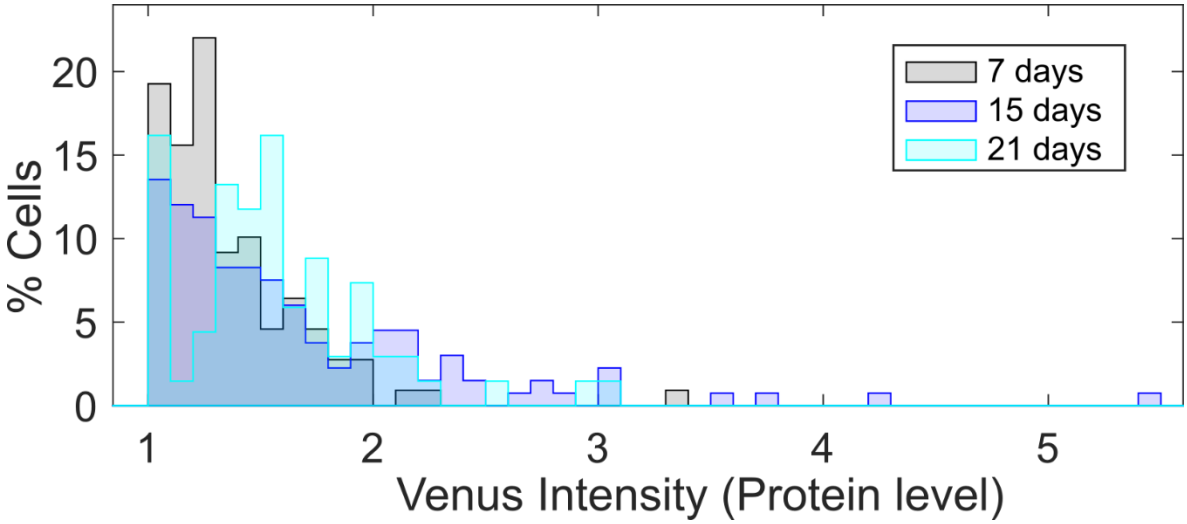


Figure 2–13 Fluorescence intensity of FLC-Venus in ON cells does not change over time in *fca-3*[14].
Histogram of ON cell tail in *fca-3* at three timepoints normalised by the total number of cells in each. Only cells with FLC-Venus intensity above 1 are shown. For 7 days n=109 cells from 17 roots, 15 days n=133 cells from 14 roots, 21 days n=68 cells from 6 roots. Data combined from three independent experiments.

2.5 All epidermal cells have the potential to switch

The files of ON cells in the epidermis hint at the heritability of the transcriptional state, but the identity of the switching cells remains somewhat unclear. We imaged FLC-Venus levels in the epidermal root initials, which are adjacent to the pluripotent stem cells of the root meristem (Quiescent Center) located the centre of the root tip. Those initials divide to form two different types of cells: Epidermis and root cap. Importantly, all epidermal cells in one file come from the same initial. If all initials showed higher FLC-Venus intensity compared to the mean epidermal signal, this would suggest that initials are ON, and switching would occur and then be maintained in the progenitor epidermal daughter cells. On the other hand, if initials show high variability of FLC-Venus intensity suggesting ON and OFF cells are both present there, this could indicate that switching exclusively occurs in the initials, with the transcriptional state maintained in the subsequent daughter cells.

To address these issues, after segmentation, the mean fluorescence intensity of the brightest sphere in the cell, corresponding to the nucleus, was measured in epidermal initials as well as in the epidermal cells of that file (Fig.2-14A, orange arrows point to initials, grey arrows point to epidermis cells; note, as the imaging depth is at the centre of the root, the signal is much weaker compared to previous epidermal images, as the laser had to first penetrate through cell layers on top).

[14] Image segmentation and preparation of graphs performed by SvR and RAK

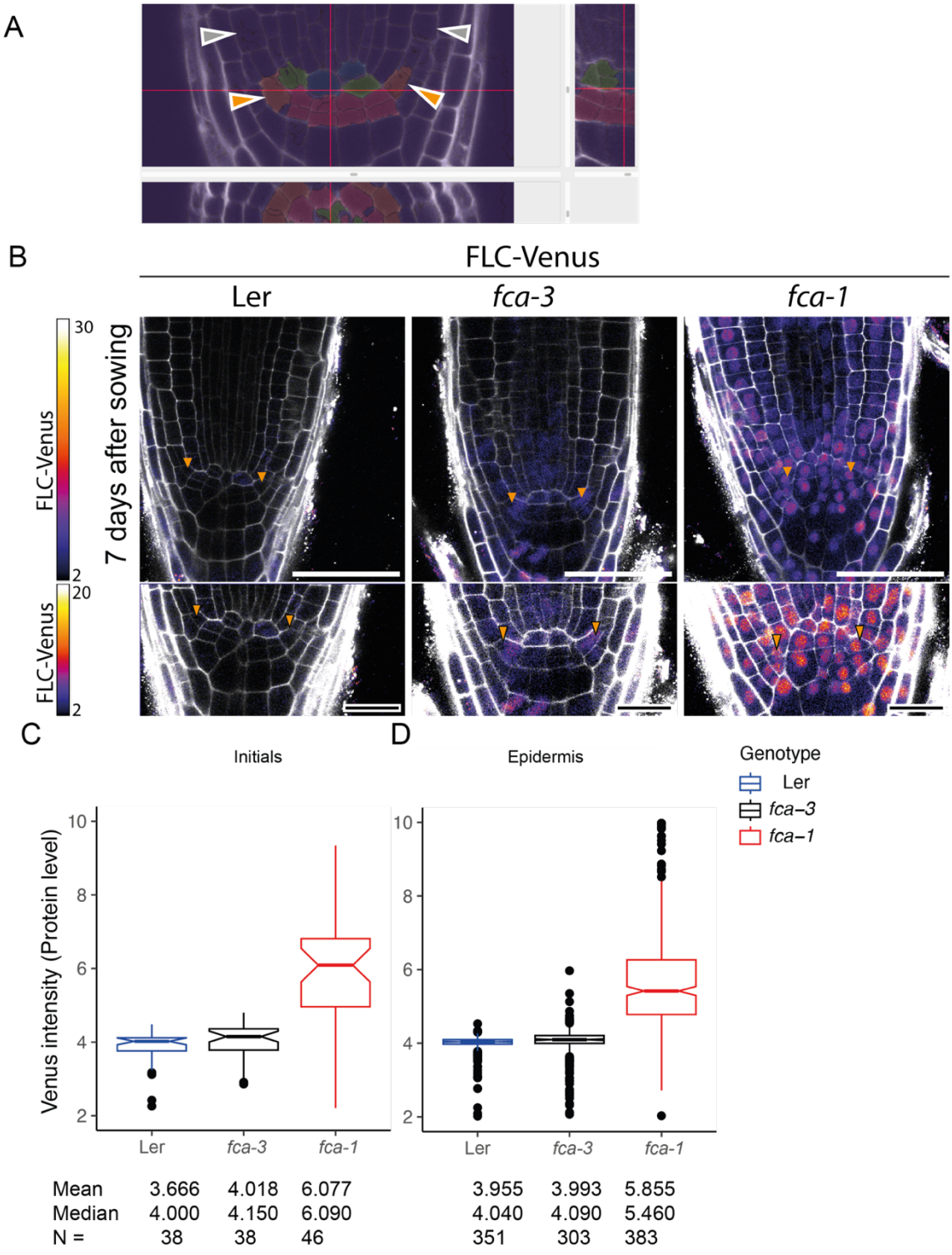


Figure 2–14 Root meristem middle section in *fca-1* and *fca-3* with FLC-Venus.[15]

(A) Manual 3D-segmented middle section of root tip with initials coloured in orange. **(B)** Representative images of FLC-Venus signal in the middle section of the root tip in three genotypes (Ler, *fca-3*, *fca-1*) covering several different cell types. FLC-Venus intensity indicated by colour map; Grey shows the propidium iodide (PI)- stained cell wall, epidermal initials highlighted by orange arrows (A-B), grey arrows point to epidermis (A). **(C)** Boxplots of FLC-Venus intensity per initial for each genotype. **(D)** Boxplots of FLC-Venus intensity per epidermal for each genotype. Boxes represent 25th – 75th percentile of the data, line within indicates median (50th percentile), whiskers extend to the furthest data point that is within 1.5 times the 25th – 75th percentile. Any data

[15] Experiments and image analysis performed by SvR

beyond whiskers are considered outliers (displayed as black dots). Means, medians and cell number (N) of the distributions are indicated under each respective boxplot. Data from 2 independent biological replicates from 19 roots. (C-D) Mann–Whitney U test to compare epidermal and initial Venus intensities showed no statistical significance (*fca-1* = 0.13, *fca-3* = 0.6, Ler = 0.06). Initials: Kruskal-Wallis = 46.88, p = 6.60e−11 with pairwise Dunn-Sidak (bonferroni-adjusted) p-values Ler–*fca-1*: $2.57 \cdot 10^{-10}$, *fca-3* – *fca-1*: $3.16 \cdot 10^{-6}$. Epidermis: Kruskal-Wallis = 593.50, p = 1.32e−129 with pairwise Dunn-Sidak (bonferroni-adjusted) p-values Ler–*fca-1*: $1.161 \cdot 10^{-113}$, Ler-*fca-3* : $1.76 \cdot 10^{-3}$, *fca-3* – *fca-1*: $2.41 \cdot 10^{-74}$

Firstly, the mean and median of quantified FLC-Venus signal in the initials are different between genotypes in line with previous epidermal imaging. Strong nuclear FLC-Venus intensity was detected in the initials of *fca-1*, with initials having a slightly higher although not statistically significant mean and median compared to epidermal cells (p-value = 0.13, Fig.2-14B-D). No significant differences in FLC-Venus intensities detected between initials and epidermal cells in Ler and *fca-3*. There was a significant differences between the FLC-Venus signal in *fca-1* and *fca-3* and Ler (p-values Ler–*fca-1*: $2.57 \cdot 10^{-10}$, *fca-3* – *fca-1*: $3.16 \cdot 10^{-6}$). If FLC-Venus signal was present in Ler and *fca-3*, it was not easily distinguishable from background in both initials and epidermal cells. Therefore, I cannot conclude that the initials were ON, although the boxplot (Fig.2-14C) suggests a possible proportion of cells showing higher FLC-Venus fluorescence.

Taken together, the limitations of deep-tissue imaging cannot give us the desired details to observe differences in the fluorescence intensity between initials and epidermal cells in the intermediate genotype *fca-3*. Moreover, I cannot exclude the possibility that switching is taking place in the initials.

2.6 Time lapse imaging reveals ON to OFF switching in *fca-3* and *fca-4*

To better understand which cells are switching and when, time lapse imaging was performed on the same roots over time in the intermediate genotypes. Two qualitative observations of switching (Fig. 2-15) from ON to OFF in *fca–3* (cells labelled 1 and 2) and *fca-4* (cell labelled 3) were made. These observed switches did not co-occur with cell division. All observed cell divisions led to two daughter cells with similar FLC-Venus levels as their progenitor cell, suggesting that the transcriptional states are inherited through cell division (see dashed boxes in Fig. 2-15, yellow ON cells; blue: OFF cells.) Note, nuclear signal in ON cells disappears during division, most likely due to the breaking down of the nuclear envelope during cell division. Consequently, to confidently assess if switching occurred, an additional sequential time point was imaged.

The experiment was performed 10 times with a total of ~20 roots at the start and several different imaging methods were tested (see Methods) with only 2 roots showing active cell cycle progression. One prevalent limitation was the strong laser power required to detect the very low signal in *fca-3*, damaging the root over long acquisition times. In this environment, the roots stop dividing and instead only elongate (Fig. 2-16). Interestingly,

when we did not observe an active cell cycle via cell division, we also could not observe any switching, hinting at a dependency between cell cycle progression and the potential to switch. However, while the number of switching events observed in this study is far too low to make justified statements about any switching behaviour, the fact that ON to OFF switching was observed three times, whereas the opposite switch from OFF to ON was not observed at all, agrees with our assumptions of either a one-directional switch or where the OFF switching rate exceeds the ON switching rate by some margin.

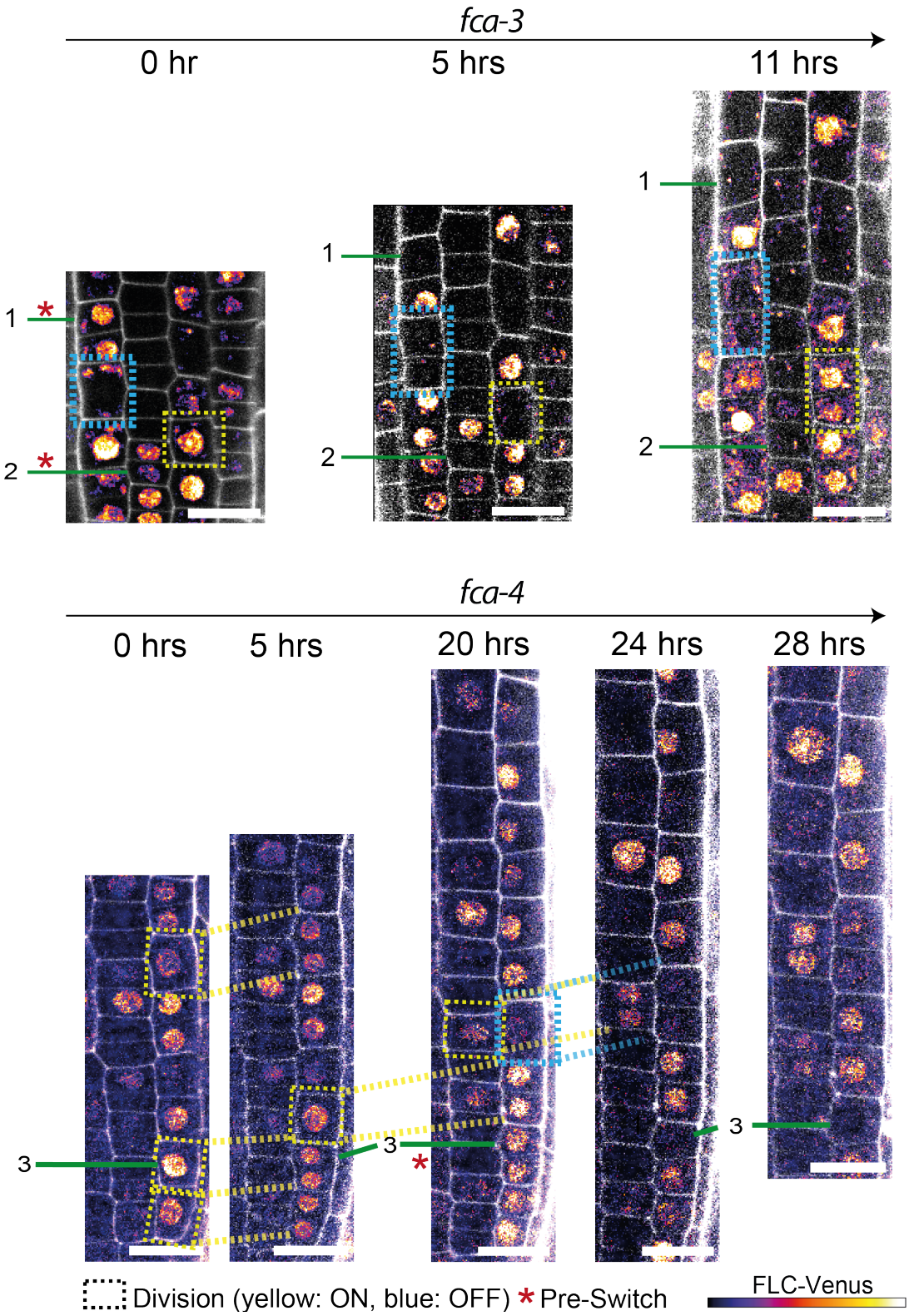


Figure 2–15 Qualitative observations of switching by time-lapse imaging in *fca-3* and *fca-4*.[16]

fca-3 (top) and *fca-4* (bottom), (1-3, green lines) indicate switches from ON to OFF, asterisk indicates pre-switch time point. Dashed boxes indicate cell division of mother cells into two daughter cells. Nuclear signal in ON cells disappears during the second imaging time point, most likely due to the breaking down of the nuclear envelope during cell division. Yellow boxes indicate ON cells and blue boxes indicate OFF cells. The same roots were imaged at each time point indicated on top. In maximum intensity projections, the FLC-Venus signal (indicated by colour map) does not directly correlate to protein level as it has been enhanced at subsequent time points because of bleaching. Scale bars indicate 20 μm.

[16] Experiments performed by SvR

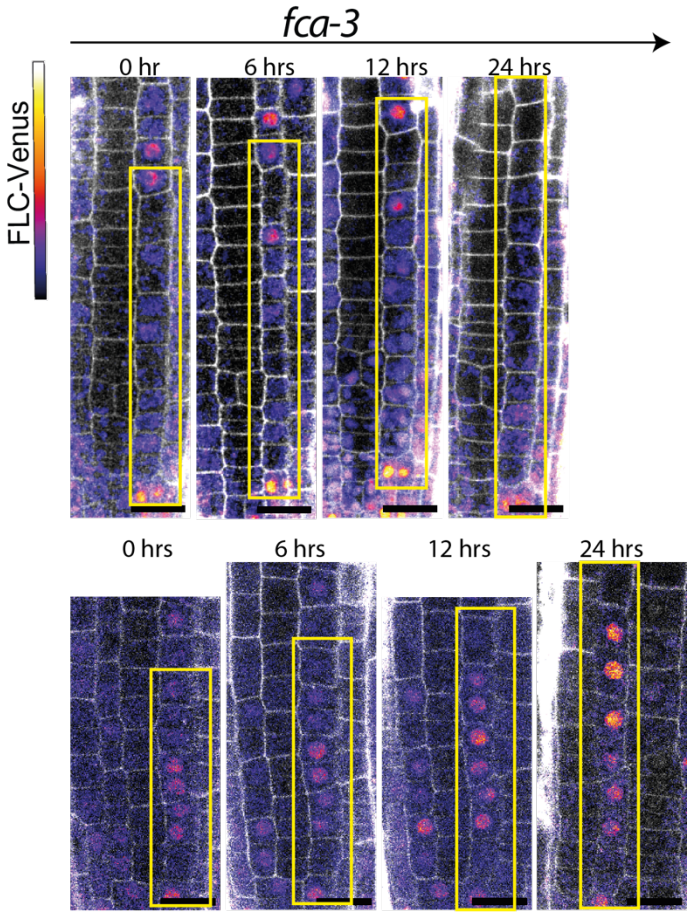


Figure 2–16 Two Examples of time lapse imaging in *fca-3* roots with no division and no switching.[17]

Two roots (top and bottom panel) were imaged at each time point over the course of 24 hrs as indicated on top. The maximum intensity projections of the FLC-Venus signal (indicated by colour map) do not directly correlate to protein level as it has been enhanced at subsequent time points because of bleaching. PI- stained cell wall; grey. Yellow boxes highlight one cell file without cell division or switching, but rather growth by elongation. Scale bars indicate 20 μm.

2.7 Mathematical model incorporating digital switching can recapitulate the *FLC* distribution over time in *fca-3*

To determine whether our interpretations of the data on analogue and digital components in *FLC* regulation are sufficient to reproduce the experimentally observed patterns in the root, Rea Antoniou-Kourounioti built a mathematical model for *FLC* RNA/protein levels in the root. She constructed a growing and dividing model incorporating digital *FLC* state switching embedded in simulated dividing cells in the root. We allowed the *FLC* state in the model to switch either ON-to-OFF, or OFF-to-ON. We focused on modelling the switching dynamics and used distributions for the protein levels in cells with two ON loci, one ON and one OFF, and two OFF empirically fitted to our data (Fig. 2-17A, Methods).

[17] Experiments performed by SvR

Possible effects from cell size and burstiness are incorporated into these empirical distributions implicitly through this fitting.

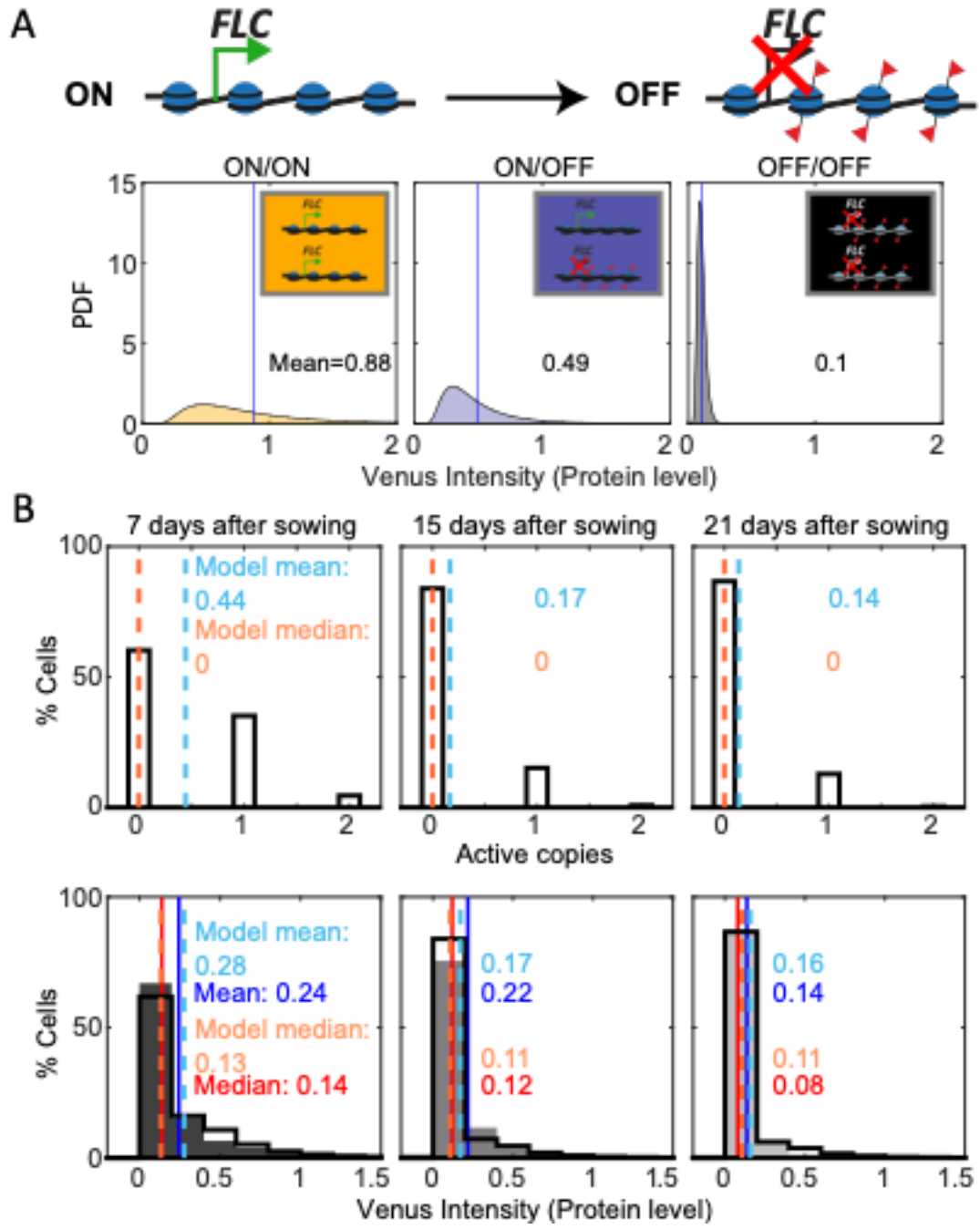


Figure 2–17 Mathematical model captures *FLC* regulation.[18]

(A) Diagram of mathematical model. Individual *FLC* gene copies can be ON or OFF, such that a cell can be in one of three states depending on the combination of ON and OFF gene copies within it (ON/ON, ON/OFF, OFF/OFF). Venus intensity (corresponding to amount of protein) within a cell was sampled from the Probability Density Function (PDF) distributions shown (described in Methods section), depending on the cell state. The means of the distributions are indicated in each panel. **(B)** Histograms of active *FLC* copies per cell (top) and Venus intensity per cell (bottom) at the indicated timepoints. Model histograms are plotted with black lines around empty bars and experimental data is shown as filled grey histograms with no outline. The means and medians of the distributions are indicated in each panel for the data (solid lines and matching colour text) and the model (dashed lines and matching colour text). Model simulated 1000 *fca*-3 cell files and histograms shown exclude bottom 4 cells of each file, as cells near the QC are also not included in the imaging.

[18] Model and Figure generated by RAK

In the plant root there is a mix of dividing and differentiated cells. Our experimental observations capture cells primarily in the division zone of the root. To avoid differences FLC-Venus imaging due to tissue-depth, cells were imaged when they were parallel to the coverslip and perpendicular to the imaging direction excluding the initials and first adjacent epidermis cells. Even within the imaging region, cell cycle times are not the same in all cells. To generate the model cell files, we used cell cycle lengths based on the literature (Rahni and Birnbaum, 2019) in line with our own few observations from Fig. 2-15.

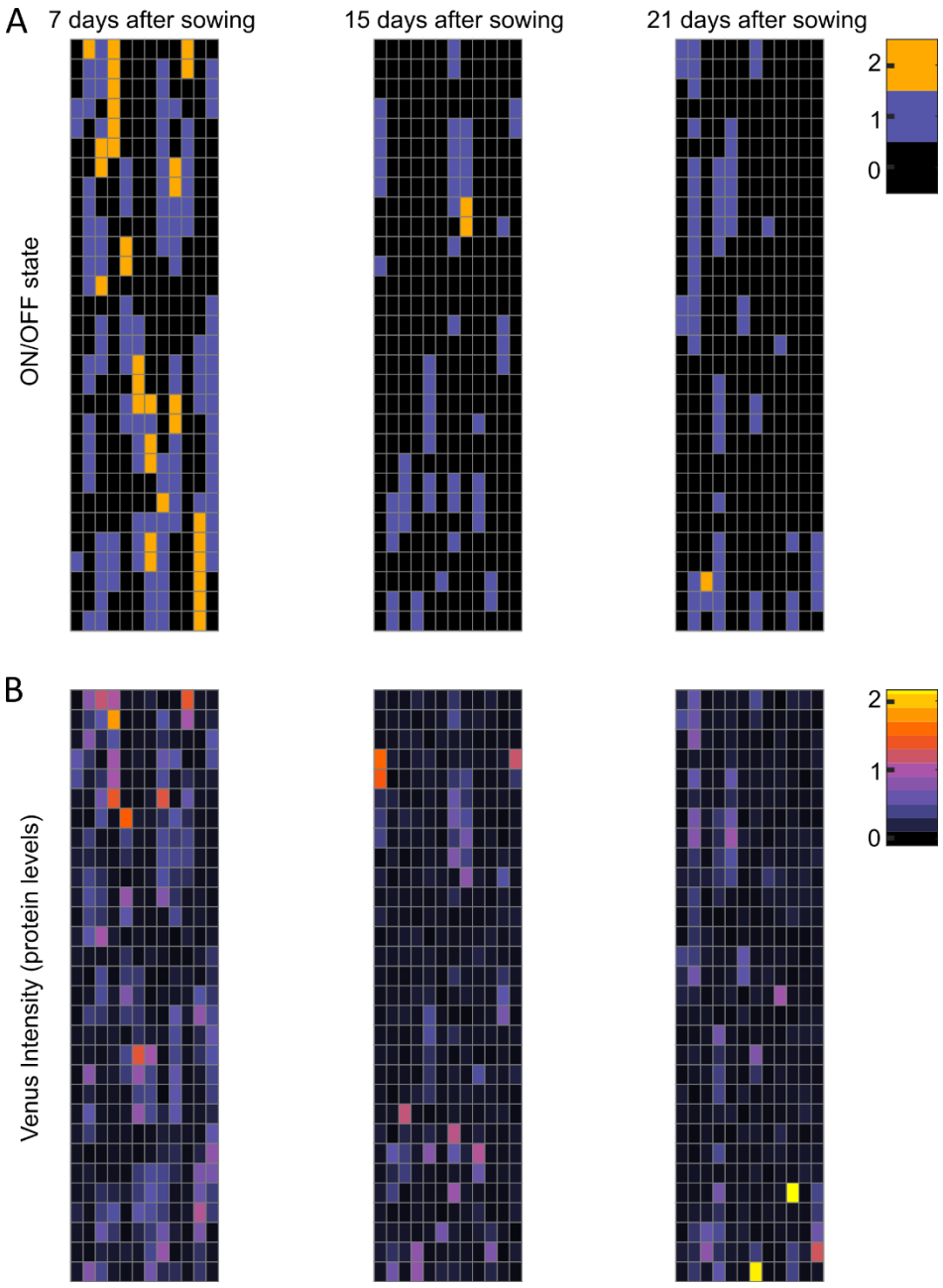
With these assumptions, our model could be fit to replicate the observed pattern of increasing OFF cells in *fca-3* roots, as well as the quantitative histograms for protein levels in *fca-3* (Fig. 2-17B, 2-18)(Details of model in Appendix 1, Chapter 8). In terms of the switching, we found the best fit where the OFF-to-ON rate is approximately 14 times slower than the ON-to-OFF rate (Fig. 2-17B), supporting predominantly one-way switching into a digitally silenced state. In addition, the model could capture the longer files present in the later timepoints in the data (21 days after sowing, Fig. 2-9C, Fig. 2-11), but unlike in the data these were not more prevalent at later timepoints than earlier. Therefore, the altered prevalence of this effect in the data may suggest additional developmental influence on the heritability of the ON/OFF state at later times in the plant. Overall, however, the model can faithfully recapitulate the developmental dynamics of *FLC* observed in *fca-3* and *fca-4*.

For Ler and *fca-1*, the switching rates can be adjusted to generate patterns of roots cells that switch off very fast/efficient (Ler) or very inefficient (*fca-1)*.

Figure 2–18 Modelling *FLC* regulation in clonal cell files in the root. [19] *(next page)*

Representative model output for *fca-3* root, showing 12 simulated cell files of length 30 cells each, at different timepoints, matching sampling times. **(A)** The state of the loci (both copies ON – orange, one copy ON – purple, both copies OFF – black). **(B)** FLC-Venus intensity (proxy for protein concentration) per cell, for the same cells.

[19] Model and Figure generated by RAK


A
7 days after sowing
15 days after sowing
21 days after sowing
ON/OFF state
2
1
0
B
Venus Intensity (protein levels)
2
1
0

Discussion

In this chapter, we have uncovered a combination of analogue and digital transcriptional regulation for the gene *FLC*: analogue regulation arises through the autonomous pathway, as illustrated by the *fca* mutants. Depending on the strength of the mutation, the functionality of the *FCA* protein is impacted, setting different analogue transcriptional states at each observed genotype. An additional component of digital switching into a heritable silenced state was observed. In the case of *fca-3* with intermediate *FLC* expression, the digital switching occurs slowly and so can be observed in our time series experiment, underlining a clear temporal separation of analogue and digital transcriptional control. However, for Ler, the switch appears to mostly occur already before our first experimental measurements at 7 days, whereas in *fca-1* the switch does not occur at all.

The silenced state is likely Polycomb dependent, given the similarity of our data to the vernalized state and the role of the Polycomb mark H3K27me3 in silencing *FLC* in wildtype Ler and Col-0 in the absence of cold treatment. In Col-0, in a mutant of the PRC2 component and H3K27 methyltransferase, *clf*, it has also been observed that *FLC* is upregulated and H3K27me3 reduced (Shu et al., 2019), though this is only a partial effect due to the presence of an active SWN, another PRC2 H3K27 methyltransferase.

The analogue autonomous pathway component of *FLC* regulation controls transcription through an *FCA*-mediated mechanism. The digital component appears to be temporally dependent on the analogue transcriptional level first set by the autonomous pathway. Based on previous studies on the interplay between transcription and PRC2 silencing, we propose that digital Polycomb silencing is prevented by high transcription in *fca-1*, while low transcription is not sufficient to significantly oppose silencing in Ler. In *fca-3*, the dynamical antagonism between Polycomb silencing and an intermediate level of transcription is still ongoing at our experimentally accessible timepoints. In this mutant, an initial transcriptionally active ON state with intermediate transcription levels controlled by an analogue mode of autonomous pathway regulation is eventually overcome and stochastically switched at each gene copy into a digital Polycomb silenced OFF state. In this way, both analogue and digital regulation are combined at *FLC*, with the timescale for switching between these modes controlled by the strength of initial analogue transcription. In *fca-3* such switching to the OFF state causes a gradual reduction in *FLC* expression at the whole plant level. Furthermore, we emphasise that there may be additional developmentally regulated processes occurring at *FLC* on top of switching with constant probability in each genotype. This possibility is underscored by the model not completely recapitulating the long cell files we observe experimentally at the 21-day timepoint. From the time-lapse imaging it is possible that an active cell cycle is required for switching, indicating that different developmental stages could have implications for switching

dynamics. More research is needed to better understand the dynamics of *FLC* switching, e.g. by using more sensitive imaging techniques such as light sheet microscopy to track *FLC* switching in real-time.

A prolonged environmental cold signal can also result in a switch from cells expressing *FLC* to non-expressing OFF cells, a process which occurs independently at each gene copy of *FLC* (Berry et al., 2015). These ON/OFF states are then maintained through cell divisions leading to epigenetic memory. This cold-dependent silencing shows only digital characteristics, with highly *FLC*-expressing cells present both before and after intermediate periods of cold, albeit in a lower proportion after. Here, we have addressed how *FLC* is regulated in the absence of cold and we show that a similar mode of regulation occurs. Furthermore, at least in the case of *fca-3* with initially intermediate transcription levels, silencing at the cell population-level also happens gradually over many cell cycles, similar to the behaviour over prolonged periods of cold. Moreover, images of roots at 21 days after sowing without cold, with fluorescently labelled FLC, are visually strikingly similar to images of vernalized roots, where in both cases we find whole files of either all ON or all OFF cells. These comparisons suggest similarities in the molecular characteristics of the Polycomb silenced state at *FLC* generated either with or without cold treatment. We hypothesise that in the absence of active Polycomb, the *fca* alleles would display only analogue differences, without the possibility of digital silencing.

The main antagonist of the autonomous pathway is the *FRIGIDA* gene, a transcriptional activator of *FLC*, which was not present in the lines used in this study. In fact, it is the antagonism between FRI and the autonomous pathway that determines the starting *FLC* levels, and so we would expect that FRI controls the analogue component directly and digital component indirectly. For future work, it would therefore be interesting to investigate transcriptional responses in a *FRI* allelic series. This could be particularly important because natural accessions show mutations in *FRI* rather than in autonomous pathway components, possibly because the autonomous components regulate many more target genes.

In this chapter, we have focused on the early regulation of *FLC* by analogue and digital pathways, rather than the effect of these pathways on flowering. However, natural variation at *FLC* largely affects expression as the plants are germinating in autumn (Hepworth et al., 2020). The analogue component of transcriptional regulation could therefore have widespread ecological relevance. For example, the high *FLC* genotype Lov-1, a natural accession from Sweden, has higher *FLC* expression upon germination in early autumn compared to Col*FRI* (Coustham et al., 2012a). The analogue component could be dominating in Lov-1, with all cells having a higher ON level. Consistent with this, Lov-1 also has lower levels of the Polycomb mark H3K27me3 at the *FLC* locus, with potentially few or no OFF cells in the warm (Qüesta et al., 2020).

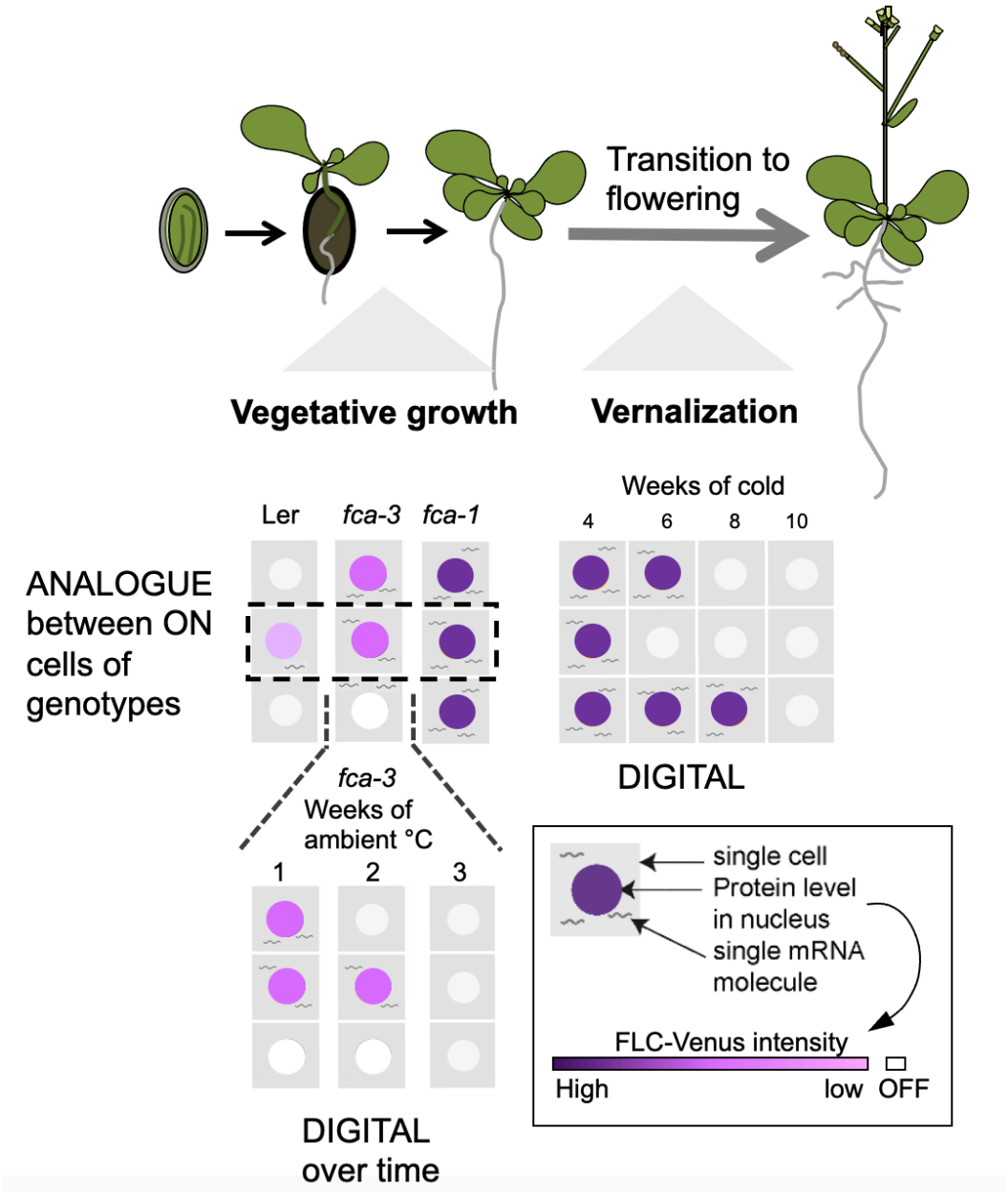


Figure 2–19 Graphical summary of findings in Chapter 2.

In this chapter, we found that a combination of analogue and digital modes of regulation are involved at Arabidopsis *FLC* before cold. Through a combination of quantitative fluorescent imaging of *FLC* mRNA and protein, as well as mathematical modelling, we uncovered the involvement of both analogue and digital modes of regulation, with a temporal separation between them where analogue precedes digital regulation. First, we have different analogue transcriptional states in the different genotypes that can switch OFF with different efficiencies depending on that state. In Ler, this occurs very fast (days), in *fca-1*, very slow (not observed). In *fca-4* and *fca-3*, it occurs gradually over weeks.

We propose that the strength of transcriptional activity, primarily influenced by the initial analogue autonomous pathway, plays a crucial role in determining the timescale for transitioning into the subsequent digital Polycomb silenced state. Importantly, we observed

that intermediate analogue *FLC* expression levels facilitated a gradual population-level digital silencing process, whereby individual gene copies exhibited stochastic and heritable switching OFF without the need for cold stimulus.

These discoveries provide a novel perspective on the gradual repression of *FLC* and shed light on the intricate integration of analogue transcriptional and digital epigenetic memory pathways. Our results highlight the complex interplay between these regulatory mechanisms and offer valuable insights into the dynamics of gene expression regulation in response to environmental cues.

CHAPTER 3 : THE COMPETENCE TO DIGITALLY SILENCE FLC IS GAINED POST-EMBRYONICALLY

Abstract (with statement of **collaboration**)

The correct timing of developmental transitions is important for a successful life cycle. One key developmental regulator is *FLOWERING LOCUS C* (*FLC*) which is silenced by prolonged cold subsequently allowing floral integrators to transition to reproductive growth.

We have previously shown (see Chapter 2) that the combination of analogue transcriptional and digital epigenetic memory pathways accounts for gradual repression in which the analogue state precedes the digital. In this chapter, I perform time course imaging during embryo development. *FLC* resetting dynamics differed between the genotypes, but at the mature embryo stage, they reached the same levels as shown by all cells being digitally ON and without detectable differences in analogue levels.

Live imaging in the germinated seedling showed differences in the analogue ON levels according to the strength of the autonomous pathway as previously assessed in the *fca* allelic mutations. At the same time point, we also observed the first OFF cells indicating a digital response in Ler and *fca-3*, suggesting that the competence to digitally switch-off also only occurs between the mature embryo and after germination.

The competence to switch OFF correlates with the upregulation of *FLC*'s antisense transcript, the lncRNA *COOLAIR*. Yet, later developmental *COOLAR* levels do not show a straightforward connection between both.

These findings imply three temporally separate well-defined developmental states at *FLC*: Firstly, *FLC* transcription – once fully reset – is ON in all cells in the mature embryo; secondly, analogue levels are set in a time window between the mature embryo and after germination; and thirdly, the competence to digitally silence *FLC* is then gained depending on the analogue levels. This leads to digital switching occurring either efficiently during early vegetative development (Ler, Col-0) or requires another environmental stimulus by vernalisation (ColFRI, *fca-1*). In the intermediate mutant *fca-3*, we observe gradual switching.

Using a mutant in the PRC2 methyltransferase subunit *clf-2*, we show that the digital component is indeed dependent on PRC2 (experiments performed in collaboration with **Eduardo Mateo-Bonmati**).

This research highlights the importance of the correct timing in the regulation of PRC2-targets.

Introduction

The well-characterized flowering repressor and Polycomb-target gene, *FLOWERING LOCUS C*, is epigenetically silenced by prolonged cold (vernalization) during its vegetative growth period to align flowering with spring. Quantitative variation (different levels) of *FLC* expression influences if there is a vernalization requirement (Col*FRI*) or if the plant has a rapid cycling lifestyle (like the accessions Col-0 or Ler). These levels are determined by an interplay of the activating FRIGIDA (FRI, part of the COMPASS-like complex) and the repressing autonomous pathway (RNA binding proteins, 3′ end processors, chromatin modifiers) setting the chromatin environment at *FLC* (Wu et al., 2019b). It has been suggested that the establishment of the initial basal *FLC* expression levels occurs during early development in the embryo (Schon et al., 2021) and is maintained during further development unless environmentally silenced by cold.

However, as we have seen, both these cases, with high (Col*FRI*) or low (Ler, Col-0) *FLC* expression, are extreme examples, where essentially all *FLC* loci are either expressed to a high level or not, before cold. It is therefore difficult to use these genotypes to understand any possible interplay between graded (analogue) or all-or-nothing (digital) control, as only the extremes are exhibited. Instead, what is required is a genotype with intermediate *FLC* expression at a cell population level. Such a genotype at a single gene copy level might exhibit either a digital or analogue *FLC* expression before cold, thereby allowing dissection of whether analogue or digital regulation is at work.

One main autonomous pathway component, the RNA binding factor FLOWERING CONTROL LOCUS A (*FCA*), promotes proximal polyadenylation of/at many gene targets. It acts on *FLC* sense and antisense, the lncRNA *COOLAIR*, at different developmental times. Additionally, *FCA* has been shown to untangle RNA-DNA hybrids and Pol II stalls in which transcription termination is required to be resolved prior to replication (Baxter et al., 2021; Xu et al., 2021).

Previously, we exploited three mutant alleles of *FCA* thereby systematically varying overall *FLC* levels. One of these mutants, *fca-1* (Koornneef et al., 1991), is a complete loss of function and therefore exhibits late flowering and high *FLC* expression before cold, similar to Col*FRI*. The wildtype, L*er*, has a fully functional autonomous pathway and so exhibits low *FLC* levels before cold and early flowering, with the *FLC* gene covered by the silencing histone mark H3K27me3 (Wu et al., 2019b). Crucially, however, the *fca-3* and *fca-4* mutants (Koornneef et al., 1991) display intermediate cell population-level *FLC* expression and flowering time. Those quantitative differences in *FLC* levels allowed us to systematically dissect the interplay between analogue and digital regulation at *FLC* in Chapter 2, leading to the conclusion that the analogue mode precedes the digital mode (Antoniou-Kourounioti et al., 2023). PRC2 is a protein complex that deposits repressive histone modifications (H3K27me3) on *FLC*, making DNA less accessible for transcription machinery, resulting in an OFF switch. On the other hand, if the locus is highly transcribed, PRC2 cannot bind to the *FLC* region and fails to repress it (Berry et al., 2017). Here, we propose that the strength

of the initial analogue transcriptional level controls the timescale for the switch into the subsequent digital, Polycomb silenced state.

These results prompted us to look at earlier stages in development to see if we can (1) detect analogue levels without digital impact (all cells ON with analogue differences between genotypes) and (2) determine a time point of the digital shutdown (by detecting OFF cells) in *fca-3*.

FLC resetting is independent of FRI

To assess purely analogue levels, we need to consider that vernalized *FLC* needs to be reactivated to allow an adequate comparison between genotypes. For *FLC* reactivation, the LEC1 transcription factor is required to bind to *FLC* approximately 24 h after fertilization. This changes the chromatin environment to be accessible for LEC2 and FUS3 at the RY motifs, thereby allowing transcription (Tian et al., 2019). Preliminary data of FLC-GUS (personal communication with Claudia Köhler) in vernalized Ler and Ler*FRI* showed that reactivation had occurred at the globular stage in both genotypes, and hence is independent of FRI.

Role of PRC2 in the embryo

Whereas it is well known that PRC2 is essential for normal development during mammalian embryogenesis (Meng et al., 2020; O'Carroll et al., 2001), its essentiality is not that clear during embryo development in Arabidopsis. The embryo-specific FIS(MEA)-PRC2 is required for endosperm development (nutritive tissue that supports embryo growth) by coordinating the expression of imprinted genes (genes that are expressed predominantly from one parental allele)(Grossniklaus et al., 1998). In the seedling, all embryonic maternally or paternally imprinted genes are expressed from both parental copies, suggesting that the imprint is erased during late embryogenesis or early vegetative development (Raissig et al., 2013).

Allele-specific gene expression data has shown that *FLC* is maternally imprinted (Choi et al., 2009; Sheldon et al., 2008). In rapid-cycling genotypes (no functional FRI), all core PRC2 units are expressed throughout embryo development, peaking at the heart stage (Spillane et al., 2007 (Ler), Zhao et al., 2019 (Col-0), Belmonte et al., 2013 (WS-0), Hofmann et al., 2019 (Col-0)), including the VRN2 subunit known to be involved in silencing *FLC* during vernalization. There is little known about PRC2 binding at *FLC* in the embryo, due to the high amount of plant material required for ChIP analysis. One study has shown that CLF PRC2 is present at the *FLC* nucleation region in embryo tissue 2 days after pollination (DAP) (Luo et al., 2020). At 6DAP, *FLC* is reset and activated according to trans factors such as LEC1, FRI, *FCA* and independent of pre-generational vernalization.

In many cases, mutations in major PRC2 components still produced viable embryos and established a normal cellularization/ body plan, as summarized in Table 3-1.

Table 3–1 PRC2 mutants in the Arabidopsis embryo.

Mutant	Phenotype	Normal embryo	K27me3	Later development	Reference
fie +/+ (endosperm FIE in *cdk1*) fertilisation independent endosperm required for all PRC2 compositions	If the endosperm has FIE, the embryo develops normally. Aberrant organ formation in later development	YES	NO	Deficiency in embryo to seedling transition	(Bouyer et al., 2011)
clf/swn	Callus-like structures rather than leaf-like structures	YES	NO	Deficiency in embryo to seedling transition	(Bouyer et al., 2011)
emf2/vrn2	Callus-like structures rather than leaf-like structures	YES	NO	Deficiency in embryo to seedling transition	(Bouyer et al., 2011)
fis	Autonomous endosperm formation without fertilization. Maternal *fis* mutants are lethal due to endosperm requirement	NO	-	paternally expressed imprinted genes, mis-regulation in endosperm	(Guitton et al., 2004; Hsieh et al.; Kiyosue et al., 1999; Wolff et al., 2011)
mea +/+ (endosperm MEA in *kpl*)	The effect on early embryo patterning and cell lineage determination was suggested by impacting cell cycle progression	NO	NO		(Simonini et al., 2021)

The fact that in most cases mutations in PRC2-components did not influence the development of the embryo could be partly explained by redundancy of the core PRC2 units compensating for the respective loss. However, this cannot be the case in a *fie* mutant, which is an essential component present in all PRC2 complexes, but which still showed no effect on embryonic patterning (Bouyer et al., 2011). While it is challenging to draw a

conclusion about the necessity of PRC2 activity in the embryo based on these mixed phenotypes, there does seem to be a direct effect of lack of core PRC2 mutants on failure to stably establish the embryo-to-seedling phase transition to end embryogenesis.

To investigate if we can observe an analogue effect on *FLC* during embryo development lacking the digital OFF cells, I performed live-imaging in the embryo (6 DAP and onwards).

Results

3.1 Digital effect already visible in 4-day-old roots

Previously in Chapter 2, using an FLC-Venus reporter line, we observed analogue differences in fluorescence intensity per cell between genotypes, as well as patterns of ON/OFF cells in the intermediate *fca-3* and *fca-4* genotypes, suggesting heritability of the transcriptional state. Strikingly, we observed a decline in the fraction of ON cells rather than in the fluorescence intensity of protein signal per ON cell over time when comparing 7- 15- and 21-day old Arabidopsis roots. Here, we look at 4-day old roots (~1-day after visible seed coat rupture) of several genotypes with fully functional *FCA* (Col-0 and Ler), and hence low *FLC* expression, intermediate *FLC* expression (*fca-3* and *fca-4*), and the high expressing *fca-1*. Interestingly, the epidermal layer of wildtype Ler and Col-0 showed some ON cells, which was observed only very rarely in 7-day old roots (see Chapter 2, Figure 2-6). In addition, in *fca-3* and *fca-4*, most cells showed ON signal. Taken together, 4-day old roots showed a similar pattern of analogue/digital FLC-Venus signal as compared to later time points, with the difference that in *fca-3* we found a higher fraction of ON cells compared to later time points (Figure 3-1).

Overall, those results indicate that earlier time points need to be investigated to potentially separate the timings of analogue and digital regulation. For that reason, we decided to look at different stages in embryo development.

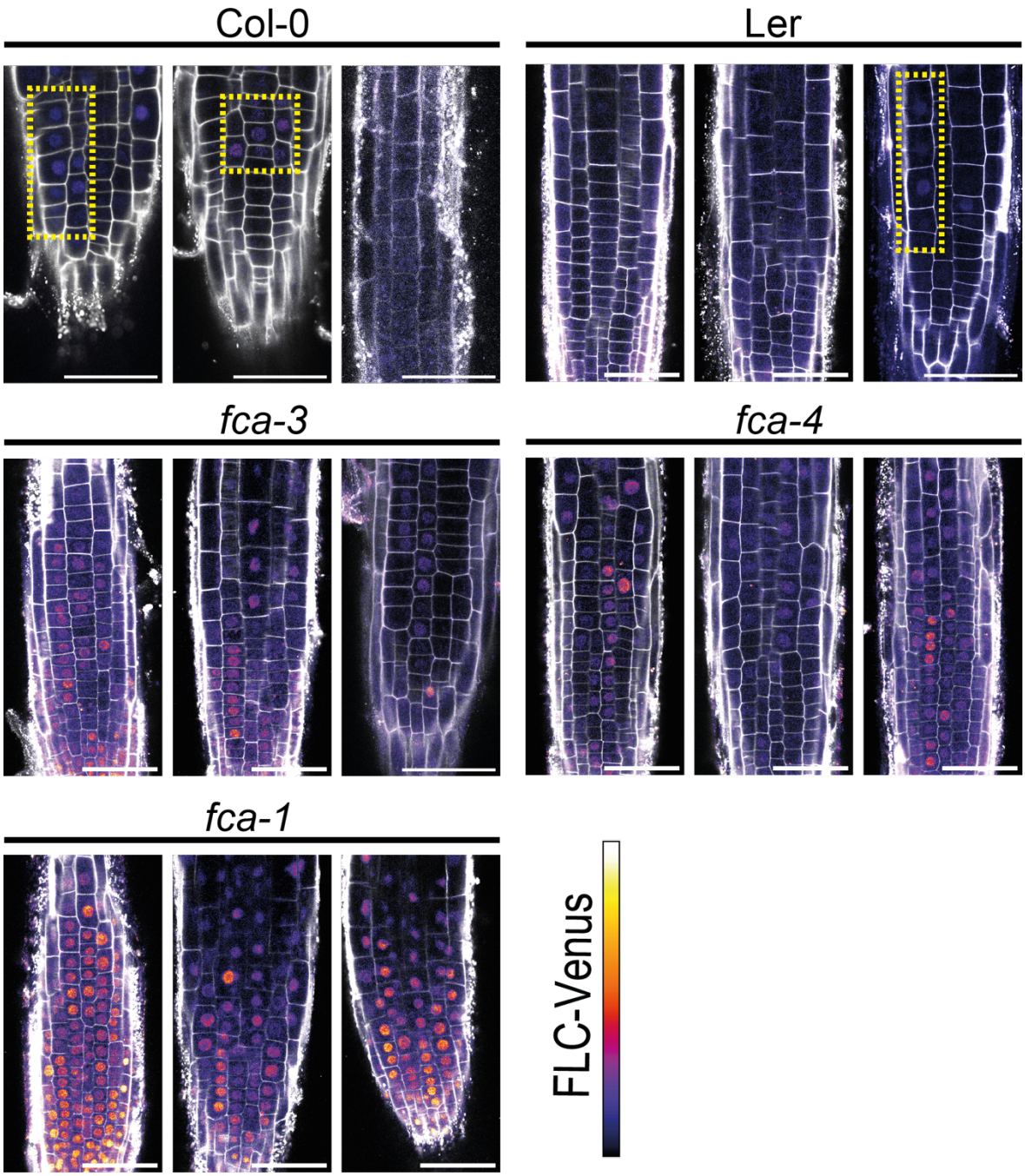


Figure 3–1 More ON cells during earlier imaging time points suggest gradual switching.

Three representative images of FLC-Venus live-imaging in 4-day old roots in genotypes with differences in analogue transcriptional levels. Wildtype, low FLC levels: Col-0, Ler (top); intermediate: *fca-3* and *fca-4* (middle); high: *fca-1* (bottom). FLC-Venus intensity is indicated by colour map (back/blue low - yellow high), PI-stained cell wall (grey). Yellow dashed boxes indicate ON cells in low FLC expressing wildtype roots. Scale bar is 50 μm., n=3 biological replicates with 10 roots.

3.2 *FLC* resetting occurred by the early heart stage but in mature embryos is independent of *FCA*

Live-imaging of embryos at early heart stage of the three genotypes (Ler, *fca-3*, *fca-1*) showed that all cells were ON (Fig. 3-2A), suggesting that resetting has occurred by the early heart stage. All ON cells were also observed in Col-0 at this stage.

In the seedling, we previously observed intrinsically different fluorescence intensity levels of FLC-Venus in ON cells between the genotypes Ler, *fca-3* and *fca-1* (Chapter 2). Similarly, at the early heart stage, although all cells were ON, they were again not at the same FLC-Venus intensity. In particular, in *fca-3* and *fca-1*, there appeared to be some cells that are brighter than the average cell of Ler. Quantification of the fluorescence intensity per genotype at heart stage revealed that *fca-3* and *fca-1* exhibit similar FLC-Venus fluorescence intensity values (mean ± SEM, *fca-3:* 27.3 ± 2.1, *fca-1*: 24.9 ± 1.4), whereas Ler is significantly lower (Ler: 10.7 ± 1.2, Figure 3-2B). Furthermore, the distribution of ON levels in *fca-3* and *fca-1* seemed to have two subpopulations of cells, whereas only one population, exhibiting lower FLC levels, was detected in Ler cells. This indicates that *FLC* resetting does not occur at the same speed or to the same level in all genotypes. Note, extra care was taken to only use embryos that did not experience previous vernalization treatment for direct comparability of resetting of *FLC* transcriptional levels.

However, in mature embryos, the fluorescence intensity values between all observed genotypes were at the same level (Ler: 7.8 ± 0.3 *fca-3:* 7.2 ± 0.3 *fca-1:* 8.0 ± 0.4) without detectable differences in analogue levels. (Figure 3-2C). We suggest that this is due to additional embryo-specific factors regulating further upregulating *FLC* or *FCA* in later embryo stages. ABSCISIC ACID-INSENSITIVE 3 (ABI3) and ABI5 positively regulate *FLC* expression during embryo maturation by binding to a ABA-responsive cis-element in the *FLC* promoter region (Xu et al., 2022), generating an overall higher transcriptional state in the embryo and thereby potentially masking analogue differences. The lack of OFF cells suggests that there is no digital OFF switching in the embryo which might be due to the high transcriptional levels opposing the digital switch-off. Further, the comparable levels of FLC-Venus indicate that *FCA* is not affecting the FLC-Venus levels in mature embryos. Rather, *FCA* is relevant for the subsequent analogue set-up of the transcriptional state at *FLC* at a later developmental stage (see below). As all cells were ON, both analogue and digital differences could only be observed upon or post germination as seen in the 4-day old root (Fig. 3-1). Due to the digital switch-off depending on analogue levels, this set-up results in the very efficient off-switching observed in Ler one day after germination. Consequently, those two processes are challenging to temporally separate, and appear to happen almost simultaneously when analogue transcription levels are low.

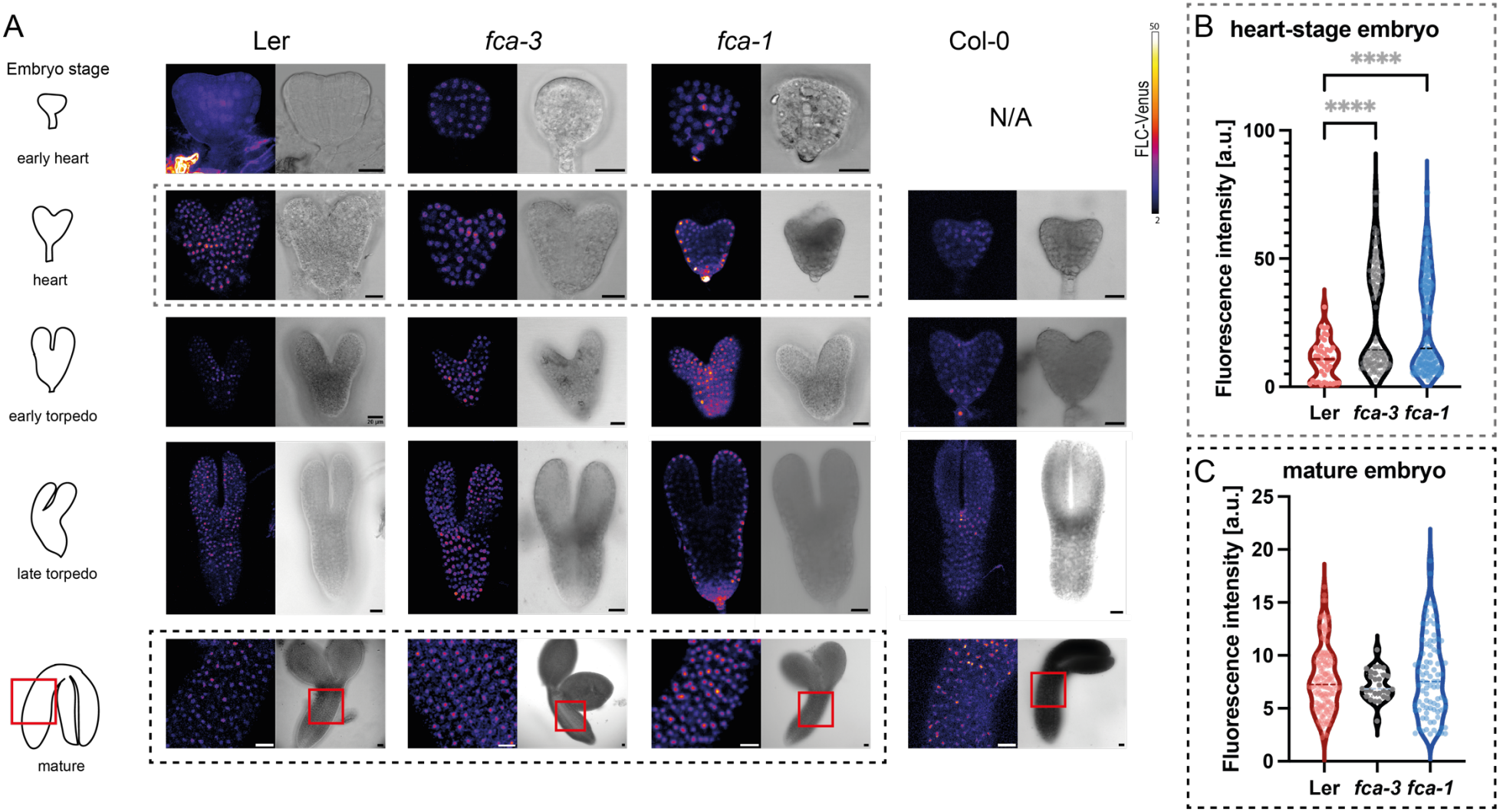

A
Ler
fca-3
fca-1
Col-0
Embryo stage
early heart
heart
early torpedo
late torpedo
mature
N/A
FLC-Venus
50
2
B heart-stage embryo
****
****
Fluorescence intensity [a.u.]
100
50
0
Ler
fca-3
fca-1
C mature embryo
Fluorescence intensity [a.u.]
25
20
15
10
5
0
Ler
fca-3
fca-1

Figure 3–2 FLC-Venus levels in Ler, *fca-3* and *fca-1* are not different in mature embryos. *(previous page)*

(A) Dissected embryos in different developmental stages from non-vernalized parents. FLC-Venus intensity is indicated by colour map, PI-stained cell wall (grey). Approximate days after pollination (DAP) relative to embryo developmental stage: Early heart: 5 DAP, Heart: 6 DAP, Early torpedo: 7-10 DAP, Torpedo: 10-12 DAP, mature: 12+ DAP. **(B-C)** Quantification of the fluorescence intensity values per nucleus of the outer cell layer of the embryo in the **(B)** heart-stage embryo and **(C)** mature embryo. Red box indicates region used for image quantification. One way ANOVA was performed to compare variability of means with p-values of Ler vs. *fca-3* and Ler vs. *fca-1* of <0.0001 in (B). Scale bar represents 20 μm, n = 5 biological replicates with >20 embryos. For quantification, 5 cells of the outer layer of each embryo were used that were in the same imaging plane between images.

While the molecular mechanism of resetting is not fully understood, the shift in abundance of early terminated *FLC* transcript to full-length transcripts has been observed at the heart stage embryo (Schon et al., 2021), resulting in the observed increase in FLC-Venus protein levels in cells. A related process has been observed at *FLC*'s antisense, the lncRNA *COOLAIR*, during early cold (Swiezewski et al., 2009). Coinciding with early transcriptional downregulation of *FLC* is the upregulation of total *COOLAIR* with changes in isoform abundances as proximal poly(A) becomes more prominent than distal Poly(A). *FCA* promotes proximal polyadenylation of sense and antisense of *FLC* with implications of the choice of poly(A) site on the chromatin environment and *FLC's* regulation. Moreover, this co-transcriptional coupling of *COOLAIR* and *FLC* processing occurs at developmentally distinct times.

To check if we can observe differences in the embryo in transcript abundance of *FLC-Venus* mRNA and *COOLAIR*, smFISH was performed in mature embryos in all genotypes (Fig. 3-3). In line with the protein observations, all cells express *FLC* mRNA further supporting our previous assumption that the digital OFF state (no transcription) does not exists in the embryo yet. Although, I could detect *FLC-Venus* mRNA in all genotypes and in all cells (white arrows pointing to white/grey foci, the background autofluorescence was so strong that quantification was not possible.

Total *COOLAIR* smFISH signal was detected in only a few cells in the mature embryo of most genotypes (Fig. 3-3, cells with *COOLAIR* signal were selected for representation in the figure). Surprisingly, in *fca-3,* no *COOLAIR* signal was detected in the embryo. The other intermediate genotype *fca-4* showed *COOLAIR* levels comparable to Ler and *fca-1*, suggesting that the effect on *COOLAIR* is *fca-3*-specific.

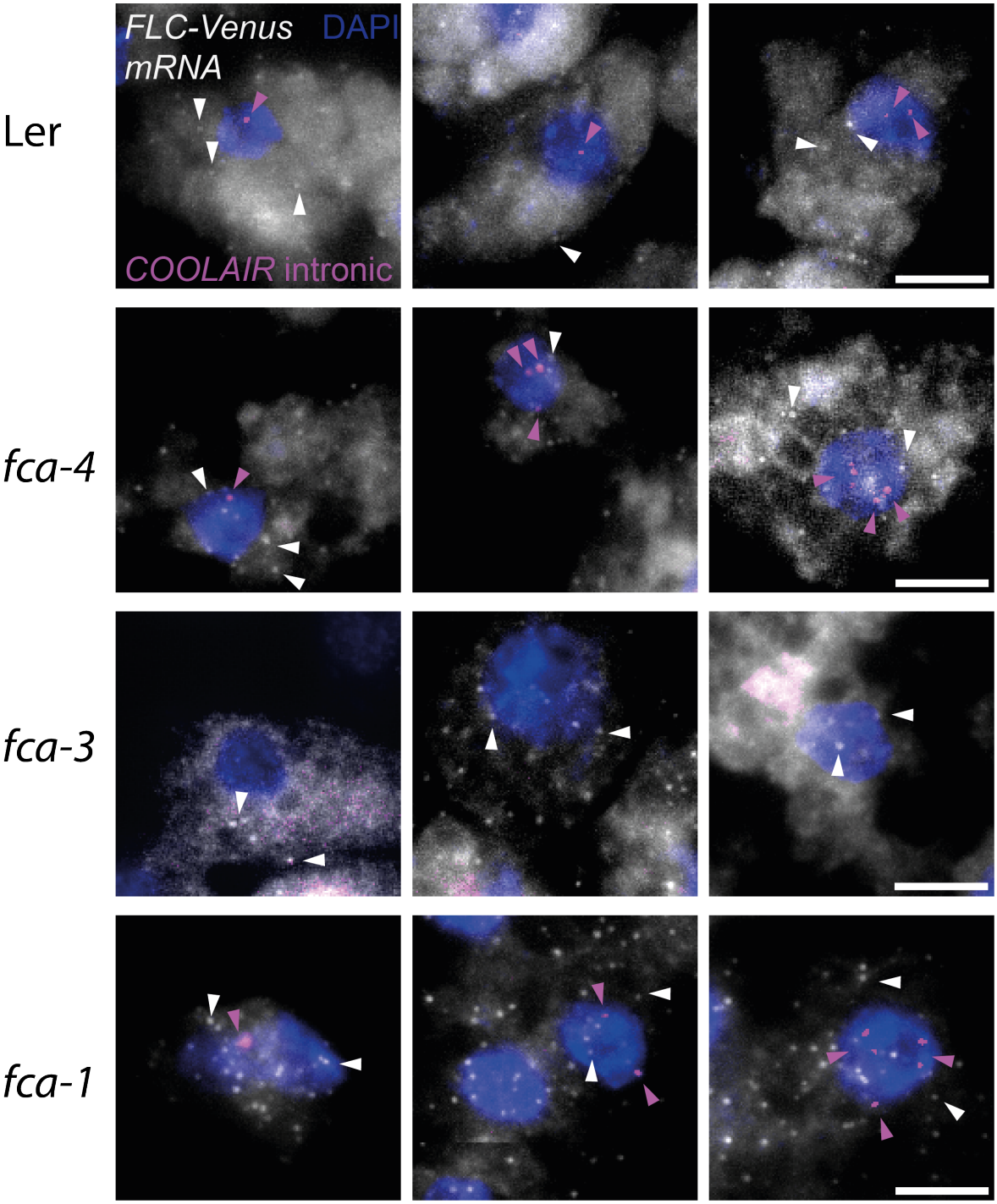


Figure 3–3 CaViAR (*COOLAIR*: visualization of Antisense RNA in the embryo).

Three representative single-cell images of *FLC* and *COOLAIR* RNA in the mature embryo. Ler, *fca-4*, *fca-3* and *fca-1* single cells in squashed mature embryo tissue. White arrows highlight *FLC-Venus* mRNA, purple arrows highlight *COOLAIR* nascent signal (magenta, intronic probes), the nucleus is DAPI-stained (blue). n = 2 biological replicates with 4 embryos imaged per replicate, scale bar is 10 μm.

To further test if the lack of *COOLAIR* signal is *fca-3*-specific and to potentially investigate the role of antisense transcription, we examined the levels of various *COOLAIR* isoforms in *fca* mutants in seedlings over time (Fig. 3-4). Interestingly, in the seedling *fca-3* showed intermediate *COOLAIR* levels between Ler and *fca-1*. Distal *COOLAIR* correlates with an actively transcribed *FLC* locus, while the proximal poly(A) is associated with a chromatin

environment of slowly transcribing polymerase II, and a transition to silenced *FLC*. Changes in those ratios (e.g., increased proximal versus distal antisense transcription) could indicate a change in the chromatin environment coinciding with the different isoforms of *COOLAIR*. Moreover, *FCA* affects the polyadenylation site of *COOLAIR* so we wondered if mutations in *fca-3* and *fca-1* would impact *FLC* transcription via antisense poly(A) site choice. At all timepoints, *fca-3* shows intermediate *COOLAIR* levels, while Ler shows the lowest and *fca-1* the highest for all isoforms, reflecting levels of sense transcription. Rather than a shift in poly(A) sites in the mutants, we observed a gradual increase of proximal and total *COOLAIR* in *fca-1* while maintaining constant levels in Ler and *fca-3* (Fig. 3-4A,B). On the other hand, levels of distal *COOLAIR* appeared to reflect the chromatin environment and sense transcription (Fig. 3-4C). In particular, *COOLAIR* is present in *fca-3* in seedlings despite its absence in the mature embryo.

On top of that, the increase of total and proximal *COOLAIR* in *fca-1* is accompanied by an overall reduction of *FLC* spliced and unspliced transcript over time but is not sufficient to fully silence *FLC*, as shown by our earlier FLC-Venus time course imaging (see *FLC* transcription levels in Chapter 2, Fig 2-6).

To assess if there is a shift in the proportion of spliced variants between proximal and distal *COOLAIR*, the ratio between both was assessed (Fig. 3-4D), showing that there is relatively higher proximal termination in Ler, correlating with functional *FCA*.

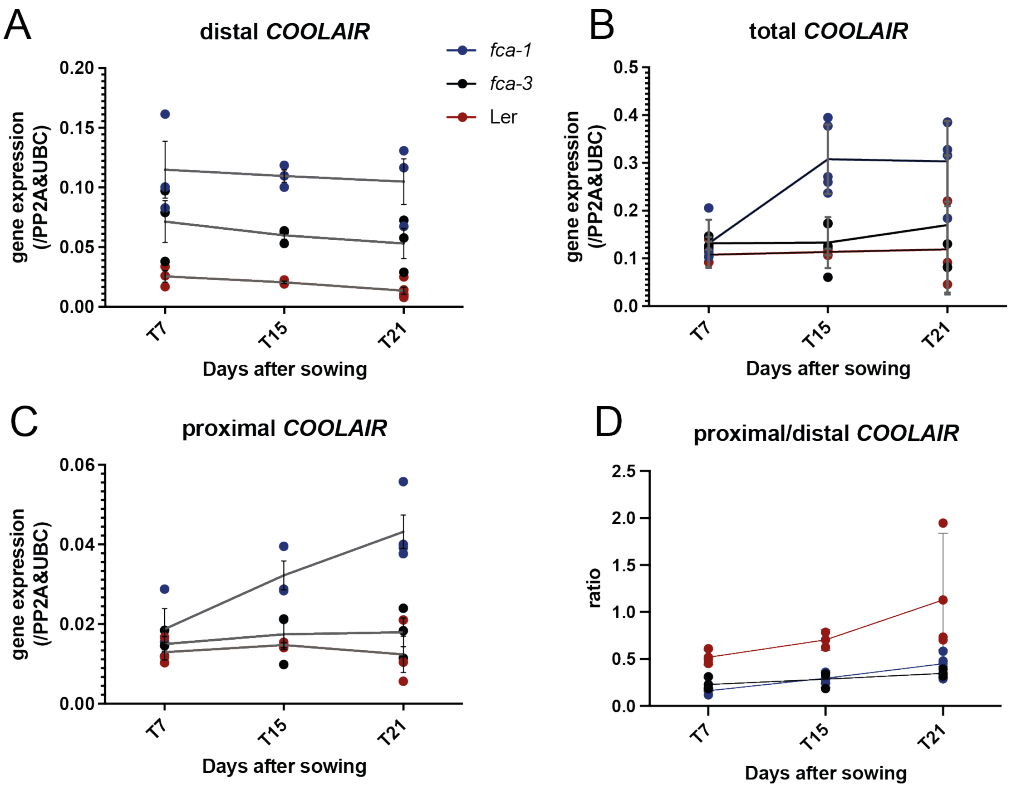


Figure 3–4 Distal *COOLAIR* reflects *FLC* transcription.

Gene expression analysis time course of *COOLAIR* transcription in seedlings in the L*er* (red) genotype and *fca* mutants (*fca-3*: dark grey, *fca-1:* blue) **(A)** distal *COOLAIR*, **(B)** total *COOLAIR*, **(C)**

proximal *COOLAIR*, **(D)** ratio of proximal to distal *COOLAIR*. Error bars represent ± SD, n=3 biological replicates. Note, corresponding *FLC* mRNA levels can be found in Figure 2-10.

Overall, this suggests that *COOLAIR* levels mostly reflect the chromatin environment and sense transcription, and that *FCA* functionality in *fca-3* and *fca-1* impacts but does not impair proximal poly(A) choice, perhaps through redundancy of other autonomous pathway components.

3.3 Chromatin environment reflects transcriptional state

To investigate if we can observe changes in the chromatin environment of *FLC*, we looked at differences in histone marks associated with the repressive, Polycomb-silenced state (H3K27me3) and with active transcription (H3K4me1 and H3K36me3), as assessed by Chromatin Immuno-precipitation over time (Fig. 3-5). Histone marks were determined in seedlings at 7, 14 and 21-day-old plants. At all timepoints, *fca-3* shows intermediate levels of H3K4me1, H3K36me3 and H3K27me3, while Ler shows the lowest H3K4me1/H3K36me3 levels with the highest H3K27me3 levels, and *fca-1* the highest H3K4me1/H3K36me3 and lowest H3K27me3 levels.

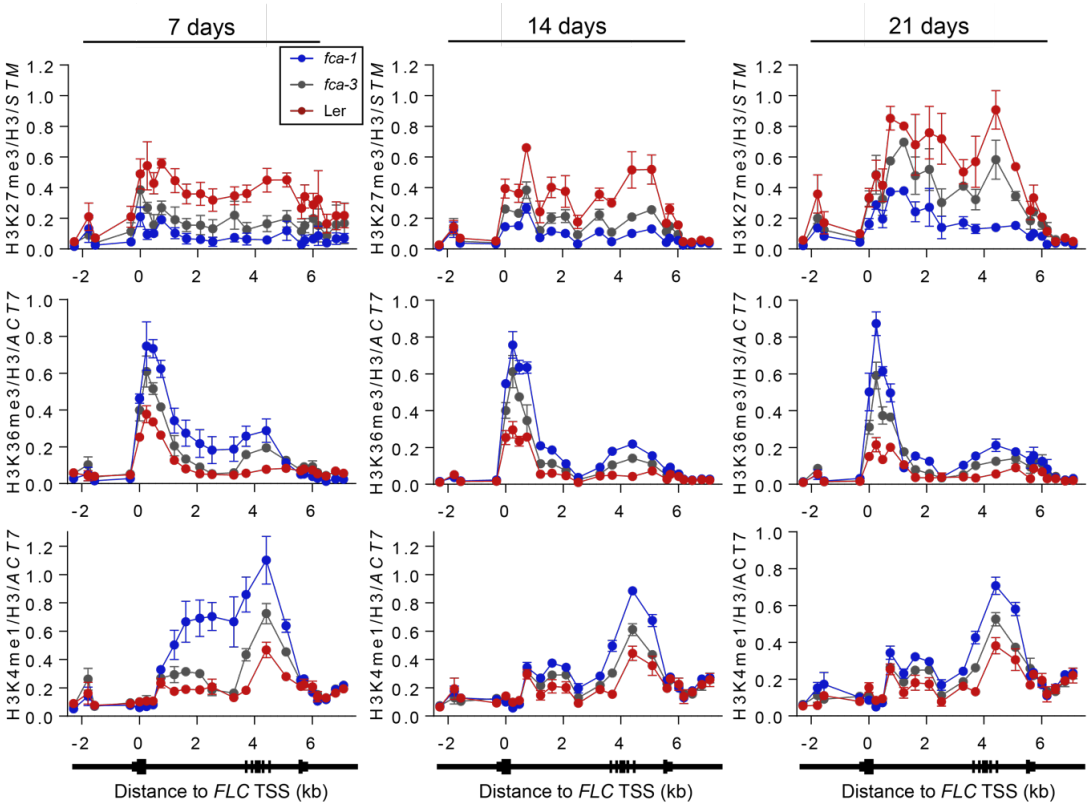


Figure 3–5 Chromatin environment reflects transcriptional state. [20]
ChIP-qPCR time course measurements of H3K27me3 (top), H3K36me3 (middle) and H3K4me1 (bottom) across the *FLC* locus in the L*er* (red) and *fca* mutants (*fca-3*: dark grey, *fca-1:* blue). Error bars represent mean ± s.e.m, n=3 biological replicates. For exact primer pair binding sites, see Table 6-2 in the Methods section.

[20] ChIP experiment: 3 replicates performed by EMB, one replicate of H3K27me3 by SvR

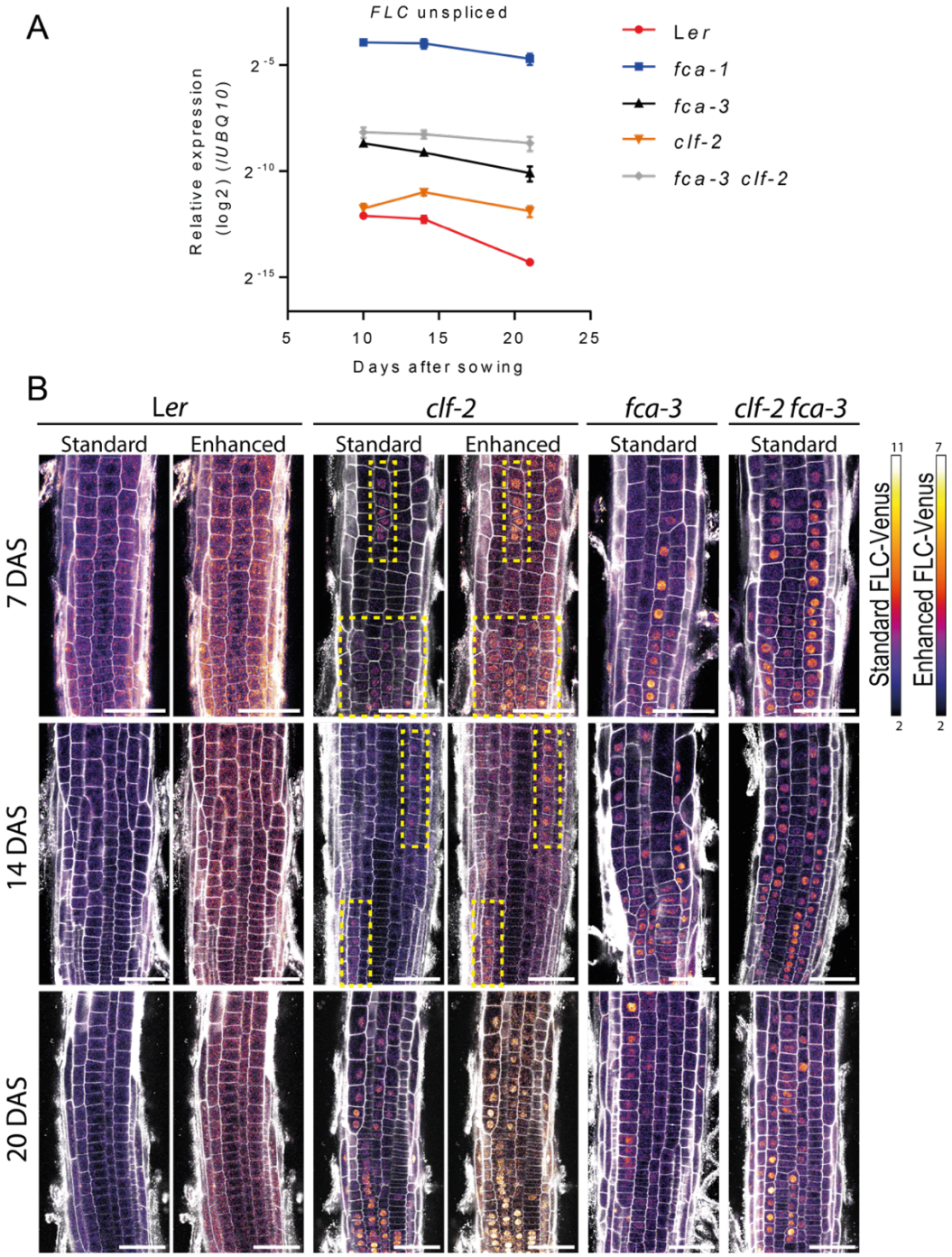


Figure 3–6 FLC-Venus time-course imaging in the different genetic backgrounds show steady state of ON cells in *clf*-2 mutants.[21]

(A) Time-course of unspliced *FLC* expression 10, 14 and 21 days after sowing (DAS), in L*er*, the mutant *clf-2*, the mutants *fca-3* and *fca-1*, and the double mutant *fca-3 clf-2* (crosses of respective lines). Error bars represent mean ± s.e.m. (n=3 biological replicates). **(B)** Representative confocal images of FLC-Venus signal at three timepoints (7, 14 and 20 DAS) in the epidermis of the root meristem of Ler, single mutants *clf-2*, and *fca-3*, and the double mutant *fca-3 clf-2*. Grey shows the cell wall dye propidium iodide; FLC-Venus intensity is indicated by colour maps for standard and enhanced images, respectively. The same settings were used for imaging of all time points. The same image presentation is used for standard and enhanced images. Brightness and contrast of FLC-Venus signal were increased to generate the enhanced images (Note different scale on colour maps). Yellow boxes in *clf-2* highlight ON cell files at 7/ 14 DAS Scale bar, 50 μm, n = 2 biological replicates.

[21] Gene expression analysis (A) was performed by EMB

Consistent with an increased fraction of OFF cells observed over time (Chapter 2), the chromatin state of all genotypes showed increased H3K27me3 levels over time while at the same time reduced H3K4me1 levels between 7-day-old and 14-day-old plants. These differences in the chromatin environment of the genotypes suggest an epigenetic component in the regulation of *FLC*.

3.4 CLF is required for stable digital shut-off

Previous genetic analysis has shown that the Polycomb methyltransferase CURLY LEAF (CLF) involved in the maintenance of repressive H3K27me3 is essential for *FCA*-mediated silencing of *FLC* (Fang et al., 2020). Here, by using an impaired Polycomb silencing mutant, *clf-2*, we wanted to investigate if we can observe a steady-state in the switching in *fca-3* (Fig. 3-6). If the OFF switching is mediated by PRC2, the *FLC* levels in *fca-3 clf-2* either decrease slower or stay at a similar level over time as non-stably silenced cells reactivate. The *clf-2* mutant was crossed into *fca-3* FLC-Venus. Gene expression analysis of *clf-2* and the *clf-2 fca-3* double mutant shows that while unspliced *FLC* is reducing in Ler, *fca-3* and *fca-1*, the *clf-2* mutants reduce at a slower rate (Fig. 3-6A). This strongly suggests that PRC2-silencing is also involved in the stable silencing of *FLC* in ambient temperatures.

We additionally complemented the gene expression analysis with time course live-imaging in the root (Fig. 3-6B), showing that indeed in *clf-2* mutants, some roots show no stable switching off. Moreover, in particular in the later time points, more ON cells were observed in the double mutant *fca-3 clf-2* compared to *fca-3* suggesting a non-stable switch OFF. Overall, there was a huge variability between roots in both morphology and FLC-Venus levels (Fig. 3-7) most likely due to the detrimental effects of Polycomb mutants on other targets.

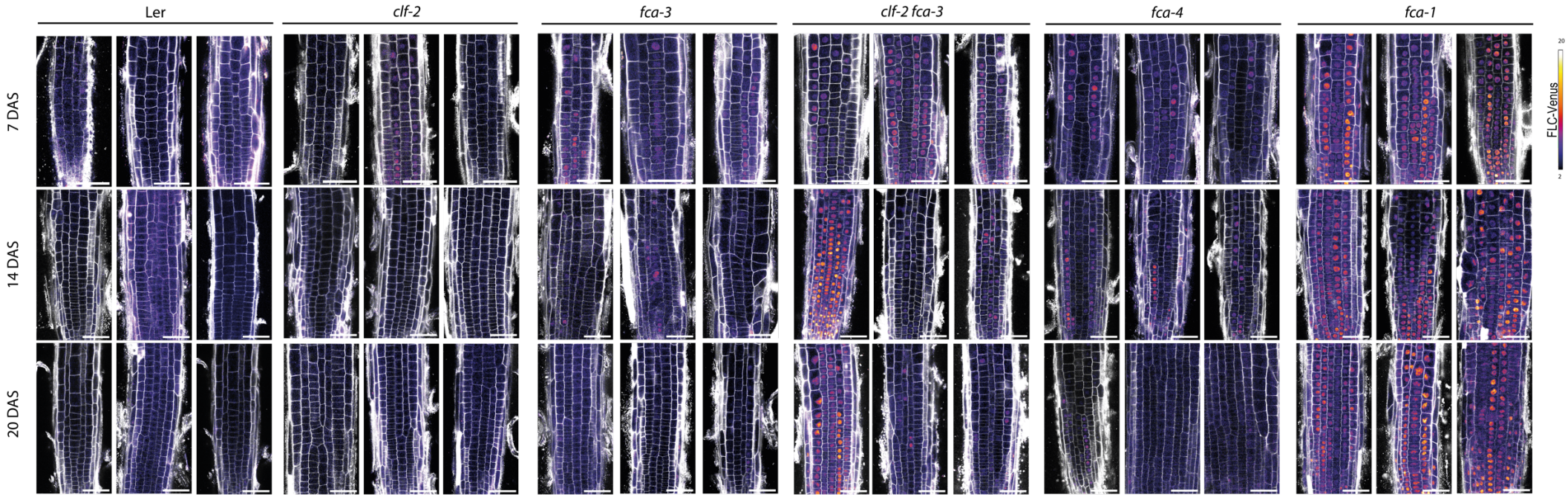


Figure 3–7 Replicates of FLC-Venus time-course imaging in the different genetic backgrounds.

Representative confocal images of FLC-Venus signal at three timepoints in the epidermis of the root meristem. Grey shows the cell wall dye propidium iodide; FLC-Venus intensity is indicated by colour map on right. The same settings were used for imaging of all time points. The same image presentation is used for all images. Scale bar, 50 μm, n=2 biological replicates with 10 roots imaged each replicate.

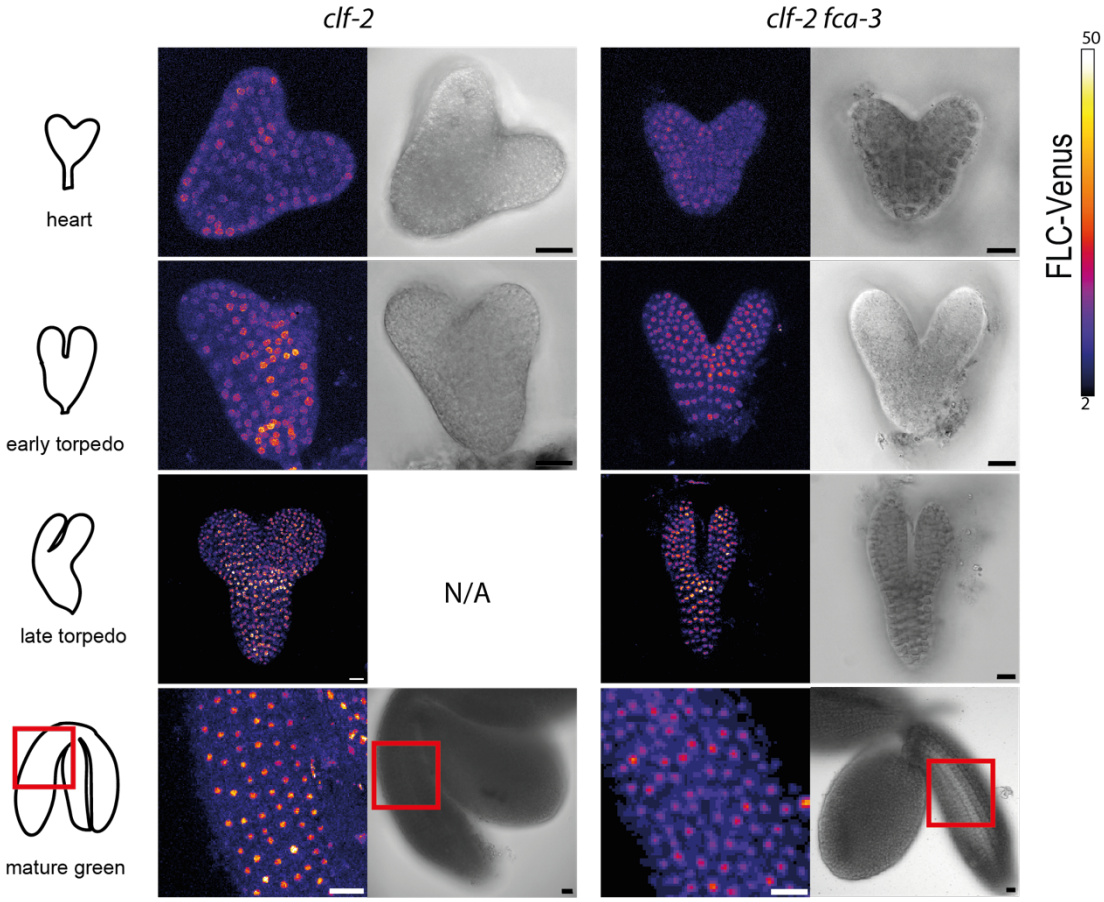


Figure 3–8 Embryo imaging in *clf-2* and the *clf-2 fca-3* double mutant show that all cells are ON at the mature embryo stage.

Live-imaging of FLC-Venus (left) and Brightfield (right) of the embryo. The red box on the brightfield image indicates the selection of cells shown for FLC-Venus intensity. FLC-Venus intensity is indicated by the colour map on right. Scale bar represents 20 μm, n=3 biological replicates with >5 embryos.

For completeness, I performed embryo imaging of *clf-2* and the *clf-2 fca-3* double mutant (Fig. 3-8) to confirm the timing of digital resetting and PRC2's involvement in subsequent switching. Similar to previously shown embryo imaging in Ler, *fca-3* and *fca-1*, all cells showed an ON signal at the mature embryo stage. This suggests that digital (CLF-mediated) switching only gains the competence to silence *FLC* between the mature embryo and 4-day-old (germinated) seedlings. Overall, this confirms that stable digital switching occurs post-embryonically.

3.5 *FCA* is not required for *FLC* silencing

PRC2-silencing is opposed by the transcriptional state of the locus. Similar to what has been observed in the *clf-2* mutant, in the extreme case of *fca-1*, the active state is stable in the warm. By exposing *fca-1* to cold-induced silencing, we observe that an *FLC* switch-off can be established and stably maintained, as after 6 weeks of vernalisation, both active and silenced states can be stably maintained (Fig. 3-9), similar to what has been previously observed in Col*FRI*. Further, this suggests that *FCA* is relevant in the establishment of the transcriptional state but not in the maintenance of silencing.

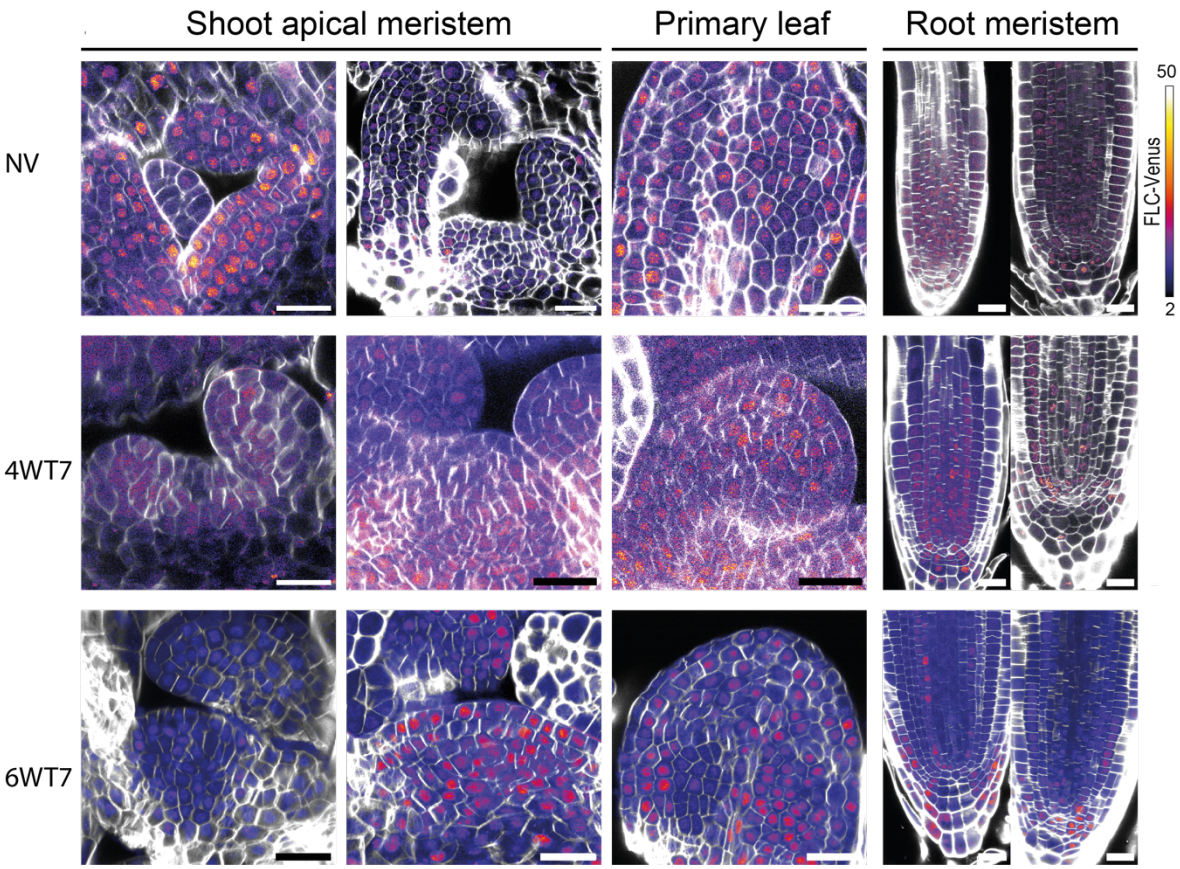


Figure 3–9 High transcriptional states in *fca-1* can be silenced by vernalization.

Representative confocal images of FLC-Venus imaging in different plant organs (shoot apical meristem, primary leaf and root meristem) at three timepoints after 0, 4 or 6 weeks of vernalization, and a subsequent 7 days of growth at 20°C. Grey shows the cell wall dye Renaissance 2200; FLC-Venus intensity is indicated by the colour map. Scale bar represents 20 μm, n=2 biological replicates with 5 roots. Note, the wildtype control here is not presented as no FLC-Venus signal (ON cells) was detected at 7 days.

Discussion

In this chapter, we have investigated the timing of the various switches occurring in *FLC* expression both in the embryo and during very early development after germination. We found that resetting of *FLC* expression occurs in the embryo and by the time of the mature embryo, all cells show an ON signal, which is similar between genotypes (although FLC-Venus levels are not equally high in every cell for each genotype). Subsequently, at a time point before 1-day post-germination, both an analogue reduction in expression level between the genotypes and a digital switch to OFF start to occur. The simplest scenario would be that the analogue reduction starts to occur immediately after germination and lowered transcription allows a switch to digital OFF to then begin to occur. We also found from our analysis of *clf-2*, that the digital switch OFF is mediated by Polycomb, as without this factor, the silenced state is not stable.

The analogue levels are set with the resetting of *FLC* in the embryo based on the presence of several autonomous pathway components such as FCA and (or the lack of) FRI. During embryogenesis, embryo specific proteins further drive transcription at *FLC* which has shown to have an effect in germination (Chiang et al., 2009). Post germination, the embryo-specific drivers of *FLC* are not expressed anymore and the autonomous pathway players fall back to action.

The requirement for *FLC* to resume in an active state during embryogenesis is further underlined by potentially having a function in seed dormancy. Chen and Penfield (2018) have shown that FT, via *COOLAIR*, regulates *FLC* in the embryo impacting seed dormancy by influencing seed coat development and seed hormone levels (Chen and Penfield, 2018). During embryogenesis, it is well known that PRC2 functions at specific developmental stages. Preventing an epigenetic OFF switch occurring too early might have many benefits for the plant. A similar mechanism has been proposed in embryonic stomatal development, where a developmental block in the stomatal precursor cells prevents lineage progression until after germination (Smit et al., 2023). Once germinated, the stomatal precursor cells perform fast stomatal differentiation for immediate photosynthetic activity. For *FLC*, the competence of PRC2 to silence *FLC* only upon germination might be a way to control quantitative variability between genetically identical plants allowing for a more successful reproductive strategy. Recent advances in methodologies requiring less material (e.g., CUT&RUN) are promising tools to answer questions regarding PRC2 function in the embryo.

Overall, our results remain consistent with a hypothesis of transcription opposing PRC2-silencing (Fig. 3-10). By using a system in which we use mutations in *FCA* to quantitatively modulate *FLC* transcription, we have found that the timing of the Polycomb OFF switch is influenced. In particular, we inferred that the probability of switching is negatively

correlated with the amount of transcription at the locus, as previously observed in a mammalian system (Holoch et al., 2021). The timing of the first digital response is by 1-day post-germination, co-occurring with the upregulation of *FLCs* antisense transcript, the lncRNA *COOLAIR*, possibly via changes in the local chromatin environment. Here, alternative termination levels of productive transcriptional activity set by an analogue mechanism determine the level of transcriptional antagonism to Polycomb silencing, as recently suggested in (Menon et al., 2024). *FCA*-mediated proximal cleavage and Poly(A) of *FLC* promote the removal of H3K4me1 by FLD, the histone demethylase and homologue of the mammalian H3K4 and H3K9 demethylase LSD1, in a complex that is associated with RNA Pol II (Mateo-Bonmatí et al., 2023). Furthermore, we observed that in a high *FLC* transcribing *fca*-1 mutant, cold-induced Polycomb silencing can be maintained in line with mathematical modelling predicting that *FCA* and FLD are relevant in the establishment of the silencing but not in its maintenance (Menon et al., 2024).

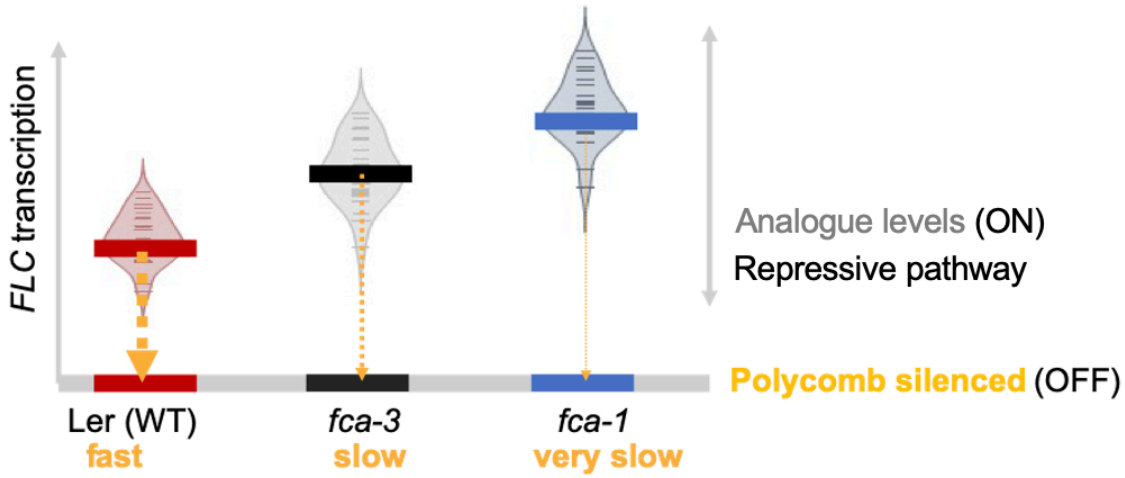


Figure 3–10 Graphical Model of Transcriptional state opposing probability of Polycomb switch to OFF state.

FLC transcriptional state in Ler, *fca-3* and *fca-1* is determined in early development when embryo-specific drivers of *FLC* transcription fade. The switching OFF probability depends on the analogue levels as set by the repressive autonomous pathway (including *FCA*). In Ler, the OFF switching happens fast, within the first divisions (days). In *fca-1*, the switching is very slow (but can be induced by vernalization) and in the intermediate *fca-3*, the switching can be observed in the root over a series of weeks.

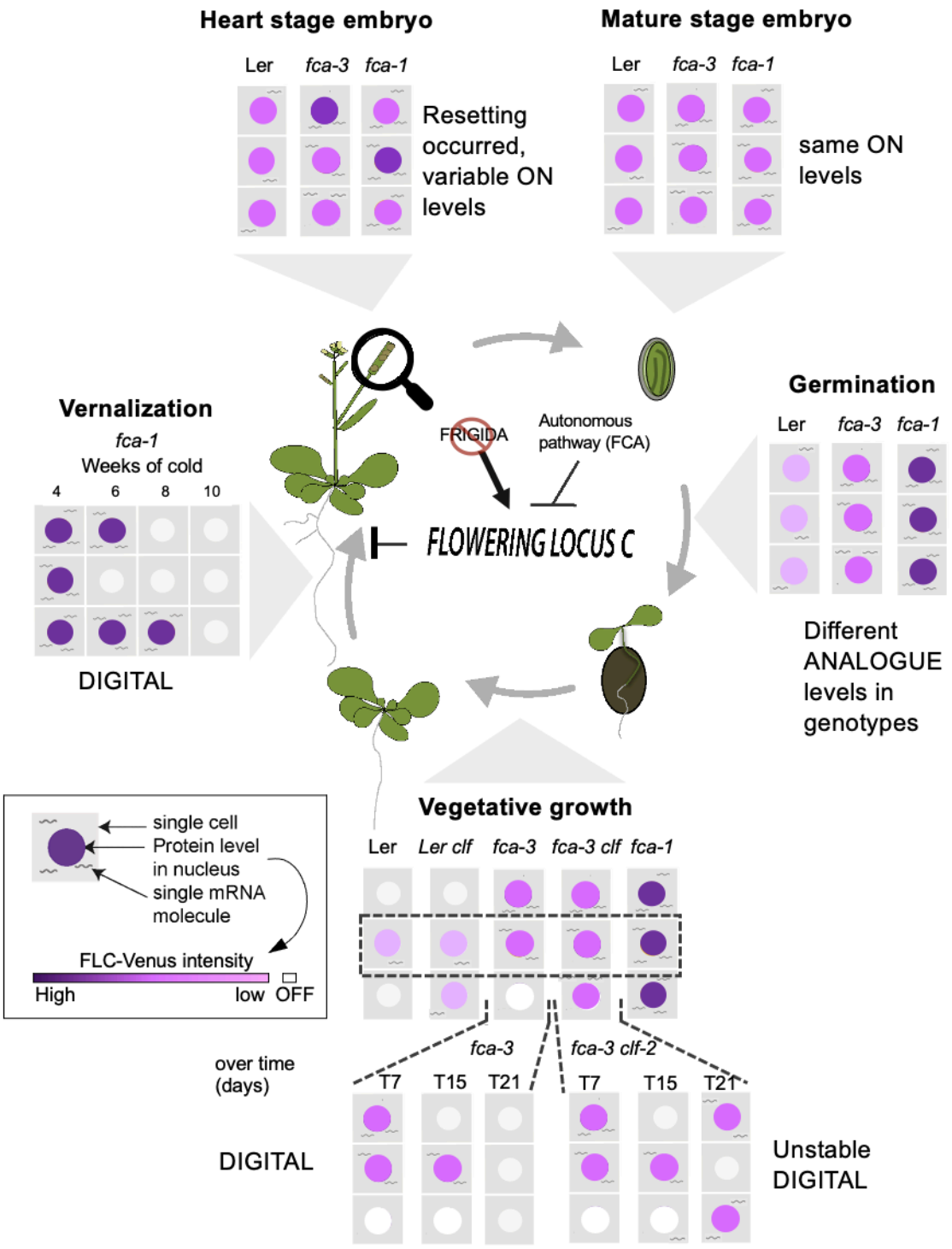


Figure 3–11 Graphical summary of findings in this Chapter.
Throughout the Arabidopsis life cycle, *FLC* regulation differs at different developmental stages. We observed a difference in FLC-Venus intensities between the heart stage embryo and mature embryo, with all cells being ON. The first digitally OFF cells were observed at 4 days after sowing (DAS), with germination occurring 3 days after sowing. Maintenance of *FLC* expression in extreme cases such as Ler and *fca-1* remained unchanged during vegetative development. Overall, however, including the *fca-3* and *fca-4* genotypes, cells switched digitally OFF with a probability depending on their prior transcriptional levels.

In accordance with what has been observed during vernalization, this switch-off depends on PRC2 and is not maintained in a methyltransferase mutant of CLF. In a highly *FLC* transcribing *fca*-1 mutant, vernalization is sufficient in silencing *FLC,* indicating that FCA is not required for the switching OFF.

CHAPTER 4 : SPATIAL DISTRIBUTION OF COOLAIR IN COLD TEMPERATURES

Abstract

Growth experiments with fluctuating temperature profiles reflecting natural temperature conditions in winter have shown that night temperatures below 0°C can accelerate *FLC* transcriptional shutdown (Zhao et al., 2021). Concomitantly, a set of *FLC* antisense transcripts, collectively called *COOLAIR*, is upregulated.

In this chapter, I characterise *COOLAIR* upregulation on a single-cell level using Fluorescence In Situ Hybridization (FISH), in which short oligonucleotides tagged with a fluorophore hybridize to the RNA molecule enabling it to be visualised as a focus.

With the goal of distinguishing between functionally distinct *COOLAIR* isoforms in cell types and identifying their subcellular localisation, an amplification-based FISH technique involving a double hybridisation was tested. However, the amplification step resulted in poor transcript specificity. I, therefore, used different probe sets to determine which *COOLAIR* isoforms contributed to *COOLAIR* upregulation in the fluctuating cold conditions. Detailed characterisation of *COOLAIR* foci was performed to dissect if *COOLAIR* upregulation can be explained by a larger fraction of cells expressing *COOLAIR* or if the *COOLAIR*-expressing cells contain more transcripts per cell. In root cells, I observed a wider expression zone plus larger foci. *COOLAIR* transcript upregulation did not depend on pre-vernalization *FLC* transcription states.

Moreover, whole mount FISH in root (and shoot) identified specific meristematic cells as being enriched with intronic *COOLAIR*. Interestingly, there seems to be a size limitation of the foci, as the maximum focus in the *COOLAIR* overexpressing NTM1-like 8-D3 mutant (*ntl8*-D3) was comparable to other genotypes, but the number of foci per cell increased.

A significant shift in the phases of the cell cycle was detected by gene expression analysis and by FISH in plants exposed to vernalization- temperatures. However, the strong fluctuating treatments did not show additional changes relative to constant conditions for the time of *COOLAIR* upregulation. More importantly, the preferential expression of *COOLAIR* in G1/S-phase in ambient temperatures is disrupted in the cold, suggesting that *COOLAIR* up-regulation is independent of cell-cycle stages during vernalisation.

Overall, these data implicate a role for antisense transcription in the response to freezing temperatures, acting to oppose sense transcription in a spatially dependent manner.

Introduction

Plants register changes in their environment to monitor the progression of seasons and adjust their growth accordingly. It has been shown that the silencing of the flowering repressor *FLC* in response to winter cold happens at each locus independently and consists of two concurrent yet interacting events: a gradual downregulation of transcription correlating with an upregulation of the *FLC* antisense transcripts, and an epigenetic change of the chromatin environment followed by a stable memory of the silenced state at the locus.

In cold conditions, *FLC* antisense transcripts collectively called *COOLAIR* are upregulated (Csorba et al., 2014; Rosa et al., 2016; Swiezewski et al., 2009). The significant up-regulation (seven- to eightfold) in *COOLAIR* RNA levels involves increased transcription and an increase in RNA half-life ($t_{1/2}$20°C= 100 min, $t_{1/2}$5°C = 240 min) (Csorba et al., 2014). Cold-induced or stabilised factors such as C-repeat containing Binding Factors (CBFs) and the NAC domain-containing transcription factor NTL8 bind to regions in the *COOLAIR* promoter, activating transcription (Jeon et al., 2023; Zhao et al., 2020a). A dominant mutant, *ntl8-D3* has a mutation in the transmembrane domain of the protein, leading to a constant relocation of the protein to the nucleus and constitutive induction of *COOLAIR* expression (Zhao et al., 2021).

COOLAIR upregulation is not a requirement for *FLC* silencing: plants lacking the *COOLAIR* transcript still vernalize normally when exposed to a prolonged period of cold (Jeon et al., 2023; Zhao et al., 2021). Although there are no canonically spliced antisense transcripts detectable in those lines, they are not fully deficient in antisense transcription as shown by the presence of cryptic antisense transcripts originating from intragenic regions in *FLC* (Zhao et al., 2021). Consecutive nights of temperature below 0°C can accelerate the *FLC* downregulation along with a super-induction of *COOLAIR* RNA (Zhao et al., 2021). Jeon et. al showed that *COOLAIR* can be induced by only one night of temperatures below 0°C, however this fails to accelerate the transcriptional shutdown of *FLC*. This suggests that while epigenetic changes of the chromatin state are working on a time scale of days to weeks (*FLC* shutdown), the transcriptional response of *COOLAIR* occurs within hours. However, it raises the question of how can the induction of *COOLAIR* contribute to the downregulation of *FLC*?

There are two possibilities for how the early upregulation of *COOLAIR* influences the transcriptional output of *FLC*. In the first, the *COOLAIR* transcript itself has an active function as shown for the specific cold-upregulated distal *COOLAIR* isoform (Class II-ii) which sequesters the *FLC* activator FRIGIDA away from the locus, reducing *FLC* initiation (Zhu et al., 2021). The second possibility is that *COOLAIR* transcription is the determining factor. This could for example be mediated by RNA polymerase II associated chromatin modifiers changing the chromatin environment (Mateo-Bonmatí et al., 2023), or by antisense transcription preventing sense initiation as suggested by an observed mutual exclusivity of transcription of either *COOLAIR* or *FLC* at the locus (Rosa et al., 2016).

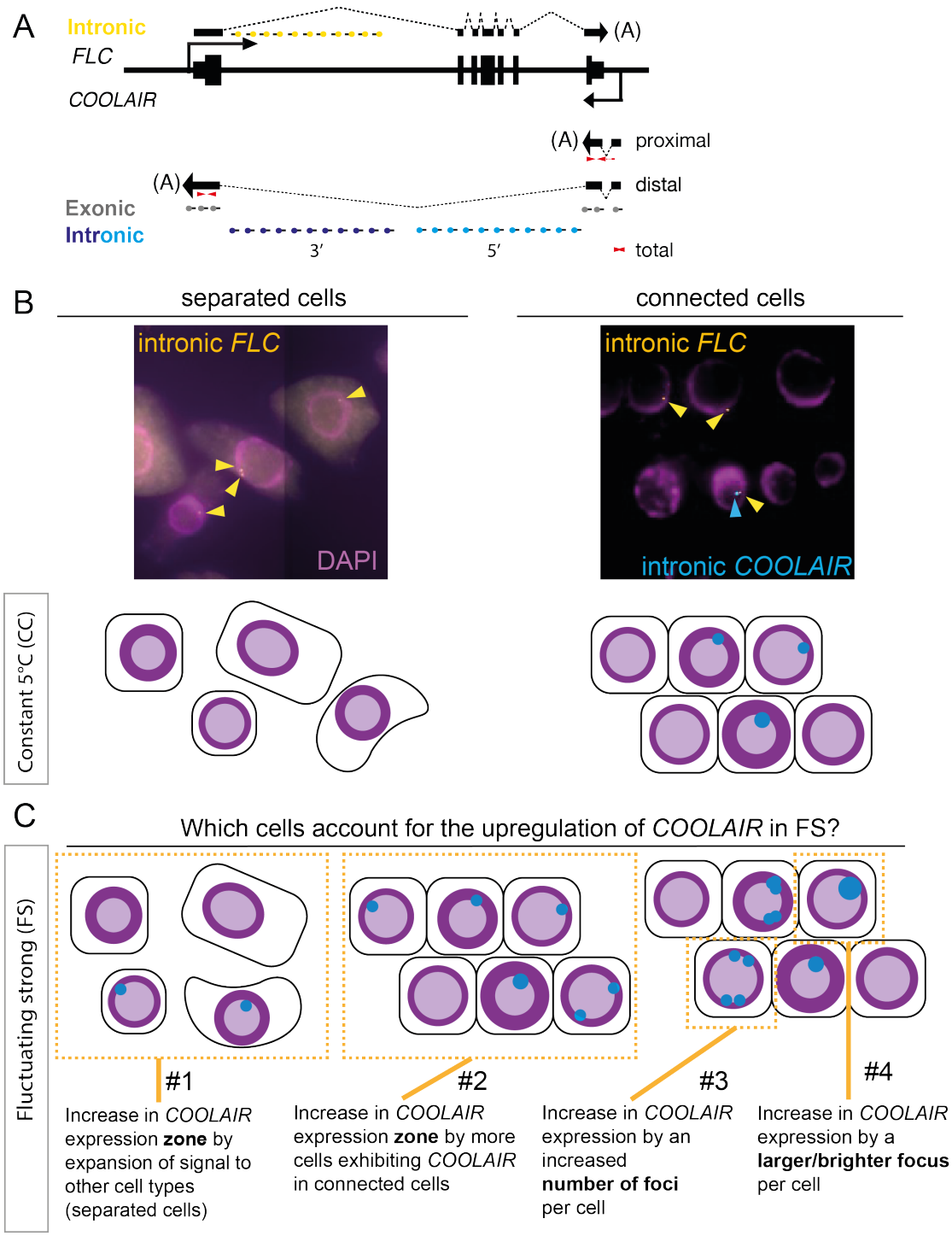


Figure 4–1 Concept of testing in which cells *COOLAIR* is expressed by FISH.

(A) Schematic representation of the *FLC* locus with probe and primer binding sites (FISH probes: *FLC* intronic: yellow, *COOLAIR* intronic: cyan (5′), light blue (3′), *COOLAIR* exonic: grey. qPCR primer to discriminate between *COOLAIR* isoforms specified on the right: red). **(B)** Example images of FISH in a single cell layer of root squashes. Cells are classified into separated cells (left) and connected cells (right) with *COOLAIR* being present in connected cells in Constant Cold (CC, cyan foci). Cell nuclei are indicated by DAPI staining (shown in purple). (Below) Schematic presentation of the observed separated and connected cells with cell outlines in black, purple cell nuclei and cyan *COOLAIR* foci. **(C)** Schematic of possible single-cell scenarios accounting for *COOLAIR* upregulation in Fluctuating Strong (FS) vernalization conditions (see Fig. 4-2A for details). #1 and #2 show an increase in the *COOLAIR* expression zone. This could expand to #1 other cell types (separated cells) or #2 a larger fraction of cells exhibiting *COOLAIR* in the connected cells. Increased signal per cell can be explained by an increased number of foci per cell #3 or #4, more signal per focus as estimated by fluorescence intensity times area of the focus.

To gain a better understanding of either case, we first need to understand which cells contribute to the *COOLAIR* upregulation in a spatial context. Previous RNA FISH (FISH) imaging targeting intronic *COOLAIR* transcript (Fig. 4-1A, blue intronic probes sets) identified signal in ~3% of ambiently-grown plant root cells. After 2 weeks of constant 5°C (Constant Cold - CC), the fraction of cells expressing *COOLAIR* increased to ~30% (Rosa et al., 2016). As FISH is performed on root squashes, the identity of those cells cannot be precisely determined. This 'squashing' of the root was performed to achieve a single-cell layer to facilitate imaging. Upon squashing, 2 populations of cells can be observed - evenly dispersed 'separated cells' and cells that do not break apart when the mechanical force is applied ('connected cells', Fig. 4-1B example image). *COOLAIR* signal was observed predominantly in connected cells. Little is known about the spatial context of the upregulation of *COOLAIR* in plants exposed to natural temperature fluctuations, as observed during the autumn-to-winter seasonal transition including freezing temperatures. In this chapter, smFISH is used to assess which type of cells account for the upregulation of *COOLAIR* that could be achieved by several different single-cell events graphically demonstrated in Fig. 4-1C with each individual scenario highlighted by a yellow dashed box.

In scenarios #1 and #2 the *COOLAIR* expression zone expands with more cells expressing the transcript. This could extend to #1 other cell types (separated cells) or #2 a larger fraction of cells exhibiting *COOLAIR* in the connected cells.

The alternative way in which upregulation could be achieved is by increasing the transcript number per cell. Signal per cell was counted to assess if there are #3 more foci per cell or #4, more signal per focus as estimated by fluorescence intensity times area of the focus (integrated density).

Differentiating between those scenarios provides an understanding of the upregulation of *COOLAIR* in the cold relevant for understanding the biological function of *COOLAIR* on a whole plant level.

Results

4.1 *COOLAIR* expression is highly induced by freezing temperatures experienced in plate-grown seedlings

Roots grow in soil containing microorganisms, the primary source of RNases capable of degrading the sensitive RNA molecules. For FISH, an elevated level of sterility is required to prevent RNA degradation during the fixation and sample preparation process as well as a transparent tissue such as the root. As previous fluctuating temperature (including freezing) experiments have been performed on rosette tissue of soil-grown plants (Zhao et al., 2021), I first confirmed if plants grown on sterile plates respond similarly to fluctuating

conditions by observing changes in gene expression via quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR) (Fig.4-2, soil [grey], plate [purple]).

All plants underwent a pre-growth period of 20°C for 7 days, followed by a temperature treatment with the below-specified conditions for 2 weeks. As previously reported in Zhao et al (2021), temperature profiles (Fig.4-2A) include plants grown in constant ambient temperature (NV – non-vernalized, 20°C, [purple line]) and plants grown in different frigid conditions here collectively called "vernalization treatments". Besides a constant 5°C (CC– constant cold, [blue line]), two fluctuating temperature profiles with an average temperature of 5°C were compared with either strong fluctuating conditions (Fluctuating Severe, FS) ranging from -1–12°C [black line] or mild fluctuating conditions (Fluctuating Mild, FM) ranging from 3–9°C [pink line]. In both FM and FS conditions, plants experience the same average temperature of 5°C. These two temperature schemes allow us to distinguish if *COOLAIR* upregulation is induced by temperatures below zero (0– -1°C) or by the fluctuation profile itself. Additionally, one night of FS treatment, here referred to as triggering (T) stimulus (abbreviated as CC+T and FM+T), was performed to distinguish between *COOLAIR* upregulation after one night of freezing temperatures and a potential adaptation to re-occurring freezing temperatures.

In soil or on plate, *FLC* spliced and unspliced RNA levels are reduced when exposed to cold (vernalization, Fig. 4-2BC). For spliced *FLC* mRNA gene expression, all vernalization treatments resulted in significantly decreased levels compared to NV (with $P<0.0156$). While in both growth media, the reduction of *FLC* appears stronger in FS compared to CC and FM, no significant difference was detected using 2-way ANOVA with multiple comparison testing. Comparing spliced and unspliced RNA levels can give us a better understanding of changes in transcription or splicing efficiency. To gain a better understanding of the changes of *FLC* levels, the expression of unspliced *FLC* was measured. Unspliced *FLC* expression levels are significantly reduced only in FS compared to NV supporting our overall observation that FS accelerates *FLC* shutdown compared to CC.

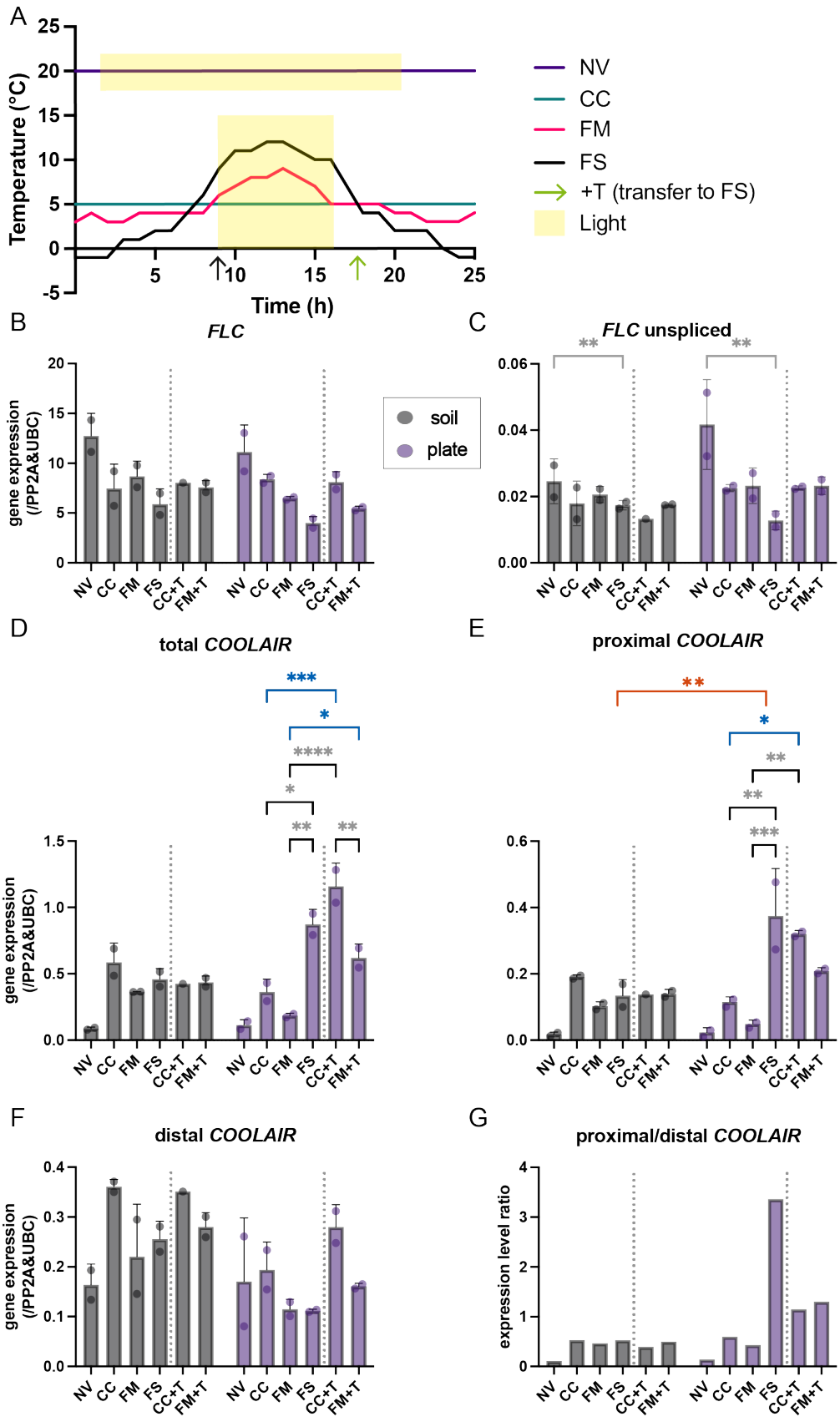
To assess change in *COOLAIR* transcription, primers were designed to target specific sequences that are unique to different isoforms (see binding sites in Fig. 4-1A, red arrows). In line with the reduction in *FLC* expression, total and proximal *COOLAIR* was significantly upregulated in all vernalization treatments compared to NV (Fig. 4-2DE). On plates, FS-grown plants showed the strongest up-regulation of proximal and total *COOLAIR*. Although following the same trend, the response on plates was overall stronger than on soil as indicated by significant differences for plate-grown but not soil-grown seedlings between different vernalization treatments. This is most likely due to the growth conditions in the soil buffering the severe temperature conditions. For distal *COOLAIR*, no significant differences to NV were detected for any vernalization-treatments. Using ANOVA with subsequent Šídák's multiple comparisons test between both growth media revealed no significant difference between the responses to vernalization-treatments, apart from a

stronger response in proximal *COOLAIR* in FS. To capture the shift from distal to proximal termination associated with early cold response (Swiezewski et al., 2009), the ratio of proximal/distal *COOLAIR* was assessed (Fig. 4-2F). Again, the expected differences were more prominent on plates compared to soil-grown plants. As the overall responses were not significantly different between plate and soil-grown plants, we conclude that both in both growth substrates, the plants respond in an analogous way to respective vernalization treatments. Hence, plate-grown plants can be used for FISH analysis to determine if more *COOLAIR* per cell or more cells expressing *COOLAIR* are contributing to the FS increase observed on the whole plant level.

Figure 4–2 Comparison of *FLC* gene expression after vernalization-treatments on plate and soil.

(next page)

(A) Temperature profile of growth conditions. NV – non-vernalized (plants grown in ambient temperature (20°C) for 7 days), CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, CC+T – constant 5°C with one night of FS, FM+T - fluctuating mild + one night of FS trigger. Yellow shadow represents the period in which the light is on in the growth chambers, indicating daytime. Note, CC, FM and FS (+T) plants grow in short day (8 hrs light/16 hrs dark) with a light intensity of 20 μmol m^{-2} s^{-1}, NV plants grow in long day conditions (16 hrs light/8 hrs dark) with a light intensity of 87 μmol m^{-2} s^{-1}. Black arrow represents the harvesting time point for all fluctuating experiments. The green arrow represents the time when plants were transferred to the triggering temperature profile the prior night as "cold" temperature intercept (6 p.m.). Note, 25 hr timepoint is identical to 1 hr. Gene expression analysis of *FLC* **(B-C)** and *COOLAIR* **(D-G)** in ColFRI after respective vernalization treatment (2 weeks) on plate- and soil-grown plants. **(B)** Gene expression of *FLC* mRNA **(C)** Gene expression data of *FLC* un-spliced **(D-F)** Gene expression of *COOLAIR* isoforms **(G)** Ratios of *COOLAIR* isoforms (proximal/distal). Šídák's multiple comparisons test comparing soil and plate (red asterisk) show a significant difference between FS for proximal *COOLAIR* with an adjusted p-value of 0.0019 for plate-soil FS comparison. Within each vernalization treatment (black and blue asterisks), a mixed-effects analysis with Tukey's multiple comparisons test was performed (significance levels of comparisons to NV are not shown in B, D, and E as all comparisons were different with $p<0.0155$). Blue asterisks highlight significant differences between vernalization treatments with and without one night of freezing (trigger). Error bars represent the SD of two biological replicates. Asterisks indicate level of significance with p: ***<0.001, **< 0.01, * <0.05. Error bars represent the SD of two biological replicates (n=2).


A
Temperature (°C)
Time (h)
NV
CC
FM
FS
+T (transfer to FS)
Light
B
FLC
gene expression (/PP2A&UBC)
soil
plate
C
FLC unspliced
D
total COOLAIR
E
proximal COOLAIR
F
distal COOLAIR
G
proximal/distal COOLAIR
expression level ratio
NV CC FM FS CC+T FM+T

4.2 Several cycles of night freezing are required to accelerate transcriptional reduction of *FLC*

One night of freezing is sufficient to induce proximal *COOLAIR* (Jeon et al., 2023). To account for the possibility of different transcriptional response dynamics, we included a triggering one-night freezing treatment (Fig-4-2BC, triggering treatment separated in graphs by grey dashed line). *COOLAIR* upregulation shows a faster response to triggering FS most prominently when the plants were grown on plates (Fig.4-2D-F). Total *COOLAIR* levels in CC+T and FM+T are significantly different to non-triggered (CC and FM) reaching levels similar or higher than FS plants (Fig4-1D, blue asterisks), suggesting that indeed one night of FS is sufficient to superinduce *COOLAIR*. Proximal *COOLAIR* reaches expression levels almost as high as observed after 2 -weeks of FS. This suggests that one night is sufficient to induce a fast responsive *COOLAIR*, reoccurring freezing does not accelerate *COOLAIR* levels.

For proximal *COOLAIR*, only the CC+T treatment showed a significant difference to CC, whereas distal *COOLAIR* was variable but overall unaffected by one-night freezing temperatures (Fig4-2E,F). Interestingly, although overall *COOLAIR* levels are different in CC+T and FM+T, the shift of distal to proximal *COOLAIR* after the triggering night is comparable between both treatments as indicated by the similar proximal/distal ratio (Fig4-2G).

In contrast to *COOLAIR*, unspliced and spliced *FLC* show no differences in expression levels compared to respective CC or FM control after one night of triggering cold (Fig-4-2BC). This shows that one night of FS is not sufficient to induce a response in *FLC* reduction.

The anti-correlation of induced *COOLAIR* levels with the reduction in *FLC* levels is more prominent in unspliced *FLC*, which was not observed after only one night of freezing (triggering) (Fig4-2C). In conclusion, several nights of FS are required to accelerate *FLC* shut down.

4.3 Testing an amplification-based FISH technique

Before performing smFISH to test if the elevated *COOLAIR* expression in FS resulted from increased expression in the same cells or a larger proportion of cells expressing *COOLAIR* (Fig. 4-1A), I considered different methods for *COOLAIR* detection.

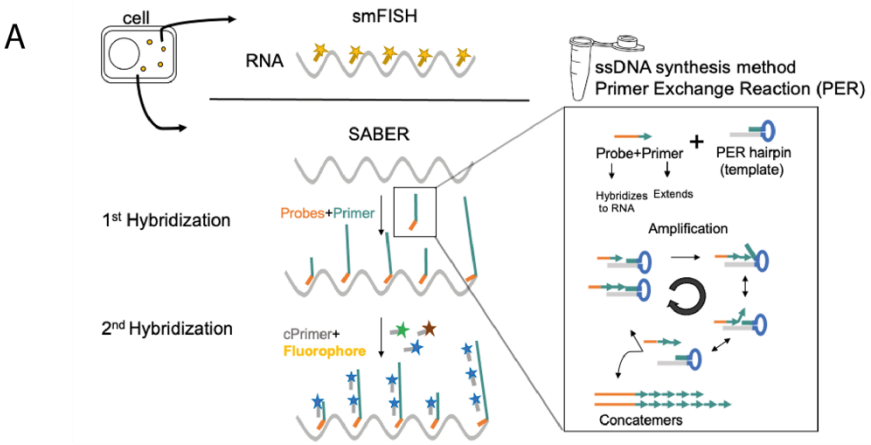
Current fluorescent in situ hybridization methods have two limitations regarding directly capturing the highly-FS induced proximal *COOLAIR*. First, they require a relatively long RNA molecule as a probe hybridization template (~750 base pairs [bp])). About 35 probes of ~ 20 nucleotides (nt) length are required to bind to one RNA molecule for it to be detected as a diffraction-limited spot. This limitation makes it impossible to visually discriminate between specific isoforms such as proximal (~400 nt) and distal *COOLAIR* (600-750 nt) (Fig.4-1A) and even more difficult for isoforms that only differ in a few nucleotides (such

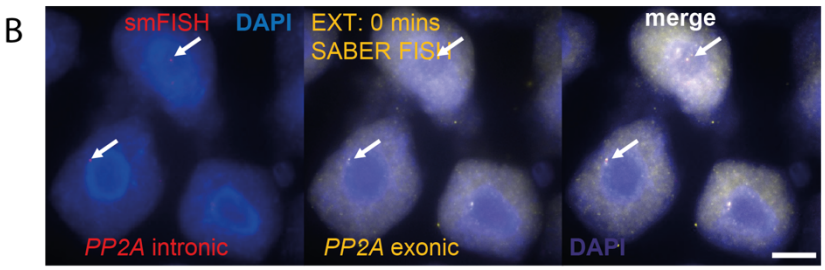
as *COOLAIR* Class II-II). Secondly, there is little flexibility/availability of fluorophores that are spectrally distinct from each other to sufficiently differentiate between probe sets.

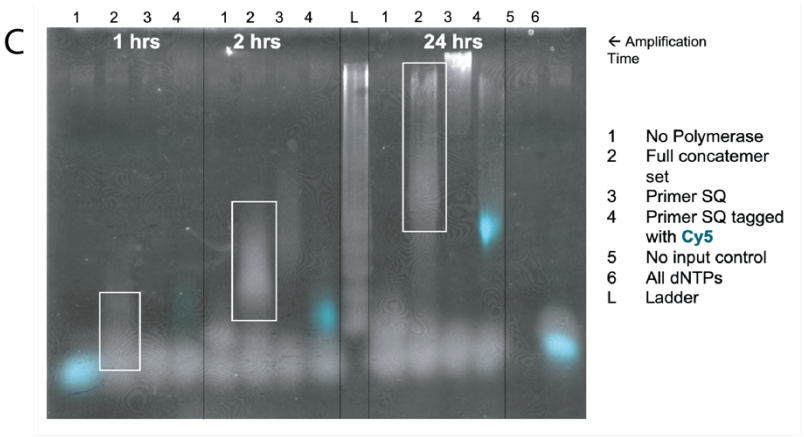
To address both limitations, I tested the Single-stranded DNA Amplification By Exchange Reaction (SABER) FISH method to amplify the signal of a shorter RNA sequence (Kishi et al., 2019)(Fig.4-3A). The method is similar to single molecule FISH as it relies on the hybridisation of DNA oligonucleotides to a target sequence. Whereas in the standard smFISH protocols the fluorophore is covalently bound to the hybridizing probe, in SABER FISH two consecutive hybridizations steps are performed. In the first step, the hybridising oligonucleotide consists of the probe, which hybridizes to the specific RNA sequence, and an externally extended primer sequence. In the second hybridisation step, a complementary primer sequence connects a fluorophore to the first hybridisation result. The extended primer of the first hybridization is produced in a rolling-cycle Primer Exchange Reaction (PER). By varying the length of PER amplification to produce concatemers of varying lengths, the number of fluorophore binding sites can be modified to enable optimisation for optimal signal/noise ratio and probe length (whilst maintaining cell penetration potential). Taken together, the 2-step system allows flexibility in adding probe numbers, in concatemer extension length and testing different fluorophores.

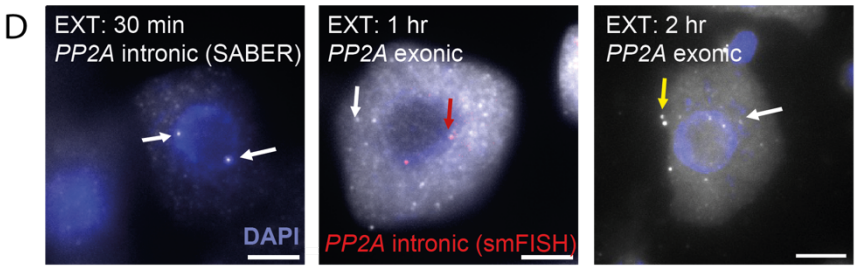
First, I confirmed that the double hybridization (without any PER amplification) yields comparable results as previously observed for smFISH probes. To test the method in plants, colocalization of exonic SABER oligonucleotides (35 oligonucleotides) to intronic smFISH probes (48 oligonucleotides) both targeting mRNA of the abundant housekeeping gene *PP2A* subunit A3 (AT1G13320.1), was assessed (Fig. 4-3B) in root meristem cells. The SABER signal in the nucleus colocalized with the intronic smFISH signal confirming successful double hybridization and specificity of the probes to *PP2A* target. Importantly, no effect on background noise levels was observed and the two cycles of hybridization seemed to enable stable detection of mRNA foci. These results indicate that SABER can successfully detect single molecule RNA signals.

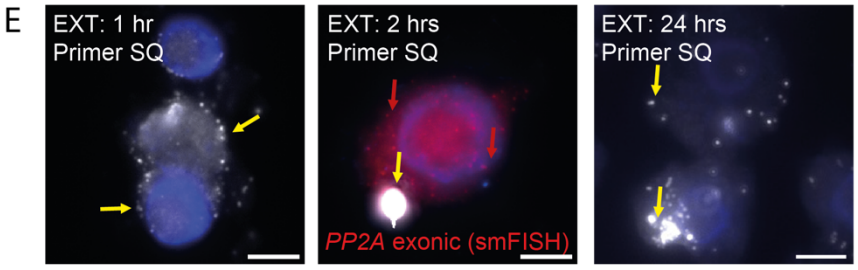
To apply it to shorter *COOLAIR* isoforms, I first needed to optimise the concatemer extension length. Different PER amplification times were tested to determine primer extension length (Fig. 4-3C, see white boxes). For improved detection of single-stranded DNA (most DNA dyes are DNA intercalating) an additional primer set (Fig. 4-3C, **4**) tagged with Cy5 was run as a proxy for primer extension, as the fluorophore-tagging enables size detection on an agarose gel. Amplified concatemers were purified and used for hybridization on root squashes. Using specific intronic and exonic probes, I was able to detect *PP2A* RNA (Fig. 4-3D) as indicated by the colocalization of smFISH intronic (red arrow) with SABER exonic probes. With increased amplification times, I observed uneven probe signal strength (2h SABER, compare yellow to white arrows).


A
cell
smFISH
RNA
ssDNA synthesis method
Primer Exchange Reaction (PER)
SABER
1st Hybridization
Probes+Primer
2nd Hybridization
cPrimer+
Fluorophore
Probe+Primer
PER hairpin
(template)
Hybridizes to RNA
Extends
Amplification
Concatemers


B
smFISH
DAPI
EXT: 0 mins
SABER FISH
merge
PP2A intronic
PP2A exonic
DAPI


C
1 2 3 4 1 2 3 4 L 1 2 3 4 5 6
1 hrs
2 hrs
24 hrs
← Amplification Time
1 No Polymerase
2 Full concatemer set
3 Primer SQ
4 Primer SQ tagged with Cy5
5 No input control
6 All dNTPs
L Ladder


D
EXT: 30 min
PP2A intronic (SABER)
DAPI
EXT: 1 hr
PP2A exonic
PP2A intronic (smFISH)
EXT: 2 hr
PP2A exonic


E
EXT: 1 hr
Primer SQ
EXT: 2 hrs
Primer SQ
PP2A exonic (smFISH)
EXT: 24 hrs
Primer SQ

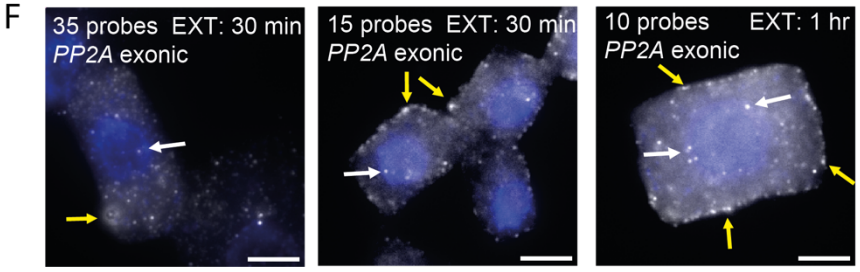

F
35 probes EXT: 30 min
PP2A exonic
15 probes EXT: 30 min
PP2A exonic
10 probes EXT: 1 hr
PP2A exonic

Figure 4–3 Comparison of FISH methods for mRNA detection in squashed root.

(previous page)

SABER probe extension and in situ hybridization on root meristem cells.
(A) Comparison of smFISH (top) and SABER FISH (bottom) methods. In the latter, concatemers (Primer+Probes) are produced with repeated Primer exchange reactions (PERs) **(B)** The bright white dots (white arrows) represent single mRNA molecules of PP2A mRNA single plane image of unamplified concatemers. Red = unspliced (intronic, Stellaris 49 probes) probes and yellow = full (exonic, SABER 35) probes. Arrows show co-localization of probe sets and indicate active transcribing PP2A. Single plane images, blue= DAPI, bright white dots= mRNA. **(C)** SABER probe sets were amplified by PER for 1, 2 and 24 hrs. White boxes indicate amplification of concatemer after respective time. White boxes show extended SABER primers 1. No Polymerase, 2. Exonic probe set, 3. Primer SQ (without RNA hybridization sequence (control to detect off-target binding during imaging), 4. Primer without RNA hybridization sequence tagged with Cy5 Fluorophore (control to reliably detect primer extension on the gel), 5. No input control, 6. all dNTPs (primer extension control, since it is dependent on the absence of dGTP), L is the Ladder. DNA (grey) was stained with SYBRGold. **(D-F)** Representative images of SABER probes after Primer Exchange Reaction *(*PER*)* - amplifications on single root cells for 30 mins, 1, 2 and 24 hr(s) (PER-amplification length specified on each image *as EXT*). White arrows indicate amplified SABER mRNA signal, yellow arrows point to unspecific SABER signal, red arrow shows smFISH signal **(D)** Extended SABER concatemers. **(E)** Extended SABER Primer Sequence (SQ). Note, with increasing length of PER-amplification, extended probes bound non-specifically **(F)** Extended SABER concatemers with reduced number of probes.

To test off-target binding, the sample was incubated with the extended primer sequence (without sequence-specific probe (hybridization) sequence) (Fig. 4-3E). For all PER amplification lengths (1 hr, 2 hrs, 24 hrs), an accumulation of signal on the outer cell wall was observed. Efforts to wash them off were unsuccessful. Note, during the amplification, the primer only sequence extends faster compared to the concatemer (Primer+Probe) (Fig. 4-3C, lane **2** compared to lane **3**). Trials using the full amplified concatemers with reduced probe number (35, 15, 10 oligonucleotides, Fig. 4-3F) were able to detect some specific PP2A mRNA molecules but did not reproduce reliable foci numbers as observed with the smFISH probe set (~120 molecules per cell). Moreover, the amplified signal seems to be more abundant in the outer cytoplasm and not in the nucleus, the region we are most interested in for detecting *COOLAIR* isoforms.

Overall, the lack of specificity, uneven amplification and the absence of reliable signal in the nucleus showed that this method would not give us the desired specificity. Hence, I did not proceed with the SABER method but rather made use of the several different probe sets (Fig. 4-1A) that might allow us to infer information about the isoform and spatial distribution of *COOLAIR* within the nucleus.

4.4 Single-molecule smFISH gives insight into *COOLAIR* isoforms in foci

Considering the observed shift in proximal/distal *COOLAIR* in FS in the gene expression analysis (Fig.4-2G), a comparison of different smFISH probe were compared to better characterize *COOLAIR* foci. Three probe sets for *COOLAIR* (Rosa et al., 2016) were used:

(1) Exonic full-length (distal) *COOLAIR* is consisting of 46 oligonucleotides. Among these probes, 23 are designed to bind to the 5'sequence present in (some but not all) proximal and distal isoforms, while the remaining 23 probes bind to distal-specific sequences (Fig 4-1A, distal probe set schematically portrayed in grey).
(2) Intronic (5') probe set, consisting of 48 probes specific to the 5' of *COOLAIR* intron 2 (Fig 4-1A, probe set portrayed in cyan (light blue)).
(3) Intronic (3') probe set, consisting of 48 probes specific to 3' of *COOLAIR* intron 2 (Fig 4-1A, dark blue probe set).

Several experiments using either 5' and 3' intronic probe sets (sets 2,3) resulted in the detection of equal numbers of foci (compare Fig. 2-6, A1-3) suggesting that the whole locus is transcribed.

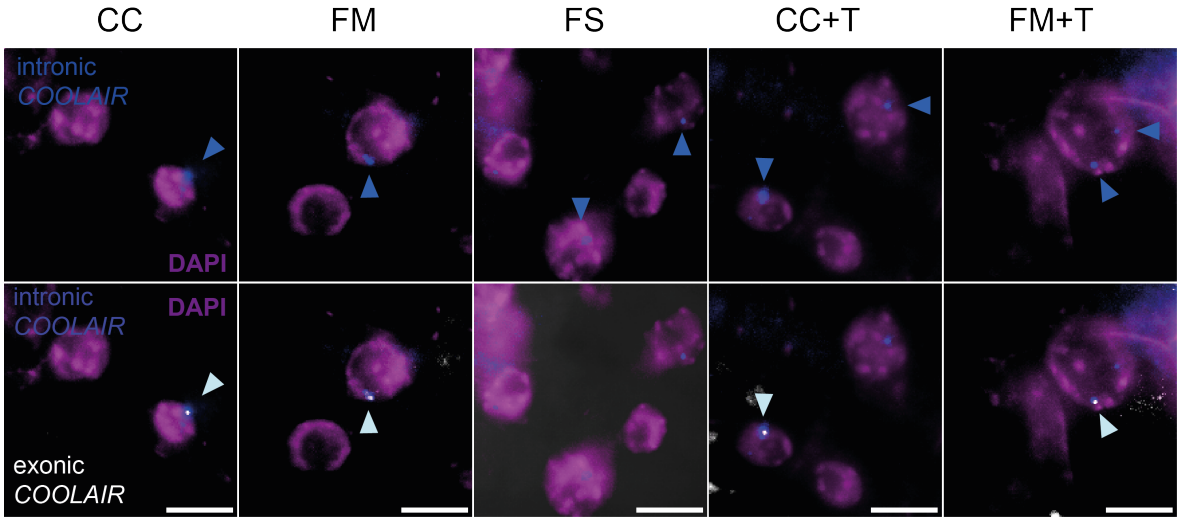


Figure 4–4 *COOLAIR* accumulation in cold consists of full-length transcript.

Comparison of *COOLAIR* foci as detected by smFISH probe sets in response to 2 weeks of vernalization treatment: CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, CC+T – constant 5°C with one night of FS, FM+T - fluctuating mild + one night of FS trigger. 5' intronic *COOLAIR* (blue) and exonic *COOLAIR* (white), DAPI (DNA dye, magenta). Arrows point to representative smFISH signal (blue: intronic signal, light grey: intronic and exonic signal). Upper row only shown intronic *COOLAIR* signal whereas bottom row shows colocalization of intronic and exonic *COOLAIR*. Note, all exonic *COOLAIR* signal colocalized with intronic signal. No exonic *COOLAIR* was detected outside the nucleus. Scale bar represents 5 μm.

Throughout all vernalization treatments, foci observed using the intronic probe sets were larger than for the exonic probe set (Fig4-4). The exonic signal (set 1) always colocalised with the intronic signal (set 2) but was not always detected at foci observed with the intronic probe sets. This suggests that distal spliced *COOLAIR* is not retained at the active transcription site. In an earlier study, it was observed that most spliced *COOLAIR* is in the cytoplasmic fraction (Csorba et al., 2014). Here, solely exonic signal without any co-localising signal for the intronic probe sets - corresponding to spliced molecules- was detected in both the nucleus or the cytoplasm. One possibility for the lack of cytoplasmic or nuclear signal in the smFISH analysis is the complex secondary structure of the spliced

COOLAIR transcript which could impede efficient probe binding (Hawkes et al., 2016; Yang et al., 2022).

As shown above, we observed the strongest upregulation of proximal *COOLAIR* by gene expression analysis in FS conditions (2 weeks of FS or one-night triggering) (Fig. 4-2E). While the smFISH probes lack the sensitivity to identify individual proximal *COOLAIR* transcripts, we can deduce that proximal *COOLAIR* is not accumulating at the *FLC* locus in FS. This reasoning is based on the fact that the distal exonic smFISH set, containing 23 sequence-specific probes for the proximal *COOLAIR*, would likely detect the proximal transcript if multiple proximal *COOLAIR* transcripts were the major component of the foci (similar observations in Rosa et al., 2016). Here, we did not detect more exonic signal in FS when proximal *COOLAIR* is upregulated. In essence, the gathered evidence leads us to the conclusion that the observed foci do not consist solely of accumulated proximal *COOLAIR* transcripts, but instead constitute the full-length transcript.

As shown above, the intronic *COOLAIR* probes were most suitable to detect cold-induced *COOLAIR*-upregulation. The constitution of the foci after vernalization-treatments did not differ between treatments apart from the fact that less exonic signal was detected in the FS treated. In the following experiment, smFISH probes hybridizing to the 5′ or 3′ region of *COOLAIR* intron 2 not covering proximal *COOLAIR* were used to further establish the cellular context of this increase.

4.5 More *COOLAIR* molecules and more cells expressing *COOLAIR* contribute to increased levels

Representative images (Fig. 4-5) of separated and connected Col*FRI* root cells after temperature treatments show intronic foci for *FLC* (yellow arrows) and *COOLAIR (*cyan arrows). No foci were detected in *FLClean*, a genotype with a CRISPR deletion of the *FLC* locus (Nielsen et al.) (shown here only for the FS-treatment). As expected, the quantification of single cells expressing *COOLAIR* (Fig. 4-6) showed a difference between separated and connected cells, with the latter showing a larger fraction of cells with *COOLAIR* in all tested conditions (Fig. 4-6A, see only *COOLAIR* containing cells in Fig. 4-4A.0). The response of separated cells to different vernalization treatments, however, showed a slightly larger fraction of cells with signal, whereas no difference between vernalization treatments was detected for the connected tissue. Surprisingly, the response of FM in separated cells was lower compared to the other vernalization treatments. The analysis indicates a slight trend for a wider expression zone of *COOLAIR* (scenario #1, Fig. 4-1).

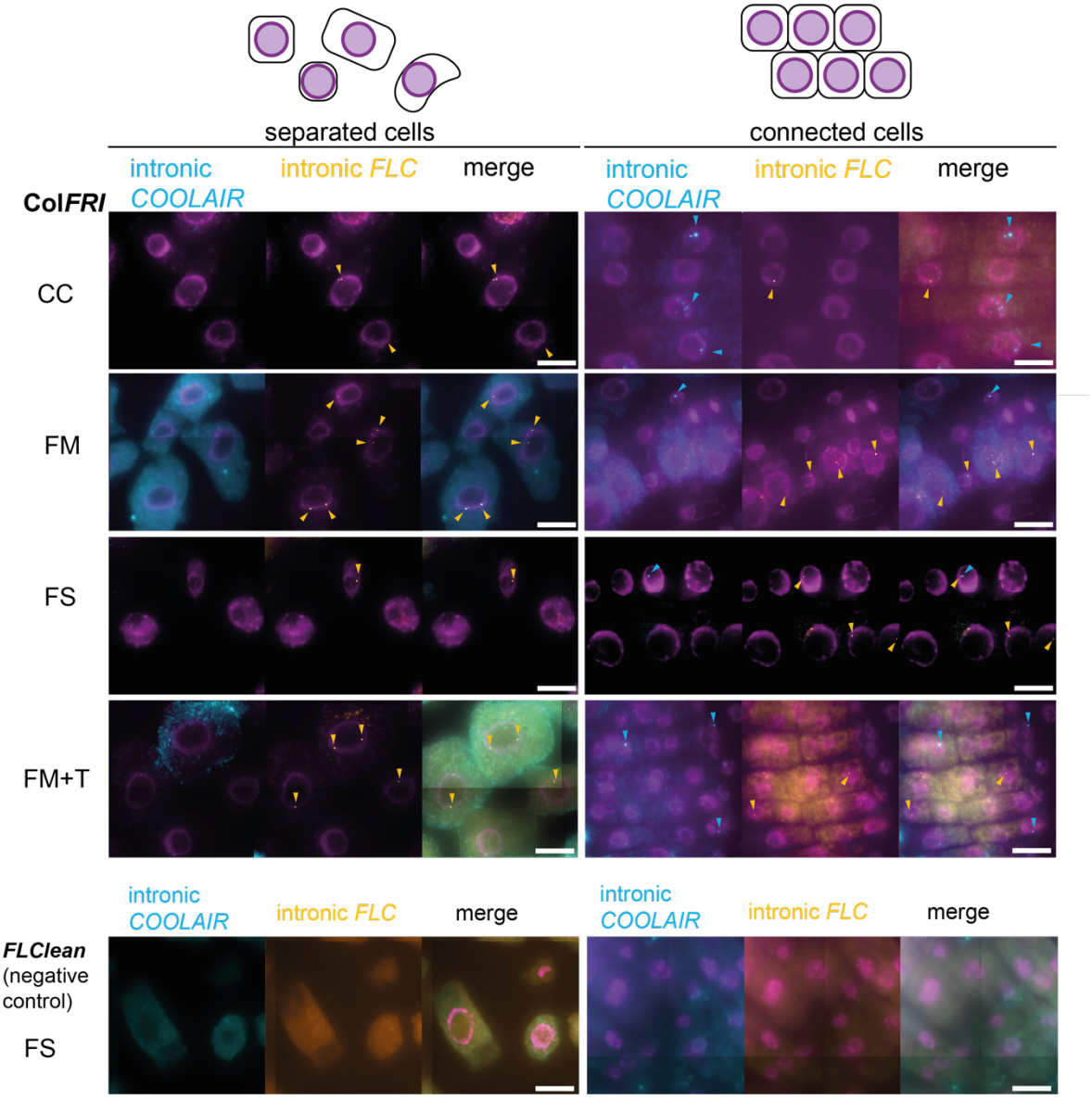


Figure 4–5 Representative smFISH images of intronic *COOLAIR* (5′ intronic probes, cyan) and intronic *FLC* (yellow) signal in Col*FRI* and *FLClean* after exposure to different vernalization treatments.

(CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, FM+T - fluctuating mild + one night of FS trigger). Separated (left) and connected (right) cells in Col*FRI* and *FLClean* after specified temperature treatment. Cyan arrows point to representative intronic *COOLAIR* signal, yellow arrows point to intronic *FLC* signal, Cell nuclei are indicated by DAPI staining (shown in purple). Scale bar 5 μm.

Additionally, *COOLAIR* up-regulation could be accounted for by more foci per cell (#3) or more signal per focus (#4) in the *COOLAIR*-expressing connected cells (Fig. 4-6CD). To assess #3, foci per cell were counted, as highlighted in Fig. 4-6A.0 (magnified region of Fig. 4-6A excluding cells with zero counts). Interestingly, foci number per cell in FS does not increase compared to CC. We next asked if the number of transcripts per focus changed by quantifying the mean fluorescence intensity of an active transcription site (two brightest foci, covering each *FLC* locus) (Rosa et al. 2016). Indeed, when fluorescence intensity/area was compared between vernalization treatments, higher fluorescence intensity values as shown by several data points displaying values between 200-300 a.u. were detected in FS

(Fig.4-6B), whereas in CC and FM, less fluorescence intensity values in that range were detected.

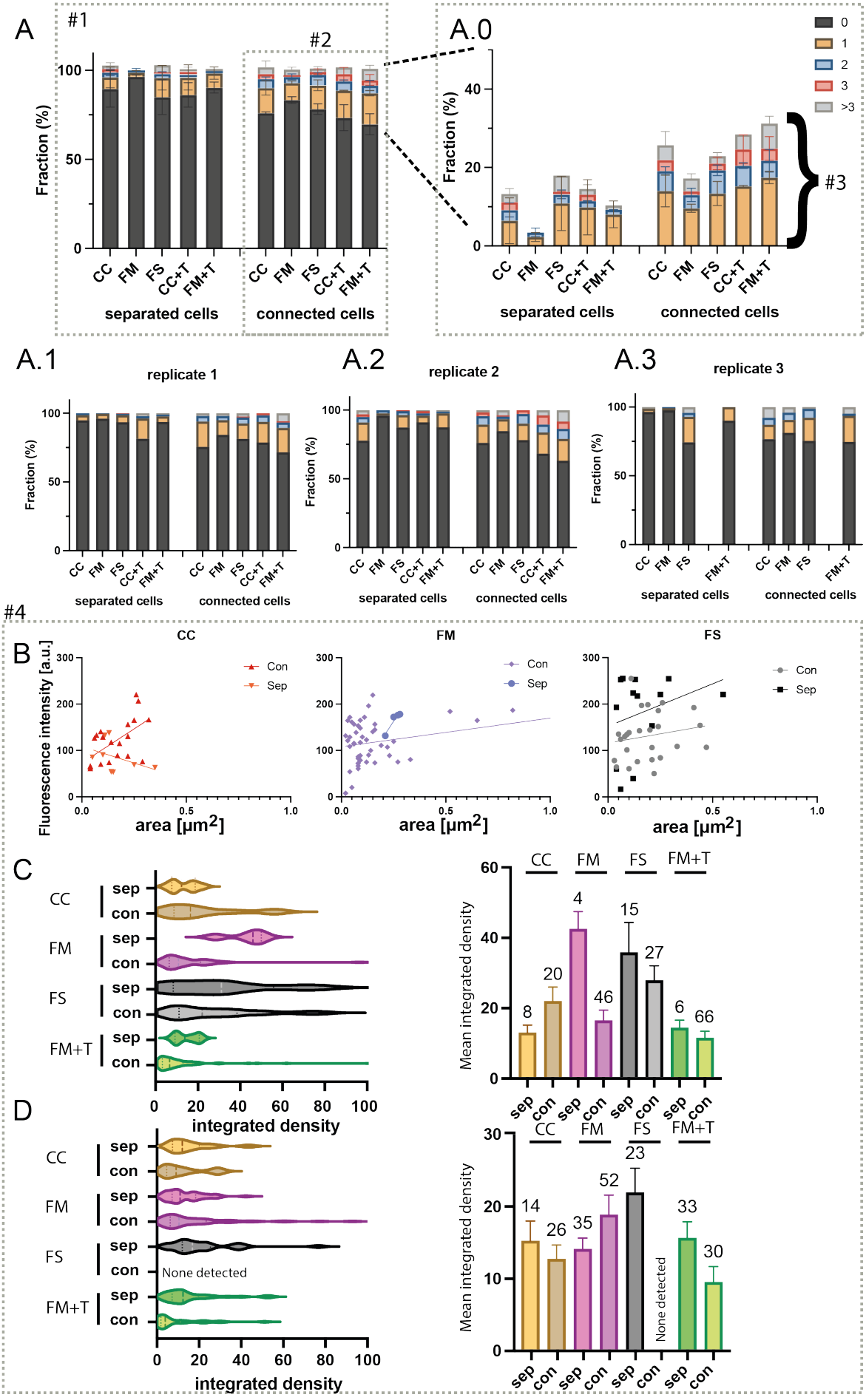


Figure 4–6 Temperature fluctuations including freezing leads to an expansion of *COOLAIR* expression zone and increased molecules per focus *(previous page)*

Quantification of smFISH intronic *COOLAIR* foci in Col*FRI* after vernalization treatments (CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, FM+T - fluctuating mild + one night of FS trigger. **(A)** Fraction of cells expressing intronic *COOLAIR* (as detected by smFISH signal) after vernalization treatments in separated or connected cells (compare Fig. 4-5). **(A.1-3)** Quantification per biological replicate with N>150 per replicate. A1. and A.3: 5‘ specific *COOLAIR* probes were used, A.2: 3’ specific *COOLAIR* probes were used. Colours represent different fractions of *COOLAIR* foci numbers observed in cells (0->3) **(B)** Fluorescence intensity of two brightest foci per cell (assumed to be the active transcription sites (ATS)) per focus area (μm^3) in connected (con) and separated (sep) cells. Lines represent linear regression for either con or sep cells (**C-D**) Integrated density (fluorescence intensity x area) of *COOLAIR* ATS. (C+D) Violin plot of the respective distribution of detected foci (after thresholding) used in the analysis in 200 (randomly picked) cells of respective vernalisation treatment (D+F) Bar graph of mean integrated density ± SEM. Numbers on top of graph represent detected foci (n=x in 200 cells) (E-F) Integrated density of *FLC* ATS (E) Violin plot (F) Bar graph of mean integrated density. Numbers represent detected *FLC* foci (after thresholding) used in the analysis in 200 (randomly picked) cells of respective verbalisation treatment.
Grey dotted boxes 1-4 represent different scenarios that could lead to an increase of *COOLAIR* in FS (see Fig. 4-1). #1 *COOLAIR* expression zone expands to other cell types (e.g., separated cells, [sep]), #2 *COOLAIR* expression zone expands in the same cell type (connected cells [con]). #1 or #2 results in a larger fraction of cells exhibiting *COOLAIR*. Alternatively, an increasing transcript number per cell could lead to a *COOLAIR* upregulation (#3 and #4). In scenario #3 foci number per cell is increased, in #4 fluorescence intensity per focus is increased.

When distinguishing between separated (sep) and connected (con) cells, the same upwards trend of increasing foci fluorescence intensity with area size in connected cells was observed. Strikingly, however, this increase in fluorescence intensity per area was not observed in separated cells in CC. The integrated density (mean fluorescence intensity x size of foci, Fig.4-6C) also revealed subtle differences between the vernalisation treatments. Whereas in CC, brighter foci were detected in connected cells, in FM and FS there were also some in separated cells. This wider distribution was in line with our previous observation of a wider expression zone of *COOLAIR* in FS. In FM+T, no increase in integrated density was observed. While in FM the foci in separated cells were brighter in intensity, they were also rare (n=4/200 cells). The few detectable smFISH foci in FM in combination with increased gene expression levels compared to CC could suggest that either different plant tissues are responding differently to FM-treatment or that the fluctuating temperatures and lack of exposure to freezing impacts splicing efficiency, resulting in an efficient processing of *COOLAIR* at the locus. Interestingly, sense transcription negatively correlated (Fig.4-6D) with antisense abundance.

In conclusion, one-night of freezing is not sufficient to increase foci number or accumulate more signal per nucleus, whereas 2 weeks of FS expands the expression area as well as *COOLAIR* transcript number per active transcription site (brightest foci). FLC/ *COOLAIR* transcription appears to be antagonistic in fluctuating cold conditions, as previously reported for CC (Rosa et al. 2016).

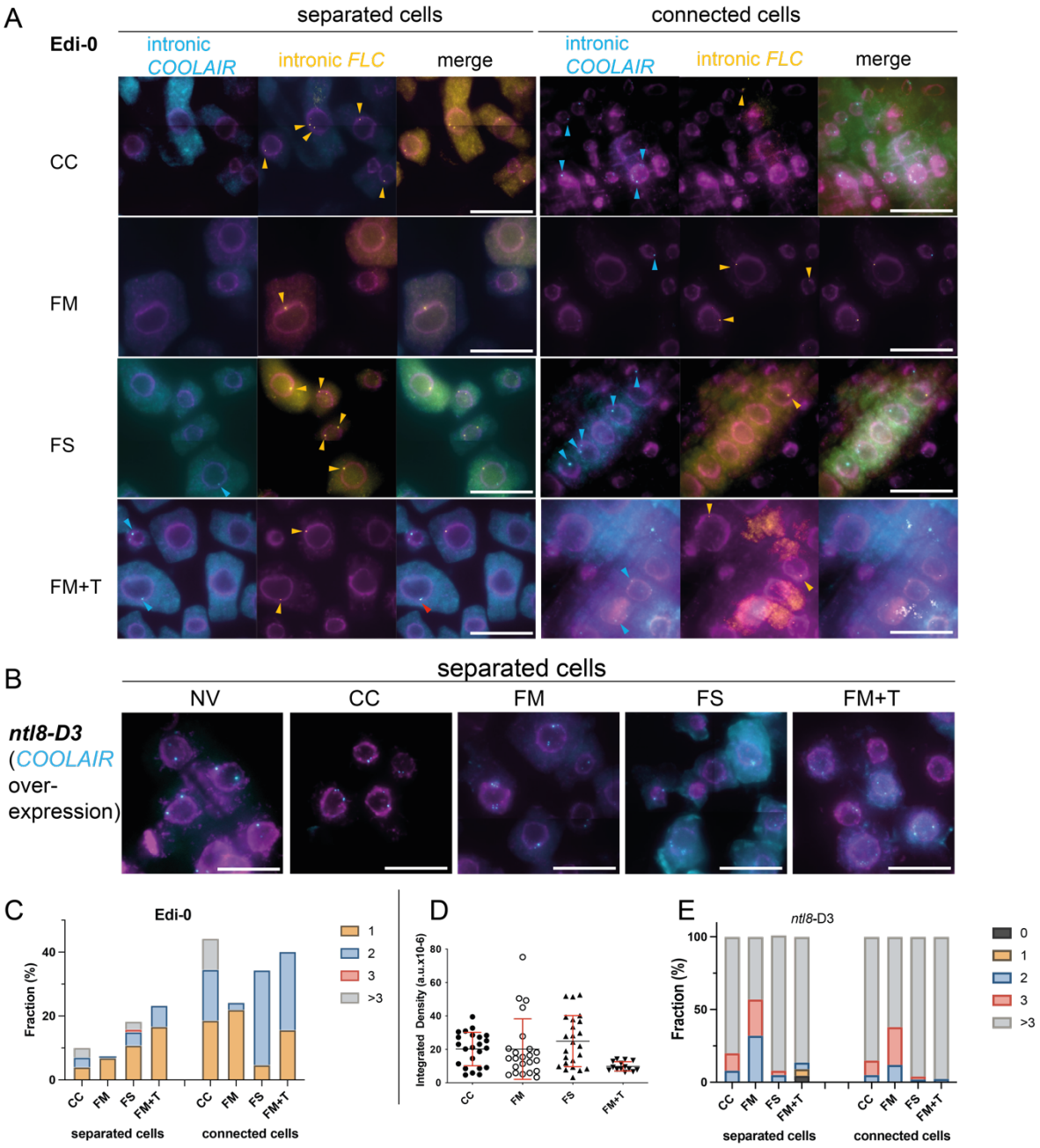


Figure 4–7 Wider expression zone observed in other genotypes.

Representative images and quantification of intronic *COOLAIR* signal in Edi-0 and *ntl8*-D3 (Col*FRI*) in vernalization treated plants (CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, FM+T - fluctuating mild + one night of FS trigger) **(A)** Representative images of nuclei of each temperature profile in separated (left) and connected (right) cells in Edi-0 **(B)** Representative images of nuclei of each temperature profile in separated cells in *ntl8*-D3 **(C)** Quantification of foci per cell in Edi-0 **(D)** Integrated density of active transcription sites (ATS) as determined by the two brightest foci per cell in *ntl8*-D3 **(E)** Quantification of foci number per cell in *ntl8*-D3.
Red error bars represent the SD of the integrated density signal. Cyan arrow points to intronic *COOLAIR* signal, yellow arrow points to intronic *FLC* signal, red arrow points to adjacent but not colocalising intronic *COOLAIR* and *FLC* signal. Scale bar 10 μm.

To see if the *COOLAIR* response as observed in Col*FRI* is representative of rapid-vernalizing genotypes, the response to FS in another rapid-vernalizing accession, Edi-0, was tested (Fig. 4-7A). In line with FISH results in Col*FRI*, higher *COOLAIR* levels were observed in connected cells compared to separated cells in vernalization-treated plants.

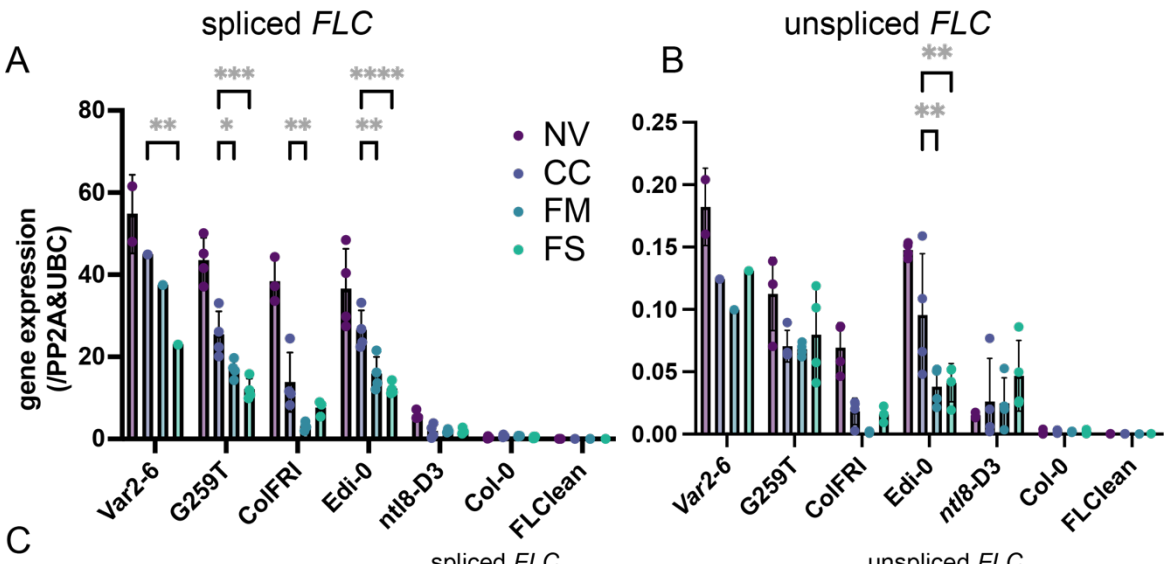
Further, the *COOLAIR* expression zone was expanded to separated cells in FS and FM+T. Since *COOLAIR* upregulation in Col*FRI* and Edi-0 respond the same to FS, it is likely to be a general phenomenon for rapid-vernalizing genotypes (Fig. 4-7C).

We wondered if we could observe the increase of transcripts per focus in a mutant with an expanded expression zone of *COOLAIR* (Fig. 4-7B). *ntl8-D3* expresses a protein that is prematurely truncated, thus lacking its transmembrane domain. This leads to a relocation of the protein into the nucleus constitutively overexpressing *COOLAIR* in all tissues of the root as shown by smFISH (Zhao et al., 2021). Surprisingly, the active transcription sites do not seem to vary immensely in size between vernalisation treatments (Fig. 4-7D). Instead, we observed an increase in the number of smaller foci per cell (Fig. 4-7E). This could suggest that once there are too many molecules in a focus, they disassociate and accumulate to additional, smaller foci detected in the nucleus (not colocalising with an active transcription site).

The increased *COOLAIR* FISH signal was confirmed with q-RT-PCR analysis. The two rapid vernalising accessions *ColFRI* and Edi-0, and the *COOLAIR* overexpressing line *ntl8*-D3, additional mutants and natural accessions with differences in *FLC* and *COOLAIR* transcription levels were included (Fig 4-8, Fig. 4-9 for *FLC* and *COOLAIR*, respectively).

The natural accession Var2-6 (Li et al., 2014) carries a G/T *FLC* variant influencing the choice of splice acceptor site in the large *COOLAIR* intron, associated with elevated *FLC* transcription levels. In Var2-6, the distal *COOLAIR* isoform, Class II.iv, is increased, coinciding with increased levels of *FLC* 5'end capped nascent transcript. An additional line carrying a mutation in the Var2--6 Single Nucleotide Polymorphism (SNP259) mutation in the Col*FRI* background, called G259T, was included as well as the low *FLC* expressing Col--0 (no active FRI, no vernalization requirement) and the *FLC* null allele (*FLClean*).

The same trend of *FLC* reduction after respective vernalization treatments was observed in different genotypes even with varying NV levels of *FLC*. Apart from Col*FRI,* spliced *FLC* decreases fastest in FS, with significant differences between NV and vernalization-treated plants for all *FLC* mRNA. The response to CC compared to FS/FM in spliced and unspliced *FLC* is significantly lower in Edi-0, and significantly reduced for spliced *FLC* in G259T. This indicates a faster *FLC* transcriptional repression in those genotypes in the fluctuating temperature conditions; most prominent in FS but weakly observed in FM. Overall, unspliced *FLC* shows similar patterns. Statistical analyses comparing the overall responses between genotypes (Fig. 4-8C), indicated differences of *FLC* reduction in different accessions. These might reflect adaptive variation aligning *FLC* transcription to specific environments.



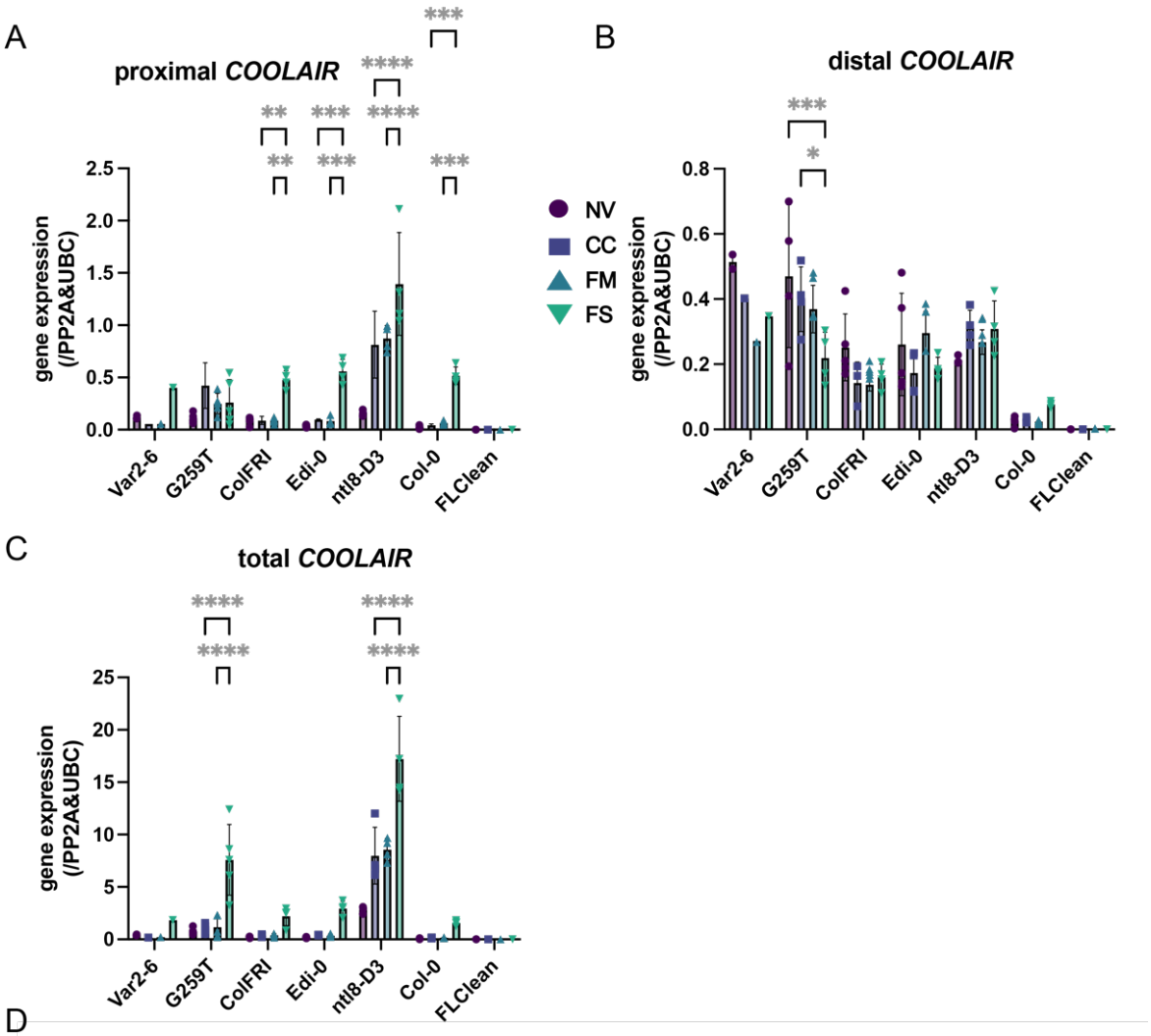
C

	spliced *FLC*			unspliced *FLC*		
Tukey's multiple comparisons test	Predicted (LS) mean diff.	Summary	Adjusted P Value	Predicted (LS) mean diff.	Summary	Adjusted P Value
G259T vs. ColFRI	8.766	****	<0.0001	0.05642	****	<0.0001
G259T vs. Edi-0	1.722	ns	0.8941	0.002039	ns	>0.9999
G259T vs. ntl8-D3	21.6	****	<0.0001	0.05451	****	<0.0001
G259T vs. Col-0	23.98	****	<0.0001	0.08084	****	<0.0001
G259T vs. FLClean	24.5	****	<0.0001	0.08238	****	<0.0001
ColFRI vs. Edi-0	-7.044	***	0.0003	-0.05439	****	<0.0001
ColFRI vs. ntl8-D3	12.84	****	<0.0001	-0.001917	ns	>0.9999
ColFRI vs. Col-0	15.21	****	<0.0001	0.02442	*	0.0265
ColFRI vs. FLClean	15.73	****	<0.0001	0.02596	ns	0.179
Edi-0 vs. ntl8-D3	19.88	****	<0.0001	0.05247	****	<0.0001
Edi-0 vs. Col-0	22.25	****	<0.0001	0.0788	****	<0.0001
Edi-0 vs. FLClean	22.77	****	<0.0001	0.08034	****	<0.0001
ntl8-D3 vs. Col-0	2.373	ns	0.6959	0.02633	*	0.0127
ntl8-D3 vs. FLClean	2.892	ns	0.8439	0.02787	ns	0.121
Col-0 vs. FLClean	0.5191	ns	>0.9999	0.001539	ns	>0.9999

Figure 4–8 Freezing night temperatures (FS) accelerate FLC reduction in most genotypes.

Gene expression of *FLC* spliced (left) and unspliced (right) in several genotypes exposed to different vernalization treatments.

(A) *FLC* spliced **(B)** *FLC* unspliced. 2-way ANOVA with multiple comparisons between temperature treatments was performed followed by a Tukey's post hoc test. Note, for *FLC* spliced, statistical analysis within genotypes is only plotted for comparison between cold-treated samples, as all were significantly different to NV after 2 weeks of cold. **(C)** Statistical analysis (ANOVA with subsequent Tukey's multiple comparison test) comparing if the overall response on gene expression is different from one genotype to another used in this chapter. Asterisks indicate level of significance with p: ***< 0.001, **< 0.01, *<0.05. Error bars represent the SEM with n=4 (Except for Var2-6 with n=1, which is included in the graph for completeness but due to germination issues, were not able to perform all required biological replicates).



D

Tukey's multiple comparisons test	proximal *COOLAIR* Predicted (LS) mean diff.	Summary	Adjusted P Value	distal *COOLAIR* Predicted (LS) mean diff.	Summary	Adjusted P Value	total *COOLAIR* Predicted (LS) mean diff.	Summary	Adjusted P Value
G259T vs. ColFRI	0.07607	ns	0.76	0.1915	****	<0.0001	1.856	**	0.0031
G259T vs. Edi-0	0.06767	ns	0.8392	0.1348	***	0.0003	1.683	**	0.0088
G259T vs. ntl8-D3	-0.5493	****	<0.0001	0.08891	ns	0.0588	-6.494	****	<0.0001
G259T vs. Col-0	0.09893	ns	0.4838	0.3261	****	<0.0001	2.176	***	0.0003
G259T vs. FLClean	0.2617	*	0.0276	0.3646	****	<0.0001	2.657	**	0.0082
ColFRI vs. Edi-0	-0.008392	ns	>0.9999	-0.05664	ns	0.4422	-0.1721	ns	0.9998
ColFRI vs. ntl8-D3	-0.6254	****	<0.0001	-0.1025	*	0.0164	-8.349	****	<0.0001
ColFRI vs. Col-0	0.02287	ns	0.9996	0.1347	***	0.0004	0.3203	ns	0.9946
ColFRI vs. FLClean	0.1856	ns	0.2761	0.1732	**	0.0035	0.8011	ns	0.9302
Edi-0 vs. ntl8-D3	-0.617	****	<0.0001	-0.0459	ns	0.7087	-8.177	****	<0.0001
Edi-0 vs. Col-0	0.03126	ns	0.9972	0.1913	****	<0.0001	0.4924	ns	0.9496
Edi-0 vs. FLClean	0.194	ns	0.2247	0.2298	****	<0.0001	0.9732	ns	0.8385
ntl8-D3 vs. Col-0	0.6482	****	<0.0001	0.2372	****	<0.0001	8.67	****	<0.0001
ntl8-D3 vs. FLClean	0.811	****	<0.0001	0.2757	****	<0.0001	9.15	****	<0.0001
Col-0 vs. FLClean	0.1627	ns	0.4351	0.03848	ns	0.9746	0.4808	ns	0.9947

Figure 4–9 Freezing night temperatures (FS) superinduce total *COOLAIR*.

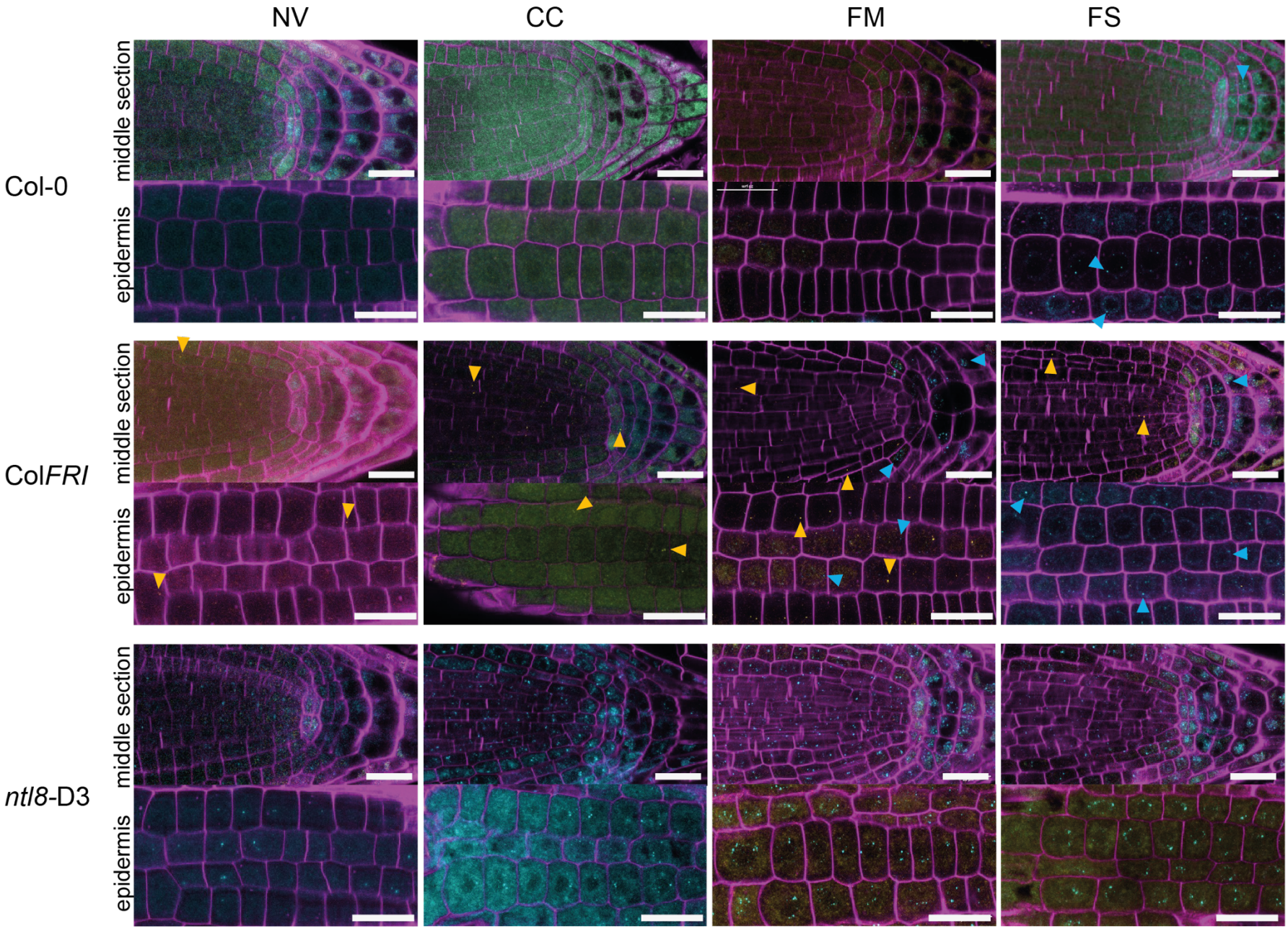
Gene expression of *COOLAIR* isoforms in several genotypes exposed to different vernalization treatments. **(A)** proximal *COOLAIR* **(B)** distal *COOLAIR*. **(C)** total *COOLAIR*. 2-way ANOVA with multiple comparisons between temperature treatments was performed followed by a Tukey's post hoc test. Note, for proximal and total *COOLAIR*, statistical analysis within genotypes is only plotted for comparison between cold-treated samples, as all were significantly different to NV after 2 weeks of cold. **(D)** Statistical analysis (ANOVA with subsequent Tukey's multiple comparison test) comparing if the overall response on gene expression is different from one genotype to another used in this chapter. Asterisks indicate level of significance with p: ***< 0.001, **< 0.01, * <0.05. Error bars represent the SEM with n=4 (Except for Var2-6 with n=1, which is included in the graph for completeness but due to germination issues, were not able to perform all required biological replicates).

I analysed whether the faster reduction of *FLC* was associated with an increase in *COOLAIR* (Fig. 4-9). For most genotypes –except for Var2-6, G259T and the negative control *FLClean*– I observed a significant upregulation of proximal *COOLAIR* in FS. Interestingly in *ntl8*D3, I see an additive effect of the vernalization treatment on antisense transcription as suggested by the smFISH results (Fig. 4-7B), with total and proximal *COOLAIR* levels increasing further (Fig. 4-9A-C). Also, the summer-annual, rapid cycling life cycle accession Col-0 shows that even if *FLC* is silenced prior to FS exposure, proximal *COOLAIR* antisense transcription is significantly up-regulated, with a similar trend also visible for distal and total *COOLAIR*. This suggests that the transcriptional push for antisense transcription at *FLC* can overrule a previously silenced locus. Compared to the other genotypes, I observed a significantly different response in genotypes that impacted RNA stability (G259T, total *COOLAIR*) or showed higher antisense levels (*ntl8*D3, for proximal and total *COOLAIR*) (Fig. 4-9D). Overall, although *COOLAIR* upregulation itself does not depend on the previous transcriptional state in most genotypes the extent of its upregulation does.

The extremely high expression of sense (G259T) and antisense (*ntl8*D3) transcription was then exploited to investigate further the spatial and cellular changes after vernalization treatments. For this purpose, whole mount FISH (Zhao et al., 2023) was used to see if the previous conclusion: cold induction was through a combination of increased *COOLAIR* per cell and a wider expression range, was confirmed.

4.6 Whole tissue smFISH confirms wider tissue response

Recent advances in tissue clearing enabled me to perform whole mount smFISH, with cellular resolution in intact tissue (Zhao et al., 2023). After 2 weeks of clearing in ClearSee (Kurihara et al., 2015), the strong background previously observed in cells was strongly reduced giving the sensitivity to observe intronic *COOLAIR* (cyan) and intronic *FLC* (yellow) signal in tissue including the epidermis, initials quiescent centre, initials, root cap, prevasculature (Fig. 4-10). In NV, the low *FLC* expressing Col-0 accession did not show any intronic *FLC* signal in any cell types, whilst the higher expressing Col*FRI* and Edi-0 showed some signal in all cell types. The G259T mutant showed the highest intronic *FLC* expression with almost every cell exhibiting *FLC* active transcription sites. Surprisingly, *COOLAIR* was also detected in the genotypes that show increased *FLC* expression in columella cells, a specific cell type in the root cap. *COOLAIR* transcription was observed in all cells in *ntl8*-D3 and a similar expression range was observed in G259T. Overall, this is in line with *COOLAIR* transcription in NV reflecting the transcriptional state of *FLC*.


NV
CC
FM
FS
Col-0
ColFRI
ntl8-D3
middle section
epidermis

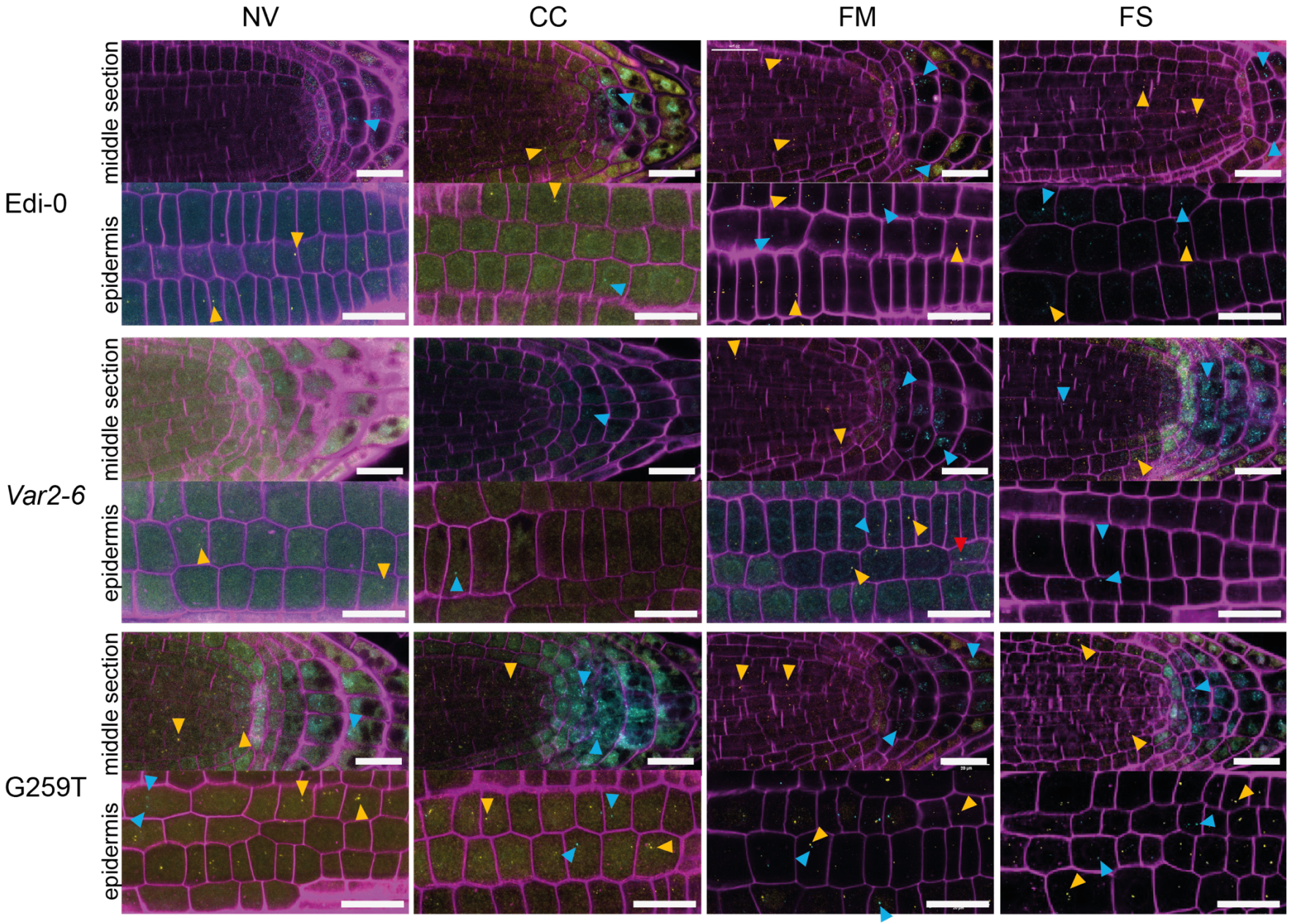

NV
CC
FM
FS
Edi-0
Var2-6
G259T
middle section
epidermis

Figure 4–10 *COOLAIR* expression zone expands in vernalization-treated roots and reveals specific *COOLAIR* cells. (*previous 2 pages*)

Representative images of whole mount smFISH in middle section (including the columella, QC (Quiescent Centre), initials and prevasculature) and epidermis of several genotypes after vernalization-treatment (NV- not vernalized, CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong). Yellow arrows point to representative foci of yellow intronic *FLC* signal, cyan arrows point to representative foci of cyan intronic *COOLAIR* signal, red arrow points to co-localisation of intronic *FLC* and *COOLAIR* in Var2-6. The cell wall is stained with SCI Renaissance 2200 stain shown in magenta. n=2 independent biological replicates with 3 roots per replicate, scale bar 20 μm.

In CC, in line with what has been previously reported (Ietswaart et al., 2017; Rosa et al., 2016), I observed *COOLAIR* foci in occasional prevasculature cells. However, the accumulation of *COOLAIR* was more prominent in the columella cells. No *FLC* transcript was detected resulting in a spatial separation between *COOLAIR* and *FLC*-expressing cells in CC in the root. *COOLAIR* expression was rarely observed in the epidermis in CC in most genotypes, apart from in *ntl8*-D3 where the epidermal signal increased in FM and even further in FS (see blue arrow). This confirmed my previous observations of an expanded expression zone in separated cells (Fig.4-6).

We had previously concluded that the FS *COOLAIR* increase included an increase in transcripts per foci. As previously observed in *ntl8*-D3, instead of *COOLAIR* foci increasing in size, more foci per cell were observed (similar in Fig.4-7), contributing to the observed increase in *COOLAIR (*Fig.4-9C).

In the G259T genotype the pattern of *FLC*/*COOLAIR* signal illustrated the mutual exclusivity of transcription at the locus in epidermal cells; the foci did not co-localize but rather shifted in fraction, whilst in NV, most foci were *FLC,* in FS most foci were *COOLAIR*.

Overall, by looking at cellular context and maintaining subcellular resolution we were able to observe an increase in cell types expressing *COOLAIR*.

4.7 Specific *COOLAIR*-expressing cells in shoot apex

Since FLC functions in the transition to flowering, I tested the cell-specific response to cold in the Shoot Apical Meristem (SAM) (Fig. 4-11). Clearing of the SAM of Col*FRI* carrying an FLC-Venus after 2-weeks of vernalization (CC) enabled simultaneous detection of *FLC* mRNA and protein levels. I observed *COOLAIR* and *FLC* FISH signals. L1 central zone cells showed *COOLAIR* in CC. L1 is the uppermost monolayer characterised by predominantly dividing anticlinal to become the epidermis. The cell layers of the central zone maintain a pool of stem cells which divide occasionally and replace cells in the neighbouring peripheral zone and rib meristem. Cells in the central zone divide at a slower rate compared to the peripheral zone, where rapid cell division contributes to cell differentiation and the formation of new plant organs (primordia) (Meyerowitz, 1997; Reddy et al., 2004).

Interestingly, a spatial separation of *FLC* in the primordia (Fig. 4-11, dark grey box) and *COOLAIR* (light grey box) in the central zone of the SAM is reminiscent of the spatial separation in *COOLAIR*-expressing cells in the root - columella compared to other rapidly dividing meristematic cell types (see Fig. 4-10) (Rahni and Birnbaum, 2019). This might indicate that some meristematic cells are responding to environmental cues by promoting *COOLAIR* transcription.

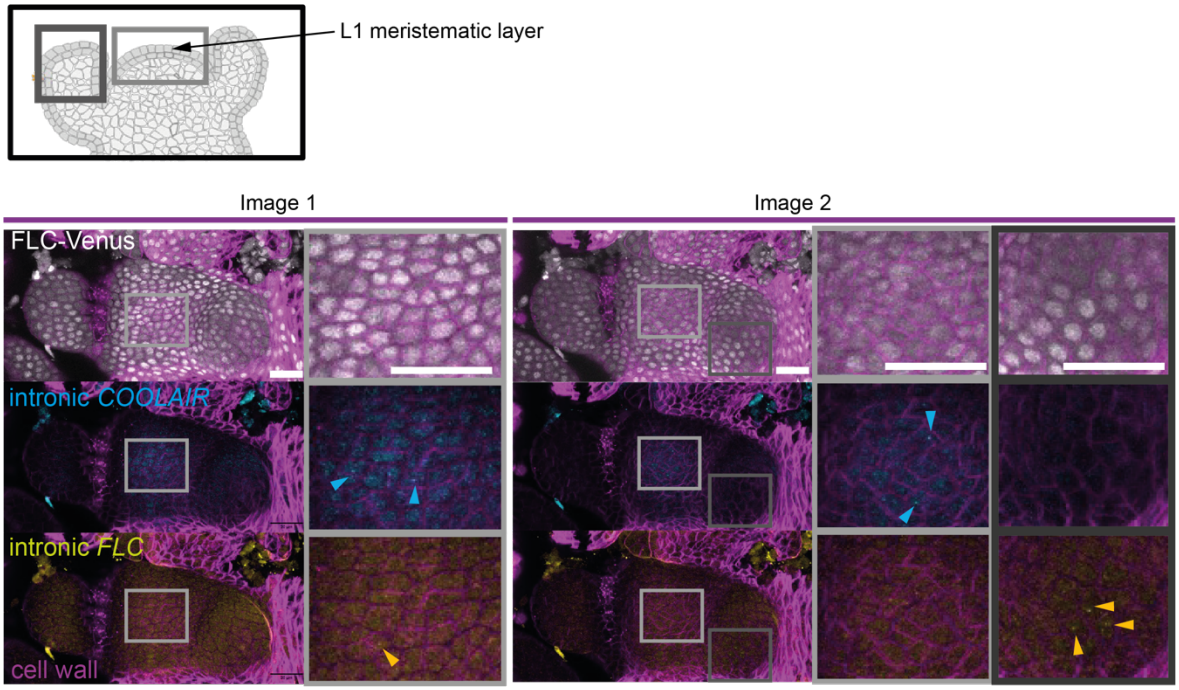


Figure 4–11 *COOLAIR* specific cells identified in apex.

Representative images of FISH in shoot apical meristem (SAM) after 2-weeks in CC in Col*FRI* carrying a FLC-Venus reporter (grey). Black box shows a schematic drawing of the SAM. Dark grey box indicates primordium, light grey box indicates central zone of the SAM. Yellow arrows point to representative yellow intronic *FLC* signal, cyan arrows point to representative cyan intronic *COOLAIR* signal. The cell wall is stained with SCI Renaissance 2200 stain shown in magenta. Images 1 and 2 are each maximum projections of 5 planes of the same z-stack with 500 nm z-distance per image covering 2.5 μm. Acquisition of image 1 started at the outside of SAM covering 0-2.5 μm imaging depth, image 2 covers 2.5-5 μm imaging depth. Scale bar represents 20 μm, n=2 with 10 apices.

This raised two questions:

(1) Does the difference in cell cycle rate between stem cell-like and primordia-forming cell indicate a cell-cycle dependency of *COOLAIR* expression?

(2) Does the expansion of the *COOLAIR* expression zone to primordia-forming (SAM)/ division zone (RAM) in FS account for the faster reduction of *FLC?*

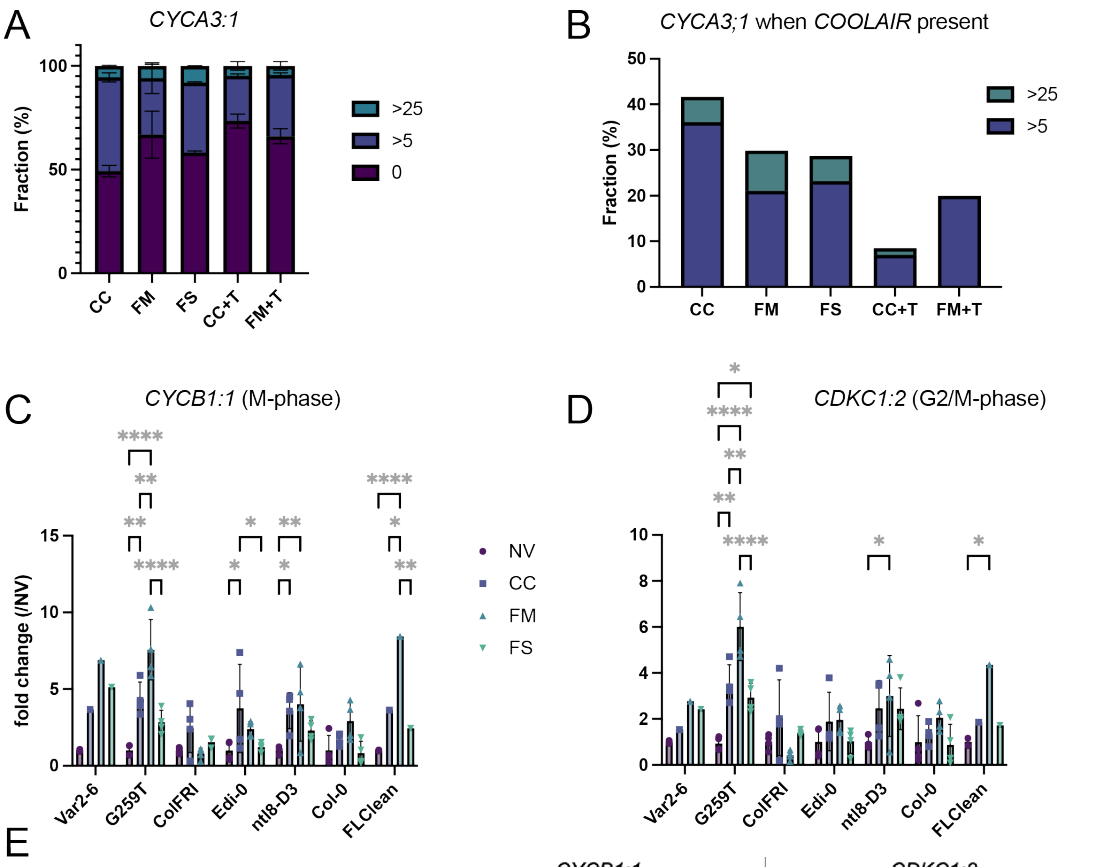
4.8 FS reduces G1 cell cycle phase and undermines *COOLAIR*'s G1-preferential expression

As thermodynamic processes differ depending on temperature, we first wanted to test if FS generally impacts cell cycle progression. If exposure to 0°C induces cell cycle arrest as previously shown in mammalian systems (Rieder and Cole), a synchronization of cells in a specific cell cycle phase might be observed. Interestingly, it has been previously reported in maize leaves that low night temperatures inhibit cell cycle progression (Rymen et al., 2007).

By performing single-molecule FISH (smFISH) using probes for the cell cycle-specific mRNA *CYCA3;1*, peaking from late G1 to early S phase (Cools et al., 2010; Menges et al., 2005), the cells in G1/S can be identified. Using G1/S as a proxy for the impact on cell cycle fractions, I analysed differences between vernalization treatments (Fig. 4-12A). *CYCA3;1* mRNA counts were classified in low and high number of transcripts per cell (>5 mRNA molecules, >25 mRNA molecules). The fraction of cells with *CYCA3;1* is slightly reduced in FS and FM. This reduction was mostly in the fraction of cells expressing >5 molecules, while the fraction of cells with > 25 mRNA molecules was unchanged or reduced. One night of triggering freezing had the strongest effect on *CYCA3;1* reduction when transferred from CC. Overall, the reduction in *CYCA3;1* suggests that the fluctuations and the freezing downregulate *CYCA3;1* abundance. Further, the stronger reduction in triggered plants suggests that several nights of FS allow the cell cycle to adapt to the freezing.

We have previously shown that *COOLAIR* is predominantly expressed in G1/S-phase of the cell cycle in Col*FRI* in NV conditions (personal communications with Congyao Xu [JIC, now Senior Editor at Nature Communications]). I wondered if *COOLAIR* is associated with G1/S during vernalization treatments as previously observed for NV. I performed multiplex smFISH in Col*FRI* with *CYCA3;1* and intronic *COOLAIR* probes and counted the overlap of G1/S marker and any *COOLAIR* signal (Fig. 4-12B). Whilst with vernalization-treatment, *COOLAIR* levels rise, the cells exhibiting *CYCA3;1* and *COOLAIR* decrease. This suggests that in the vernalization-treated cells, the cell-cycle preference for *COOLAIR* transcription in G1/S-phase is not maintained.

Further, the decrease of cells at G1-phase suggests that they might arrest at another cell cycle stage. To test this, I performed gene expression analysis using *CYCB1;1* specific for M-Phase and *CDKC1:2* specific for G2-M-phase (Fig. 4-12CD). By additionally including different genotypes, similar patterns between vernalization-treatments could be assessed. *CYCB1;1* showed an increase in vernalization-treated compared to NV plants in most genotypes. This was most striking after the FM treatment, apart from Col*FRI* and Col-0, in which no effect in cell cycle marker was observed (Fig. 4-12C). *CDKC1:2* did not show an observable trend between genotypes.



E

Tukey's multiple comparisons test	CYCB1:1 Predicted (LS) mean diff.	CYCB1:1 Sum mary	CYCB1:1 Adjusted P Value	CDKC1:2 Predicted (LS) mean diff.	CDKC1:2 Sum mary	CDKC1:2 Adjusted P Value
G259T vs. ColFRI	-2.345	****	<0.0001	2.131	****	<0.0001
G259T vs. Edi-0	-0.4325	ns	0.9076	1.874	****	<0.0001
G259T vs. ntl8-D3	-1.231	*	0.0317	1.12	*	0.0215
G259T vs. Col-0	-1.53	**	0.0024	2.027	****	<0.0001
G259T vs. FLClean	-2.251	**	0.0031	1.11	ns	0.275
ColFRI vs. Edi-0	1.913	***	0.0001	-0.2571	ns	0.9867
ColFRI vs. ntl8-D3	1.114	ns	0.0801	-1.011	ns	0.0596
ColFRI vs. Col-0	0.8154	ns	0.3465	-0.1042	ns	>0.9999
ColFRI vs. FLClean	0.09404	ns	>0.9999	-1.021	ns	0.3844
Edi-0 vs. ntl8-D3	-0.7988	ns	0.3711	-0.7542	ns	0.2769
Edi-0 vs. Col-0	-1.098	ns	0.0672	0.1529	ns	0.9991
Edi-0 vs. FLClean	-1.819	*	0.0304	-0.7634	ns	0.7067
ntl8-D3 vs. Col-0	-0.2988	ns	0.9858	0.9071	ns	0.1092
ntl8-D3 vs. FLClean	-1.02	ns	0.5529	-0.009251	ns	>0.9999
Col-0 vs. FLClean	-0.7213	ns	0.8561	-0.9163	ns	0.5049

Figure 4–12 Cell cycle does not affect *COOLAIR* transcription during vernalization-treatments.

(A-B) smFISH quantification *CYCA3;1* (G1/S-phase mRNA marker) and intronic *COOLAIR* in vernalization-treated plants (CC - constant 5°C, FM - fluctuating mild, FS - fluctuating strong, FM+T - fluctuating mild + one night of FS trigger). **(A)** Fraction of cells expressing 0, >5 or >25 mRNA molecules of *CYCA3;1* **(B)** Fraction of cells expressing >5 or >25 mRNA molecules of *CYCA3;1* while also showing intronic *COOLAIR* signal **(C-D)** Gene expression of **(C)** *CYCB1:1* (M-phase) and **(D)** *CDKC1:2* (G2/M-phase) in different genotypes with different vernalization treatments as indicated by colour map. 2-way ANOVA with multiple comparisons between temperature treatments was performed followed by a Tukey's post hoc test. **(E)** Statistical analysis (ANOVA with subsequent Tukey's multiple comparison test) comparing the response of genotypes to vernalization treatments of mutants to each.

Asterisks indicate level of significance with p: ***< 0.001, **< 0.01, * <0.05. Error bars represent the SEM with n=4 (Except for Var2-6 with n=1, which is included in the graph for completeness but due to germination issues, were not able to perform all required biological replicates).

Only G259T showed significant changes in gene expression for all vernalization-treatments, as well as *ntl8*-D3 and *FLClean* between NV and FM (Fig. 4-12C). The lack of significant differences in Col*FRI* and Col-0 suggests that there is not a major effect on the cell cycle length of the G2/M phase at the time of the analysis.

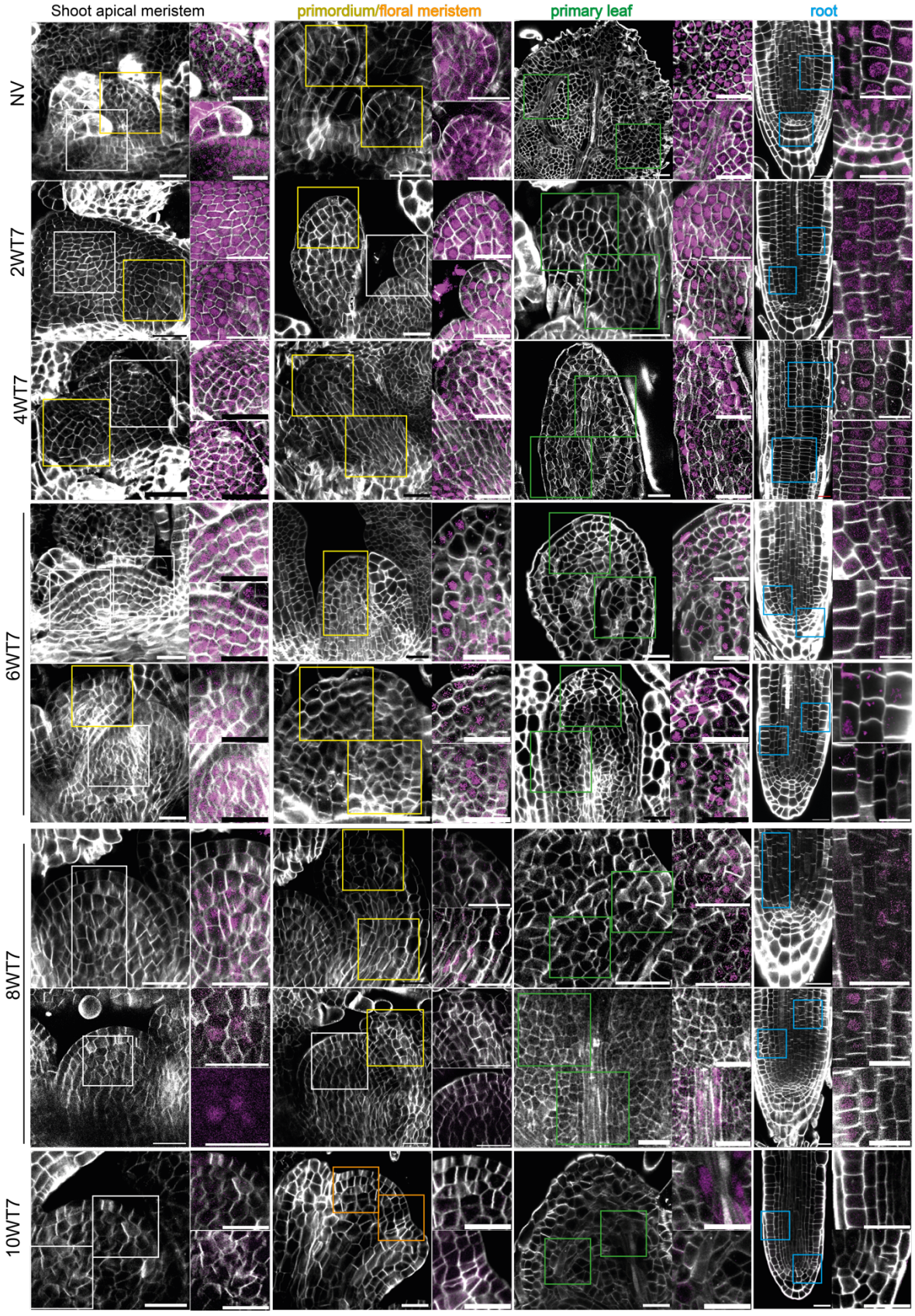
Taken together, there is a slight trend of cells having a shorter G1/S-phase and potentially a longer G2/M phase. Moreover, changes in cell cycle length, as estimated by the abundance of cells at a specific cell cycle stage, do not appear to drive the increase in *COOLAIR*.

4.9 *COOLAIR*-expressing cells might switch off later in CC

To answer which cells in the meristem are switching-off *FLC* first and potentially link switching to the abundance of *COOLAIR* in the plant, I performed time course imaging of an FLC-Venus fluorescent protein reporter line in different tissues of the plant at CC (Fig. 4--13, SAM [white box]; primordium [yellow box]; floral meristem [after the transition to flowering, [orange box]; primary leaf [green box]; root [blue box]. For the time course of NV-10 weeks vernalization, the imaging was performed 7 days after return to ambient temperatures [20°C]). FLC expression switched off in all tissues of the plant, confirming the population-level measuring of time when exposed to cold. By 8-weeks of vernalization, meristem transitioning to flowering was observed, as indicated by the elongation of the apex. After 10-weeks of vernalization, the meristem had transitioned into a floral meristem. The SAM (including the cells showing *COOLAIR* signal in CC) maintained FLC-Venus signal the longest when compared to primordia, primary leaves, and RAM. Whilst after six weeks of vernalization and 7 days post growth in ambient temperature (6WT7), most SAM cells still show nuclear FLC-Venus signal (white boxes). Other tissues, such as primary leaves and roots only showed occasional cells with signal. This suggests that *COOLAIR* abundance (as measured by intronic smFISH probes) might not influence the probability of Polycomb-mediated switching in CC.

Figure 4–13 FLC-Venus switches off later in meristems (*next page*)

Time course imaging of FLC-Venus (magenta) in Col*FRI* of different tissue types after being exposed to different lengths of constant 5 °C (CC). Shoot apical meristem (SAM, white box), primordium (yellow box) floral meristem (orange box), primary leaf (green box), root (blue box). The cell wall is stained with SCI Renaissance 2200 stain shown in grey. Imaging was performed 7 days after return to ambient temperatures (22°C) to avoid imaging Venus-artefacts as previously reported in Lövkvist et al. (2021). Scale bar represents 20 μm.


Shoot apical meristem
primordium/floral meristem
primary leaf
root
NV
2WT7
4WT7
6WT7
8WT7
10WT7

Discussion

In this chapter, I set out to characterize how the upregulation of *COOLAIR* in different vernalization-treatments can be accounted for on a cellular level. I have shown that 2 weeks of FS expanded the expression area as well as *COOLAIR* transcript number per active transcription site (brightest foci). The reoccurring nightly freezing temperatures observed in FS are required for a reduction of *FLC* during early vernalisation. Sense transcription is negatively correlated with antisense abundance, suggesting that FS may accelerate *FLC* silencing in a wider expressing zone than previously reported (Duncan et al., 2016; Ietswaart et al., 2017) as observed by an increase of *COOLAIR* expression in separated or epidermal cells. This was most obvious in the outer epidermal cells due to a reduction in imaging sensitivity with increased tissue depths. Nevertheless, similar behaviour was observed in images from the middle section of the root including other cell types, such as cortex and endodermis. Interestingly, the foci number per cell in FS does not appear to be affected compared to CC. Considering that FS reduces *FLC* levels faster, an earlier time point would be a sensible addition to observe *COOLAIR* foci to account for the possibility of a faster response of *COOLAIR* in FS.

With the described smFISH method, I have shown that *COOLAIR* transcripts covering the whole locus accumulate. Their colocalization with the *FLC* locus and upregulation in FS suggest a cis-regulatory role for *COOLAIR* in *FLC* regulation. Whole-mount imaging confirmed the expansion of *COOLAIR*-expressing cells, as we observed more signal in epidermal cells. The spatial separation of *FLC* in the primordia/fast dividing tissue and *COOLAIR* in the SAM/stem cells is in line with this, suggesting that the potential function of *COOLAIR* in antagonising sense transcription is most important in stem cells. The increase of intronic *COOLAIR* is independent of the prior transcription state as indicated by the strong upregulation of *COOLAIR* in genotypes that have very low *FLC* levels prior to vernalization requirement (Col-0). The ability of plants to respond to both acute and ambient cold temperatures may be crucial for their adaptation to different climates. We observed differences in *COOLAIR* upregulation in Var2-6 NIL (Near Isogenic Line) and G259T, suggesting that plants have evolved to modulate transcriptional states to fine-tune flowering time (Coustham et al., 2012b; Hepworth et al., 2020; Qüesta et al., 2020; Shindo et al., 2006). In most mutants, the pattern of *FLC*/*COOLAIR* signal (intronic probes) beautifully illustrates the mutual exclusivity of sense/antisense transcription at the locus. Most prominent in G259T, the foci did not colocalize upon FS but rather shifted in fraction from *FLC* in NV to *COOLAIR* in FS. Potentially, transcription is driven from one to the other strand. As previously observed, FS induced a faster reduction of *FLC*. How this links to Polycomb nucleation remains to be fully established, and the current view is that *COOLAIR*-induced transcriptional down-regulation and nucleation of Polycomb silencing are parallel pathways interconnected by control through NTL8 (Nielsen et al.). I view

COOLAIR transcription as another layer for fine-tuning the response to the environment. This could be especially relevant in the context of climate change where increased temperature fluctuations are predicted. In the future, there will be increasing pressure for plants to correctly integrate these environmental cues and ensure reproductive success.

Total RNAs obtained from Quant-Sequencing, a technique where RNA molecules are captured by their 3' poly(A) tail (Moll et al., 2014), revealed that there is mainly proximally poly-adenylated *COOLAIR* accumulating at the chromatin (internal communications with Minglei Yang). In contrast to the Quant-Sequencing data, my FISH data showed that the *COOLAIR* foci at the locus consist mainly of intronic and exonic *COOLAIR* signal which could indicate that it is full-length transcript. Polymerase 2 (Pol II) pausing or polyadenylation does not automatically mean Pol II is terminated and released. Tang et al. (2022) have identified several genes in human cell lines where the distally poly-adenylated transcripts are retained at the locus, and the proximal poly(A) form is released from the locus (Tang et al., 2022). We can hypothesise that a similar event is taking place at *COOLAIR:* Although the proximally polyadenylated *COOLAIR* transcript is released from the locus, the intron-containing transcripts are retained at the chromatin, detectable by smFISH intronic probes. That would additionally explain why we do not detect nucleoplasmic/cytoplasmic RNAs with exonic probes, as the number of probes binding to proximal *COOLAIR* is too low to be detectable as diffraction-limited foci. Our use of alternative image-based methods for detecting specific *COOLAIR* isoforms have had poor success. A similar amplification system to SABER, called ClampFISH (Rouhanifard et al., 2019) using a padlock system with circular DNA backbone that covalently locks when bound to its target RNA also resulted in non-specific labelling (internal communications with Stefanie Rosa and Miguel Coelho Montez). Yet, amplification-based systems have a massive potential for distinguishing between isoforms and will potentially be relevant in elucidating co-transcriptional regulation by lncRNAs in the future as shown by the recent advances in performing 3D multiplexed spatial gene expression analysis using PHYTOMap in the *Arabidopsis* root (Nobori et al., 2023).

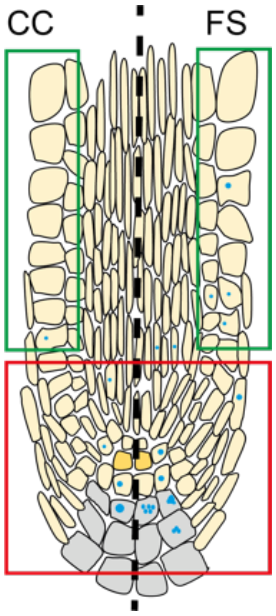


Figure 4–14 *COOLAIR* expression zone expands when exposed to fluctuating freezing conditions.
Graphical summary of the findings in this Chapter. Schematic model of the middle section of the *Arabidopsis* root comparing constant cold (CC) and fluctuating strong (FS) vernalization-treatments. *COOLAIR* signal indicated by cyan (blue) foci. Boxes highlight different tissues (green: epidermis, red: core meristem with cells in dark yellow: QC, grey: Columella).

Nevertheless, the use of intronic smFISH probes to investigate *COOLAIR* upregulation in the strong fluctuating cold including freezing temperatures revealed that this is accounted for by a wider expression zone as observed in separated and epidermal cells with more transcript per active transcription site. This does not appear to be connected to the cell cycle. Although *COOLAIR*-upregulation correlates with a faster repression of *FLC,* the different spatial distribution of cells suggests that may function in parallel to Polycomb-nucleation.

CHAPTER 5 : GENERAL DISCUSSION

Analogue and digital gene regulation for fine-tuning transcriptional states

At *FLC*, several players contribute to its complex, multi-layered and finely-tuned regulation.

The opposing autonomous and FRIGIDA pathways first set up a basic transcriptional level most likely during germination (Chapter 3). The vernalization pathway subsequently represses *FLC* by first involving antisense transcription and then epigenetically silences it, mediated by the Polycomb Repressive Complex II.

In this thesis, I have shown that there are two modes of gene regulation integrated to determine the plants' reproductive strategy: analogue and digital. Of course, there are other known influences that impact the transcriptional state of the *FLC* locus: DNA-sequence modifications affect the basal levels (as seen in the natural variations), whereas the cold acts via the vernalization pathway. Plant age can circumvent the vernalization pathway by activating floral integrators to transition to flowering independent of *FLC* levels (Teotia and Tang, 2015). However, we have shown here that *FLC* levels can also decrease digitally over time in a genotype with intermediate expression levels *fca-3* (Chapter 2), suggesting that there are transcriptional state transitions, from an active to a repressed state, at *FLC*. Chromatin state transitions at *FLC* have been previously associated with cold exposure, where the time spent in the cold is measured by stochastic switching of more loci from an ON to an OFF state (digital) leading to a collective population level counting of time exposed to cold (Angel et al., 2015). In this thesis, I demonstrated that in the intermediate *fca-3* and *fca-4* mutants, these chromatin switches occur even in ambient conditions (20°C), unrelated to cold exposure.

Using *fca-3* and *fca-4*, examination of quantitative differences in *FLC* levels allowed us to systematically dissect the interplay of analogue (graded) and digital (ON or OFF) regulation at *FLC* before cold concluding that the analogue precedes the digital mode (Chapter 2). Furthermore, consistent with the one-directional nature of vernalization-induced switching, our findings indicate a similar unidirectional pattern depending on the analogue transcriptional state of the *FLC* locus (Chapter 2&3). PRC2-silencing is opposed by active transcription (Berry et al., 2017) and depends on the trans factors present, such as

FCA (Chapter 3), resulting in analogue transcriptional levels controlling the timescale of the switch to the subsequent digital, Polycomb-mediated silencing.

While the purpose of the switch OFF in the cold is to register and store information about noisy environmental conditions to align the transition to flowering with spring, any purpose in the artificial system of allelic mutations in the *FCA* protein is not so obvious. Furthermore, it is important to note that there are no naturally occurring allelic variations in genes involved in the autonomous pathway. Indeed, its highly conserved nature suggests its functional importance. Instead, there is genetic variability in FRIGIDA, which also results in shifting the intricate balance between activating and repressing pathways vital for setting up analogue levels. Nevertheless, study of the *fca* alleles has allowed us to gain insight into autonomous pathway function and to understand in general how analogue and digital modes of gene expression interact. We also note that analogue levels leave room for flexibility in response to the environment, while once *FLC* is PRC2-silenced, this state seems to be stably inherited.

Transcriptional response at *FLC* in early cold

Stochastic PRC2-silencing results in a flexibility in switching that is strategically valuable as it allows the plant to filter noisy temperature cues, generating stable quantitative silencing of *FLC* even in a highly fluctuating temperature environment. The vernalization-induced environmental switch to a Polycomb silenced memory state is digital, but little is known about earlier responding components in the cold. In the early cold, there is a lagging phase in which digital switching occurs only infrequently (Angel et al., 2015). During that first response, it has been shown that early cold reduces *FLC* expression independently from Polycomb switching (VIN3-independent, VRN2-independent).

Interestingly, this early repression is not due to a hyper-functional autonomous pathway but rather due to cold-induced factors such as NTL8, CBFs or WRKY63 contributing to drive *COOLAIR* transcription in a concentration-dependent manner in response to cold exposure (Hung et al., 2022; Jeon et al., 2023; Zhao et al., 2020a). This concentration-dependent increase could be considered an analogue induction of *COOLAIR* with prolonged cold. Given the mutual exclusivity of either *FLC* or *COOLAIR* transcription at a given time at the *FLC* locus (Rosa et al., 2016), we could hypothesise that by activating *COOLAIR*, also the analogue *FLC* levels are modified. Yet, this does not directly lead to digital switching as found in the case of *fca-3* and *fca-4*, as transcription of any kind (including antisense) interferes with Polycomb repression. Rather, the *COOLAIR*-mediated pathway sets up a chromatin environment that could facilitate a digital switch by

disrupting a transcriptional chromatin loop (Crevillén et al., 2013) and removing H3K36me3 from the transcription start site of *FLC* (Csorba et al., 2014).

COOLAIR's function has been widely debated in the field. The complex intersection of sense and antisense transcription makes it very difficult to adequately distinguish their differing roles, as each tends not to come without the other, comparable to the chicken or egg paradox. There are several examples of cases during vernalization in which the transcript has an active role in affecting *FLC* levels e.g. by sequestering FRI away from the promoter (Zhu et al., 2021) or in which the act of transcription plays a role in impacting the chromatin environment (Wang et al., 2014). The latter often involves autonomous pathway proteins in promoting a weak proximal splicing site.

In the mutant line *ntl8*-D3, both *COOLAIR* and the PRC2 accessory protein VIN3 are ectopically expressed (Zhao et al., 2020a) accompanied by an accumulation of H3K27me2 modification but not the H3K27me3 at the nucleation region (Nielsen et al.) required for Polycomb switching. In Chapter 4, I showed that in severe fluctuating (FS) temperature conditions ranging from 12 -1°C, *FLC* unspliced is upregulated indicating that the presence of antisense transcription alone is not sufficient for switching states.

Furthermore, in the cold, I identified a specific cell type in the root tip, the columella cells, that hyperaccumulate *COOLAIR* (Chapter 4) known to be involved in gravity sensing (Yoder et al., 2001). Their location at the tip of the root is to protect and shield the root apical meristem, and to guide it. As they are often the first contact point to the soil, it is important that they are fast responders to environmental cues. At *FLC,* the accumulation of *COOLAIR* might be a way of maintaining an open chromatin and therefore to prevent silencing of *FLC*.

Overall, a picture is starting to emerge in which *COOLAIR* can be a fast responder to a changing environment, thereby aiding in the modulation of a transcriptional state, while at the same time buffering the system from switching too fast. In Chapter 4, I have shown that this can go either way by preventing early switch-off in a milder fluctuating environment, thereby buffering the system or promoting a faster switch-off as in the case of freezing temperature fluctuations. This response is mediated by an analogue and digital components. Taken together, the **theoretical benefit** of a combination of digital and analogue modes of gene regulation, allows plants to buffer environmental signals and respond accordingly.

The combination of analogue and digital gene regulation might be more relevant for fine-tuning transcriptional states in plants, because of the constant exposure to changing environments requiring different time scales of feedback regulatory loops. It will be interesting to further unravel differences in *COOLAIR* responses in other natural variants.

In a recent preprint, experimental work in combination with mathematical modelling investigated how the adaptive immune system can integrate information about the nature and severity of a pathogen attack in its memory T cell population. Interestingly, here the cis-epigenetic switch is reversible allowing T cells to adapt to the pathogen behaviour (Abadie et al., 2023). Although by a different mechanism, it also highlights the importance of flexibility and stochasticity in cellular decision-making processes.

FCA mutants in Ler not in Col-0

Most autonomous pathway mutants including the *fca* alleles (*FCA* mutants) with intermediate effects on *FLC,* have been identified in Ler mutants. Interestingly, no intermediate mutants have been detected in any Col-0 screens. Out of curiosity, I backcrossed the *fca-3* mutation into Col-0 and observed that the intermediate effect of *fca-3* on *FLC* transcript levels and flowering time was lost. Instead, *fca-3*$^{Col-0}$ was even later in flowering compared to *fca-9*, the Col-0 null mutant. While *FLC-Venus* mRNA levels in *fca-3*^{Ler} as detected by smFISH had an average of three molecules per cell (Chapter 2, Fig. 2-4), *fca-3*$^{Col-0}$ was indistinguishable from *fca-1*-Ler and *fca-9*. As we used the Col-0 transgene in the Ler background, we can assume that it is not an *FLC*-specific response but rather indicates a difference in co-transcriptional gene regulation between Ler and Col-0, e.g. in the splicing machinery. Additionally, the high levels of mRNA per cell in *fca-3* are similar to the mRNA levels detected in the mature embryo.

The histone demethylase FLD, responsible for the removal of H3K4me1 from the *FLC* locus and indirectly promoting the addition of H3K27me's, has only been shown to interact with the *FLC* locus in Col-0 but not in Ler (Liu et al., 2007; Wu et al., 2016). These considerations suggest that there might be genotype-specific differences in the autonomous pathway players and their redundancy. Nevertheless, the great potential of the natural accessions in Arabidopsis is that they can be used to get to the bottom of the mechanism.

By using in situ hybridization methods, I have unravelled a spatial separation of antisense response at *FLC*. While this is a significant addition to our current understanding of *COOLAIR* function, the question of isoform specificity is still pending. Amplification-based systems have a massive potential for distinguishing between isoforms and will potentially be relevant in elucidating co-transcriptional regulation by lncRNAs in the future as shown by the recent advances in successfully performing 3D multiplexed spatial gene expression analysis using PHYTOMap in the *Arabidopsis* root (Nobori et al., 2023). This will allow us to locate specific *COOLAIR* isoforms by multiplexing several isoform-specific probes and interpret their co-localisation.

While these methods are a powerful tool to understand transcriptional differences in isoforms in different cells, they are based on fixed material which makes inferring transcriptional dynamics more challenging. Alternatively, real-time quantification of transcription by labelling RNA has been successfully used (Alamos et al., 2020; Hani et al., 2021) to reveal single-cell differences in response to the environment. RNAs can be fluorescently labelled using the bacteriophage MS2 or the bacterial transcription antitermination protein BglG, both binding to palindromic stem-loop sequences formed by an RNA secondary structure (reviewed in Urbanek et al., 2014). RNA-binding proteins fused to a detector protein (e.g. GFP) bind to the RNA tags.

I previously tried to visualize single *FLC* RNAs in living cells with the help of Yusheng Zhao. We transformed two different versions with 18x BglG stem-loops inserted in Intron 1 of *FLC*, and in the same construct either Version 1 with 24x MS2 stem loops inserted in Exon 1 or Version 2: with 24x MS2 stem loops fused to Intron 1 and a NOS terminator into *FLClean*. Unfortunately, RNA expression and flowering time estimations showed that none of the single copy lines complemented *FLC* levels, suggesting that the stem loops lead to either heterochromatinization of the locus or immediate degradation of either transcript as previously reported by Pichon et al (2020).

One big drawback of the RNA-binding protein tagged with a fluorophore was the background fluorescence of unbound detector molecules in the cell. As a response to this difficulty, a system arose that uses monovalent and highly specific genetic tags (Snap- and Halo-Tags) for attaching an externally added fluorescent label to a protein of interest (RNA-binding protein). This method allows better fine-tuning of signal/noise because it allows the washing away of unbound fluorophore molecules.

Another potential live imaging approach that would reduce background signal is using RNA aptamers. Similar to RNA-stem loops, they are inserted in the target RNA sequence. Upon contact with a fluorogenic ligand, those aptamers become a fluorescent complex with increased fluorogenic properties (Paige et al., 2011; Wu et al., 2019a).

Apart from the fact that an RNA tag can interfere with the stability of the RNA, these live imaging approaches have the drawback that they are not able to reflect the exact chromatin environment at the endogenous locus because they are transgenes randomly inserted in the genome. As discussed in the sections above, transcriptional regulation at *FLC* is tightly regulated by the local environment. Changes of the insertion sites can have a big impact on how *FLC* is regulated. New approaches such as CRISPR LiveFISH use catalytically inactive dCas9 or dCas13 as binding proteins and short complementary guide RNAs for targeting unlabelled RNA expressed from the endogenous locus (Wang et al.,2019).

In the future, cell biological approaches using advanced imaging techniques will undoubtedly help in understanding transcriptional regulation dynamics. By visualizing RNA in living cells, and isoform-specific labelling we will uncover the precise mechanisms and interactions involved in transcriptional state switching at the *FLC* locus.

Limitatons and potential of the root system for imaging FLC

While FLC expression is not exclusive to roots and may not directly affect flowering there, it is still a useful system for live-imaging due to its transparency and its relatively easy access and its defined cell organisation allowing tracing of heritability in clonal files. Yet, the root is not the primary tissue contributing to a transition to flowering. Therefore, FLC-imaging in the root might not provide a complete picture of its role in the flowering process. Nevertheless, FLC-Venus Protein has been shown to respond in the same manner to cold in primary leaves compared to roots (Chapter2, Fig.2-7).

Imaging a very low expressing reporter-line, such as FLC-Venus in *fca-3* mutant, in green tissues is extremely challenging. In this thesis, we have managed to obtain images of FLC-Venus in fixed young leaf tissues using the tissue-clearing agent ClearSee. These images show that FLC in young leaves does indeed behave similarly to the roots. While the results from fixed images were reassuring, using live-imaging is crucial for understanding the dynamics of the switch. Attempts of using spectral imaging, in which the signal of the fluorophore is separated from the background based on its spectral properties and lifetime, did not deliver sufficient results. Alternatively, a different fluorophore can be used with increased brightness and lifetime.

As shown before in the squashed root cells, *COOLAIR* intronic smFISH signal was found abundantly in a focus in constant 5°C conditions. Whole mount *COOLAIR* imaging using improved clearing protocols identified specific cells in the meristematic regions in root and shoot that accumulate intronic *COOLAIR* signal in constant 5°C conditions, indicating a spatial cell-type specific distribution in both tissues.

Overall, this suggests that the response at the *FLC* locus in the root can be representative of the shoot.

List of priorities

There are still many open questions to be answered. In the following list of future experiments

Table 5–1 List of priorities for future work

Experiment	Expected outcome
Quantification of WM (whole mount) FISH	In these root squashes, exact cell types could not be attributed yet. Using an imaging pipeline that combined 3D root segmentation with foci detection will allow quantification of cells in a cellular context.
Life-imaging in intermediate *fca-3*/*fca-4* mutant using a brighter fluorophore and a reporter-cell wall marker	Using an imaging method that does not require the use of relatively strong lasers during time course imaging. Additionally, a cell wall reporter-line will make the use of genotoxic dyes (although used at low concentrations) redundant. Both will reduce the impact of the imaging conditions on growth autofluorescence which will allow the detection of switching of cells. The use of a lights-sheet microscope will reduce phototoxicity and photobleaching as it only excites as excites a thin plane of the sample at a time.
Confirm analogue/digital in allelic FRI series	There are several naturally occurring mutations of FRI leading to quantitative variability in the *FLC* levels. This tool could be used to confirm that our observed analogue responses is a result of the tug-of-war between the autonomous pathway and FRI, which results in the frequency of a digital switch.
COOLAIR response to fluctuating temperatures in shoot apex tissue	First investigations have shown that *COOLAIR* also accumulates in the shoot apex tissue in the form of bright foci in specific cells in constant cold (5°C). This suggests a function of *COOLAIR* in specific meristematic cells. By further investigating the response to fluctuating temperatures conditions and observing the response in mutant analysis, we could gain a better understanding of the function of the accumulations.
Reducing autofluorescence in embryo by using ClearSee with subsequent mRNA Quantification	Although detectable, the *FLC* mRNA in the embryos was not detectable by our used isolated cell RNA detection pipeline due to poor signal/noise ratio. By using ClearSee, the transparency will be increased.

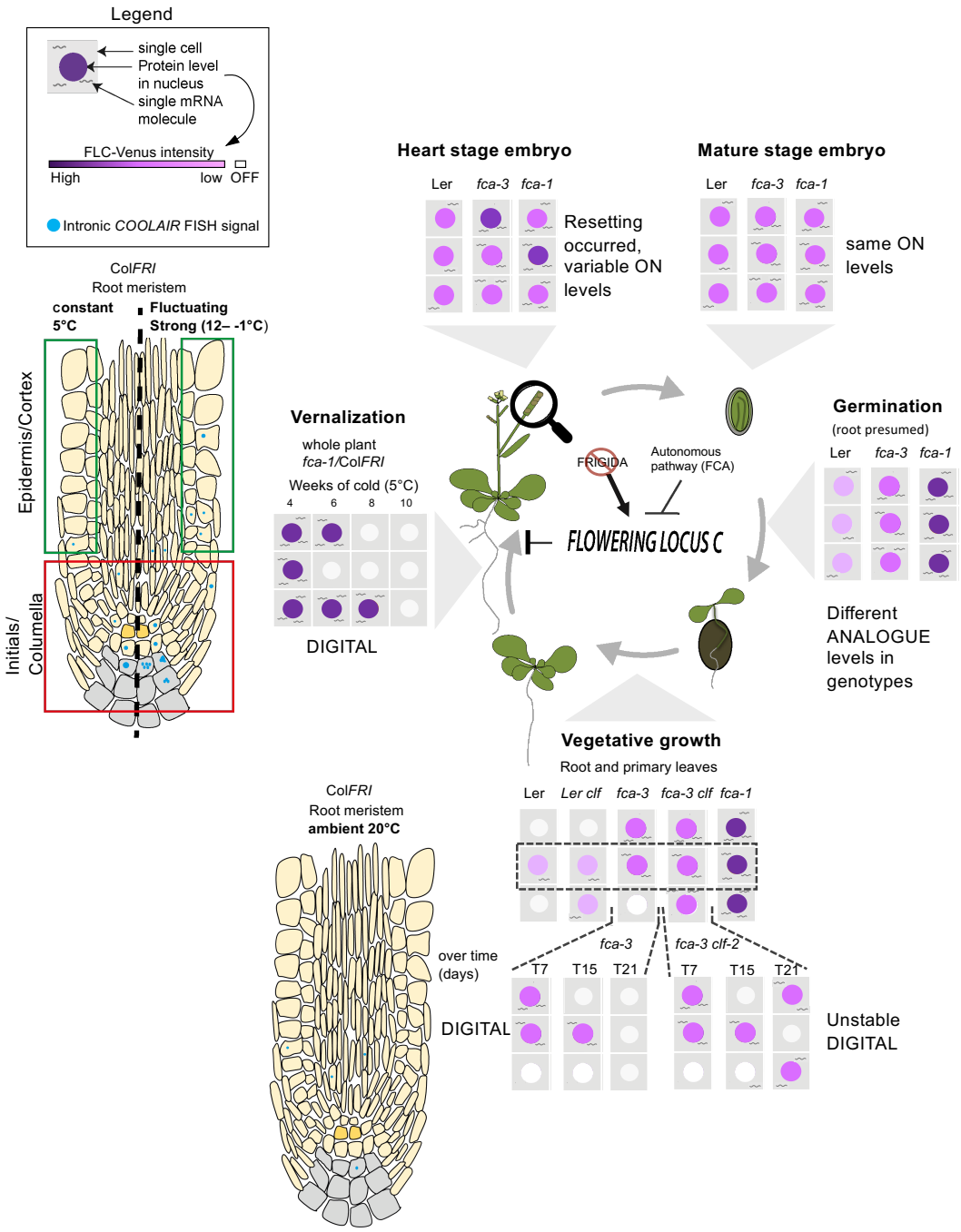


Figure 5–1 Schematic overview of finding of this work

Throughout the Arabidopsis life cycle, *FLC* regulation differs at different developmental stages. At the single cell level, differences in FLC-Venus intensities in the embryo were observed, between heart and mature stage, with overall ON cells in the latter. After germination, analogue differences were observed between mutants in the cold-independent autonomous pathway. Additionally, the first digitally OFF cells (silenced *FLC)* were observed at 4 days. During vegetative development the fraction of OFF cells increased in the intermediate genotypes *fca-3* and *fca-4*, while in the extreme cases they remained unchanged with either OFF (Ler) or ON (*fca-1).*

The digital switching probability depend on the analog transcriptional levels of *FLC* after germination of the genotype. In accordance with what has been observed during vernalization, this switch-off depends on PRC2 and is not maintained in a methyltransferase mutant of CLF. In a highly *FLC* transcribing *fca*-1 mutant, vernalization is sufficient in silencing *FLC,* indicating that FCA is not required for the switching OFF.

FLC antisense transcripts, collectively named *COOLAIR*, are found in few cells during vegetative growth in the root meristem. Upregulation of *COOLAIR* in the cold was shown to occur in the

meristematic vasculature, columella and initials. In strong fluctuating temperature conditions, further upregulation can be accounted for by an increased number of cells expressing *COOLAIR* and more *COOLAIR* transcripts accumulating in a defined focus (For Legend, see top left corner of the figure).

These findings provide insights into the complexity of gene regulation throughout the live cycle of Arabidopsis involving adaptations, long non-coding RNAs, epigenetic modifiers as well as environmental stimuli. We have shown an additional layer to the regulation of a polycomb target with implications for other Polycomb targets, which –when misregulated– are often associated with diseases such as cancer.

Understanding the response of the major flowering regulator *FLC* in severe fluctuating temperature conditions is especially relevant in the future for estimating the impact of extreme weather patterns on crop yield due to to climate change.

In summary, this research contributes to a broader understanding of epigenetic regulation and environmental adaptation in plants, offering potential applications in agriculture for developing climate-resilient crops.

Final Conclusions

Overall, what I have learned from my work is that analogue and digital modes of gene regulation are integrated at *FLC* to determine a plant's reproductive strategy. Two modes combined provide flexibility and stochasticity for cellular decision-making processes, particularly in response to changing environmental conditions or developmental transitions. Antisense transcription aids in fine-tuning this response, allowing the plants to buffer environmental signals and respond accordingly.

CHAPTER 6 : METHODS AND MATERIAL

Collaborative work is specified in the text and names are **highlighted** in bold letters for transparency.

6.1 Plant material

Plant lines used in this thesis are *Arabidopsis thaliana* accessions of either Landsberg *erecta* (Ler), Columbia-0 (Col-0) as well as Col-0 introgressed with a FRIGIDA (FRI SF, active FRI *sant feliu* allele introgression) Col-*FRI*^{SF} (Col*FRI*) background (Lee et al., 1994) or are natural accessions that are backcrossed to Col-0. If not specified otherwise, lines used in this thesis have been described before and are summarised in table 6-1 as well as briefly discussed below.

Table 6–1 *Arabidopsis* lines used in this thesis

Plant line	Background	Reference
fca-3	Ler	(Koornneef et al., 1991)
fca-4	Ler	(Koornneef et al., 1991)
fca-1	Ler	(Koornneef et al., 1991)
clf-2	Ler	(Goodrich et al.)
Col*FRI*^{SF2}	Col*FRI*	(Lee et al., 1994)
FLClean	Col*FRI*	(Nielsen et al.)
ntl8-D3	Col*FRI*	(Zhao et al., 2020a)
Var2-6 (NIL)	Col*FRI*	(Hepworth et al., 2020)
G259T	Col*FRI*	(Li et al., 2014)
Edi-0	Edi-0	(Shindo et al., 2005)

Generation of the *fca* alleles in L*er* carrying an FLC-Venus (Col-0) reporter construct was performed by **Scott Berry** and **Huamei Wang** and specified in Antoniou-Kourounioti et al. (2023). Briefly, the FLC-Venus construct was transformed into Ler, a single copy insertion line was selected and crossed into the *fca* allelic mutants. *clf-2* was crossed into *fca-3* FLC-Venus by **Eduardo Mateo-Bonmati** and selected by PCR-genotyping (Genotyping primer sequences in Table 6-2).

To be able to compare different natural accessions of *Arabidopsis* within a consistent genetic context, near-isogenic lines (NILs) were used. In these lines, the natural accession was backcrossed to Col*FRI* multiple rounds and selected for the accession *FLC* (Hepworth et al., 2020).

6.2 Growth conditions

Seeds were surface-sterilized with 5% (v/v) sodium hypochlorite for 5 min, rinsed with sterile water 4x for 1 min and stratified at 4°C for 3 days in water and in the dark. Days after sowing (DAS) are counted after stratification. Seeds were plated on MES buffered Murashige and Skoog (GM-MS) media pH 5.8 for 7 days in long-day conditions (16 h light/8 h dark), 20°C on shelves lit by 3 tubes (Philips, TLD 36W/865) producing a light intensity of 87 µmol m^{-2} s^{-1}.

For experiments in non-vernalized (NV) conditions, plants were grown on vertically oriented squared petri plates when used for imaging, and horizontally oriented round petri dishes for molecular experiments.

For the fluctuating temperature experiments, plants germinated on horizontal plates. At 4 days after sowing (when the root emerged in the growth media), plates were shifted to a vertical orientation. This allowed the roots to grow inside the media which was important to avoid autofluorescence for subsequent imaging. Roots were harvested for FISH experiments. Plants were transferred to temperature chamber (Gen 2000 Conviron) in the morning and exposed for 2 weeks with either constant 5°C, daily fluctuating 5°C (3°C–9°C) or daily fluctuating 5°C (−1°C to 12°C) as previously described in (Zhao et al., 2021). To resemble winter conditions, short day conditions (8 h light/16 h dark) were used and the light intensity was set to 20 µmol m^{-2} s^{-1}. Additionally, some plants were shifted for one night (when temperatures aligned) to the strong fluctuating chamber, where they experienced one night with 1°C fluctuations. Detailed temperatures regimes are shown in Figure 4-1.

Harvest of fluctuating temperature plants was performed in a 2-hours window between 10:00 and 12:00 (ZT1-3). Soil-grown plants were washed 2x with PBS for 2 seconds to wash off soil particles before fixation (smFISH) or drying and immerging in liquid nitrogen (RNA extraction).

6.3 Copy number analysis

The number of transgenic insertions in the newly generated lines was determined via a quantitative PCR (qPCR)-based assay for the BASTA-resistance gene (Bartlett et al., 2008) performed by **IDna Genetics** (Norwich, UK).

6.4 Embryo dissection

Embryos at different developmental stages were isolated by first glueing the siliques to a slide using double-sided sticky tape. Embryo stages could be approximated by the position of the silique along the inflorescence stem. Surgical needles were used to open siliques on either side of the middle rip. Seeds were taken out of the silique using forceps and placed in a drop of embryo dissection media (0.3 M Sorbitol, 0.5 mM MES, pH 5.8). Embryos were dissected from the seed with surgical needles.

6.5 Gene expression analysis

For gene expression analysis, 20+ plants were harvested at each indicated timepoint.

For *fca* time series, plants were harvested at 15:00 (ZT12). Harvest of plants exposed to fluctuating temperature conditions was performed between 10:00-12:00 (ZT1-3). Soil-grown plants were washed 2x with PBS for 2 seconds to wash off soil particles before being dried on a tissue.

Plant material was snap frozen in liquid nitrogen, ground and RNA was extracted as previously described (Yang et al., 2014). Specifically, 400 μL of Phenol:Chloroform:Isoamyl alcohol (25:24:1, Sigma, P3803) and 400 µL homogenisation buffer (100 mM Tris pH 8.0, 5 mM EDTA, 100 mM NaCl, 0.5 % SDS) pre-warmed to 60°C was added to the frozen plant powder and shaken for 15 min before being centrifuge at max speed for 10 minutes. The upper aqueous layer was transferred and the same amount of isopropanol as well as 50 µL 3M Sodium Acetate (RNase free) was added. Tubes were mixed briefly and transferred to -70°C for >15 min. Samples were centrifuged at maximum speed for 30 mins at 4°C, nucleic acids were resuspended in H_2O. To separate RNA from genomic DNA, Lithium Chloride (LiCl, Invitrogen, AM9480) was added to a final concentration of 2 M and incubated at 4°C for overnight. RNA pellets were obtained via centrifugation (20,000 × g, 4°C, 30 minutes) and washed with 75% ethanol before being resuspended in H_2O. The concentration of RNA was determined using a Nanodrop 1000 (Thermo-Scientific).

Contaminating genomic DNA was removed using TurboDNase (Ambion, AM1907) according to the manufacturer's instructions.

RNA was reverse transcribed into cDNA using SuperScript IV Reverse transcriptase (Invitrogen, 18090010) and RT primers for genes of interest. Gene expression was measured by qPCR, data was normalized to protein phosphatase 2A *(PP2A)* and Ubiquitin-conjugating enzyme 21 (*UBC).* Primer sequences are summarized in Table 6-2.

Table 6–2 Primers used in this thesis.

Primer name	Target	Sequence (5′-3′)	Purpose
SDB_FLC_4548_F	Unspliced FLC	ACTTGATCGGTAATTGCTGAACA	qPCR
SDB_FLC_4701_R	Unspliced FLC	CATTGGACACACAACACGCA	qPCR + RT
FLC_spliced_F	Spliced FLC	AGCCAAGAAGACCGAACTCA	qPCR
FLC_spliced_R	Spliced FLC	TTTGTCCAGCAGGTGACATC	qPCR + RT
set6_new_JH_LP	Total COOLAIR	TGCATCGAGATCTTGAGTGTATGT	qPCR
set6_new_JH_RP	Total COOLAIR	ACGTCCCTGTTGCAAAATAAGC	qPCR
Set1 new LP	RT for proximal COOLAIR	TGGTTGTTATTTGGTGGTGTG	RT
set2_new_2_LP	Proximal COOLAIR	CCTGCTGGACAAATCTCCGA	qPCR
set2_new_2_RP	Proximal COOLAIR	TCACACGAATAAGGTGGCTAATTAAG	qPCR
FLC -158_F	RT for distal COOLAIR	GCCCGACGAAGAAAAAGTAG	RT
Set4-new-F-195	Distal COOLAIR	GTATCTCCGGCGACTTGAAC	qPCR
Set4-new-R-195	Distal COOLAIR	GGATGCGTCACAGAGAACAG	qPCR
SDB_FLC-VENUS_ex6_cDNA_744_R	FLC-Venus	GTGCAGATCAGCTTCAGGGT	qPCR + RT
UBC_qPCR_F	Spliced UBC	CTGCGACTCAGGGAATCTTCTAA	qPCR
UBC_qPCR_R	Spliced UBC	TTGTGCCATTGAATTGAACCC	qPCR + RT
PP2A F2	Spliced PP2AA3	ACTGCATCTAAAGACAGAGTTCC	qPCR
PP2A R2	Spliced PP2AA3	CCAAGCATGGCCGTATCATGT	qPCR + RT
gamma.QPCR.F	γ*FCA*	AACTCCAGGGGCCAGCCATAATACT	qPCR
gamma.QPCR.R	γ*FCA*	TGTGTCCGAGCATAATCCTGAGCAT	qPCR + RT
SDB_*fca*-3_F1	*fca*-3 GT	TCCAACTGTTTGACTGTCAT	Genotyping
SDB_*fca*-3_R1	*fca*-3 GT	ACATACCACGACTCTGTCTA	Genotyping
fca-3 GTF	*fca*-3 GT	AACGGCTACCTTCCCAGTTC	Genotyping

fca-3 GTR	*fca*-3 GT	CCATTGCCGTCTCTTTGCTT	Genotyping
fca-1_GTF	*fca*-1 GT	ACCTCTTCACAGTCCACAGG	Genotyping
fca-1_GTR	*fca*-1 GT	TGAGTTTGCTGCGACAAGAG	Genotyping
fca-4 GTR	*fca*-4 GT	GGTAGGAGCTTATACAGTATGC	Genotyping
fca-4 GTF	*fca*-4 GT	ATGAGTTATCTTGCCCATAAC	Genotyping
fca-4 (for WT) GTR	*fca*-4 GT	CTCCCCGTCAGCATATCGAA	Genotyping
clf-2_RP	clf-2 GT	TTAACCCGGACCCGCATTTGTTTCGG	Genotyping
clf-2_LP	clf-2 GT WT	AGAGAAGCTCAAACAAGCCATCGA	Genotyping
clf-2_Ds	clf-2 GT mutation	GTCGGCGTGCGGCTGGCGGCG	Genotyping
-392F	FLC locus	actatgtaggcacgactttggtaac	ChIP
-272R	FLC locus	tgcagaaagaacctccactctac	ChIP
-49f	FLC locus	GCCCGACGAAGAAAAAGTAG	ChIP
58r	FLC locus	TTCAAGTCGCCGGAGATACT	ChIP
670f	FLC locus	TGAAGTTTCAAGCCATCTTTGA	ChIP
779r	FLC locus	TCACTCTGAAAAGAGACATTAATCA	ChIP
1533f	FLC locus	TTGACAATCCACAACCTCAATC	ChIP
1670r	FLC locus	TCAATTTCCTAGAGGCACCAA	ChIP
2465F	FLC locus	AGTTTGGCTTCCTCATACTTATGG	ChIP
2579R	FLC locus	CAATGAACCTTGAGGACAAGG	ChIP
3197f	FLC locus	ggggctgcgtttacatttta	ChIP
3353r	FLC locus	gtgatagcgctggctttgat	ChIP
3643F	FLC locus	TGAAATGTTACGAATACTAGCGTGT	ChIP
3752 R	FLC locus	GGATCAAAACTACTAGCTAACCCTTG	ChIP
5030f	FLC locus	ccggttgttggacataactagg	ChIP
5153r	FLC locus	ccaaacccagacttaaccagac	ChIP
5672F	FLC locus	CCTGCTGGACAAATCTCCGA	ChIP
5757R	FLC locus	GGATTTTGATTTCAACCGCCGA	ChIP
5948f	FLC locus	cgtgtgagaattgcatcgag	ChIP
6088r	FLC locus	aaaaacgcgcagagagagag	ChIP

STM F	STM locus	GCCCATCATGACATCACATC	ChIP
STM R	STM locus	GGGAACTACTTTGTTGGTGGTG	ChIP

6.6 Chromatin Immunoprecipitation

Histone Chromatin Immunoprecipitation (ChIP) was performed for *fca* allelic mutants as previously described (Angel et al., 2011).

2.5 gr of 7-, 14-, or 21-day-old seedlings were crosslinked in 1% formaldehyde and ground to fine powder for 15 min in liquid nitrogen. Nuclei were extracted using Honda buffer as described previously Sun et al. 2013. In all histone ChIP reactions, sonication, immunoprecipitation, DNA recovery and purification were performed as previously described in Wu et al. 2016. The antibodies used were anti-H3 (Abcam, ab176842), anti-H3K27me3 (Merck, 07-449 and abcam, ab192985), anti-H3K36me3 (Abcam, ab9050), anti-H3K4me1 (Abcam, ab8895). Values were normalized to H3 and internal control in either *STM* or *ACT7*. I performed the first T7 replicate, and the following experiments were performed by **Eduardo Mateo-Bonmati**. The DNA was quantified by qPCR with primers specific for *FLC* regions (Table 2).

6.7 Protein detection (Immunoblots)

Nuclear protein extraction was carried out on 1 g of ground plant tissue. Pre-cooled Cell Lysis Buffer (20 mM Tris-HCl, pH 7.5, 250 mM Sucrose, 25% glycerol, 20 mM KCl, 2.5 mM MgCl2, 0.1 % Nonidet P-40, 5 mM Dithiothreitol (DTT), 1× Protease Inhibitor Cocktail (Sigma-Aldrich, P9599), 1 mM Phenylmethylsulfonylfluorid (PMSF)) was added and mixed, before centrifugation at 4000 rpm for 15 min at 4°C. Pellet was resuspend and 2x washed in 1 mL of Nuclear wash buffer (20 mM Tris-HCl, pH 7.5, 25% glycerol, 2.5 mM MgCl2, 0.3% Triton X-100, 5 mM DTT, 1× Protease Inhibitor Cocktail, 1 mM PMSF) by centrifuging at 4000 rpm for 15 min at 4°C. Chromatin pellet is resuspended in 100 µL of Nuclear Lysis Buffer (50 mM Tris-HCl pH8, 10 mM EDTA, 1% Sodium dodecyl sulfate (SDS), 1× Protease Inhibitor, 1 mM PMSF). The extracted proteins were then separated on a NuPAGE 4-12% Bis-Tris Gel (Invitrogen, NP0326BOX), with the PageRuler Prestained Protein Ladder (10 to 180 kDa) (Thermo Scientific, 26616) serving as a reference standard. The gels were subsequently transferred onto a 0.45 μm PVDF membrane (GE Healthcare). For blocking, the membrane was immersed in a solution containing 5% (w/v) dried milk and 2% BSA in PBST (phosphate-buffered saline (PBS) with 0.1% Tween-20). The anti-*FCA* rabbit serum antibody (KL4) is a polyclonal antibody against a C-terminal FCA protein that has been expressed in *Escherichia coli*. The C-terminal fragment extends from just after the second RNA-recognition motif to the end of the coding region (Macknight et al., 2002). KL4

antibody was applied overnight at 4°C at a dilution of 1:500. The membrane was washed 4x with PBST for 5 minutes each on rocking platform.

For chemiluminescence detection, the membrane was incubated with anti-rabbit IgG horseradish peroxidase (1:10,000, GE Healthcare, NA935) and imaged on ImageQuant LAS 500 (Avantor).

FLC-Venus protein detection to assay protein stability was performed by **Yusheng Zhao** using the de novo protein synthesis inhibitor cycloheximide (Sigma-Aldrich, C1988) following the procedures in (Zhao, Antoniou-Kourounioti et al. 2020), 50 µl total protein was first enriched using magnetic GFP-trap beads (Chromo Tek, GTMA-20) before an antibody specific against GFP (Roche, 11814460001). The chemiluminescence signal was obtained by the FUJI Medical X-ray film (FUJI, 4741019289). Quantification was performed with ImageJ after the films were scanned with a printer scanner (RICOH).

Ponceau staining was performed with commercial Ponceau buffer (Sigma-Aldrich, P7170) and used as the processing control. All of the Immunoblots were performed with equal weight of whole seedlings.

6.8 Fluorescence in situ hybridisation for RNA detection

6.8.1 FISH on squashed material

A detailed smFISH protocol is described by Duncan et al. (2017), with small modifications. All chemicals/reagents used for smFISH were prepared in RNase-free solutions.

Respective plant parts were harvested and fixed in 4% paraformaldehyde (PFA). Fixation times vary depending on the sample. Embryos were fixed for 15 min, roots for 30 mins. 4% PFA was prepared by combining 16% EM-grade PFA (Electron Microscopy Sciences, 15710) with 10XPBS (Invitrogen, AM9625), and adjusting the final volume RNase-free water (Qiagen, 129117). After fixation, samples were washed 2x in 1xPBS. For squashed samples, samples were mounted on slide and 1xPBS and squashed between slides and cover slip gently applying pressure to the coverslip.

Samples were permeabilized for a minimum of 4 hrs with 70% ethanol in RNase-free water. Samples were then cleared in a Coplin jar, 30 min 100% MeOH and 100% EtOH and air dried. Samples were first washed for 5 mins in wash buffer (2X saline-sodium citrate (SSC, Invitrogen, AM9763), 10% formamide (Sigma, F9037)) before being incubated in a humidity chamber overnight at 37°C in the dark with smFISH probes to the final concentration of 150 nM. Probes were applied to samples in FISH hybridization solution consisting of: 2X SSC,

10% dextran sulfate (Thermo Scientific, AC433240050), 10% formamide, 5 μM oligonucleotide probes conjugated to fluorophores (Stellaris, LGC Biosearch Technologies) dissolved in RNase-free water. Individual probe sequences are in Table 3 (FLC-Venus) or published (PP2A, (Duncan et al., 2016)).

After hybridization, samples were washed 2x with freshly prepared wash buffer on the slide before 30 min incubation in a choplin jar at 37°C in the dark. 4,6-diamidino-2-phenylindole (DAPI, Thermo scientific, 62248) nuclear staining was applied in wash buffer at a final concentration of 100 ng/mL for 10 mins in the dark. DAPI was carefully removed and samples were washed in wash buffer and 2xSSC for 5 mins on the slide.

Samples were mounted in GLOX buffer using high precision cover glasses (Marienfeld, 0107032).

Single-molecule RNA Fluorescence in situ Hybridization (smFISH) of *FLC-Venus* on *Arabidopsis* roots in the *fca* allelic mutants was performed by **Anis Meschichi** (SLU, Uppsala) detailed in (Antoniou-Kourounioti et al., 2023).

6.8.2 Whole mount FISH

Whole mount FISH was performed as previously described in Zhao et al. (2023) (Zhao et al., 2023).

Plant material was dissected and fixed in 4% PFA. 1 cm of the root tip were cut and fully immersed in 4% PFA. Shoot apices were first dissected under a stereo microscope and incubated for 45 mins applying 2x vacuum for 15 mins. After Fixation and 2x washes in 1xPBS for 1 min, tissues are permeabilized in 75% Ethanol (EtOH) for > 16 hrs, 100% Methanol (MeOH) 2x for 15 mins, 2x in 100% EtOH. After being washed in 1xPBS, samples were cleared using ClearSee (10% (w/v) Xylitol powder (Sigma-Aldrich, X3375], 15% (w/v) Sodium deoxycholate [Sigma-Aldrich, 30970], 25% (w/v) Urea [Sigma-Aldrich, U5378] (Kurihara et al., 2015) at 4°C for specified times. Root samples were cleared for 7 days, and shoot samples were cleared for 14 days.

After, samples were first washed with wash buffer before being embedded in a hydrogel according to Gordillo et al., 2020(Gordillo et al., 2020). Hydrogel is a gel-plug that is used to keep sample hydrated and in place consisting of 30% acrylamide–bisacrylamide (29:1) (Bio-Rad, 1610158), 10% of Ammonium Persulfate, 5% of N,N,N′,N′-Tetramethylethylendiamin (TEMED, Thermo Scientific, 17919) in H2O. Adding TEMED will activate polymerization, sample embedding must occur immediately after. Roots were arranged in wash buffer on a poly-L slide, 20 μL of the activated embedding solution was added and covered with a cover slip. After 20-30 min of polymerization the cover slip was

removed carefully by inserting a razor blade between gel and coverslip and lifting the coverlip.

Hybridization and washes were performed as described in the FISH protocol (on squashed material. 0.1% (v/v) SCRI Renaissance 2200 (Renaissance Chemicals, SR2200) (Musielak et al., 2015) in wash buffer was used to stain the cell walls. After cell wall staining, slides were washed in wash buffer 2x on the slide and in a slide staining dish for 15 min before being mounted with 20 µl of vectashield (BioNordika, H-1000-10). Samples were either imaged immediately or on the next day.

6.8.3 SABER FISH

Signal amplification by exchange reaction (SABER) to amplify FISH signal was performed as described in Kishi et al (2019) with modifications (Kishi et al., 2019).

Briefly, SABER FISH hybridising oligonucleotides consist of two parts, the RNA hybridizing probe and the extending primer sequence collectively called concatemer, when extended.

To test the method in plants, I designed 65 probes for the full genomic sequence of the abundant housekeeping gene PP2A subunit A3 (AT1G13320.1) using the python script of OligoMiner (Beliveau et al., 2018) against the TAIR10 *Arabidopsis* genome with the parameters: Length: 35-40, Tm (in 50% formamide): 38-50 °C, Spacing: 0. Probes were chemically synthesized with primer sequence at the 3' ends which will extend to Primer Exchange Reaction (PER) concatemers by PCR amplification (shown in Fig 6-1). The most suitable 18 nt primer sequences were determined using BLAST against the whole genome and RNA sequences in *A. thaliana*. Two primer sets (60 and 65) with E-values of >10 were tested and primer sequences without probes were used as negative control. 1 nmol Probes+Primer were re-dissolved in 100 µL of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0, ThermoFisher, AM9849) to create a probe stock of 10 µM, mixed by pipetting, vortexed and centrifuged briefly.

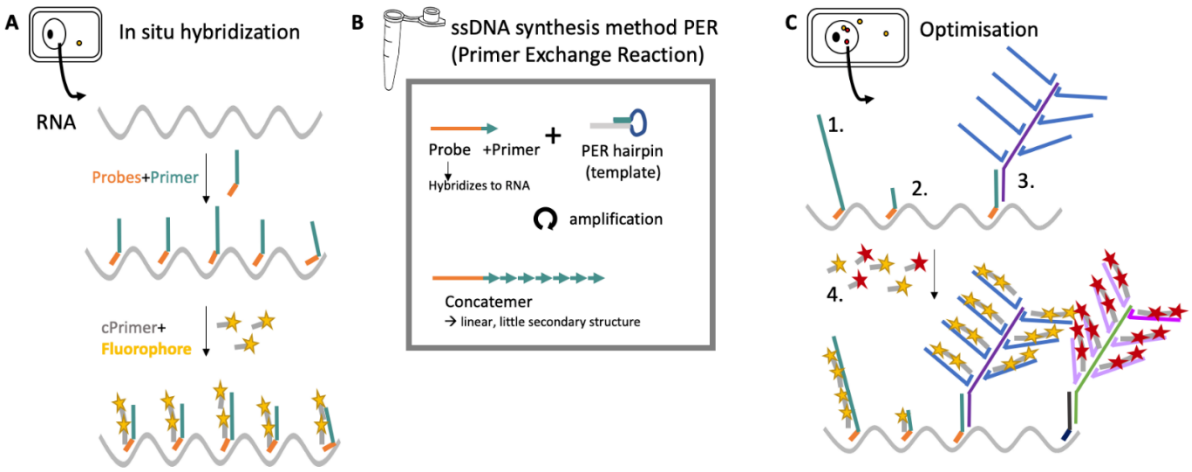


Figure 6–1 Schematic overview of the SABER FISH method by Kishi et al (2019)

(A) Two-step in situ hybridization of Probes (orange)+Primers (turquoise) followed by a secondary hybridization of the fluorophore (yellow) and complementary primer sequence to the primer. **(B)** Extension of the Primer-tail (concatemer) by rolling cycle amplification. **(C)** Optimisation of the in situ hybridization. 1. Represents different concatemer lengths, 2. Represents Adjustment of probe numbers, 3. Represents the addition of a bridge sequence (purple), 4. Represents two-colour labelling.

Individual probes, Hairpin (used for PER amplification) and Fluorophore were ordered in a 96 plate for oligonucleotides from Integrated DNA technologies (IDT) with normal desalting. Hairpin and Fluorophore were additionally HPLC purified and carry an inverted dT (idT) sequence at the 3′ end to inhibit degradation by 3′ exonucleases and extension by DNA polymerases.

In the Primer Exchange reaction (PER) the primer will bind to the complementary sequence of the corresponding PER hairpin to repeatedly add the same sequence (primer) domain onto the single-stranded primer (see Fig 6-1B). The primers and hairpins are designed so that the primer sequence does not contain the nucleotide G. PER amplification was performed by mixing 1×PBS (ThermoFisher, AM9625), 10mM MgSO4 (NEB, B1003S), dNTP mix* (A, C, T only 6mM each, NEB, N0446S) CleanG (1μM). 5 μL/mL Klenow DNA Polymerase I (large fragment) (Invitrogen, 18012-039) was added last, mixed and transferred to PCR tubes containing 5 μM of the respective hairpin. The strand-displacing polymerase incorporates the nucleotides A, T and C present in the reaction mix until it hits a CCC terminator sequence in the hairpin template unable to add G. Besides functioning as a terminator sequence, the lack of G in the available nucleotides also minimizes secondary structure. An additional CleanG hairpin is added to incorporate contaminating Gs. The mixture was incubated for 15 min at 37°C in a thermocycler before the unextended probe+primer oligomer mix (full probe set for the start) was added. A time series ranging from 30 min to 24 hrs was performed to track amplification/extension. Samples were exposed to 80°C for 20 min to deactivate the Klenow polymerase and kept at 4°C. The extension lengths of the concatemer were checked via agarose gel-electrophoresis (10 μL on

3 % agarose gel). 1x SYBR Gold (Invitrogen, S11494) in TE buffer was used to stain the concatemer extension, imaged on Typhoon FLA 9500(GE Healthcar). 90 μL of the concatemer reaction product was purified using a MinElute PCR purification column (Qiagen, 28004) and resuspended in 25 μL in nuclease-free H_2O. Concentration was determined using nanodrop (ssDNA settings). Concatemers are stored at -20°C before probes are used for hybridization. Individual probe+primer, hairpin, and imager sequences are in Table 6-3.

Sample preparation was performed as previously described in section 6.7.1. All sample handling was performed on ice when possible, and all solutions used were RNase-free to minimize RNA degradation.

In the first hybridization step (overnight at 37°C), the concatemer will hybridize to the target RNA. Visualization was achieved by a secondary hybridization of the imager oligonucleotide consisting of a Fluorophore ATTO565 and complementary primer sequence (Figure 6-1). The secondary hybridization was performed in WB without formamide. Samples were mounted in SlowFade Diamond (Invitrogen, S36967) or Vectashield (Vecta laboratories, H-1000-10).

Table 6–3 Primer and Hybridization sequences used for SABER FISH

Name (	Sequence
PER hairpin	ATTATTCACTG**GGCCTTTTGGCCC**AGTGAATAATAGTGAATAA/3InvdT/
Clean.G	CCCCGAAAGTGGCCTCG**GGCCTTTTGGCCC**GAGGCCACTTTCG
Imager	/5ATTO565N/TT**AGTGAATAA**T**AGTGAATAA**T
PP2A FISH sequence (Duncan et al., 2017)	SABER sequences (primer+probe)
ccgagcgatctatcaatcag	AAAAGCAAATACGCCCAACGAACAAATCAtttTTATTCACT
gacatcctcaccaaaactca	AACATTAGTAGCAGAGAGAAACAATGGTCACGtttTTATTCACT
tcgggtataaaggctcatca	AAATCTCATGAGAACCAAACACAATTCGTTGCtttTTATTCACT
tagctcgtcgataagcacag	TTAGCTCCACCAAGCATGGCCGTATCtttTTATTCACT
ccaagagcacgagcaatgat	ATGTTCTCCACAACCTTTCACAACCAACAtttTTATTCACT
atcaactcttttcttgtcct	TATATGGACGTTTAGATTCCCTGATGAGGAGGtttTTATTCACT
catcgtcattgttctcacta	TCATAAGCTGTTTCTTTACTTGTGAGAGCCAAtttTTATTCACT
atagccaaaagcacctcatc	CAAAGAATGTATACTAACCGCTTGGTCGACTtttTTATTCACT

atacagaataaaacccccca	ATCGGAATGAGAGATTGCATCATTTTGGCCtttTTATTCACT
caagtttcctcaacagtgga	AGTTATTACCGCAGGTAAGAGTTTGGAACATGtttTTATTCACT
tcatctgagcaccaattcta	TGATCTCGGATCCCATTACTGGAGCGtttTTATTCACT
tagccagaggagtgaaatgc	AGAAGCGATACTGCACGAAGAATCGTCtttTTATTCACT
cattcaccagctgaaagtcg	AGTGTGGGTTGTTAATCATCTCTAGAACCTGAtttTTATTCACT
ggaaaatcccacatgctgat	GCAAGACGCTTCAGATTGTTTGCAGCtttTTATTCACT
atattgatcttagctccgtc	AGCTTCACGGATTGAGTGAACCTGACtttTTATTCACT
attggcatgtcatcttgaca	TGGATGATTCAATGAAGGTAGAGAATGAGGGAtttTTATTCACT
aaattagttgctgcagctct	CCTAATTGACTGGCCAACAAGGGAATATACTCtttTTATTCACT
gctgattcaattgtagcagc	GATTATAGCCAGACGTACTCTCCAGTGCCtttTTATTCACT
ccgaatcttgatcatcttgc	CATTGACTTGACCAAAACATGATACAGCTCCTtttTTATTCACT
caaccctcaacagccaataa	ACCCAATGTCGTACAAAGAGATGAGTAACTTGtttTTATTCACT
ctccaacaatttcccaagag	ATCAATGCCAAGTTAAAAGCTGCAAAATCCAtttTTATTCACT
caaccatataacgcacacgc	ATATAGATTATTGTTGCGAGAACCTTCCAGGCtttTTATTCACT
agtagacgagcatatgcagg	ATTCCTAAAGATGACAGTGACTAGGACGAAGAtttTTATTCACT
gaacttctgcctcattatca	AGGTAAGCTTAGCAAACACCGACAGTATTTTtttTTATTCACT
cacagggaagaatgtgctgg	CACTCTTCTTAAGTGTGTACCTTTACACAGGGtttTTATTCACT
tgacgtgctgagaagagtct	AAGAATGTGCTGGATAGCGAGTTCAGGtttTTATTCACT
cccattataactgatgccaa	AGCTGCTATCCGAACTTCTGCCTCATtttTTATTCACT
tggttcacttggtcaagttt	TATCACAAAGTAGACGAGCATATGCAGGCtttTTATTCACT
tctacaatggctggcagtaa	ACCAGATCCGTCCTTAATACATAAGGCACAtttTTATTCACT
cgattatagccagacgtact	AAGATGACTCAACAAACATGGACTTCCAAGTtttTTATTCACT
gactggccaacaagggaata	ACCATAACTAACTTCAACCATCACTGCTTACCtttTTATTCACT
catcaaagaagcctacacct	TAGTTGGCTCCGGTCCTACAGCTTCAtttTTATTCACT
ttgcatgcaaagagcaccaa	ATAAAGCAGCTCATACATATCTGCCTTACAGCtttTTATTCACT
acggattgagtgaaccttgt	AGAATGTGTGCAACACAGTCCTGGGGtttTTATTCACT

cttcagattgtttgcagcag	AAAAGACACTGAAGCTGAGATGTAGACAACCtttTTATTCACT
ggaccaaactcttcagcaag	AATTCTGCTTAACTTCAGCTAATGGTCAGTGTtttTTATTCACT
ggaactatatgctgcattgc	AACCCTAATACTTGCCTATTACACATCAGCCAtttTTATTCACT
gtgggttgttaatcatctct	TAATAAGGGAGAGCATACCATCTTGCGTAAGAtttTTATTCACT
tgcacgaagaatcgtcatcc	ACATGGACATAATGTCTGTCTTCAAATGAGCTtttTTATTCACT
ttactggagcgagaagcgat	GATTCAATTGTAGCAGCAAACTTCCCCAAATtttTTATTCACT
gaacatgtgatctcggatcc	TAGTTGCTGCAGCTCTGCGCACCATTtttTTATTCACT
ctctgtctttagatgcagtt	GGCATGTCATCTTGACAAAGCTGACCAtttTTATTCACT
catcattttggccacgttaa	ATTGATCTTAGCTCCGTCTTTAGCACATCTGGtttTTATTCACT
cgtatcatgttctccacaac	ACATGCTGATACTCTGGCTGTGAACCAtttTTATTCACT
atcaacatctgggtcttcac	TTCACCAGCTGAAAGTCGCTGTTACACtttTTATTCACT
ttggagagcttgatttgcga	AATGACAGTGTTCAGTAATGCTTAAGTCGGTTtttTTATTCACT
acacaattcgttgctgtctt	TCACATTACGAGCACCTATACCGAGGTtttTTATTCACT
cgcccaacgaacaaatcaca	ATGTATGGTCACATACACAACCACAACCCtttTTATTCACT
	AAAAGAGCGCTAAAGCCAGAGCGATGtttTTATTCACT
	AGGGGAACACTGAACAAGAAAAGCATATTTCCtttTTATTCACT
	GGGGACAATTGACAAAGGACCCCAAAtttTTATTCACT
	AGAGCATGGAAAGGAATCAATAAAATCGACCAtttTTATTCACT
	AAGTAAACCAGCCTTATCTAACCTTAGCCAGAtttTTATTCACT
	CCCTCATCTGAGCACCAATTCTACAAAGTGtttTTATTCACT
	AATCCACAGCTTTCTCCCTCACGCAAtttTTATTCACT
	GTTTCCTCAACAGTGGATAGAGTCTCCAAAGtttTTATTCACT
	CCCCAACTCTTCCGCCATAGCCAAAAtttTTATTCACT
	CCAAGAGCACGAGCAATGATAGAAAGCCtttTTATTCACT
	TTTTAGCTCGTCGATAAGCACAGCAATCGtttTTATTCACT
	ATAAAGGCTCATCAACCATAGACATCCTCACCtttTTATTCACT
	ACAAACTTTAAACTCAATTCCGCCTCTGAGAtttTTATTCACT
	TGATTAGAATGGAGAACTGAAACTTTGTGCCGtttTTATTCACT
	AGAAATGATCGAGAATTTACCGATTACGCTGCtttTTATTCACT
	TTTCCAAGTTCCGAGCGATCTATCAATCAGtttTTATTCACT

6.9 Fixed-sample imaging

6.9.1 FISH

The imaging was conducted using a Zeiss Elyra PS1 wide-field epifluorescence microscope equipped with a 100x oil-immersion objective (1.46 NA) and a cooled electron-multiplying CCD (charge-coupled device) camera Andor iXon 897 (512 x 512 resolution, QE > 90%) in standard mode. Z-stacks covering the entire nucleus were acquired with an optical sectioning size of 0.2 μm. For DAPI staining, an excitation line at 405 nm was used, and signals were detected within the wavelength range of 420-480 nm. For probes labelled with Quasar570, an excitation line at 561 nm was used, and signal detection occurred within the range of 570-640 nm. Probes labelled with Quasar670, an excitation line of 642 nm was used, signal was detected at 655–710 nm.

For SABER probes labelled with ATTO565 and Alexa546, an excitation line of 561 nm was used, and signals were detected at 570-640 nm, with a 30 ms exposure time for DAPI, and 200-500 ms for the ATTO565 and Alexa546 dyes.

Isolated single cells, as well as single-layered connected cells, were selected for fluctuating temperature experiments for analysis.

6.9.2 Whole-mount FISH

Confocal microscopy was performed using the Stellaris 8 FALCON (Leica microsystems) using an HC PL Apo CS2 63X water-immersion objective (1.20 NA) in photon-counting mode. Optical sections with z-steps of 0.3 μm were collected. A laser speed of 400 Hz was used. The pinhole of FISH probes was opened to 2 Airy units.

Images were captured with an OPSL 561 nm excitation line (2%) and emission was collected at 535 - 650 nm using a HyD X detector for probes labelled with Quasar570 with a line a line accumulation of 4 and frame averaging of 2. The white light laser (WWL) at 80 MHz frequency with a laser power of 85% at an intensity of 5% was used to excite Quasar670 at 645 nm, signal was detected using the HyD X detector at 660-730 with a line accumulation of 6 and frame averaging of 2. For SCRI Renaissance 2200 imaging we used a Diode 405 nm laser line and emission was collected at 410 -503 nm.

6.9.3 Vernalization timecourse

Confocal microscopy was performed using the Stellaris 8 (Leica microsystems) using an HC PL APO CS2 63X water-immersion objective (1.20 NA) in photon-counting mode. A scan speed of 400 Hz was used with a line accumulation of 4. The excitation wavelength for SCRI Renaissance 2200 was 405 and 514 nm (OPSL 514 laser at 5%) for FLC-Venus imaging. The pinhole size was 1 Airy Unit. The image collection was performed at 410–503 nm for

SCRI Renaissance 2200 using the HyD S1 detector and 518–560 nm for FLC-Venus using the HyD S 2 detector. Optical sections with z-steps of 0.5 µm were collected.

6.10 Live imaging

6.10.1 FLC-Venus protein levels

FLC-Venus imaging of root samples in Figure 2-5 was performed by **Anis Meschichi**, optical sections were collected with a Zeiss LSM780 microscope equipped with a Channel Spectral GaAsP detector with a 20x objective (0.8 NA). For z-stacks, the step size was 0.5 µm with a pinhole aperture of 1.5 AU. The overall Z size varied between 45 and 75 slices depending on the orientation of the root. Roots from FLC-Venus lines were immersed in 1 µg/mL propidium iodide (PI, Sigma-Aldrich, P4864) to label the cell wall. For visualization of roots stained with PI, an excitation line of 514 nm was used, and the signal was detected at wavelengths between 611–656 nm. For observation of FLC-Venus signal, we used a 514 nm excitation line and detected from 518 to 535 nm. To allow comparison between treatments, the same laser power and detector settings were used for all FLC-Venus images.

In FLC-Venus time-course imaging of epidermal root meristems, the Leica SP8X or Stellaris 8 were used with 20x multi-immersion objective (0.75 NA). The Argon (SP8X) or OPSL 514 (Stellaris 8) lasers were used at 5% to excite FLC-Venus and PI at a 514 nm wavelength in bidirectional mode (PI signal was used for set-up). Venus was detected between 518 and 550 nm with the HyD SMD2 detector in photon counting mode; PI was detected at 600–675 nm. Epidermal images were obtained in photon- counting mode with laser speed 200, line accumulation of 6 (pixel dwell time of 2.43 µs), and a Z-step size of 0.95 µm and a pinhole size of 1 AU. For representative images, these were projected such that one single middle slice from the PI channel was used to show the cell outline, onto which 10 slices of FLC-Venus channel were average intensity projected (T7, LSM780 imaging, Figure 2-5) or sum projected for time-course imaging (T7/T15/T21, SP8X imaging, Figure 2-6). The dynamic range of the FLC-Venus signal was pushed from 0 to 255 to 5–22 (for LSM780 images) and 30–255 (for SP8X images) for all images apart from where enhanced. In enhanced *fca-3*/Ler images, the dynamic range was further pushed to 5–10 (for LSM780 images) and to 30–160 (for SP8X images) to obtain a similarly strong signal as observed in *fca-1*.
For *fca-4* FLC-Venus imaging, the Zeiss LSM880 was used with a 25x multi-immersion objective (0.8 NA). The Argon laser at 514 nm and 3% was used for excitation of FLC-Venus and PI, with a detection range between 520–550 nm and 600–650 nm, respectively. Z-stacks were taken covering the epidermis, with a step size of 1 µm. For representative images, five slices of FLC-Venus covering the nucleus were average projected and merged to the middle PI slice. FLC-Venus signal was enhanced to a similar intensity level as in *fca-1* by shifting the dynamic range of the FLC-Venus signal from 0 to 255 to 16–120.

FLC-Venus imaging of young leaves was performed by **Anis Meschichi** and **Lihua Zhao** (SLU, Uppsala) on 9-day-old seedlings. Seedlings were dissected by removing the cotyledons and keeping only the young leaves and shoot meristem. Young leaves were fixed in 4% paraformaldehyde at 4°C overnight. The samples were then treated with 100% methanol and 100% ethanol twice for 15 min each. ClearSee (Kurihara et al., 2015) solution was then used to clear the samples for 1 week. Next, two washes were carried out in 1× PBS. Samples were embedded in a hydrogel according to Gordillo et al. (2020). After rinsing with 1× PBS, samples were stained with Renaissance 2200 solution for 15 mins at room temperature. Slides were mounted in vectashield. FLC-Venus levels of the leaf images were quantified manually using ImageJ. The mean fluorescence intensity of a circular region covering the nucleus in each cell was measured, and the mean background of each image was subtracted.

6.10.2 Time course

Time course imaging of FLC-Venus protein levels in epidermal and cortex cells as well as the initials in *fca* allelic mutants was performed using the Leica confocal SP8X or Stellaris 8 microscope, using the argon laser at 5% power with 514 nm excitation for FLC-Venus and Propidium Iodide, using emission range of 518-550 nm and 600-650 nm or the OPSL 514 at 5%, respectively. An HC PL APO 20x multi-immersion objective (0.70 NA) was used.

Embryo time course was imaged on the Stellaris 8 FALCON (Leica microsystems) using an HC PL Apo CS2 63X water-immersion objective (1.20 NA) in photon-counting mode. 5% power with OPSL 514 nm excitation for FLC-Venus using the HyD S detector or the WWL was used. 200 Hz laser speed, line accumulation of 6. Transmitted light (PTM) was detected using the same detector.

6.10.3 Time-lapse imaging

Several methodologies have been tested to perform root time lapse imaging. Overall, roots were grown on top of the media of vertically oriented plates. Imaging was performed directly on the plate, by adding 20 μL PI (1 μg/mL) on the root tip, incubating for 10 min and carefully replacing it with H_2O before gently adding a cover glass on top of the root. Imaging was performed on 7-day-old roots in 5- or 6-hour intervals over a period of 48 hours, the cover slip was removed after each imaging. Switching was observed rarely (2 roots only, 1 in *fca-3* and 1 in *fca-4*). For the other imaged roots (+150 roots with varying conditions and experimental set-ups), the ON pattern was unchanged throughout all imaging time points. At the same time, no cell division was observed indicating that there was no progression of the cell cycle. Switching events were observed using the Zeiss LSM 880, in fast scan mode (800 Hz), Argon laser at 1%, no averaging using a 20x water objective (1.20 NA).

Additional efforts in imaging switching using Petri dishes with coverslip bottom (MatTek 35 mm, P35G-1.5-14-C) (with roots growing between the cover glass and a block of agar),

and an agar-block chamber containing Perfluorodecalin (Sigma-Aldrich, P9900) to maintain constant levels of oxygen while imaging as described in (Kirchhelle and Moore, 2017) were not successful.

6.11 Image analysis

6.11.1 FISH

Quantification of FISH probes was performed with the aid of **Matthew Harley** (EMBL, Cambridge), **Rea Antoniou-Kourounioti** and **Anis Meschichi** and took place in two stages:
1. Identification of probe locations in the whole 3D image, excluding the top and bottom z-slice from each z-stack due to light reflection at the plant cell wall.
2. Assignment of identified probes to specific cells via segmentation of the image into regions.
To detect probes, a white tophat filter was applied to the probe channel, followed by image normalization and thresholding. Individual probes were then identified by connected component segmentation. The centroids of each segmented region were assigned as the probe's location.
Images were manually annotated with markers to indicate positions of nuclei and whether cells were fully visible or occluded. A combination of the visibility markers and nuclei were used to seed a watershed segmentation of the image, thus dividing the image into regions representing individual cells. Probes within each region were counted to generate the per-cell counts. Occluded cells were excluded from the analysis (probe counts for those regions were ignored). Segmentation and probe detection parameters were optimized, with different thresholds used for different replicates and genotypes. This was necessary due to the difference in signal and background of the probes in each of these cases.
Visual inspection of smFISH images revealed high rates of false positive detection in Ler, likely due to very low levels of expression. In the absence of real signals, unspecific signals can be counted through our automated quantification pipeline as a consequence of unspecific probe binding or general background. When expression levels are high, this effect is negligible, but when expression levels are very low, it can become significant. Therefore, differences at the cellular mRNA level between *fca-3* and Ler are likely higher than the ones presented in Figure 2-4A. The similar count range observed in Ler and *fca-3* ON cells (Fig. 2-4) suggests that there is an overlap between the distribution of real ON cells in *fca-3* and background signal (in both Ler and *fca-3*). However, the higher frequency of such cells in *fca-3* compared to Ler confirms that these are not simply the result of background measurements. This is reflected in the significantly different distribution of Ler and *fca-3* cells in Figure 2-2. Furthermore, visual inspection of the images confirmed that in *fca-3* we can confidently detect real signals (higher than in Ler and lower than *fca-1*), confirming the analogue component.
Custom code is available at https://github.com/JIC-Image-Analysis/fishtools

COOLAIR fluorescence intensity of FISH foci has been quantified by first creating maximum projections of nucleus-spanning stacks for each nucleus. A threshold was defined to select the brightest foci (0-4) reflecting active transcription sites of *COOLAIR*. The signal was corrected for background noise (average of 3 regions of Interest that did not show signal) and the integrated density was calculated using ImageJ.

6.11.2 FLC-Venus fluorescence intensity

To measure per-cell FLC-Venus intensities in roots, we developed a custom image analysis pipeline with **Dr Matthew Hartley** (EMBL-EBI, UK) to extract cell structure information from the PI cell wall information and use this to measure per-cell nuclear fluorescence. Images were initially segmented by using the SimpleITK implementation of the morphological watershed algorithm (Beare and Lehmann, 2006). Reconstructed regions touching the image boundaries and those below a certain size threshold were removed. Segmentations were then manually curated by **Dr Rea Antoniou-Kourounioti** (UoG, UK), **Terri Holmes** (JIC/UEA, UK) with a segmentation tool in python written by **Dr John Fozard** (JIC, UK) to merge over-segmented regions, and to assign file identities to the resulting segmented cells. This curation was performed using custom software, able to merge segmented cells, resegment cells from specified seeds, and split cells along user-defined planes. To approximate the nuclear position, we fitted a fixed size spherical volume of 15 voxels radius to the point of maximal FLC-Venus intensity in each reconstructed cell. Cells in which the sphere overlap fraction (between sphere and cell) was less than 55% were excluded from the analysis. Per-voxel mean FLC-Venus intensities for each cell were then calculated by dividing the summed intensity within the intersection of the spherical region and the cell, by the fixed sphere volume. A background correction was performed by the following calculation to estimate the FLC-Venus intensity in the nucleus ("sphere_filled", also see Figure 6-2):

sphere_filled = (mean_in_sphere / overlap_fraction) - mean_outside_sphere

The "mean_in_sphere" intensity was first divided by the "overlap_fraction", the fraction of the sphere which overlaps the cell, to correct for regions of the sphere outside the cell. The mean intensity of the region outside the sphere ("mean_outside_sphere"), which is the background signal, was then subtracted from this overlap-corrected sphere intensity. This background subtraction is also the cause for the negative values in a fraction of OFF cells where the signal inside and outside the sphere are both background noise. Because of the sphere's prescribed size, in some cases even the brightest sphere inside a cell will have lower intensity than the remaining cell (e.g., by necessarily including the darkest region of the cell), leading to these negative values. The intensity levels of each cell were manually compared to the algorithm output levels using ImageJ. Histograms of the distributions of cells were composed in collaboration with **Dr Rea Antoniou-Kourounioti**.

Custom code for initial segmentation and Venus intensity measurement is written in the Python language and is available at https://github.com/JIC-Image-Analysis/root_measurement. The code for the segmentation curation software is available at https://github.com/jfozard/segcorrect.

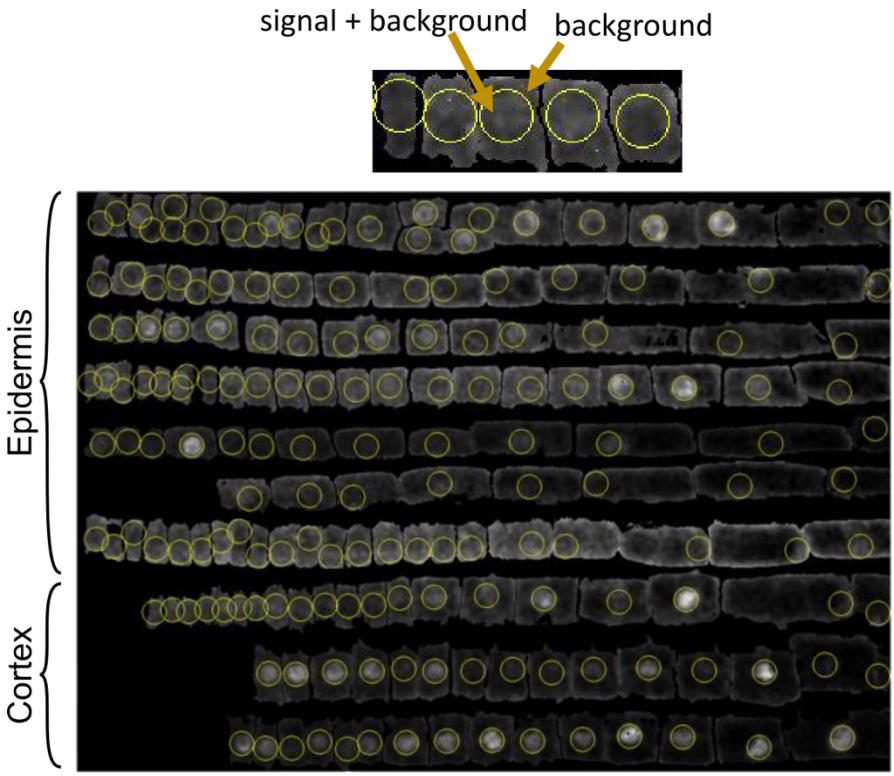


Figure 6–2 Segmentation and quantification method for FLC-Venus protein intensity per nucleus

Custom visualisation and segmentation software was used to manually correct automatic segmentations and annotate cells according to cell file. Top: Diagram explaining quantification of FLC-Venus intensity. In ON cells, the sphere contains the nucleus and therefore the FLC-Venus signal as well as the background signal. In OFF cells, the sphere will contain only background signal. The mean intensity outside the sphere was used as the background signal and this was subtracted from the sphere intensity. Bottom: Cell files from a single root separated according to tissue type demonstrating the result of the nucleus finding algorithm. Longitudinal root section shown. In each cell, the sphere of constant radius that maximised fluorescent intensity inside the sphere was determined. For each cell, the position of the sphere is shown, corresponding to the nucleus in ON cells.

6.12 Mathematical model

A mathematical model for *FLC* chromatin states and protein levels in the root was constructed by **Rea Antoniou-Kourounioti** used to simulate root cell files and generate simulated protein level histograms. Details for this model are specified in Antoniou-Kourounioti et al. (2023). The custom code is available at: https://github.com/ReaAntKour/*fca*_alleles_root_model.

6.13 Statistical analysis

Statistical tests are specified in each figure. Data were analysed using R software (http://r-project.org/) or using GraphPad Prism (https://www.graphpad.com).

6.14 Use of AI during the writing of the thesis

The Grammarly plug-in (https://app.grammarly.com) was used as a spelling correction tool. ChatGPT 3.5 (https://chat.openai.com) was consulted for feedback on the writing of the abstract and parts of the general discussion.

CHAPTER 7 : BIBLIOGRAPHY

Abadie, K., Clark, E.C., Ukogu, O., Yang, W., Daza, R.M., Ng, K.K.H., Fathima, J., Wang, A.L., Bhandoola, A., Nourmohammad, A., et al. (2023). Flexible and scalable control of T cell memory by a reversible epigenetic switch. BioRxiv 2022.12.31.521782.

Aguilar-Martínez, J.A., Poza-Carrión, C., and Cubas, P. (2007). Arabidopsis BRANCHED1 Acts as an Integrator of Branching Signals within Axillary Buds. Plant Cell *19*, 458–472.

Ahmad, K., Henikoff, S., and Ramachandran, S. (2022). Managing the Steady State Chromatin Landscape by Nucleosome Dynamics. Annu Rev Biochem *91*.

Alamos, S., Reimer, A., Niyogi, K.K., and Garcia, H.G. (2020). Quantitative imaging of RNA polymerase II activity in plants reveals the single-cell basis of tissue-wide transcriptional dynamics. BioRxiv.

Alberts B, Johnson A, Lewis J, et al. M.B. of the C. 4th edition. N.Y.G.S. 2002. P.D.A. from: https://www. ncbi. nlm. nih. gov/books/NBK26922. (2010). Molecular Biology of the Cell. Dev. Biol. *91*, 1–10.

Angel, A., Song, J., Dean, C., and Howard, M. (2011). A Polycomb-based switch underlying quantitative epigenetic memory. Nature *476*, 105–108.

Angel, A., Song, J., Yang, H., Questa, J.I., Dean, C., and Howard, M. (2015). Vernalizing cold is registered digitally at FLC. Proc. Natl. Acad. Sci. U. S. A. *112*, 4146–4151.

Antoniou-Kourounioti, R.L., Meschichi, A., Reeck, S., Berry, S., Menon, G., Zhao, Y., Fozard, J., Holmes, T., Zhao, L., Wang, H., et al. (2023). Integrating analog and digital modes of gene expression at Arabidopsis FLC. Elife *12*.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Bartlett, J.G., Alves, S.C., Smedley, M., Snape, J.W., and Harwood, W.A. (2008). High-throughput Agrobacterium-mediated barley transformation. Plant Methods *4*, 1–12.

Baxter, C.L., Šviković, S., Sale, J.E., Dean, C., and Costa, S. (2021). The intersection of DNA replication with antisense 3′ RNA processing in Arabidopsis FLC chromatin silencing. Proc. Natl. Acad. Sci. U. S. A. *118*, 2–4.

Beare, R., and Lehmann, G. (2006). The watershed transform in ITK - discussion and new developments. Insight J. *202*, 1–24.

Beliveau, B.J., Kishi, J.Y., Nir, G., Sasaki, H.M., Saka, S.K., Nguyen, S.C., Wu, C.-T., and Yin, P. (2018). OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. Proc. Natl. Acad. Sci. U. S. A. *115*, E2183–E2192.

Bell, L.R., Horabin, J.I., Schedl, P., and Cline, T.W. (1991). Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in Drosophila. Cell *65*, 229–239.

Berry, S., and Dean, C. (2015). Environmental perception and epigenetic memory: mechanistic insight through FLC. Plant J.

Berry, S., Hartley, M., Olsson, T.S.G., Dean, C., and Howard, M. (2015). Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. Elife *4*.

Berry, S., Dean, C., and Howard, M. (2017). Slow Chromatin Dynamics Allow Polycomb Target Genes to Filter Fluctuations in Transcription Factor Activity. Cell Syst. *4*, 445-457.e8.

Blair, L., Auge, G., and Donohue, K. (2017). Effect of FLOWERING LOCUS C on seed germination depends on dormancy. Funct. Plant Biol. *44*, 493–506.

Bloomer, R.H., and Dean, C. (2017). Fine-tuning timing: Natural variation informs the mechanistic basis of the switch to flowering in Arabidopsis thaliana. J. Exp. Bot. *68*, 5439–5452.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular Signals of Epigenetic States. Science *330*, 612.

Bouyer, D., Roudier, F., Heese, M., Andersen, E.D., Gey, D., Nowack, M.K., Goodrich, J., Renou, J.P., Grini, P.E., Colot, V., et al. (2011). Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. PLoS Genet. *7*, 1002014.

Buitrago, D., Labrador, M., Arcon, J.P., Lema, R., Flores, O., Esteve-Codina, A., Blanc, J., Villegas, N., Bellido, D., Gut, M., et al. (2021). Impact of DNA methylation on 3D genome structure. Nat. Commun. *12*.

Chen, M., and Penfield, S. (2018). Feedback regulation of COOLAIR expression controls seed dormancy and flowering time. Science (80-.). *360*, 1014–1017.

Chiang, G.C.K., Barua, D., Kramera, E.M., Amasino, R.M., and Donohue, K. (2009). Major flowering time gene, FLOWERING LOCUS C, regulates seed germination in Arabidopsis thaliana. Proc. Natl. Acad. Sci. U. S. A. *106*, 11661–11666.

Choi, J., Hyun, Y., Kang, M.J., In Yun, H., Yun, J.Y., Lister, C., Dean, C., Amasino, R.M., Noh, B., Noh, Y.S., et al. (2009). Resetting and regulation of FLOWERING LOCUS C expression during Arabidopsis reproductive development. Plant J. *57*, 918–931.

Choi, K., Kim, J., Hwang, H.J., Kim, S., Park, C., Kim, S.Y., and Lee, I. (2011). The FRIGIDA complex activates transcription of FLC, a strong flowering repressor in Arabidopsis, by recruiting chromatin modification factors. Plant Cell *23*, 289–303.

Cools, T., Iantcheva, A., Maes, S., Van Den Daele, H., and De Veylder, L. (2010). A replication stress-induced synchronization method for Arabidopsis thaliana root meristems. Plant J. *64*, 705–714.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., et al. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science (80-.). *316*, 1030–1033.

Le Corre, V., Roux, F., and Reboud, X. (2002). DNA polymorphism at the FRIGIDA gene in Arabidopsis thaliana: Extensive nonsynonymous variation is consistent with local selection for flowering time. Mol. Biol. Evol. *19*, 1261–1271.

Cortijo, S., Aydin, Z., Ahnert, S., and Locke, J.C. (2019). Widespread inter-individual gene expression variability in Arabidopsis thaliana. Mol. Syst. Biol. *15*, 8591.

Coulon, A., Ferguson, M.L., de Turris, V., Palangat, M., Chow, C.C., and Larson, D.R. (2014). Kinetic competition during the transcription cycle results in stochastic RNA processing. Elife *3*, e03939.

Coustham, V., Li, P., Strange, A., Lister, C., Song, J., and Dean, C. (2012a). Quantitative modulation of Polycomb silencing underlies natural variation in vernalization. Science (80-.). *337*, 584–587.

Coustham, V., Li, P., Strange, A., Lister, C., Song, J., and Dean, C. (2012b). Quantitative modulation of Polycomb silencing underlies natural variation in vernalization. Science (80-.). *337*, 584–587.

Crevillén, P., Sonmez, C., Wu, Z., and Dean, C. (2013). A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J. *32*, 140.

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., Cao, X., and Dean, C. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. Nature *515*, 587–590.

Csorba, T., Questa, J.I., Sun, Q., and Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. Proc. Natl. Acad. Sci. U. S. A. *111*, 16160–16165.

Deng, W., Ying, H., Helliwell, C.A., Taylor, J.M., Peacock, W.J., and Dennis, E.S. (2011). FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *108*, 6680–6685.

Dillon, N. (2006). Gene regulation and large-scale chromatin organization in the nucleus. Chromosom. Res. *14*, 117–126.

Dolan, L., Janmaat, K., Willemsen, V., Linstead, P., Poethig, S., Roberts, K., and Scheres, B. (1993). Cellular organisation of the Arabidopsis thaliana root. Development *119*, 71–84.

Dolan, L., Duckett, C.M., Grierson, C., Linstead, P., Schneider, K., Lawson, E., Dean, C., Poethig, S., and Roberts, K. (1994). Clonal relationships and cell patterning in the root epidermis of Arabidopsis. Development *120*.

Duncan, S., Olsson, T.S.G., Hartley, M., Dean, C., and Rosa, S. (2016). A method for detecting single mRNA molecules in Arabidopsis thaliana. Plant Methods *12*, 13.

Duncan, S., Olsson, T., Hartley, M., Dean, C., and Rosa, S. (2017). Single Molecule RNA FISH in Arabidopsis Root Cells. BIO-PROTOCOL.

Edwards, K.D., Anderson, P.E., Hall, A., Salathia, N.S., Locke, J.C.W., Lynn, J.R., Straume, M., Smith, J.Q., and Millar, A.J. (2006). FLOWERING LOCUS C Mediates Natural Variation in the High-Temperature Response of the Arabidopsis Circadian Clock. Plant Cell *18*, 639–650.

Fang, X., Wang, L., Ishikawa, R., Li, Y., Fiedler, M., Liu, F., Calder, G., Rowan, B., Weigel, D., Li, P., et al. (2019). Arabidopsis FLL2 promotes liquid–liquid phase separation of polyadenylation complexes. Nature *569*, 265–269.

Fang, X., Wu, Z., Raitskin, O., Webb, K., Voigt, P., Lu, T., Howard, M., and Dean, C. (2020). The 3 processing of antisense RNAs physically links to chromatin-based transcriptional control. Proc. Natl. Acad. Sci. U. S. A. *117*, 15316–15321.

Fenderson, B. (2008). Molecular Biology of the Cell,5th Edition.

Fiedler, M., Franco-Echevarría, E., Schulten, A., Nielsen, M., Rutherford, T.J., Yeates, A., Ahsan, B., Dean, C., and Bienz, M. (2022). Head-to-tail polymerization by VEL proteins underpins cold-induced Polycomb silencing in flowering control. Cell Rep. *41*.

Foltman, M., Evrin, C., De Piccoli, G., Jones, R.C., Edmondson, R.D., Katou, Y., Nakato, R., Shirahige, K., and Labib, K. (2013). Eukaryotic Replisome Components Cooperate to Process Histones During Chromosome Replication. Cell Rep. *3*, 892–904.

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., and Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. Genes Dev. *25*, 2041–2056.

Geraldo, N., Bäurle, I., Kidou, S.I., Hu, X., and Dean, C. (2009). FRIGIDA Delays Flowering in Arabidopsis via a Cotranscriptional Mechanism Involving Direct Interaction with the Nuclear Cap-Binding Complex. Plant Physiol. *150*, 1611–1618.

Goodnight, D., and Rine, J. (2020). S-phase-independent silencing establishment in Saccharomyces cerevisiae. Elife *9*, e58910.

Goodrich, J., Puangsomlee, P., Martint, M., Longt, D., and Meyerowitz+, E.M. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis.

Gordillo, S.V.G., Escobar-Guzman, R., Rodriguez-Leal, D., Vielle-Calzada, J.P., and Ronceret, A. (2020). Whole-Mount Immunolocalization Procedure for Plant Female Meiocytes. Methods Mol Biol *2061*, 13–24.

Grossniklaus, U., Vielle-Calzada, J.P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal control of embryogenesis by MEDEA, a Polycomb group gene in Arabidopsis. Science (80-.). *280*.

Groth, A., Corpet, A., Cook, A.J.L., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007). Regulation of Replication Fork Progression Through Histone Supply and Demand. Science (80-.). *318*, 1928–1931.

Guitton, A.E., Page, D.R., Chambrier, P., Lionnet, C., Faure, J.E., Grossniklaus, U., and Berger, F. (2004). Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis

thaliana. Development *131*, 2971–2981.

Hani, S., Cuyas, L., David, P., Secco, D., Whelan, J., Thibaud, M.C., Merret, R., Mueller, F., Pochon, N., Javot, H., et al. (2021). Live single-cell transcriptional dynamics via RNA labelling during the phosphate response in plants. Nat. Plants *7*, 1050–1064.

Hawkes, E.J., Hennelly, S.P., Novikova, I. V, Irwin, J.A., Dean, C., and Sanbonmatsu, K.Y. (2016). COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. Cell Rep. *16*, 3087–3096.

Hayashi, K., Hasegawa, J., and Matsunaga, S. (2013). The boundary of the meristematic and elongation zones in roots: endoreduplication precedes rapid cell expansion.

He, Y., Michaels, S.D., and Amasino, R.M. (2003). Regulation of Flowering Time by Histone Acetylation in Arabidopsis. Science (80-.). *302*, 1751–1754.

Hepworth, J., Antoniou-Kourounioti, R.L., Berggren, K., Selga, C., Tudor, E.H., Yates, B., Cox, D., Harris, B.R.C., Irwin, J.A., Howard, M., et al. (2020). Natural variation in autumn expression is the major adaptive determinant distinguishing Arabidopsis FLC haplotypes. Elife *9*, e57671.

Hofmann, F., Schon, M.A., and Nodine, M.D. (2019). The embryonic transcriptome of Arabidopsis thaliana. Plant Reprod. *32*, 77–91.

Holoch, D., Wassef, M., Lövkvist, C., Zielinski, D., Aflaki, S., Lombard, B., Héry, T., Loew, D., Howard, M., and Margueron, R. (2021). A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. Nat. Genet. 2021 5312 *53*, 1686–1697.

Hsieh, T.-F., Shin, J., Uzawa, R., Silva, P., Cohen, S., Bauer, M.J., Hashimoto, M., Kirkbride, R.C., Harada, J.J., Zilberman, D., et al. Regulation of imprinted gene expression in Arabidopsis endosperm.

Huang, X., Ding, J., Effgen, S., Turck, F., and Koornneef, M. (2013). Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of Arabidopsis. New Phytol. *199*, 843–857.

Hung, F.Y., Shih, Y.H., Lin, P.Y., Feng, Y.R., Li, C., and Wu, K. (2022). WRKY63 transcriptional activation of COOLAIR and COLDAIR regulates vernalization-induced flowering. Plant Physiol. *190*, 532–547.

Hyun, K. gi, Noh, Y.S., and Song, J.J. (2017). Arabidopsis FRIGIDA stimulates EFS histone H3 Lys36 methyltransferase activity. Plant Cell Rep. *36*, 1183–1185.

Ietswaart, R., Rosa, S., Wu, Z., Dean, C., and Howard, M. (2017). Cell-Size-Dependent Transcription of FLC and Its Antisense Long Non-coding RNA COOLAIR Explain Cell-to-Cell Expression Variation. Cell Syst.

Ingouff, M., Rademacher, S., Holec, S., Šoljić, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S., and Berger, F. (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in arabidopsis. Curr. Biol. *20*, 2137–2143.

Jeon, M., Jeong, G., Yang, Y., Luo, X., Jeong, D., Kyung, J., Hyun, Y., He, Y., and Lee, I. (2023). Vernalization-triggered expression of the antisense transcript COOLAIR is mediated by CBF genes. Elife *12*.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science (80-.). *290*, 344–347.

Jones, P.A., and Taylor, S.M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85–93.

Joubès, J., and Chevalier, C. (2000). Endoreduplication in higher plants.

Kirchhelle, C., and Moore, I. (2017). Chamber for Long-term Confocal Imaging of Root and Hypocotyl Development. J. Vis. Exp 55331.

Kishi, J.Y., Lapan, S.W., Beliveau, B.J., West, E.R., Zhu, A., Sasaki, H.M., Saka, S.K., Wang, Y., Cepko, C.L., and Yin, P. (2019). SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat. Methods *16*, 533–544.

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J.J., Goldberg, R.B., et al. (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *96*.

Koornneef, M., Hanhart, C.J., and van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. MGG Mol. Gen. Genet. *229*, 57–66.

Koornneef, M., Alonso-Blanco, C., Blankestijn-De Vries, H., Hanhart, C.J., and Peeters, A.J.M. (1998). Genetic Interactions Among Late-Flowering Mutants of Arabidopsis.

Kueh, H.Y., Champhekar, A., Nutt, S.L., Elowitz, M.B., and Rothenberg, E. V (2013). Positive Feedback Between PU.1 and the Cell Cycle Controls Myeloid Differentiation. Science (80-.). *341*, 670–673.

Kurihara, D., Mizuta, Y., Sato, Y., and Higashiyama, T. (2015). ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. Development *142*, 4168–4179.

Lee, I., Michaels, S.D., Masshardt, A.S., and Amasino, R.M. (1994). The late-flowering phenotype of FRIGIDA and mutations in LUMINIDEPENDENS is suppressed in the Landsberg erecta strain of Arabidopsis. Plant J. *6*.

Lee, L.R., Wengier, D.L., and Bergmann, D.C. (2019). Cell-type–specific transcriptome and histone modification dynamics during cellular reprogramming in the Arabidopsis stomatal lineage. Proc. Natl. Acad. Sci. U. S. A. *116*, 21914–21924.

Lee, Y.P., Fleming, A.J., Körner, C., and Meins, F. (2009). Differential expression of the CBF pathway and cell cycle-related genes in Arabidopsis accessions in response to chronic low-temperature exposure. Plant Biol. *11*, 273–283.

Li, C., Gu, M., Shi, N., Zhang, H., Yang, X., Osman, T., Liu, Y., Wang, H., Vatish, M., Jackson, S., et al. (2011). Mobile FT mRNA contributes to the systemic florigen signalling in floral induction. Sci. Reports 2011 11 *1*, 1–6.

Li, P., Tao, Z., and Dean, C. (2014). Phenotypic evolution through variation in splicing of the noncoding RNA COOLAIR. Genes Dev. 696–701.

Lim, M.-H., Kim, J., Kim, Y.-S., Chung, K.-S., Seo, Y.-H., Lee, I., Kim, J., Hong, C.B., Kim, H.-J., and Park, C.-M. (2004). A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. Plant Cell *16*, 731–740.

Liu, F., Quesada, V., Crevillén, P., Bäurle, I., Swiezewski, S., and Dean, C. (2007). The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC. Mol. Cell *28*, 398–407.

Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3′ processing of antisense transcripts triggers arabidopsis FLC chromatin silencing. Science (80-.). *327*, 94–97.

Liu, J., He, Y., Amasino, R., and Chen, X. (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes Dev. *18*, 2873–2878.

Lövkvist, C., Mikulski, P., Reeck, S., Hartley, M., Dean, C., and Howard, M. (2021). Hybrid protein assembly-histone modification mechanism for PRC2-based epigenetic switching and memory. Elife *10*, e66454.

Lu, F., Cui, X., Zhang, S., Jenuwein, T., and Cao, X. (2011). Arabidopsis REF6 is a histone H3 lysine 27 demethylase. Nat. Genet. 2011 437 *43*, 715–719.

Luo, X., Chen, T., Zeng, X., He, D., and He, Y. (2019). Feedback Regulation of FLC by FLOWERING LOCUS T (FT) and FD through a 5′ FLC Promoter Region in

Arabidopsis. Mol. Plant *12*, 285–288.

Luo, X., Ou, Y., Li, R., and He, Y. (2020). Maternal transmission of the epigenetic 'memory of winter cold' in Arabidopsis. Nat. Plants *6*, 1211–1218.

Macknight, R., Bancroft, I., Page, T., Lister, C., Schmidt, R., Love, K., Westphal, L., Murphy, G., Sherson, S., Cobbett, C., et al. (1997). FCA, a gene controlling flowering time in arabidopsis, encodes a protein containing RNA-binding domains. Cell *89*, 737–745.

Macknight, R., Duroux, M., Laurie, R., Dijkwel, P., Simpson, G., and Dean, C. (2002). Functional Significance of the Alternative Transcript Processing of the Arabidopsis Floral Promoter FCA. Plant Cell *14*, 877-88.

Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., and Dean, C. (2014). Functional Consequences of Splicing of the Antisense Transcript COOLAIR on FLC Transcription. Mol. Cell *54*, 156–165.

Mateo-Bonmatí, E., Fang, X., Maple, R., Fiedler, M., Passmore, L.A., and Dean, C. (2023). The CPSF phosphatase module links transcription termination to chromatin silencing. BioRxiv 2023.07.06.547976.

Medford, J. 1 (1992). Vegetative Apical Meristems.

Melaragno, J.E., Mehrotra, B., and Coleman, A.W. (1993). Relationship between endopolyploidy and cell size in epidermal tissue of arabidopsis. Plant Cell *5*.

Meng, T.G., Zhou, Q., Ma, X.S., Liu, X.Y., Meng, Q.R., Huang, X.J., Liu, H.L., Lei, W.L., Zhao, Z.H., Ouyang, Y.C., et al. (2020). PRC2 and EHMT1 regulate H3K27me2 and H3K27me3 establishment across the zygote genome. Nat. Commun. *11*, 1–12.

Menges, M., De Jager, S.M., Gruissem, W., and Murray, J.A.H. (2005). Global analysis of the core cell cycle regulators of Arabidopsis identifies novel genes, reveals multiple and highly specific profiles of expression and provides a coherent model for plant cell cycle control. Plant J. *41*, 546–566.

Menon, G., Schulten, A., Dean, C., and Howard, M. (2021). Digital paradigm for Polycomb epigenetic switching and memory. Curr. Opin. Plant Biol. *61*, 102012.

Menon, G., Mateo-Bonmati, E., Reeck, S., Maple, R., Wu, Z., Ietswaart, R., Dean, C., and Howard, M. (2024). Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing. Mol. Cell *2*, https://doi.org/10.1016/j.molcel.2024.05.014.

Meyerowitz, E.M. (1997). Genetic control of cell division patterns in developing plants. Cell *88*, 299–308.

Michaels, S.D., and Amasino, R.M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell *11*, 949–956.

Moazed, D. (2011). Mechanisms for the Inheritance of Chromatin States. Cell *146*, 510–518.

Moll, P., Ante, M., Seitz, A., and Reda, T. (2014). QuantSeq 3′ mRNA sequencing for RNA quantification. Nat. Methods 2014 1112 *11*, i–iii.

Munsky, B., and Neuert, G. (2015). From analog to digital models of gene regulation. Phys Biol *12*, 45004.

Musielak, T.J., Schenkel, L., Kolb, M., Henschen, A., and Bayer, M. (2015). A simple and versatile cell wall staining protocol to study plant reproduction. Plant Reprod. *28*.

Nielsen, M., Menon, G., Zhao, Y., Mateo-Bonmati, E., Wolff, P., Zhou, S., Howard, M., and Dean, C. Antisense transcription and PRC2 repression function in parallel during vernalization.

Nobori, T., Oliva, M., Lister, R., and Ecker, J.R. (2023). Multiplexed single-cell 3D spatial gene expression analysis in plant tissue using PHYTOMap. Nat. Plants.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001).

The Polycomb -Group Gene Ezh2 Is Required for Early Mouse Development . Mol. Cell. Biol. *21*, 4330–4336.

Oya, S., Takahashi, M., Takashima, K., Kakutani, T., and Inagaki, S. (2022). Transcription-coupled and epigenome-encoded mechanisms direct H3K4 methylation. Nat. Commun. *13*.

Page, T., Macknight, R., Yang, C.H., and Dean, C. (1999). Genetic interactions of the Arabidopsis flowering time gene FCA, with genes regulating floral initiation. Plant J *17*, 231–239.

Paige, J.S., Wu, K., and Jaffrey, S.R. (2011). RNA mimics of green fluorescent protein. Science *333*, 642.

Probst, A. V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.

Quesada, V., Macknight, R., Dean, C., and Simpson, G.G. (2003). Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. EMBO J. *22*, 3142.

Qüesta, J.I., Antoniou-Kourounioti, R.L., Rosa, S., Li, P., Duncan, S., Whittaker, C., Howard, M., and Dean, C. (2020). Non-coding SNPs influence a distinct phase of Polycomb silencing to destabilise long-term epigenetic memory at Arabidopsis FLC. Genes Dev *34*, 446–461.

Rahni, R., and Birnbaum, K.D. (2019). Week-long imaging of cell divisions in the Arabidopsis root meristem. Plant Methods *15*, 1–14.

Raissig, M.T., Bemer, M., Baroux, C., and Grossniklaus, U. (2013). Genomic Imprinting in the Arabidopsis Embryo Is Partly Regulated by PRC2. PLOS Genet. *9*, e1003862.

Reddy, G.V., Heisler, M.G., Ehrhardt, D.W., and Meyerowitz, E.M. (2004). Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. Development *131*, 4225–4237.

Rieder, C.L., and Cole, R.W. Cell Cycle Cold-Shock and the Mammalian Cell Cycle. Cell Cycle *1*, 169–175.

Roeder, A.H.K., and Yanofsky, M.F. (2006). Fruit Development in Arabidopsis. Arab. B. *4*.

Rosa, S., Duncan, S., and Dean, C. (2016). Mutually exclusive sense-antisense transcription at FLC facilitates environmentally induced gene repression. Nat. Commun.

Rouhanifard, S.H., Mellis, I.A., Dunagin, M., Bayatpour, S., Jiang, C.L., Dardani, I., Symmons, O., Emert, B., Torre, E., Cote, A., et al. (2019). ClampFISH detects individual nucleic acid molecules using click chemistry–based amplification. Nat. Biotechnol. *37*, 84–94.

Rymen, B., Fiorani, F., Kartal, F., Vandepoele, K., Inzé, D., and Beemster, G.T.S. (2007). Cold Nights Impair Leaf Growth and Cell Cycle Progression in Maize through Transcriptional Changes of Cell Cycle Genes. Plant Physiol. *143*, 1429.

Schmalenbach, I., Zhang, L., Ryngajllo, M., and Jiménez-Gómez, J.M. (2014). Functional analysis of the Landsberg erecta allele of FRIGIDA. BMC Plant Biol. *14*, 1–11.

Schomburg, F.M., Patton, D.A., Meinke, D.W., and Amasino, R.M. (2001). FPA, a Gene Involved in Floral Induction in Arabidopsis, Encodes a Protein Containing RNA-Recognition Motifs. Plant Cell *13*, 1427.

Schon, M., Baxter, C., Xu, C., Enugutti, B., Nodine, M.D., and Dean, C. (2021). Antagonistic activities of cotranscriptional regulators within an early developmental window set FLC expression level. Proc. Natl. Acad. Sci. U. S. A. *118*, e2102753118.

Sheldon, C.C., Hills, M.J., Lister, C., Dean, C., Dennis, E.S., and Peacock, W.J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. Proc. Natl. Acad. Sci. U. S. A. *105*, 2214–2219.

Shindo, C., Aranzana, M.J., Lister, C., Baxter, C., Nicholls, C., Nordborg, M., and Dean, C. (2005). Role of FRIGIDA and FLOWERING LOCUS C in Determining Variation in Flowering Time of Arabidopsis. Plant Physiol. *138*, 1163–1173.

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., and Dean, C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response. Genes Dev. *20*, 3079–3083.

Shishkova, S., Rost, T.L., and Dubrovsky, J.G. (2008). Determinate Root Growth and Meristem Maintenance in Angiosperms. Ann. Bot. *101*, 319.

Shu, J., Chen, C., Thapa, R.K., Bian, S., Nguyen, V., Yu, K., Yuan, Z.-C., Liu, J., Kohalmi, S.E., Li, C., et al. (2019). Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in Arabidopsis seedlings. Plant Direct *3*, e00100.

Simonini, S., Bemer, M., Bencivenga, S., Gagliardini, V., Pires, N.D., Desvoyes, B., van der Graaff, E., Gutierrez, C., and Grossniklaus, U. (2021). The Polycomb group protein MEDEA controls cell proliferation and embryonic patterning in Arabidopsis. Dev. Cell *56*.

Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY Is an RNA 3′ End-Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition. Cell *113*, 777–787.

Smit, M.E., Vaté, A., Mair, A., Northover, C.A.M., Bergmann Correspondence, D.C., and Bergmann, D.C. (2023). Extensive embryonic patterning without cellular differentiation primes the plant epidermis for efficient post-embryonic stomatal activities. Dev. Cell *58*, 506-521.e5.

Spensley, M., Kim, J.Y., Picot, E., Reid, J., Ott, S., Helliwell, C., and Carré, I.A. (2009). Evolutionary conserved regulatory motifs in the promoter of the arabidopsis clock gene late elongated hypocotyl. Plant Cell *21*, 2606–2623.

Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. Nat. Rev. Mol. Cell Biol.

Stirling, P.C., Chan, Y.A., Minaker, S.W., Aristizabal, M.J., Barrett, I., Sipahimalani, P., Kobor, M.S., and Hieter, P. (2012). R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants. Genes Dev. *26*, 163–175.

Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D.J., and Jacobsen, S.E. (2014). Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat. Struct. Mol. Biol. *21*, 64–72.

Sun, B., Zhou, Y., Cai, J., Shang, E., Yamaguchi, N., Xiao, J., Looi, L.S., Wee, W.Y., Gao, X., Wagner, D., et al. (2019). Integration of transcriptional repression and polycomb-mediated silencing of WUSCHEL in floral meristems. Plant Cell *31*, 1488–1505.

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., and Dean, C. (2013). R-loop stabilization represses antisense transcription at the arabidopsis FLC locus. Science (80-.). *340*, 619–621.

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

Taiz, L., and Zeiger, E. (2014). Plant Physiology and Development.

Tamada, Y., Yun, J.-Y., chul Woo, S., and Amasino, R.M. ARABIDOPSIS TRITHORAX-RELATED7 Is Required for Methylation of Lysine 4 of Histone H3 and for Transcriptional Activation of FLOWERING LOCUS C C W.

Tang, P., Yang, Y., Li, G., Huang, L., Wen, M., Ruan, W., Guo, X., Zhang, C., Zuo, X., Luo, D., et al. (2022). Alternative polyadenylation by sequential activation of distal and proximal PolyA sites. Nat. Struct. Mol. Biol. 2022 291 *29*, 21–31.

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

Teotia, S., and Tang, G. (2015). To Bloom or Not to Bloom: Role of MicroRNAs in Plant Flowering. Mol. Plant *8*, 359–377.

Tian, Y., Zheng, H., Zhang, F., Wang, S., Ji, X., Xu, C., He, Y., and Ding, Y. (2019). PRC2 recruitment and H3K27me3 deposition at FLC require FCA binding of COOLAIR. Sci. Adv.

5, 1–13.

Urbanek, M.O., Galka-Marciniak, P., Olejniczak, M., and Krzyzosiak, W.J. (2014). RNA imaging in living cells – methods and applications. RNA Biol. *11*, 1083.

Wang, Z., Wu, Z., Raitskin, O., Sun, Q., and Dean, C. (2014). Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor. 1–6.

Wolff, P., Weinhofer, I., Seguin, J., Roszak, P., Beisel, C., Donoghue, M.T.A., Spillane, C., Nordborg, M., Rehmsmeier, M., and Köhler, C. (2011). High-Resolution Analysis of Parent-of-Origin Allelic Expression in the Arabidopsis Endosperm. PLoS Genet. *7*.

Wu, J., Zaccara, S., Khuperkar, D., Kim, H., Tanenbaum, M.E., and Jaffrey, S.R. (2019a). Live imaging of mRNA using RNA-stabilized fluorogenic proteins. Nat. Methods *16*, 862.

Wu, Z., Ietswaart, R., Liu, F., Yang, H., Howard, M., and Dean, C. (2016). Quantitative regulation of FLC via coordinated transcriptional initiation and elongation. Proc. Natl. Acad. Sci. *113*, 218–223.

Wu, Z., Fang, X., Zhu, D., and Dean, C. (2019b). Autonomous Pathway: FLOWERING LOCUS C Repression through an Antisense-Mediated Chromatin-Silencing Mechanism. Plant Physiol. *182*, 27–37.

Xu, C., Wu, Z., Duan, H.C., Fang, X., Jia, G., and Dean, C. (2021). R-loop resolution promotes co-transcriptional chromatin silencing. Nat. Commun. *12*.

Xu, G., Tao, Z., and He, Y. (2022). Embryonic reactivation of FLOWERING LOCUS C by ABSCISIC ACID-INSENSITIVE 3 establishes the vernalization requirement in each Arabidopsis generation . Plant Cell *34*, 2205–2221.

Yang, H., Howard, M., and Dean, C. (2014). Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at Arabidopsis FLC. Curr Biol *24*, 1793–1797.

Yang, H., Berry, S., Olsson, T.S.G., Hartley, M., Howard, M., and Dean, C. (2017). Distinct phases of Polycomb silencing to hold epigenetic memory of cold in Arabidopsis. *1145*, 1142–1145.

Yang, M., Zhu, P., Cheema, J., Bloomer, R., Mikulski, P., Liu, Q., Zhang, Y., Dean, C., and Ding, Y. (2022). In vivo single-molecule analysis reveals COOLAIR RNA structural diversity. Nature *609*, 394.

Yoder, T.L., Zheng, H.Q., Todd, P., and Staehelin, L.A. (2001). Amyloplast Sedimentation Dynamics in Maize Columella Cells Support a New Model for the Gravity-Sensing Apparatus of Roots. Plant Physiol. *125*, 1045–1060.

Zhang, L., and Jiménez-Gómez, J.M. (2020). Functional analysis of FRIGIDA using naturally occurring variation in Arabidopsis thaliana. Plant J. *103*, 154–165.

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

Zhao, L., Fonseca, A., Meschichi, A., Sicard, A., and Rosa, S. (2023). Whole-mount smFISH allows combining RNA and protein quantification at cellular and subcellular resolution. Nat. Plants *9*, 1094.

Zhao, Y., Antoniou-Kourounioti, R.L., Calder, G., Dean, C., and Howard, M. (2020a). Temperature-dependent growth contributes to long-term cold sensing. Nature *583*, 825–829.

Zhao, Y., Antoniou-Kourounioti, R.L., Calder, G., Dean, C., and Howard, M. (2020b). Temperature-dependent growth contributes to long-term cold sensing. Nature *583*, 825–829.

Zhao, Y., Zhu, P., Hepworth, J., Bloomer, R., Antoniou-Kourounioti, R.L., Doughty, J., Heckmann, A., Xu, C., Yang, H., and Dean, C. (2021). Natural temperature fluctuations promote COOLAIR regulation of FLC. Genes Dev. *38*, 1–11.

Zheng, S., Hu, H., Ren, H., Yang, Z., Qiu, Q., Qi, W., Liu, X., Chen, X., Cui, X., Li, S., et al. (2019). The Arabidopsis H3K27me3 demethylase JUMONJI 13 is a temperature and photoperiod dependent flowering repressor. Nat. Commun. 2019 101 *10*, 1–11.

Zhu, P., Lister, C., and Dean, C. (2021). Cold-induced Arabidopsis FRIGIDA nuclear condensates for FLC repression. Nature *599*, 657–661.

CHAPTER 8 : APPENDIX

Publications

Integrating analog and digital modes of gene expression at Arabidopsis *FLC*

Proximal Termination Generates a Transcriptional State that Determines the Rate of Establishment of Polycomb Silencing

RESEARCH ARTICLE

Integrating analog and digital modes of gene expression at *Arabidopsis FLC*

Rea L Antoniou-Kourounioti[1,2†], Anis Meschichi[3†], Svenja Reeck[4†], Scott Berry[5], Govind Menon[1], Yusheng Zhao[6], John Fozard[1], Terri Holmes[7], Lihua Zhao[3], Huamei Wang[8], Matthew Hartley[9], Caroline Dean[4*], Stefanie Rosa[3*], Martin Howard[1*]

[1]Department of Computational and Systems Biology, John Innes Centre, Norwich, United Kingdom; [2]School of Molecular Biosciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom; [3]Swedish University of Agricultural Sciences, Plant Biology Department, Uppsala, Sweden; [4]Department of Cell and Developmental Biology, John Innes Centre, Norwich, United Kingdom; [5]EMBL Australia Node in Single Molecule Science, School of Medical Sciences, University of New South Wales, Sydney, Australia; [6]State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China; [7]Faculty of Medicine and Health Sciences, Norwich Medical School, University of East Anglia, Norwich, United Kingdom; [8]College of Life Sciences, Wuhan University, Wuhan, China; [9]European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, United Kingdom

***For correspondence:** caroline.dean@jic.ac.uk (CD); stefanie.rosa@slu.se (SR); martin.howard@jic.ac.uk (MH)

[†]These authors contributed equally to this work

Competing interest: The authors declare that no competing interests exist.

Funding: See page 18

Received: 25 April 2022
Preprinted: 05 July 2022
Accepted: 25 June 2023
Published: 19 July 2023

Reviewing Editor: Hao Yu, National University of Singapore & Temasek Life Sciences Laboratory, Singapore

Abstract Quantitative gene regulation at the cell population level can be achieved by two fundamentally different modes of regulation at individual gene copies. A 'digital' mode involves binary ON/OFF expression states, with population-level variation arising from the proportion of gene copies in each state, while an 'analog' mode involves graded expression levels at each gene copy. At the *Arabidopsis* floral repressor *FLOWERING LOCUS C (FLC),* 'digital' Polycomb silencing is known to facilitate quantitative epigenetic memory in response to cold. However, whether *FLC* regulation before cold involves analog or digital modes is unknown. Using quantitative fluorescent imaging of *FLC* mRNA and protein, together with mathematical modeling, we find that *FLC* expression before cold is regulated by both analog and digital modes. We observe a temporal separation between the two modes, with analog preceding digital. The analog mode can maintain intermediate expression levels at individual *FLC* gene copies, before subsequent digital silencing, consistent with the copies switching OFF stochastically and heritably without cold. This switch leads to a slow reduction in *FLC* expression at the cell population level. These data present a new paradigm for gradual repression, elucidating how analog transcriptional and digital epigenetic memory pathways can be integrated.

Editor's evaluation

Regulation of gene expression in many biological systems occurs either digitally where gene expression is either on or off or through an analog mode with graded modulation of gene expression. In this study, the authors report how these two regulatory modes are integrated into a one-way switch pattern to control the expression of the *Arabidopsis* floral repressor gene *FLOWERING LOCUS C* (*FLC*). The results of their work lead the authors to propose that analog regulation in the autonomous flowering pathway precedes digital regulation conferred by Polycomb silencing before cold

exposure, and that this temporal switch correlates with the strength of transcription at the *FLC* locus in different genetic backgrounds.

Introduction

One of the most fundamental questions in molecular biology is how quantitative gene expression is achieved. Traditionally, such regulation is ascribed to sequence-specific transcription factors that bind to regulatory DNA elements. According to the concentration of the transcription factors, gene expression can then be quantitatively up- or downregulated. While such regulation undoubtedly occurs in many systems, it has become abundantly clear in recent years that this paradigm is fundamentally incomplete. This is especially so in eukaryotes where quantitative transcriptional regulation can arise from modulation of the local chromatin environment of a gene. For example, by varying the type and level of histone modifications, DNA accessibility can be radically altered (***Ahmad et al., 2022***). In one scenario, nucleosome positioning affects the ability of transcription factors to bind. Another possibility is that alteration of the chromatin environment directly affects the kinetics of transcription (***Coulon et al., 2014***) by altering how fast the RNA polymerase elongates.

For the *Arabidopsis* floral repressor gene *FLOWERING LOCUS C* (*FLC*), it has been shown that expression levels are quantitatively reduced by a prolonged duration of cold. This quantitative response is achieved through individual *FLC* gene copies making a *cis*-mediated, digital switch from an 'ON' (expressing) state to an 'OFF' (silenced) state (***Angel et al., 2011***). This switch is asynchronous between gene copies, even in the same cell, with the number switched OFF increasing over time in the cold. This results in a gradual decrease in *FLC* expression over time at a whole plant level with silenced gene copies covered by high levels of the silencing histone mark H3K27me3 controlled by the Polycomb system through Polycomb Repressive Complex 2 (PRC2). This mode of regulation is called 'digital,' to highlight the discrete ON/OFF states for each gene copy (***Figure 1A***, Digital Regulation, ***Munsky and Neuert, 2015***). Such digital regulation has also been observed in many other systems, both natural (***Saxton and Rine, 2022***) and engineered (***Bintu et al., 2016***).

An alternative mode of quantitative gene regulation is one that allows graded expression levels to be maintained at each gene copy, rather than just an ON or an OFF state. Quantitative regulation at the cell population level can then be achieved by tuning the expression level uniformly at all gene copies, as in inducible gene expression systems. A well-characterized example of such behavior is in the level of stress-responsive gene expression as controlled by the transcription factor Msn2 in budding yeast (***Stewart-Ornstein et al., 2013***). This graded mode of regulation is called 'analog' (***Figure 1A***, Analog Regulation,), in contrast to the digital alternative.

While *FLC* loci are known to have digital behavior during and after a cold treatment, allowing them to robustly hold epigenetic memory of cold exposure, it remains unknown how the starting expression levels (prior to cold) are regulated quantitatively in terms of analog versus digital control. After plants germinate, they grow as seedlings that have not experienced any cold exposure, or digital Polycomb switching, so quantitative variation in *FLC* expression seen in young seedlings could represent different cellular proportions of digitally regulated *FLC*, as well as graded transcriptional changes. More generally, elucidating the interplay between analog and digital control is essential for a more in-depth understanding of quantitative gene regulation. Although digital and analog modes of repression have been separately studied in the past (see, e.g.,), how these two fundamentally different modes of regulation might be combined has not been considered. *FLC* is an ideal system to study this question due to its digital control after cold, as well as the wealth of knowledge about its regulation at all stages (***Berry and Dean, 2015***; ***Wu et al., 2020***).

FLC levels are set during embryogenesis by competition between an *FLC* activator called *FRIGIDA* (*FRI*) and the so-called autonomous repressive pathway (***Figure 1B***, ***Li et al., 2018***; ***Schon et al., 2021***). The commonly used *Arabidopsis* accessions Ler and Col-0, have mutations in the *FRI* gene, allowing the autonomous pathway to dominate, thereby repressing *FLC* during vegetative development and resulting in a rapidly cycling summer annual lifestyle (***Johanson et al., 2000***). On the other hand, genetic introduction of an active *FRI* allele (Col*FRI*, ***Lee and Amasino, 1995***) generates an initially high *FLC* expression state, which then requires cold for *FLC* repression and subsequent flowering. However, both these cases, with high (Col*FRI*) or low (Ler, Col-0) *FLC* expression, are extreme examples, where essentially all *FLC* loci are either expressed to a high level or not, before cold. It is

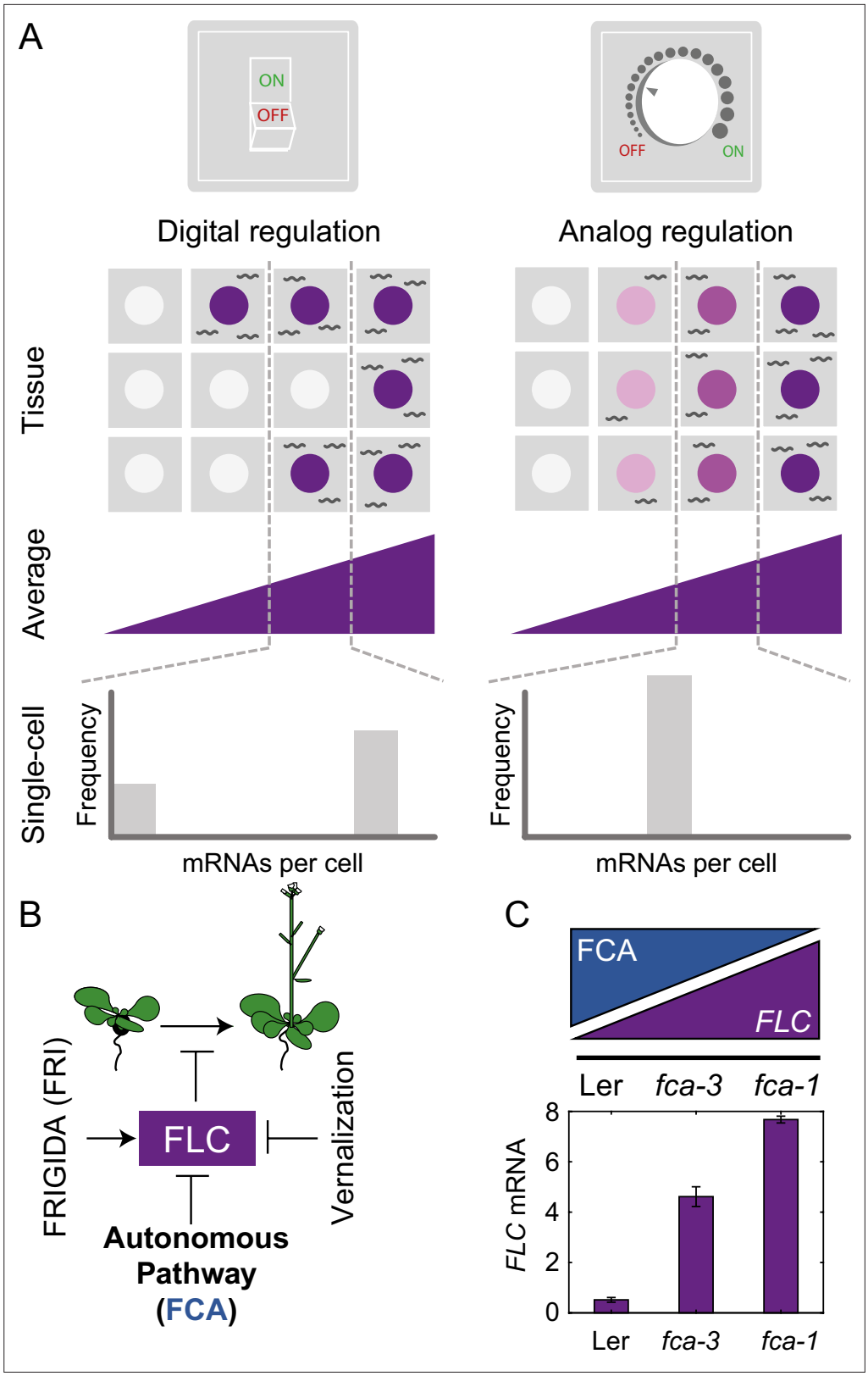


Figure 1. Schematic of digital and analog gene regulation. (**A**) Digital regulation (left) corresponds to loci being in an 'ON' state (purple) or 'OFF' state (white), where we assume for simplicity that there is only one gene copy per cell. At the tissue level, moving from low to high average expression (columns left to right) is achieved by a change in the fraction of cells in each of the two states. This mode is distinct from analog regulation (right), where each

Figure 1 continued on next page

Figure 1 continued

cell has a graded expression level that roughly corresponds to the overall population average. (**B**) *FLC* represses the transition to flowering and is controlled by *FRIGIDA*, the autonomous pathway and the vernalization pathway (inducing digital epigenetic silencing in the cold). (**C**) Reduced *FCA* activity leads to higher cell population-level *FLC* expression in *fca* mutants. Wildtype Ler has lowest *FLC* expression, *fca-3* intermediate and *fca-1* highest. Expression is measured by qPCR relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC*). Error bars show SEM of n = 3 biological replicates measured 7 d after sowing. Statistical tests: multiple comparisons following ANOVA (F-value = 214.62, p-value=$2.6 \cdot 10^{-6}$) with Tukey HSD post hoc tests for *fca-3* – *fca-1*: p-value=0.00029; *fca-3* – Ler: p-value=$5.5 \cdot 10^{-5}$; *fca-1* – Ler: p-value<$2.6 \cdot 10^{-6}$.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Characterization of *fca* alleles and FLC-Venus transgene.

therefore difficult to use these genotypes to understand any possible interplay between digital and analog control as only the extremes are exhibited. Instead, what is required is a genotype with intermediate *FLC* expression at a cell population level during embryogenesis. Such a genotype at a single gene copy level might exhibit either a graded (analog) or all-or-nothing (digital) *FLC* expression before cold, thereby allowing dissection of whether analog or digital regulation is at work.

In this work, we exploited mutant alleles in *FLOWERING CONTROL LOCUS A* (*FCA*), which is part of the *FLC*-repressive autonomous pathway, thereby systematically varying overall *FLC* levels. One of these mutants, *fca-1* (***Koornneef et al., 1991***), is a complete loss of function and therefore exhibits late flowering (***Figure 1—figure supplement 1A***) and high *FLC* expression before cold (***Figure 1C***, ***Figure 1—figure supplement 1B–D***), similar to Col*FRI*. The wildtype, Ler, has a fully functional autonomous pathway and so exhibits low *FLC* levels before cold (***Figure 1C***, ***Figure 1—figure supplement 1B–D***) and early flowering (***Figure 1—figure supplement 1A***), with the *FLC* gene covered by the silencing histone mark H3K27me3 (***Wu et al., 2020***). Crucially, however, the *fca-3* and *fca-4* mutants lead to compromised FCA function (see 'Materials and methods' and ***Koornneef et al., 1991***). In *fca-3*, there is a splice site mutation that leads to changed exon structure through variable use of alternative splice sites with either loss or misfolding of the C-terminus (***Macknight et al., 2002***). The *fca-4* allele is the result of a large genomic inversion that disrupts the *FCA* gene at the 3′ end of exon 4 (***Page et al., 1999***). The expressed 3′ fragment contains the second RNA-binding domain and the C-terminal region of the protein including the WW protein interaction domain, sufficient to give an intermediate flowering time phenotype (***Page et al., 1999***). Both mutants display intermediate cell population-level *FLC* expression (***Figure 1C***, ***Figure 1—figure supplement 1B–D***) and flowering time (***Figure 1—figure supplement 1A***), indicating partial functionality of the FCA protein in these mutants. Crucially, this intermediate property allows us to systematically dissect the interplay of analog and digital regulation at *FLC* before cold using a combination of single cell and whole plant assays, together with mathematical modeling, revealing a temporal separation between the two regulatory modes.

Results

Analysis of *fca* alleles reveals both analog and digital regulation at *FLC*

To investigate the mode of repression arising from regulation by the autonomous pathway, we utilized the *fca-1* and *fca-3* mutants, as well as the parental Ler genotype, and assayed how individual cells varied in their *FLC* expression in these three genotypes at 7 d after sowing. We quantified the number of individual mRNAs per cell using single-molecule fluorescence in situ hybridization (smFISH) (***Figure 2A and B***, ***Figure 2—figure supplement 1A and B***). The endogenous Ler *FLC* carries a Mutator transposable element (TE) in intron 1, silencing *FLC* expression (***Liu et al., 2004***). Possibly because of the TE, we observed *FLC* mRNA accumulation in the nucleolus in this background (***Figure 2—figure supplement 1A***). To avoid this complication, we transformed a Venus-tagged *FLC* into Ler (***Figure 2A***,). The transgenic *FLC* sequence was from Col-0, which does not contain the TE. We then crossed the FLC-Venus into our mutant genotypes (*fca-1*, *fca-3, fca-4*), thus ensuring that the transgene is in the same genomic location and therefore that changes in its expression will be due to the *FCA* mutations in these lines. We also used smFISH probes agaisnt the Venus and *FLC* sequence. These probes generated a signal specific to the transgenic *FLC* copy in these plants

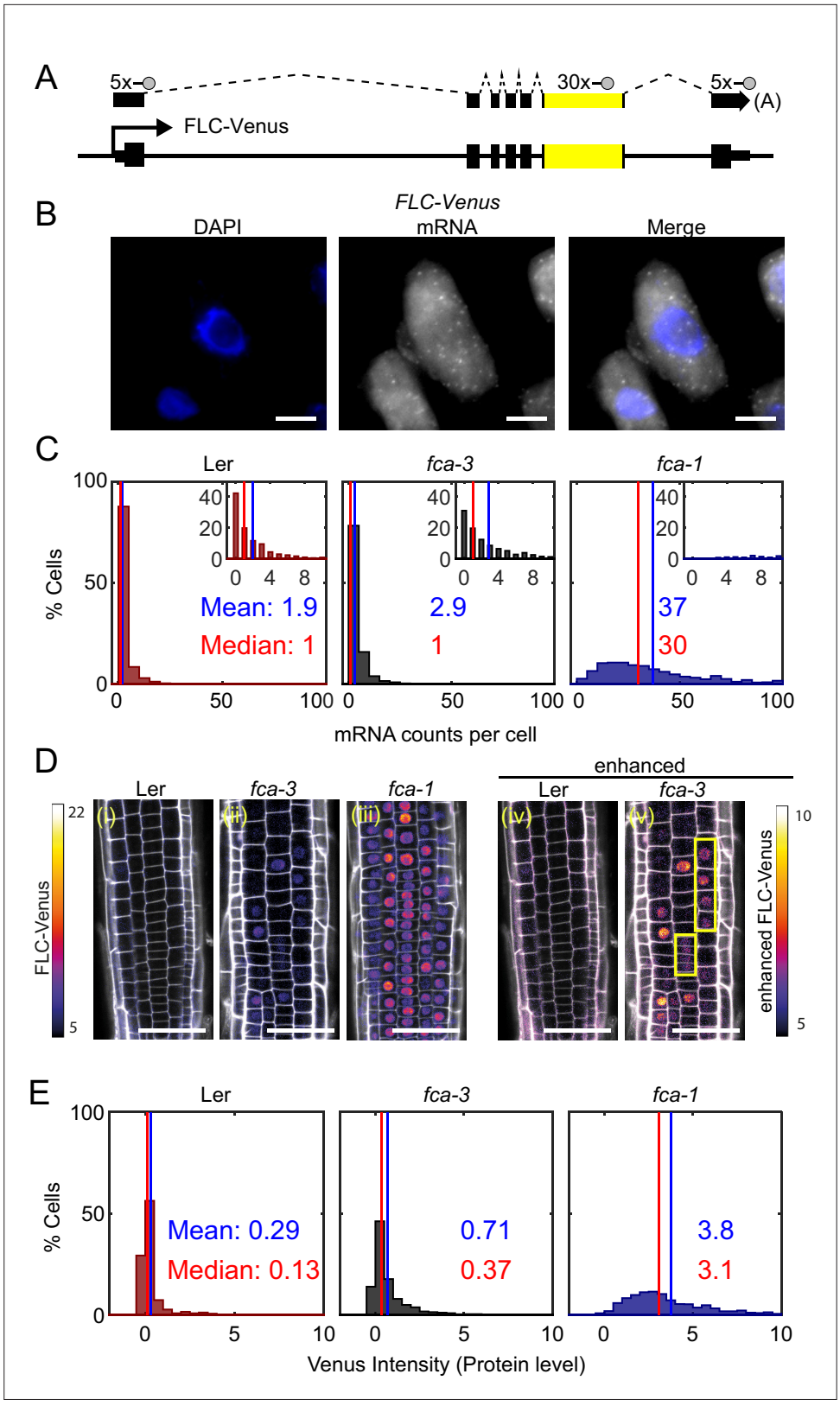


Figure 2. *FLC* expression per cell in *fca* mutants. (**A**) Schematic diagram of *FLC-Venus* locus with transcript and exonic probe position indicated. A total of 40 probes were designed, 10 against the *FLC* sequence and 30 for the *Venus* sequence. (**B**) Detection of *FLC-Venus* transcripts in single cells. Representative images of isolated cells with DAPI staining (blue) and *FLC-Venus* mRNA (gray) obtained from *Arabidopsis thaliana* root squashes. Scale

Figure 2 continued on next page

Figure 2 continued

bar, 5 µm. (**C**) Histograms of single-molecule fluorescence in situ hybridization (smFISH) results for each genotype (Ler, *fca-3*, *fca-1*) showing the number of single-molecule *FLC-Venus* RNAs detected per cell. We show the distribution for Ler, but note that this is likely indistinguishable from background using our methods. The means and medians of the distributions are indicated in each panel. Insets show the same data for mRNA counts between 0 and 10 (where 10 is the 95th percentile of the *fca-3* data). For Ler n = 853, *fca-3* n = 1088, *fca-1* n = 792; data from 7 d after sowing (two independent experiments). Statistical tests: three-way comparison with Kruskal–Wallis ($\chi^2(2) = 1611.21$, p-value=0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: adj. p-valu =$5.5 \cdot 10^{-275}$; *fca-3* – Ler: adj. p-value $1.2 \cdot 10^{-9}$; *fca-1* – Ler: adj. p-value=$2.9 \cdot 10^{-257}$. (**D**) Representative confocal images of roots for each genotype. FLC-Venus intensity indicated by color maps; gray shows the propidium iodide (PI) channel. Same settings were used for imaging and image presentation in (**i–iii**). Images in (**iv**) and (**v**) are the same as (**i**) and (**ii**), respectively, but adjusted to enhance the Venus signal by changing brightness and contrast (please note different scale of color map). Yellow boxes in (**v**) show short files of ON cells. Scale bar, 50 µm. (**E**) Histograms of FLC-Venus intensity per cell in each genotype. The means and medians of the distributions are indicated in each panel. For Ler n = 537 cells from 6 roots, *fca-3* n = 1515 cells from 14 roots, *fca-1* n = 1031 cells from 11 roots; data from 7 d after sowing (two independent experiments). Statistical tests: three-way comparison with Kruskal–Wallis ($\chi^2(2) = 1607.56$, p-value=0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: adj. p-value=$8.7 \cdot 10^{-281}$; *fca-3* – Ler: adj. p-value=$1.7 \cdot 10^{-32}$; *fca-1* – Ler: adj. p-value=$1.2 \cdot 10^{-201}$.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Single-molecule fluorescence in situ hybridization (smFISH) method for FLC-Venus imaging.

Figure supplement 2. Threshold for ON/OFF state of cells in single-molecule fluorescence in situ hybridization (smFISH) experiments.

Figure supplement 3. FLC-Venus imaging in *fca* alleles – root replicates and *fca-4*.

Figure supplement 4. FLC-Venus imaging in *fca* mutants and wildtype, in young leaf tissue.

Figure supplement 5. Segmentation and quantification method for FLC-Venus protein intensity per nucleus.

since we observed no signal in lines without the transgenic *FLC* (***Figure 2—figure supplement 1B***). Comparison of whole-plant gene expression showed similar behavior in the mutants for both the endogenous and transformed *FLC* (***Figure 1—figure supplement 1C and D***). Focusing on the *Venus* sequence conferred the additional advantage that we were able to use the same lines for both mRNA and protein level quantifications (see below).

We observed that the mean number of *FLC* mRNA transcripts per cell was highest in *fca-1*, lowest in Ler, and intermediate in *fca-3*, with the distributions being significantly different (***Figure 2C***, all differences significant with Bonferroni-adjusted $p \leq 1.2 \cdot 10^{-9}$), with 89% of *fca-1* cells having higher expression than the 95th percentile of *fca-3* cells. Furthermore, when looking only at the ON cells (***Figure 2—figure supplement 1C***; ON cells defined here as cells with more than three mRNAs counted, though the conclusion is unchanged with alternative thresholds; ***Figure 2—figure supplement 2***), the mRNA numbers in *fca-3* were only around 1/5 of those in *fca-1*, much less than one-half, ruling out that the reduced levels in *fca-3* were solely due to silencing of one of the two gene copies. The mRNA distribution of *fca-3* cells is significantly different from Ler, confirming that *fca-3* is not simply all digitally OFF. However, we note that the Ler images contain noisy signals due to background (see 'Materials and methods'), resulting in the detection of ON cells with similar mRNA counts as in *fca-3* (***Figure 2—figure supplement 1***). Despite this, *fca-3* has considerably more ON cells than in Ler (***Figure 2—figure supplement 1***), indicating that counts for *fca-3* are not simply background. Visual inspection and manual counting of images further confirmed these conclusions ('Materials and methods'). Nevertheless, mRNA numbers in many *fca-3* cells were close to zero, suggesting that digital silencing may also be relevant in this case. Hence, the differences in overall expression between the mutant genotypes appeared to have two components. Firstly, there were different fractions of apparently silenced cells (digitally OFF) without any appreciable expression (three or fewer mRNAs counted): 72% in *fca-3* and 1.5% in *fca-1*. Secondly, in cells that did express *FLC* (digitally ON), the *FCA* mutations lead to an analog change in their expression levels, so that there was more *FLC* mRNA in *fca-1* ON cells than in the *fca-3* ON cells (***Figure 2—figure supplement 1C***). This behavior clearly differed from the digital regulation observed at *FLC* after cold treatment (***Figure 1A***, ***Angel et al., 2011***; ***Rosa et al., 2016***).

We additionally used confocal live imaging to investigate FLC protein levels in individual cells in the root tip (***Figure 2D***, ***Figure 2—figure supplement 3***) and in young leaves (***Figure 2—figure supplement 4***). We found similar protein expression patterns in roots and leaves, suggesting that each mutation was having similar effects in different tissues. Therefore, conclusions we draw from root experiments can be extended to *FLC* regulation in leaves. Imaging in roots has clear technical advantages, as well as the presence of clonal cell files which can inform heritability, and so further microscopy experiments are in root tissue only. After cell segmentation, we quantified the FLC-Venus intensity within the nuclei, comparing the different genotypes (***Figure 2E***, ***Figure 2—figure supplement 5***; all differences are significant with p). This procedure allowed us to combine information of relative protein levels with the cell positions in the root. Median intensity levels of FLC -Venus per cell and the overall histogram distribution revealed again intermediate levels of FLC protein in *fca-3*, relative to Ler and *fca-1*. At the same time, there was again strong evidence in favor of a digital component for *FLC* regulation (***Figure 2D(v)***): in cells with the lowest protein levels, these levels were similar in *fca-3* and in Ler, again supporting a digital OFF state. Furthermore, we could see a mix of distinct ON and OFF cells by enhancing the *fca-3* images to increase the Venus intensity to a similar level as in *fca-1*. By contrast, in *fca-1* all cells were ON, whereas in Ler, cells appeared OFF even with an equivalent adjustment (***Figure 2D(iv)***, ***Figure 2—figure supplement 3***). We could also infer a potentially heritable component in the ON/OFF states as short files of ON cells could be observed in *fca-3* (***Figure 2D(v)***, ***Figure 2—figure supplement 3***, yellow boxes). In *fca-4*, we observed similar short ON files (***Figure 2—figure supplement 3***, yellow boxes), suggesting that this is a general feature that is not specific to *fca-3*. Overall, our results support a combination of analog and digital regulation for *FLC*: in Ler most cells were digitally OFF, in *fca-1* all cells were digitally ON, while in *fca-3* a fraction of cells were OFF, but for those cells that were ON, the level of *FLC* expression was reduced in an analog way relative to *fca-1*.

FLC RNA and protein are degraded quickly relative to the cell cycle duration

Our results strongly pointed toward a digital switching component being important for *FLC* regulation by the autonomous pathway, but with an analog component too for those loci that remain ON. A study in yeast previously reported on the expected RNA distributions for the two cases of analog and digital control (***Goodnight and Rine, 2020***). An important additional consideration when interpreting the analog/digital nature of the regulation concerns the half-lives of the mRNA and protein. In a digital scenario, long half-lives (***Rahni and Birnbaum, 2019***) are expected to broaden histograms of mRNA/protein levels due to the extended times needed for mRNA/protein levels to increase/decrease after state switching. This could lead to a possible misinterpretation of analog regulation, for example, if intermediate levels of mRNA/protein remain in OFF cells that are descended from ON cells. In contrast, short half-lives will lead to clearer bimodality. Furthermore, what might look like a heritable transcription state could also appear due to slow dilution of a stable protein, as observed in other cases (***Kueh et al., 2013***; ***Zhao et al., 2020***). We therefore needed to measure the half-lives of the mRNA and protein to interpret our observations appropriately.

FLC mRNA has previously been shown to have a half-life of approximately 6 hr (***Ietswaart et al., 2017***) in a different genotype (Col*FRI*) to that used here. We measured the half-lives of both RNA and protein in our highly expressing *FLC* line, *fca-1* (***Figure 3—figure supplement 1***). The RNA half-life measurement used actinomycin D treatment, inhibiting transcription, whereas the protein measurement used cycloheximide, arresting protein synthesis. The half-lives were then extracted from the subsequent decay in mRNA/protein levels. We found that both degradation rates were quite fast, with half-lives of ~5 and ~1.5 hr, respectively, for the mRNA and protein (***Figure 3—figure supplement 1***). These timescales are short compared to the cell cycle duration, here of ~1 d (***Rahni and Birnbaum, 2019***), let alone compared to the timescale of development. Therefore, slow degradation is unlikely to be the cause of the apparent analog regulation and of the observed heritability seen in our root images. Furthermore, the short protein half-life indicates that any potential effects from growth causing dilution, and thus a reduction in protein concentrations, will also be small.

Switching of *FLC* loci to an OFF state over time in *fca-3*

To understand the nature of potential digital switching, it is important to determine whether switching occurs from ON to OFF, OFF to ON, or in both directions. If most loci are switching one-way only, in either direction, this would lead to a gradual change of overall *FLC* expression over time. Alternatively, two-way switching or non-switching in at least a few cells would be necessary to have a constant concentration of *FLC* mRNA/protein over time. These considerations therefore raised the related question of whether cell population-level silencing is at steady state at the time of observation or whether we are capturing a snapshot of a transient behavior, with cells continuing to switch over developmental time.

To test if *FLC* expression is changing over time, we sampled *FLC* expression in the intermediate *fca-3* mutant, as well as *fca-1* and Ler, at 7, 15, and 21 d after sowing. This experiment revealed a decreasing trend in *fca-3* and Ler (***Figure 3A***), which did not seem to be due to a change in *FCA* expression over the same timescale (***Figure 3B***). This was particularly clear when comparing *FLC* and *FCA* in *fca-3* between 7 and 15 d: since FCA is a repressor of *FLC* expression, a decrease in both suggests that FCA is not the cause of the *FLC* expression decrease. We therefore concluded that the most likely explanation was primarily one-way switching from an otherwise heritable *FLC* ON state, to the heritable silenced OFF state occurring digitally and independently at individual loci in *fca-3* and Ler. We note that Ler was already mostly OFF at the starting time of 7 d, and so the significant but very slow rate of decrease (slope: –0.038, p-value: $4.8 \cdot 10^{-6}$, ***Figure 3A***; slope: –0.023, p-value: 0.0030, ***Figure 3—figure supplement 2A***) is interpreted as a small number of remaining ON cells continuing to switch OFF. However, at the later timepoints Ler had transitioned to flowering and so no biological conclusions were drawn from the *FLC* dynamics at those times in this genotype. Rather, this data was used as a negative control for the root imaging experiments below. In *fca-1* a downward trend was not statistically significant (p=0.12) in ***Figure 3A***, likely due to the wide error bars at the first timepoint. In another experiment, however (***Figure 3—figure supplement 2A***), there was slow but significant decrease also in *fca-1*, suggesting that there might be some cells switching OFF also in that case, but more slowly compared to *fca-3*.

By imaging FLC-Venus, we observed that the fraction of ON cells was indeed decreasing in *fca-3*, over the time course (***Figure 3C and D***, ***Figure 3—figure supplement 3***). In fact, after 21 d the pattern of ON/OFF cells in the *fca-3* roots was very similar to that of plants that had experienced cold leading to partial cell population-level *FLC* shutdown, with the majority of cell files being stably repressed, but still with some long files of cells in which *FLC* was ON (compare timepoint 21 in ***Figure 3C*** and figures in ***Berry et al., 2015***). It remains possible that there is a low level of switching in the opposite direction, from OFF to ON, in *fca-3* (see below). The other genotypes did not show this statistically significant decreasing trend in the FLC-Venus data. In *fca-1*, the fraction of ON cells did not show a consistent pattern, while in Ler, the fraction of ON cells remained essentially constant (at a very low level). These results differed from the qPCR data (see above), suggesting in Ler, for example, that epidermis cells in the root tip could switch off early, while a small number of ON cells remain in other tissues of the plant that continue to switch off, thereby explaining the slow decrease in the qPCR experiment (***Figure 3A***, ***Figure 3—figure supplement 2A***).

We also examined whether the change over time in *fca-3* could be due to an analog change in the expression of the ON cells rather than a decreasing number of ON cells. By setting a threshold at an intensity of 1, we look at only the tail of the ON cell distribution (normalized by the total number of these cells in each condition) (***Figure 3—figure supplement 4***). We see that high-intensity cells are still present at later times and do not show a reduction in intensity relative to earlier timepoints. This finding is consistent with digital regulation: a decrease in the number of ON cells, but with all remaining ON cells expressing *FLC* at similarly high levels over time. Overall, our results in *fca-3* are consistent with progressive digital switching of *FLC* loci, primarily in the ON-to-OFF direction.

Mathematical model incorporating digital switching can recapitulate the *FLC* distribution over time in *fca-3*

We finally developed a mathematical model for *FLC* ON/OFF state and protein levels in the root (***Figure 4A***, ***Figure 4—figure supplement 1***, 'Materials and methods') to test if the analog and digital components inferred from the data were sufficient to reproduce the experimentally observed patterns. The model incorporated digital *FLC* state switching embedded in simulated dividing cells

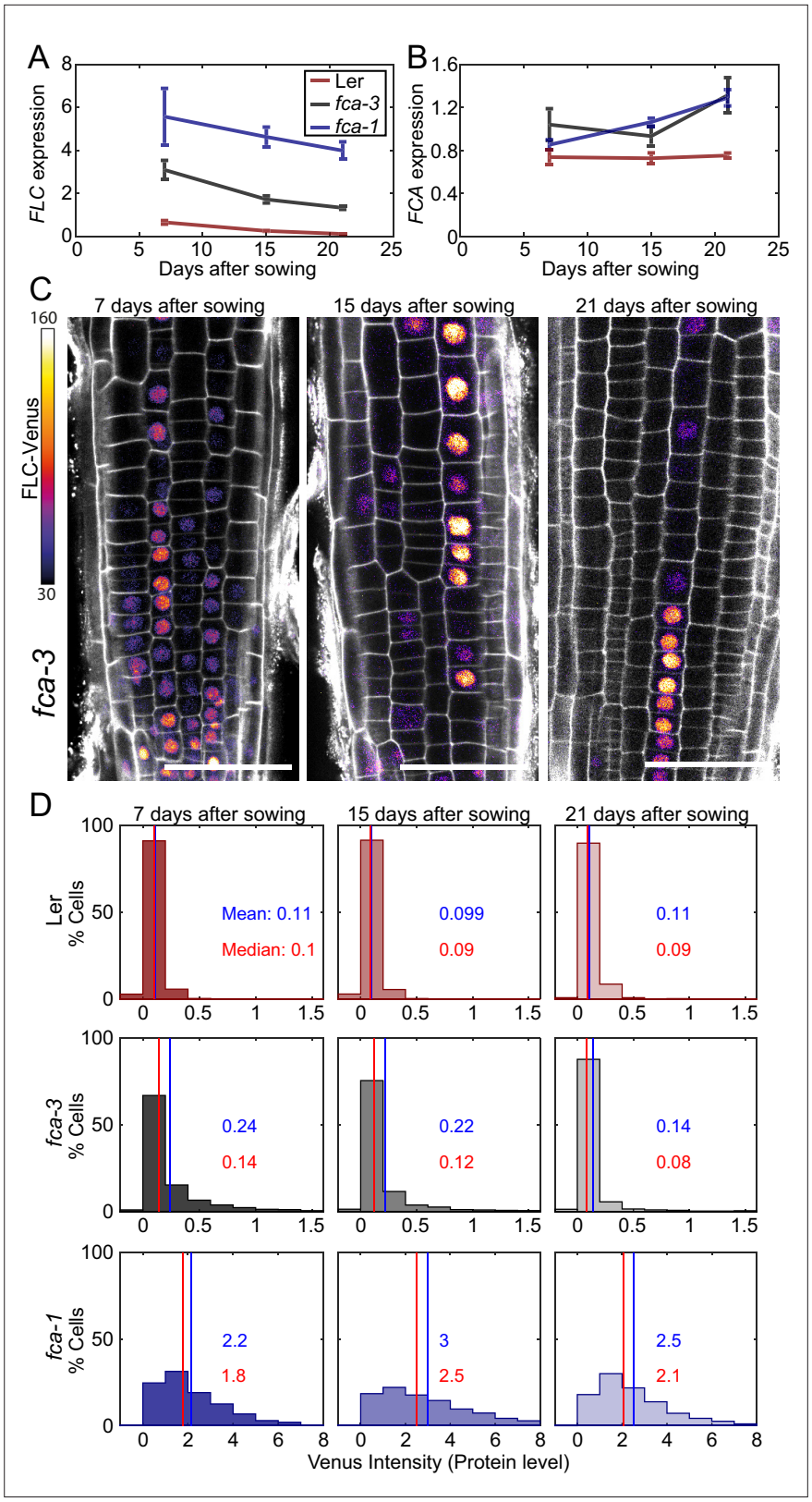


Figure 3. Experimental observation of gradual *FLC* silencing. (**A**) Timeseries of *FLC* expression in *fca* mutant alleles transformed with the FLC-Venus construct. Expression is measured by qPCR in whole seedlings relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC*). Error bars show SEM of n = 3 biological replicates. Statistical tests: samples were excluded as outliers based on Grubbs' test with alpha = 0.05. Linear

Figure 3 continued on next page

Figure 3 continued

regression on timeseries for each genotype. Slope for *fca-3* = −0.13, p-value=$2.0 \cdot 10^{-4}$; slope for *fca-1* = −0.11, p-value=0.12; slope for Ler = −0.038, p-value=$4.8 \cdot 10^{-6}$. (**B**) Timeseries of *FCA* expression, otherwise as in (**A**). Statistical tests: samples were excluded as outliers based on Grubbs' test with alpha = 0.05. Linear regression on timeseries for each genotype. Slope for *fca-3* = 0.021, p-value=0.22; slope for *fca-1* = 0.032, p-value=$3.8 \cdot 10^{-4}$; slope for Ler = $1.1 \cdot 10^{-3}$, p-value=0.84. (**C**) Representative images of *fca-3* roots by confocal microscopy. FLC-Venus intensity indicated by color map; gray shows the PI channel. Same settings were used for imaging and image presentation. Scale bar, 50 μm. (**D**) Histograms of FLC-Venus intensity per cell at each timepoint. The means and medians of the distributions are indicated in each panel. For Ler: 7 d, n = 1121 cells from 10 roots (three independent experiments); 15 d, n = 1311 cells from 9 roots (two independent experiments); 21 d, n = 1679 cells from 12 roots (three independent experiments). For *fca-3*: 7 d, n = 2875 cells from 24 roots (three independent experiments); 15 d, n = 3553 cells from 23 roots (three independent experiments); 21 d, n = 3663 cells from 21 roots (three independent experiments). For *fca-1*: 7 d, n = 1022 cells from 9 roots (three independent experiments); 15 d, n = 1770 cells from 12 roots (three independent experiments); 21 d, n = 2124 cells from 12 roots (three independent experiments). Statistical tests: linear regression on timeseries for each genotype. Slope for *fca-3* = −0.0077, p-value=$4.0 \cdot 10^{-46}$; slope for *fca-1* = 0.018, p-value=0.00064; slope for Ler = -0.00015, p-value=0.44.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Degradation rate of *FLC* mRNA and protein levels.

Figure supplement 2. *FLC* and *FCA* expression in whole seedlings over time .

Figure supplement 3. FLC-Venus time-course replicates in *fca* alleles.

Figure supplement 4. Experimental intensity of FLC-Venus in ON cells does not change over time in *fca-3*.

in the root. We allowed the *FLC* state in the model to switch either ON to OFF, or OFF-to-ON. We focused on modeling the switching dynamics and used distributions for the protein levels in cells with two ON loci, one ON and one OFF, and two OFF empirically fitted to our data (***Figure 4A***, 'Materials and methods'). Possible effects from cell size and burstiness are incorporated into these empirical distributions implicitly through this fitting.

In the plant root, there is a mix of dividing and differentiated cells. Our experimental observations capture cells primarily in the division zone of the root, but even within this region, cell cycle times are not the same in all cells (***Rahni and Birnbaum, 2019***). To generate the model cell files, we used cell cycle lengths based on the literature (***Rahni and Birnbaum, 2019***).

With these assumptions, our model could be fit ('Materials and methods') to replicate the observed pattern of increasing OFF cells in *fca-3* roots, as well as the quantitative histograms for protein levels in *fca-3* (***Figure 4B***). In terms of the switching, we found the best fit where the OFF-to-ON rate is over 10 times slower than the ON-to-OFF rate (***Figure 4B***, ***Supplementary file 1***), supporting predominantly one-way switching from an analog into a digitally silenced state. In addition, the model could capture the longer files present in the later timepoints in the data (T21, ***Figure 3C***, ***Figure 4—figure supplement 1***), but unlike in the data these were not more prevalent at later timepoints than earlier. Therefore, the altered prevalence of this effect in the data may suggest additional developmental influence on the heritability of the ON/OFF state at later times in the plant. Overall, however, the model can faithfully recapitulate the developmental dynamics of *FLC* in *fca-3*.

Discussion

In this work, we have uncovered a combination of analog and digital transcriptional regulation for the gene *FLC*: analog regulation arises through the autonomous pathway, as illustrated by the *fca* mutants, before digital switching into a heritable silenced state. The silenced state is Polycomb dependent, given the similarity of our data to the vernalized state and the role of the Polycomb mark H3K27me3 in silencing *FLC* in wildtype Ler and Col-0 in the absence of cold treatment. In Col-0, in a mutant of the PRC2 component and H3K27 methyltransferase, *clf*, it has also been observed that *FLC* is upregulated and H3K27me3 reduced (***Shu et al., 2019***), though this is only a partial effect due to the presence of an active SWN, another PRC2 H3K27 methyltransferase. Based on previous studies on the interplay between transcription and PRC2 silencing (***Beltran et al., 2016***; ***Berry et al., 2017***; ***Holoch et al., 2021***; ***Lövkvist et al., 2021***), we propose that the rate of digital Polycomb silencing is dependent on the transcription levels. These transcription levels are genetically controlled and

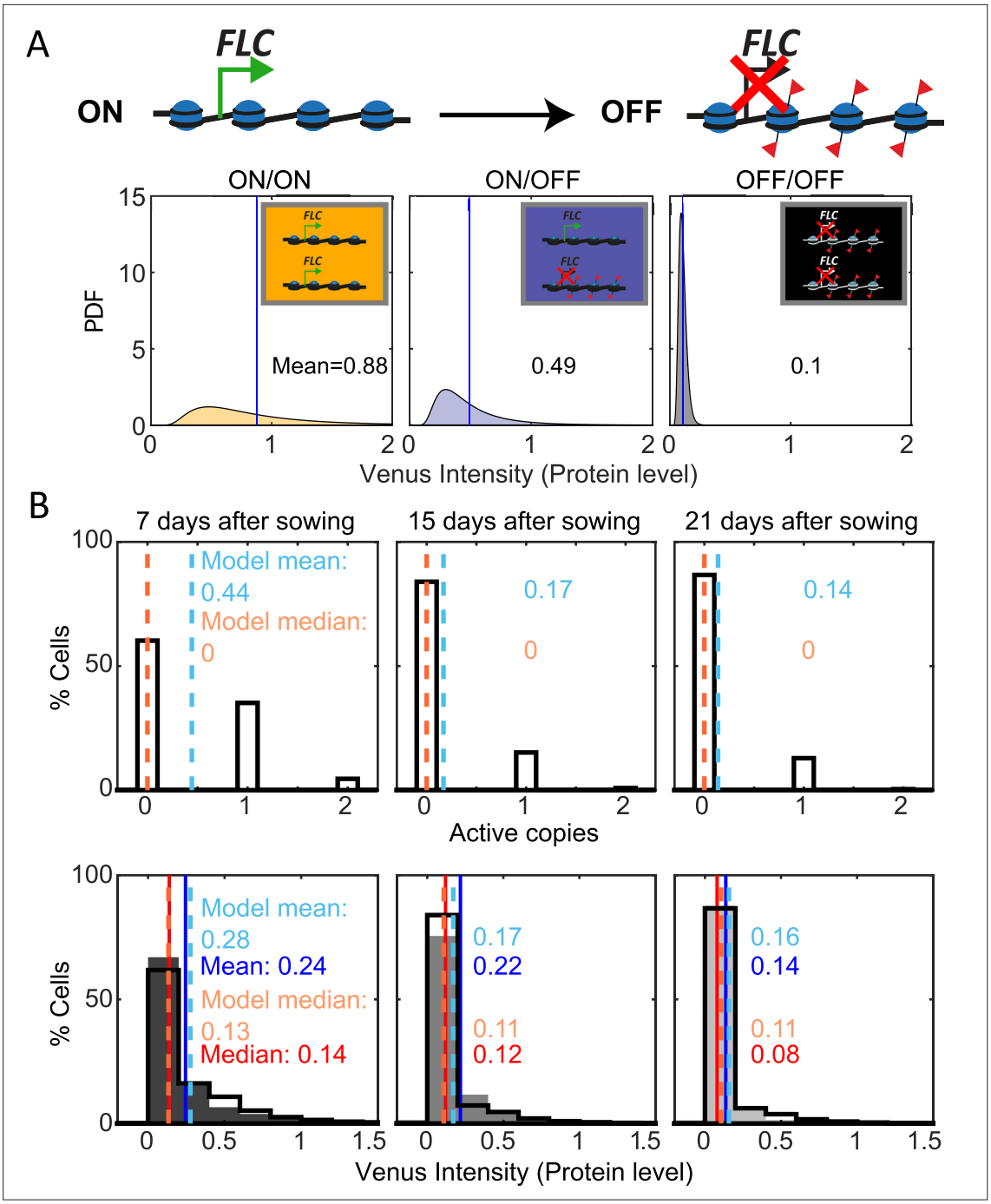


Figure 4. Mathematical model captures *FLC* regulation. (**A**) Diagram of mathematical model. Individual *FLC* gene copies can be ON or OFF, such that a cell can be in one of three states depending on the combination of ON and OFF gene copies within it (ON/ON, ON/OFF, OFF/OFF). Venus intensity (corresponding to amount of protein) within a cell was sampled from the distributions shown (described in 'Materials and methods' section), depending on the cell state. The means of the distributions are indicated in each panel. (**B**) Histograms of active *FLC* copies per cell (top) and Venus intensity per cell (bottom) at the indicated timepoints. Model histograms are plotted with black lines around empty bars and experimental data is shown as filled gray histograms with no outline. The means and medians of the distributions are indicated in each panel for the data (solid lines and matching color text) and the model (dashed lines and matching color text). Model simulated 1000 *fca-3* cell files and histograms shown exclude bottom four cells of each file as cells near the QC are also not included in the imaging.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Modeling *FLC* regulation in clonal cell files in the root.

constitute analog regulation since the pathways governing this process are capable of graded expression control. Analog regulation is set within each genotype, as illustrated by the *fca* mutants, before possible digital switching at each gene copy into a heritable silenced OFF state.

In the case of Ler, low analog transcription is not sufficient to significantly oppose silencing and so the switch occurs before our first experimental measurements at 7 d. In *fca-1,* the switch does not occur at all, prevented by high analog transcription in that genotype. However, for *fca-3,* with intermediate analog *FLC* expression, the digital switching occurs slowly and so can be observed in our timeseries experiments, underlining a clear temporal separation of analog and digital transcriptional control. In this way, both analog and digital regulation are combined at *FLC*, with the timescale for switching between these states controlled by the strength of initial analog transcription. In *fca-3,* such switching to the OFF state causes a gradual reduction in *FLC* expression at the whole plant level. Furthermore, we emphasize that there may be additional developmentally regulated processes occurring at *FLC,* in addition to the constant rate of switching. This possibility is underscored by the model not completely recapitulating the long cell files we observe experimentally at the 21-day timepoint.

A prolonged environmental cold signal can also result in a switch from cells expressing *FLC* to nonexpressing OFF cells, a process that occurs independently at each gene copy of *FLC* (***Berry et al., 2015***). These ON/OFF states are then maintained through cell divisions leading to epigenetic memory. This cold-dependent silencing shows only digital characteristics, with highly *FLC*-expressing cells present both before and after intermediate periods of cold, albeit in a lower proportion after. Here, we have addressed how *FLC* is regulated in the absence of cold and show that a similar mode of regulation occurs. Furthermore, at least in the case of *fca-3* with initially intermediate transcription levels, silencing at the cell population level also happens gradually over many cell cycles, like the behavior over prolonged periods of cold. Moreover, images of roots at 21 d after sowing without cold, with fluorescently labeled FLC, are visually strikingly like images of vernalized roots, where in both cases we find whole files of either all ON or all OFF cells. These comparisons suggest similarities in the molecular characteristics of the Polycomb silenced state at *FLC* generated either with or without cold treatment. We hypothesize that in the absence of active Polycomb the *fca* alleles would display only analog differences, without the possibility of digital silencing.

In this work, we have focused on the early regulation of *FLC* by analog and digital pathways, rather than the effect of these on flowering. However, natural variation at *FLC* largely affects expression as the plants are germinating in autumn (***Hepworth et al., 2020***). The analog component of transcriptional regulation could therefore have widespread ecological relevance. For example, the high *FLC* genotype Lov-1, a natural accession from Sweden, has higher *FLC* expression upon germination in early autumn compared to Col*FRI* (***Coustham et al., 2012***). The analog component could be dominating in Lov-1, with all cells having a higher ON level. Consistent with this, Lov-1 also has lower levels of the Polycomb mark H3K27me3 at the *FLC* locus, with potentially few or no OFF cellsn the warm (***Qüesta et al., 2020***).

The main antagonist of the autonomous pathway is the *FRIGIDA* gene, a transcriptional activator of *FLC* (***Figure 1B***), which was not present in the lines used in this study. In fact, it is the antagonism between FRI and the autonomous pathway that determines the starting *FLC* levels, and so we would expect that FRI controls the analog component directly and digital component indirectly. For future work, it would therefore be interesting to investigate transcriptional responses in an *FRI* allelic series. This could be particularly important because natural accessions show mutations in *FRI* rather than in autonomous pathway components, possibly because the autonomous components regulate many more target genes.

Overall, our work has revealed a combination of both analog and digital modes of regulation at *Arabidopsis FLC* before cold, with analog preceding digital. We further propose it is the strength of the initial analog, autonomous pathway transcriptional level that controls the timescale for the switch into the subsequent digital, Polycomb silenced state. Future work is needed to address how this integrated analog and digital regulation affects flowering time, including further exploration of this pathway in leaves and the use of cell-tracking to capture the switching process in action.

Materials and methods

Plant material

fca alleles

fca-3 and *fca-4* are the result of X-ray mutagenesis, while *fca-1* was induced by EMS (**Koornneef et al., 1991**). For Ler, *fca-3* and *fca-1*, the transcript levels of *FCA* determined with a primer pair binding between exons 16 and 19 are comparable (**Figure 3B**). The primer pair used is specific to the gamma isoform in a wildtype context, which so far has been the only FCA isoform known to generate a functional protein (**Macknight et al., 2002**). The *fca-1* C-T point mutation generates an in-frame early stop codon (**Macknight et al., 1997**), whereas *fca-3* is a splice site mutation from G to A at the 3′ splice site of intron 6 (Chr4:9,210,672). The closest predicted weak alternative splice site (Chr4:9,210,582; 90 bp) is within intron 7, the next constitutive acceptor splice site of intron 7 (Chr4:9,210,518; 154 bp downstream). This suggests that either intron 6 is not spliced out or exon 7 is skipped. Using a C-terminal polyclonal serum antibody, no protein was detected as previously shown in either *fca-3* or *fca-1* (**Macknight et al., 2002**). However, more characterization is needed to determine the exact protein that is made in *fca-3. fca-4* contains a breakpoint within intron 4 of *FCA* leading to a chromosomal rearrangement, with the 3′ fragment of *FCA* fused to another gene, leading to partial protein (**Page et al., 1999**). *fca-3* and *fca-4* were empirically selected as intermediate mutants for this study based on their effects on *FLC* levels and on flowering time (**Figure 1C**, **Figure 1—figure supplement 1**).

Generation of FLC-Venus transgenic lines

pSLJ-755I6 FLC-Venus contains a 12.7 kb genomic fragment containing *FLC* from the Col-0 accession with the Venus coding sequence inserted into the NheI site of *FLC* exon 6, as previously described (**Berry et al., 2015**). This *FLC-Venus* construct was transformed into Ler using *Agrobacterium tumefaciens* and single-copy transgenic lines were selected. We crossed plants that were homozygous for FLC-Venus with the mutant genotypes (*fca-1, fca-3,* and *fca-4*) to obtain F2 plants homozygous for both *FLC-Venus* and *fca-1, fca-3,* or *fca-4*, by PCR-based genotyping (primers in **Supplementary file 2**) and copy number analysis. *fca-4* genotyping was performed using specific forward primers for the *fca-4* mutation (600 bp) and WT (720 bp). The *fca-3* mutation introduces a recognition sequence for the restriction enzyme BglII, generating fragments of 87 bp and 195 bp (*fca-3*) compared to 282 bp (WT). The *fca-1* mutation generates an additional restriction enzyme recognition site for MseI generating fragments of 157 bp, 125 bp, 18 bp (*fca-1*), and 175 bp, 125 bp for WT. F3 plants derived from these were used for all experiments. We verified by qPCR that *FLC-Venus* showed similar changes in expression as endogenous *FLC* in the *fca-1, fca-3,* and *fca-4* backgrounds (**Figure 1—figure supplement 1**).

Plant growth

Seeds were surface sterilized in 5% v/v sodium hypochlorite for 5 min and rinsed three times in sterile distilled water. Seeds were stratified for 2 d at 4°C in Petri dishes containing MS media without glucose. The plates were placed vertically in a growth cabinet (16 hr light, 22°C) for 1 wk.

Gene expression analysis

For gene expression timeseries, 20+ whole seedlings were harvested at each timepoint (7-, 15-, and 21-day-old plants). Plant material was snap frozen with liquid nitrogen, ground, and RNA was extracted using the phenol:chloroform:isoamyl alcohol (25:24:1) protocol (**Yang et al., 2014**). RNA was purified with the TurboDNase (Ambion) kit to remove DNA contamination and reverse transcribed into cDNA using SuperScript IV Reverse transcriptase (Invitrogen) and RT primers for genes of interest. Gene expression was measured by qPCR, and data was normalized to PP2A and UBC, unless specified otherwise. Primer sequences are summarized in **Supplementary file 2**.

Primer pairs:

FLC spliced: FLC_spliced_F / FLC_spliced_R
FLC unspliced: SDB_FLC_4548_F / SDB_FLC_4701_R
FLC-Venus spliced: FLC_spliced_F / SDB_FLC-VENUS_ex6_cDNA_744_R
FLC-Venus unspliced: SDB_FLC_4548_F / SDB_FLC-VENUS_ex6_cDNA_744_R

Actinomycin D treatment

For actinomycin D (ActD) experiments, 6-day-old plants were initially germinated in non-supplemented media and were transferred to new plates containing ActD. The plants were kept in darkness during ActD treatment. Stock solution of ActD (1 mg/mL dissolved in DMSO) was added to molten MS media to a final concentration of 20 µg/mL. ActD was obtained from Sigma (Cat# A4262-2MG).

Cycloheximide (CHX) treatment

FLC-Venus protein stability was assayed with the de novo protein synthesis inhibitor cycloheximide (C1988, Sigma-Aldrich) following the procedures in ***Zhao et al., 2020***. Briefly, 7-day-old *fca-1* seedlings carrying the *FLC-Venus* transgene were treated in liquid MS medium containing 100 µM CHX. The seedlings were sampled after 0, 1.5, 3, 6, 12, and 24 hr of treatment. A nontreatment control is also included, in which seedlings were soaked in liquid MS medium without the inhibitor CHX for 24 hr. Approximately 1.0 ng seedlings were ground to a fine powder with Geno/Grinder. Total protein was extracted with 1.5 mL buffer (50 mM Tris-Cl pH 8.0, 15 4 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.3% NP-40, 1% Triton-100, 5 mM DTT, and protease inhibitor). Then each sample was cleaned by centrifugation at 16,000 × *g* at 4°C for 15 min, 50 µL total protein was taken as input before enrichment with magnetic GFP-trap beads (GTMA-20, Chromo Tek). The input samples were run on a separate gel and used as a processing loading control for the starting level for each sample. The enriched FLC-Venus protein was detected by western blot assay with the antibody anti-GFP (11814460001, Roche). Signals were visualized with chemiluminescence (34095, Pierce) with a secondary antibody conjugated to horseradish peroxidase (NXA931V, GE Healthcare). The chemiluminescence signal was obtained by the FUJI Medical X-ray film (4741019289, FUJI). Quantification was performed with ImageJ after the films were scanned with a printer scanner (RICOH). Ponceau staining was performed with commercial Ponceau buffer (P7170, Sigma-Aldrich) and used as the processing controls. All of the western blot assays were performed with equal weight of whole seedlings.

smFISH

smFISH was carried out on root squashes as described by ***Duncan et al., 2017***. Briefly, root tips from 7-day-old seedlings were cut using a razor blade and placed into glass wells containing 4% paraformaldehyde and fixed for 30 min. Roots were then removed from the fixative and washed twice with nuclease free 1× PBS. Several roots were then arranged on a microscope slide and squashed between the slide and coverslip. This procedure is required to generate a monolayer of cells, thereby decreasing the background, which is necessary to detect the intrinsically low signals from single molecules of mRNA (***Duncan et al., 2016***). Slides were submerged (together with the coverslips) for a few seconds in liquid nitrogen until frozen. The coverslips were then removed, and the roots were left to dry at room temperature for 30 min.

Tissue permeabilization and clearing were achieved by immersing sequentially the samples in 100% methanol for 1 hr, 100% ethanol for 1 hr, and 70% ethanol for a minimum of 1 hr. The ethanol was left to evaporate at room temperature for 5 min and slides were then washed with Stellaris RNA FISH Wash Buffer A (Biosearch Technologies; Cat# SMF-WA1-60). 100 µL of hybridization solution (containing 10% dextran sulfate, 2× SSC, and 10% formamide), with each probe set at a final concentration of 125 nM, was then added to each slide. The slides were left to hybridize at 37°C overnight in the dark.

The hybridization solution containing unbound probes was pipetted out the following morning. Each sample was then washed twice with Stellaris RNA FISH Wash Buffer B (Biosearch Technologies; Cat# SMF-WB1-20) with the second wash left to incubate for 30 min at 37°C. 100 µL of the nuclear stain DAPI (100 ng/mL) was then added to each slide and left to incubate at 37°C for 10 minutes. Slides were then quickly washed with 2× SSC. 100 µL GLOX buffer minus enzymes (0.4% glucose in 10 mM Tris, 2× SSC) was added to the slides and left to equilibrate for 2 min. Finally, this was removed and replaced with 100 µL of GLOX buffer (containing 1 µL of each of the enzymes glucose oxidase [#G0543 from Sigma] and catalase [#C3155 from Sigma]). The samples were then covered by 22 mm × 22 mm No. 1,5 coverslips (VWR), sealed with nail varnish, and immediately imaged.

smFISH probe synthesis

We used the online program Stellaris Probe Designer version 2.0 from Biosearch Technologies to design probe sequences for *FLC-Venus*. For probe sequences, see ***Supplementary file 3*** for *FLC-Venus*. For unspliced *PP2A* and for *FLC* exonic, see ***Duncan et al., 2016*** (reproduced in ***Supplementary file 4***).

Image acquisition

The smFISH slides were imaged using a Zeiss LSM800 inverted microscope, with a 63x water-immersion objective (1.20 NA) and Microscopy Camera Axiocam 503 mono. The following wavelengths were used for fluorescence detection: for probes labeled with Quasar570, an excitation filter 533–558 nm was used and signal was detected at 570–640 nm; for probes labeled with Quasar670, an excitation filter 625–655 nm was used and signal was detected at 665–715 nm; for DAPI, an excitation filter 335–383 nm was used and signal was detected at 420–470 nm; for GFP, an excitation filter 450–490 nm was used and signal was detected at 500–550 nm.

For the FLC-Venus protein level quantification of root samples, optical sections were collected with a Zeiss LSM780 microscope equipped with a Channel Spectral GaAsP detector with a 20x objective (0.8 NA). For z-stacks, the step size was 0.5 µm with a pinhole aperture of 1.5 AU. The overall Z size varied between 45 and 75 slices depending on the orientation of the root. Roots from FLC-Venus lines were immersed in 1 µg/mL propidium iodide (PI, Sigma-Aldrich, P4864) to label the cell wall. For visualization of roots stained with PI, an excitation line of 514 nm was used, and signal was detected at wavelengths of 611–656 nm. For observation of Venus signal, we used a 514 nm excitation line and detected from 518 to 535 nm. To allow comparison between treatments, the same laser power and detector settings were used for all FLC-Venus images.

In FLC-Venus time-course imaging of epidermal root meristems, the Leica SP8X or Stellaris 8 were used with 20x multi-immersion objective (0.75 NA). The Argon (SP8X) or OPSL 514 (Stellaris 8) lasers were used at 5% to excite FLC-Venus and PI at a 514 nm wavelength in bidirectional mode (PI signal was used for set-up). Venus was detected between 518 and 550 nm with the HyD SMD2 detector in photon counting mode; PI was detected at 600–675 nm. Epidermal images were obtained in photon-counting mode with laser speed 200, line accumulation of 6 (pixel dwell time of 2.43 µs), and a Z-step size of 0.95 µm and a pinhole size of 1 AU. For representative images, these were projected such that one single middle slice from the PI channel was used to show the cell outline, onto which 10 slices of FLC-Venus channel were average intensity projected (T7, LSM780 imaging, ***Figure 2***) or sum projected for time-course imaging (T7/T15/T21, SP8X imaging, ***Figure 3***). The dynamic range of the FLC-Venus signal was pushed from 0 to 255 to 5–22 (for LSM780 images, ***Figure 2***) and 30–255 (for SP8X images, ***Figure 3***) for all images apart from where enhanced. In enhanced *fca-3*/Ler images, the dynamic range was further pushed to 5–10 (for LSM780 images, ***Figure 2***) and to 30–160 (for SP8X images, ***Figure 3***) to obtain a similarly strong signal as observed in *fca-1*.

For *fca-4* FLC-Venus imaging, the Zeiss LSM880 was used with a 25x multi-immersion objective (0.8 NA). The Argon laser at 514 nm and 3% was used for excitation of FLC-Venus and PI, with a detection range between 520–550 nm and 600–650 nm, respectively. Z-stacks were taken covering the epidermis, with a step size of 1 µm. In the top two images, the signal was line-averaged four times; for the bottom image, no averaging was performed. For representative images (***Figure 2—figure supplement 3***), five slices of FLC-Venus covering the nucleus were average projected and merged to the middle PI slice. FLC-Venus signal was enhanced to a similar intensity level as in *fca-1* by shifting the dynamic range of the FLC-Venus signal from 0 to 255 to 16–120.

FLC-Venus imaging of young leaves was performed on 9-day-old seedlings. Seedlings were dissected by removing the cotyledons and keeping only the young leaves and shoot meristem. Young leaves were fixed in 4% paraformaldehyde at 4°C overnight. The samples were then treated with 100% methanol and 100% ethanol twice for 15 min each. ClearSee (***Kurihara et al., 2015***) solution was then used to clear the samples for 1 wk. Next, two washes were carried out in 1× PBS. Samples were embedded in a hydrogel according to ***Gordillo et al., 2020***. After rinsing with 1× PBS, samples were stained with Renaissance 2200 solution for 15 mins at room temperature. Slides were mounted in vectashield. FLC-Venus levels of the leaf images were quantified manually using ImageJ. The mean fluorescence intensity of a circular region covering the nucleus in each cell was measured, and the mean background of each image was subtracted.

Image analysis

FISH analysis

Quantification of FISH probes took place in two stages:

1. Identification of probe locations in the whole 3D image, excluding the top and bottom z-slice from each z-stack due to light reflection at the plant cell wall.
2. Assignment of identified probes to specific cells via segmentation of the image into regions.

To detect probes, a white tophat filter was applied to the probe channel, followed by image normalization and thresholding. Individual probes were then identified by connected component segmentation. The centroids of each segmented region were assigned as the probe's location.

Images were manually annotated with markers to indicate positions of nuclei and whether cells were fully visible or occluded. A combination of the visibility markers and nuclei were used to seed a watershed segmentation of the image, thus dividing the image into regions representing individual cells. Probes within each region were counted to generate the per-cell counts. Occluded cells were excluded from the analysis (probe counts for those regions were ignored). Segmentation and probe detection parameters were optimized, with different thresholds used for different replicates and genotypes. This was necessary due to the difference in signal and background of the probes in each of these cases.

Visual inspection of smFISH images revealed high rates of false positive detection in Ler, likely due to very low levels of expression. In the absence of real signals, unspecific signals can be counted through our automated quantification pipeline as a consequence of unspecific probe binding or general background. When expression levels are high, this effect is negligible, but when expression levels are very low, it can become significant. Therefore, differences at the cellular mRNA level between *fca-3* and Ler are likely higher than the ones presented in ***Figure 2C***. The similar count range observed in Ler and *fca-3* 'ON' cells (***Figure 2—figure supplements 1C and 2***) suggests that there is an overlap between the distribution of real ON cells in *fca-3* and background signal (in both Ler and *fca-3*). However, the higher frequency of such cells in *fca-3* compared to Ler confirms that these are not simply the result of background measurements. This is reflected in the significantly different distribution of Ler and *fca-3* cells in ***Figure 2C***. Furthermore, visual inspection of the images confirmed that in *fca-3* we can confidently detect real signals (higher than in Ler and lower than *fca-1*), confirming the analog component.

Custom code is available at https://github.com/JIC-Image-Analysis/fishtools (copy archived at ***Hartley and Antoniou-Kourounioti, 2022a***).

FLC-Venus fluorescence intensity in roots

To measure per-cell FLC-Venus intensities in roots, we developed a custom image analysis pipeline to extract cell structure information from the PI cell wall information and use this to measure per-cell nuclear fluorescence. Images were initially segmented using the SimpleITK (***Beare and Lehmann, 2006***) implementation of the morphological watershed algorithm. Reconstructed regions touching the image boundaries and those below a certain size threshold were removed. Segmentations were then manually curated to merge over-segmented regions and assign file identities to the resulting segmented cells. This curation was performed using custom software, able to merge segmented cells, resegment cells from specified seeds, and split cells along user-defined planes. To approximate the nuclear position, we fitted a fixed size spherical volume of 15 voxels radius to the point of maximal FLC-Venus intensity in each reconstructed cell (***Figure 2—figure supplement 5***). Cells where the sphere overlap fraction (between sphere and cell) was less than 55% were excluded from the analysis. Per-voxel mean FLC-Venus intensities for each cell were then calculated by dividing the summed intensity within the intersection of the spherical region and the cell, by the fixed sphere volume. A background correction was performed by the following calculation to estimate the FLC-Venus intensity in the nucleus ('sphere_filled,' also see ***Supplementary file 5***): sphere_filled = (mean_in_sphere / overlap_fraction) - mean_outside_sphere.

The 'mean_in_sphere' intensity was first divided by the 'overlap_fraction,' the fraction of the sphere which overlaps the cell, to correct for regions of the sphere outside the cell. The mean intensity of the region outside the sphere ('mean_outside_sphere'), which is the background signal, was then subtracted from this overlap-corrected sphere intensity. This background subtraction is also the cause

for the negative values in a fraction of OFF cells where the signal inside and outside the sphere are both background noise. Because of the sphere's prescribed size, in some cases even the brightest sphere inside a cell will have lower intensity than the remaining cell (e.g., by necessarily including the darkest region of the cell), leading to these negative values.

Custom code for initial segmentation and Venus intensity measurement is written in the Python language (***Python Software Foundation, 2023***), and is available at https://github.com/JIC-Image-Analysis/root_measurement (copy archived at ***Hartley and Antoniou-Kourounioti, 2022b***).

The code for the segmentation curation software is available at https://github.com/jfozard/segcorrect (copy archived at ***Fozard, 2022***).

Mathematical model

We constructed a model for *FLC* chromatin states and protein levels in the root (***Figure 4A***) and used it to simulate root cell files and generate simulated protein-level histograms (***Figure 4B***, ***Figure 4—figure supplement 1***).

The root was represented as a collection of cell files (such as the 12 representative cell files shown in ***Figure 4—figure supplement 1***), and 1000 cell files were simulated in total. Each cell file consisted of a list of 30 cells (***Supplementary file 1***), ordered from the root tip toward the rest of the plant. This region was intended to approximately match the division zone and imaging region. Cells were able to divide, giving rise to two daughter cells and pushing cells that were further from the tip upward. Cells escaping the 30 cell limit were removed from the simulation. Cells in a given file were clonally related to each other, and eventually all originated from divisions of the 'initial cells' (the first cells of the cell file, which are adjacent to the quiescent center [QC] of the root apical meristem). Each cell was described by its index along the cell file, its digital chromatin state, its protein concentration (expressed as the Venus intensity to match experimental observations), and its remaining cell cycle duration.

With regard to the digital chromatin state, a cell could be in an ON/ON (both *FLC* copies in the active chromatin state), ON/OFF (one active and one inactive), or OFF/OFF state (both inactive). All cells at the start of the simulation at 0 d are in the ON/ON configuration. The model also included a digital switching process between these cell states. Based on our data, we expected switching to occur at least primarily from a heritable ON state to a heritable OFF state. However, we also simulated switching from OFF to ON in the model to test if we could explain the data also in that case. At each timestep and depending on its digital chromatin state, a single *FLC* copy could switch from an ON to an OFF state with probability p_{OFF} , or from an OFF to an ON state with probability p_{ON} (values in ***Supplementary file 1***).

In order to compare the simulated cell states against the FLC-Venus data, we processed the model outputs using an additional step. This step gave each ON *FLC* copy an associated protein level, which was sampled from a log-normal distribution with parameters: μ_{ON}, σ^2_{ON}. The background within each cell was sampled from a log-normal distribution with parameters: $\mu_{OFF}, \sigma^2_{OFF}$. These four parameters and the switching probabilities p_{OFF} and p_{ON} were manually fitted to the experimental Venus distributions for *fca-3* at 7, 15, and 21 d (***Figures 3D and 4***, ***Supplementary file 1***) by adjusting the parameters and visually inspecting until the model fits were judged to be satisfactory. Protein levels for each cell, according to the combination of ON and OFF *FLC* copies (***Figure 4A***), were given by appropriate combinations of random variables, specifically,

- for ON/ON cells: $2e^{X_{ON}} + e^{X_{OFF}}$ (cell background and signal from two ON copies);
- for ON/OFF cells: $e^{X_{ON}} + e^{X_{OFF}}$ (cell background and signal from one ON copy); and
- for OFF/OFF cells: $e^{X_{OFF}}$ (only the cell background),

where $X_{ON} \sim N\left(\mu_{ON}, \sigma^2_{ON}\right)$ and $X_{OFF} \sim N\left(\mu_{OFF}, \sigma^2_{OFF}\right)$. The noisy nucleus background signal (X_{OFF}) was assumed to be higher than any real signal originating from OFF *FLC* copies. The log-normal distribution was selected for simplicity and automatic non-negativity, and we found empirically that it gave reasonable fits.

The duration of the cell cycle of each cell was determined by the cell's position along the cell file at the time of the division that created it. These cell cycle times according to position were based on the literature (***Rahni and Birnbaum, 2019***). We used a truncated normal distribution with the mean and standard deviation matching the measured values for epidermis cells. A minimum value was set such

that the cell cycle could not be shorter than 13 hr (the lowest value observed for epidermis or cortex cells in ***Rahni and Birnbaum, 2019***).

The simulation process was as follows in each cell file, at each timestep:

Division:

- Remaining cell cycle duration is reduced by timestep in all cells.
- Cells with ≤ 0 remaining cell cycle duration divide, such that the positional index of cells with higher index than the dividing cells is increased by 1.
- For each dividing cell, the two daughter cells (dividing cell and cell with positional index +1 relative to it) are assigned new cell cycle durations according to their position.
- Daughter cells are assigned the same digital chromatin state as the mother cell.
- Cells with positional index greater than 30 are removed from the simulation.

Switching:

- For each cell in the cell file, a random number (r_i) is generated, which is compared to the switching probability as followsL
- If the cell is in the OFF/OFF state:
  - With probability p_{ON}^2 it will switch to the ON/ON state – if $r_i < p_{ON}^2$
  - With probability $2p_{ON}\left(1-p_{ON}\right)$ it will switch to the ON/OFF state – if $r_i \geq p_{ON}^2$ and $r_i < \left(p_{ON}^2 + 2p_{ON}\left(1-p_{ON}\right)\right)$
  - With probability $\left(1-p_{ON}\right)^2$ it will stay in the OFF/OFF state – if $r_i \geq \left(p_{ON}^2 + 2p_{ON}\left(1-p_{ON}\right)\right)$
- If the cell is in the ON/OFF state:
  - With probability $p_{OFF}\left(1-p_{ON}\right)$ it will switch to the OFF/OFF state – if $r_i < \left(p_{OFF}\left(1-p_{ON}\right)\right)$
  - With probability $p_{ON}\left(1-p_{OFF}\right)$ it will switch to the ON/ON state – if $r_i \geq \left(p_{OFF}\left(1-p_{ON}\right)\right)$ and $r_i < \left(p_{OFF}\left(1-p_{ON}\right) + p_{ON}\left(1-p_{OFF}\right)\right)$
  - With probability $\left(p_{ON}p_{OFF} + \left(1-p_{ON}\right)\left(1-p_{OFF}\right)\right)$ it will stay in the ON/OFF state – if $r_i \geq \left(p_{OFF}\left(1-p_{ON}\right) + p_{ON}\left(1-p_{OFF}\right)\right)$
- If the cell is in the ON/ON state:
  - With probability p_{OFF}^2 it will switch to the OFF/OFF state – if $r_i < p_{OFF}^2$
  - With probability $2p_{OFF}\left(1-p_{OFF}\right)$ it will switch to the ON/OFF state – if $r_i \geq p_{OFF}^2$ and $r_i < \left(p_{OFF}^2 + 2p_{OFF}\left(1-p_{OFF}\right)\right)$
  - With probability $\left(1-p_{OFF}\right)^2$ it will stay in the ON/ON state – if $r_i \geq \left(p_{OFF}^2 + 2p_{OFF}\left(1-p_{OFF}\right)\right)$

Custom code is available at https://github.com/ReaAntKour/fca_alleles_root_model (copy archived at ***Antoniou-Kourounioti, 2023***).

Acknowledgements

We thank all members of the Dean, Rosa and Howard groups for excellent discussions. Special thanks to Dr. Cecilia Lövkvist for project coordination. Additionally, we would like to thank all the funders listed below.

Additional information

Funding

Funder	Grant reference number	Author
Biotechnology and Biological Sciences Research Council	BB/P020380/1	Caroline Dean Caroline Dean
Vetenskapsrådet	2018-04101	Anis Meschichi Anis Meschichi
Biotechnology and Biological Sciences Research Council	BB/P013511/1	Caroline Dean Caroline Dean

Funder	Grant reference number	Author
HORIZON EUROPE Marie Sklodowska-Curie Actions	813282	Svenja Reeck Svenja Reeck
HORIZON EUROPE Marie Sklodowska-Curie Actions	MSCA-IF 101032710	Lihua Zhao

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Rea L Antoniou-Kourounioti, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing; Anis Meschichi, Formal analysis, Investigation, Methodology, Writing – original draft; Svenja Reeck, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing; Scott Berry, Conceptualization, Investigation, Writing – review and editing; Govind Menon, Validation, Investigation; Yusheng Zhao, Resources, Investigation, Methodology; John Fozard, Formal analysis, Software; Terri Holmes, Huamei Wang, Investigation; Lihua Zhao, Methodology; Matthew Hartley, Conceptualization, Resources, Software, Formal analysis, Validation, Investigation, Writing – review and editing; Caroline Dean, Conceptualization, Resources, Supervision, Funding acquisition, Investigation, Writing – original draft, Writing – review and editing; Stefanie Rosa, Conceptualization, Formal analysis, Supervision, Funding acquisition, Methodology, Writing – original draft, Project administration, Writing – review and editing; Martin Howard, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Rea L Antoniou-Kourounioti http://orcid.org/0000-0001-5226-521X
Anis Meschichi http://orcid.org/0000-0001-8946-6023
Svenja Reeck https://orcid.org/0000-0002-6362-5310
Scott Berry https://orcid.org/0000-0002-1838-4976
Govind Menon https://orcid.org/0000-0002-1028-5463
Yusheng Zhao https://orcid.org/0000-0002-5893-504X
John Fozard https://orcid.org/0000-0001-9181-8083
Terri Holmes https://orcid.org/0000-0002-8480-0755
Lihua Zhao https://orcid.org/0000-0002-1758-9873
Matthew Hartley https://orcid.org/0000-0001-6178-2884
Caroline Dean https://orcid.org/0000-0002-6555-3525
Stefanie Rosa https://orcid.org/0000-0002-8100-1253
Martin Howard https://orcid.org/0000-0001-7670-0781

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.79743.sa1
Author response https://doi.org/10.7554/eLife.79743.sa2

Additional files

Supplementary files

- Supplementary file 1. Mathematical model parameters.
- Supplementary file 2. Primers used in this study.
- Supplementary file 3. smFISH probe sequences used to detect FLC Venus transcripts. These probes were labeled with Quasar570.
- Supplementary file 4. smFISH probe sequences from ***Duncan et al., 2016*** (**A**) used to detect unspliced PP2A transcripts and (**B**) used to detect *FLC* sense spliced. Both probe sets were labeled with Quasar670.
- Supplementary file 5. Raw data from this study.
- MDAR checklist

Data availability

All microscopy data has been deposited to the BioImage Archive with Accession code S-BIAD425 (https://www.ebi.ac.uk/biostudies/studies/S-BIAD425). Derived quantification data from images, as well as qPCR data, are provided as ***Supplementary file 5***. All code is available through GitHub repositories, as described in the 'Materials and methods'.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Antoniou-Kourounioti R, Meschischi A, Reeck S, Berry S, Menon G, Zhao Y, Fozard JA, Holmes T, Wang H, Hartley M, Dean C, Rosa S, Howard M	2022	Integrating analog and digital modes of gene expression at Arabidopsis FLC	https://www.ebi.ac.uk/biostudies/BioImages/studies/S-BIAD425?query=S-BIAD425	BioImage Archive, S-BIAD425

References

Ahmad K, Henikoff S, Ramachandran S. 2022. Managing the steady state chromatin landscape by nucleosome dynamics. *Annual Review of Biochemistry* **91**:183–195. DOI: https://doi.org/10.1146/annurev-biochem-032620-104508, PMID: 35303789

Angel A, Song J, Dean C, Howard M. 2011. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476**:105–108. DOI: https://doi.org/10.1038/nature10241, PMID: 21785438

Antoniou-Kourounioti R. 2023. Fca_Alleles_Root_Model. swh:1:rev:d9915b5637c17bf0f0234b5353fcd48d7382062e. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:87018302934fc2d1582439551fd5a12169e126c3;origin=https://github.com/ReaAntKour/fca_alleles_root_model;visit=swh:1:snp:784fbaba0071b5f5ad21aea41cec6fcd24b5d84f;anchor=swh:1:rev:d9915b5637c17bf0f0234b5353fcd48d7382062e

Beare R, Lehmann G. 2006. The watershed transform in ITK - discussion and new developments. *The Insight Journal* **202**:1–24. DOI: https://doi.org/10.54294/lf8u75

Beltran M, Yates CM, Skalska L, Dawson M, Reis FP, Viiri K, Fisher CL, Sibley CR, Foster BM, Bartke T, Ule J, Jenner RG. 2016. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Research* **26**:896–907. DOI: https://doi.org/10.1101/gr.197632.115, PMID: 27197219

Berry S, Dean C. 2015. Environmental perception and epigenetic memory: mechanistic insight through FLC. *The Plant Journal* **83**:133–148. DOI: https://doi.org/10.1111/tpj.12869, PMID: 25929799

Berry S, Hartley M, Olsson TSG, Dean C, Howard M. 2015. Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. *eLife* **4**:e07205. DOI: https://doi.org/10.7554/eLife.07205, PMID: 25955967

Berry S, Dean C, Howard M. 2017. Slow Chromatin Dynamics allow Polycomb target genes to filter fluctuations in transcription factor activity. *Cell Systems* **4**:445–457. DOI: https://doi.org/10.1016/j.cels.2017.02.013, PMID: 28342717

Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M, Elowitz MB. 2016. Dynamics of epigenetic regulation at the single-cell level. *Science* **351**:720–724. DOI: https://doi.org/10.1126/science.aab2956, PMID: 26912859

Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, Larson DR. 2014. Kinetic competition during the transcription cycle results in stochastic RNA processing. *eLife* **3**:e03939. DOI: https://doi.org/10.7554/eLife.03939, PMID: 25271374

Coustham V, Li P, Strange A, Lister C, Song J, Dean C. 2012. Quantitative modulation of polycomb silencing underlies natural variation in vernalization. *Science* **337**:584–587. DOI: https://doi.org/10.1126/science.1221881, PMID: 22798408

Duncan S, Olsson TSG, Hartley M, Dean C, Rosa S. 2016. A method for detecting single mRNA molecules in *Arabidopsis thaliana Plant Methods* **12**:13. DOI: https://doi.org/10.1186/s13007-016-0114-x, PMID: 28035231

Duncan S, Olsson TSG, Hartley M, Dean C, Rosa S. 2017. Single molecule RNA FISH in Arabidopsis root cells. *BIO-PROTOCOL* **7**:e2240. DOI: https://doi.org/10.21769/BioProtoc.2240

FozardJ. 2022. Segcorrect. swh:1:rev:e8d9b79982a0beae2a765e949679281cc91ffe3a. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:c469a18d85ac6c5a25a080f0d3705213b8f58ea8;origin=https://github.com/jfozard/segcorrect;visit=swh:1:snp:ae45790bb006eab36f0c8d8d004c5b9c20e64b11;anchor=swh:1:rev:e8d9b79982a0beae2a765e949679281cc91ffe3a

Goodnight D, Rine J. 2020. S-phase-independent silencing establishment in *Saccharomyces cerevisiae eLife* **9**:e58910. DOI: https://doi.org/10.7554/eLife.58910, PMID: 32687055

Gordillo SVG, Escobar-Guzman R, Rodriguez-Leal D, Vielle-Calzada JP, Ronceret A. 2020. Whole-mount immunolocalization procedure for plant female meiocytes. *Methods in Molecular Biology* **2061**:13–24. DOI: https://doi.org/10.1007/978-1-4939-9818-0_2, PMID: 31583649

HartleyM, Antoniou-Kourounioti R. 2022a. Fishtools. swh:1:rev:5a0c5415cf074ff5a725e25b87abe9bc5b1e2e71. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:9b50eb9240900a01b72dc5112689384bfc1c7ba4;origin=https://github.com/JIC-Image-Analysis/fishtools;visit=swh:1:snp:ec231aa406bc68115285c02f9ff111934de188ef;anchor=swh:1:rev:5a0c5415cf074ff5a725e25b87abe9bc5b1e2e71

HartleyM, Antoniou-Kourounioti R. 2022b. Root_Measurement. swh:1:rev:6d7fda6f57199d17dee98b70ff16e32a386864f1. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:5c8099a5d81acace2d36d92f1b5287209365bebc;origin=https://github.com/JIC-Image-Analysis/root_measurement;visit=swh:1:snp:b0649defb67a6f680ef36d0393e78b8d3357f09a;anchor=swh:1:rev:6d7fda6f57199d17dee98b70ff16e32a386864f1

Hepworth J, Antoniou-Kourounioti RL, Berggren K, Selga C, Tudor EH, Yates B, Cox D, Collier Harris BR, Irwin JA, Howard M, Säll T, Holm S, Dean C. 2020. Natural variation in autumn expression is the major adaptive determinant distinguishing Arabidopsis FLC haplotypes. *eLife* **9**:e57671. DOI: https://doi.org/10.7554/eLife.57671, PMID: 32902380

Holoch D, Wassef M, Lövkvist C, Zielinski D, Aflaki S, Lombard B, Héry T, Loew D, Howard M, Margueron R. 2021. A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nature Genetics* **53**:1686–1697. DOI: https://doi.org/10.1038/s41588-021-00964-2, PMID: 34782763

Ietswaart R, Rosa S, Wu Z, Dean C, Howard M. 2017. Cell-size-dependent transcription of FLC and its antisense long non-coding RNA COOLAIR explain cell-to-cell expression variation. *Cell Systems* **4**:622–635. DOI: https://doi.org/10.1016/j.cels.2017.05.010, PMID: 28624615

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C. 2000. Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science* **290**:344–347. DOI: https://doi.org/10.1126/science.290.5490.344, PMID: 11030654

Koornneef M, Hanhart CJ, van der Veen JH. 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Molecular & General Genetics* **229**:57–66. DOI: https://doi.org/10.1007/BF00264213, PMID: 1896021

Kueh HY, Champhekar A, Nutt SL, Elowitz MB, Rothenberg EV. 2013. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**:670–673. DOI: https://doi.org/10.1126/science.1240831, PMID: 23868921

Kurihara D, Mizuta Y, Sato Y, Higashiyama T. 2015. ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. *Development* **142**:4168–4179. DOI: https://doi.org/10.1242/dev.127613, PMID: 26493404

Lee I, Amasino RM. 1995. Effect of vernalization, photoperiod and light quality on the flowering phenotype of Arabidopsis plants containing the FRIGIDA gene. *Plant Physiology* **108**:157–162. DOI: https://doi.org/10.1104/pp.108.1.157, PMID: 12228459

Li Z, Jiang D, He Y. 2018. FRIGIDA establishes a local chromosomal environment for FLOWERING LOCUS C mRNA production. *Nature Plants* **4**:836–846. DOI: https://doi.org/10.1038/s41477-018-0250-6, PMID: 30224662

Liu J, He Y, Amasino R, Chen X. 2004. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. *Genes & Development* **18**:2873–2878. DOI: https://doi.org/10.1101/gad.1217304, PMID: 15545622

Lövkvist C, Mikulski P, Reeck S, Hartley M, Dean C, Howard M. 2021. Hybrid protein assembly-histone modification mechanism for PRC2-based epigenetic switching and memory. *eLife* **10**:e66454. DOI: https://doi.org/10.7554/eLife.66454, PMID: 34473050

Macknight R, Bancroft I, Page T, Lister C, Schmidt R, Love K, Westphal L, Murphy G, Sherson S, Cobbett C, Dean C. 1997. FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. *Cell* **89**:737–745. DOI: https://doi.org/10.1016/s0092-8674(00)80256-1, PMID: 9182761

Macknight R, Duroux M, Laurie R, Dijkwel P, Simpson G, Dean C. 2002. Functional significance of the alternative transcript processing of the Arabidopsis floral promoter FCA. *The Plant Cell* **14**:877–888. DOI: https://doi.org/10.1105/tpc.010456, PMID: 11971142

Munsky B, Neuert G. 2015. From analog to digital models of gene regulation. *Physical Biology* **12**:045004. DOI: https://doi.org/10.1088/1478-3975/12/4/045004, PMID: 26086470

Page T, Macknight R, Yang CH, Dean C. 1999. Genetic interactions of the Arabidopsis flowering time gene FCA, with genes regulating floral initiation. *The Plant Journal* **17**:231–239. DOI: https://doi.org/10.1046/j.1365-313x.1999.00364.x, PMID: 10097382

Python Software Foundation. 2023. The Python Language Reference. 3. Python. https://docs.python.org/3/reference/

Qüesta JI, Antoniou-Kourounioti RL, Rosa S, Li P, Duncan S, Whittaker C, Howard M, Dean C. 2020. Noncoding SNPs influence a distinct phase of Polycomb silencing to destabilize long-term epigenetic memory at Arabidopsis FLC. *Genes & Development* **34**:446–461. DOI: https://doi.org/10.1101/gad.333245.119, PMID: 32001513

Rahni R, Birnbaum KD. 2019. Week-long imaging of cell divisions in the Arabidopsis root meristem. *Plant Methods* **15**:30. DOI: https://doi.org/10.1186/s13007-019-0417-9, PMID: 30988691

Rosa S, Duncan S, Dean C. 2016. Mutually exclusive sense-antisense transcription at FLC facilitates environmentally induced gene repression. *Nature Communications* **7**:13031. DOI: https://doi.org/10.1038/ncomms13031, PMID: 27713408

Saxton DS, Rine J. 2022. Distinct silencer states generate epigenetic states of heterochromatin. *bioRxiv*. DOI: https://doi.org/10.1101/2022.02.01.478725

Schon M, Baxter C, Xu C, Enugutti B, Nodine MD, Dean C. 2021. Antagonistic activities of cotranscriptional regulators within an early developmental window set FLC expression level. *PNAS* **118**:e2102753118. DOI: https://doi.org/10.1073/pnas.2102753118, PMID: 33879620

Shu J, Chen C, Thapa RK, Bian S, Nguyen V, Yu K, Yuan ZC, Liu J, Kohalmi SE, Li C, Cui Y. 2019. Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in Arabidopsis seedlings. *Plant Direct* **3**:e00100. DOI: https://doi.org/10.1002/pld3.100, PMID: 31245749

Stewart-Ornstein J, Nelson C, DeRisi J, Weissman JS, El-Samad H. 2013. Msn2 coordinates a stoichiometric gene expression program. *Current Biology* **23**:2336–2345. DOI: https://doi.org/10.1016/j.cub.2013.09.043, PMID: 24210615

Wu Z, Fang X, Zhu D, Dean C. 2020. Autonomous pathway: FLOWERING LOCUS C repression through an antisense-mediated chromatin-silencing mechanism. *Plant Physiology* **182**:27–37. DOI: https://doi.org/10.1104/pp.19.01009, PMID: 31740502

Yang H, Howard M, Dean C. 2014. Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at Arabidopsis FLC. *Current Biology* **24**:1793–1797. DOI: https://doi.org/10.1016/j.cub.2014.06.047, PMID: 25065750

Zhao Y, Antoniou-Kourounioti RL, Calder G, Dean C, Howard M. 2020. Temperature-dependent growth contributes to long-term cold sensing. *Nature* **583**:825–829. DOI: https://doi.org/10.1038/s41586-020-2485-4, PMID: 32669706

RESEARCH ARTICLE

Figures and figure supplements

Integrating analog and digital modes of gene expression at *Arabidopsis FLC*

Rea L Antoniou-Kourounioti, Anis Meschichi and Svenja Reeck *et al*.

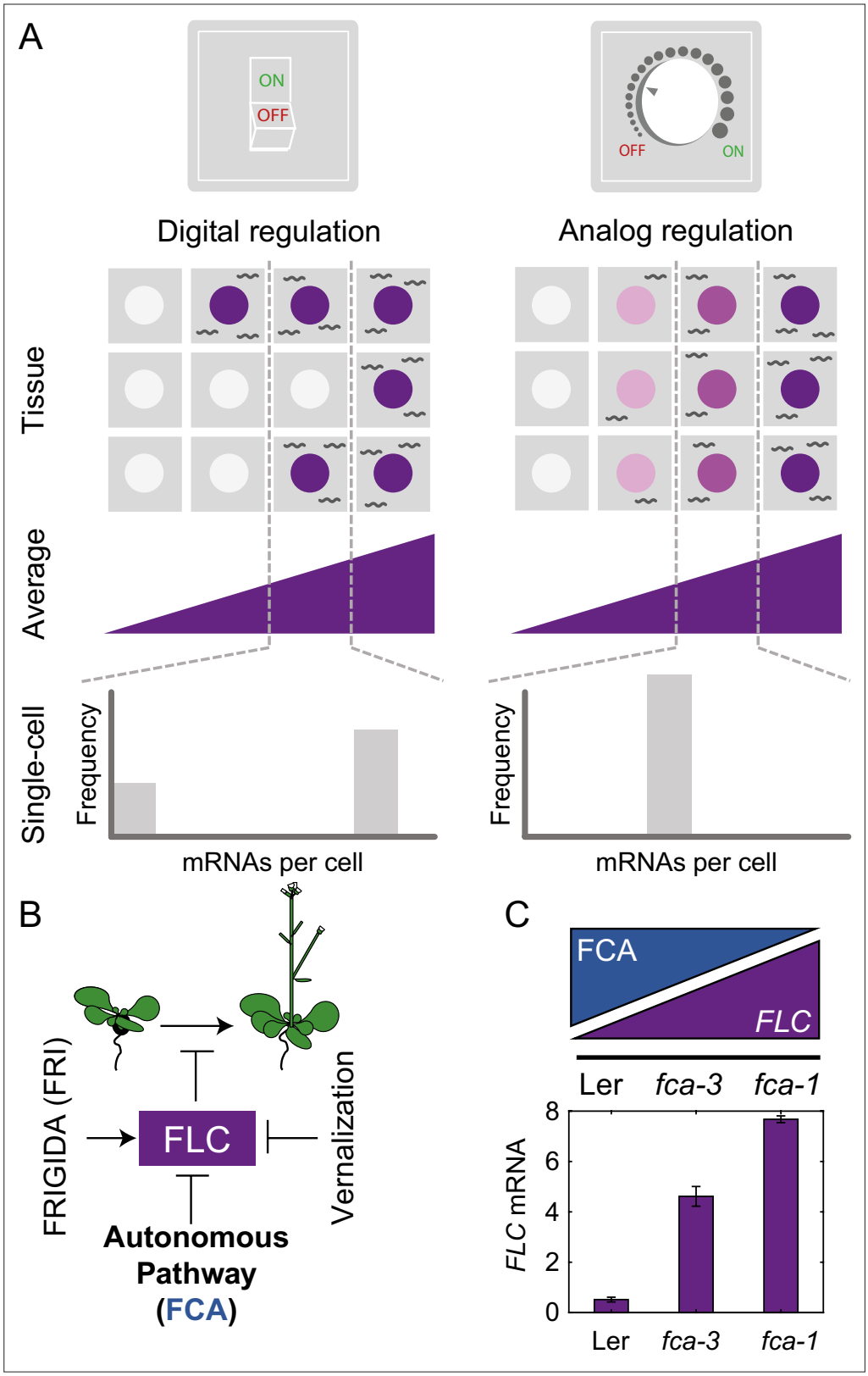


Figure 1. Schematic of digital and analog gene regulation. (**A**) Digital regulation (left) corresponds to loci being in an 'ON' state (purple) or 'OFF' state (white), where we assume for simplicity that there is only one gene copy per cell. At the tissue level, moving from low to high average expression (columns left to right) is achieved by a change in the fraction of cells in each of the two states. This mode is distinct from analog regulation (right), where each

Figure 1 continued on next page

Figure 1 continued

cell has a graded expression level that roughly corresponds to the overall population average. (**B**) *FLC* represses the transition to flowering and is controlled by *FRIGIDA*, the autonomous pathway and the vernalization pathway (inducing digital epigenetic silencing in the cold). (**C**) Reduced *FCA* activity leads to higher cell population-level *FLC* expression in *fca* mutants. Wildtype Ler has lowest *FLC* expression, *fca-3* intermediate and *fca-1* highest. Expression is measured by qPCR relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC*). Error bars show SEM of n = 3 biological replicates measured 7 d after sowing. Statistical tests: multiple comparisons following ANOVA (F-value = 214.62, p-value=$2.6 \cdot 10^{-6}$) with Tukey HSD post hoc tests for *fca-3* – *fca-1*: p-value=0.00029; *fca-3* – Ler: p-value=$5.5 \cdot 10^{-5}$; *fca-1* – Ler: p-value<$2.6 \cdot 10^{-6}$.

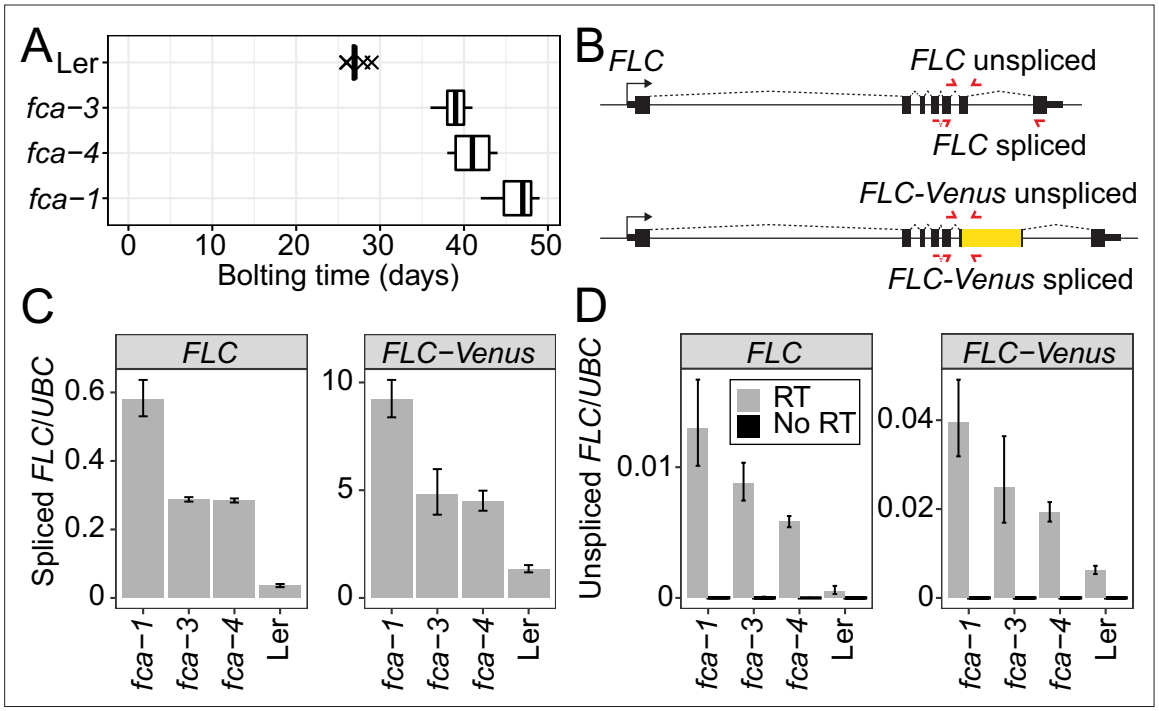


Figure 1—figure supplement 1. Characterization of *fca* alleles and FLC-Venus transgene. (**A**) Flowering time measured in days from sowing until first appearance of a floral meristem (bolting). 12 plants were measured for each genotype. Boxplots show median and interquartile range (IQR), with data points outside median ±1.5 IQR shown as crosses. (**B**) Diagram showing positions of qPCR primers (red) to quantify spliced and unspliced endogenous *FLC* and transgenic *FLC-Venus* RNA. (**C, D**) *FLC* and *FLC-Venus* RNA measured by qPCR from whole seedlings, relative to *UBC* RNA. No RT (reverse transcription)-negative control in (**D**) shows no expression with same unspliced primers. Error bars show SEM of n = 2 biological replicates for *fca-3* and n = 3 biological replicates for other genotypes, harvested 14 d after sowing.

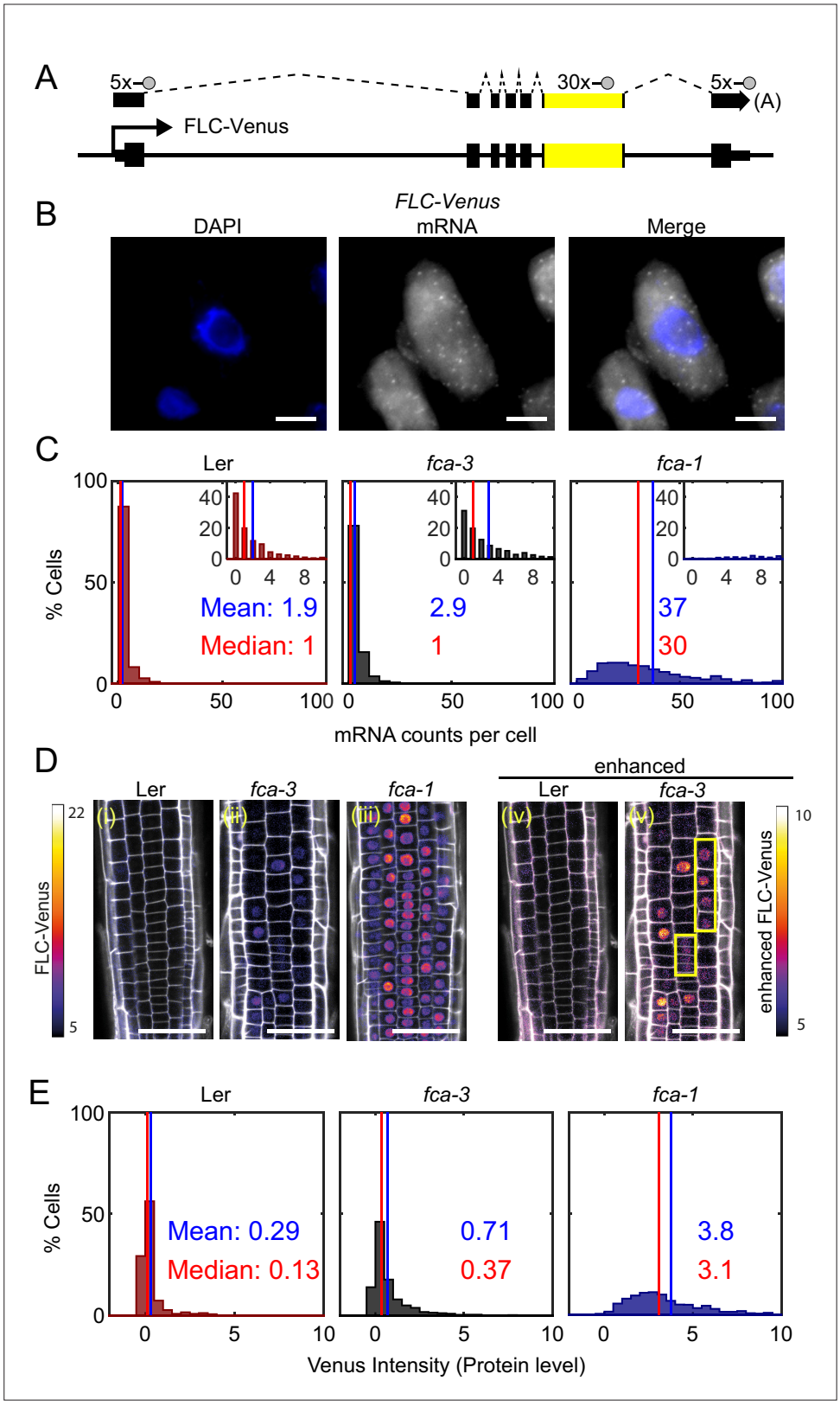


Figure 2. *FLC* expression per cell in *fca* mutants. (**A**) Schematic diagram of *FLC-Venus* locus with transcript and exonic probe position indicated. A total of 40 probes were designed, 10 against the *FLC* sequence and 30 for the *Venus* sequence. (**B**) Detection of *FLC-Venus* transcripts in single cells. Representative images of isolated cells with DAPI staining (blue) and *FLC-Venus* mRNA (gray) obtained from *Arabidopsis thaliana* root squashes. Scale

Figure 2 continued on next page

Figure 2 continued

bar, 5 µm. (**C**) Histograms of single-molecule fluorescence in situ hybridization (smFISH) results for each genotype (Ler, *fca-3*, *fca-1*) showing the number of single-molecule *FLC-Venus* RNAs detected per cell. We show the distribution for Ler, but note that this is likely indistinguishable from background using our methods. The means and medians of the distributions are indicated in each panel. Insets show the same data for mRNA counts between 0 and 10 (where 10 is the 95th percentile of the *fca-3* data). For Ler n = 853, *fca-3* n = 1088, *fca-1* n = 792; data from 7 d after sowing (two independent experiments). Statistical tests: three-way comparison with Kruskal–Wallis ($\chi^2\left(2\right) = 1611.21$, p-value=0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: adj. p-valu =5.5 · 10^{-275}; *fca-3* – Ler: adj. p-value 1.2 · 10^{-9}; *fca-1* – Ler: adj. p-value=2.9 · 10^{-257}. (**D**) Representative confocal images of roots for each genotype. FLC-Venus intensity indicated by color maps; gray shows the propidium iodide (PI) channel. Same settings were used for imaging and image presentation in (**i–iii**). Images in (**iv**) and (**v**) are the same as (**i**) and (**ii**), respectively, but adjusted to enhance the Venus signal by changing brightness and contrast (please note different scale of color map). Yellow boxes in (**v**) show short files of ON cells. Scale bar, 50 µm. (**E**) Histograms of FLC-Venus intensity per cell in each genotype. The means and medians of the distributions are indicated in each panel. For Ler n = 537 cells from 6 roots, *fca-3* n = 1515 cells from 14 roots, *fca-1* n = 1031 cells from 11 roots; data from 7 d after sowing (two independent experiments). Statistical tests: three-way comparison with Kruskal–Wallis ($\chi^2\left(2\right) = 1607.56$, p-value=0) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: adj. p-value=8.7 · 10^{-281}; *fca-3* – Ler: adj. p-value=1.7 · 10^{-32}; *fca-1* – Ler: adj. p-value=1.2 · 10^{-201}.

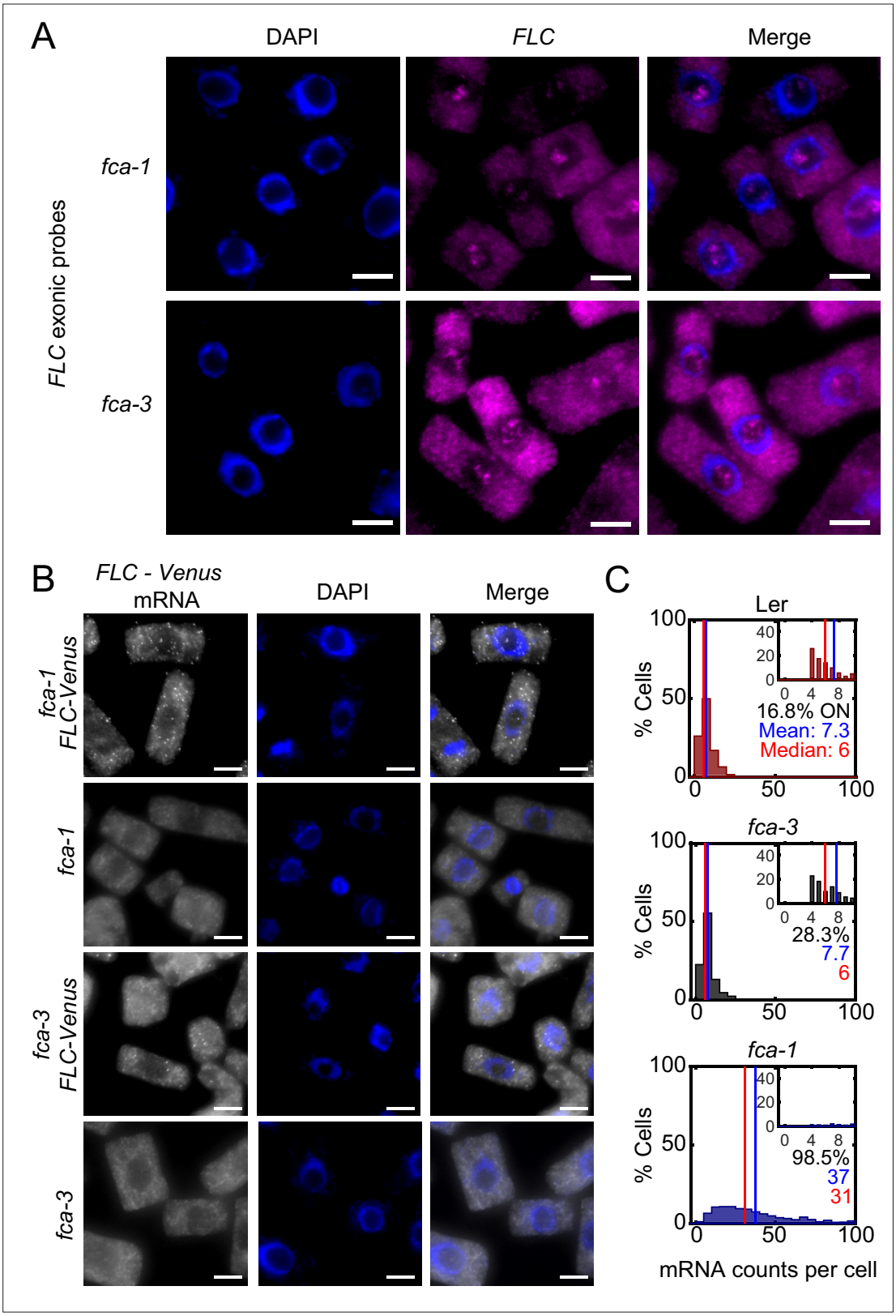


Figure 2—figure supplement 1. Single-molecule fluorescence in situ hybridization (smFISH) method for FLC-Venus imaging. (**A**) Detection of nucleolar signal using *FLC*-specific probes (magenta) in Ler background *fca-1* and *fca-3* plants in *Arabidopsis thaliana* root cells (monolayer of cells derived from root squashes). The nucleus is stained with DAPI (blue). Scale bar, 5 μm. (**B**) FLC-Venus probes (gray) bind specifically to the transgenic *FLC-Venus* RNA and not to the endogenous *FLC* transcript. All data from seedlings 7 d after sowing. Scale bar, 5 μm. (**C**) The data of ***Figure 2C*** plotted for the three

Figure 2—figure supplement 1 continued on next page

Figure 2—figure supplement 1 continued

genotypes but showing only measurements in 'ON cells.' These are defined as cells with more than three *FLC-Venus* mRNA. The means and medians of the distributions are indicated in each panel, as well as the percentage of ON cells. For Ler n = 143, *fca-3* n = 308, *fca-1* n = 780, from 7 d post sowing. Statistical tests: three-way comparison with Kruskal–Wallis ($\chi^2\left(2\right) = 673.43$, p-value=$5.9 \cdot 10^{-147}$) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: p-value=$1.4 \cdot 10^{-113}$; *fca-3* – Ler: p-value=1; *fca-1* – Ler: p-value=$1.7 \cdot 10^{-64}$. We note that the 'ON cells' in Ler come primarily from background signal.

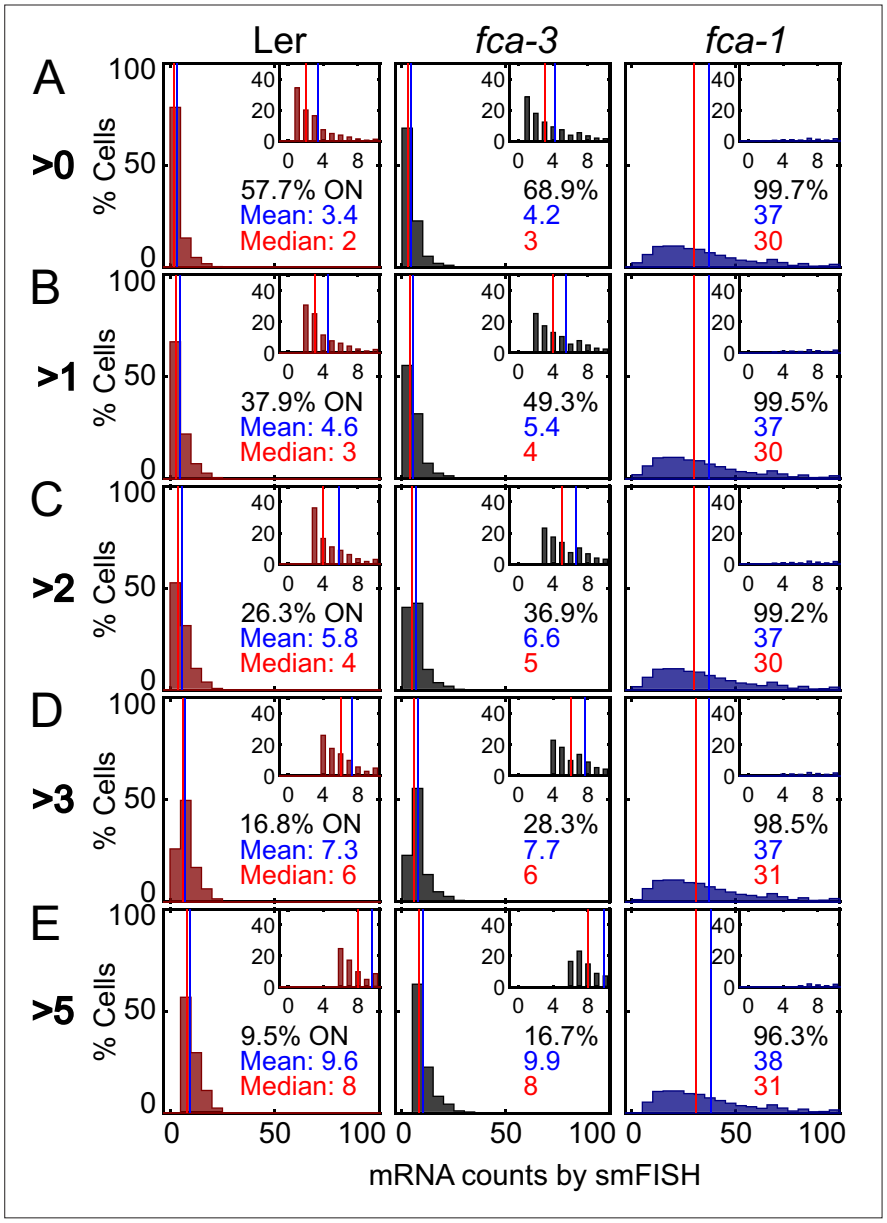


Figure 2—figure supplement 2. Threshold for ON/OFF state of cells in single-molecule fluorescence in situ hybridization (smFISH) experiments. ***Figure 2—figure supplement 1C*** replotted for different thresholds 0 (**A**), >1 (**B**), >2 (**C**), >3 (**D**) and >5 (**E**) mRNAs. Red number indicates the median of the ON cells, while the blue number indicates the mean, and the black number is the percentage of cells that satisfy the threshold condition. We note that the 'ON cells' in Ler come primarily from background signal and the consistently lower percentage of such cells compared to *fca-3* reflects this.

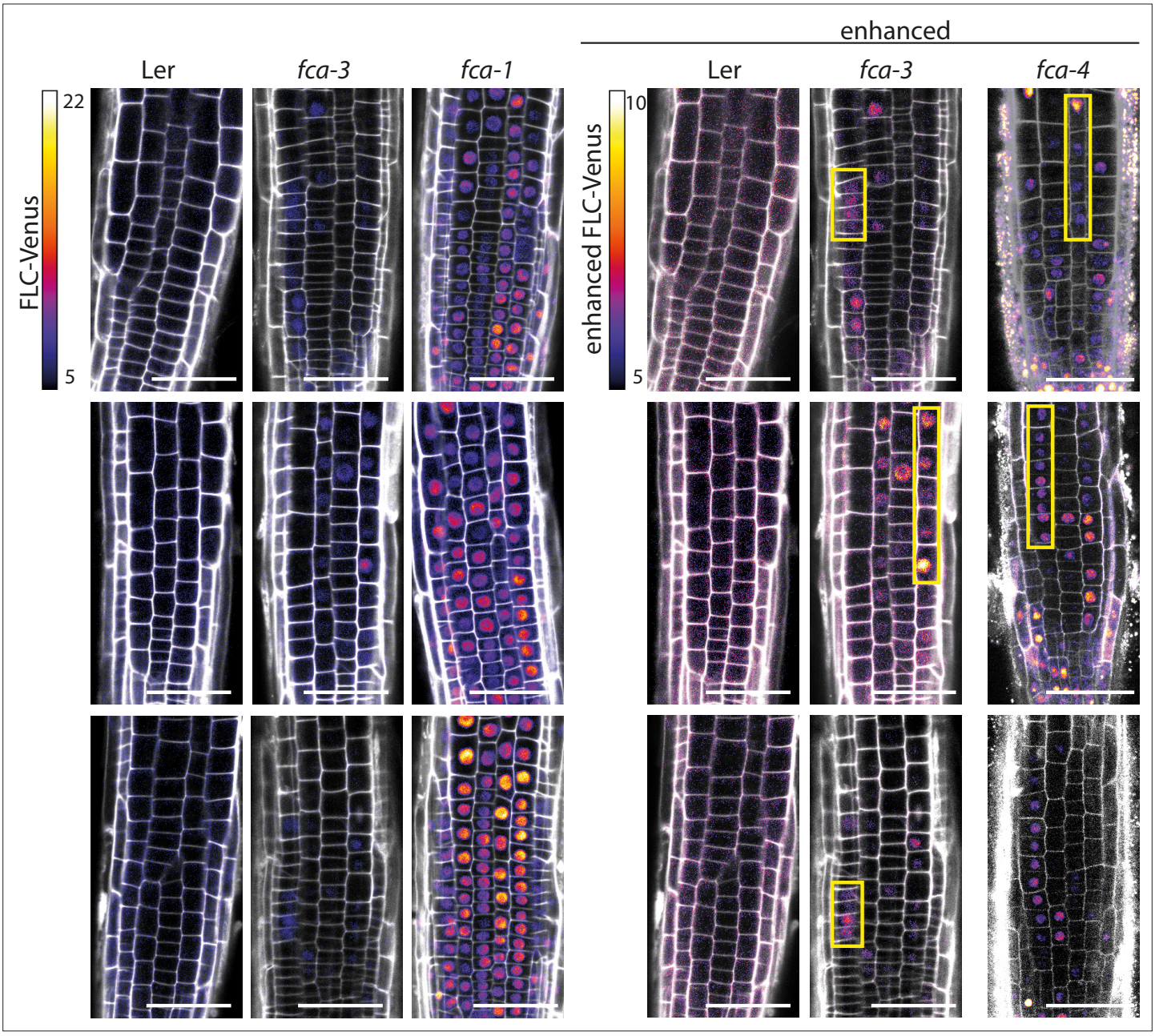


Figure 2—figure supplement 3. FLC-Venus imaging in *fca* alleles – root replicates and *fca-4*. Representative images of FLC-Venus signal in four genotypes (Ler, *fca-3*, *fca-1*, *fca-4*) from 7 d post sowing. FLC-Venus intensity indicated by color maps; gray shows the propidium iodide (PI) channel. Same settings were used for imaging and image presentation in left three columns. 'Enhanced' images (three columns on right) are the same respective images for Ler and *fca-3*, adjusted to enhance FLC-Venus signal by changing brightness and contrast (please note different scale of color map). *fca-4* images were acquired on a different microscope and were enhanced to a comparable level as in Ler and *fca-3* (enhanced). Yellow boxes show files of ON cells in *fca-3* and *fca-4*. Scale bar, 50 μm.

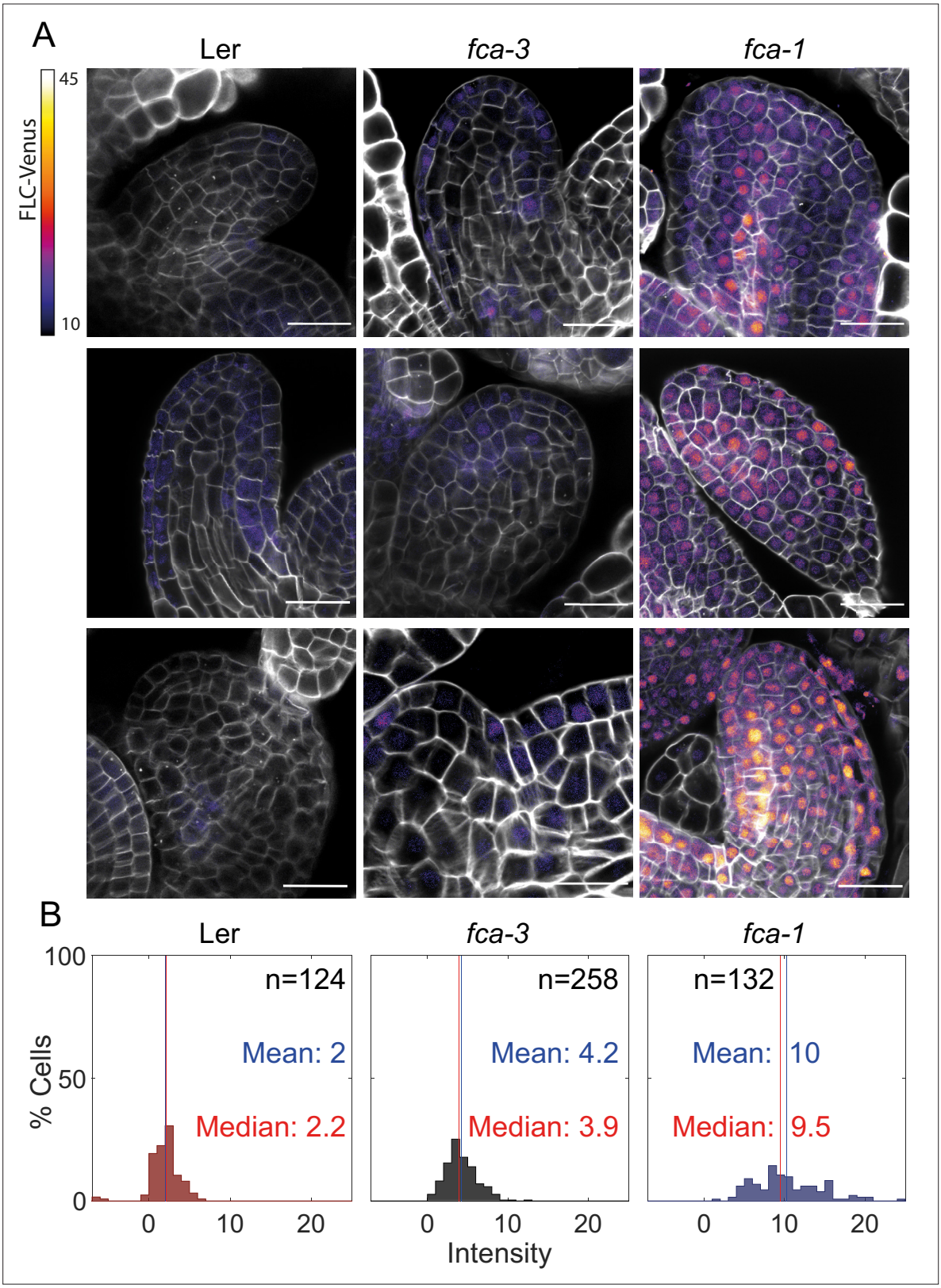


Figure 2—figure supplement 4. FLC-Venus imaging in *fca* mutants and wildtype, in young leaf tissue. (**A**) Three representative images of FLC-Venus signal in the three genotypes (*Ler*, *fca-3*, *fca-1*) from 9 d post sowing. FLC-Venus intensity indicated by color map; gray shows the cell wall dye (SR2200). Scale bar, 20 µm. (**B**) Histograms of FLC-Venus intensity per cell in each genotype. The cell numbers (n), means, and medians of the distributions are indicated in each panel. Data from 9 d after sowing (two experiments). Statistical tests: three-way comparison with Kruskal–Wallis ($X^2(2) = 292.27$, p-

Figure 2—figure supplement 4 continued on next page

Figure 2—figure supplement 4 continued

value=3.4 · 10^{-64}) and pairwise comparisons with Wilcoxon rank sum tests with Bonferroni-adjusted p-values for *fca-3* – *fca-1*: adj. p-value=2.7 · 10^{-42}; *fca-3* – Ler: adj. p-value=3.2 · 10^{-22}; *fca-1* – Ler: adj. p-value=1.4 · 10^{-40}.

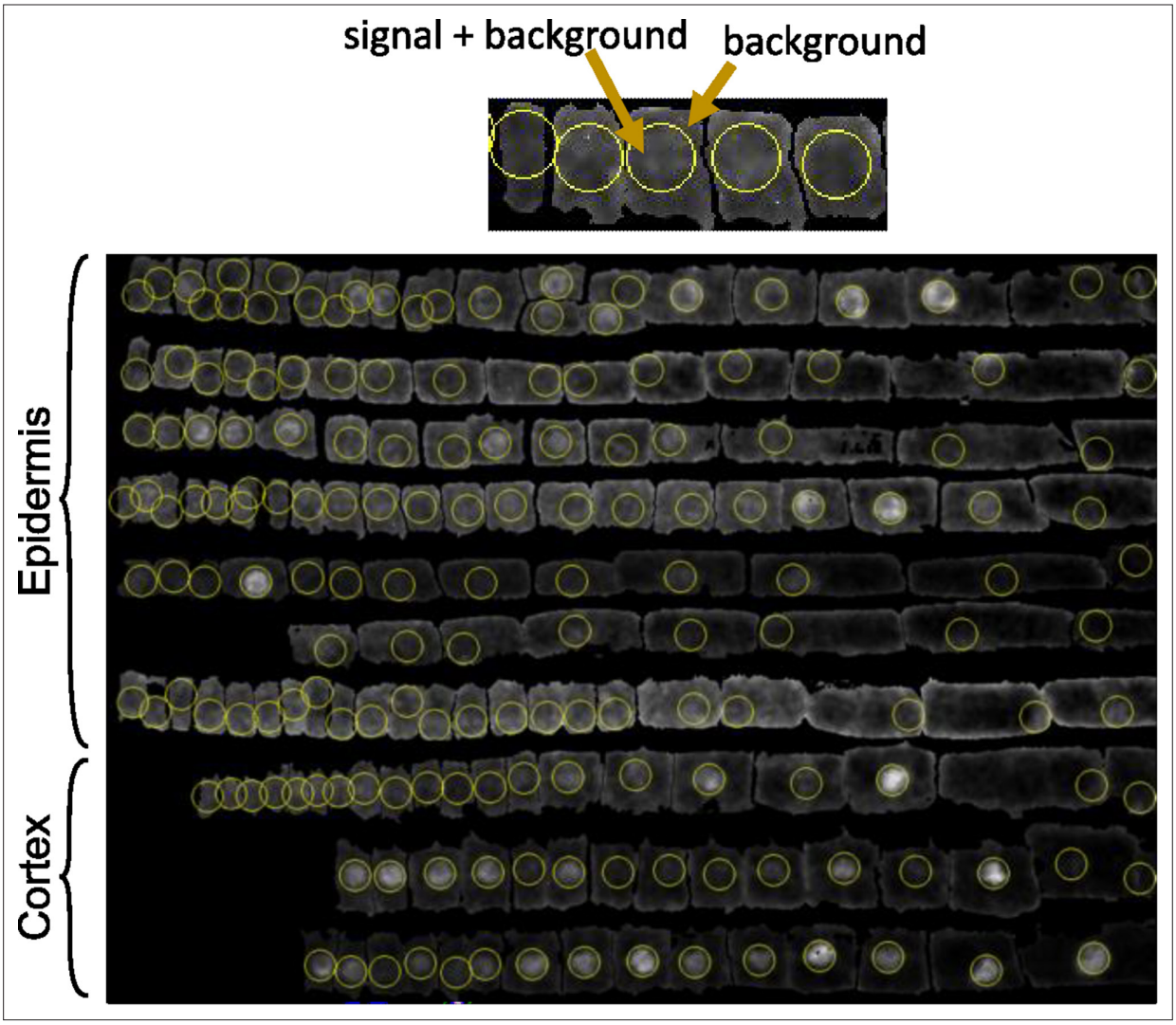


Figure 2—figure supplement 5. Segmentation and quantification method for FLC-Venus protein intensity per nucleus. Custom visualization and segmentation software was used to manually correct automatic segmentations and annotate cells according to cell file. Top: diagram explaining quantification of FLC-Venus intensity. In ON cells, the sphere contains the nucleus and therefore the FLC-Venus signal as well as the background signal. In OFF cells, the sphere will contain only background signal. The mean intensity outside the sphere was used as the background signal, and this was subtracted from the sphere intensity. Bottom: cell files from a single root separated according to tissue type demonstrating the result of the nucleus finding algorithm. Longitudinal section shown. In each cell, the sphere of constant radius that maximized fluorescent intensity inside the sphere was determined. For each cell, the position of the sphere is shown, corresponding to the nucleus in ON cells.

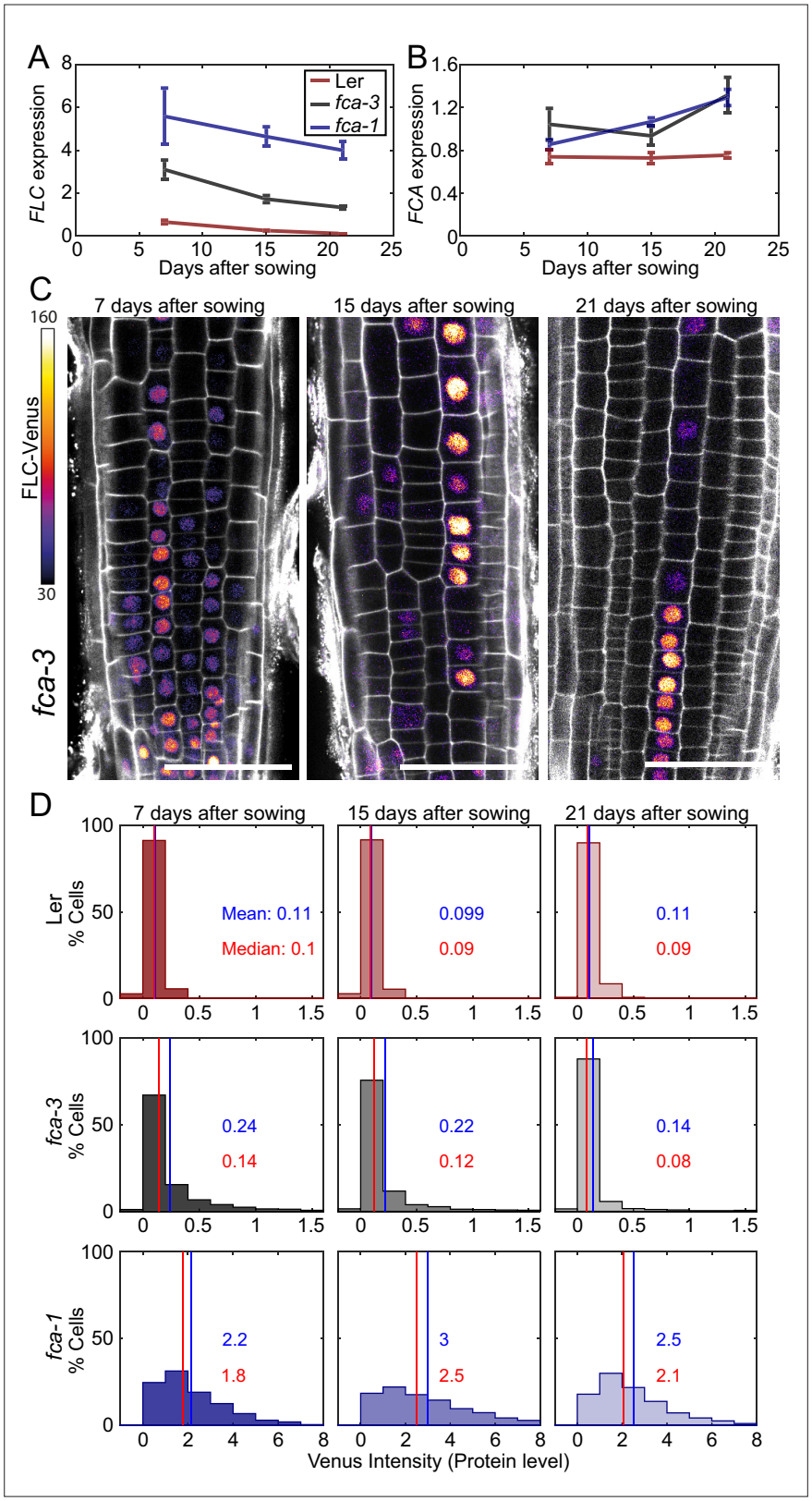


Figure 3. Experimental observation of gradual *FLC* silencing. (**A**) Timeseries of *FLC* expression in *fca* mutant alleles transformed with the FLC-Venus construct. Expression is measured by qPCR in whole seedlings relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC*). Error bars show SEM of n = 3 biological replicates. Statistical tests: samples were excluded as outliers based on Grubbs' test with alpha = 0.05. Linear

Figure 3 continued on next page

Figure 3 continued

regression on timeseries for each genotype. Slope for *fca-3* = −0.13, p-value=$2.0 \cdot 10^{-4}$; slope for *fca-1* = −0.11, p-value=0.12; slope for Ler = −0.038, p-value=$4.8 \cdot 10^{-6}$. (**B**) Timeseries of *FCA* expression, otherwise as in (**A**). Statistical tests: samples were excluded as outliers based on Grubbs' test with alpha = 0.05. Linear regression on timeseries for each genotype. Slope for *fca-3* = 0.021, p-value=0.22; slope for *fca-1* = 0.032, p-value=$3.8 \cdot 10^{-4}$; slope for Ler = $1.1 \cdot 10^{-3}$, p-value=0.84. (**C**) Representative images of *fca-3* roots by confocal microscopy. FLC-Venus intensity indicated by color map; gray shows the PI channel. Same settings were used for imaging and image presentation. Scale bar, 50 μm. (**D**) Histograms of FLC-Venus intensity per cell at each timepoint. The means and medians of the distributions are indicated in each panel. For Ler: 7 d, n = 1121 cells from 10 roots (three independent experiments); 15 d, n = 1311 cells from 9 roots (two independent experiments); 21 d, n = 1679 cells from 12 roots (three independent experiments). For *fca-3*: 7 d, n = 2875 cells from 24 roots (three independent experiments); 15 d, n = 3553 cells from 23 roots (three independent experiments); 21 d, n = 3663 cells from 21 roots (three independent experiments). For *fca-1*: 7 d, n = 1022 cells from 9 roots (three independent experiments); 15 d, n = 1770 cells from 12 roots (three independent experiments); 21 d, n = 2124 cells from 12 roots (three independent experiments). Statistical tests: linear regression on timeseries for each genotype. Slope for *fca-3* = −0.0077, p-value=$4.0 \cdot 10^{-46}$; slope for *fca-1* = 0.018, p-value=0.00064; slope for Ler = -0.00015, p-value=0.44.

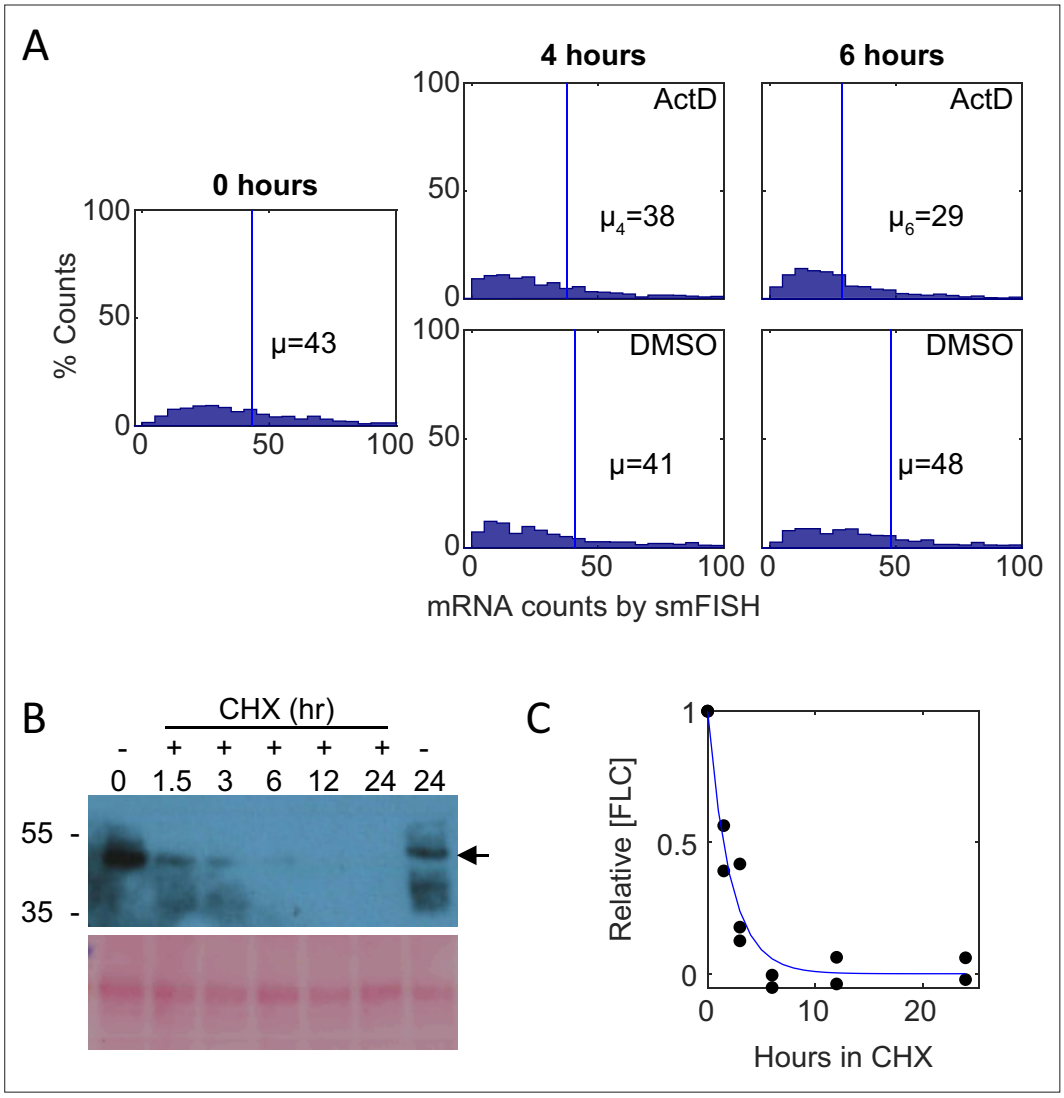


Figure 3—figure supplement 1. Degradation rate of *FLC* mRNA and protein levels. (**A**) Histogram of *FLC-Venus* mRNAs per cell measured by single-molecule fluorescence in situ hybridization (smFISH) after 4 hr or 6 hr of ActD treatment (top) or mock DMSO treatment (bottom). Intronic PP2A probes were used to confirm the effect of ActD on transcription. Mean mRNA number per cell, μ, is indicated in each panel. Half-life $t_{1/2}$ of 5 hr calculated as in ***Ietswaart et al., 2017*** using the formula $t_{1/2} = ln\left(2\right)/d$, with $d = ln\left(\mu_4/\mu_6\right)/2$, where d is the degradation rate and μ_4, μ_6 are the mean RNA counts after 4 and 6 hr, respectively. For 0 hr, n = 1554; for ActD 4 hr, n = 820; for ActD 6 hr, n = 784; for DMSO 4 hr, n = 759; for DMSO 6 hr, n = 776, in seedlings 7 d after sowing. Three independent experiments were performed and combined. (**B**) Analysis of FLC-Venus protein stability with cycloheximide (CHX) treatment. 7-day-old seedlings of *fca-1* plants carrying FLC-Venus reporter were subjected to liquid MS solution supplemented with 100 μM CHX for 0, 1.5, 3, 6, 12, or 24 hr. Blot shows representative replicate. Nontreatment control was also included (right-most well). Arrow indicates FLC-Venus. Ponceau staining was used as processing controls. (**C**) Quantification of the FLC-Venus from blot shown in (**B**) and additional replicates. Combined data from n = 3 biological replicates are shown with exponential decay profile fitted to the first three timepoints (half-life = 1.45 hr).

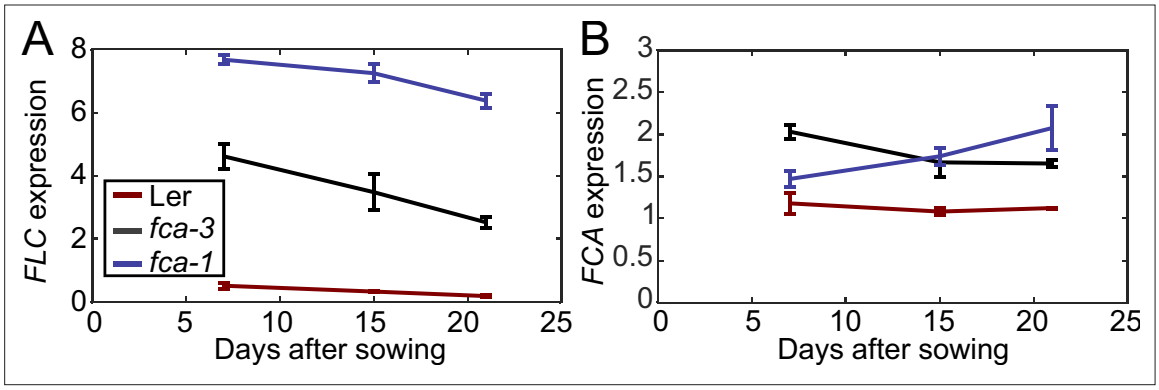


Figure 3—figure supplement 2. *FLC* and *FCA* expression in whole seedlings over time . (**A**) Timeseries of *FLC* expression in *fca* mutant alleles. Expression is measured by qPCR relative to the house-keeping gene index (geometric mean of *PP2A* and *UBC*). Error bars show SEM of n = 3 biological replicates. *FLC* 7d timepoint is repeated from ***Figure 1C***. Statistical tests: linear regression on timeseries for each genotype. Slope for *fca-3* = −0.15, p-value=0.0063; slope for *fca-1* = −0.091, p-value=0.0054; slope for Ler = −0.023, p-value=0.0030. (**B**) Timeseries of *FCA* expression, otherwise as in (**A**). Statistical tests: linear regression on timeseries for each genotype. Slope for *fca-3* = −0.028, p-value=0.043; slope for *fca-1* = 0.043, p-value=0.031; slope for Ler = −0.00445, p-value=0.56.

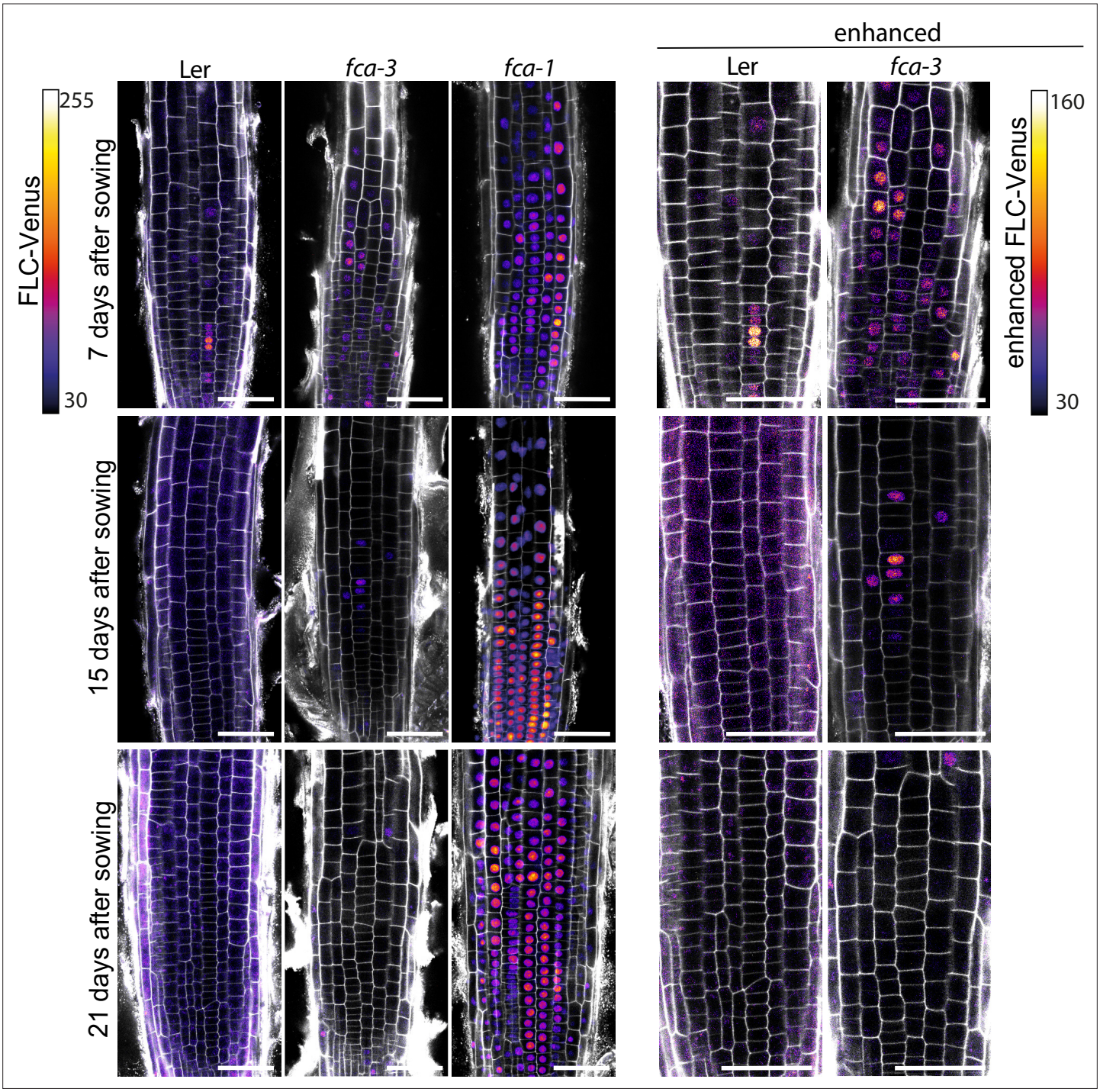


Figure 3—figure supplement 3. FLC-Venus time-course replicates in *fca* alleles. Additional representative images of FLC-Venus signal in the three genotypes (Ler, *fca-3*, *fca-1*) at the three timepoints. FLC-Venus intensity is indicated by color maps; gray shows the PI channel. Same settings were used for imaging and image presentation in left three columns. 'Enhanced' images (two columns on right) are a zoomed-in area of the same respective images, adjusted to enhance FLC-Venus signal to similar level as *fca-1* by changing brightness and contrast (please note different scale of color map). Scale bar, 50 μm.

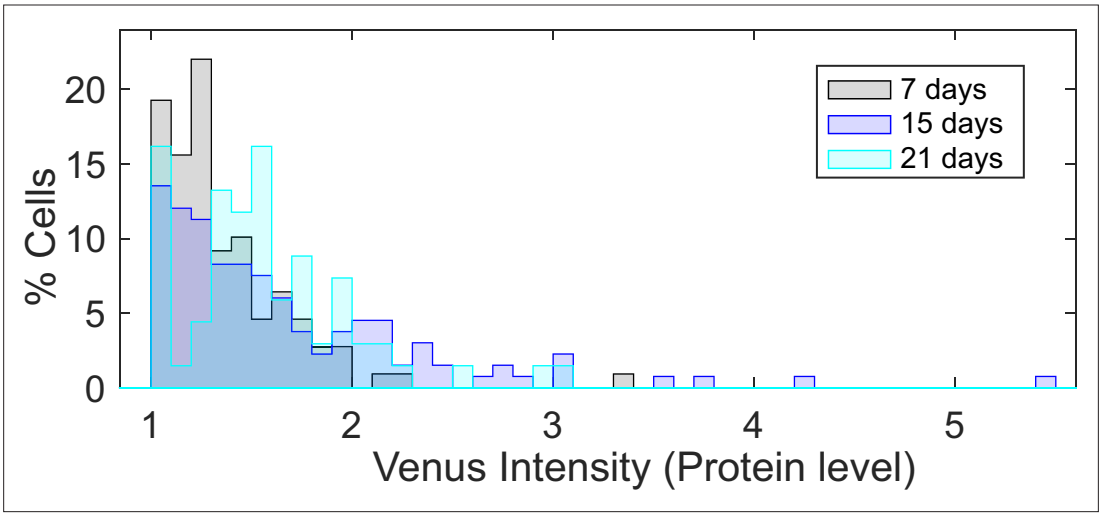


Figure 3—figure supplement 4. Experimental intensity of FLC-Venus in ON cells does not change over time in *fca-3*. Histogram of ON cell tail in *fca-3* at three timepoints normalized by the total number of cells in each. Only cells with FLC-Venus intensity above 1 are shown. For 7 d n = 109 cells from 17 roots, 15 d n = 133 cells from 14 roots, 21 d n = 68 cells from 6 roots. Data combined from three independent experiments.

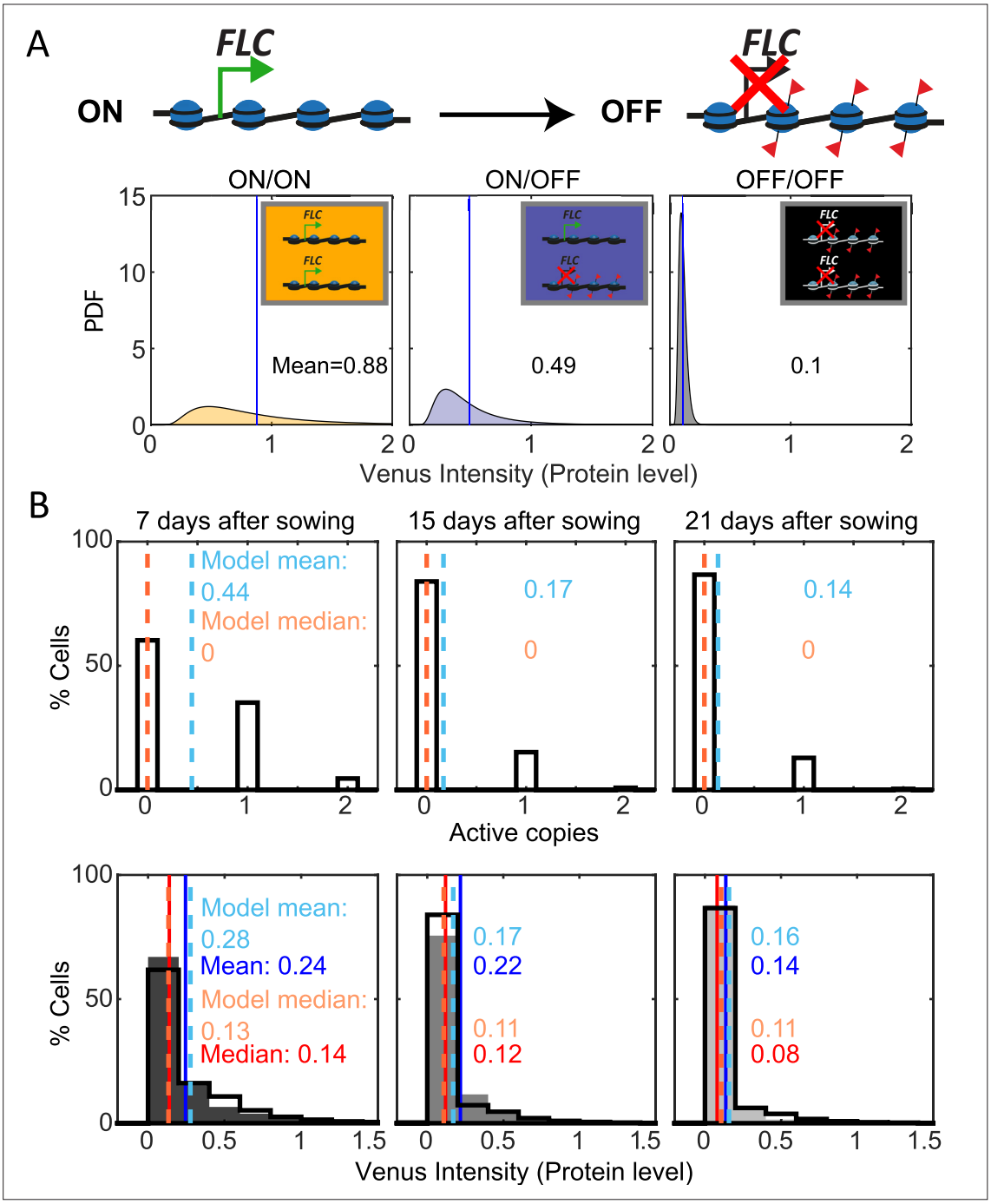


Figure 4. Mathematical model captures *FLC* regulation. (**A**) Diagram of mathematical model. Individual *FLC* gene copies can be ON or OFF, such that a cell can be in one of three states depending on the combination of ON and OFF gene copies within it (ON/ON, ON/OFF, OFF/OFF). Venus intensity (corresponding to amount of protein) within a cell was sampled from the distributions shown (described in 'Materials and methods' section), depending on the cell state. The means of the distributions are indicated in each panel. (**B**) Histograms of active *FLC* copies per cell (top) and Venus intensity per cell (bottom) at the indicated timepoints. Model histograms are plotted with black lines around empty bars and experimental data is shown as filled gray histograms with no outline. The means and medians of the distributions are indicated in each panel for the data (solid lines and matching color text) and the model (dashed lines and matching color text). Model simulated 1000 *fca-3* cell files and histograms shown exclude bottom four cells of each file as cells near the QC are also not included in the imaging.

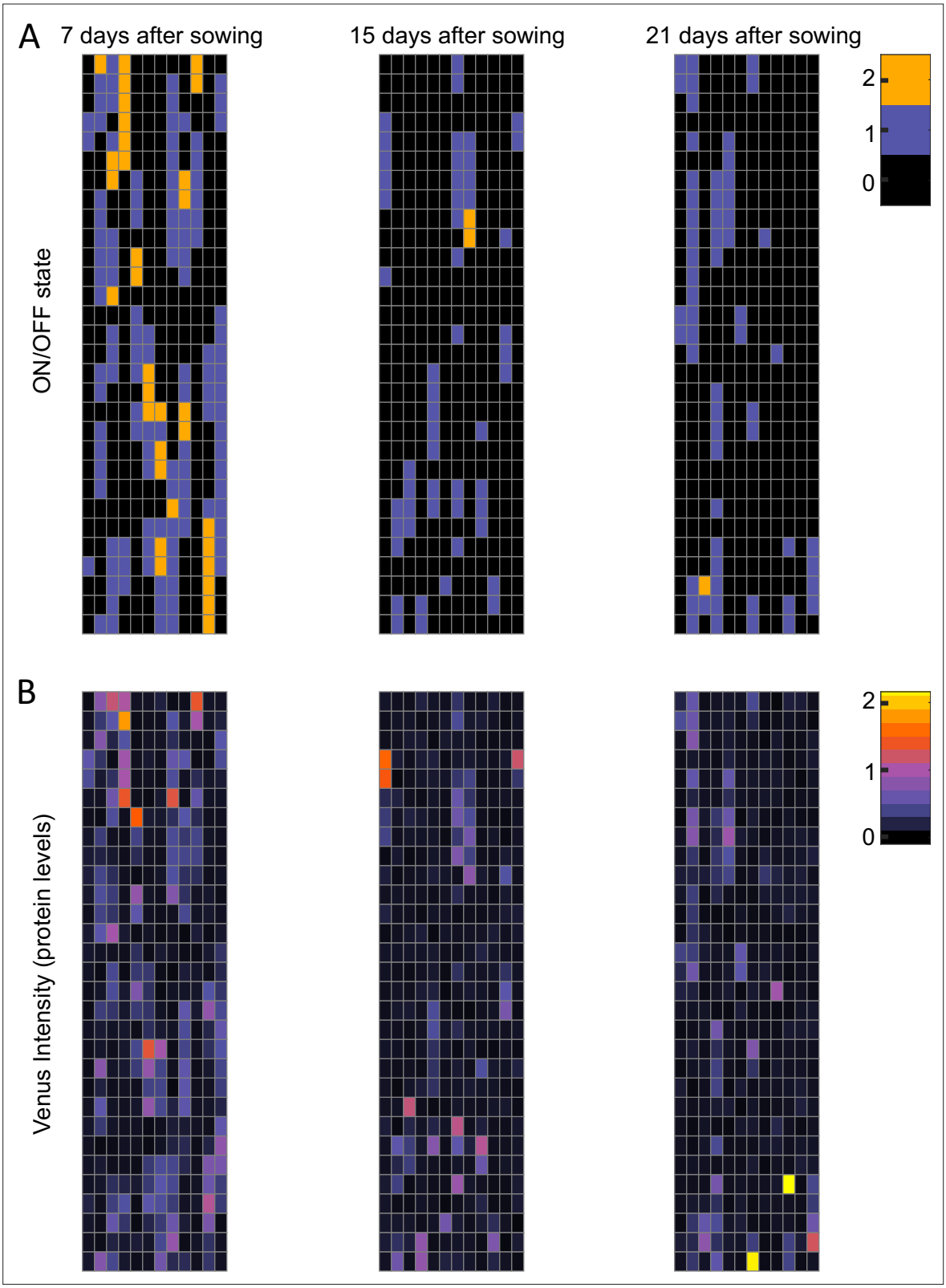


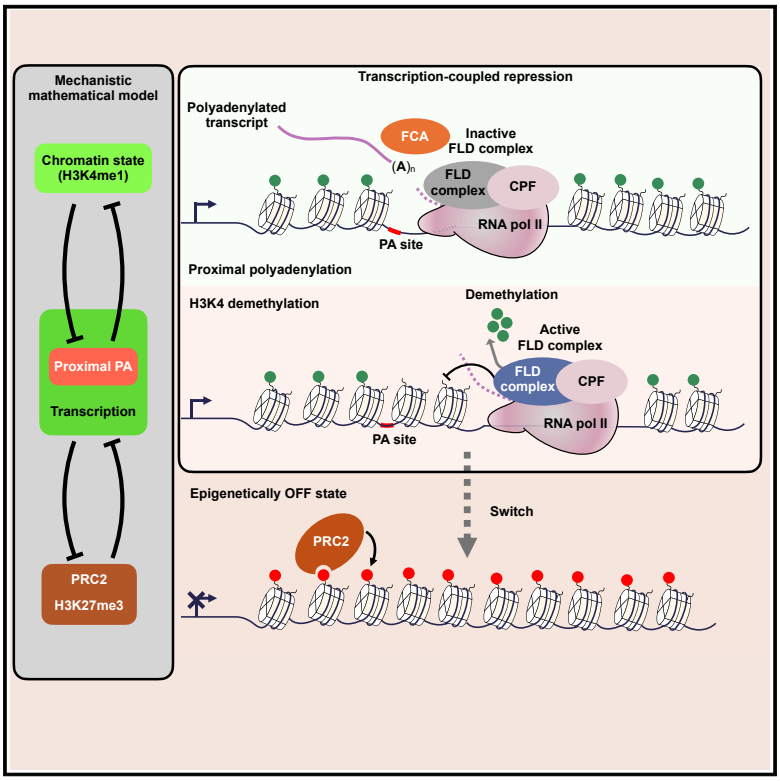
Figure 4—figure supplement 1. Modeling *FLC* regulation in clonal cell files in the root. Representative model output for *fca-3* root, showing 12 simulated cell files of length 30 cells each, at different timepoints, matching sampling times. (**A**) The state of the loci (both copies ON – orange, one copy ON – purple, both copies OFF – black). (**B**) FLC-Venus intensity (proxy for protein concentration) per cell, for the same cells.

Molecular Cell

Article

Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing

Graphical abstract



Highlights

- Transcription, co-transcriptional processing, and chromatin state are tightly linked
- Transcription-coupled repression links H3K4me1 removal to proximal polyadenylation
- Mathematical modeling reveals how this mechanistically links to silencing by PRC2
- More effective transcription-coupled repression enables faster epigenetic silencing

Authors

Govind Menon, Eduardo Mateo-Bonmati, Svenja Reeck, ..., Robert Ietswaart, Caroline Dean, Martin Howard

Correspondence

caroline.dean@jic.ac.uk (C.D.), martin.howard@jic.ac.uk (M.H.)

In brief

Mechanistic links between transcription, co-transcriptional processing, chromatin state, and epigenetic silencing remain poorly understood. Menon et al. used mathematical modeling and experiments to reveal a transcription-coupled repression mechanism involving feedback between chromatin state and proximal polyadenylation, which controls transcriptional antagonism to PRC2, thereby determining the establishment rate of PRC2 silencing at a locus.

Menon et al., 2024, Molecular Cell *84*, 1–17
June 20, 2024
https://doi.org/10.1016/j.molcel.2024.05.014

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing

Govind Menon,[1,6] Eduardo Mateo-Bonmati,[2,5,6] Svenja Reeck,[2] Robert Maple,[2] Zhe Wu,[3] Robert Ietswaart,[4] Caroline Dean,[2,*] and Martin Howard[1,7,*]

[1]Department of Computational and Systems Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK
[2]Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK
[3]Key Laboratory of Molecular Design for Plant Cell Factory of Guangdong Higher Education Institutes, Institute of Plant and Food Science, Department of Biology, School of Life Sciences, Southern University of Science and Technology, Shenzhen, China
[4]Harvard Medical School, Department of Genetics, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
[5]Present address: Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM), Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)/CSIC, 28223 Pozuelo de Alarcón, Madrid, Spain
[6]These authors contributed equally
[7]Lead contact
*Correspondence: caroline.dean@jic.ac.uk (C.D.), martin.howard@jic.ac.uk (M.H.)
https://doi.org/10.1016/j.molcel.2024.05.014

SUMMARY

The mechanisms and timescales controlling *de novo* establishment of chromatin-mediated transcriptional silencing by Polycomb repressive complex 2 (PRC2) are unclear. Here, we investigate PRC2 silencing at *Arabidopsis FLOWERING LOCUS C* (*FLC*), known to involve co-transcriptional RNA processing, histone demethylation activity, and PRC2 function, but so far not mechanistically connected. We develop and test a computational model describing proximal polyadenylation/termination mediated by the RNA-binding protein FCA that induces H3K4me1 removal by the histone demethylase FLD. H3K4me1 removal feeds back to reduce RNA polymerase II (RNA Pol II) processivity and thus enhance early termination, thereby repressing productive transcription. The model predicts that this transcription-coupled repression controls the level of transcriptional antagonism to PRC2 action. Thus, the effectiveness of this repression dictates the timescale for establishment of PRC2/H3K27me3 silencing. We experimentally validate these mechanistic model predictions, revealing that co-transcriptional processing sets the level of productive transcription at the locus, which then determines the rate of the ON-to-OFF switch to PRC2 silencing.

INTRODUCTION

Transcriptional regulation is often mediated by chromatin states that are key to developmental changes and environmental responses in many organisms. These states are characterized by histone modifications and *trans*-factors that recognize these modifications, which together influence the level of productive transcription. Chromatin states can be heritable across DNA replication and mitosis, with controlled switching of developmental loci between active and silenced states underpinning changes in gene expression programs driving differentiation.

Polycomb repressive complex 2 (PRC2)-mediated H3K27me3 is a system that generates such heritable and switchable chromatin states. Mechanisms enabling inheritance include read-write reinforcement of existing histone modifications and antagonism between transcription and silencing modifications.[1–3] Such mechanisms can generate bistability producing all or nothing (ON or OFF) transcriptional states.[1,4] Each gene copy can be stably maintained in either an ON or OFF state, even in dividing cells, and copies can be independently switched between states. Such switching has been achieved through experimental perturbations in various studies,[3] but there are few examples where the mechanisms are studied in natural developmental contexts.

One such example is PRC2-mediated regulation at the *Arabidopsis* floral repressor gene *FLOWERING LOCUS C* (*FLC*).[5] The winter cold-induced epigenetic silencing of *FLC* is a cell-autonomous, stochastic ON/OFF switching of individual copies into a Polycomb-silenced state (Figure 1A).[5–9] In this process, the silencing histone modification H3K27me3 is first nucleated (deposited at a specific intragenic site) during the cold and subsequently spread across the locus through a cell-cycle-dependent mechanism.[10]

In rapid-cycling *Arabidopsis* accessions *FLC* silencing is constitutive, occurring early in development. This silencing requires the RNA-binding protein FCA (encoded by *FLOWERING*

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

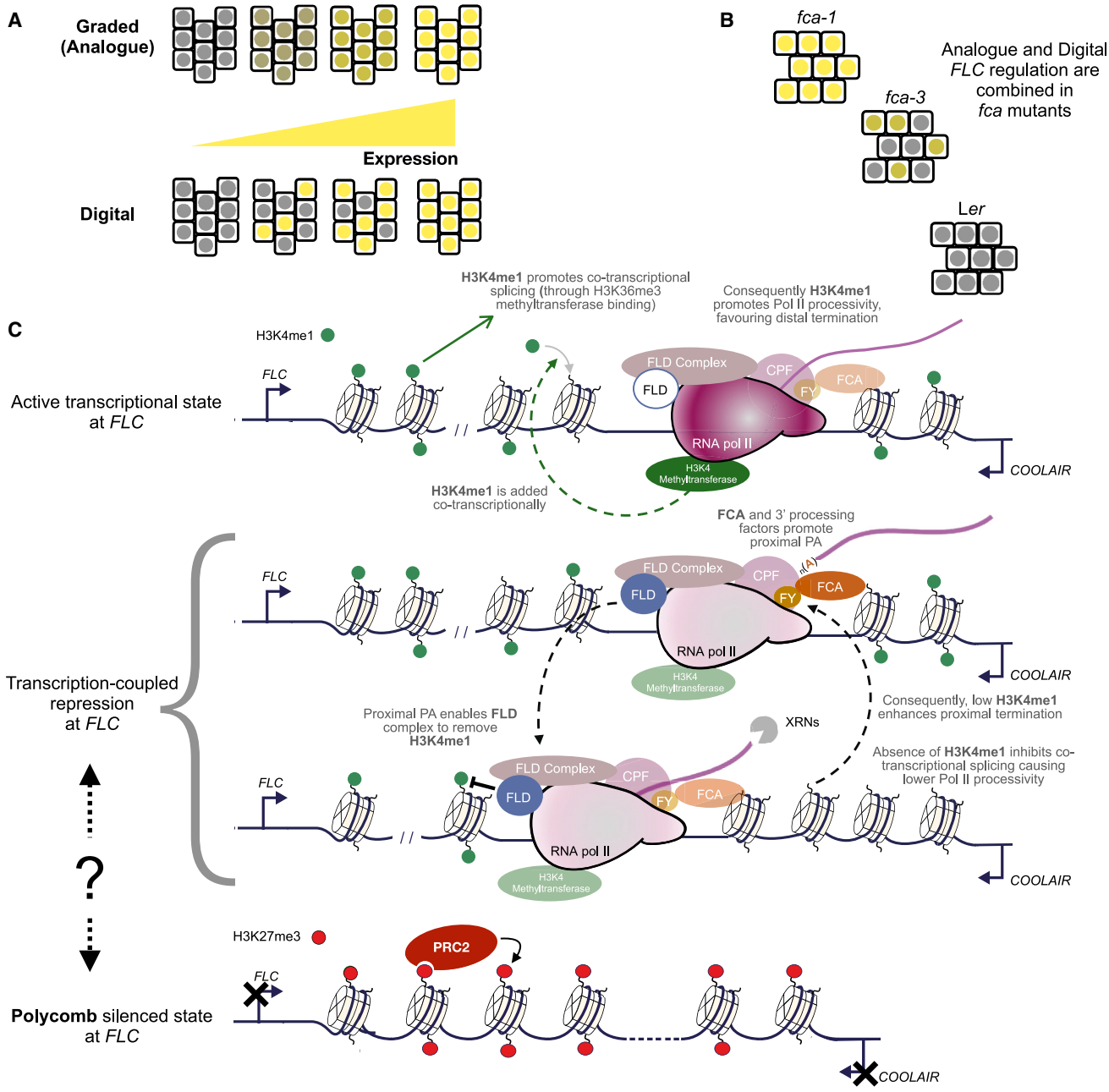


Figure 1. Transcription-coupled repression mechanism and modes of regulation at *FLC*

(A) Two possible modes of regulating population-level expression—analog and digital.

(B) Combined digital and analog regulation found in *fca* mutants.

(C) Transcription-coupled repression mechanism. Activating H3K4me1 is added during productive transcription (green dashed arrow). FCA promotes proximal termination, which enables removal of transcriptionally activating H3K4me1 across the locus by FLD. H3K4me1 removal inhibits co-transcriptional processing by preventing addition of H3K36me3, further promoting proximal termination. A feedback loop is formed, where proximal termination is self-reinforcing. Also shown is the Polycomb-silenced state at *FLC*—its link to transcription-coupled repression has been unclear. See Video S1.

CONTROL LOCUS A), which interacts with cleavage and polyadenylation specificity factor (CPF) machinery.[11] FCA promotes proximal termination of *FLC* sense transcription in the developing embryo,[12] and of *FLC* antisense transcription in seedlings.[13] *FLC* silencing also requires FLD (*FLOWERING LOCUS D*),[14] a plant homolog of the conserved mammalian histone K4 demethylase LSD1.[15] The role of FCA and other co-transcriptional regulators in this mode of *FLC* silencing, and FCA interactions with termination and 3′ processing factors, suggests a transcription-coupled mechanism that links to Polycomb silencing.

Indeed, the accompanying paper identifies a role for the CPF phosphatase module, a key regulator of RNA polymerase II (RNA Pol II) termination, in this mode of *FLC* silencing.[16] Hence, this mechanism could potentially reveal the link between transcription-coupled repression and Polycomb silencing.

Transcription-coupled mechanisms have been implicated in heterochromatin formation in many organisms through 3′ processing and non-coding RNA, for instance in *Schizosaccharomyces pombe* (*S. pombe*).[17] Clr4-mediated heterochromatin assembly by the RNAi-independent pathway involves non-canonical polyadenylation and 3′ processing factors.[18–20] Non-canonical termination site usage mediated by the conserved CPF and the RNA-binding protein Mmi1 drive Clr4-mediated heterochromatin assembly at meiotic and environmental response genes. Another RNA-binding protein, Seb1, promotes heterochromatin formation through termination of non-coding transcription.[21,22] For PRC2-mediated silencing, early termination has been proposed to reduce nascent RNA inhibition of PRC2 by limiting the physical length of the transcript.[23] Nevertheless, the exact mechanisms linking co-transcriptional processing to chromatin regulation are still poorly understood.

Exploiting an allelic series of *fca* mutations, we recently showed that the FCA-mediated mechanism can generate a graded regulation of *FLC* expression.[24] Thus, *FLC* can be regulated both in a graded manner (or *analog* regulation) or by an ON/OFF mode of regulation (or *digital* regulation) as in the cold-induced silencing (Figure 1A). The intermediate *fca-3* allele revealed that FCA-mediated analog regulation precedes an ON/OFF developmental switch at *FLC* (Figure 1B).[24] Thus, the study highlighted the switch from analog to digital regulation but did not mechanistically address *how* that switch occurred. In this study, we asked: how can co-transcriptional processing mediated by FCA and the linked function of the chromatin modifier FLD achieve the observed *analog* mode/graded manner of *FLC* regulation? Does the observed ON-to-OFF switch involve PRC2? If so, what is the mechanistic link between this analog transcription-coupled mode of regulation and a Polycomb ON/OFF *digital* mode of regulation? Specifically, how is an FCA/FLD-mediated repressed state (with enhanced early termination of transcription enabling graded regulation of the amount of full-length transcript produced) linked to a PRC2-mediated silenced state (OFF state with a low frequency of transcription initiation)?

Using mathematical modeling combined with molecular approaches, we dissect the interactions between co-transcriptional processing and chromatin regulation that underpin how transcription-coupled repression can lead to Polycomb-mediated silencing of *FLC*. Specifically, the model describes sense and antisense transcription, alternative transcriptional termination, the dynamics of activating (H3K4me1) and silencing (H3K27me) modifications, and the interactions between these components (Figure 1C and STAR Methods). Importantly, the complexity of these feedbacks could not be interpreted without modeling approaches. We first showed that the model can generate intermediate transcriptional outputs at non-Polycomb-silenced *FLC* copies (analog/graded regulation) while allowing Polycomb nucleation to occur slowly over time (ON/OFF switch). We used the model to qualitatively predict changes in activating H3K4me1 and silencing H3K27me3 modifications at *FLC* over time and experimentally validated these predictions by measuring changes in these modifications over a developmental time course, showing that the model can be quantitatively fit to capture these changes. Further validation was provided by measuring time course changes in *FLC* expression, where the model parameter fit obtained from histone modification changes also achieves a satisfactory quantitative match with no additional changes. The separate roles of the co-transcriptional repression and Polycomb silencing as predicted by the model were further validated by examining a mutant of a Polycomb methyltransferase. The model clearly elaborates how FCA/FLD-mediated transcription-coupled repression, induced by co-transcriptional processing steps as described in the accompanying paper,[16] can influence the timing of the digital switch to Polycomb silencing. The combination of molecular genetics and the theoretical modeling described here provides a quantitative description of how co-transcriptional processing can quantitatively set transcriptional states that determine rates of Polycomb silencing, a mechanism that is likely to be generally relevant in many organisms.

RESULTS

Molecular framework governing the *FLC* transcription-coupled chromatin repression pathway

The RNA-binding protein FCA promotes proximal polyadenylation/termination at many targets in the *Arabidopsis* genome.[25] FCA is known to directly interact *in vivo* with FY (flowering time control protein FY), a CPF component needed for silencing homologous to yeast Pfs2 and mammalian WDR33.[11] FCA also indirectly interacts with other cleavage and polyadenylation specificity factors,[26] as well as the 3′ processing factors CstF64 and CstF77.[13] FCA-mediated silencing involves CDKC2, a component of the transcription elongation factor P-TEFb,[27] as well as the core spliceosome component PRP8,[28] suggesting it is the process of transcription, co-transcriptional processing, and proximal termination, rather than just the transcripts themselves, that drive silencing. FCA is proposed to recognize any situation where RNA Pol II has stalled and transcription termination is required to clear a ‘‘tangle’’ in the chromatin.[29,30]

FCA-mediated *FLC* silencing requires the protein FLD, a homolog of the mammalian H3K4/K9 demethylase LSD1.[14] FLD is thought to demethylate H3K4me1 in *Arabidopsis*.[15,31] FLD associates *in vivo* with the SET domain protein SDG26, which interacts with FY, indicating a physical link between processing of antisense transcripts and chromatin remodeling.[32] FLD also associates *in vivo* with LD, a TFII-S domain containing protein and homolog of the CPF phosphatase subunit, PNUTS in human and Ref2 in yeast. This strong association of chromatin modifiers such as FLD with RNA Pol II regulators shown in the accompanying paper,[16] identifies a mechanism where FCA-mediated proximal cleavage and polyadenylation triggers FLD-mediated removal of H3K4me1. The *Arabidopsis* H3K36 methyltransferase SDG8 is known to recognize H3K4me1, so H3K4me1 removal may be expected to affect H3K36 methylation. *FLC* upregulation

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

in an *fca* mutant is rescued in an *sdg8 fca* double mutant,[32] indicating that H3K36me3 may be essential to set up an active transcriptional state at *FLC*. This is consistent with our previous chromatin immunoprecipitation (ChIP)-qPCR data at *FLC*, showing high levels of H3K4me1 and H3K36me3 at the locus in a high transcriptional state and low levels of both these modifications in a low transcriptional state.[32] Evidence from *Saccharomyces cerevisiae* shows that H3K36me3 is required for the recruitment of splicing factors and effective co-transcriptional splicing.[33] Based on this evidence, FLD-mediated removal of H3K4me1 may cause reduced H3K36 methylation, thus inhibiting co-transcriptional splicing, consequently reducing Pol II processivity, and promoting further early transcription termination. The FLD mechanism thus forms a feedback loop where early termination enhances itself through H3K4me1 removal (Figure 1C).

FCA promotes premature termination of both sense and antisense transcription at *FLC*

The *FLC* sense proximal transcripts are detectable in young embryos[12] but less so in vegetative tissue, likely due to rapid nascent RNA turnover. By contrast, the *FLC* antisense RNAs transcribed from the 3′ end of *FLC* (collectively called *COOLAIR*), which are alternatively polyadenylated (PA) at a proximal PA site within the *FLC* gene or a distal PA site in the vicinity of the *FLC* promoter,[13,34] are both readily detectable. The mechanism underlying the differential turnover of proximally PA transcripts is currently unknown, but differential fates of coding and non-coding transcripts are not uncommon.[35] To assess the relative roles of FCA-mediated premature termination of sense/antisense transcripts in silencing, we compared repression of a wild-type *FLC* transgene (*FLC-15*), with that of an *FLC* transgene where the antisense promoter has been replaced (*TEX1.0*) altering the type and quantity of antisense transcript production.[8,27,28] We performed the comparison in two genetic backgrounds: one with the endogenous FCA carrying a null mutation (*fca-9* mutant in *Col-0*), thus having no functional FCA, and the other with the same mutant of the endogenous FCA but carrying a transgenic 35S::FCA construct,[14] giving FCA overexpression. We made two observations (Figure 2A): (1) the wild-type *FLC* transgene is significantly repressed by FCA overexpression (as previously reported[14]); (2) this FCA-dependent repression of the wild-type *FLC* transgene is significantly attenuated ~10 fold in *TEX1.0*, but repression is not lost completely. Thus, the antisense transcription mechanism only partially accounts for *FLC* repression, a fact supported by a clustered regularly interspaced short palindromic repeats (CRISPR) deletion of the antisense promoter at the endogenous *FLC* locus, causing a small *FLC* upregulation.[36] To further explore the importance of FCA-mediated proximal termination, we analyzed the production of sense PA *FLC* transcripts by Quant-seq (Moll et al.[37]) in the *fca* allelic series.[24,38] Three genotypes give different levels of *FLC* expression at the whole-plant level: the wild-type L*er* (low expression), the *fca-1* mutant (high expression), and the *fca-3* mutant (intermediate expression). Proximally PA *FLC* transcripts were detected for all genotypes analyzed at sites close to those reported in embryos (Figure 2B).[12] These data support a model whereby FCA-mediated proximal termination of both sense and antisense transcription contributes to repression of *FLC* transcription.

Computational model of *FLC* chromatin silencing through alternative termination

The silenced state of *FLC* is associated with H3K27me3 coverage of the whole locus.[39] Genetic analysis also shows that the plant Polycomb methyltransferase CURLY LEAF (CLF) is essential for FCA-mediated silencing of *FLC*.[32] To address how these factors are linked, we constructed a computational model (Figure 3A, details in STAR Methods), using our prior knowledge of Polycomb dynamics known to generate digital ON/OFF memory states. We used this model to dissect the underlying mechanism of the transition, elucidating how digital silencing might be integrated into FCA-mediated 3′ processing.

FCA and FLD mediate a transcription-coupled repression mechanism that is sufficient to generate analog regulation

Our previous analysis of the *fca* allelic series highlighted that differences in overall expression between genotypes appeared to have two components: a digital ON/OFF mode where we observed clear differences in the number of *FLC* OFF cells, and an analog mode, with quantitatively different levels of *FLC* in the ON cells (Figure 1B). Intermediate expression was associated with a slower switch off over time into the digital OFF state. This demonstrates that FCA quantitatively modulates transcription and influences the timing of the switch into the digital OFF state.

We examined whether the model could capture quantitative variation in *FLC* expression (analog mode) in the absence of the Polycomb module (see model schematic in Figure 3A). To represent differences in FCA function between the three genotypes, we used a single parameter controlling the probability of FCA-mediated proximal termination of transcription (see STAR Methods). We simulated the model with different values for this parameter, representing different levels of functionality of the FCA protein. We set the parameter to its highest value for L*er*, an intermediate value for *fca-3*, and the lowest for *fca-1* (Figure 3B). With these parameter settings, the model, with just transcription-coupled repression (Figure 3A) could recapitulate the analog mode of regulation, with simulations showing the lowest frequency of productive transcription events for L*er*, the highest for *fca-1*, and intermediate for *fca-3* (Figure 3C). Consistent with the relationship between productive transcription and H3K4me1 in the model, the levels of this modification were predicted to be lowest in L*er*, intermediate in *fca-3*, and highest in *fca-1* (Figure 3C). The specific model parameter values used in Figure 3C are chosen to obtain a quantitative fit for the full model to experimentally measured histone modification levels, as we explain below.

Model predicts differences in the rate of Polycomb silencing establishment in the different genotypes

With the analog module constrained in this way, we simulated the full model, including the Polycomb module (Figure 3A). This model recapitulated the experimentally observed reduction in *FLC* transcriptional output over time (Figure 3D), as well as the digital switch to stable silencing (one-way switch) (Figure 3E) observed using FLC-Venus imaging in root meristematic cells.[24]

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

CellPress
OPEN ACCESS

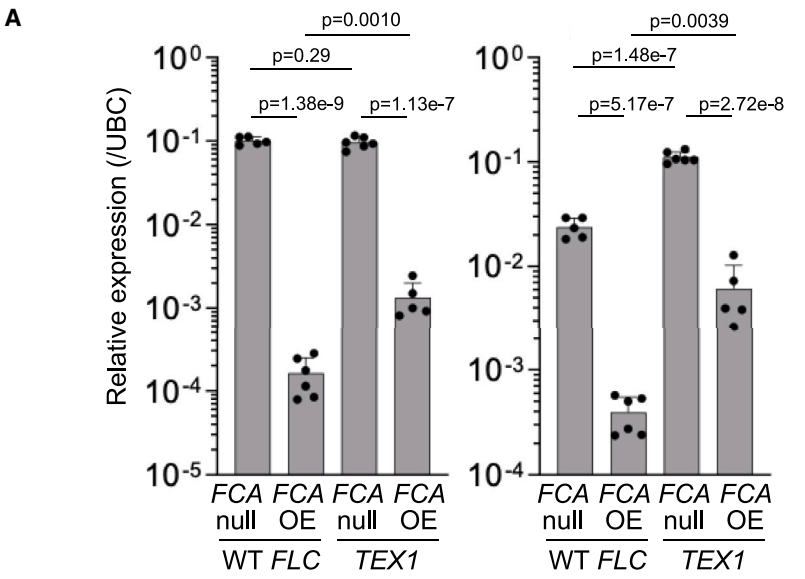
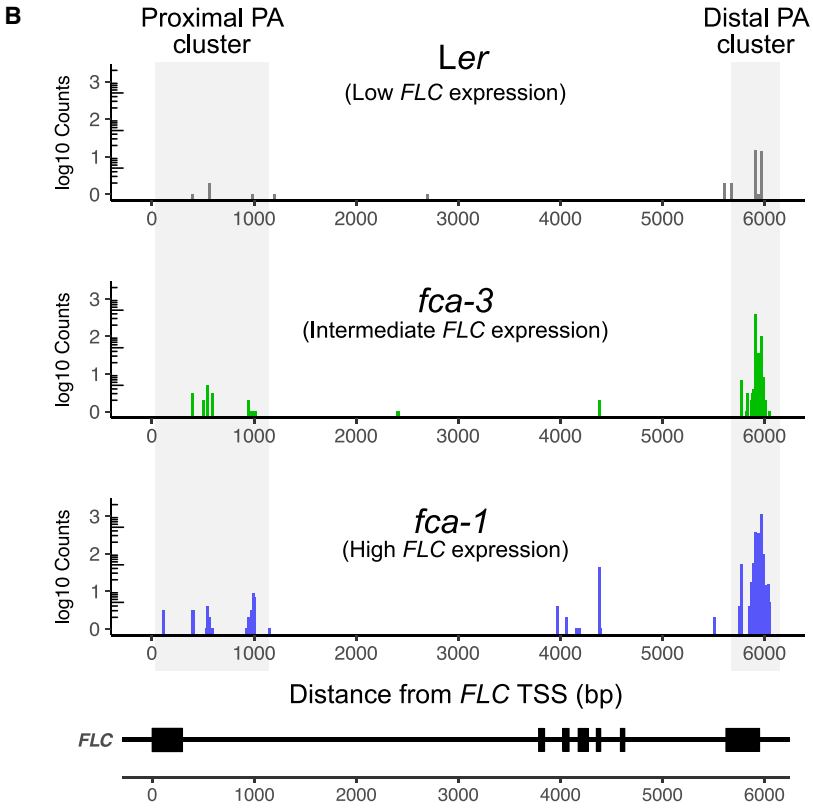


Figure 2. Transcription-coupled repression mechanism works through sense and antisense transcription

(A) qPCR data showing expression of transgenic wild-type *FLC* and antisense disrupted *FLC* constructs (*FLC-15* and *TEX1* respectively) with and without FCA overexpression (*35S::FCA* indicated by FCA OE) in 14-day-old seedlings. All genotypes are in *fca-9* (*FCA* null mutant) background. Error bars: mean ± SEM. *p* values shown for comparison of means using one-tailed Student's t test.

(B) Quant-seq data for 7-day-old seedlings quantifying polyadenylated transcripts at *FLC* in parental Landsberg (*Ler*) genotype and *fca* mutants. Individual replicates in Data S3.



Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

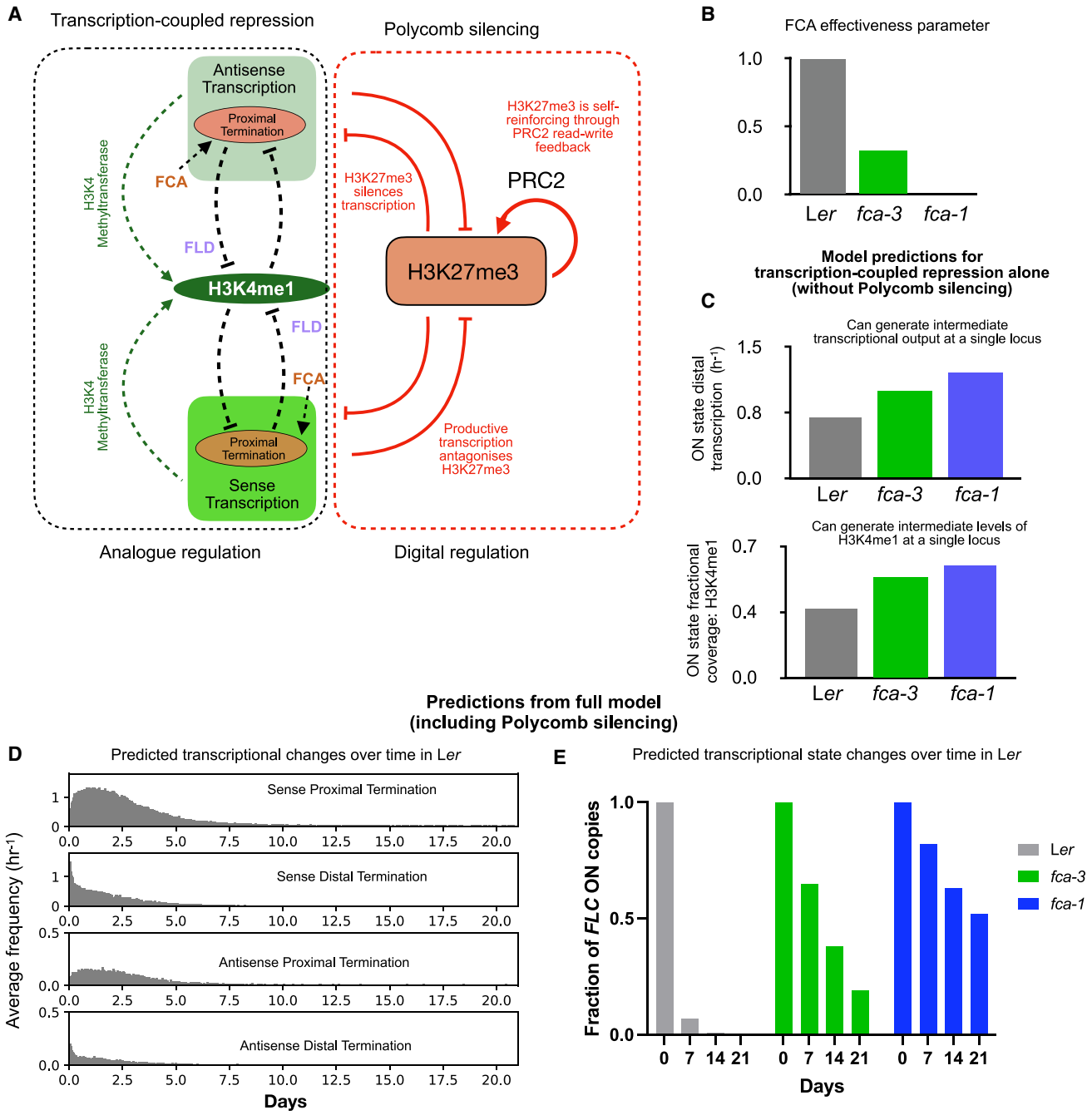


Figure 3. Mathematical model incorporating transcription-coupled repression mechanism and PRC2 silencing

(A) Schematic: model includes sense and antisense transcription, co-transcriptional addition of H3K4me1 (green dashed arrows), FCA-mediated proximal termination, FLD-mediated removal of H3K4me1, and promotion of proximal termination by low H3K4me1 coverage. (Black dashed lines indicate feedback). Model also includes H3K27 methylation, PRC2-mediated read-write maintenance of these modifications, and mutual antagonism between H3K27me3 and transcription.

(B) FCA effectiveness parameter values for the three genotypes.

(C) Top panel: model-predicted frequency of distal transcription events (steady-state average over 1,000 simulated loci) in an ON state for model of transcription-coupled repression mechanism shown in (A) but without Polycomb silencing module. Each simulated locus was started in an active state with full H3K4me1 coverage and allowed to equilibrate over a 5 cell cycle duration. Averaging performed over duration of following 10 cell cycles. Bottom panel: model-predicted fractional coverage of H3K4me1 in ON state from above simulations.

(D) Predicted transcriptional changes (quantified by frequency of transcription termination events) over time from combined model (A). Frequencies of different termination events averaged over 1,000 simulated copies of *FLC* locus, all started in an active state with full H3K4me1 coverage.

(E) Predicted changes in fraction of ON copies in *fca* mutants, computed from simulations of 1,000 copies over 30 cell cycles, started in a full H3K4me1 covered ON state. Fractions are shown at starting time point (0) and at the end of 7, 14, and 21 days from start. See STAR Methods for definition of ON and OFF states.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

For L*er*, where productive transcription is lowest, the model predicted a rapid switch to the digital OFF/Polycomb-silenced state (Figures 3E and S4Bi). For *fca-3*, the model predicted a slow switch to the digital OFF state, while even slower switching is predicted for *fca-1* (Figures 3E and S4Bi). In all cases, the switching was predicted to be one-way, with the Polycomb-silenced state being highly stable.

Thus, the combined model of transcription-coupled repression and Polycomb silencing predicts that the level of productive transcriptional activity set by FCA/co-transcriptional silencing/FLD (analog module) determines the level of transcriptional antagonism to H3K27me3. The level of transcriptional antagonism in turn dictates the rate of establishment of the Polycomb-silenced state.

Model predictions are consistent with observed chromatin changes over time

We next used the combined model to predict average levels of H3K4me1 and H3K27me3 at *FLC* in the three genotypes. Starting from an active chromatin state with full H3K4me1 coverage, the model predicted a progressive reduction in H3K4me1 over time in all genotypes, with a low steady-state level quickly attained in *Ler* and a slower reduction in *fca-3* and *fca-1*, with *fca-1* being the slowest (see red bars in Figure 4D). In all genotypes, this reduction was combined with an accumulation of H3K27me3 as more loci switch to the Polycomb-silenced state (see red bars in Figure 4C). At each time point (corresponding to our experimental time points—see below), the model predicted differences in average H3K4me1 and H3K27me3 levels at *FLC* between the three genotypes. L*er* was predicted to have the lowest H3K4me1 and highest H3K27me3, *fca-3* to have intermediate levels of both modifications and *fca-1* having the highest H3K4me1 and lowest H3K27me3.

To validate these predictions, we performed a time course measurement of silencing H3K27me3 and the activating modifications H3K4me1 and H3K36me3 over the *FLC* locus by ChIP followed by quantitative PCR (ChIP-qPCR), examining their levels in whole seedlings at 7, 14, and 21 days after sowing. For each genotype and time point, we use the Polycomb-silenced *STM* locus as the positive control for H3K27me3 and the actively expressed *ACT7* locus as the positive control for H3K4me1 (conversely, *STM* and ACT7 function as negative controls for the activating modifications and H3K27me3, respectively). See Figures S4Cii–S4Civ) for a comparison of the levels measured at *FLC* and these control loci at the 7-day time point. Consistent with model predictions, at all time points we see that *fca-3* shows intermediate levels of H3K4me1 and H3K27me3, as compared with *fca-1* and Ler, with primers across the *FLC* locus showing a consistent trend (Figures 4A and 4B). This trend is reinforced by a statistical comparison between the three genotypes of the averaged H3K27me3 and H3K4me1 levels across the locus (Figures S4Ai and S4Aii).

Qualitatively consistent with model predictions, as more loci switch to the digital OFF state, in *fca-3* and *fca-1*, measured H3K4me1 levels fall over time (statistically significant for both *fca-3* and *fca-1* between 14 and 21 days, Figure S4Aiv). On the other hand, H3K27me3 levels rise (the increase is statistically significant in *fca-3* between 14 and 21 days, though narrowly not for *fca-1*: $p = 0.048$, sig. threshold = 0.0253 (Holm-Sidak corrected), Figure S4Aiii). In Ler, H3K4me1 shows essentially no change over the experimental time points, while the H3K27me3 also does not show a statistically significant change on average across the locus (Figures S4Aiii and S4Aiv). The lack of significant change in measured H3K27me3 and H3K4me1 in L*er* is consistent with steady-state levels having been reached by the 7-day time point. The model predicts faster dynamics of H3K4me1 removal (through more frequent proximal termination) and faster establishment of H3K27me3 in L*er*, so a faster approach to steady state is consistent. Activating H3K36me3 also shows a similar pattern, with *fca-3* exhibiting intermediate levels and overall levels decreasing over time (Figure S4Biii). Thus, measured histone modification changes in all three genotypes are qualitatively consistent with the model predictions.

In addition to qualitative consistency between model predictions and experiment, we examined whether predictions quantitatively match the data by adjusting specific parameters. Fitted parameters are indicated in Tables S1–S4. All other parameters were held fixed. A direct quantitative comparison between the model predictions and the ChIP-qPCR data is difficult because the simplest model would assume that the switching dynamics happen at all loci, while the ChIP-qPCR analysis uses whole-seedling tissue, including tissue where the switching may be inhibited by other means. Therefore, when comparing the model predictions to the ChIP-seq data, we include a subpopulation of active *FLC* copies with full H3K4me1 coverage and no H3K27me3 coverage to compute a corrected model-predicted level for both modifications (see STAR Methods for details). Model-predicted levels are indicated in Figures 4C and 4D (red bars), showing overall consistency with the data. Hence, the combined model of transcription-coupled repression, mediated by proximal termination of transcription and Polycomb silencing, successfully predicts both the relative levels of the modifications between genotypes and how these change over time.

Model predicts relative differences in proximal termination between genotypes

To validate the model-predicted *FCA*-dependent differences in proximal termination, we compared these predictions to our Quant-seq data (data shown in Figure 2B). We compared the predicted ratio of proximal to distal termination events to the measured PA reads. To compute this ratio for the Quant-seq data, we defined a proximal termination cluster covering a range of termination sites close to the 5′ end of the gene.[12] A quantitative fit of these ratios is not attempted here, as we do not know turnover rates of the different transcripts. However, predicted trends in the differences in the ratio of proximal to distal polyadenylation of sense transcripts between the *FCA* mutants and the parental genotype qualitatively matched the trend for measured ratios of *FLC* sense transcription in our Quant-seq dataset (Figures 4E, 4F, and S4Ci; Data S3).

Disrupting a Polycomb methyltransferase compromises the one-way switch to a Polycomb OFF state but does not disrupt FCA-mediated regulation

The model predicts that the transcription-coupled repression mechanism is sufficient to generate the analog repression and

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

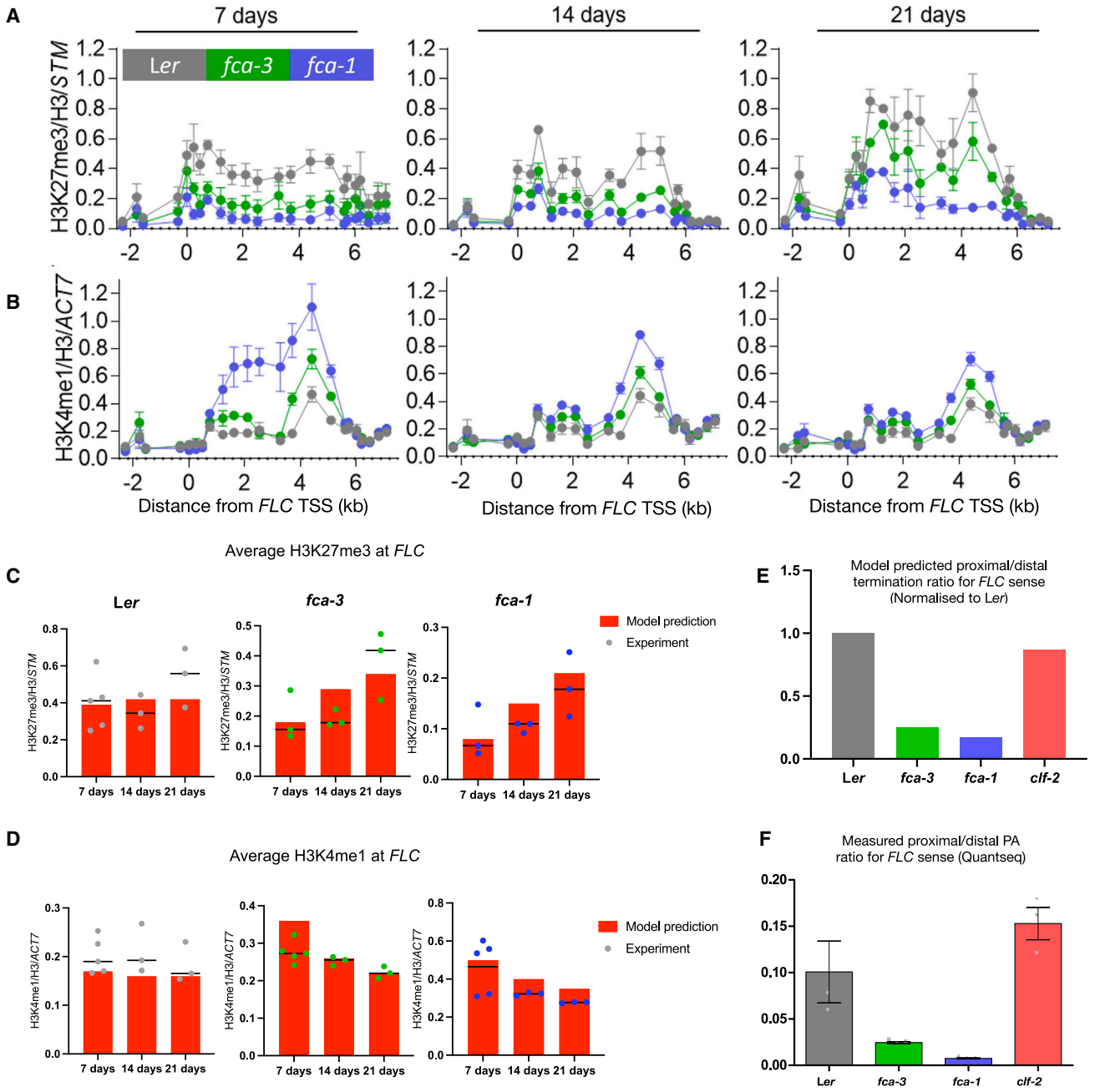


Figure 4. Experimental validation of model predictions

(A and B) ChIP-qPCR time course measurements of H3K27me3 (A) and H3K4me1 (B) across the *FLC* locus in parental L*er* genotype and *fca* mutants. H3K27me3 levels are shown normalized to H3 and to control locus *STM*. H3K4me1 levels shown normalized to H3 and to control locus *ACT7*. Each dot represents an amplicon. Error bars in (A) represent mean ± SEM. (n = 5 biological replicates only for L*er* at 7 days and n = 3 biological replicates in all other cases). Error bars in (B) represent mean ± SEM. (n = 5 biological replicates for L*er*, *fca-3*, and *fca-1* at 7 days and n = 3 biological replicates in all other cases).

(C and D) Vertical red bars indicate model-predicted level. All model-predicted histone modification levels averaged over the locus and averaged over 1,000 simulated copies started in an active state with full H3K4me1 coverage. Horizontal black bars indicate median of biological replicates. (C) Comparison of model-predicted H3K27me3 levels and changes over time in different genotypes with measured levels from (A), averaged across 15 primers covering the *FLC* locus. (D) Comparison of model-predicted H3K4me1 levels with experimental data shown in (B).

(E and F) Model-predicted (E) and measured (Quant-seq) (F) ratios of proximal to distal termination for *FLC* sense transcription in parental L*er* genotype, *fca* mutants, and *clf* mutant. Measured ratios computed using Quant-seq data for 7-day-old seedlings (Figures 2B and S4Ci). Error bars in (F) represent mean ± SEM. (n = 3 biological replicates). Genomic coordinates for proximal and distal termination clusters provided in Data S1. Read counts provided in Data S2. Model-predicted ratios normalized to the predicted ratio for L*er*. Model-predicted ratios computed at 7-day time point (averaged over 6 h preceding this time point) for all genotypes (average over 1,000 simulated copies started in an active state with full H3K4me1 coverage).

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

CellPress
OPEN ACCESS

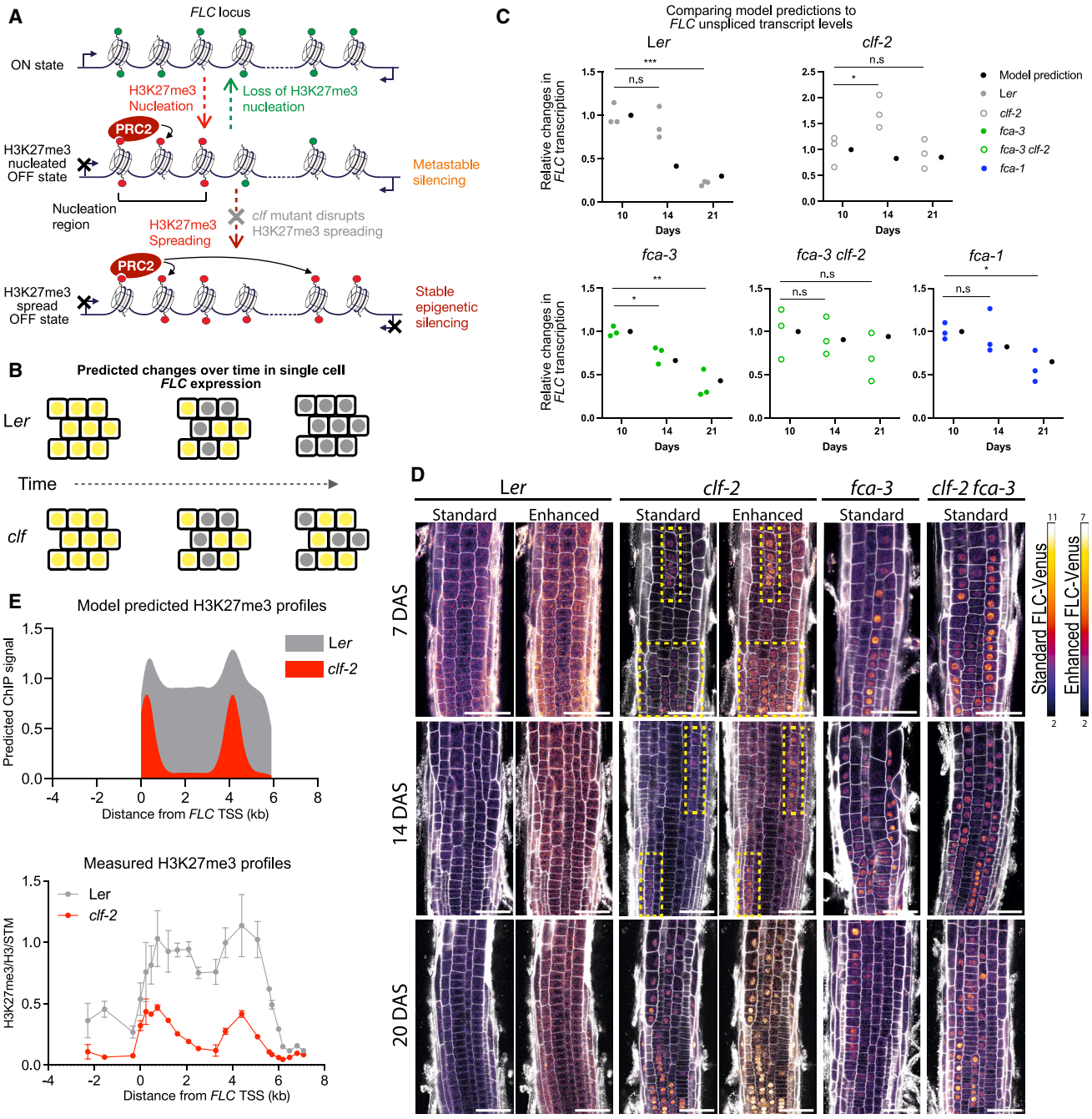


Figure 5. Experimental validation of model predictions

(A) Schematic showing rationale for examining *clf* mutant. Transition from PRC2 nucleated state to spread state is prevented in *clf* mutant. The lack of stability of a nucleated state allows switching back to a digital ON state. Green circles: active modifications (H3K4me1); red circles: repressive modifications (H3K27me3).

(B) Reduction over time in number of ON cells observed in parental *Ler* genotype is expected to be disrupted in *clf* mutant.

(C) Comparison of model-predicted changes in *FLC* transcription, with whole-seedling *FLC* unspliced transcript levels measured by qPCR in *fca* genotypes (*Ler*, *fca-3*, *fca-1*) and *clf* mutants (*clf-2* and *fca-3 clf-2* double mutant). Data normalized to *UBQ10*. Comparisons used one-tailed Student's t test and corrected for multiple comparisons using the Holm-Sidak approach (*, **, and *** indicate significance thresholds of 0.05, 0.01, and 0.001, respectively). Model predictions correspond to predicted frequency of *FLC* sense distal transcription events at corresponding time points, averaged over 1,000 simulated loci. Simulations for *Ler*, *fca-3*, *fca-1* identical to those used to predict changes in histone modification levels (Figure 4). Simulations for *clf* mutants (*clf-2*, *fca-3 clf-2*) use changed parameter values for PRC2-mediated H3K27 methylation (STAR Methods and Table S2). For comparison of relative changes, both model predictions and data normalized to level at 10 days.

(legend continued on next page)

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

that this repression controls the antagonism of transcription to Polycomb, thus enabling the switch to Polycomb digital silencing. We further tested these predictions using a Polycomb mutant—a loss-of-function mutant of *CLF*. This methyltransferase is known to be required for spreading of H3K27me3 from Polycomb nucleation peaks at many *Arabidopsis* targets,[40] and specifically at *FLC*, both in developmental[40] and cold-induced silencing contexts.[10] This mutant is not expected to completely disrupt the ON/OFF silencing mode, as H3K27me3 nucleation alone, which still occurs in *clf*, gives a metastable Polycomb-silenced state.[10] However, only after the H3K27me3 spreads across the whole locus is the Polycomb-silenced state fully stable in rapidly dividing cells.[10] Thus, the maintenance of the OFF state is predicted to be compromised in *clf* (Figures 5A and 5B). Therefore, in a *clf* mutant background, the model makes two predictions: (1) reduction over time in overall *FLC* expression seen in L*er, fca-3*, and *fca-1*, i.e., an overall reduction in the number of actively transcribed *FLC* copies—due to a fully stable Polycomb digital switch—will be missing in a *clf* mutant background due to leakage out of the metastable Polycomb OFF state (Figures 5A and 5B); (2) analog differences generated by *fca* mutants will persist since the FCA-mediated repression can function independently of H3K27me3.

To validate these predictions, we measured *FLC* expression (measuring spliced and unspliced *FLC* transcripts by qPCR) at the whole-plant level, over the same developmental time course in five different genotypes (Figure 5C shows unspliced *FLC*; Figure S5A shows spliced *FLC*): in the parental L*er*, the *CLF* mutant *clf-2*,[41] *fca-3*, *fca-1*, and the *fca-3 clf-2* double mutant. Consistent with previous data[24] and observed chromatin changes (Figures 4A–4D), the FCA genotypes *Ler, fca-3*, and *fca-1* exhibited statistically significant changes in *FLC* expression over the time course (Figures 5C and S5A). As predicted by the model, this slow reduction in expression is compromised in the *clf-2* and *fca-3 clf-2* mutants—neither exhibits a statistically significant *reduction* in *FLC* expression. *clf-2* exhibits an increase at the 14-day time point (statistically significant only for unspliced *FLC*) that the model cannot capture. Consistent with predictions for the *clf* mutants, the analog differences associated with FCA are maintained, i.e., the expression level in the *clf-2* mutant is clearly lower than in *clf-2 fca-3* at all time points (Figure S5'B), indicating that FCA-mediated transcription-coupled repression is the main factor underlying analog differences.

We quantitatively compared these data to model-predicted *FLC* sense transcriptional output (i.e., frequency of distally terminated sense transcription events) over time. Model predictions are indicated by filled black circles in Figures 5C and S5A. To do this, we simulated the full model for L*er, fca-3*, and *fca-1*. We did not attempt to fit to absolute levels of *FLC* transcripts since the measured levels also depend on RNA stability, which we do not measure experimentally or include in our model. Therefore, we compared changes relative to the measured levels at the first (10-day) time point. The relative changes predicted by the model are overall quantitatively consistent with the measured changes in both unspliced (Figure 5C) and spliced *FLC* (Figure S5A) (note that the model predictions are identical for spliced and unspliced—the model does not distinguish between the two). To capture the *clf* mutants (*clf-2* and *fca-3 clf-2*), the model was simulated again for L*er* and *fca-3* with only the PRC2 read-write feedback parameters changed to weaken spreading (Table S2).

In some cases, between 10 and 14 days, the model-predicted changes are larger than the measured changes, notably, for unspliced *FLC* in L*er* (Figure 5C) and spliced *FLC* in *fca-3* and *fca-1* (Figure S5A). However, since the model was only parametrized to fit the ChIP time course data (Figure 4), the overall satisfactory quantitative fit observed for *FLC* expression with the same parameter values provides further support for the validity of the model.

In addition to *FLC* expression at the whole-plant level, we conducted time course imaging of Venus tagged FLC in the roots in the *clf* mutants (Figure 5D), following the same approach previously used to analyze the *fca* alleles.[24] As predicted, *clf-2* showed clear differences to the L*er* genotype, with ON cells being visible at all three time points, consistent with *FLC* switching both ON to OFF and OFF to ON as described above, and the observed lack of reduction in *FLC* expression at the whole-plant level (Figure 5C). Both *clf-2* and *fca-3 clf-2* exhibited large variability between roots, some roots showing essentially no *FLC* ON cells (Data S4). This is possibly due to the widespread misregulation of growth and developmental genes resulting from the compromised Polycomb silencing in this mutant. Nevertheless, consistent with the model predictions, the slow Polycomb digital switch to OFF is disrupted in *fca-3 clf-2*: when comparing roots exhibiting ON cells in the double mutant to *fca-3* roots, the double mutant overall shows more ON cells at later time points.

The model predicts a high ratio of proximal to distal termination in the *clf* mutant similar to the L*er* genotype (Figure 4E), even though overall transcription is lower in L*er*. This is because the FCA/FLD-mediated transcription-coupled mechanism is not disrupted in the *clf* mutant model, so the promotion of proximal termination by this mechanism and the resulting feedback remain intact. Consistent with model predictions, the ratio of proximal to distal termination of the *FLC* sense transcript in the *clf* mutant, using our strand-specific Quant-seq data (Figure S4Ci), exhibits a similar ratio to the L*er* genotype (Figure 4F).

As further model validation, we compared predicted H3K27me3 levels at the locus in a *clf* mutant to experimentally

(D) FLC-Venus time course imaging in different genetic backgrounds. Representative confocal images of FLC-Venus signal at three time points in epidermis of the root meristem of parental genotype Ler, single mutants *clf-2* and *fca-3*, and double mutant *fca-3 clf-2*. Gray shows cell-wall-dye propidium iodide; FLC-Venus intensity indicated by color maps for standard and enhanced images, respectively. Same imaging settings used for all time points. Same image presentation used for standard and enhanced images. Brightness and contrast of FLC-Venus signal were increased to generate enhanced images (note different scale on color maps). Yellow boxes in *clf-2* highlight ON cell files at 7 and 14 DAS. Scale bars, 50 μm. DAS, days after sowing.

(E) Top panel: model-predicted profiles of H3K27me3 coverage (see STAR Methods). Model-predicted profile is steady-state average over 100 simulated loci, started with full H3K4me1 and equilibrated over 10 cell cycles. Bottom panel: H3K27me3 profiles measured by ChIP-qPCR in 14-day-old seedlings. Error bars: mean ± SEM. (n = 3 biological replicates). Each dot represents an amplicon.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

measured levels. The lack of stability of the nucleated state means that the model predicts an H3K27me3 peak, but at significantly lower levels relative to a spread state. The incorporation of the *FLC* gene loop in the model, as outlined above, causes the model to generate a second peak in the *FLC* 3′ region, in the region that interacts strongly with the nucleation region. Consistent with predictions, we find that the ChIP profile of H3K27me3 in *clf-2* exhibits two peaks, at significantly lower levels relative to the spread state in L*er* (Figure 5E). Our results also aligned well with previously published genome-wide H3K27me3 data for a *clf* mutant in a Col-0 background.[40] Overall, predictions from the model for *clf* mutants match well with our experimental data.

The repressor FLD is co-transcriptionally targeted to *FLC* but requires FCA-mediated proximal termination for its repressive function

We next investigated the mechanism of FLD functionality in the co-transcriptional repression mechanism. In apparent contradiction to its role as a repressor, FLD occupancy was reported to be significantly higher in a high transcriptional state of *FLC*, relative to a Polycomb-silenced low transcriptional state.[31] Specifically, high FLD levels were assayed by ChIP-seq of a transgenic *fld-4; 3xFLAG-FLD* in the Col*FRI* background (Col-0 with an introgressed active *FRIGIDA* allele), where the *FLC* activator FRIGIDA (FRI) sets up a high transcriptional state. By contrast, in a Col-0 background, where FRI is not functional, *FLC* is Polycomb-silenced, and FLD shows low enrichment.

The model allows us to explain this counter-intuitive behavior of FLD. Since FLD function is coupled to transcription and termination in the model, FLD targeting must be co-transcriptional, consistent with the molecular and genetic analysis in the accompanying paper.[16] FLD levels at the locus must therefore be low in a Polycomb-silenced state, where transcription events are infrequent. Such co-transcriptional FLD targeting is consistent with the reported genome-wide correlation between FLD enrichment and transcription as measured by Pol II ChIP and transcript amounts.[31] By capturing the digital switch off in L*er* and *fca-3*, the model predicts that FLD enrichment would reduce over time in these genotypes. In essence, FLD is only needed transiently to generate a lowly transcribed state capable of nucleating Polycomb silencing, as transcription antagonizes Polycomb. Following Polycomb nucleation and spreading of H3K27me3, FLD recruitment to the locus is reduced due to the lack of transcription.

The experimental validation of the model assumption of FLD targeting coupled to transcription termination is presented in the accompanying paper,[16] where we elucidate the links between the FLD complex[26] and CPF components. By ChIP-qPCR analysis of FLD enrichment at *FLC*, comparing high transcription states induced by two very different genetic backgrounds (*fca-9* and *FRI*) with a Polycomb-silenced state (Col-0), we show that high FLD enrichment at *FLC* accompanies high transcription, irrespective of FCA function. Our model also supports the scenario proposed by Inagaki et al.[31] in which FLD downregulates histone H3K4me1 around termination sites in regions with convergent overlapping transcription genome-wide in *Arabidopsis*.

We tested whether FLD enrichment would change over time, as expected from the model-predicted switch to the Polycomb-silenced state. We measured 3xFLAG-FLD enrichment by ChIP-qPCR over a time course in a Col-0 background (as high-expressing *fca-9* and FRI are not expected to exhibit a digital switch OFF). Consistent with model predictions, we observed a reduction over time in FLD enrichment across *FLC* (Figure 6B), mirroring the reduction in *FLC* transcription levels over time also observed in Col-0 by qPCR (Figure 6A). Levels at the initial 7-day time point were high and declined to very low levels (almost background) by 21 days. Consistently, both 7- and 14-day average FLD levels across the locus showed statistically significant differences to the control sample, while 21-day levels did not (Figure S5Avi).

These observations are consistent with co-transcriptional FLD targeting to the *FLC* locus, with its repressive activity dependent on FCA-mediated proximal polyadenylation of sense and antisense transcription. In genotypes such as *fca-9*, where FCA is non-functional, or in the presence of the transcriptional activator FRIGIDA, which antagonizes proximal PA by FCA,[12] FLD continues to be co-transcriptionally targeted but fails to carry out its repressive function.

FCA-mediated transcription-coupled repression potentially controls many Polycomb targets

We analyzed our genome-wide Quant-seq data to examine whether the FCA-mediated transcription-coupled mechanism could be applicable to a wider range of targets beyond *FLC* and whether this mechanism is also linked to Polycomb silencing at other targets. Analysis of differential upregulation in *fca-1* (null mutant) relative to L*er* showed that 130 targets were significantly upregulated in *fca-1*. Of these 130 targets, 107 were upregulated to an intermediate level in *fca-3* (without necessarily showing statistically significant upregulation in *fca-3*, probably because of the intermediate nature of this genotype), hence behaving similar to *FLC* (Figure 6C). We examined whether these 107 targets were Polycomb targets. We used published ChIP-seq data for the *Arabidopsis* PRC2 methyltransferases CLF and SWN[40] (*SWINGER*) in the Col-0 genotype and generated a metagene plot for CLF and SWN enrichment across the 107 targets, comparing it to the whole genome metaplot for CLF and SWN enrichment. This showed a clear higher enrichment of these Polycomb components at these 107 putative FCA targets, particularly close to the transcription start site (TSS), compared with the genome-wide average (Figure 6D).

To examine whether FCA and FLD function overlap in general, we used genome-wide Quant-seq data in the null mutant genotypes *fca-9* and *fld-4* and the corresponding parental genotype Col-0 to carry out a similar analysis. We found that 91 genes were significantly upregulated in *fca-9* compared with the parental genotype, Col-0. We first examined whether these targets overlap with those observed in *fca*-1. Our data show extensive differential gene expression between the Col-0 and L*er* parental genotypes (not shown), and these differences are reflected in the respective mutants in these backgrounds. It is therefore not surprising that we find a relatively small overlap of 10 targets between the *fca-1*-upregulated set and the *fca-9*-upregulated set, even though these are both null mutants of *FCA*.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

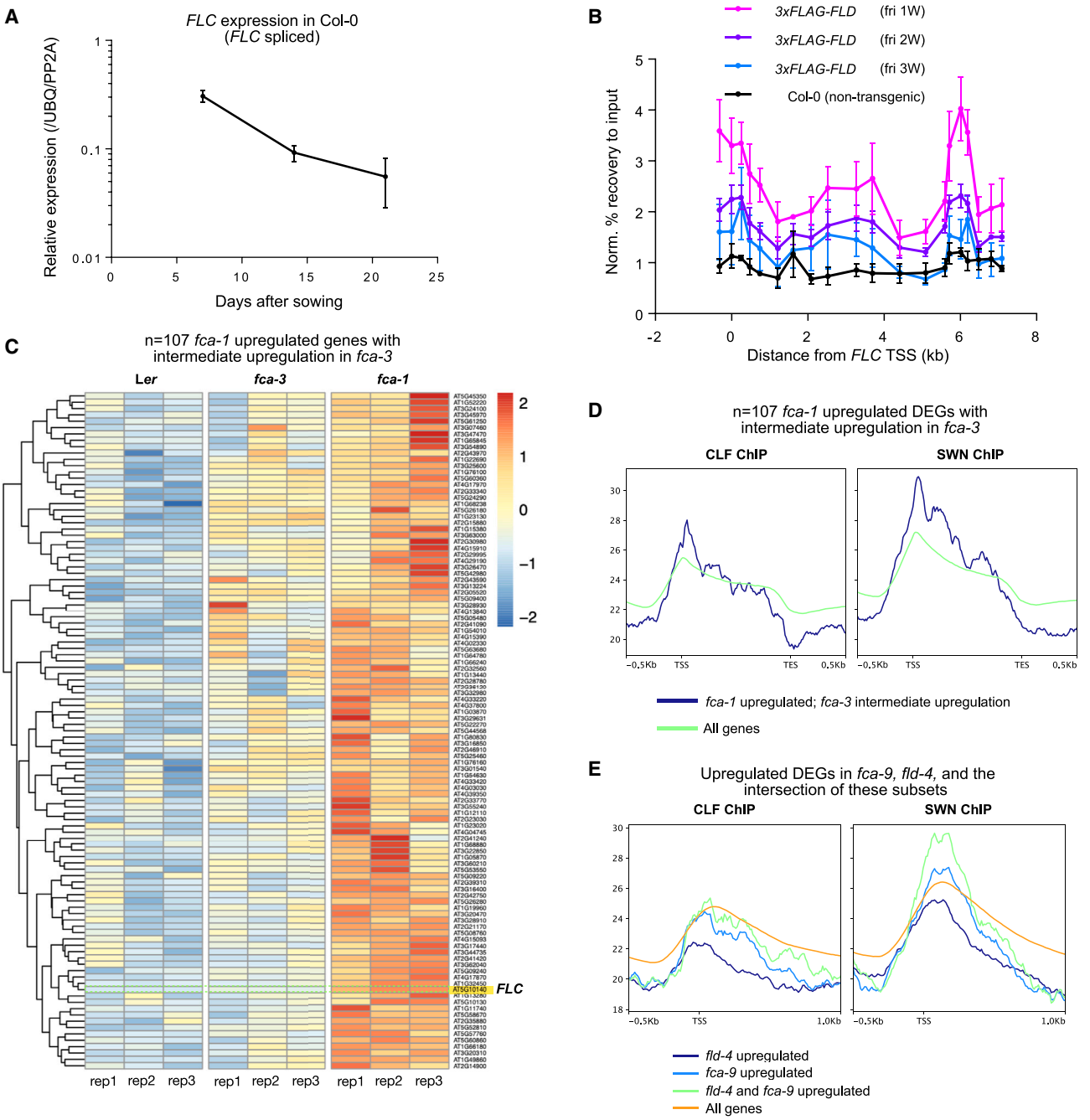


Figure 6. Co-transcriptional targeting of FLD and the potential genome-wide role of FCA/FLD-mediated transcription-coupled repression

(A) Changes over time in spliced *FLC* expression in Col-0 seedlings, measured by qPCR. Error bars: mean ± SEM. (*n* = 3 biological replicates). Data shown normalized to *UBQ10* and *PP2A*.

(B) Changes over time in FLD targeting to *FLC* locus, measured by time course ChIP-qPCR quantification of 3xFLAG-FLD in seedlings (Col-0 background) (normalized to IGN5). Error bars represent mean ± SEM. (*n* = 3 biological replicates).

(C) Heatmap generated from analyzing genome-wide Quant-seq data, showing changes in transcriptional output at 107 targets significantly upregulated in *fca-1* and which show intermediate (but not necessarily statistically significant) upregulation in *fca-3*.

(D) Metagene plots showing enrichment of PRC2 methyltransferases SWN and CLF at the 107 targets (differentially expressed genes, DEGs) shown in (C), compared with enrichment genome-wide.

(E) Metagene plots comparing SWN and CLF enrichment at *fca-9*-upregulated, *fld-4*-upregulated, and overlapping DEGs, with their genome-wide enrichment.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

CellPress
OPEN ACCESS

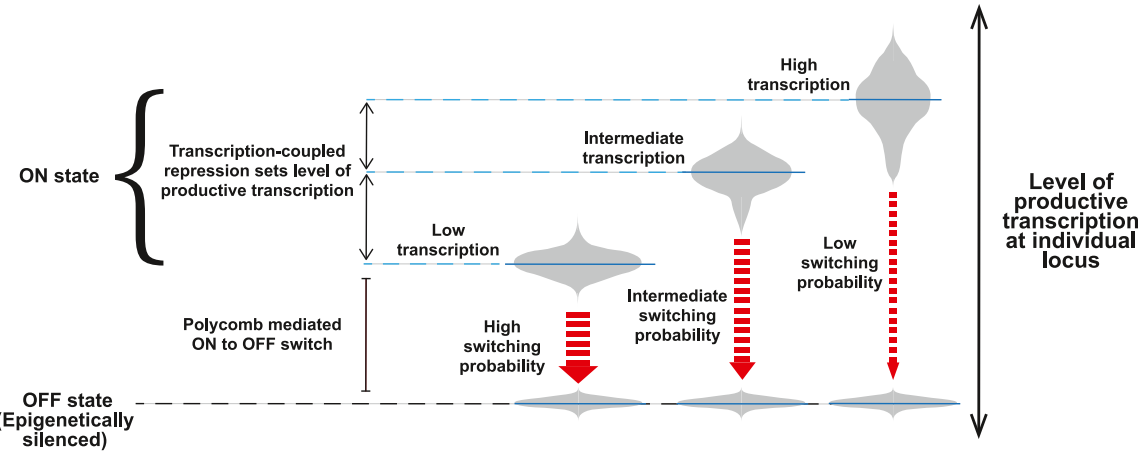


Figure 7. Schematic showing how the level of transcriptional activity set by the transcription-coupled repression can dictate the timescale for establishment of Polycomb silencing

Differences in red arrow thicknesses indicate different rates of switching to Polycomb-silenced OFF state.

Analysis of *fld-4* relative to Col-0 showed a larger number of upregulated targets (193), consistent with the proposed wider role for FLD in coordinating convergent transcription genome-wide.[31] We found that 48 upregulated targets overlapped between *fca-9* and *fld-4*, indicating that these factors may be working together at a number of genes. As above, we compared metagene plots of CLF and SWN enrichment across these 48 targets, as well as for the *fca-9* and *fld-4* singly upregulated subsets, in a whole genome metagene plot for CLF and SWN enrichment. For the *fca-9*-upregulated genes, we saw slightly higher enrichment of SWN but not of CLF around the TSS compared with the genome-wide average (Figure 6E). Interestingly, for the *fld-4*-upregulated subset, we observe lower enrichment of both SWN and CLF compared with the genome-wide average (Figure 6E). This is consistent with co-transcriptional targeting of FLD, but where the majority of targets are not Polycomb targets. However, the metagene plot for the 48 overlapping targets compared with the genome-wide-average shows a further increase in SWN enrichment (and a smaller increase for CLF) around the TSS than seen for the *fca-9*-upregulated targets alone (Figure 6E). This suggests that FLD function may be linked to silencing by PRC2 in specific contexts, where FLD action is facilitated by FCA-mediated proximal polyadenylation.

These analyses indicate that FCA/FLD-mediated transcription-coupled repression may be working at a broad range of targets and that at these targets, as at *FLC*, this repression works in concert with the Polycomb system. We also note that multiple genes of general interest are present in the *fca-1* and *fca-9*/*fld-4*-upregulated lists, including a Trithorax homolog (ATX1) implicated in development and multiple stress response genes such as GRP5[42] and GASA9[43] showing highly cell-type-specific expression (see Table S6 for gene names of selected targets), all features that are consistent with Polycomb regulation.

DISCUSSION

How co-transcriptional processing mechanistically influences transcriptional output and how this is linked to epigenetically stable Polycomb silencing is still not understood in any system. We have exploited the *Arabidopsis FLC* system to investigate these questions, taming the system's complexity by building a computational model to reveal how the level of productive transcriptional activity set by the transcription-coupled mechanism (FCA/FLD) determines the level of transcriptional antagonism to Polycomb silencing. The different levels of transcriptional activity set by the transcription-coupled repression can dictate the timescale for establishment of Polycomb silencing (Figure 7). We validated the model by a variety of means, including gene expression, histone modification, and FLC-Venus imaging time courses in the wild-type and in cases disrupted via a Polycomb mutation. Our model also resolves the apparent contradiction between the observed FLD targeting to a transcriptionally active *FLC* locus and its role as a repressor: FLD's co-transcriptional recruitment, needed to lower transcriptional activity to pave the way for Polycomb silencing, is eliminated once full Polycomb silencing is achieved. Hence, factors localizing at activated genes can in some instances be paradoxically essential for shutting down transcription.

The transcriptional antagonism to Polycomb silencing, a fundamental aspect of the silencing mechanism, is also important in the mammalian and *Drosophila* contexts.[3,44] In the plant context, this antagonism sets the timescale for the slow switch into the digital Polycomb-silenced state, in an interlinked three-way association between chromatin, transcription, and co-transcriptional processing. Importantly, our model predicts slow/inefficient establishment of Polycomb silencing in the *fca-1* mutant. As a result, if silencing can be established in *fca-1* (independently of FCA), for example, mediated by cold-specific PRC2 accessories in the case of cold-induced silencing, we predict that both silenced and active states can each be stably maintained. Hence, our conceptual framework also explains why cold-induced Polycomb silencing can be maintained in *fca* (or *fld*) mutants. The model predicts that these proteins are needed only for Polycomb silencing establishment during development but not for its maintenance. In future work, it would be very

interesting to specifically test this prediction by removing FCA and/or FLD at a later developmental stage after silencing is established.

The putative H3K4 demethylase FLD is known to be a key component of the FCA-mediated transcription-coupled repression mechanism, indicating an important role for removal of H3K4me associated with and downstream of proximal termination. A key mechanistic feature of the model is the existence of a link also in the opposite direction: H3K4me1 levels influencing the processivity of RNA Pol II—with low H3K4me1 levels reducing processivity. Thus, the model contains a positive feedback loop (mutual inhibition) between proximal termination and H3K4me1. Although chromatin modifications have been mechanistically linked to co-transcriptional splicing,[33] links to Pol II processivity are only beginning to be explored, as seen from recent findings on Pol II regulation by the Integrator complex.[45,46] These links could also involve an effect on Pol II speed, which can influence termination.[47]

Overall, we have revealed the mechanism of co-transcriptional repression at *FLC* and why it is an essential precursor for slow switching into the Polycomb-silenced state. These features are summarized in an animation to help explain these molecular feedbacks and how this mechanism can lead to expression of *FLC* in one of two stable expression states (Video S1). It is particularly intriguing that many elements of this pathway, such as FLD and FCA, are conserved transcriptional/splicing regulators that likely have functions at many more targets. Their core role in an integrated sense/antisense transcriptional circuitry likely singles out *FLC* as being acutely sensitive to their action. Indeed, our analysis of genome-wide Quant-seq data and overlap with ChIP-seq data for Polycomb components indicates that this transcription-coupled mechanism targets many genes and potentially works in concert with Polycomb at these targets. We anticipate these mechanisms will have broad relevance.

Limitations of the study

The model assumes that both sense and antisense transcription occur, but proximal sense transcripts have been difficult to detect. Establishing whether this results from differential turnover requires further work. Our assumption that sense and antisense transcription occurs at all *FLC* copies and at all developmental stages is a simplification, since there may be heterogeneity at the level of tissues and developmental stages; for instance, antisense transcripts may be detected in specific cell types.[48] Model comparisons to experimental data are limited by the experimental data for expression and histone modifications being at the whole-seedling level. To better define the initial conditions for the model, it will be necessary to measure the *FLC* chromatin state in the embryo. Although partial disruption of the Polycomb module through the *clf-2* mutant supports that transcription-coupled repression and Polycomb silencing work in parallel, we currently do not have a system to study the transcription-coupled mechanism in the absence of Polycomb targeting to *FLC*. A core feature of the proposed model—the link between H3K4me1 and Pol II processivity—also needs further mechanistic experimental validation. Finally, although the genome-wide analysis of the Quant-seq data suggests a potential wider role for the mechanism, demonstrating its mechanistic function at targets other than *FLC* requires further work.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Gene expression analyses
  - ChIP
  - Imaging
  - Quant-seq
  - Accession numbers
  - Bioinformatic analysis
  - Computational model summary
  - Model features
  - Additional considerations for the model
  - Model implementation
  - Simulations
  - Initial conditions
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of simulation output
  - Comparing simulation output to experimental data
  - Preventing proximal termination results in bistability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2024.05.014.

ACKNOWLEDGMENTS

We thank all members of the Dean and Howard labs and Prof. Lori A Passmore (MRC-LMB) for thoughtful inputs, Shuqin Chen for excellent technical assistance, and Soichi Inagaki for sharing their *3xFLAG-FLD* transgenic line.[31] Video S1 was produced by Margot Riggi and Janet Iwasa, U. of Utah. This work was funded by European Research Council Advanced Grant (EPISWITCH, 833254), Wellcome Trust (210654/Z/18/Z), and a Royal Society Professorship (RP\R1\180002) to C.D. and BBSRC Institute Strategic Programmes (BB/J004588/1 and BB/P013511/1) and EPSRC/BBSRC Physics of Life grant (EP/T00214X/1) to M.H. and C.D. E.M.B. acknowledges grant RYC2021-030895-I, funded by MCIN/AEI and by the European Union NextGenerationEU/PRTR. S.R. was funded by the HORIZON EUROPE Marie Sklodowska-Curie Actions PEP-NET (Predictive Epigenetics: Fusing Theory and Experiment) grant 813282.

AUTHOR CONTRIBUTIONS

Conceptualization, G.M., E.M.B., R.I., C.D., and M.H.; methodology, G.M., E.M.B., S.R., R.M., C.D., and M.H.; software, G.M.; validation, E.M.B., S.R., and R.M.; formal analysis, G.M. and R.M.; investigation, G.M., E.M.B., S.R., R.M., and Z.W.; data curation, G.M. and R.M.; writing – original draft, G.M., E.M.B., S.R., R.M., C.D., and M.H.; writing – review & editing, G.M., E.M.B., R.M., C.D., and M.H.; visualization, G.M., E.M.B., S.R., and R.M.; supervision, C.D. and M.H.; funding acquisition, C.D. and M.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 10, 2023
Revised: February 28, 2024
Accepted: May 14, 2024
Published: June 7, 2024

REFERENCES

1. Berry, S., Dean, C., and Howard, M. (2017). Slow Chromatin Dynamics Allow Polycomb Target Genes to Filter Fluctuations in Transcription Factor Activity. Cell Syst. *4*, 445–457.e8. https://doi.org/10.1016/j.cels.2017.02.013.
2. Dodd, I.B., Micheelsen, M.A., Sneppen, K., and Thon, G. (2007). Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell *129*, 813–822. https://doi.org/10.1016/j.cell.2007.02.053.
3. Holoch, D., Wassef, M., Lövkvist, C., Zielinski, D., Aflaki, S., Lombard, B., Héry, T., Loew, D., Howard, M., and Margueron, R. (2021). A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. Nat. Genet. *53*, 1686–1697. https://doi.org/10.1038/s41588-021-00964-2.
4. Pease, N.A., Nguyen, P.H.B., Woodworth, M.A., Ng, K.K.H., Irwin, B., Vaughan, J.C., and Kueh, H.Y. (2021). Tunable, division-independent control of gene activation timing by a polycomb switch. Cell Rep. *34*, 108888. https://doi.org/10.1016/j.celrep.2021.108888.
5. Menon, G., Schulten, A., Dean, C., and Howard, M. (2021). Digital paradigm for Polycomb epigenetic switching and memory. Curr. Opin. Plant Biol. *61*, 102012. https://doi.org/10.1016/j.pbi.2021.102012.
6. Angel, A., Song, J., Dean, C., and Howard, M. (2011). A Polycomb-based switch underlying quantitative epigenetic memory. Nature *476*, 105–108. https://doi.org/10.1038/nature10241.
7. Berry, S., Hartley, M., Olsson, T.S.G., Dean, C., and Howard, M. (2015). Local chromatin environment of a Polycomb target gene instructs its own epigenetic inheritance. eLife *4*, e07205. https://doi.org/10.7554/eLife.07205.
8. Csorba, T., Questa, J.I., Sun, Q., and Dean, C. (2014). Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. Proc. Natl. Acad. Sci. USA *111*, 16160–16165. https://doi.org/10.1073/pnas.1419030111.
9. Lövkvist, C., and Howard, M. (2021). Using computational modelling to reveal mechanisms of epigenetic Polycomb control. Biochem. Soc. Trans. *49*, 71–77. https://doi.org/10.1042/BST20190955.
10. Yang, H., Berry, S., Olsson, T.S.G., Hartley, M., Howard, M., and Dean, C. (2017). Distinct phases of Polycomb silencing to hold epigenetic memory of cold in *Arabidopsis*. Science *357*, 1142–1145. https://doi.org/10.1126/science.aan1121.
11. Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C. (2003). FY Is an RNA 3′ End-Processing Factor that Interacts with FCA to Control the Arabidopsis Floral Transition. Cell *113*, 777–787. https://doi.org/10.1016/S0092-8674(03)00425-2.
12. Schon, M., Baxter, C., Xu, C., Enugutti, B., Nodine, M.D., and Dean, C. (2021). Antagonistic activities of cotranscriptional regulators within an early developmental window set FLC expression level. Proc. Natl. Acad. Sci. USA *118*, e2102753118. https://doi.org/10.1073/pnas.2102753118.
13. Liu, F., Marquardt, S., Lister, C., Swiezewski, S., and Dean, C. (2010). Targeted 3′ Processing of Antisense Transcripts Triggers *Arabidopsis FLC* Chromatin Silencing. Science *327*, 94–97. https://doi.org/10.1126/science.1180278.
14. Liu, F., Quesada, V., Crevillén, P., Bäurle, I., Swiezewski, S., and Dean, C. (2007). The Arabidopsis RNA-Binding Protein FCA Requires a Lysine-Specific Demethylase 1 Homolog to Downregulate FLC. Mol. Cell *28*, 398–407. https://doi.org/10.1016/j.molcel.2007.10.018.
15. Martignago, D., Bernardini, B., Polticelli, F., Salvi, D., Cona, A., Angelini, R., and Tavladoraki, P. (2019). The Four FAD-Dependent Histone Demethylases of Arabidopsis Are Differently Involved in the Control of Flowering Time. Front. Plant Sci. *10*, 669. https://doi.org/10.3389/fpls.2019.00669.
16. Mateo-Bonmatí, E., Montez, M., Maple, R., Fiedler, M., Fang, X., Saalbach, G., Passmore, A.L., and Dean, C. (2024). A CPF-like phosphatase module links transcription termination to chromatin silencing. Mol. Cell. *84*. https://doi.org/10.1101/2023.07.06.547976.
17. Shimada, Y., Carl, S.H., Skribbe, M., Flury, V., Kuzdere, T., Kempf, G., and Bühler, M. (2021). An enhancer screen identifies new suppressors of small-RNA-mediated epigenetic gene silencing. PLOS Genet. *17*, e1009645. https://doi.org/10.1371/journal.pgen.1009645.
18. Chalamcharla, V.R., Folco, H.D., Dhakshnamoorthy, J., and Grewal, S.I.S. (2015). Conserved factor Dhp1/Rat1/Xrn2 triggers premature transcription termination and nucleates heterochromatin to promote gene silencing. Proc. Natl. Acad. Sci. USA *112*, 15548–15555. https://doi.org/10.1073/pnas.1522127112.
19. Tucker, J.F., Ohle, C., Schermann, G., Bendrin, K., Zhang, W., Fischer, T., and Zhang, K. (2016). A Novel Epigenetic Silencing Pathway Involving the Highly Conserved 5'-3' Exoribonuclease Dhp1/Rat1/Xrn2 in Schizosaccharomyces pombe. PLOS Genet. *12*, e1005873. https://doi.org/10.1371/journal.pgen.1005873.
20. Vo, T.V., Dhakshnamoorthy, J., Larkin, M., Zofall, M., Thillainadesan, G., Balachandran, V., Holla, S., Wheeler, D., and Grewal, S.I.S. (2019). CPF Recruitment to Non-canonical Transcription Termination Sites Triggers Heterochromatin Assembly and Gene Silencing. Cell Rep. *28*, 267–281.e5. https://doi.org/10.1016/j.celrep.2019.05.107.
21. Marina, D.B., Shankar, S., Natarajan, P., Finn, K.J., and Madhani, H.D. (2013). A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. Genes Dev. *27*, 1851–1856. https://doi.org/10.1101/gad.226019.113.
22. Wittmann, S., Renner, M., Watts, B.R., Adams, O., Huseyin, M., Baejen, C., El Omari, K., Kilchert, C., Heo, D.-H., Kecman, T., et al. (2017). The conserved protein Seb1 drives transcription termination by binding RNA polymerase II and nascent RNA. Nat. Commun. *8*, 14861. https://doi.org/10.1038/ncomms14861.
23. Kaneko, S., Son, J., Bonasio, R., Shen, S.S., and Reinberg, D. (2014). Nascent RNA interaction keeps PRC2 activity poised and in check. Genes Dev. *28*, 1983–1988. https://doi.org/10.1101/gad.247940.114.
24. Antoniou-Kourounioti, R.L., Meschichi, A., Reeck, S., Berry, S., Menon, G., Zhao, Y., Fozard, J., Holmes, T., Zhao, L., Wang, H., et al. (2023). Integrating analog and digital modes of gene expression at Arabidopsis FLC. eLife *12*, e79743. https://doi.org/10.7554/eLife.79743.
25. Sonmez, C., Bäurle, I., Magusin, A., Dreos, R., Laubinger, S., Weigel, D., and Dean, C. (2011). RNA 3′ processing functions of *Arabidopsis* FCA and FPA limit intergenic transcription. Proc Natl Acad Sci USA *108*, 8508–8513. https://doi.org/10.1073/pnas.1105334108.
26. Fang, X., Wang, L., Ishikawa, R., Li, Y., Fiedler, M., Liu, F., Calder, G., Rowan, B., Weigel, D., Li, P., and Dean, C. (2019). Arabidopsis FLL2 promotes liquid-liquid phase separation of polyadenylation complexes. Nature *569*, 265–269. https://doi.org/10.1038/s41586-019-1165-8.
27. Wang, Z.W., Wu, Z., Raitskin, O., Sun, Q., and Dean, C. (2014). Antisense-mediated FLC transcriptional repression requires the P-TEFb transcription elongation factor. Proc. Natl. Acad. Sci. USA *111*, 7468–7473. https://doi.org/10.1073/pnas.1406635111.
28. Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., and Dean, C. (2014). Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. Mol. Cell *54*, 156–165. https://doi.org/10.1016/j.molcel.2014.03.026.
29. Baxter, C.L., Šviković, S., Sale, J.E., Dean, C., and Costa, S. (2021). The intersection of DNA replication with antisense 3′ RNA processing in *Arabidopsis FLC* chromatin silencing. Proc. Natl. Acad. Sci. USA *118*, e2107483118. https://doi.org/10.1073/pnas.2107483118.

30. Xu, C., Wu, Z., Duan, H.C., Fang, X., Jia, G., and Dean, C. (2021). R-loop resolution promotes co-transcriptional chromatin silencing. Nat. Commun. *12*, 1790. https://doi.org/10.1038/s41467-021-22083-6.

31. Inagaki, S., Takahashi, M., Takashima, K., Oya, S., and Kakutani, T. (2021). Chromatin-based mechanisms to coordinate convergent overlapping transcription. Nat. Plants *7*, 295–302. https://doi.org/10.1038/s41477-021-00868-3.

32. Fang, X., Wu, Z., Raitskin, O., Webb, K., Voigt, P., Lu, T., Howard, M., and Dean, C. (2020). The 3′ processing of antisense RNAs physically links to chromatin-based transcriptional control. Proc. Natl. Acad. Sci. USA *117*, 15316–15321. https://doi.org/10.1073/pnas.2007268117.

33. Leung, C.S., Douglass, S.M., Morselli, M., Obusan, M.B., Pavlyukov, M.S., Pellegrini, M., and Johnson, T.L. (2019). H3K36 Methylation and the Chromodomain Protein Eaf3 Are Required for Proper Cotranscriptional Spliceosome Assembly. Cell Rep. *27*, 3760–3769.e4. https://doi.org/10.1016/j.celrep.2019.05.100.

34. Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature *462*, 799–802. https://doi.org/10.1038/nature08618.

35. Nojima, T., and Proudfoot, N.J. (2022). Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. Nat. Rev. Mol. Cell Biol. *23*, 389–406. https://doi.org/10.1038/s41580-021-00447-6.

36. Zhao, Y., Zhu, P., Hepworth, J., Bloomer, R., Antoniou-Kourounioti, R.L., Doughty, J., Heckmann, A., Xu, C., Yang, H., and Dean, C. (2021). Natural temperature fluctuations promote *COOLAIR* regulation of *FLC*. Genes Dev. *35*, 888–898. https://doi.org/10.1101/gad.348362.121.

37. Moll, P., Ante, M., Seitz, A., et al. (2014). QuantSeq 3′ mRNA sequencing for RNA quantification. Nat. Methods *11*. i–iii. https://doi.org/10.1038/nmeth.f.376.

38. Koornneef, M., Hanhart, C.J., and van der Veen, J.H. (1991). A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Mol. Gen. Genet. *229*, 57–66. https://doi.org/10.1007/BF00264213.

39. Wu, Z., Ietswaart, R., Liu, F., Yang, H., Howard, M., and Dean, C. (2016). Quantitative regulation of *FLC* via coordinated transcriptional initiation and elongation. Proc. Natl. Acad. Sci. USA *113*, 218–223. https://doi.org/10.1073/pnas.1518369112.

40. Shu, J., Chen, C., Thapa, R.K., Bian, S., Nguyen, V., Yu, K., Yuan, Z.-C., Liu, J., Kohalmi, S.E., Li, C., and Cui, Y. (2019). Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in Arabidopsis seedlings. Plant Direct *3*, e00100. https://doi.org/10.1002/pld3.100.

41. Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature *386*, 44–51. https://doi.org/10.1038/386044a0.

42. Mangeon, A., Magioli, C., Tarré, E., Cardeal, V., Araujo, C., Falkenbach, E., Rocha, C.A.B., Rangel-Lima, C., and Sachetto-Martins, G. (2010). The tissue expression pattern of the AtGRP5 regulatory region is controlled by a combination of positive and negative elements. Plant Cell Rep. *29*, 461–471. https://doi.org/10.1007/s00299-010-0835-7.

43. Yang, Y., Costa, A., Leonhardt, N., Siegel, R.S., and Schroeder, J.I. (2008). Isolation of a strong Arabidopsis guard cell promoter and its potential as a research tool. Plant Methods *4*, 6. https://doi.org/10.1186/1746-4811-4-6.

44. Klose, R.J., Cooper, S., Farcas, A.M., Blackledge, N.P., and Brockdorff, N. (2013). Chromatin Sampling—An Emerging Perspective on Targeting Polycomb Repressor Proteins. PLOS Genet. *9*, e1003717. https://doi.org/10.1371/journal.pgen.1003717.

45. Stein, C.B., Field, A.R., Mimoso, C.A., Zhao, C., Huang, K.L., Wagner, E.J., and Adelman, K. (2022). Integrator endonuclease drives promoter-proximal termination at all RNA polymerase II-transcribed loci. Mol. Cell *82*, 4232–4245.e11. https://doi.org/10.1016/j.molcel.2022.10.004.

46. Wang, H., Fan, Z., Shliaha, P.V., Miele, M., Hendrickson, R.C., Jiang, X., and Helin, K. (2023). H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. Nature *615*, 339–348. https://doi.org/10.1038/s41586-023-05780-8.

47. Cortazar, M.A., Sheridan, R.M., Erickson, B., Fong, N., Glover-Cutter, K., Brannan, K., and Bentley, D.L. (2019). Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "sitting duck torpedo" mechanism. Mol. Cell *76*, 896–908.e4. https://doi.org/10.1016/j.molcel.2019.09.031.

48. Ietswaart, R., Rosa, S., Wu, Z., Dean, C., and Howard, M. (2017). Cell-Size-Dependent Transcription of *FLC* and Its Antisense Long Non-coding RNA *COOLAIR* Explain Cell-to-Cell Expression Variation. Cell Syst. *4*, 622–635.e9. https://doi.org/10.1016/j.cels.2017.05.010.

49. Andrews, S. (2010). FastQC: a Quality Control Tool for High Throughput Sequence Data. Cambridge, United Kingdom.

50. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. j. *17*, 3. https://doi.org/10.14806/ej.17.1.200.

51. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

52. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

53. Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. *81*, 2340–2361. https://doi.org/10.1021/j100540a008.

54. Box, M.S., Coustham, V., Dean, C., Protocol, et al. (2011). A simple phenol-based method for 96-well extraction of high quality RNA from Arabidopsis. Plant Methods *7*, 7. https://doi.org/10.1186/1746-4811-7-7.

55. Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N.J., and Dean, C. (2013). R-Loop Stabilization Represses Antisense Transcription at the *Arabidopsis FLC* Locus. Science *340*, 619–621. https://doi.org/10.1126/science.1234848.

56. Crevillén, P., Sonmez, C., Wu, Z., and Dean, C. (2013). A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J. *32*, 140–148. https://doi.org/10.1038/emboj.2012.324.

57. Sun, L., Zhou, J., Xu, X., Liu, Y., Ma, N., Liu, Y., Nie, W., Zou, L., Deng, X.W., and He, H. (2024). Mapping nucleosome-resolution chromatin organization and enhancer-promoter loops in plants using Micro-C-XL. Nat. Commun. *15*, 35. https://doi.org/10.1038/s41467-023-44347-z.

58. Shema, E., Jones, D., Shoresh, N., Donohue, L., Ram, O., and Bernstein, B.E. (2016). Single-molecule decoding of combinatorially modified nucleosomes. Science *352*, 717–721. https://doi.org/10.1126/science.aad7701.

59. Voigt, P., LeRoy, G., Drury, W.J., III, Zee, B.M., Son, J., Beck, D.B., Young, N.L., Garcia, B.A., and Reinberg, D. (2012). Asymmetrically Modified Nucleosomes. Cell *151*, 181–193. https://doi.org/10.1016/j.cell.2012.09.002.

60. Schmitges, F.W., Prusty, A.B., Faty, M., Stützer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone Methylation by PRC2 Is Inhibited by Active Chromatin Marks. Mol. Cell *42*, 330–341. https://doi.org/10.1016/j.molcel.2011.03.025.

61. Kim, D.-H., Tang, Z., Shimada, M., Fierz, B., Houck-Loomis, B., Bar-Dagen, M., Lee, S., Lee, S.-K., Muir, T.W., Roeder, R.G., and Lee, J.W. (2013). Histone H3K27 Trimethylation Inhibits H3 Binding and Function of SET1-Like H3K4 Methyltransferase Complexes. Mol. Cell. Biol. *33*, 4936–4946. https://doi.org/10.1128/MCB.00601-13.

62. McCabe, M.T., Graves, A.P., Ganji, G., Diaz, E., Halsey, W.S., Jiang, Y., Smitheman, K.N., Ott, H.M., Pappalardi, M.B., Allen, K.E., et al. (2012). Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proc. Natl. Acad. Sci. USA *109*, 2989–2994. https://doi.org/10.1073/pnas.1116418109.

63. Soares, L.M., He, P.C., Chun, Y., Suh, H., Kim, T., and Buratowski, S. (2017). Determinants of Histone H3K4 Methylation Patterns. Mol. Cell *68*, 773–785.e6. https://doi.org/10.1016/j.molcel.2017.10.013.

64. Blackledge, N.P., and Klose, R.J. (2021). The molecular principles of gene regulation by Polycomb repressive complexes. Nat. Rev. Mol. Cell Biol. *22*, 815–833. https://doi.org/10.1038/s41580-021-00398-y.

65. Zhang, Y.-Z., Yuan, J., Zhang, L., Chen, C., Wang, Y., Zhang, G., Peng, L., Xie, S.-S., Jiang, J., Zhu, J.-K., et al. (2020). Coupling of H3K27me3 recognition with transcriptional repression through the BAH-PHD-CPL2 complex in Arabidopsis. Nat. Commun. *11*, 6212. https://doi.org/10.1038/s41467-020-20089-0.

66. Brookes, E., de Santiago, I., Hebenstreit, D., Morris, K.J., Carroll, T., Xie, S.Q., Stock, J.K., Heidemann, M., Eick, D., Nozaki, N., et al. (2012). Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell Stem Cell *10*, 157–170. https://doi.org/10.1016/j.stem.2011.12.017.

67. Chan, J., Kumar, A., and Kono, H. (2022). RNAPII driven post-translational modifications of nucleosomal histones. Trends Genet. *38*, 1076–1095. https://doi.org/10.1016/j.tig.2022.04.010.

68. Oya, S., Takahashi, M., Takashima, K., Kakutani, T., and Inagaki, S. (2022). Transcription-coupled and epigenome-encoded mechanisms direct H3K4 methylation. Nat. Commun. *13*, 4521. https://doi.org/10.1038/s41467-022-32165-8.

69. Tamada, Y., Yun, J.-Y., Woo, S.C., and Amasino, R.M. (2009). ARABIDOPSIS TRITHORAX-RELATED7 Is Required for Methylation of Lysine 4 of Histone H3 and for Transcriptional Activation of *FLOWERING LOCUS C*. Plant Cell *21*, 3257–3269. https://doi.org/10.1105/tpc.109.070060.

70. Li, Z., Jiang, D., and He, Y. (2018). *FRIGIDA* establishes a local chromosomal environment for *FLOWERING LOCUS C* mRNA production. Nat. Plants *4*, 836–846. https://doi.org/10.1038/s41477-018-0250-6.

71. Smith, E.R., Lee, M.G., Winter, B., Droz, N.M., Eissenberg, J.C., Shiekhattar, R., and Shilatifard, A. (2008). *Drosophila* UTX Is a Histone H3 Lys27 Demethylase That Colocalizes with the Elongating Form of RNA Polymerase II. Mol. Cell. Biol. *28*, 1041–1046. https://doi.org/10.1128/MCB.01504-07.

72. Yang, H., Howard, M., and Dean, C. (2016). Physical coupling of activation and derepression activities to maintain an active transcriptional state at *FLC*. Proc. Natl. Acad. Sci. USA *113*, 9369–9374. https://doi.org/10.1073/pnas.1605733113.

73. Rosa, S., Duncan, S., and Dean, C. (2016). Mutually exclusive sense–antisense transcription at *FLC* facilitates environmentally induced gene repression. Nat. Commun. *7*, 13031. https://doi.org/10.1038/ncomms13031.

74. Escobar, T.M., Loyola, A., and Reinberg, D. (2021). Parental nucleosome segregation and the inheritance of cellular identity. Nat. Rev. Genet. *22*, 379–392. https://doi.org/10.1038/s41576-020-00312-w.

75. Michaels, S.D., He, Y., Scortecci, K.C., and Amasino, R.M. (2003). Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc. Natl. Acad. Sci. USA *100*, 10102–10107. https://doi.org/10.1073/pnas.1531467100.

76. Liu, J., He, Y., Amasino, R., and Chen, X. (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes Dev. *18*, 2873–2878. https://doi.org/10.1101/gad.1217304.

77. Tian, Y., Zheng, H., Zhang, F., Wang, S., Ji, X., Xu, C., He, Y., and Ding, Y. (2019). PRC2 recruitment and H3K27me3 deposition at *FLC* require FCA binding of *COOLAIR*. Sci. Adv. *5*, eaau7246. https://doi.org/10.1126/sciadv.aau7246.

78. Liang, S.C., Hartwig, B., Perera, P., Mora-García, S., de Leau, E., Thornton, H., de Lima Alves, F.L., Rappsilber, J., Yang, S., James, G.V., et al. (2015). Kicking against the PRCs – A Domesticated Transposase Antagonises Silencing Mediated by Polycomb Group Proteins and Is an Accessory Component of Polycomb Repressive Complex 2. PLOS Genet. *11*, e1005660. https://doi.org/10.1371/journal.pgen.1005660.

79. Zee, B.M., Levin, R.S., Xu, B., LeRoy, G., Wingreen, N.S., and Garcia, B.A. (2010). In vivo residue-specific histone methylation dynamics. J. Biol. Chem. *285*, 3341–3350. https://doi.org/10.1074/jbc.M109.063784.

80. Lövkvist, C., Mikulski, P., Reeck, S., Hartley, M., Dean, C., and Howard, M. (2021). Hybrid protein assembly-histone modification mechanism for PRC2-based epigenetic switching and memory. eLife *10*. https://doi.org/10.7554/eLife.66454.

81. Nielsen, M., Menon, G., Zhao, Y., Mateo-Bonmati, E., Wolff, P., Zhou, S., Howard, M., and Dean, C. (2024). COOLAIR and PRC2 function in parallel to silence FLC during vernalization. Proc. Natl. Acad. Sci. USA *121*, e2311474121. https://doi.org/10.1073/pnas.2311474121.

82. Rahni, R., and Birnbaum, K.D. (2019). Week-long imaging of cell divisions in the Arabidopsis root meristem. Plant Methods *15*, 30. https://doi.org/10.1186/s13007-019-0417-9.

83. Matsumoto, M., and Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Model. Comput. Simul. *8*, 3–30. https://doi.org/10.1145/272991.272995.

84. Qüesta, J.I., Antoniou-Kourounioti, R.L., Rosa, S., Li, P., Duncan, S., Whittaker, C., Howard, M., and Dean, C. (2020). Noncoding SNPs influence a distinct phase of Polycomb silencing to destabilize long-term epigenetic memory at Arabidopsis *FLC*. Genes Dev. *34*, 446–461. https://doi.org/10.1101/gad.333245.119.

85. Gregersen, L.H., Mitter, R., Ugalde, A.P., Nojima, T., Proudfoot, N.J., Agami, R., Stewart, A., and Svejstrup, J.Q. (2019). SCAF4 and SCAF8, mRNA Anti-Terminator Proteins. Cell *177*, 1797–1813.e18. https://doi.org/10.1016/j.cell.2019.04.038.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-H3 antibody	Abcam	Cat: ab176842; RRID: AB_2493104
Anti-H3K27me3 antibody	Abcam	Cat: ab192985; RRID: AB_2650559
Anti-H3K36me3 antibody	Abcam	Cat: ab9050; RRID: AB_306996
Anti-H3K4me1 antibody	Abcam	Cat: ab8895; RRID: AB_306847
Anti-FLAG ® antibody	Merck	Cat: F1804; RRID: AB_262044
Chemical, peptides, and recombinant proteins		
Turbo DNase	Ambion	Cat: AM1907
SuperScript IV	Invitrogen	Cat: 18090050
Phenol solution saturated with 0.1 M Citrate	Merk Life Science UK Ltd	Cat: P4682
RNaseOUT RNase Inhibitor	ThermoFisher Scientifc	Cat: 10777019
Lightcycler 480 Sybr Green I	Roche Diagnostics Ltd	Cat: 04887352001
Propidium iodide	Merck	Cat: P4864
cOmplete protease inhibitors	Merck	Cat: 11697498001
Protein A-coated Dynabeads	ThermoFisher Scientific	Cat: 10001D
M-270 Epoxy Dynabeads	ThermoFisher Scientific	Cat: 14311D
EGS (ethylene glycol bis(succinimidyl succinate))	ThermoFisher Scientific	Cat: 21565
Chelex resin	BioRad	Cat: 1421253
Proteinase K	Merck	Cat: 3115887001
Deposited data		
Quant-seq Col-0, *fld-4, fca-9*	SRA	SRA: PRJNA1076161
Quant-seq L*er*, *fca-1, fca-3, clf-2*	SRA	SRA: PRJNA980626
FLC-Venus Images	Mendeley Data	Mendeley Data: https://doi.org/10.17632/963wwcgs33.2
Experimental models: Organisms/strains		
Arabidopsis thaliana Ler	Standard accession	N/A
fca-1	Koornneef et al.[38]	N/A
fca-3	Koornneef et al.[38]	N/A
clf-2	Goodrich et al.[41]	N/A
fca-4	Koornneef et al.[38]	N/A
Arabidopsis thaliana Col-0	Standard accession	N/A
fca-9	Fang et al.[32]	N/A
fld-4	Fang et al.[32]	N/A
fld-4; FLD_{pro}:3xFLAG-FLD; fri	Inagaki et al.[31]	N/A
Oligonucleotides	Dean lab	N/A
Primers used in this study are listed in Table S5.	–	N/A
Software and algorithms		
Simulation and Analysis code	This study	https://doi.org/10.5281/zenodo.11186918
Microsoft Excel	Microsoft	https://www.microsoft.com
GraphPad Prism 10	GraphPad	https://www.graphpad.com
FastQC v0.11.7	Andrews[49]	https://github.com/s-andrews/FastQC
Cutadapt 1.18	Martin[50]	https://github.com/marcelm/cutadapt
STAR v2.6.1a	Dobin et al.[51]	https://github.com/alexdobin/STAR
Trimmomatic v0.39	Bolger et al.[52]	https://github.com/usadellab/Trimmomatic
Gillespie exact stochastic simulation algorithm	Gillespie[53]	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Martin Howard (martin.howard@jic.ac.uk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The Quant-seq data has been deposited at SRA: PRJNA980626 and SRA: PRJNA1076161, and will be publicly available at the time of publication.
- The code for model simulations and analysis can be accessed at https://doi.org/10.5281/zenodo.11186918. *FLC*-Venus time-course imaging data can be accessed at Mendeley Data: https://doi.org/10.17632/963wwcgs33.2.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Arabidopsis thaliana plants were used in this study. All plants were homozygous for the indicated genotype. The mutants *fca-1, fca-3, fca-4, clf-2* were in a Landsberg (*Ler*) background. The *fca* mutants used in this study were previously described in Antoniou-Kourounioti et al.[24] The *clf-2* mutant was previously described in Goodrich et al. [41] FLC-Venus transgenic lines in *FCA* mutant backgrounds were previously described in Antoniou-Kourounioti et al. [24] The *fca-9* and *fld-4* mutants were in Columbia (Col) background. The transgenic *fld-4*; FLD_{pro}*:3xFLAG-FLD* line was previously described[31] and was shared by Soichi Inagaki.

Seeds were surface sterilized in 40 % v/v commercial bleach for 10 min and rinsed 4 times with sterile distilled water. Seeds were then sown on standard half-strength Murashige and Skoog (MS) medium (0.22% MS, 8% plant agar) media plates and kept at 4°C in darkness for 3 days before being transferred to long day photoperiod conditions (16 h of light, 8 h dark).

METHOD DETAILS

Gene expression analyses

Seedlings were harvested and total RNA was extracted as previously described (Box et al.[54]). Genomic DNA was digested with TURBO DNA-free (Ambion Turbo Dnase kit AM1907) according to manufacturer's guidelines, before reverse transcription (RT) was performed. The RT reactions were performed with SuperScript IV reverse transcriptase (ThermoFisher, 18090010) following the manufacturer's instructions using gene-specific primers. Primers are listed in Table S5. qPCR analyses were performed, and data was normalized to the indicated housekeeping gene or genes.

ChIP

For the timecourse measurement of histone modification levels at FLC, 2.5 g of 7-, 14-, or 21-day-old seedlings were crosslinked in 1% formaldehyde and ground to fine powder. Nuclei were extracted using Honda buffer as described previously in Sun et al.[55] In all histone ChIP reactions, sonication, immunoprecipitation, DNA recovery and purification were performed as previously described in Wu et al.[39] The antibodies used were anti-H3 (ab176842), anti-H3K27me3 (Merck 07-449 and ab192985), anti-H3K36me3 (ab9050), anti-H3K4me1 (ab8895). All ChIP data were quantified by qPCR with primers listed in Table S5. Values were normalized to H3 and internal control to either STM or ACT7.

For H3K27me3 ChIP in the *clf-2* mutant, 2.0 g of 14-day-old seedlings were crosslinked in 1% formaldehyde and ground to fine powder. Nuclei were extracted using Honda buffer as described previously in Sun et al.[55] In all histone ChIP reactions, sonication, immunoprecipitation, DNA recovery and purification were performed as previously described in Wu et al.[39] The antibodies used were anti-H3 (ab176842), and anti-H3K27me3 (ab192985). All ChIP data were quantified by qPCR with primers listed in Table S5. Values were normalized to H3 and internal control *STM*.

For FLAG-FLD ChIP, 2.5 g seedlings were crosslinked in 0.685 mg/mL of EGS [ethylene glycol bis (succinimidyl succinate)] and 1% formaldehyde and ground to fine powder. Nuclei were extracted using Honda buffer as described for histone ChIP. Nuclei pellets were resuspended in 600 μl of RIPA buffer (1% NP-40, 0.5% Sodium Deoxycholate, 0.1 % SDS in PBS) and sonicated 5 times for 5 min (30s ON/ 30s OFF) at high mode in a Bioruptor Pico (Diagenode). IP were performed overnight at 4°C using 1.5 mg of pre-coated epoxy Epoxy Dynabeads (M-270) with 1.5 μl of anti-FLAG (Sigma, F1804) per reaction. Beads were then washed for 5 min twice with Low Salt Buffer (150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl), twice with Hight Salt Buffer (500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl), and twice with TE buffer (1 mM EDTA, 10 mM Tris-HCl). DNA recovery and purifications were performed as for Histone ChIP. Results in Figure 6B were normalized to IGN5.

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

Imaging

Seeds were surface-sterilized with 5% (v/v) sodium hypochlorite for 5 min, rinsed with sterile water 4x for 1 min and stratified at 4°C for 2 days in water and in the dark. Seeds were plated on MES buffered Murashige and Skoog (GM-MS) media, pH 5.8 and grown on vertically oriented petri plates for 7 days in long-day conditions (16 h light/8 h dark, 22/20°C).

Time course imaging of FLC-Venus protein levels in the epidermis of L*er*, *fca* allelic mutants, the *clf-2* mutant as well as the *fca-3 clf-2* mutant, was performed using a Leica confocal Stellaris 8 microscope with a 20x multi-immersion objective (0.75 NA). Root tips (2 cm) were cut off and immersed in 1.5 μg/mL Propidium Iodide (PI, Sigma–Aldrich, P4864) in 1x PBS before being imaged immediately. The OPSL 514 laser was used at 5% power with 514 nm excitation for FLC-Venus and PI. In photon counting mode, Venus was detected between 518-550 nm with the HyD SMD2 detector, PI was detected between ~600-650 nm. Images were acquired with a laser speed of 400 Hz, line accumulation of 6 (pixel dwell time of 2.43 μs), a Z-step size of 0.95 μm and a pinhole size of 1 AU. The same settings were used at all imaging time points to allow direct comparison between genotypes and time points. The representative images were projected using a single middle slice from the PI channel to show the cell outline, and 10 slices of FLC-Venus channel were maximum intensity projected. In standard image presentation, the dynamic range of the FLC-Venus channel was pushed from 0-255 to 2-11, in enhanced image presentation, the dynamic ranged was pushed from 2-7. In Data S4, the dynamic range of FLC-Venus channel was enhanced from 0-255 to 2-20 for all images.

To confirm the generality of our results for *fca-3*, we also observed similar behaviour of *FLC* expression in single cells for a different intermediate mutant of *FCA*, called *fca-4*[24] (Data S4), showing these effects are not allele-specific, but related to the regulatory function of FCA at *FLC*.

Quant-seq

For Quant-seq experiments, total RNA was isolated as for qPCR and further cleaned up using the Qiagen RNeasy miniprep kit (74106). Library preparation and sequencing was performed by Lexogen GmbH. RNA integrity was assessed on a Fragment Analyzer System using the DNF-471 RNA Kit (15 nt) (Agilent). Sequencing-ready libraries were generated from 100ng of input RNA using a QuantSeq 3' mRNA-Seq Library Prep Kit REV for Illumina (015UG009V0271) following standard procedures. Indexed libraries were assessed for quality on a Fragment Analyzer device (Agilent), using the HS-DNA assay and quantified using a Qubit dsDNA HS assay (Thermo Fisher). A sequencing-ready pool of indexed libraries were sequenced on an Illumina NextSeq 2000 with a 100-cycle cartridge using the Custom Sequencing Primer (CSP) at Lexogen GmbH.

Accession numbers

The genes mentioned in this article, in alphabetical order, have been given the following accession numbers by the Arabidopsis Information Resource (TAIR; https://www.arabidopsis.org) database: *ACT7* (AT5G09810), *CLF* (AT2G23380), *COOLAIR* (AT5G01675), *FCA* (AT4G16280), *FLC* (AT5G10140), *FLD* (AT3G10390), *FRI* (AT4G00650), *PP2A* (AT1G69960), *SWN* (AT4G02020), *UBC* (AT5G25760).

Bioinformatic analysis

Data analysis of RNAseq data was performed by Lexogen GmbH using their standard pipeline. Read Quality was assessed by FastQC[49] and trimmed for quality and adapter content with cutadapt[50] version 1.18. Clean reads were mapped to the Arabidopsis reference genome (TAIR10) with STAR[51] version 2.6.1a. Raw reads have been deposited on Sequence Read Archive (SRA) in SRA: PRJNA980626, and SRA: PRJNA1076161. To calculate proximal to distal count ratios, read ends were counted for each identified polyA cluster[12] (Data S1). Proximal polyadenylation clusters are defined as termination sites within the first Exon and Intron of *FLC* while distal clusters are in close proximity to the canonical polyadenylation site. For each biological sample, ratios were calculated from the summed proximal and distal cluster counts (Data S2).

Computational model summary

We built a stochastic model describing the dynamics of the activating and repressive histone modifications of interest, across a single *FLC* locus. The model includes activating (H3K4me1) and repressive (H3K27me3) modifications at each histone across the *FLC* locus, assumed mutually exclusive on the same H3 tail. The model also incorporates sense and antisense transcription at the locus. Importantly, the model assumes distinct roles for the two kinds of modifications in transcription regulation: H3K4me1 is assumed to affect RNA Pol II processivity rather than transcription initiation/transition to elongation. Higher overall coverage of H3K4me1 across the locus therefore promotes distal termination of transcription by reducing the probability of early termination. H3K4me1 modifications are added co-transcriptionally, allowing distal termination to feedback onto further H3K4me1 addition. Conversely, proximal termination of both sense and antisense transcription leads to removal of H3K4me1, in line with the potential function of FLD as a demethylase, allowing feedback onto further proximal termination. Higher H3K27me3 coverage, rather than affecting processivity, is assumed to decrease the frequency of both sense and antisense transcription events (representing either transcription initiation or transition to elongation). The model describes constitutive deposition of H3K27me by PRC2 in the *FLC* nucleation region. H3K27me spreads to the rest of the locus through interactions with the nucleation region, through PRC2 read-write feedback interactions. An experimentally determined gene loop mediating an interaction between the 5' and 3' ends of *FLC*[56,57] is also incorporated in the model. Finally, in line with experimental evidence from mammalian systems and previously developed models for PRC2 silencing,[1]

the model assumes that productive transcription promotes removal of H3K27me through both histone exchange and co-transcriptional demethylation, thereby implementing antagonism between transcription and Polycomb silencing. These aspects are integrated into a fully stochastic spatiotemporal simulation of the *FLC* locus, where transcription events (sense and antisense) and histone modification states (H3K4me1, unmodified, or H3K27me1/2/3 at each H3 histone across the locus) are simulation outputs. Full details of the model are below.

Model features

Activating and repressive histone modifications

The mathematical model describes the modification state of each H3 histone at the *FLC* locus, with the whole locus represented by an array of 30 nucleosomes, consistent with the approximately 6 kilobases extent of the locus (nucleosomal DNA + linker DNA spans approximately 200 bp). The modifications described are: (1) H3K4me1, an activating modification associated with gene bodies in Arabidopsis, and known to exhibit significant differences between active and silenced states of *FLC*; (2) the Polycomb-mediated repressive modification H3K27me3 as well as possible H3K27me2, H3K27me1 and me0 states. The model assumes mutual exclusivity of K4me and K27me on the same H3 tail in line with experimentally observed mutual exclusivity in mESCs[58–61]. Evidence from mammalian and yeast studies indicate that H3K27me and H3K4me are added non-processively[62,63]. Therefore, the model assumes stepwise addition of these modifications.

Sense and Antisense transcription

The model describes sense and antisense transcription events at the locus. The frequency of these events (assumed to represent transcription initiation or transition to elongation) is determined by the coverage of repressive H3K27me3 across the locus. This is consistent with Polycomb (PRC1) binding to H3K27me3 and controlling accessibility to gene regulatory elements by mediating compaction,[64] as well as other potential H3K27me3-mediated mechanisms, such as inhibition of RNA Pol II release from the Transcription Start Site (TSS),[65] thus preventing transitions into an actively transcribing Pol II state[66]. Our data indicates that both sense and antisense transcription at *FLC* are silenced by Polycomb – both types of transcripts are present at higher levels in the absence of Polycomb silencing, as observed in genotypes containing the *FLC* activator FRIGIDA, and in loss of function mutants of FCA and FLD.

Co-transcriptional addition of activating modifications

Activating modifications such as H3K4me and H3K36me in many contexts and across organisms are known to be added co-transcriptionally[67]. The association of the H3K4 methyltransferase complex with elongating RNA Pol II and the addition of this modification during transcriptional elongation has been demonstrated in *S. cerevisiae*[63]. In the case of Arabidopsis H3K4me1, a genome wide analysis of the localization patterns of H3K4me1, of RNA Pol II, and of the different H3K4 methyltransferases, indicates that the co-transcriptional addition of H3K4me1 over gene bodies is mediated mainly by the methyltransferase ATXR7[68]. Consistent with this role, an *ATXR7* mutant has been shown to partially suppress *FLC* activation by the activator FRIGIDA[69].

Alternative polyadenylation mediated by coverage of activating modifications

The model assumes that higher overall coverage of H3K4me1 across the locus promotes distal termination of transcription. This assumption is in line with a potential feedback mechanism where H3K4me1 enables binding of the H3K36me3 methyltransferase (SDG8) and consequent H3K36me3-mediated assembly of splicing factors[33] which facilitate co-transcriptional processing. Such a role for H3K4 methylation is also consistent with evidence in yeast[63] and mammalian cells,[46] indicating a role in facilitating productive transcriptional elongation.

Proximal termination mediated removal of activating modifications

The model assumes proximal termination of both sense and antisense mediates removal of H3K4me1. This is in line with the potential function of FLD as a demethylase,[31] and the fact that genetically, FLD relies on FCA-mediated proximal termination of transcription[14]. The link between FLD and proximal termination is further supported by the molecular and genetic analysis in the accompanying paper[16]. The removal of H3K4me1 resulting from proximal PA is assumed to take place uniformly across the whole locus, as the data we currently have does not allow us to observe the spatial details of this process.

Polycomb nucleation and spreading

The model describes nucleation of H3K27me by Polycomb in the *FLC* nucleation region, where Polycomb silencing is known to be nucleated during cold-induced chromatin silencing of *FLC*. Subsequent H3K27me3 spreading across the locus in this context is known to depend on the plant PRC2 methyltransferase CLF and on an active cell cycle, while nucleation does not depend on either. Experimental evidence from a genome-wide ChIP-seq study showing H3K27me3 patterns in a *clf* mutant[40] indicates that a similar mechanism is involved in the early developmental silencing of *FLC* – the *clf* mutant shows a clear H3K27me3 peak in the *FLC* nucleation region, while spreading is compromised. Consistent with this, in the model, H3K27me3 is assumed to spread to the rest of the locus through interactions with the nucleation region, through read-write feedback interactions. The spreading is assumed to rely on an active cell cycle – in the model, this is achieved by having the interactions with the nucleation region stronger in S/G2 phases.

Looping interactions mediating Polycomb activity

Experimental evidence indicates the existence of a gene loop mediating an interaction between the 5' and 3' ends of *FLC*,[56,57] with the level of interaction potentially depending the transcriptional state of the locus.[70] Significantly, this looping interaction is present in the silenced *FLC* state in the reference genotype Col-0, and consistently, a secondary peak of H3K27me3 is observed in published ChIP-seq data in *clf.*[40] Therefore, we capture this interaction in the model by allowing for stronger interactions with the nucleation region for a designated "looping region" within the gene body.

Productive transcription antagonises H3K27 methylation

Experimental evidence from mammalian systems indicates that productive transcription can promote removal of H3K27me through histone exchange. In addition, the H3K27me demethylases have been shown to physically interact with and co-localise on chromatin with elongating RNA Pol II, in *Drosophila*[71] and more specifically at the *FLC* locus.[72] Both features have been captured in previous models of Polycomb silencing,[1] and they are also incorporated in the current model. These features allow the rate of H3K27me turnover across the locus to be determined by the frequency of transcription events. Our model also incorporates alternative PA of both sense and antisense transcription, allowing proximal PA and termination of RNA Pol II. The model assumes that only distal PA and termination (i.e., productive transcription events) can remove H3K27me through the above mechanisms.

Mutual exclusivity of sense and antisense at the same locus

We previously found that distal antisense transcription and sense *FLC* transcription are mutually exclusive at individual *FLC* loci.[73] Consistent with this mutual exclusivity, we have observed an anticorrelation between the levels of distal antisense transcripts and sense transcripts in multiple contexts, including different chromatin states of the *FLC* locus.[36,48] While this evidence demonstrates a mutual exclusivity between full length (distal) sense and antisense transcription events, we currently do not have data to determine whether the mutual exclusivity applies to proximally terminated transcription events. However, for simplicity, the model assumes mutual exclusivity for both types of transcription events. This is realised by having a refractory period after a sense (antisense) transcription event, during which antisense (sense) initiation is not allowed to occur. The model assumes equal refractory periods for sense and antisense transcription, as well as for distal and proximal termination.

Inheritance of modified histones during DNA replication

In line with recent evidence indicating that only H3-H4 tetramers carrying repressive modifications are reliably inherited across DNA replication,[74] the model assumes that H3 pairs corresponding to an individual nucleosome are inherited with probability half to a daughter strand, only if at least one of the H3 carries K27 methylation.

Additional considerations for the model

Absence of a role for the Ler Mutator-like transposable element at FLC

The *FLC* allele in the L*er* accession is known to contain a *Mutator*-like transposable element (TE) in intron 1 that causes lower expression from this allele relative to *FLC* alleles in other accessions.[75] This lower expression results from RNAi-mediated chromatin silencing of the TE element.[76] However, experimental evidence indicates that the intermediate *FLC* expression levels in *fca* mutants are unlikely to be mediated by the TE. The TE has been shown to act *in cis*,[75] as expression of an *FLC* allele from the Columbia-0 (Col-0) accession is unaffected in F1 plants generated by a L*er* to Col-0 cross, while the L*er FLC* allele expression is attenuated by the TE. In our case, the intermediate expression in *fca* mutants could be observed for both the endogenous *FLC* (L*er FLC* containing TE) and the transgenic Venus tagged *FLC* (Col-0 *FLC* sequence), arguing against a role for the *cis*-acting TE mechanism in generating intermediate transcription states.

Role for H3K36me3

The above model only considers the histone modifications H3K4me1 and H3K27me and does not explicitly consider H3K36me3. As discussed in the text, the Arabidopsis H3K36 methyltransferase SDG8 is known to recognise H3K4me1, thus mechanistically linking the two modifications. Combined with the H3K36me3 requirement for effective co-transcriptional splicing and recruitment of splicing factors as reported in *S. cerevisiae*,[33] this suggests a mechanism where FLD-mediated removal of H3K4me1 may cause reduced H3K36 methylation, which in turn promotes further early transcription termination by disrupting co-transcriptional splicing. For simplicity, the model omits the H3K36me3 step and assumes that H3K4me1implicitly mediates this effect on termination.

Direct physical interaction between transcription-coupled repression and Polycomb

FCA has been reported to interact directly with the Arabidopsis PRC2 methyltransferase CLF *in vivo*.[77] However, we have found no evidence of this interaction in our FCA immunoprecipitation followed by mass spectrometry (IP-MS) data or other published CLF IP-MS data,[26,78] and hence this interaction is not further considered in our model or other analyses.

Model implementation

Locus description

The whole locus represented by an array of 30 nucleosomes (60 H3 histones), consistent with the approximately 6 kilobase extent of the locus (nucleosomal DNA + linker DNA spans approximately 200 bp).

Mutually exclusive H3 modifications

Each H3 histone can be in one of five modification mutually exclusive states: {-1,0,1,2,3} representing H3K4me1, H3 unmodified on K4 and K27, H3K27me1, me2, and me3 respectively. This means addition of K27me can take place only in the absence of H3K4me1 and vice versa. All transitions are assumed to be non-processive: only transitions involving addition/removal of a single methyl group are allowed. The modification state of the locus is represented by an array S containing the modification states of the 60 histones. Each entry $S_i$can take the values {-1,0,1,2,3}. The array S is used to store the current state of all 60 histones over the course of the simulation. Note that we do not describe H3K36me3 in the model, as the evidence indicates it is downstream of H3K4me1, as described above and in the main text.

Alternative termination of transcription

Each transcription event (sense or antisense) has a probability of resulting in proximally termination. This probability is assumed to be determined by both the current H3K4me1 fractional coverage (over the whole locus) and the FCA effectiveness parameter, as follows:

$$P(\text{Proximal}) = \delta_{Proximal} + (p_0 + p_{FCA})(1 - P_{K4me1})(1 - \delta_{Proximal} - \mu_{Distal}).$$

The probability of distal termination is given by:

$$P(\text{Distal}) = 1 - P(\text{Proximal}).$$

Here p_{FCA} represents FCA effectiveness parameter and p_0 captures possible FCA independent proximal termination. It is important to account for FCA independent proximal termination since proximal termination is observed even in the null *fca-1* mutant (Figure 2B). Therefore, we keep $p_0 > 0$, while constraining $(p_0 + p_{FCA})$ to be less than or equal to 1. The parameter p_{FCA} is assigned different values depending on the genotype being simulated: L*er*, *fca-3*, or *fca-1* (see Table S3).

P_{K4me1} represents the fractional coverage of H3K4me1, $\delta_{Proximal}$ is a residual proximal termination probability (some proximal termination occurs even when $P_{K4me1} = 1$), and μ_{Distal} is a residual probability for distal termination (some distal termination occurs even when $p_{FCA} = 1$ or $P_{K4me1} = 0$).

Transcriptional antagonism of H3K27me3

Two modes of H3K27me3 removal[1] are implemented in the model: (1) Histone exchange: during each distally terminated transcription event, iterating over the histones (entries of array S), for each histone, we generate a pseudo-random number from a uniform distribution (0,1) and set both entries corresponding to the current nucleosome to 0 if the generated number is less than $p_{ex}/2$ (see Table S4). (2) Co-transcriptional Demethylation: During each distally terminated transcription event, iterating over the entries of array *S*, for each histone, we generate a pseudo-random number from a uniform distribution (0,1) and reduce the methylation state by one if the current state is {1, 2, or 3} AND the generated number is less than p_{dem} (see Table S4).

Transcription mediated effects on H3K4 methylation

In the model, transcription affects H3K4me1 in two ways: (1) Co-transcriptional addition of H3K4me1: Each unmodified H3 histone has a fixed probability p_{K4me}^{Transc}, of being methylated during a productive (distally terminated) transcription event (not in case of proximal termination). (2) H3K4 demethylation by FLD: Each H3 histone carrying H3K4me1 has a fixed probability p_{dem}^{FLD} of losing this modification during a proximally terminated transcription event.

H3K4me1 dynamics

In addition to the transcription mediated effects on H3K4me1, we also assume background (transcription independent) addition and turnover. Including transcription independent addition of H3K4me1 allows for the transcription independent function of some Arabidopsis H3K4 methyltransferases as recently reported.[68] The background turnover rate (independent of proximal termination and FLD) of H3K4me1 is set so that overall turnover rates of H3K27me3 and H3K4me1 are in the ratio 3.128/0.959, as measured in HeLa cells by a SILAC technique.[79] The H3K4 methylation propensities (background and transcription mediated) are chosen to be significantly higher than the K27me3 propensities, consistent with faster H3K4 methylation seen in time course evidence from mammalian systems.

H3K27me3 dynamics

Following Berry et al.[1] the model implements both de novo and PRC2 read-write feedback mediated addition of H3K27me. Values for methylation rate constants are mostly chosen as in Berry et al.[1] The overall rate alone is adjusted from its value in Berry et al.[1] to allow for a changed interaction paradigm – instead of just nearest neighbour interactions, the current model allows long ranged interactions between nucleosomes across the whole locus, with any nucleosome being able to interact with any other within the locus at some rate (see below). To account for nucleation, spreading and the looping interaction, the nucleosome array representing the whole locus is divided into three subsets: nucleation region (NR, nucleosomes 1 to 3, i.e., histones 1 to 6), looping region (LR, nucleosomes 20 to 22, i.e., histones 39 to 44), and body region (Body, remaining nucleosomes). The choice of nucleosomes 19 to 21 for the looping region is consistent with the observed secondary nucleation peak observed around 4kb downstream of the *FLC* TSS in published data for the *clf* mutant.[40]

The methylation propensities at each histone *i* are computed as:

$$r_i^{me} = \beta\left(\delta_{S_i,0}\left(\gamma_{me0-1} + k_{me0-1}E_i\right) + \delta_{S_i,1}\left(\gamma_{me1-2} + k_{me1-2}E_i\right) + \delta_{S_i,2}\left(\gamma_{me2-3} + k_{me2-3}E_i\right)\right)$$

where E_i represents the PRC2 read-write feedback contribution.

For nucleation region histones, the PRC2 read-write feedback mediated contribution to the H3K27 methylation propensity is computed as:

$$E_i = \xi_{NR}^{NR}\sum_{j\in NR}\left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{LR}^{NR}\sum_{j\in LR}\left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{Body}^{NR}\sum_{j\in Body}\left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right),$$

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

where ρ_{me2} is the relative activation of PRC2 by H3K27me2 as compared to H3K27me3. For the looping region histones, the PRC2 read-write feedback mediated contribution to the H3K27 methylation propensity is computed as:

$$E_i = \xi_{NR}^{LR} \sum_{j \in NR} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{LR}^{LR} \sum_{j \in LR} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{Body}^{LR} \sum_{j \in Body} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right).$$

For the remaining gene body histones, the PRC2 read-write feedback mediated contribution to the H3K27 methylation propensity is computed as:

$$E_i = \xi_{NR}^{Body} \sum_{j \in NR} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{LR}^{Body} \sum_{j \in LR} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right) + \xi_{Body}^{Body} \sum_{j \in Body} \left(\rho_{me2}\delta_{S_j,2} + \delta_{S_j,3}\right).$$

Here the parameters ξ are chosen so that interactions between nucleosomes within the nucleation and looping region histones produce the highest contribution to feedback methylation, followed by the contribution of NR and LR nucleosomes to methylation of other gene body nucleosomes. This is consistent with the PRC2 enzymatic machinery localising mainly to the nucleation region. Also consistent with this, we assume lower contributions from gene body histones to the methylation of NR and LR nucleosomes, and the lowest contribution from interactions between gene body histones.

Hence, we set the above parameters as follows:

$$\xi_{NR}^{NR} = \xi_{LR}^{NR} = \xi_{NR}^{LR} = \xi_{LR}^{LR} > \xi_{NR}^{Body} = \xi_{LR}^{Body} > \xi_{Body}^{NR} = \xi_{Body}^{LR} > \xi_{Body}^{Body}.$$

Following,[1] a background removal of H3K27 methylation is included in the model, represented by the rate constant γ_{dem}. This is in addition to the transcription mediated removal of H3K4me described above.

Dependence on an active cell cycle

To implement the dependence of H3K27me3 spreading on an active cell cycle (as observed experimentally[10]), the total cell cycle duration is divided into two parts: G1 and S/G2/M, with replication of the locus occurring at the midpoint of S/G2/M. The duration of G1 is assumed to be approximately half the total duration of the cell-cycle. The four parameters corresponding to interactions between the NR/LR and the body region (ξ_{NR}^{Body}, ξ_{LR}^{Body}, ξ_{Body}^{NR}, ξ_{Body}^{LR}) are reduced by a factor of 10 during the part of the cell cycle corresponding to G1 phase.

H3K27me3 spreading mutant

For simulating the H3K27me3 spreading mutant (*clf-2*), we lower the read-write feedback contributions from the NR and LR nucleosomes to methylation of gene body nucleosomes (i.e., lower ξ_{NR}^{Body} and ξ_{LR}^{Body}), as well as the contribution from the gene body nucleosomes to methylation of the NR and LR nucleosomes (i.e., lower ξ_{Body}^{NR} and ξ_{Body}^{LR}). Since the loss of one of the PRC2 methyltransferases may also be expected to partially affect H3K27me3 addition within the nucleation region itself, we also slightly reduce the feedback contributions between nucleosomes inside the NR and LR regions (i.e., lower ξ_{NR}^{NR}, ξ_{LR}^{NR}, ξ_{NR}^{LR}, ξ_{LR}^{LR}). These parameter values are specified in Table S2.

De novo addition of H3K27me3

In addition to the differences in PRC2 feedback mediated contributions, the nucleation region (first three nucleosomes) is assumed to have a significantly higher *de novo* contribution to the H3K27 methylation propensity (γ_{me} parameters), consistent with evidence for *de novo* PRC2 targeting to this region during early developmental silencing of *FLC*.[40]

H3K27me3 fractional coverage measure

Our previous studies of *FLC* silencing indicate that nucleation region H3K27me3 alone can significantly repress transcription, with the body region H3K27me3 making a relatively smaller additional contribution to silencing. This finding is different to the assumptions of Berry et al.,[1] where both H3K27me2 and H3K27me3 could repress. Furthermore, nucleation of H3K27me3 in the cold is sufficient to switch the locus to a digitally OFF state, as seen by FLC Venus imaging in the roots.[80] Potentially different contributions of the nucleation and looping region H3K27me3 versus the body region H3K27me3 to the transcriptional silencing are captured by allowing a higher contribution from these NR and LR nucleosomes when computing the overall fractional coverage of H3K27me3:

$$P_{K27me3} = \eta P_{K27me3}^{NR+LR} + (1 - \eta) P_{K27me3}^{Body}$$

Here P_{K27me3}^{NR+LR} and P_{K27me3}^{Body} represent fractional coverage of H3K27me3 in the nucleation region/looping region and the rest of the locus, respectively, and $0 < \eta < 1$. The overall fractional coverage, P_{K27me3} computed above, is then used to determine the transcription propensity (see below).

Polycomb silencing of transcription

H3K27me3 coverage across the locus is assumed to determine the "transcription initiation" frequency. This frequency is assumed to be a piecewise linear function of H3K27me3 coverage with a saturation threshold (similar to the model in Berry et al.[1]). Note that this frequency can also represent the transition of initiated Pol II into the elongating phase. Nucleation region and looping region contributions to the fractional coverage are assumed to be higher than the rest of the locus (see above). H3K27me3 coverage also has the same effect on sense and antisense initiation frequency.

$$f = \alpha\left(f_{max} - \frac{P_{K27me3}}{P_T}\left(f_{max} - f_{min}\right)\right) \text{ for } P_{K27me3} < P_T$$

Please cite this article in press as: Menon et al., Proximal termination generates a transcriptional state that determines the rate of establishment of Polycomb silencing, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.05.014

and

$$f = \alpha f_{min} \text{ for } P_{K27me3} \geq P_T$$

We also allow for a slightly higher transcription frequency in the H3K27me3 nucleated state compared to a fully H3K27me3 spread state. This is consistent with our previously published data for cold-induced *FLC* silencing, where *FLC* transcriptional output is observed to reduce further during the post-cold transition to a spread state following nucleation.[81] To enable this behaviour, we set $\eta < P_T$ so that the full silencing threshold is not reached even with full coverage of H3K27me3 in the nucleation and looping regions (i.e., even with $P_{K27me3}^{NR+LR} = 1$).

Published evidence based on sequencing of nascent RNA as well as our own qPCR data[39] indicates lower frequency of antisense transcription compared to sense transcription. This difference between sense and antisense transcription is captured by using different values of α for sense and antisense initiation frequency. Note that H3K4me1 does not affect the initiation frequency.

Histone inheritance during DNA replication

For each individual simulation, the DNA replication times are fixed deterministically, assuming a cell cycle duration of 22h.[82] At each DNA replication timepoint, the Gillespie Stochastic Simulation Algorithm (SSA) is interrupted. We then iterate over the entries of the array S, where for each nucleosome at the locus carrying at least one H3K27me, its modification state is set to 0, i.e., the pair of entries corresponding to the nucleosome is set to 0, with a probability of 0.5. We assume here random inheritance of intact H3-H4 nucleosomes, with equal probability to each daughter strand. For each nucleosome, this is done by generating a pseudo-random number from a uniform distribution (0,1) and setting the corresponding histone modification states to 0 if the generated number is less than 0.5. In contrast, for each nucleosome carrying no H3K27me, the modification state is always set to zero, i.e., no inheritance of nucleosomes carrying only H3K4me1.

Simulations

Model simulations were carried out using the Gillespie exact SSA.[53] Simulations were coded in C++ and compiled using Clang/LLVM (version 14.0.0). The Mersenne Twister 19937 algorithm,[83] implemented in the GNU Scientific Library (GSL version 2.7), was used to generate pseudorandom numbers for the Gillespie SSA. The simulation code can be accessed at: https://doi.org/10.5281/zenodo.11186917

Initial conditions

In all cases – both for the transcription-coupled repression module in isolation and the full model, the initial state was set to full coverage of H3K4me1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of simulation output

Average levels of fractional coverage of histone modifications and average frequency of different transcription events and termination site usage (sense distal, antisense distal, sense proximal, antisense proximal) were computed by averaging over 500 individual trajectories. Individual trajectories were started at random (uniformly distributed) times in the cell cycle to avoid having all trajectories synchronised. To quantify switching to the digital OFF state (which in the model can correspond either to a Polycomb nucleated state or a spread state), we have to define what qualifies as an OFF state. In our analysis we classify an OFF state as any state for which:

$$P_{K27me3} > \chi P_T$$

where the parameter χ is chosen so that $\chi P_T < \eta$, and all the other parameters/quantities are as defined above. This condition on χ ensures that an H3K27me3 build-up in the nucleation region (even if this modification is absent across the rest of the locus), can be sufficient to qualify as a digitally OFF state. In all simulation results shown, χ is set to 0.25.

Comparing simulation output to experimental data

Comparison to timecourse ChIP-qPCR data

In Figure 4 we compare the model predicted levels of H3K27me3 and H3K4me1 to the normalised levels measured by ChIP-qPCR (averaged over primers across the locus). The model predicted levels consist of an average over multiple simulated copies (1000 copies in all cases). While the model only describes a set of *FLC* copies in dividing cells, the ChIP-qPCR measurement is performed on whole-seedling tissue. Therefore, we compare a model predicted population average to the experimental data. Note that we expect there to be heterogeneity in the modification levels between tissues, and differences in the dynamics between dividing and non-dividing cells. We have previously addressed some of these differences at *FLC* in detail using a population level Ordinary Differential Equation model of *FLC* copies switching between different states.[81,84] However, capturing these aspects is beyond the scope of the current study. We therefore make a simpler assumption: that at all timepoints and across genotypes, there exist two subpopulations of *FLC* copies (at the whole plant level) in addition to the population of copies represented in the model. We assume that one of these subpopulations consists entirely of active *FLC* copies (with high H3K4me1 coverage) with no possibility of PRC2 silencing, and the other subpopulation consists of PRC2 silenced copies (with high H3K27me3 coverage), with no possibility

of switching to an active state. We assume that these copies form a fixed fraction of the total *FLC* copies, fixed across timepoints and across genotypes (hence independent of FCA). Note that these subpopulations must also consist of dividing cells in order for their fractions to be maintained constant over time. We do not explicitly simulate these subpopulations. Instead, we account for their contribution to the averaged histone modification levels. The contribution from each copy in the active subpopulation to H3K4me1 is assumed to be the same as the average steady state level at non-Polycomb silenced copies in the *fca-1* (null mutant) model, and with no contribution to the H3K27me3 level. Similarly, the contribution from each copy in the silenced subpopulation to H3K27me3 is assumed to be the same as the average level at PRC2 silenced copies predicted by the L*er* model (these levels are similar in the L*er, fca-3, and fca-1* models). We assume that these silenced copies make no contribution to the H3K4me1 level. The sizes of these two subpopulations relative to the population of copies represented by the model form two model parameters that we change to fit the data. The fits presented in Figures 4C and 4D assume that the active subpopulation is 0.2 times the size of the modelled population, while the silenced subpopulation is 0.1 times the size of the modelled population.

Comparison to ChIP profile across FLC

In Figure 5E, we compare the model predicted H3K27me3 profiles at *FLC* in the L*er* parental genotype and the Polycomb spreading mutant *clf-2* to the ChIP-qPCR measured profiles in these genotypes. Note that the model predicted profile is an average across multiple simulated copies (1000 copies). To do this we used an approach previously developed[39] for comparing model predicted Pol II profiles to those measured by ChIP. In our case, this method generates a predicted H3K27me3 ChIP profile by combining the fully spatially resolved model predicted profile with the resolution limiting effects arising from the sonication generated fragment size distribution inherent in ChIP based measurements. Mathematically, this method involves a convolution of the full model predicted profile with the probability distribution of fragment sizes generated during the ChIP experiment. Here we use the same fragment size distribution estimated in Wu et al.[39] The code used to carry out this analysis is also provided at: https://doi.org/10.5281/zenodo.11186917

Comparison of predicted FLC *shutdown*

In Figures 5C and S5A, we compare the model predicted *FLC* transcriptional shutdown over time to changes in *FLC* unspliced (Figure 5C) and *FLC* spliced (Figure S5A) RNA levels, measured by qPCR at the whole plant level. The model predicted *FLC* transcriptional output is computed as the average frequency of *FLC* sense distal transcription events (transcription events per hour). This frequency is computed by averaging sense distal transcription events over a 6hr time window (an approximation based on our estimate of ~6hr *FLC* mRNA half-life) leading to the respective sampling timepoints (to match the experiments, we choose 10 day, 14 day, and 21 day timepoints). We note that the model is not designed to capture absolute RNA levels in the different genotypes, since this may depend on differences in their stability, as well as heterogeneity between tissue types which are not described in the model. What the model is intended to capture, is the quantitative trends, including differences between genotypes and changes over time, i.e. a reduction over time in some genotypes which may be disrupted in others. A whole-plant level measurement of *FLC* RNA is an appropriate quantity for capturing these types of differences, and we therefore use these measurements for comparison to the model predictions. Since the model does not capture absolute RNA levels, for comparison of experimental timecourse data to model predictions, we consider changes relative to the first experimental timepoint (i.e., fold changes).

Preventing proximal termination results in bistability

The model predicts that the locus is bistable in the absence of proximal-termination mediated repression, so that both active and Polycomb silenced states can be stably maintained (Figure S4Bii). This is consistent with the FCA mediated mechanism not having a role in maintaining the Polycomb silenced state, as evidenced by the stable maintenance of cold-induced Polycomb silencing at *FLC* in both *fca* mutants and genotypes with active FRIGIDA. Hence our model also explains how the *FLC* transcriptional activator FRIGIDA can set up an active transcriptional state, preventing Polycomb silencing by functioning as an anti-terminator. Anti-termination factors that prevent the use of early termination sites have been implicated in tuning transcriptional output from eukaryotic polymerases.[85] In the case of *FLC*, the transcriptional activator FRIGIDA has recently been shown to function at least partly by antagonising FCA mediated proximal cleavage and polyadenylation of *FLC* sense transcription at specific stages of embryo development.[12] The presence of active FRIGIDA promotes distal polyadenylation and prevents *FLC* silencing, further supporting the view that FCA-mediated proximal termination is essential for repression. Interestingly, FRIGIDA expression outside of a specific developmental stage fails to activate *FLC* expression[12]: this reinforces the view that once Polycomb epigenetic silencing has been established, reactivating transcription is difficult, with simple upregulation of an activator being insufficient to overcome robust digital silencing.