Quantitative imaging of RNA polymerase II activity in plants reveals the single-cell basis of tissue-wide transcriptional dynamics

Simon Alamos[1], **Armando Reimer**[2], **Krishna K. Niyogi**[1,3,4,*], **Hernan G. Garcia**[5,6,7,*]

[1]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California

[2]Biophysics Graduate Group, University of California Berkeley, Berkeley, California

[3]Howard Hughes Medical Institute, University of California Berkeley, Berkeley, California

[4]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California

[5]Department of Molecular and Cell Biology, University of CaliforniaBerkeley, Berkeley, California

[6]Department of Physics, University of California Berkeley, Berkeley, California

[7]Institute for Quantitative Biosciences-QB3, University of California Berkeley, Berkeley, CA, USA

Abstract

The responses of plants to their environment often hinge on the spatiotemporal dynamics of transcriptional regulation. While live-imaging tools have been used extensively to quantitatively capture rapid transcriptional dynamics in living animal cells, lack of implementation of these technologies in plants has limited concomitant quantitative studies. Here, we applied the PP7 and MS2 RNA-labeling technologies for the quantitative imaging of RNA polymerase II activity dynamics in single cells of living plants as they respond to experimental treatments. Using this technology, we count nascent RNA transcripts in real-time in *Nicotiana benthamiana* (tobacco) and *Arabidopsis thaliana* (Arabidopsis). Examination of heat shock reporters revealed that plant tissues respond to external signals by modulating the number of cells engaged in transcription rather than the transcription rate of active cells. This switch-like behavior, combined with cell-to-cell variability in transcription rate, results in mRNA production variability spanning three orders of magnitude. We determined that cellular heterogeneity stems mainly from the stochasticity intrinsic to individual alleles. Taken together, our results demonstrate that it is now possible to quantitatively study the dynamics of transcriptional programs in single cells of living plants.

Introduction

Plant growth and development depends on rapid and sensitive signaling networks that monitor environmental fluctuations and transduce this information into transcriptional changes that lead to physiological adaptation. Gene regulation in plants can be extremely

[*]**For correspondence:** hggarcia@berkeley.edu (HG); niyogi@berkeley.edu (KN).

fast, with changes in mRNA abundance detectable in seconds to minutes, for example in response to modulations in light intensity (Suzuki et al., 2015; Crisp et al., 2017), light quality (Leivar et al., 2009), the axis of gravity (Kimbrough et al., 2004), nutrient concentration (Krouk et al., 2010) or temperature (Zandalinas et al., 2020).

A first step toward understanding how plant transcriptional programs unfold in time and space is to quantify gene activity in individual living cells as they respond to external stimuli. Protein reporters have been used in plants to measure the dynamics of single-cell gene activity in live tissues over hours to days (Gould et al., 2018). However, fluorescent proteins mature at timescales that are long (>30 min) compared to the rates that characterize stress-responsive transcription (~1 min) (Kollist et al., 2019), particularly in organisms grown at moderate temperatures such as plants (Balleza et al., 2018). In addition, protein reporter signals convolve processes such as transcription, RNA processing, RNA transport, translation, and protein degradation, often making it challenging to precisely identify where and how regulatory control is being applied along the central dogma.

In the last few years, our understanding of transcriptional regulation in animals has been transformed by techniques that quantify transcriptional activity in single cells of living embryos (Ferraro et al., 2016; Lucas et al., 2013; Gregor et al., 2014; Garcia et al., 2020) and adult mice (Das et al., 2018). Here, nascent RNA is fluorescently labeled by tagging genes of interest with RNA aptamers such as MS2 or PP7 that recruit fluorescent proteins to transcriptional loci, revealing real-time transcriptional activity at the single-cell level. However, research into the equally diverse and important gene regulatory aspects of plant development and physiology has remained relatively isolated from these technological breakthroughs.

Here we bridged this technological gap by developing and implementing the PP7 and MS2 technologies for labeling nascent RNA in *Arabidopsis thaliana* (Arabidopsis) and *Nicotiana benthamiana* (tobacco). Through state-of-the-art quantitative imaging, we counted the absolute number of elongating RNA polymerase II (RNAP) molecules at individual genes and measured how this number is regulated dynamically in response to heat stress. We used this stress response in leaf tissue as a model to determine how tissue-level patterns of mRNA accumulation arise from the dynamical transcriptional behavior of individual cells. We uncovered previously unknown modes of gene regulation in plants by which tissues respond to external signals by modulating the fraction of cells engaged in transcription, but leave the single-cell transcription rate unchanged. Further, we determined how these regulatory layers give rise to a surprising level of cellular heterogeneity. The resolution afforded by PP7 and MS2 made it possible to characterize the sources of this cell-to-cell variability, revealing that stochastic processes intrinsic to individual alleles contribute to differences of three orders of magnitude in mRNA production between neighboring cells. Together, these results highlight the potential of live-imaging techniques for uncovering and quantitatively describing regulatory processes with spatiotemporal resolutions that cannot be achieved with methods such as traditional protein reporters or single-cell RNA sequencing. We envision that this approach will open new avenues of inquiry in plant cell and developmental biology.

Results

Establishment of the PP7 and MS2 systems for single-cell live imaging of transcription in plants

To quantitatively measure transcriptional dynamics in tobacco and Arabidopsis, we implemented an mRNA fluorescent-tagging approach previously used in animal cells in culture (Golding et al., 2005; Chubb et al., 2006; Darzacq et al., 2007; Larson et al., 2011), *D. melanogaster* embryos (Garcia et al., 2013; Lucas et al., 2013), the mouse brain (Park et al., 2014), and *Caenorhabditis elegans* (Lee et al., 2019) in which the gene of interest is tagged with tandem repeats of the PP7 DNA sequence that, when transcribed, form RNA stem-loops (Fig. 1A) (Chao et al., 2008; Larson et al., 2011). The PP7 loop RNA is bound by the PP7 bacteriophage coat protein (PCP) (Chao et al., 2008) expressed under a ubiquitous promoter. Fusing PCP to a fluorescent protein results in the fluorescent labeling of nascent RNA molecules. By virtue of the relatively slow movement of genomic loci in the nucleus and the accumulation of fluorophores in the diffraction-limited volume of the gene, sites of active transcription appear as bright fluorescent puncta over the background of nuclear PCP fluorescence in a laser-scanning confocal microscope. The fluorescence intensity of these spots reports on the number of RNAP molecules actively transcribing the gene at any given time (Garcia et al., 2013) and is proportional to the instantaneous rate of transcription (Lammers et al., 2020; Bothma et al., 2014).

To optimize this imaging strategy for plants, we generated two classes of constructs (Fig. 1B): (1) coat protein constructs that fuse PCP to a fluorescent protein such as GFP under a constitutive and ubiquitously expressed Arabidopsis promoter, and (2) reporter constructs that contain a neutral DNA sequence consisting of a firefly luciferase-*β*-glucoronidase fusion with 24 PP7 stem loop repeats inserted in the 5' end of this gene, under the control of the promoter of interest. To aid in the automated segmentation of nuclei, reporter constructs also contain a nuclear label consisting of the mScarlet red fluorescent protein (Bindels et al., 2016) fused to the Arabidopsis histone 2B coding region driven by a ubiquitous promoter (Federici et al., 2012). These two constructs confer resistance to different antibiotics, allowing sequential and combinatorial transformation into plants.

We tested this system in tobacco by simultaneously infiltrating leaves with two *Agrobacterium* strains, one strain carrying a PCP-GFP plasmid and a second strain carrying a reporter plasmid lacking a functional promoter, yielding homogeneous GFP nuclear and cytoplasmic fluorescence (Fig. 1C, top left). When the strong and constitutive 35S promoter was used to drive the reporter construct, nuclear GFP puncta became visible (Fig. 1C, top right). These results suggest that spots correspond to sites of active transcription and rule out potential PCP-GFP nuclear aggregation artifacts. Analogous results were obtained in stably transformed transgenic Arabidopsis plants (Fig. 1C, bottom).

We next sought to confirm that spot fluorescence constitutes a dynamical readout of transcriptional activity. To this end, we asked whether spot fluorescence dynamics in tobacco qualitatively recapitulate previous observations performed on the same promoters in Arabidopsis with orthogonal techniques. This comparison is made possible by the strong conservation of transcriptional regulation in plants (Wilhelmsson et al., 2017), in particular

the heat shock response (Mittler et al., 2012). We measured the transcriptional activity of two well-known constitutive and heat shock-inducible Arabidopsis genes (*GAPC2* and *HSP70*, respectively (Czechowski et al., 2005; Dong Yul Sung et al., 2001)) before and during a heat shock treatment. GAPC2-PP7 expression was detectable at 25 °C (Fig. 1D, top left, Movie S1). The presence of multiple spots per nucleus is likely due to multiple transgene transfer events; the number of spots did not change with treatment (Figure 1D, bottom left). Further, the fluorescence of one of these spots over time did not change upon heat shock (Fig. 1E), in accordance with the constitutive expression of *GAPC2* in Arabidopsis (Czechowski et al., 2005). Consistent with the heat shock inducibility of the *HSP70* gene in Arabidopsis (Dong Yul Sung et al., 2001), HSP70-PP7 transcription was hardly detectable at 25 °C in tobacco (Fig. 1D, top right). However, upon increasing the temperature to 39 °C, multiple fluorescent puncta rapidly appeared (Fig. 1D, bottom right, Movie S1), and their fluorescence increased with time (Fig. 1E). Thus, we conclude that the PP7 system reliably recapitulates previous qualitative knowledge of transcriptional dynamics in plants.

Simultaneously tagging multiple mRNA species or multiple locations of the same mRNA species with different fluorescent proteins has revealed regulatory and physical interactions between loci and uncovered the regulation of distinct steps of the transcription cycle in cells in culture and animals (Hocine et al., 2012; Coulon et al., 2014; Fukaya et al., 2016, 2017; Lim et al., 2018b,a). To enable such multiplexing in plants, we also implemented the MS2 system, which is analogous and orthogonal to the PP7 system. Here, MS2 loops are specifically recognized by an MCP coat protein (MCP) (Bertrand et al., 1998). We tested the MS2 system in tobacco and obtained results comparable to those obtained for PP7 (Fig. S1), allowing us to track the expression dynamics of two transgenes in a single cell (Fig. 1F).

Quantitative characterization of the PP7 system in Arabidopsis

To study transcriptional regulation at the single-cell level in populations of genetically identical leaf cells, we next generated stably transformed lines of Arabidopsis carrying PCP-GFP and a PP7 reporter construct driven by the promoter of the stress-inducible HSP101 gene. Stably transformed lines are preferable to the transient transformation in tabacco, because agroinfiltration inserts a variable number of transgenes randomly throughout the genome. A line carrying a single reporter locus (hereafter referred to as HSP101-PP7–1) was used for the following experiments unless stated otherwise; for details, see Materials and Methods: Generation of transgenic Arabidopsis lines.

A key step toward establishing PP7 as a reporter of single-cell transcriptional activity in Arabidopsis is to demonstrate that the observed spot fluorescence dynamics quantitatively recapitulate this activity. We therefore sought to cross-validate PP7 measurements with RT-qPCR quantifications of mRNA abundance in our stably transformed Arabidopsis plants. The *HSP101* mRNA is hardly detectable across vegetative tissues under standard growth conditions (Queitsch et al., 2000) and rapidly accumulates to high levels upon treatments inducing cytosolic protein misfolding such as heat shock (Charng et al., 2007). As previous experiments have shown that, upon induction, *HSP101* is expressed uniformly throughout plant tissues (Winter et al., 2007; Jean-Baptiste et al., 2018), we compared the average

transcriptional activity of a few hundred leaf cells obtained by microscopy with that of the whole plant in bulk reported by RT-qPCR.

As expected, we did not detect actively transcribing cells in HSP101-PP7–1 plants imaged for 1 h at room temperature (Fig. S2), but shifting the microscope stage from 22 °C to 39 °C resulted in the rapid appearance of transcription spots (Fig. 2A, Movie S1). To compare the instantaneous metric of transcriptional activity reported by spot fluorescence with the number of accumulated reporter mRNA molecules captured by RT-qPCR, we converted spot fluorescence to number of produced mRNA molecules by integrating the fluorescence of all spots in the field of view over time (Fig. S3; Garcia et al. (2013).

Controls for GFP photobleaching ruled out the possibility that we underestimated the produced mRNA calculated by microscopy (Fig. S4). Finally, we measured HSP1010 reporter mRNA abundance by RT-qPCR using whole plants treated with heat shock (see Materials and Methods: Heat shock treatments). These measurements were strongly correlated with each other (Fig. 2B), confirming that spot fluorescence directly reports on the rate of mRNA production. This conclusion held regardless of mRNA degradation rate (Fig. S5).

While our measurements so far have shown that PP7 fluorescence is *proportional* to the number of actively transcribing RNAP molecules, it does not, by itself, report on their *absolute number*. Expressing measurements in terms of absolute number of active RNAP molecules instead of arbitrary fluorescence units is necessary for directly comparing data across microscopy setups and laboratories, and for integration with other quantitative measurements and theoretical models (Rosenfeld et al., 2005; Cai et al., 2006; Garcia and Phillips, 2011; Garcia et al., 2013; Xu et al., 2015). In order to turn the PP7 system into such a precision tool, we calibrated its arbitrary fluorescence units to report on the number of RNAP molecules actively transcribing the reporter gene. We followed a recently established approach to measure the fluorescence of individual GFP molecules arranged in 60-meric nanocages *in vitro* (Hsia et al., 2016) and *in vivo* (Akamatsu et al., 2020). We fused GFP to a monomer that forms these 60-meric nanocages and expressed it in tobacco leaves (Fig. 2C) to obtain a distribution of fluorescence intensity values for the resulting GFP punctae (Fig. 2D, left).

Fusing two GFP molecules to each nanocage monomer yielded the fluorescence distribution of nanocages containing 120 GFP (Fig. 2D, left). A linear fit of the means of these distributions passing through the origin shows that the mean fluorescence of 120 GFP is almost exactly twice that of 60 GFP (Fig. 2D, right), confirming the validity of this approach. The slope of this fit is an estimate of the average number of arbitrary units of fluorescence corresponding to a single GFP molecule in our microscopy setup, making it possible to report PP7 measurements in absolute units.

Our absolute calibration also provided the opportunity to determine the limits of applicability of the PP7 technology. Specifically, there is a minimum number of actively transcribing RNAP molecules below which no reliable detection is possible. Figure 2E compares histograms of the calibrated number of RNAP molecules in the weakest detectable

spots and the corresponding fluctuations in background fluorescence in the data from Figure 1F. Consistent with previous measurements (Garcia et al., 2013; Lammers et al., 2020), these histograms overlap at approximately 3 RNAP molecules, marking the level at which PP7 fluorescent spots become undetectable. The average gene length in Arabidopsis is about 2 kbp (The Arabidopsis Genome Iniative, 2000) and the footprint of an elongating RNAP molecule is ≈ 35 bp (Tornaletti et al., 1999). As a result, an average gene can accommodate a maximum of 2 kbp/35 bp≈60 RNAP molecules, well above the minimum 3 RNAP molecules that constitute this detection limit. An alternative way to view this detection limit is to consider the minimum detectable rate of transcription initiation. Given an elongation rate of 1.5 kbp/min (Ardehali and Lis, 2009), an RNAP molecule takes 3 min to transcribe an average Arabidopsis gene. Thus, to ensure at least 3 RNAP molecules on the gene and signal detectability at any time point, transcription needs to initiate at a minimum rate of 1 RNAP/min.

Uncovering single-cell transcriptional responses to heat shock

While static snapshots of tissues have provided profound lessons about the spatial control of transcription in animals and plants alike (Birnbaum, 2018; Taylor-Teeples et al., 2011), these approaches have not revealed how single-cell transcriptional dynamics dictate the temporal modulation of gene expression patterns. We sought to bridge this gap between single-cell and tissue-wide transcriptional dynamics by tracking individual nuclei and measuring the fluorescence of their corresponding transcription spot over time. To expand our range of inquiry, we generated two additional reporter lines under the control of a second heat shock-inducible promoter (HsfA2-PP7, Movie S1) or of a constitutive promoter (EF-Tu-PP7, Movie S1). In order to simplify the experiment, we imaged diploid cells of hemizygous Arabidopsis derived from the first generation of single-insertion transgenic plants (i.e., T2 individuals) such that each nucleus contained at most one spot (See Materials and Methods: Microscopy setup and image acquisition).

A striking feature of the single-cell response is the existence of a reproducible fraction of nuclei that does not show detectable expression throughout the experiment in all three assayed promoters (Fig. 3A, Fig. S6). The presence of these transcriptionally refractory cells was surprising given that endogenous *HSP101* and *HsfA2* are strongly induced and are necessary to survive heat stress in a dose-dependent manner (Queitsch et al., 2000; Charng et al., 2007). Similarly, as a highly expressed constitutive gene, *EF-Tu* would also be expected to be transcribed in every cell. Yet, this constitutive gene also presents a substantial fraction of refractory cells (Fig. 3A, right). Such refractory cells have also been identified in live-imaging studies of the early development of the fruit fly (Garcia et al., 2013; Lammers et al., 2020; Berrocal et al., 2020) and in *in vitro* cultures of animal cells (Hafner et al., 2020).

To confirm that the presence of refractory cells was not an artifact of our construct or of the PP7 technology, we examined a transgenic plant containing a HSP101-GFP fusion driven by the *HSP101* promoter that fully complements the heat-susceptibility phenotype of a *hsp101* knockout (McLoughlin et al., 2016). Treatment of HSP101-GFP plants with the conditions used in our PP7 experiments revealed the presence of two types of cells: cells

whose fluorescence was close to that of untreated cells and highly induced cells (Figure S7). These low-fluorescence cells, which can be located right next to highly expressing ones, support the existence of transcriptionally refractory cells and the ability of the PP7 technology to detect them.

Within responsive nuclei, we also found substantial heterogeneity in the instantaneous number of actively transcribing RNAP molecules. For example, at any given time, not all responsive nuclei harbored fluorescent spots; the fraction of active nuclei is modulated in response to heat shock, but remains constant for the constitutive promoter (Fig. 3B). Interestingly, individual spots do not turn on synchronously and present periods of high transcriptional activity interspersed by periods of low to no activity (Fig. 3C). This single-cell behavior is consistent with the presence of transcriptional bursts, which have been identified across organisms and are believed to emerge from the intrinsically stochastic nature of the biochemical process of transcription (Nicolas et al., 2017). Interestingly, the only plant gene probed in such detail before (to our knowledge) lacked such bursts (Ietswaart et al., 2017).

Tissue-wide transcriptional dynamics arise from the switch-like regulation of the instantaneous fraction of transcribing cells

How do tissue-level patterns of mRNA arise from the transcriptional activities of individual cells? Such tissue-level control could be implemented in two possible ways (Ko, 1992; Walters et al., 1995; Blackwood and Kadonaga, 1998; Fiering et al., 2000). One strategy consists of modulating the single-cell rate of transcription in a graded fashion (Fig. 4A, top). Alternatively, transcriptional control could work like a switch, where the fraction of actively transcribing cells is modulated across the tissue (Fig. 4A, bottom). Several *Drosophila* enhancers invoke both strategies simultaneously (Garcia et al., 2013; Bothma et al., 2014; Lammers et al., 2020; Berrocal et al., 2020). Single time-point measurements in plants(Turco et al., 2019; Angel et al., 2011) and live-imaging studies in cell culture (Hafner et al., 2020) have also provided evidence for switch-like control.

We found that, as transcriptional induction ensues, the instantaneous fraction of cells actively transcribing increases (Fig. 3B). In addition, the level of transcription in active cells can also fluctuate (Fig. 3C). We therefore sought to determine the extent to which each regulatory strategy gives rise to tissue-wide control of the mean mRNA production rate. To this end, we expressed the total bulk transcriptional activity in terms of the quantitative contribution of each regulatory strategy as

$$\underbrace{\frac{\sum_i fluo_i(t)}{N_{total}}}_{\substack{\text{mean tissue} \\ \text{transcription rate}}} = \underbrace{\frac{\sum_i fluo_i(t)}{N_{active}}}_{\substack{\text{mean transcription rate} \\ \text{of active cells}}} \times \underbrace{\frac{N_{active}}{N_{tad}}}_{\substack{\text{instantaneous fraction} \\ \text{of active cells}}} \tag{1}$$

Here, $fluo_i(t)$ is the fluorescence of the i-th cell at time point t, N_{active} is the instantaneous number of active cells, and N_{total} is the total number of cells.

In order to determine how the resulting tissue-level transcriptional dynamics arises from the two contributions on the right side of Equation 1, we first determined the tissue-wide transcription rate at each time point by adding the fluorescence of all spots in each frame and then dividing by the total number of nuclei in the field of view (Eq. 1). The tissue-wide transcription rate of HSP101-PP7–1 and HsfA2-PP7–1 rose upon induction, while that of the constitutive EF-Tu-PP7 reporter remained constant throughout the experiment (Fig. 4B, left).

To determine whether the graded modulation of the transcription rate among active cells contributes to the mean tissue transcription rate, we calculated the mean spot fluorescence across actively transcribing cells. Further, to determine the contribution of the switch-like of regulation, we computed the instantaneous fraction of cells actively transcribing the reporter. Our calculations revealed that the temporal modulation of the transcription rate among active cells remained relatively constant throughout induction (Fig. 4B, middle). In contrast, the fraction of active nuclei was strongly modulated as a result of induction (Fig. 4B, right). Interestingly, these dynamics of the fraction of active cells were qualitatively comparable to the mean tissue transcription rate (compare Fig. 4B left and right).

To quantify the relative contribution of each of these regulatory strategies to the overall transcriptional dynamics, we measured the fold-change of each term in Equation 1. We defined this fold-change as the ratio between the value of each magnitude at peak induction (blue and green arrowheads in Fig. 4B) and at 10 min, shortly after the beginning of the response (grey arrowhead in Fig. 4B). For both heat-inducible promoters, the fold-change in the mean transcription rate across active cells was close to one (Fig. 4C). In contrast, the fold-change in the instantaneous fraction of active cells was almost identical to that of the total activity (Fig. 4C).

Thus, the duration of the treatment does not impact the rate of transcription of individual actively transcribing cells—when an individual cell transcribes, it tends to do so, on average, at a characteristic, relatively stable level regardless of induction time (Fig. S8). Instead, the time under stress modulates the tissue-wide transcription rate by increasing the probability that each individual cell engages in transcription.

Allele-specific regulation underlies most tissue-wide heterogeneity in mRNA production in living plants

Although physiological responses occur at the tissue level, each cell must bear the phenotypic consequences of its individual gene regulatory behavior in response to stress. Studies of microorganisms and mammalian cells in culture have revealed that single-cell transcriptional responses to outside stimuli are often highly variable, leading researchers to posit that organisms possess mechanisms to buffer this "noise" or to leverage variability to drive the adoption of cellular fates that, for example, provide resistance against environmental insults such as antibiotics Raj and van Oudenaarden (2008); Maheshri and O'Shea (2007); Eldar and Elowitz (2010). However, remarkably little is known about the level, functional roles, and underlying molecular mechanisms of transcriptional noise in shaping stress responses in multicellular systems like plants (Cortijo and Locke, 2020).

Although, on average, the rate of transcription of our heat-responsive reporters did not change with the duration of the heat treatment (Fig. 4C), at any given time point, the levels of activity across cells spanned more than two orders of magnitude (Fig. 5A). This behavior, combined with asynchronous activation (Fig. 3A) and the presence of cells that are transiently or permanently inactive transcriptionally (Fig. 4B,D), gives rise to a wide distribution in the predicted mRNA produced per cell (Figure 5B). This distribution spans more than three orders of magnitude, with a coefficient of variation (CV, standard deviation divided by the mean) of approximately 1.6.

What are the molecular sources of this cell-to-cell variability in the amount of mRNA produced (Figure 5C)? It could be the result of stochastic processes intrinsic to the allele-specific biochemical reactions that mediate transcription. In addition, a large fraction of this noise may be due to processes extrinsic to the allele itself, such as differing molecular compositions of neighboring cells. A previous measurement of gene expression noise in Arabidopsis using constitutively expressed fluorescent proteins found that extrinsic noise explains most of the cellular heterogeneity (Araújo et al., 2017). However, it is unclear how this noise in accumulated protein relates to transcriptional variability, and whether there are differences between constitutive and regulated promoters.

To determine whether variability is intrinsic or extrinsic to the allele, it is necessary to compare the expression of alleles belonging to the same cell with that of alleles in nearby cells (Elowitz et al., 2002). To make this possible, we imaged Arabidopsis individuals homozygous for the reporter, which display up to two fluorescent spots per nucleus in diploid cells (Fig. 5D, top, Movie S1). Four traces originating from two nuclei indicate that the transcriptional activity of alleles in the same nucleus can be much more similar to each other than the activity of alleles in different nuclei (Fig. 5D, bottom), suggesting a prominent role of extrinsic noise in transcriptional variability. However, our measurements also revealed that not all alleles in all nuclei are transcriptionally active: nuclei are approximately equally divided between populations presenting two, one, or even no transcription spots (Fig. 5E). The presence of nuclei with only one active allele suggests that the decision of alleles to become active is intrinsic to each allele. Thus, qualitatively, we have identified both potentially meaningful intrinsic and extrinsic contributions to the total transcriptional noise.

In order to determine the quantitative contribution of each source of variability to the single-cell distribution of mRNA produced, we followed Elowitz et al. (2002) (see Section S2 .1). To show that the results from this analysis do not depend on the number of transgene copies per insertion, we identified additional single insertion Arabidopsis lines for which we confirmed the presence of a single transgene copy per insertion locus using qPCR (see Figure S9 and associated calculations in Section S2 .2).

Figure 5F presents the integrated spot fluorescence of alleles pairs belonging to the same nucleus in homozygous plants of HSP101-PP7–1 and two additional lines with a single transgene copy per insertion. Our calculation of the noise components revealed that intrinsic sources explain most (~ 2/3) of the variability in all of the lines tested (Fig. 5G).

In contrast to investigations of cell-to-cell variability of protein expression in constitutive promoters (Araújo et al., 2017), our results demonstrate that most of the cellular heterogeneity in the transcriptional response to heat shock is not due to cells having a different chemical composition. Instead, stochastic processes at the level of each individual allele explain most of the cell-to-cell differences in the amount of mRNA produced per cell. Importantly, while here we have focused on the noise in the amount of produced mRNA, further insights can be drawn from examining the sources of molecular variability in, for example, instantaneous transcriptional activity (Fig. S10).

Discussion

Over the last few decades, it has become clear that the averaging resulting from bulk tissue sampling obscures important details about the spatial control of cellular processes in plants and animals alike. In plants, this limitation has motivated recent advances in single-cell RNA sequencing (McFaline-Figueroa et al., 2020). However, these measurements depend on the previous history of RNA transcription and degradation and thus obscure information about regulatory dynamics. Further, single-cell sequencing technologies tend to sacrifice spatial information. While enabling technologies to light up the process of transcription and its control in real time, in single cells or whole animals, have been developed (Munsky et al., 2012; Tutucci et al., 2018), plants have remained surprisingly sidelined.

Here, by implementing the PP7 and MS2 systems to fluorescently label nascent RNA molecules in plants, we have shown, to our knowledge for the first time, that it is possible to count the number of RNAP molecules actively transcribing individual alleles in single living cells of tobacco and Arabidopsis as they respond to their environment. This technical advance yielded unprecedented access to the temporal history of activity of individual alleles, making it possible to uncover distinct modes by which single-cell transcriptional activity in plants leads to tissue-wide gene expression dynamics.

Using this technique, and consistent with similar observations in other systems (Garcia et al., 2013; Lammers et al., 2020; Hafner et al., 2020), we discovered a fraction of transcriptionally refractory cells that do not transcribe regardless of induction conditions (Fig. 4 D). Single-molecule RNA FISH experiments in Arabidopsis roots found that at any given time $\approx$ 20% of cells are transcriptionally inactive for the constitutively expressed *PP2C* gene (Duncan et al., 2016). However, unlike the live-imaging approach developed here, single-molecule RNA FISH relies on fixed samples; it cannot determine whether this inactive state was transient or stable.

We also found that tissue-wide transcriptional induction dynamics are the result of the temporal modulation in the fraction of cells that switch to a transcriptionally active state, and not of the graded control of the transcription rate of active cells (Fig. 4C). This form of regulation has been hypothesized to be at play in the regulation of the FLC gene in response to temperature (Angel et al., 2011) and in the commitment to xylem cell fate in response to the VND7 transcription factor (Turco et al., 2019). Using our technologies, it should now be possible to directly test these models.

Gene expression can vary significantly from cell to cell in microbial and animal species (Raj and van Oudenaarden, 2008). By making it possible to measure cell-to-cell transcriptional variability in real time in living plant cells, we confirmed that plants are no exception to this widespread presence of transcriptional variability. The single-locus resolution of our method allowed us to determine that cell-to-cell variability in mRNA production arises mainly from stochastic processes instrinsic to each allele (Fig. 4 G). Studies in *in-vitro* cell cultures have found that gene-expression noise can have profound consequences for cellular survival (El Meouche et al., 2016; Shaffer et al., 2017); however, the role of transcriptional noise in plant stress responses remains an open question (Cortijo and Locke, 2020; Roeder, 2018). We envision that the strategy applied here to systematically dissect transcriptional heterogeneity in Arabidopsis and tobacco will shed light on this interplay between transcriptional variability and stress response. Further, it will be interesting to examine how some unusual aspects of plant cell biology and genetics can buffer transcriptional noise. For example, cytoplasmic connections could play a role in short-range sharing of gene products (Faulkner, 2018), averaging out extrinsic noise; multiple genome copies per nucleus in mature plant cells may provide further opportunities to average out intrinsic noise across alleles (Lee et al., 2019). Similarly, we speculate that the conspicuous retention of large numbers of seemingly redundant gene paralogs in plants may also help buffer intrinsic fluctuations in individual genes (Li et al., 2015).

Our approach requires access to a confocal microscope and to transgenesis tools, and should therefore be relatively easy to apply to many biological problems in plant development and physiology. However, imaging deep into tissues with the resolution necessary to resolve diffraction-limited spots remains a challenge, particularly in plants. Advances such as multiphoton imaging, lattice light-sheet microscopy, and adaptive optics will overcome this limitation (Liu et al., 2018).

Lacking single-polymerase resolution currently limits the applicability of MS2 and PP7 to genes transcribed at relatively high rates. A transcription initiation rate of 1 RNAP/ min, corresponding to our detection limit of 3 elongating RNAP molecules on an average Arabidopsis gene, could be sufficient to sustain slow transcriptional processes operating at long developmental timescales. For example, the FLC gene, a key seasonal developmental regulator in Arabidopsis is rarely occupied by more than one elongating RNAP at a time (Ietswaart et al., 2017) which may explain why previous attempts at visualizing nascent FLC mRNAs in live Arabidopsis plants have failed (Wu et al., 2016). A growing interest in live imaging of transcription combined with advances in fluorophore chemistry (Iwatate et al., 2020) as well as in the PP7 and MS2 technologies themselves (Wu et al., 2012) offer hope for breaking this detection threshold.

It will undoubtedly be of interest to correlate the activities of genes by visualizing their transcription simultaneously. This multiplexing is already possible for two genes using MS2 and PP7. A third color could be added by implementing interlaced MS2 and PP7 loops (Hocine et al., 2012). To further extend the palette, it should be possible to engineer other orthogonal RNA-binding proteins-RNA aptamer pairs (Daigle and Ellenberg, 2007; Katz et al., 2018).

Finally, and more generally, the random integration of transgenes in plants makes it challenging to dissect the role(s) of regulatory sequences at their endogenous genomic locations. Delivery of DNA with CRISPR/Cas9 or site-specific recombinases promises to unleash the potential of quantitative reporters of gene expression.

In this study, we focused on a simple step in the plant's use of temperature as a signaling input. More complex treatments have been previously used to show that plants can mount specific responses to inputs, such as memory in response to pulses of heat shock (Charng et al., 2007) and nonlinear integration of combinations of high light and temperature stress (Zandalinas et al., 2020). By administering experimental treatments while simultaneously measuring their effects on gene regulation, it will be possible to determine how these operations are performed at the cellular level. In addition, the sub-nuclear resolution of nascent RNA tagging could make it possible to resolve long-standing issues in plant signaling, such as the role of protein aggregates or "nuclear speckles" that are pervasive in light-responsive signaling pathways in plants (Ronald and Davis, 2019).

In conclusion, by enabling the measurement of transcription at high spatiotemporal resolution, the PP7 and MS2 methods introduced here close a critical technological gap in plant biology. These new technologies open new avenues of inquiry and will make it possible to quantitatively interrogate transcriptional control in living plants and to engage in the discourse between theory and experiment that has characterized the study of gene regulation in single cells and animal tissues over the last two decades.

Materials and Methods

Plasmids and Agrobacterium strains

All plasmid sequences used in this study can be accessed from a public Benchling folder. Plasmids will be made available at Addgene. All vectors were based on pCambia derivatives (Hajdukiewicz et al., 1994) and transformed into the GV3101::pMP90 Agrobacterium strain by electroporation. Plasmids confering Kanamycin resistance in plants (i.e reporter constructs) were based on pCambia2300. Plasmids confering Hygromycin resistance in plants (i.e PCP, MCP and nanocages constructs) were based on pCambia1300. A list of the plasmids used in this study can be found in table S1. The Arabidopsis gene identifiers associated with genomic sequences used in these plasmids are listed in table S3.

Plant growth conditions

Nicotiana benthamiana (tobacco) plants were grown in a greenhouse under natural light conditions prior to agroinfiltration. Following infiltration, tobacco plants were kept under 30 μE of constant light. Arabidopsis plants used for experiments were grown in 1/2 strength MS agar containing 50 μg/μl of Kanamycin under short day conditions (8 hours of 30 μE light per day) for four to six weeks prior to imaging.

Agroinfiltration

Agrobacterium glycerol stocks were streaked on LB plates containing 50 μg/μl Kanamycin and 50 μg/μl Gentamycin. Fresh colonies were grown overnight in liquid LB containing

the same antibiotic concentrations, spun down and resuspended in an equal volume of infiltration buffer (10 mM MES pH5.6, 10 mM $MgCl_2$, 150 μM Acetosyringone). Cells were incubated for 2–4 hours in infiltration buffer shaking at room temperature after which the cultures were diluted 1:3 to an OD_{600} of approximately 0.3. In experiments that required combining strains, coat protein and reporter strains were mixed in a 3:1 ratio (the exact ratio does not qualitatively affect the results). In PP7 and MS2 experiments, infiltrated leaves were imaged approximately 2 days after infiltration. For absolute calibration experiments, plants were imaged 12–18 hours after infiltration.

Generation of transgenic Arabidopsis lines

To generate lines carrying both PCP-GFP and PP7 reporters we followed a sequential transformation approach. We first selected PCP-GFP lines in 35 μg/ml of Hygromycin and kept lines exhibiting moderate levels of fluorescence and no obvious growth phenotype. Next, we transformed T1 or T2 PCP-GFP individuals with PP7 reporter Agrobacterium strains and selected transformants in 50 μg/ml Kanamycin and 35 μg/ml Hygromycin. Individuals T1 for the PP7 construct were screened for nuclear mScarlet fluorescence and presence of transcription spots matching previous knowledge about the activity of the corresponding endogenous gene. In all cases, to select for antibiotic resistance we followed the protocol by Harrison et al. (2006). A list of the lines used in this study can be found in table S2.

Determining the number of unlinked reporter transgene insertions

To select lines carrying a single insertion reporter locus we plated approximately 60 T2 seeds in MS plates containing Kanamycin and counted the ratio of survivors. This ratio was divided by the survival ratio in plates containing no antibiotics. A x^2 test was used to determine whether the product of these two ratios was statistically different from the expected ratio of 3/4. To confirm the absence of two or more unlinked reporter loci we examined transcription spots in guard cells. Unlike other leaf cell types, these cells are exclusively diploid (Melaragno et al., 1993) and therefore the presence of a single spot per guard cell nucleus in a T1 individual confirms the absence of unlinked insertions.

Heat shock treatments

To control the sample temperature in the microscope stage we used an OkoLabs H101-LG temperature chamber calibrated to achieve a maximum of $\approx 39°C$. The temperature experienced by the sample was determined once using an electronic probe. The heat shock treatment used for the RTqPCR experiment in Figure 2A was performed as follows: whole 4–6 week-old plants were placed in 1.7 ml plastic tubes containing 200 μl of water. The sample corresponding to time = 0 minutes was immediately taken out of the tube, quickly tapped dry, transferred to a new tube containing silica beads and frozen in liquid nitrogen. The rest of the samples were transferred to a 39°C heat block and removed at set times. Plants were then quickly tapped dry and frozen in liquid nitrogen.

Microscopy setup and image acquisition

In tobacco experiments, a piece of infiltrated leaf spot was mounted in water between a glass slide and a glass coverslip with the abaxial (bottom) side facing the objective. To image the GFP nanocages in mesophyll cells, the abaxial epidermis was first removed. In Arabidopsis experiments, full 2–4 day old leaves from 4–6 week old plants were mounted in tap water between a gas permeable cellophane membrane (Lumox film; Starstedt) and and a glass coverslip with the adaxial (top) side facing the objective. All samples were imaged close to the base of the leaf blade immediately after mounting. All data was taken in a Leica SP8 confocal microscope with a white light laser using a 63X oil objective. The dimensions of the field of view were 92.26×46.09 μm using 1052×512 pixels, resulting in a pixel size of 90 nm. Z stacks consisting of 25 slices of 0.5 μm each were taken every 60 seconds accumulating fluorescence 3 times over lines. The beginning of each stack was set to the upper-most nucleus in the leaf epidermis. For GFP, excitation 488 nm and emission 498559 nm. For mScarlet, excitation 569 nm, emission 579–630 nm. For Chlorophyll, excitation 488 nm, emission 665–675 nm. To ensure quantitative consistency across experiments, the 488 nm laser power at was calibrated to 10.5 μW ($\approx$ 5% laser power) at the beginning of each imaging session using a power meter. The percentage intensity of the 569 nm laser line was kept consistent across experiments at 5%.

RT-qPCR

Total RNA was extracted using the Quiagen RNeasy kit following the manufacturer instructions. Reverse transcription was performed using the Qiagen Omniscript kit with a primer mix of random 10mers (10 μM final concentration) and 15mer oligo dT primers (1 μM final concentration). mRNA abundance was calculated by the delta CT method. Primers for endogenous HSP101 were 5'GGTCGATGGATGCAGCTAAT and 5'CTTCAAGCGTTGTAGCACCA from Yoshida et al. (2011). Primers for the Actin2 standard were 5'CGCTCTTTCTTTCCAAGCTCAT and 5'GCAAATCCAGCCTTCACCAT from Liu and Ma (2011). Primers for the reporter mRNA were 5'GGGTTCATCAGAGTGCCAGAG and 5'AGGCAGAGCGACACCTTTAG. A negative control was performed under identical conditions replacing the RT enzyme with water.

Image analysis: spot fluorescence and tracking

Raw image stacks of the coat protein channel were used to identify fluorescent punctae corresponding to transcription spots using the ImageJ implementation of the 3D Trainable Weka Segmentation toolbox (Arganda-Carreras et al., 2017). Following Lammers et al. (2020), after segmentation, spots in each z-slice were fitted to a 2D Gaussian. The z-slice with the largest Gaussian amplitude was selected for the spot fluorescence calculation. Spot fluorescence corresponds to the sum of pixel intensity values in a circle with a radius of 1.08 μm centered around the center of the fitted Gaussian minus the background fluorescence offset. Per pixel background fluorescence is calculated from the baseline of the spot Gaussian fit. The imaging error associated with each spot trace is determined from the fluctuations over time in the per pixel offset multiplied by the spot integration area. A spline

is fitted to the offset time trace to calculate offset error as the standard deviation around this spline. False negative and false positive spots were corrected manually.

Image analysis: nuclear segmentation and spot tracking

Maximum intensity projections of the nuclear marker channel were used for nuclear segmentation using the ImageJ implementation of the 2D Trainable Weka Segmentation toolbox (Arganda-Carreras et al., 2017) or a custom-written Matlab pipeline. False negative and false positive nuclei were then manually corrected. Spots were assigned to nuclei based on physical overlap. Tracking of spots over time was based on nuclear tracking and manually corrected whenever errors were found.

Image analysis: nucleus fluorescence

A binary mask of segmented nuclei was applied to the PCP-GFP or Histone 2B- mScarlet channel. For each z-slice and frame, the mean fluorescence across pixels within each nucleus area was calculated. Then, for each frame, we took the intensity of the z-slice with the maximum mean fluorescence as the metric for fluorescent protein concentration in that given frame.

Determining transgene copy number by qPCR

Genomic DNA was extracted from leaf tissue using CTAB and phenol:chlorophorm precipitation. Primers used to amplify the reporter transgene were 5'gacgcaagaaaaatcagagagatcc and 5'ggtttctacaggacggaccatacac. Primers used to amplify a region near the *Lhcb3* gene used as an internal genomic control were 5'acaggtttggtcaagtcaattacga and 5'atggtttccatgaatactgaacacg. The final concentration of genomic DNA per reaction was 0.75 ng. For a more detailed explanation of the calculations and controls related to this experiment see Section S2 .2.

Absolute calibration using nanocages

Tobacco leaves were infiltrated with agrobacterium strains containing plasmids where the promoter of the Arabidopsis UBC1 gene (1138bp upstream of the AT1G14400 start codon) was used to drive the 60mer monomer fused to either one or two mGFP5 coding sequences. The N terminus of the rabbit Cytochrome P450 CII1 was added as an N terminal tag to target the protein fusions to the cytosolic side of the ER. Mesophyll cells were imaged no later than 15 hours after infiltration since longer incubation resulted in the appearance of large GFP aggregates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Akamatsu M, Vasan R, Serwas D, Ferrin M, Rangamani P, and Drubin DG (2020). Principles of selforganization and load adaptation by the actin cytoskeleton during clathrin-mediated endocytosis. eLife, 9:1–40.

Angel A, Song J, Dean C, and Howard M. (2011). A Polycomb-based switch underlying quantitative epigenetic memory. Nature, 476(7358):105–109. [PubMed: 21785438]

Araújo IS, Pietsch JM, Keizer EM, Greese B, Balkunde R, Fleck C, and Hülskamp M. (2017). Stochastic gene expression in Arabidopsis thaliana . Nature Communications, 8(1):2132.

Ardehali MB and Lis JT (2009). Tracking rates of transcription and splicing in vivo. Nature Structural and Molecular Biology, 16(11):1123–1124.

Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, and Seung HS (2017). Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. Bioinformatics, 33(15):2424–2426. [PubMed: 28369169]

Balleza E, Kim JM, and Cluzel P. (2018). Systematic characterization of maturation time of fluorescent proteins in living cells. Nat Methods, 15(1):47–51. [PubMed: 29320486]

Berrocal A, Lammers N, Garcia HG, and Eisen MB (2020). Kinetic sculpting of the seven stripes of the drosophila even-skipped gene. bioRxiv, page 335901.

Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, and Long RM (1998). Localization of ash1 mrna particles in living yeast. Mol Cell, 2(4):437–45. [PubMed: 9809065]

Bindels DS, Haarbosch L, Weeren LV, Postma M, Wiese KE, Mastop M, Aumonier S, Gotthard G, Royant A, Hink MA, and Jr TWJG (2016). mScarlet : a bright monomeric red fluorescent protein for cellular imaging. (november).

Birnbaum KD (2018). Power in Numbers: Single-Cell RNA-Seq Strategies to Dissect Complex Tissues. Annual Review of Genetics, 52(1):203–221.

Blackwood EM and Kadonaga JT (1998). Going the distance: a current view of enhancer action. Science, 281(5373):60–3. [PubMed: 9679020]

Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, and Levine M. (2014). Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. Proceedings of the National Academy of Sciences, 111(29):10598–10603.

Cai L, Friedman N, and Xie XS (2006). Stochastic protein expression in individual cells at the single molecule level. Nature, 440(7082):358–62. [PubMed: 16541077]

Chao JA, Patskovsky Y, Almo SC, and Singer RH (2008). Structural basis for the coevolution of a viral RNA-protein complex. Nature Structural and Molecular Biology, 15(1):103–105.

Charng Y. y., Liu H. c., Liu N. y., Chi W. t., Wang C. n., Chang S. h., and Wang T. t. (2007). A Heat-Inducible Transcription Factor , HsfA2 , Is Required for Extension of Acquired Thermotolerance. 143(January):251–262.

Chubb JR, Trcek T, Shenoy SM, and Singer RH (2006). Transcriptional Pulsing of a Developmental Gene. Current Biology, 16(10):1018–1025. [PubMed: 16713960]

Cortijo S. and Locke JC (2020). Does Gene Expression Noise Play a Functional Role in Plants? Trends in Plant Science, pages 1–11. [PubMed: 31679993]

Coulon A, Ferguson ML, de Turris V, Palangat M, Chow CC, and Larson DR (2014). Kinetic competition during the transcription cycle results in stochastic rna processing. Elife, 3.

Crisp PA, Ganguly DR, Smith AB, Murray KD, Estavillo GM, Ford E, Bogdanovi O, Lister R, Borevitz JO, and Eichten SR (2017). Rapid Recovery Gene Downregulation During Excess-light Stress and Recovery in Arabidopsis.

Czechowski T, Stitt M, Altmann T, Udvardi MK, and Scheible W-R (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. Plant Physiology, 139(September):5–17. [PubMed: 16166256]

Daigle N. and Ellenberg J. (2007). AN-GFP: An RNA reporter system for live-cell imaging. Nature Methods, 4(8):633–636. [PubMed: 17603490]

Darzacq X, Shav-tal Y, Turris VD, Brody Y, Shenoy SM, Phair RD, and Singer RH (2007). In vivo dynamics of RNA polymerase II transcription. 14(9):796–806.

Das S, Moon HC, Singer RH, and Park HY (2018). A transgenic mouse for imaging activity-dependent dynamics of endogenous arc mRNA in live neurons. Science Advances, 4(6):1–14.

Dong Yul Sung Vierling, E., and Guy CL (2001). Comprehensive expression profile analysis of the Arabidopsis hsp70 gene family. Plant Physiology, 126(2):789–800. [PubMed: 11402207]

Duncan S, Olsson TSG, Hartley M, Dean C, and Rosa S. (2016). A method for detecting single mRNA molecules in Arabidopsis thaliana. pages 1–10.

El Meouche I, Siu Y, and Dunlop MJ (2016). Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells. Scientiic Reports, 6(January):1–9.

Eldar A. and Elowitz MB (2010). Functional roles for noise in genetic circuits. Nature, 467(7312):167–73. [PubMed: 20829787]

Elowitz MB, Levine AJ, Siggia ED, and Swain PS (2002). Stochastic gene expression in a single cell. Science, 297(5584):1183–1186. [PubMed: 12183631]

Faulkner C. (2018). Plasmodesmata and the symplast. Current Biology, 28(24):R1374–R1378. [PubMed: 30562524]

Federici F, Dupuy L, Laplaze L, Heisler M, and Haseloff J. (2012). Integrated genetic and computation methods for in planta cytometry. Nature methods, 9(5):483–5. [PubMed: 22466793]

Ferraro T, Lucas T, Clemot M, De Las Heras Chanes J, Desponds J, Coppey M, Walczak AM, and Dostatni N. (2016). New methods to image transcription in living fly embryos: the insights so far, and the prospects. Wiley Interdiscip Rev Dev Biol, 5(3):296–310. [PubMed: 26894441]

Fiering S, Whitelaw E, and Martin DI (2000). To be or not to be active: The stochastic nature of enhancer action. BioEssays, 22(4):381–387. [PubMed: 10723035]

Fu AQ and Pachter L. (2016). Estimating intrinsic and extrinsic noise from single-cell gene expression measurements. Statistical Applications in Genetics and Molecular Biology, 15(6):447–471. [PubMed: 27875323]

Fukaya T, Lim B, and Levine M. (2016). Enhancer control of transcriptional bursting. Cell, 166(2):358–368. [PubMed: 27293191]

Fukaya T, Lim B, and Levine M. (2017). Rapid rates of pol ii elongation in the drosophila embryo. Curr Biol, 27(9):1387–1391. [PubMed: 28457866]

Garcia H, Tikhonov M, Lin A, and Gregor T. (2013). Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning. Current Biology, 23(21):2140–2145. [PubMed: 24139738]

Garcia HG, Berrocal A, Kim YJ, Martini G, and Zhao J. (2020). Lighting up the central dogma for predictive developmental biology. Curr Top Dev Biol, 137:1–35. [PubMed: 32143740]

Garcia HG and Phillips R. (2011). Quantitative dissection of the simple repression input-output function. Proc Natl Acad Sci U S A, 108(29):12173–8. [PubMed: 21730194]

Golding I, Paulsson J, Zawilski SM, and Cox EC (2005). Real-time kinetics of gene activity in individual bacteria. Cell, 123(6):1025–1036. [PubMed: 16360033]

Gould PD, Domijan M, Greenwood M, Tokuda IT, Rees H, Kozma-Bognar L, Hall AJ, and Locke JC (2018). Coordination of robust single cell rhythms in the Arabidopsis circadian clock via spatial waves of gene expression. eLife, 7:1–20.

Gregor T, Garcia H, and Little SC (2014). The embryo as a laboratory: quantifying transcription in Drosophila. Trends in genetics : TIG, 30(8):1–12. [PubMed: 24290192]

Hafner A, Reyes J, Stewart-Ornstein J, Tsabar M, Jambhekar A, and Lahav G. (2020). Quantifying the Central Dogma in the p53 Pathway in Live Single Cells. Cell Systems, pages 1–11.

Hajdukiewicz P, Svab Z, and Maliga P. (1994). The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Molecular Biology, 25(6):989–994. [PubMed: 7919218]

Harrison SJ, Mott EK, Parsley K, Aspinall S, Gray JC, and Cottage A. (2006). A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. Plant Methods, 2(1):1–7. [PubMed: 16401339]

Hilfinger A. and Paulsson J. (2011). Separating intrinsic from extrinsic fluctuations in dynamic biological systems. Proceedings of the National Academy of Sciences of the United States of America, 108(29):12167–12172. [PubMed: 21730172]

Hocine S, Raymond P, Zenklusen D, Chao JA, and Singer RH (2012). Single-molecule analysis of gene expression using two-color rna labeling in live yeast. Nat Methods.

Hsia Y, Bale JB, Xu C, Huang P. s., Ravichandran R, Yi S, Davis TN, Gonen T, and Neil P. (2016). Design of a hyperstable 60-subunit protein icosahedron. Nature, pages 1–12.

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 Systems, 4(6):622–635.e9. [PubMed: 28624615]

Iwatate R, Yoshinari A, Yagi N, Grzybowski M, Ogasawara H, Kamiya M, Komatsu T, Taki M, Yamaguschi S, Frommer WB, and Nakamura M. (2020). Covalent Self-labeling of Tagged Proteins with Chemical Fluorescent Dyes in BY-2 Cells and Arabidopsis Seedlings. The Plant Cell, page tpc.00439.2020.

Jean-Baptiste K, McFaline-Figueroa JL, Alexandre CM, Dorrity MW, Saunders L, Bubb KL, Trapnell C, Fields S, Queitsch C, and Cuperus JT (2018). Developmental and conditional dynamics of gene expression in single root cells of A. thaliana. bioRxiv, page 448514.

Katz N, Cohen R, Solomon O, Kaufmann B, Atar O, Yakhini Z, Goldberg S, and Amit R. (2018). An in Vivo Binding Assay for RNA-Binding Proteins Based on Repression of a Reporter Gene. ACS Synthetic Biology, 7(12):2765–2774. [PubMed: 30408420]

Kimbrough JM, Salinas-mondragon R, Boss WF, and Brown CS (2004). Kimbrough et al. 2004, The Fast and Transient Transcriptional Network of.pdf. 136(September):2790–2805.

Ko MSH (1992). Induction mechanism of a single molecule: stochastic or deterministic? BioEssays, 14(5):341–346. [PubMed: 1637366]

Kollist H, Zandalinas SI, Sengupta S, Nuhkat M, Kangasjärvi J, and Mittler R. (2019). Rapid Responses to Abiotic Stress: Priming the Landscape for the Signal Transduction Network. Trends in Plant Science, 24(1):25–37. [PubMed: 30401516]

Krouk G, Mirowski P, LeCun Y, Shasha DE, and Coruzzi GM (2010). Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. Genome Biology, 11(12):R123. [PubMed: 21182762]

Lammers NC, Galstyan V, Reimer A, Medin SA, Wiggins CH, and Garcia HG (2020). Multimodal transcriptional control of pattern formation in embryonic development. Proceedings of the National Academy of Sciences of the United States of America, 117(2):836–847. [PubMed: 31882445]

Larson DR, Zenklusen D, Wu B, Chao JA, and Singer RH (2011). Real-time observation of transcription initiation and elongation on an endogenous yeast gene. Science, 332(6028):475–478. [PubMed: 21512033]

Lee CH, Shin H, and Kimble J. (2019). Dynamics of Notch-Dependent Transcriptional Bursting in Its Native Context. Developmental Cell, 50(4):426–435.e4. [PubMed: 31378588]

Leivar P, Tepperman JM, Monte E, Calderon RH, Liu TL, and Quail PH (2009). Definition of early transcriptional circuitry involved in light-induced reversal of PIF-imposed repression of photomorphogenesis in young Arabidopsis seedlings. The Plant cell, 21(11):3535–53. [PubMed: 19920208]

Li Z, Defoort J, Tasdighian S, Maere S, Van De Peer Y, and De Smet R. (2015). Gene duplicability of core genes is highly consistent across all angiosperms. Plant Cell, 28(2):326–344.

Lim B, Fukaya T, Heist T, and Levine M. (2018a). Temporal dynamics of pair-rule stripes in living drosophila embryos. Proc Natl Acad Sci U S A, 115(33):8376–8381. [PubMed: 30061421]

Lim B, Heist T, Levine M, and Fukaya T. (2018b). Visualization of Transvection in Living Drosophila Embryos. Molecular Cell, 70(2):287–296.e6. [PubMed: 29606591]

Liu J. and Ma J. (2011). Fates-shifted is an f-box protein that targets bicoid for degradation and regulates developmental fate determination in drosophila embryos. Nat Cell Biol, 13(1):22–9. [PubMed: 21170036]

Liu TL, Upadhyayula S, Milkie DE, Singh V, Wang K, Swinburne IA, Mosaliganti KR, Collins ZM, Hiscock TW, Shea J, Kohrman AQ, Medwig TN, Dambournet D, Forster R, Cunniff B, Ruan

Y, Yashiro H, Scholpp S, Meyerowitz EM, Hockemeyer D, Drubin DG, Martin BL, Matus DQ, Koyama M, Megason SG, Kirchhausen T, and Betzig E. (2018). Observing the cell in its native state: Imaging subcellular dynamics in multicellular organisms. Science, 360(6386).

Lucas T, Ferraro T, Roelens B, De Las Heras Chanes J, Walczak AM, Coppey M, Dostatni N, Las JD, Chanes H, Walczak AM, Coppey M, and Umr UMR (2013). Live imaging of bicoid-dependent transcription in Drosophila embryos. Current Biology, 23(21):2135–9. [PubMed: 24139736]

Maheshri N. and O'Shea EK (2007). Living with noisy genes: How cells function reliably with inherent variability in gene expression. Annual Review of Biophysics and Biomolecular Structure, 36:413–434.

McFaline-Figueroa JL, Trapnell C, and Cuperus JT (2020). The promise of single-cell genomics in plants. Current Opinion in Plant Biology, 54:114–121. [PubMed: 32388018]

McLoughlin F, Basha E, Fowler ME, Kim M, Bordowitz J, Katiyar-Agarwal S, and Vierling E. (2016). Class I and II small heat shock proteins together with HSP101 protect protein translation factors during heat stress. Plant Physiology, 172(2):1221–1236. [PubMed: 27474115]

Melaragno JE, Mehrotra B, and Coleman AW (1993). Relationship between endopolyploidy and cell size in epidermal tissue of Arabidopsis. Plant Cell, 5(11):1661–1668. [PubMed: 12271050]

Mittler R, Finka A, and Goloubinoff P. (2012). How do plants feel the heat? Trends in Biochemical Sciences, 37(3):118–125. [PubMed: 22236506]

Munsky B, Neuert G, and Van Oudenaarden A. (2012). Using gene expression noise to understand gene regulation. Science, 336(6078):183–187. [PubMed: 22499939]

Nicolas D, Phillips NE, and Naef F. (2017). What shapes eukaryotic transcriptional bursting? Molecular BioSystems, 13(7):1280–1290. [PubMed: 28573295]

Park HY, Lim H, Yoon YJ, Follenzi A, Nwokafor C, Lopez-Jones M, Meng X, and Singer RH (2014). Visualization of dynamics of single endogenous mrna labeled in live mouse. Science, 343(6169):422–4. [PubMed: 24458643]

Queitsch C, Hong SW, Vierling E, and Lindquist S. (2000). Heat shock protein 101 plays a crucial role in thermotolerance in Arabidopsis. Plant Cell, 12(4):479–492. [PubMed: 10760238]

Raj A. and van Oudenaarden A. (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. Cell, 135(2):216–226. [PubMed: 18957198]

Roeder AH (2018). Use it or average it: stochasticity in plant development. Current Opinion in Plant Biology, 41:8–15. [PubMed: 28837855]

Ronald J. and Davis SJ (2019). Focusing on the nuclear and subnuclear dynamics of light and circadian signalling. Plant Cell and Environment, 42(10):2871–2884.

Rosenfeld N, Young JW, Alon U, Swain PS, and Elowitz MB (2005). Gene regulation at the single-cell level. Science, 307(5717):1962–5. [PubMed: 15790856]

Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M, Eggan E, Anastopoulos IN, Vargas-garcia CA, Singh A, Nathanson KL, Herlyn M, and Raj A. (2017). reprogramming as a mode of cancer drug resistance. Nature Publishing Group, 546(7658):431–435.

Suzuki N, Devireddy AR, Inupakutika MA, Baxter A, Miller G, and Song L. (2015). Ultra-fast alterations in mRNA levels uncover multiple players in light stress acclimation in plants. pages 760–772.

Taylor-Teeples M, Ron M, and Brady SM (2011). Novel biological insights revealed from cell type-specific expression profiling. Current Opinion in Plant Biology, 14(5):601–607. [PubMed: 21704550]

The Arabidopsis Genome Iniative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature, 408(December 2000):796–815. [PubMed: 11130711]

Tornaletti S, Reines D, and Hanawalt PC (1999). Structural characterization of RNA polymerase II complexes arrested by a cyclobutane pyrimidine dimer in the transcribed strand of template DNA. Journal of Biological Chemistry, 274(34):24124–24130.

Turco GM, Rodriguez-Medina J, Siebert S, Han D, Valderrama-Gómez M, Vahldick H, Shulse CN, Cole BJ, Juliano CE, Dickel DE, Savageau MA, and Brady SM (2019). Molecular Mechanisms Driving Switch Behavior in Xylem Cell Differentiation. Cell Reports, 28(2):342–351.e4. [PubMed: 31291572]

Tutucci E, Livingston NM, Singer RH, and Wu B. (2018). Imaging mRNA In Vivo, from Birth to Death. Annual Review of Biophysics, 47(1):85–106.

Walters MC, Fieringt S, Eidemillert J, Magist W, and Groudinett M. (1995). Enhancers increase the probability but not the level of gene expression. Gene Expression, 92(July):8.

Wilhelmsson PK, Mühlich C, Ullrich KK, and Rensing SA (2017). Comprehensive Genome-Wide Classification Reveals That Many Plant-Specific Transcription Factors Evolved in Streptophyte Algae. Genome Biology and Evolution, 9(12):3384–3397. [PubMed: 29216360]

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, and Provart NJ (2007). An "electronic fluorescent pictograph" Browser for exploring and analyzing large-scale biological data sets. PLoS ONE, 2(8):1–12.

Wu B, Chao JA, and Singer RH (2012). Fluorescence fluctuation spectroscopy enables quantitative imaging of single mrnas in living cells. Biophysical Journal, 102:2936. [PubMed: 22735544]

Wu Z, Ietswaart R, Liu F, Yang H, Howard M, and Dean C. (2016). Quantitative regulation of FLC via coordinated transcriptional initiation and elongation. Proceedings of the National Academy of Sciences of the United States of America, 113(1):218–223. [PubMed: 26699513]

Xu H, Sepulveda LA, Figard L, Sokac AM, and Golding I. (2015). Combining protein and mrna quantification to decipher transcriptional regulation. Nat Methods, 12(8):739–42. [PubMed: 26098021]

Yoshida T, Ohama N, Nakajima J, Kidokoro S, Mizoi J, Nakashima K, Maruyama K, Kim JM, Seki M, Todaka D, Osakabe Y, Sakuma Y, SchöZ F, Shinozaki K, and Yamaguchi-Shinozaki K. (2011). Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. Molecular Genetics and Genomics, 286(5–6):321–332. [PubMed: 21931939]

Zandalinas SI, Fritschi FB, Mittler R, and Lawson T. (2020). Signal transduction networks during stress combination. Journal of Experimental Botany, 71(5):1734–1741. [PubMed: 31665392]
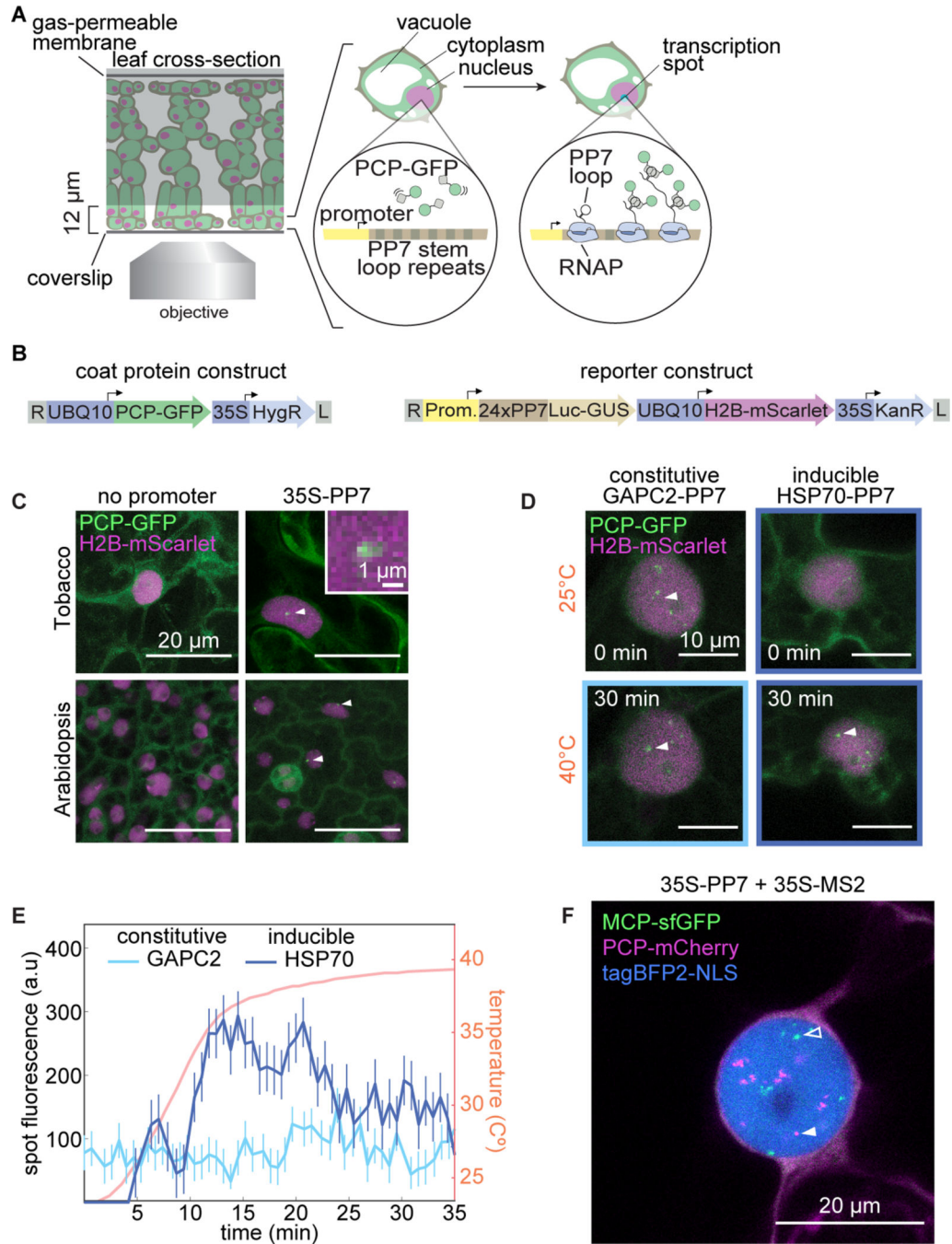
Figure 1. Fluorescence labeling of nascent RNA in tobacco and Arabidopsis reveals single-cell transcriptional dynamics in real time.
(**A**) Schematic of the live-imaging experimental setup in leaves and diagram of the PP7 RNA labeling system. (**B**) Schematic of the constructs used in this study. (UBQ10, Arabidopsis ubiquitin 10 promoter; 35S, CaMV 35S promoter; HygR, hygromycin resistance; Luc-GUS, firefly luciferase-/3-glucoronidase fusion; H2B, Arabidopsis histone 2B coding sequence; KanR, kanamycin resistance; L, T-DNA left border; R, T-DNA right border). (**C**) Maximum projection of snapshots of cells expressing PCP-GFP and the reporter construct with or

without the constitutive 35S promoter driving expression of the PP7-tagged Luc-GUS gene. White arrows indicate nuclear fluorescent puncta corresponding to transcription spots. Inset: magnification of PP7 fluorescence. **(D)** Maximum projection snapshots of tobacco cells expressing PCP-GFP and reporter constructs driven by the promoters of the Arabidopsis *GAPC2* and *HSP70* genes. Time under heat shock is indicated. White arrowheads indicate the fluorescent spots quantified in (F). **(E)** Fluorescence time traces of single nuclear GFP puncta in tobacco leaf epidermis cells expressing PCP-GFP and reporter constructs driven by various Arabidopsis promoters. Prior to spot detection, spots are assigned a fluorescence value of zero. Error bars represent the uncertainty in the spot fluorescence extraction (Materials and Methods). **(F)** Maximum projection snapshot of tobacco leaf epidermal cell expressing PCP-mCherry, MCP-GFP, H2B-tagBFP2, and two reporter constructs driven by the 35S promoter and tagged with PP7 (magenta) or MS2 (green). Open and closed arrowheads indicate MCP-tagged and PCP-tagged nascent RNAs, respectively.
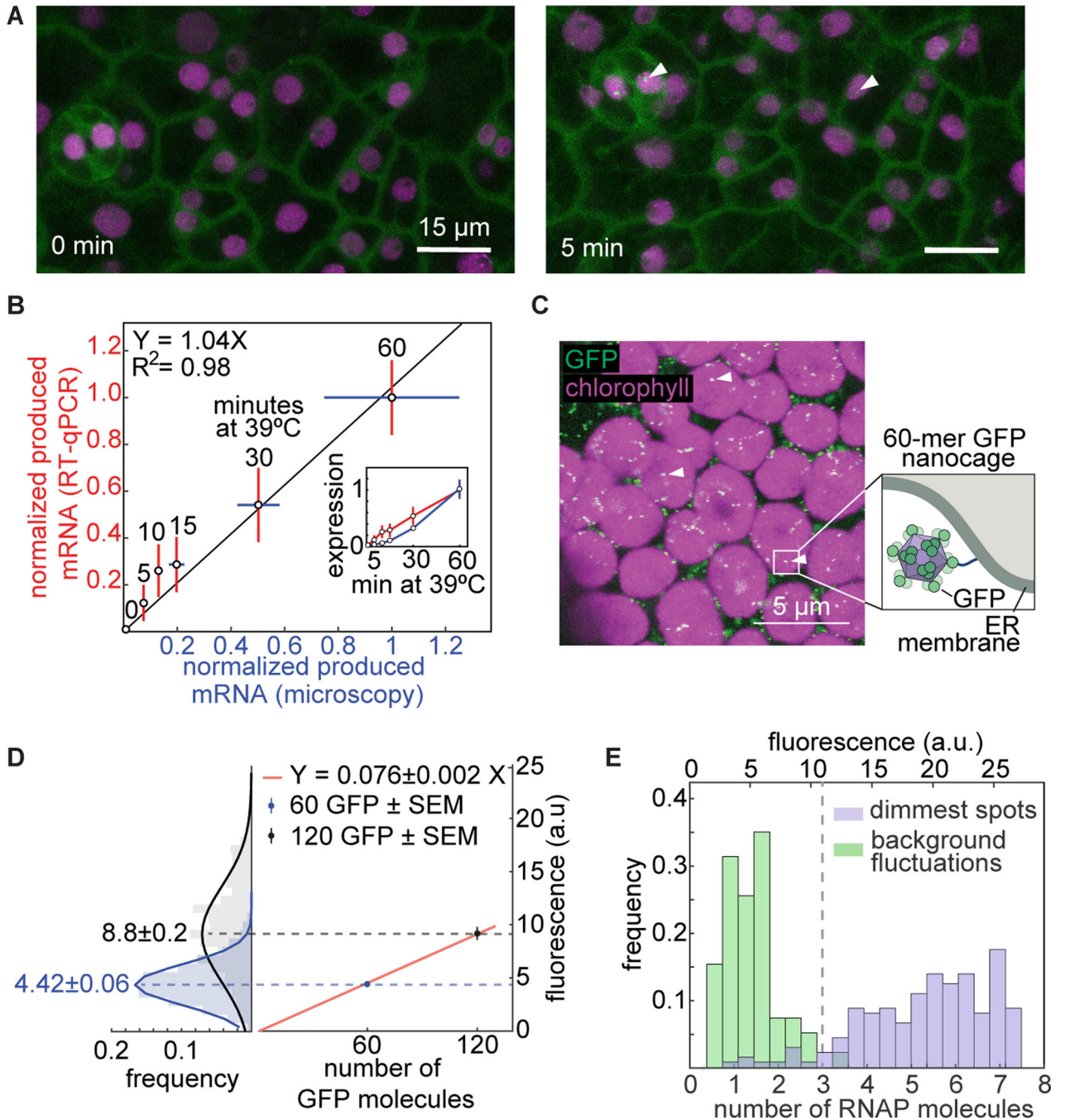
Figure 2. Cross validation, absolute calibration, and sensitivity of the PP7 reporter system.
(**A**) Maximum fluorescence projections of leaf epidermal tissue of an Arabidopsis line stably transformed with PCP-GFP and a reporter construct driven by the HSP101 promoter under heat shock. Time stamps indicate time under heat shock. Arrowheads point to transcription spots. (**B**) Comparison between total mRNA produced as reported by RT-qPCR and PCP-GFP. PCP-GFP error corresponds to the standard error of the mean over 10 biological replicates; RT-qPCR error corresponds to the standard error of the mean (SEM) across three biological replicates. Data are normalized to each corresponding signal at 60

min. The solid black line shows a linear fit to the data going through the origin. The inset shows the normalized mean and SEM of expression level as a function of time for RT-qPCR and microscopy. **(C)** Maximum fluorescence projection of a tobacco mesophyll cell expressing a construct encoding a 60 GFP nanocage tethered to the outer membrane of the endoplasmic reticulum (ER). **(D, left)** Absolute calibration of GFP fluorescence. Histograms and Gaussian fit of single-nanocage fluorescence distributions for the 60-GFP (blue) and 120-GFP (black) nanocages transiently expressed in tobacco leaves. The mean of each distribution is shown next to each histogram. As expected, the means are related by a factor of two. **(D, right)** Mean and standard error of the mean (SEM) of the nanocage fluorescence as a function of number of GFP molecules per cage. The red line is a linear fit passing through the origin, revealing a calibration factor of 0.076 ± 0.002 a.u./GFP molecule (error reporting on the 95% confidence interval). **(E)** Histograms of the calibrated number of transcribing RNAP molecules in the dimmest three frames of the weakest half of HSP101-PP7 fluorescence time traces (purple) and their associated fluorescence background fluctuations (green). The point where the distributions overlap, at 3 RNAP molecules (vertical dashed line), can be considered the detection threshold.
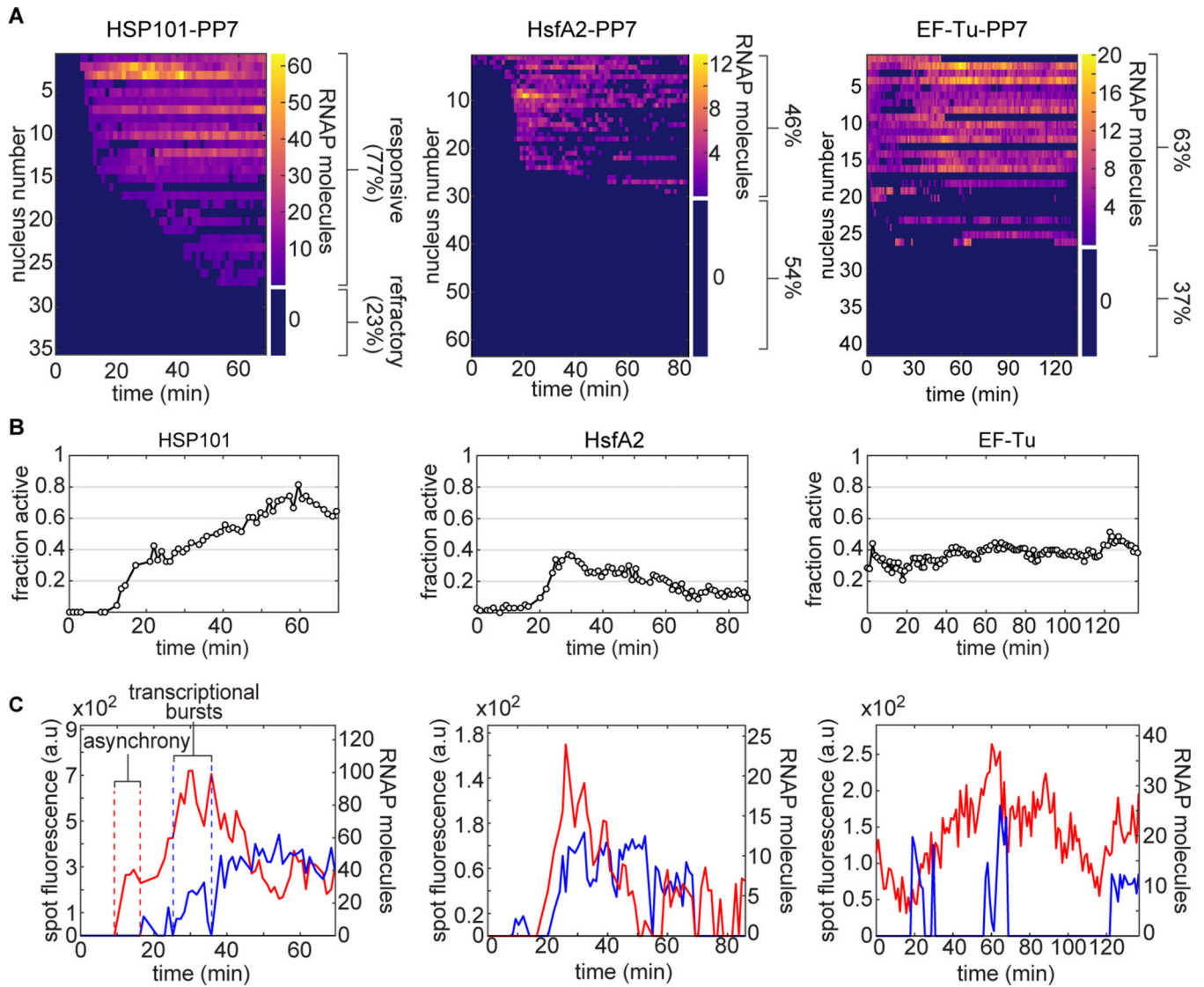
Figure 3. Single-cell control of transcriptional activity in response to heat shock in Arabidopsis.
(A) Heat maps of spot fluorescence in all nuclei (rows) over time (columns) across the the field of view in HSP101-PP7–1, HsfA2-PP7–1, and EF-Tu-PP7–1 plants. Dark blue represents the absence of detectable signal. The size of the colorbar on the right of each heatmap shows the proportion of nuclei that exhibited activity in at least one frame during the experiment (>68 min) to refractory cells that presented no spots. **(B)** Instantaneous fraction of actively transcribing nuclei measured as the number of nuclei with spots divided by the total number of nuclei in the field of view. **(C)** Representative single-spot fluorescence time traces. Upon induction, transcriptional onset can occur asynchronously and transcriptional activity occurs in bursts, modulating the instantaneous fraction of transcriptionally active nuclei in (B).
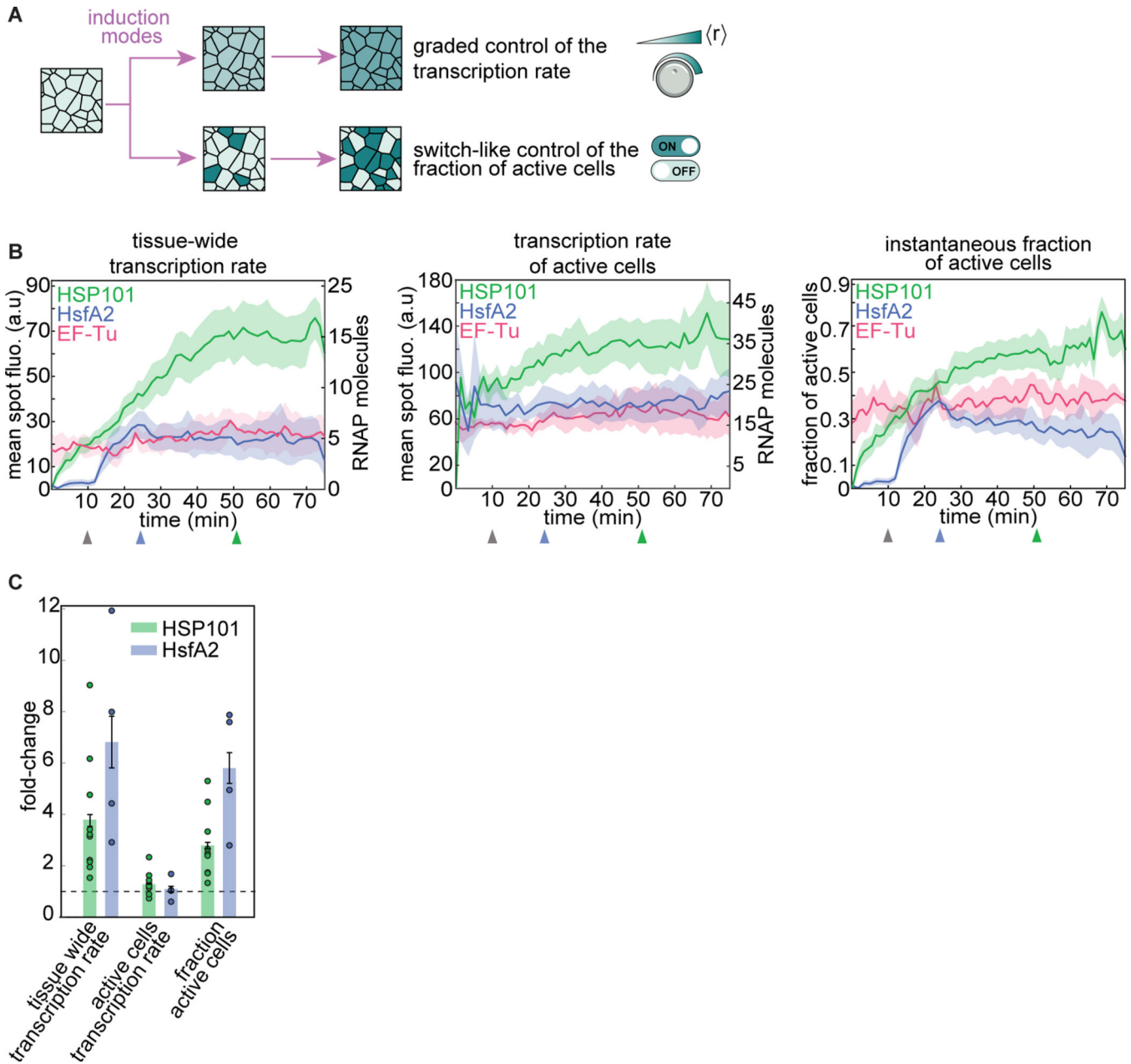
Figure 4. Single-cell regulatory strategies determining tissue-wide transcriptional dynamics.
(A) Tissue-wide transcriptional control can be achieved through two non-exclusive regulatory modes: the graded modulation of the rate of transcription across cells, or the switch-like regulation of the fraction of actively transcribing cells. (B) Mean tissue transcription rate (left), transcription rate of active cells (middle), and instantaneous fraction of actively transcribing cells (right) for Arabidopsis lines carrying inducible promoters HSP101-PP7–1 (green) and HsfA2-PP7–1 (blue), and a line with the constitutive reporter EF-Tu-PP7–1 (red). Time $t = 0$ corresponds to the frame at which spots were first detected. (C) Fold-change in the mean tissue-wide transcription rate compared to the fold-change in the mean transcription rate of active cells and in the fraction of active cells, defined

as the ratio between the value at its peak and at $t = 10min$ for HSP101-PP7–1 (gray vs. green arrowheads in B) and HsfA2-PP7 (gray vs. blue arrowheads in B). The horizontal dashed line indicates a fold change of 1. (A-C, shaded regions and error bars are SEM calculated across 10, 5, and 3 experimental replicates for HSP101-PP7–1, HsfA2-PP7–1, and EF-Tu-PP7–1, respectively.)
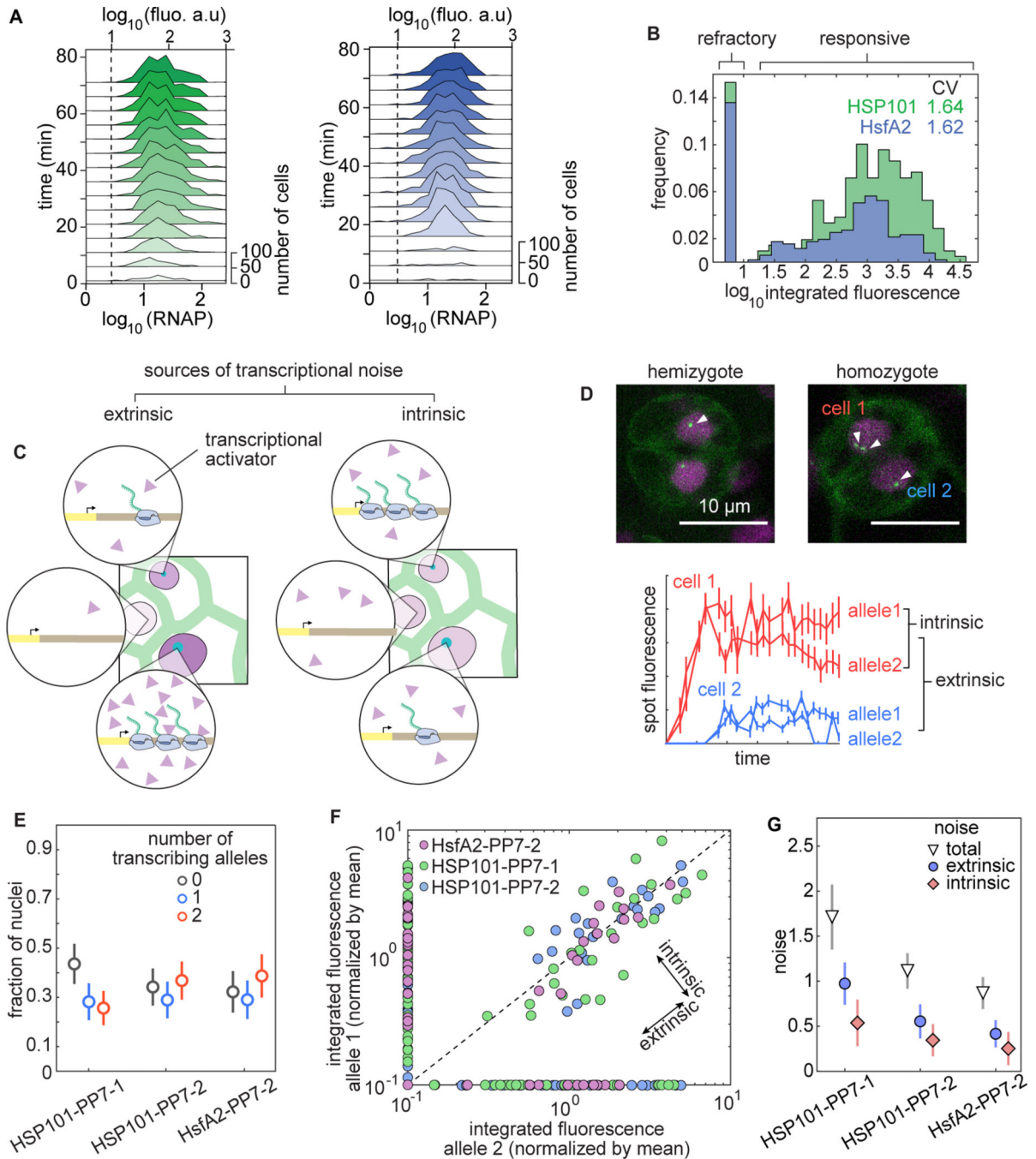
Figure 5. Allele-specific processes explain most of the cellular heterogeneity in produced mRNA in Arabdopsis.
(A) Histograms of spot fluorescence over time for the combined replicates of Figure 4. The dashed line indicates the detection threshold determined in Figure2D. **(B)** Histograms of predicted total produced mRNA per cell across all replicates from Figure 4. **(C)** Schematic of extrinsic (left) and intrinsic (right) sources of transcriptional noise. Extrinsic noise arises from cellular differences in the abundance of regulatory molecules (purple triangles) while intrinsic noise captures differences among cells with identical composition. **(D)** Two-allele

experiment to decompose the total transcriptional variability into intrinsic and extrinsic noise. Top: guard cells (obligate diploids) expressing HSP101-PP7. White arrowheads indicate transcription spots corresponding to one or two alleles of the reporter transgene in homologous chromosomes. In the homozygote it is possible for only one allele to be active in different cells. Bottom: spot fluorescence traces from homozygous cells shown on top. **(E)** Fraction of nuclei with zero, one, or two spots in heat shock-treated homozygous plants at the frame with the maximum number of visible spots. **(F)** Scatter plot of the integrated spot fluorescence normalized by the mean for alleles belonging to the same nucleus. Undetected spots were assigned a value of zero and plotted on the x- and y-axes. **(G)** Decomposition of the total variability in (F) into extrinsic and intrinsic components shows comparable contributions of both components to the total noise, with the intrinsic component explaining most of the variability. Error bars in (E) and (G) are bootstrapped errors.