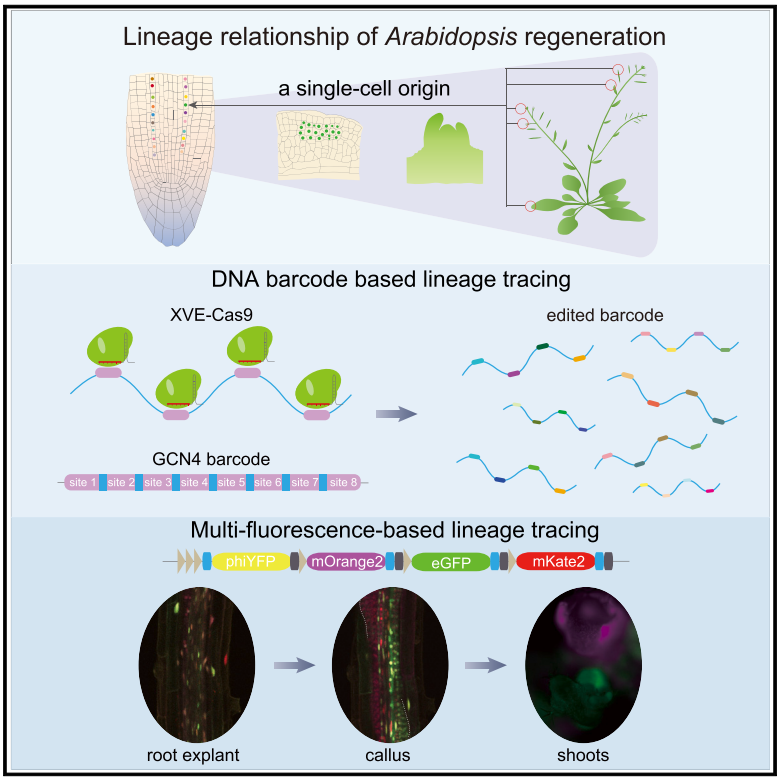
Article

Developmental Cell

Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration

Graphical abstract



Highlights

- Generation of cell-lineage tracing toolsets
- Lineage tracing reveals the single-cell origin of *Arabidopsis* regeneration
- Lineage tracing reveals the multicellular origin of lateral root development

Authors

Xinyue Lu, Qiyan Zhang, Zejia Wang, ..., Shuyi Cai, Huawei Zhang, Qikun Liu

Correspondence

qikunliu@pku.edu.cn

In brief

Lu and Zhang et al. investigate the cell-lineage relationship of regenerated *Arabidopsis* plants. They demonstrate that the regenerated plant originated from a single somatic cell within the donor tissue.

Lu et al., 2025, Developmental Cell *60*, 1–15
January 20, 2025
https://doi.org/10.1016/j.devcel.2024.10.023

CellPress

Developmental Cell

CellPress
OPEN ACCESS

Article

Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration

Xinyue Lu,[1,3] Qiyan Zhang,[1,3] Zejia Wang,[1] Xuanzhi Cheng,[1] Huiru Yan,[2] Shuyi Cai,[1] Huawei Zhang,[2] and Qikun Liu[1,4,*]
[1]State Key Laboratory of Protein and Plant Gene Research, School of Advanced Agricultural Sciences, Peking University, Beijing 100871, China
[2]Institute of Advanced Agricultural Science, Peking University, Weifang, China
[3]These authors contributed equally
[4]Lead contact
*Correspondence: qikunliu@pku.edu.cn
https://doi.org/10.1016/j.devcel.2024.10.023

SUMMARY

Plants demonstrate a high degree of developmental plasticity, capable of regenerating entire individuals from detached somatic tissues—a regenerative phenomenon rarely observed in metazoa. Consequently, elucidating the lineage relationship between somatic founder cells and descendant cells in regenerated plant organs has long been a pursuit. In this study, we developed and optimized both DNA barcode- and multi-fluorescence-based cell-lineage tracing toolsets, employing an inducible method to mark individual cells in *Arabidopsis* donor somatic tissues at the onset of regeneration. Utilizing these complementary methods, we scrutinized cell identities at the single-cell level and presented compelling evidence that all cells in the regenerated *Arabidopsis* plants, irrespective of their organ types, originated from a single progenitor cell in the donor somatic tissue. Our discovery suggests a single-cell passage directing the transition from multicellular donor tissue to regenerated plants, thereby creating opportunities for cell-cell competition during plant regeneration—a strategy for maximizing survival.

INTRODUCTION

The organizational pattern and division dynamics of stem cell lineages are fundamental to plant and animal development. Notably, contrasting developmental patterns emerge between these two kingdoms. In animals, primordial germ cells are set aside early in embryonic development for reproductive purposes. By contrast, plants maintain a post-embryonic stem cell niche (SCN) in meristematic tissues and initiate the formation of vegetative and reproductive organs at different developmental stages.[1]

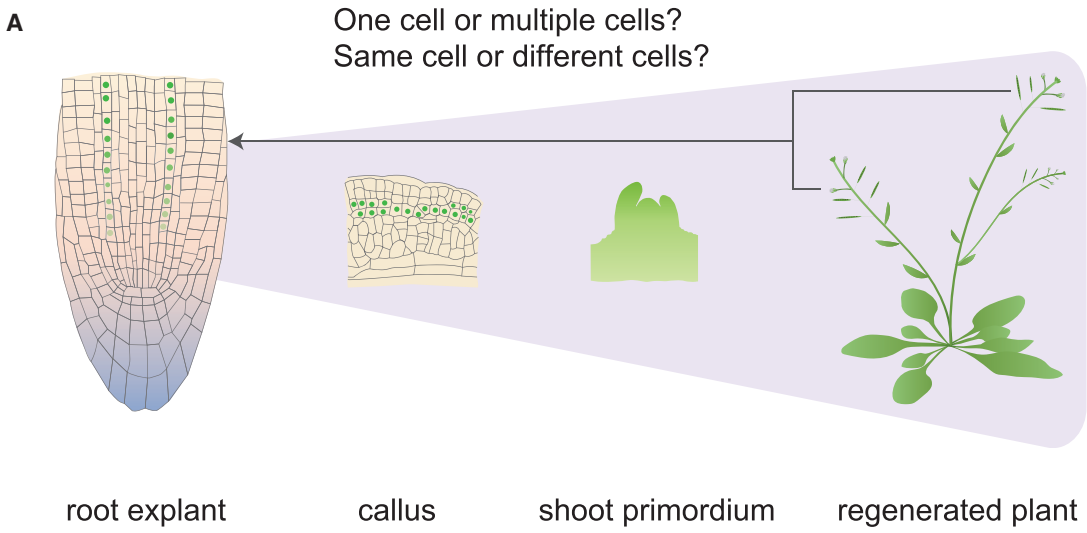
While stage-specific developmental trajectories of meristematic tissue can be computationally delineated based on cell atlases generated through single-cell sequencing,[2–8] constructing retrospective cell division histories requires traceable markers inherited through mitotic divisions. These traceable markers can be broadly categorized into two types: visible labels and DNA barcode-based markers.[9–11]

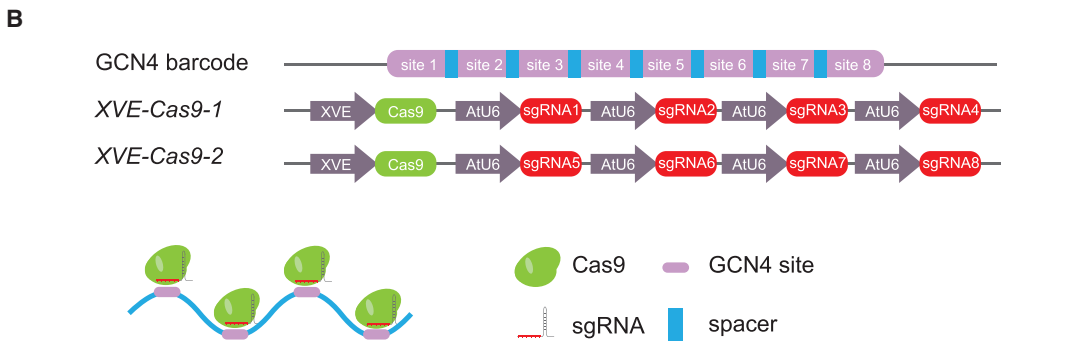
Clonal analyses of meristematic tissue using visible markers have been conducted in diverse plant species, employing various labeling strategies such as dyes, radioactive tracers, β-glucuronidase (GUS) reporters, and fluorescent proteins.[12–19] GUS reporter labeling in random cells has indicated that cell fate for nodes, axillary buds, and internodes is separately designated and emerges sequentially during rice stem development.[20] Combining a fluorescent protein with the CRISPR technique enables the study of lineage relationships in living plants. This breakthrough is vividly illustrated by the successful visualization of lineage tracing in the meristematic cells present in *Arabidopsis* shoot and *Marchantia polymorpha* apical notch.[19] Furthermore, the application of tissue-specific and/or chemically inducible promoters further increases the temporal and spatial versatility of lineage-tracing tools.[19] While clonal analysis in plants has predominantly relied on a single binary readout, multi-fluorescent labeling methods have been developed in animal systems for enhanced resolution in cell labeling.[21,22]

Alternatively, naturally occurring somatic mutations can function as endogenous DNA barcodes, providing information about the cell division history.[9] In animals, various types of genetic variations have been utilized, including single-nucleotide variants, copy-number variations, transposon insertions, and mtDNA mutations, among others.[23–26] Profiling somatic mutations in shrub willow indicated an early separation of leaf and root cell lineages, analogous to the metazoan germline development.[27] However, the scarcity of naturally occurring somatic mutations has made it challenging to profile lineage information at single-cell level.

More recently, DNA sequence variations artificially created by Cas9 have been employed as barcodes for cell-lineage identification.[28–33] The diverse range of genetic lesions and the corresponding high-resolution profiling, facilitated by high-throughput

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023





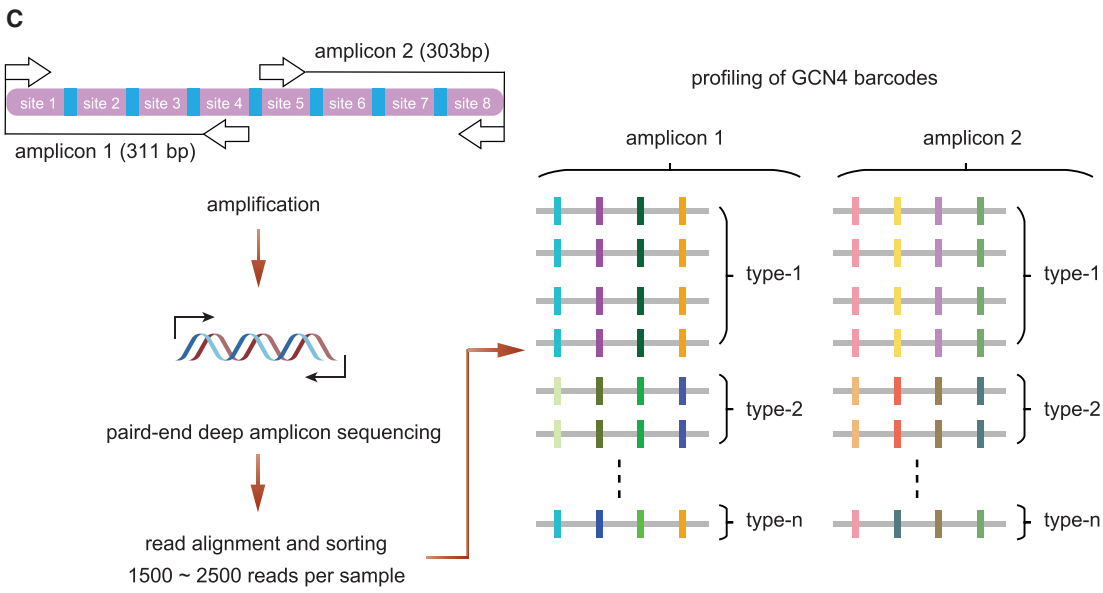


Figure 1. Design of an inducible DNA barcoding system for lineage tracing during plant regeneration

(A) A schematic model illustrates *Arabidopsis* organogenesis from multicellular root explants and the driven questions of this study.

(B) The schematic illustration depicts the inducible CRISPR-based DNA barcoding system. Eight repeats of GCN4 coding sequences are tandemly arranged and separated by 15-bp spacers. Either the first four or the second four GCN4 repeats are targeted by individually designed sgRNAs. The expression of Cas9 is driven by the β-estradiol-inducible *XVE* promoter.

(legend continued on next page)

sequencing, have significantly enhanced our capability in lineage tracing. This success has been demonstrated at both organ and whole-organism levels in various animal systems.[31,32,34–37] Furthermore, the delineation of DNA barcodes can be coupled with transcriptome analysis, providing additional cell type information for each lineage.[28,29,33,34,38,39] By contrast, no DNA barcode-based lineage tracing tools have been reported in plants so far.

The replenishment of damaged or lost organs through regeneration represents a key survival strategy for both animals and plants. Examples include planarian head and tail reemergence after wounding and plant root regeneration after tip excision.[40,41] Moreover, plants possess a high capability for the complete re-establishment of whole new individuals from detached tissues.[42]

Whole plant regeneration can occur through two different modes: somatic embryogenesis and organogenesis.[43] In somatic embryogenesis, a non-zygotic embryo is first formed, developing into a new plant in a self-autonomous way.[44] Conversely, organogenesis occurs in two consecutive steps: a mass of pluripotent cells, called callus, is initially formed,[45–47] which can then further regenerate either shoot or root.[48] High levels of cytokinin promote shoot formation, while a high auxin content promotes root formation.[49,50]

Interestingly, meristematic tissues appear dispensable for plant regeneration.[40,43,51,52] This prompts questions about the cell origin of the regenerated plant tissues. Following root tip excision, cells from multiple tissue types were found to contribute to the replenished root tips.[51] Similarly, the use of an inducible GUS reporter has uncovered the identity of founder cells contributing to adventitious roots.[52] While these studies focused on the cell origin of each regenerated organ, the lineage history of descendant cells for the entire regenerated plant has remained unknown (Figure 1A).

Furthermore, another interesting question worth exploring is the minimum number of somatic cells designated for each regenerated plant. It has been demonstrated in different plant species that, a single protoplast isolated from leaf mesophyll cells can reenter the cell cycle and form callus, which eventually regenerates into a whole plant.[53,54] On the other hand, organogenesis from multicellular explants recapitulates the process of organ initiation during post-embryonic development, which has been proposed to involve varied numbers of founder cells.[12,27,55,56] Furthermore, as intercellular communication is crucial for cell fate transitioning and organogenesis,[57] it remains unclear if plant regeneration initiated from multicellular explant tissues is also of single-cell origin.

In this study, we developed and optimized toolsets that operate in an inducible manner, capable of tracing the division history of somatic cells through DNA barcoding and multi-color fluorescence. Using these combined methods, we conducted lineage profiling in regenerated plants at the single-cell level and unexpectedly discovered that all cells in the regenerated *Arabidopsis* plant, irrespective of their tissue types, originated from a single somatic progenitor cell in the multicellular explant. Our study not only provided toolsets suitable for plant cell-lineage tracking but also suggested a single-cell passage directing the regeneration process of multicellular donor explants. This phenomenon represents a strategy for generating offspring of better fitness through cell-cell competition, optimizing the chances of plant survival.

RESULTS

Development of the inducible DNA barcoding system

To trace cell lineages in the regenerated plants, we initially synthesized a 250-bp DNA barcode sequence comprising two functional single guide RNA (sgRNA) targeting sites that are tandemly repeated five times each (Figure S1A). The sgRNA targeting sites can be effectively edited in wheat but lack sequence homology to the *Arabidopsis* genome, minimizing CRISPR off-targeting.[58] However, upon obtaining stable transgenic *Arabidopsis*, we observed frequent and spontaneous rearrangements of the DNA barcode transgene (Figure S1B). Substituting the wheat sgRNA sites with functional rice sgRNA target sequences yielded similar results (Figures S1A and S1C),[58] indicating a lack of DNA sequence stability when identical sgRNA sites are closely arranged.

To address this issue, we employed the design of the 10× GCN4 tandem array derived from the *SunTag* construct, initially developed for targeted gene activation.[59] This design has demonstrated successful applications in both animal and plant systems.[60–62] Importantly, within the 10×GCN4 tandem array, each of the 57-bp GCN4 coding sequences exhibits slight variations from one another, separated by a 15-bp spacer sequence. This characteristic makes it an ideal sgRNA target for creating DNA barcodes (Figure 1B). The sequence integrity over the entire length of the GCN4 transgene was confirmed in multiple *Arabidopsis* transgenic lines (Figures S1D and S1E). Among these plants, line #4482, which harbors multiple copies of the 10× GCN4 transgene at a single genomic locus, was selected for future experiments.

To preserve the DNA barcodes before plant regeneration, we developed a chemically inducible Cas9 barcode editor integrating the *LexA-VP16-ER* module (*XVE-Cas9*) (Figure 1B).[63] Although up to 10 GCN4 repeats are available for DNA barcode editing, we designed two 4×sgRNA expressing cassettes and restricted the usage of barcodes to either the first four (*XVE-Cas9-1*) or the second four (*XVE-Cas9-2*) GCN4 repeats (Figure 1B). This approach allows for convenient profiling of DNA barcodes through PCR followed by paired-end deep-amplicon sequencing, maximizing the resolution in identifying different types of DNA barcodes present in each sample (Figure 1C).

Multiple *Arabidopsis* T1 lines carrying the *XVE-Cas9-1* or *XVE-Cas9-2* transgene were obtained. Among them, lines

(C) Schematic illustration of the workflow of DNA barcode profiling. To investigate the genetic lesions generated at the GCN4 sites, DNA from either bulk tissue samples or isolated single nuclei is extracted and amplified. The amplicon length for *Cas9-1-GCN4* and *Cas9-2-GCN4* targeting sites is 311 bp (amplicon 1) and 303 (amplicon 2) bp, respectively, corresponding to the maximum readable length of paired-end deep-amplicon sequencing. The amplicons were sequenced at a depth of 1,500- to 2,000-fold. The sequenced reads are then sorted and aligned to the GCN4 reference to identify the types of genetic lesions created by Cas9, representing different types of DNA barcodes present in the original sample.
See also Figure S1.

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

(#4551 and #4567) exhibiting the lowest basal expression levels of Cas9 were selected for crossing with the GCN4 transformed plant (#4482, Figure S2A). The resulting plants (*Cas9-1-GCN4* and *Cas9-2-GCN4*) showed no observable developmental or regeneration abnormalities compared with wild-type plants (Figures S2B–S2H).

Characterization of DNA barcodes

We assessed the effects of β-estradiol treatment on Cas9 expression and the editing of GCN4 DNA barcodes. Root explants of *Cas9-1-GCN4* and *Cas9-2-GCN4* were placed on callus-inducing media (CIM) containing various concentrations of β-estradiol, ranging from 1 to 100 μM. No inhibitory impact on the explants' regenerative capacity was observed with the β-estradiol treatments (Figure S3). Cas9 expression levels were measured by quantitative real-time PCR on different days after the treatments. The results showed that for both *Cas9-1-GCN4* and *Cas9-2-GCN4*, treatment with 1 μM β-estradiol had minimal impact on Cas9 expression compared with the non-treatment control (Figures 2A and S2I). By contrast, Cas9 expression was significantly induced by treatment with 10 μM β-estradiol and even more so with 100 μM β-estradiol treatment (Figures 2A and S2I). Temporally, Cas9 expression of *Cas9-1-GCN4* peaked on day 3 on CIM but rapidly returned to the non-treatment level by day 5, suggesting a transient induction of Cas9 by β-estradiol treatment (Figure 2A). In *Cas9-2-GCN4* plants, Cas9 expression also peaked on day 3 but decreased to half the peak level on day 5 and remained unchanged by day 10 (Figure S2I).

We further characterized the types and frequency of genetic lesions comprising the GCN4 DNA barcode under different treatments using an equal number of root explants ($n = 36$). The results indicated that in the absence of β-estradiol treatment, very few Cas9-induced genetic lesions were detected, regardless of treatment duration, suggesting minimal background activity of *XVE-Cas9-1* (Figures 2B and 2C). As the β-estradiol concentration increased, the frequency of Cas9-induced lesions also rose, with the 100 μM β-estradiol group displaying the highest overall editing efficiency (Figures 2B and 2C). Moreover, extending the duration of β-estradiol treatment from 3 to 10 days did not appear to further increase the number and types of genetic lesions, even though editable DNA barcodes are still available (Figures 2B and 2C). This aligns with the observation that Cas9 expression was transiently induced within the initial 3 days (Figure 2A). Consequently, in subsequent experiments, samples were typically treated with 100 μM β-estradiol for 5 days unless otherwise specified.

To validate the reproducibility of our barcode profiling method, barcode amplification followed by paired-end deep-amplicon sequencing was repeated for each DNA sample. The enrichment of each type of genetic lesion was represented by the frequency of the corresponding sequenced reads and was directly compared between the two replicates. The results demonstrated a strong correlation and high reproducibility between replicates (Figure 2D). However, we also observed a relatively weak correlation for DNA barcodes with a read frequency less than 5% (Figure 2D), possibly due to inefficient amplification of certain types of edited DNA barcode. To further enhance the robustness of the analysis, barcode profiling in subsequent experiments was consistently conducted in triplicates, and the results were pooled for downstream analysis (see STAR Methods). For each sample, barcodes with a read frequency above 5% were noted as major types, whereas the rest were noted as minor types.

Next, we examined in detail the types and distribution frequency of genetic lesions in *XVE-Cas9-1* transformed calli ($n = 20$) treated with and without β-estradiol. A range of genetic lesions was observed, including nucleotide insertions between 1 and 24 bp and deletions spanning 1–233 bp (Figure 2E). Genetic lesions were observed in each of the first four GCN4 repeats comprising the DNA barcode, and the frequency of observing genetic lesions simultaneously occurring at two GCN4 repeats was the highest, followed by editing in one site and three sites (Figure 2F).

Similarly, the types and frequency of GCN4 genetic lesions introduced by *XVE-Cas9-2* were also examined using the same amount of calli, which exhibited a lower diversity compared with that introduced by *XVE-Cas9-1* (Figure 2G). Consequently, *Cas9-1-GCN4* plants were selected for cell-lineage investigation in subsequent experiments.

Barcode profiling in regenerated plant organs

After establishing the inducible barcoding system, we proceeded to compare the profiles of DNA barcodes present in various organs of the regenerated *Arabidopsis* plants. Root explants of *Cas9-1-GCN4* plants were cultured on CIM supplemented with β-estradiol to induce CRISPR editing (Figure 3A). On day 5, calli were transferred to shoot-inducing media (SIM) without β-estradiol to induce shoot formation and plant regeneration (Figure 3A). Leaves from distinct shoot branches and dissected floral organs originating from a single floral bud were collected for barcode profiling (Figure 3B).

Intriguingly, the results indicated a comparable composition and distribution pattern of major barcodes among different organs in the regenerated plant, regardless of their branch origins (Figure 3C). Repetition of the experiments using additional regenerated plants yielded similar findings (Figures S4A–S4D). Compared with the major barcodes, the distribution pattern of minor barcodes is less consistent among different samples (Figure 3C). The variation in minor barcodes may be attributed to inefficient or inconsistent amplification of certain types of genetic lesions. Alternatively, lineage-specific Cas9 editing may be present.

The overall resemblance in DNA barcode distribution across regenerated organs argues against the preservation of organ-specific somatic cell lineages in the explants (Figures 3B and 3C). Additionally, in each regenerated plant, we observed a total of 6–8 different types of DNA barcodes, much less than the average number of barcode types ($n = 31$) observed in a single callus (Figure 2G), suggesting a much-reduced cell heterogeneity in the regenerated plants compared with the callus. However, since barcode profiling was performed at the organ level, it remains uncertain whether the observed combinations of DNA barcodes stem from a few somatic cells harboring heterogeneous barcode sequences or a single ancestor cell carrying multiple types of DNA barcodes.

Barcode profiling using isolated single nuclei

To determine the exact number of founder cells giving rise to the regenerated *Arabidopsis*, we decided to conduct barcode profiling in the regenerated plants at the single-cell level. We first tested whether DNA barcodes could be precisely examined at

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

CellPress
OPEN ACCESS

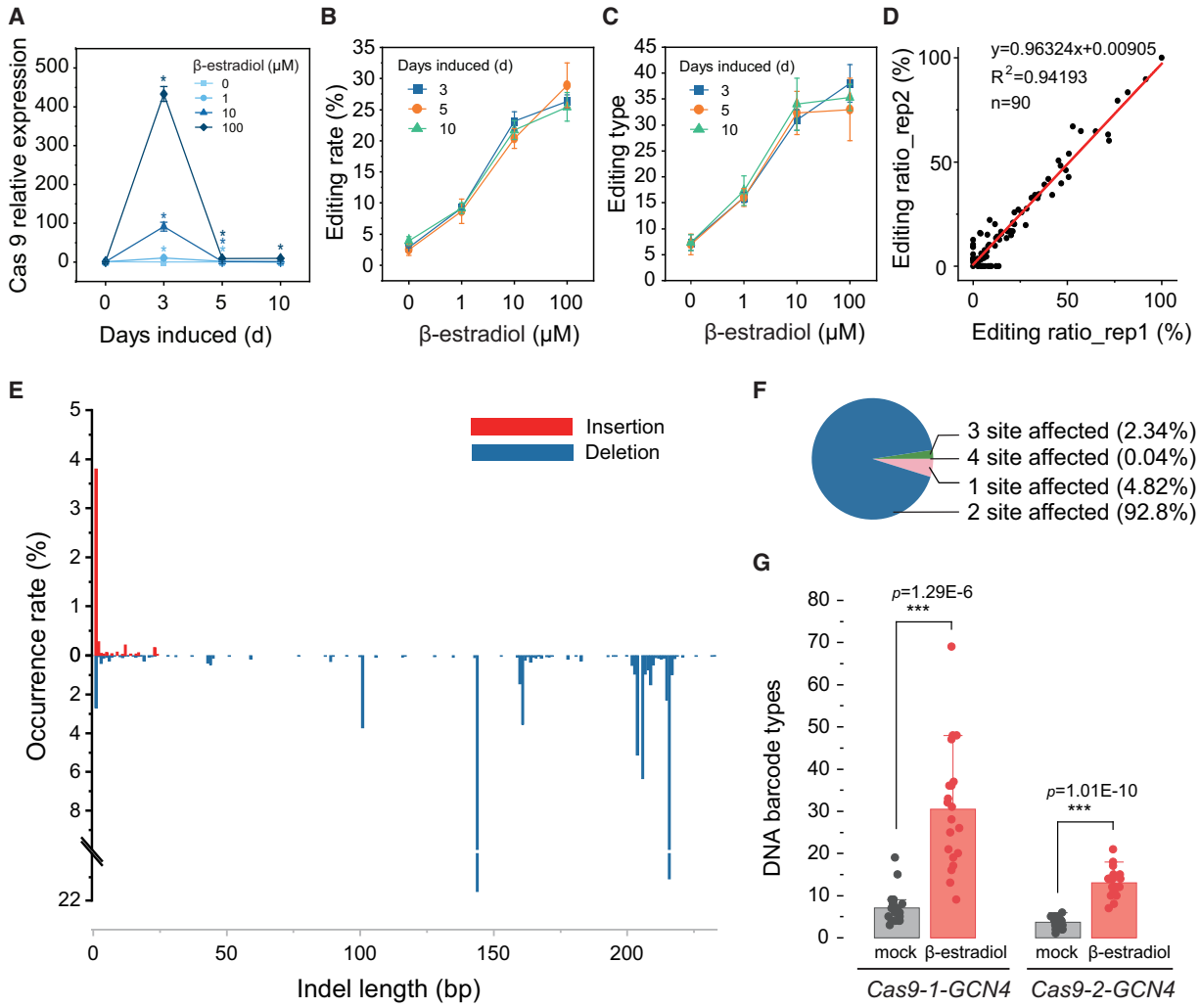


Figure 2. Characterization of DNA barcodes in *Cas9-1/2-GCN4* transgenic plants

(A) Relative expression levels of Cas9 in calli treated with different concentrations of β-estradiol on different days. Each biological replicate comprises approximately thirty pooled calli. The data represent the mean ± SEM from three biological replicates. Asterisks indicate significant differences compared with the control under 0 μM β-estradiol treatment on day 0 (Student's t test, $p < 0.05$).

(B) Editing rate of DNA barcode in *Cas9-1-GCN4* transformed calli treated with different concentrations of β-estradiol on different days. Reads were mapped to the GCN4 repeats. The ratio was calculated by normalizing the total number of reads containing Cas9-induced genetic lesions to the total number of mappable reads. Each sample contains an equal number of explants ($n = 36$). The data present the mean ± SD from three technical replicates.

(C) Data convention is the same as in (B) except for measuring the number of DNA barcode types.

(D) Reproducibility in quantification of DNA barcodes. Two independent replicates of barcode amplification and deep-amplicon sequencing were performed for each DNA sample. The frequency of each type of DNA barcode was then compared between the two replicates.

(E) A histogram depicting the size distribution of insertions (top) and deletions (bottom) in the DNA barcode. Data were collected from 20 *Cas9-1*-GCN4 transformed calli treated with 100 μM β-estradiol for 5 days on CIM.

(F) The ratio of DNA barcodes with genetic lesions occurring simultaneously at varying numbers of GCN4 sites.

(G) Diversity of DNA barcode types in *Cas9-1/2-GCN4* transformed calli. Each dot represents the number of DNA barcode types detected in a single callus in the presence (red) or absence (gray) of β-estradiol treatment. Each vertical bar represents the mean ± SD from 20 calli (Student's t test, $p < 0.05$).

See also Figure S2.

the single-cell level using multicellular root explant tissue prior to plant regeneration. Nuclei were isolated from the root explants treated on CIM with and without β-estradiol and stained with Propidium Iodide (PI). In such case, genetic lesions were independently induced by Cas9 within each individual cell, potentially creating a great diversity of DNA barcodes. Single nuclei were then isolated using fluorescence-activated cell sorting (FACS) and used directly as DNA templates for barcode profiling (Figure 4A; see the STAR Methods section). As expected, we observed a significant increase in the number of nuclei displaying XVE-Cas9-induced lesions, as well as the types of edited DNA barcodes in the β-estradiol-treated samples compared with the non-treatment control (Figures 4B and S5), confirming the feasibility of our method.

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

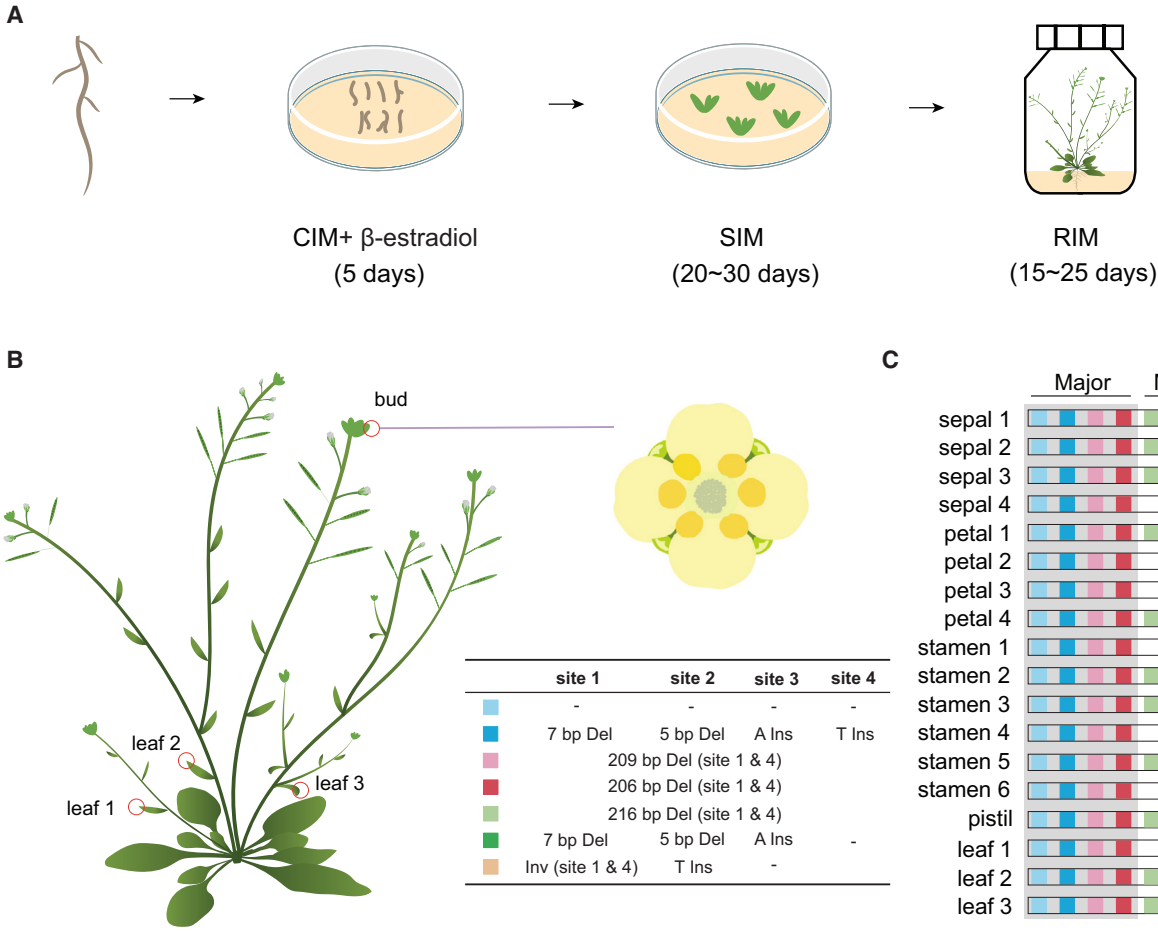


Figure 3. Characterization of DNA barcodes in the regenerated plant

(A) Schematic representation of the regeneration process and the timing of β-estradiol treatment. Root explants were cut and placed on CIM supplemented with 100 μM β-estradiol. After 5 days, the calli were transferred to SIM without β-estradiol to induce shoot formation.

(B) A schematic illustration of the regenerated plant indicating the sampling position. Three leaves and the floral organs dissected from a single flower were collected for barcode profiling.

(C) DNA barcodes identified from various organs collected from different positions of the regenerated plant as illustrated in (B). Each colored square block represents a distinct type of DNA barcode containing sequence information of four GCN4 sites. The corresponding sequence details at each GCN4 site are described in the table. The shaded area indicates major type DNA barcodes detected within each organ, with corresponding read frequency greater than 5%. DNA barcodes with corresponding read frequency less than 5% are denoted as minor types. The read frequency for each DNA barcode was calculated by normalizing the number of reads representing that DNA barcode to the total number of mappable reads of each sample. Del, deletion; Ins, insertion; Inv, inversion. See also Figure S4.

As for the regenerated plants, single nuclei were isolated from eight different organs (ten nuclei each), resulting in a total of eighty nuclei per plant (Figures 4C and 4D). Similar to our observation using different organs, barcode profiling of these isolated nuclei suggested a highly similar distribution pattern regardless of their tissue origin (Figures 4C and 4D). These results were consistent when the experiments were repeated using independently regenerated plants (Figures S4E and S4F). Our results strongly indicated that cells from various organs in the regenerated plants originate from a single ancestor cell in the root explants.

Fluorescence-based lineage analysis in live cells

In addition to the DNA barcoding method, we aimed to delineate the cell origin of regenerated *Arabidopsis* using traceable multi-fluorescence signals through live cell imaging. Therefore, we adopted and modified the Brainbow strategy (Figure 5A; see the STAR Methods), where individual cells can be randomly labeled with distinct fluorescence through Cre-triggered site-specific *loxP* recombination.[21,22] The temporal control of Cre recombinase was achieved by using the XVE regulatory module (*XVE-Cre*), as done for the *XVE-Cas9* constructs (Figure 5A). Prior to Cre induction, a 35S promoter drives the constitutive expression of *Phialidium sp*. yellow fluorescent protein (phiYFP). Upon Cre induction, the phiYFP expression will be replaced with that of a randomly chosen downstream fluorescence protein (XFP, fluorescent proteins collectively called XFP) and inherited mitotically (Figure 5A). Consequently, cell lineages in the regenerated plant tissues can be inferred based on their fluorescence labeling.

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

CellPress
OPEN ACCESS

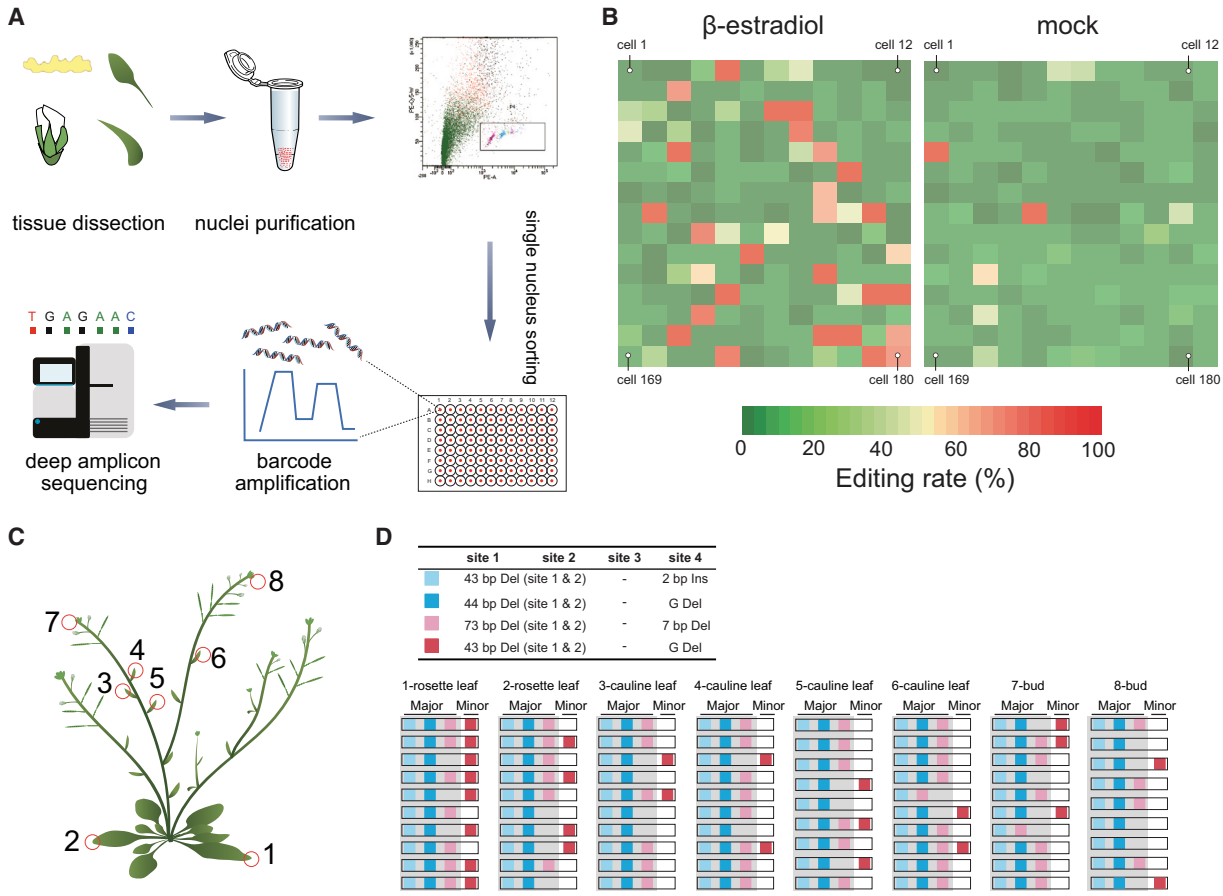


Figure 4. Single-nucleus profiling of DNA barcodes in the regenerated plant

(A) A schematic diagram depicts the experimental procedure of single-nuclei isolation and DNA barcode profiling. Samples were chopped in nuclei isolation buffer and subsequently stained with PI. Single nuclei were sorted into 96-well plates followed by DNA barcode amplification and deep-amplicon sequencing.

(B) The heatmap illustrates editing rates in each single nucleus isolated from calli treated with (left) or without (right) β-estradiol. Each square within the heatmap corresponds to an isolated single nucleus, and the color gradient represents the editing ratio as reflected by the percentage of sequenced reads containing genetic lesions.

(C) A schematic drawing of the regenerated plant indicating the sampling position.

(D) DNA barcodes identified from single nuclei isolated from different organs of the regenerated plant as illustrated in (C). A total of ten nuclei were analyzed for each organ. Each row represents a barcode combination in a single nucleus. Data convention is the same as in Figure 3C.

See also Figures S4 and S5.

Arabidopsis plants harboring both the *XVE-Cre* and *Brainbow* constructs (*XVE-Brainbow*) were generated. A uniform phiYFP fluorescence can be observed in the roots, indicating minimal, if any, Cre background expression (Figure S6A). Next, root explants were placed on CIM supplemented with β-estradiol for 5 days to induce Cre expression (Figure 3A). As expected, distinct fluorescence signals were observed in different root cells, including those emitted from either a single type of XFP or a stochastic combination of multiple XFPs (Figures 5B and S6B–S6F). We conducted temporal tracking for callus formation on CIM. Images of the same explants were captured from CIM days 0 to 5. On CIM day 0, cells exhibiting different fluorescent colors exhibited a random spatial distribution within the root explant (Figure 5B). Over the following days, as cells proliferated, clusters of cells displaying identical fluorescence began to form, tending to occupy specific regions within the explants (Figure 5B).

The root explants were subsequently transferred to standard SIM without β-estradiol to induce shoot formation. After an additional 5 days, clusters of cells exhibiting varied types of fluorescence and differential spatial distribution were observed within the callus (Figures 5C and S6G). Given that cells within each cluster displayed identical fluorescence, we deduced that these cells likely descended from the same founder cell, potentially contributing to the development of individual shoot primordia. If this hypothesis holds true, we anticipate that all cells in the later regenerated plants will also display identical fluorescence but not mosaic colors. We conducted temporal tracking of root explants on SIM during shoot formation. Since predicting the exact site of future shoot formation is challenging, we captured images of the entire explants every 2 days until new shoots emerged. As demonstrated in the representative case, by SIM day 7, two cell clusters, from which future shoots would originate, became visible (Figure 5D). Over the subsequent 4 days, these two cell

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

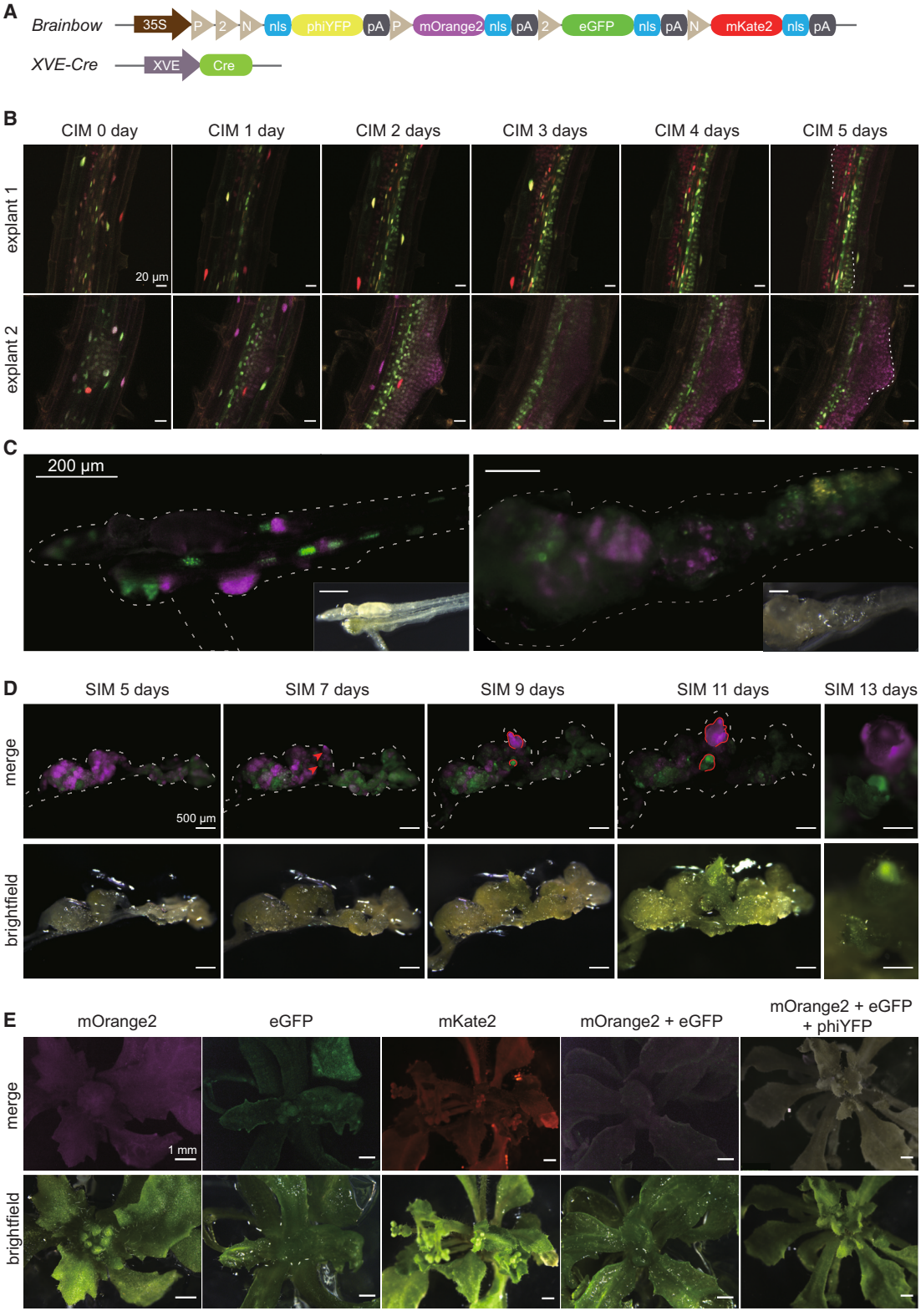


Figure 5. Multi-fluorescence-based lineage tracing of *Arabidopsis* regeneration

(A) Schematic illustration of the *XVE-Brainbow* multi-fluorescence labeling tool. P, *loxP* site; 2, *lox2272* site; N, *loxN* site; nls, nuclear localization signal; pA, polyadenylation signal. In the presence of Cre, DNA recombination occurs between a randomly chosen pair of *loxP* sites (P-, 2-, or N-type), leading to the deletion of the phiYFP coding sequence and expression of the remaining downstream fluorescence protein.

(legend continued on next page)

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

clusters gradually enlarged and became more distinct. By SIM day 13, a single shoot had formed in each of the two cell clusters, each displaying a uniform and distinct fluorescent color (Figure 5D). Despite the diverse fluorescence distribution pattern observed in *XVE-Brainbow* callus (Figures 5C and S6G), the regenerated plants (49 out of 49) uniformly exhibited the same fluorescence type across all tissue types examined (Figures 5D, 5E, S6H, and S6I). Collectively, these observations align with our earlier DNA barcode profiling using *Cas9-1-GCN4* plants, reinforcing the idea that regenerated *Arabidopsis* originates from a single progenitor cell in the root explants.

On SIM day 5, the emergence of cell clusters with identical fluorescence in callus suggests the possible initiation of division and differentiation of the single founder cell. If this holds true, extending the treatment of β-estradiol and *XVE-Brainbow* labeling beyond the founder cell differentiation initiation point may result in regenerated plants exhibiting mosaic fluorescence patterns. To examine this hypothesis, *XVE-Brainbow* explants were placed on regular CIM and subjected to β-estradiol treatment for 2 h on SIM at day 6 and then transferred to regular SIM (Figure S7A). Live tracing of the subsequent regeneration processes demonstrated the formation of chimeric regenerants with mixed fluorescent signals (Figure S7B). This phenomenon represents the majority of the examined plants (36 out of 41, Figure S7C). These patterns varied, with some plants exhibiting different fluorescence colors in distinct tissues and others showing mosaic fluorescence in a single leaf (Figure S7C). In summary, our findings indicate that during *Arabidopsis* regeneration, the differentiation of the single founder cell initiates before SIM day 5.

Cell-type-specific lineage tracing

Previous studies have shown that callus formation from root and hypocotyl explants recapitulates the developmental trajectory of lateral root (LR) initiation.[55,64] The initial stage is characterized by the emergence of LR meristem-like protuberances from 4 to 6 xylem pole pericycle (XPP) cells.[55] However, owing to swift cell fate transitions and the diffusion of the XPP marker, it has remained unclear how the descendants of XPP cells are spatially distributed within the callus and subsequently contribute to the formation of regenerated plants. It is also not clear if XPP cells constitute the sole source of the future regenerated plants.

To address this question, we developed a GFP reporter that can track divisions for specific cell types through CRISPR base editing (Figure 6A). In this design, the GFP reporter can be permanently activated in XPP cells and stably inherited in their descendant cells. This was realized by a C to T substitution at codon 66 of GFP catalyzed by the APOBEC3A-Cas9 nickase-uracil glycosylase inhibitor fusion protein (*A3A-PBE*), which is specifically expressed in XPP cells (Figure 6A).[58] To maintain temporal control over the GFP reporter, we utilized a β-estradiol-inducible sgRNA expression system (Figure 6A). The β-estradiol treatment was conducted in intact roots instead of root explants to ensure the preservation of XPP cell identities during GFP labeling.

Root explants from 7-day-old *Arabidopsis* seedlings co-transformed with the GFP reporter and the *A3A-PBE* construct were placed on 1/2 Murashige-Skoog (MS) medium supplemented with β-estradiol for 3 days to initiate XPP cell labeling. As anticipated, the majority of XPP cells exhibited GFP fluorescence upon β-estradiol treatment, while no such signal was observed before the treatment (Figure 6B). Subsequently, root explants were transferred to regular CIM without β-estradiol, and cells displaying GFP fluorescence were dynamically traced in the following days. The results indicated that periclinal division of XPP cells began on CIM day 2 (Figure 6B), and at least two rounds of cell division occurred by CIM day 5, forming up to four layers of GFP-positive cells (Figure 6B).

The calli were subsequently transferred to SIM, and the even distribution of GFP fluorescence along the root longitudinal axis was maintained until SIM day 5, when cell clusters occupying specific spatial domains began to form (Figure 6C), mirroring our earlier observation using *XVE-Brainbow* plants (Figure 5C). We then continuously live-traced these fluorescent cell clusters over the next 2 weeks until shoot formation (Figure 6C). Among the regenerated plants (n = 130), we observed both GFP-positive (n = 18) and GFP-negative (n = 112) individuals. The results confirmed XPP cells as an important source of future regenerants. However, the low ratio of GFP-positive plants among the regenerants (18 out of 130) indicated that XPP cells may not be the sole contributing source of plant regeneration. Nevertheless, no chimeric plants were observed, further supporting our discovery that each regenerated plant is derived from a single somatic cell in the root explants.

Cell-lineage analysis of LR development

The branching of LRs plays an essential role in shaping the root system architecture of vascular plants. This process exhibits high developmental plasticity, enabling plants to adapt to a variety of environmental conditions. Observing LR initiation has been difficult due to the challenges in predicting the precise initiation site of LRs, and various studies have proposed different numbers of pericycle cells involved in LR initiation.[55,65–67] To demonstrate the utility of our lineage-tracing tools in studying other plant development processes, we employed our *XVE-Brainbow* system to gain insights into LR initiation.

We treated 4-day-old seedlings carrying the *XVE-Brainbow* transgene with β-estradiol for 12 h (Figure 7A). This treatment enabled the marking of primary root cells with distinct fluorescence before LR initiation (Figure 7B). LRs began to emerge on

(B) Time-lapse tracing for callus formation on CIM. Images of the same explants were captured from CIM days 0 to 5. The white dotted lines outline the periphery of proliferated cell clusters displaying identical fluorescence. Scale bars: 20 μm.
(C) Fluorescence stereo microscopy of callus showing the emergence of cell clusters displaying identical fluorescence on SIM at day 5. The white dotted lines outline the edges of the callus. The bright-field picture is shown at the bottom right corner in each image. Scale bars: 200 μm.
(D) Time-lapse tracing for shoot formation on SIM. Images of the same callus were captured from SIM days 5 to 13. Cell clusters from which future shoots emerge are indicated by red arrowheads. The red and white dotted lines outline the edges of the newly formed shoots and the callus, respectively. Scale bars: 500 μm.
(E) Representative examples of regenerated plants exhibiting identical fluorescence signals across their entirety. The annotations above each panel specify the type of fluorescence emitted by the respective plants. Scale bars: 1 mm.
See also Figures S6 and S7.

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

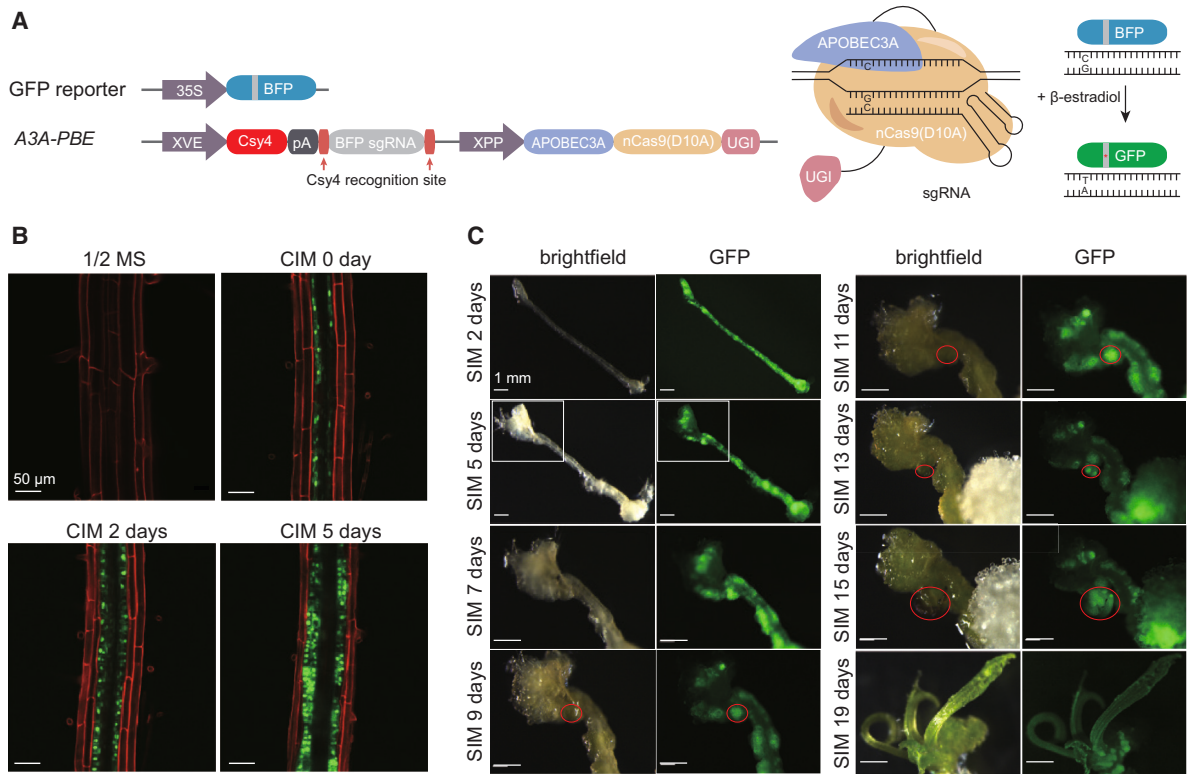


Figure 6. Lineage tracing of XPP during *Arabidopsis* regeneration

(A) Schematic illustration of cell-type-specific lineage tracing using the GFP reporter targeted by an inducible *A3A-PBE* construct.

(B) Confocal microscopy of root explants shows XPP cell-lineage tracing during callus formation. Cell walls are stained with PI. The nuclear-localized GFP fluorescence indicates XPP cells and their descendants. Scale bars: 50 μm.

(C) Fluorescence stereo microscopy shows XPP cell-lineage tracing during shoot formation. From day 7 onward, the partial enlargement of the region inside the white box is shown. The red circles indicate a GFP-positive region that shows the emergence of shoot primordia and leaf regeneration. Scale bars: 1 mm.

day 6. Upon reaching 10 days of age, we examined the fluorescence in the newly formed LRs. In almost all cases (175 out of 196, 89.3%), multiple fluorescent colors were observed in each LR (Figures 7C–7I). Cells displaying different fluorescent colors, typically 2–4 types, were found in the LR meristematic region (Figures 7C–7I). This suggests that *Arabidopsis* LR initiation likely involves multiple (2–4) pericycle cells. Cells of the same fluorescence were longitudinally aligned (Figures 7C–7I), indicating their direct lineage relationship. Meanwhile, the fluorescence exhibited in the LR quiescent center (QC) and its surrounding SCN tends to occupy a more extensive portion of the LR. This finding is consistent with the previous study, which indicates that the establishment of the LR primordium is initiated by a single founder cell and its gradual recruitment of neighboring cells to become founder cells.[67]

DISCUSSION

In this study, we explored the lineage relationship of *Arabidopsis* regeneration through the development of multiple complementary toolsets, encompassing both CRISPR barcoding and multi-color fluorescence labeling strategies. The integration of these methodologies yielded converging lines of evidence supporting a single-cell origin model for *Arabidopsis* organogenesis. Specifically, a single somatic cell from the root explant serves as the exclusive progenitor for the subsequent regeneration of each individual plant.

Our observation of cell clusters exhibiting identical fluorescence on SIM at day 5 suggests that the fate of the founder cell may have already been determined at this juncture (Figures 5C and S6G). Delaying the β-estradiol treatment until SIM day 6 led to regenerated plants displaying mosaic fluorescence in different organs, enabling us to pinpoint the latest time point for founder cell differentiation (Figure S7). *WUSCHEL* (WUS) is the key transcription factor in establishing shoot identity.[68] A previous study has indicated that *WUS* expression was identified in dispersed single cells after explants were placed on SIM for 3 days.[69] The presence of *WUS*-expressing single cells may signify the earliest time point at which the fate of the regenerative founder cell is established. Taken together, these findings consistently indicate that *Arabidopsis* regenerative founder cells are established and commence differentiation between SIM days 3 and 5.

In the CRISPR barcode assay, deep-amplicon sequencing at over 1,500-fold coverage has greatly enhanced our ability to unveil all essential types of genetic lesions present in the DNA sample. While each type of genetic lesion can be quantitatively determined based on the number of sequenced reads, caution is exercised when using this parameter as an indicator for their relative enrichments in the original bulk DNA extracts. This caution arises because DNA barcodes of varied sequence

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

CellPress
OPEN ACCESS

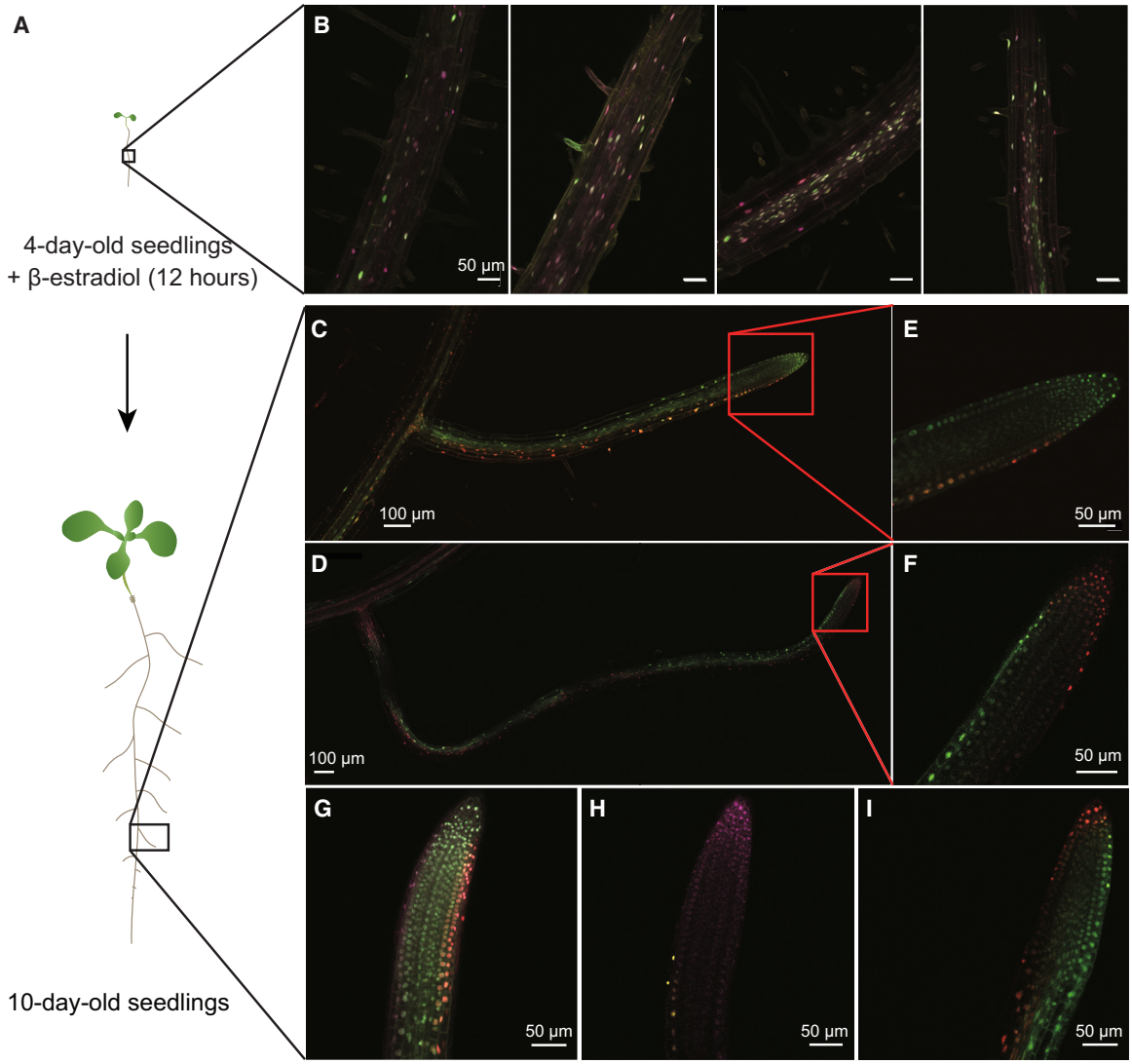


Figure 7. Cell-lineage analysis of lateral root development

(A) A schematic diagram illustrates the experimental procedure for the initial labeling of primary root cells and the subsequent examination of LR cells.
(B) Four representative samples show the detection of fluorescence in labeled primary root cells prior to the onset of LR formation. Scale bars: 50 μm.
(C–I) Examination of fluorescence in the newly formed LRs. The entire LRs (C and D) and the enlarged meristematic regions (E–I) are displayed. Scale bars: 100 μm (C and D), 50 μm (E–I).

compositions and lengths may be differentially amplified. This is evident from the analysis using single nuclei, where an unequal representation of different DNA barcodes was observed even though each type of DNA barcode is expected to be of single copy (Figure S5).

For the same reason, certain types of difficult-to-amplify DNA barcodes may be underrepresented in the sequencing results (minor barcodes) and can also be occasionally missed in certain individual samples, especially when the amount of DNA input is low, as in the case of isolated single nuclei (Figures 4D and S4E). Nevertheless, the genetic lesions representing major barcodes are almost uniformly shared among cells collected from the same regenerated plant, and we have never observed a single nucleus displaying a unique type of DNA barcode (Figures 4D, S4E, and S4F). Taking all the above reasoning into account and considering the evidence from the *XVE-Brainbow* tracing experiments, we conclude that the regenerated *Arabidopsis* plants are of single-cell origin.

A relatively small number of barcode types were identified in each regenerated plant, ranging from 2 to 8 (Figures 3, 4, and S4). By contrast, there is a broad distribution in the number of barcode types identified in each callus, ranging from 9 to 69 (Figure 2G), indicating a more heterogeneous cell composition in the callus. Supporting this notion, live tracing of the mitotic division of XPP cell lineages showed a roughly uniform periclinal cell division along the root longitudinal axis (Figure 6B). Single-cell sequencing has indicated distinct cell identities inside the callus, where middle layer cells exhibiting root QC identities will eventually generate shoots and roots.[70] In our study, the highly diverse composition of cell identities in the callus was

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

confirmed by DNA barcode profiling using isolated single nuclei (Figure S5). Therefore, while the callus is of multicellular origin, we did not observe evidence of deploying cells of mixed origins in the formation of each individually regenerated plant. This result suggests a scenario where the individual cell of better fitness compared with its neighboring cells was chosen to become the founder cell of the future regenerated plant, employing a cell-cell competition strategy to optimize the chances of survival. Supporting this notion, cell competition also appears to play key and prevalent roles in varying developmental processes in animals, including embryogenesis, tissue homeostasis, tumorigenesis, aging, etc.[71,72]

In addition to roots, plant organogenesis can be induced from callus derived from various types of somatic tissues, such as cotyledons and leaves.[64] Since the process of callus formation from these other types of organs also mirrors the development of LR initiation,[64] there is a compelling reason to believe that organogenesis from other somatic tissues may similarly originate from a single cell. Moreover, beyond organogenesis, plants possess the ability to regenerate through somatic embryos, which can be induced either directly from somatic cells or indirectly through embryonic callus.[73,74] Considering that microspores of many plants can be treated *in vitro* to generate haploid plants,[75] we propose that somatic embryogenesis may also be traced back to a single cell. Further experiments are necessary to definitively confirm this hypothesis.

Based on macroscopic observations, it has been reported that *Arabidopsis* plants regenerated from root and hypocotyl explants originate from XPP cells.[55] Through cell-type-specific labeling and lineage tracing, we identified regenerated plants that are GFP-positive, providing direct evidence that these plants are descendants of the GFP-labeled XPP cells in the root explants (Figures 6B and 6C). Interestingly, a significant portion of *Arabidopsis* regenerants is GFP-negative. This could be attributed to the fact that not all XPP cells are successfully GFP-labeled. Alternatively, other cell types besides XPP cells may also be involved in regeneration, such as pericycle cells that do not exhibit XPP promoter activity. Future investigations employing promoters with a broader expression domain among pericycle cells will be of interest. Through time-series single-cell transcriptome profiling during callus formation, researchers have identified two transitional cell types derived from XPP cells that are responsible for initiating callus formation. These include LR primordium-initiating (LRPI)-like cells and QC-like cells.[76]

In addition to LR development, the aerial parts of plants also pose many questions worth exploring, such as the establishment of the plant germ line and the initiation of axillary buds. While observing cell lineages using microscopy techniques may be limited by the size of the organ and the intensity of autofluorescence, our CRISPR barcode system may be more suitable for capturing lineage information of plant aerial organs. Therefore, our toolkit holds promise for addressing a multitude of developmental questions in plants. Researchers can select a toolkit based on their specific research purposes.

Limitations of the study

In our study, large fragment deletions (LFDs) are frequently observed between two GCN4 sites, reducing the throughput and diversity of possible barcodes. For example, in a pooled sample of 20 calli, we identified a total of 227 unique barcode types, of which 186 (81.9%) featured an LFD. Therefore, barcoding strategies that avoid creating DNA double-strand breaks, such as base editing, may further improve the efficiency of the lineage tracing system. Due to the length limitation of deep-amplicon sequencing, our barcode length was restricted to a maximum of four GCN4 sites. Therefore, coupling barcode profiling with other long-read sequencing platforms to include more GCN4 sites will further enhance the resolution of our barcoding system.

In addition, we found that the Brainbow system is not suitable for studying cell lineages in aged plant tissues due to the tissue thickness and the high intensity of autofluorescence emitted by chlorophyll and lignin. Further optimization of imaging techniques and plant tissue transparency treatments will expand its use in more tissue types in the future.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Qikun Liu (qikunliu@pku.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data reported in this work will be shared by the lead contact upon reasonable request.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

This research was supported by the State Key Laboratory for Protein and Plant Gene Research and the startup funds from the School of Advanced Agricultural Sciences at Peking University. This study was supported by grants 32200469 and 32070562 from the National Natural Science Foundation of China. The authors thank the Flow Cytometry Core at the National Center for Protein Sciences at Peking University, particularly Dr. Yinghua Guo, for technical help. We also thank Dr. Fuchou Tang (Peking University, China) and Dr. Yuan Gao for technical guidance and Dr. Xiaofeng Cao (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences), Dr. Long Wang (Nanjing University), Dr. Chongyi Xu and Dr. Yuxin Hu (Institute of Botany, Chinese Academy of Sciences), Dr. Zhaojun Ding (Shandong University), Dr. Yuling Jiao, and Dr. Yuehui He (Peking University) for their critical reading and discussions of the manuscript.

AUTHOR CONTRIBUTIONS

Q.L. conceived and oversaw the study. H.Z. advised on the experimental design. X.L., Q.Z., X.C., S.C., and H.Y. performed the experiments. Z.W., X.L., and X.C. analyzed the data. Q.L. and X.L. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS

- Tissue culture
- Plasmid construction
- β-estradiol treatment
- Amplification of DNA barcodes and deep-amplicon sequencing
- Quantitative real-time PCR
- Single-nucleus isolation using Fluorescence-activated Cell Sorting (FACS)
- Single-cell DNA barcode amplification and sequencing
- ClearSee and microscopy

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.devcel.2024.10.023.

Received: January 27, 2024
Revised: July 27, 2024
Accepted: October 30, 2024
Published: November 25, 2024

REFERENCES

1. Yang, S., Yu, Q., Zhang, Y., Jia, Y., Wan, S., Kong, X., and Ding, Z. (2018). ROS: The Fine-Tuner of Plant Stem Cell Fate. Trends Plant Sci. *23*, 850–853. https://doi.org/10.1016/j.tplants.2018.07.010.
2. Liao, R.Y., and Wang, J.W. (2023). Analysis of meristems and plant regeneration at single-cell resolution. Curr. Opin. Plant Biol. *74*, 102378. https://doi.org/10.1016/j.pbi.2023.102378.
3. Neumann, M., Xu, X., Smaczniak, C., Schumacher, J., Yan, W., Blüthgen, N., Greb, T., Jönsson, H., Traas, J., Kaufmann, K., et al. (2022). A 3D gene expression atlas of the floral meristem based on spatial reconstruction of single nucleus RNA sequencing data. Nat. Commun. *13*, 2838. https://doi.org/10.1038/s41467-022-30177-y.
4. Satterlee, J.W., Strable, J., and Scanlon, M.J. (2020). Plant stem-cell organization and differentiation at single-cell resolution. Proc. Natl. Acad. Sci. USA *117*, 33689–33699. https://doi.org/10.1073/pnas.2018788117.
5. Shahan, R., Nolan, T.M., and Benfey, P.N. (2021). Single-cell analysis of cell identity in the Arabidopsis root apical meristem: insights and opportunities. J. Exp. Bot. *72*, 6679–6686. https://doi.org/10.1093/jxb/erab228.
6. Zhang, T.Q., Xu, Z.G., Shang, G.D., and Wang, J.W. (2019). A Single-Cell RNA Sequencing Profiles the Developmental Landscape of Arabidopsis Root. Mol. Plant *12*, 648–660. https://doi.org/10.1016/j.molp.2019.04.004.
7. Zhang, T.Q., Chen, Y., and Wang, J.W. (2021). A single-cell analysis of the Arabidopsis vegetative shoot apex. Dev. Cell *56*, 1056–1074.e8. https://doi.org/10.1016/j.devcel.2021.02.021.
8. Zong, J., Wang, L., Zhu, L., Bian, L., Zhang, B., Chen, X., Huang, G., Zhang, X., Fan, J., Cao, L., et al. (2022). A rice single cell transcriptomic atlas defines the developmental trajectories of rice floret and inflorescence meristems. New Phytol. *234*, 494–512. https://doi.org/10.1111/nph.18008.
9. Baron, C.S., and van Oudenaarden, A. (2019). Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat. Rev. Mol. Cell Biol. *20*, 753–765. https://doi.org/10.1038/s41580-019-0186-3.
10. Kretzschmar, K., and Watt, F.M. (2012). Lineage tracing. Cell *148*, 33–45. https://doi.org/10.1016/j.cell.2012.01.002.
11. Yao, M., Ren, T., Pan, Y., Xue, X., Li, R., Zhang, L., Li, Y., and Huang, K. (2022). A New Generation of Lineage Tracing Dynamically Records Cell Fate Choices. Int. J. Mol. Sci. *23*, 5021. https://doi.org/10.3390/ijms23095021.
12. Irish, V.F., and Sussex, I.M. (1992). A fate map of the Arabidopsis embryonic shoot apical meristem. Development *115*, 745–753. https://doi.org/10.1242/dev.115.3.745.
13. Jegla, D.E., and Sussex, I.M. (1989). Cell lineage patterns in the shoot meristem of the sunflower embryo in the dry seed. Dev. Biol. *131*, 215–225. https://doi.org/10.1016/s0012-1606(89)80053-3.
14. Satina, S., Blakeslee, A.F., and Avery, A.G. (1940). Demonstration of the three germ layers in the shoot apex of Datura by means of induced polyploidy in periclinal chimeras. Am. J. Bot. *27*, 895–905. https://doi.org/10.1002/J.1537-2197.1940.TB13952.X.
15. Smetana, O., Mäkilä, R., Lyu, M., Amiryousefi, A., Sánchez Rodríguez, F., Wu, M.F., Solé-Gil, A., Leal Gavarrón, M., Siligato, R., Miyashima, S., et al. (2019). High levels of auxin signalling define the stem-cell organizer of the vascular cambium. Nature *565*, 485–489. https://doi.org/10.1038/s41586-018-0837-0.
16. Frieda, K.L., Linton, J.M., Hormoz, S., Choi, J., Chow, K.K., Singer, Z.S., Budde, M.W., Elowitz, M.B., and Cai, L. (2017). Synthetic recording and in situ readout of lineage information in single cells. Nature *541*, 107–111. https://doi.org/10.1038/nature20777.
17. He, L., Pu, W., Liu, X., Zhang, Z., Han, M., Li, Y., Huang, X., Han, X., Li, Y., Liu, K., et al. (2021). Proliferation tracing reveals regional hepatocyte generation in liver homeostasis and repair. Science *371*, eabc4346. https://doi.org/10.1126/science.abc4346.
18. Morita, R., Sanzen, N., Sasaki, H., Hayashi, T., Umeda, M., Yoshimura, M., Yamamoto, T., Shibata, T., Abe, T., Kiyonari, H., et al. (2021). Tracing the origin of hair follicle stem cells. Nature *594*, 547–552. https://doi.org/10.1038/s41586-021-03638-5.
19. Donà, M., Bradamante, G., Bogojevic, Z., Gutzat, R., Streubel, S., Mosiolek, M., Dolan, L., and Mittelsten Scheid, O. (2023). A versatile CRISPR-based system for lineage tracing in living plants. Plant J. *115*, 1169–1184. https://doi.org/10.1111/tpj.16378.
20. Tsuda, K., Maeno, A., and Nonomura, K.I. (2023). Heat shock-inducible clonal analysis reveals the stepwise establishment of cell fate in the rice stem. Plant Cell *35*, 4366–4382. https://doi.org/10.1093/plcell/koad241.
21. Cai, D., Cohen, K.B., Luo, T., Lichtman, J.W., and Sanes, J.R. (2013). Improved tools for the Brainbow toolbox. Nat. Methods *10*, 540–547. https://doi.org/10.1038/nmeth.2450.
22. Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature *450*, 56–62. https://doi.org/10.1038/nature06293.
23. Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K., Wu, H., et al. (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell *148*, 886–895. https://doi.org/10.1016/j.cell.2012.02.025.
24. McConnell, M.J., Lindberg, M.R., Brennand, K.J., Piper, J.C., Voet, T., Cowing-Zitron, C., Shumilina, S., Lasken, R.S., Vermeesch, J.R., Hall, I.M., et al. (2013). Mosaic copy number variation in human neurons. Science *342*, 632–637. https://doi.org/10.1126/science.1243472.
25. Hwang, B., Lee, W., Yum, S.-Y., Jeon, Y., Cho, N., Jang, G., and Bang, D. (2019). Lineage tracing using a Cas9-deaminase barcoding system targeting endogenous L1 elements. Nat. Commun. *10*, 1234. https://doi.org/10.1038/s41467-019-09203-z.
26. Ludwig, L.S., Lareau, C.A., Ulirsch, J.C., Christian, E., Muus, C., Li, L.H., Pelka, K., Ge, W., Oren, Y., Brack, A., et al. (2019). Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. Cell *176*, 1325–1339.e22. https://doi.org/10.1016/j.cell.2019.01.022.
27. Ren, Y., He, Z., Liu, P., Traw, B., Sun, S., Tian, D., Yang, S., Jia, Y., and Wang, L. (2021). Somatic Mutation Analysis in Salix suchowensis Reveals Early-Segregated Cell Lineages. Mol. Biol. Evol. *38*, 5292–5308. https://doi.org/10.1093/molbev/msab286.
28. Junker, J., Spanjaard, B., Peterson-Maduro, J., Alemany, A., Hu, B., Florescu, M., and van Oudenaarden, A. (2017). Massively parallel clonal analysis using CRISPR. Cas9 induced genetic scars. Preprint at bioRxiv. https://doi.org/10.1101/056499.
29. Li, L., Bowling, S., McGeary, S.E., Yu, Q., Lemke, B., Alcedo, K., Jia, Y., Liu, X., Ferreira, M., Klein, A.M., et al. (2023). A mouse model with high

clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. Cell *186*, 5183–5199.e22. https://doi.org/10.1016/j.cell.2023.09.019.

30. Liu, K., Deng, S., Ye, C., Yao, Z., Wang, J., Gong, H., Liu, L., and He, X. (2021). Mapping single-cell-resolution cell phylogeny reveals cell population dynamics during organ development. Nat. Methods *18*, 1506–1514. https://doi.org/10.1038/s41592-021-01325-x.
31. McKenna, A., Findlay, G.M., Gagnon, J.A., Horwitz, M.S., Schier, A.F., and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science *353*, aaf7907. https://doi.org/10.1126/science.aaf7907.
32. Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J.P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. Nat. Biotechnol. *36*, 469–473. https://doi.org/10.1038/nbt.4124.
33. Alemany, A., Florescu, M., Baron, C.S., Peterson-Maduro, J., and Van Oudenaarden, A. (2018). Whole-organism clone tracing using single-cell sequencing. Nature *556*, 108–112. https://doi.org/10.1038/nature25969.
34. Raj, B., Wagner, D.E., McKenna, A., Pandey, S., Klein, A.M., Shendure, J., Gagnon, J.A., and Schier, A.F. (2018). Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. Nat. Biotechnol. *36*, 442–450. https://doi.org/10.1038/nbt.4103.
35. Perli, S.D., Cui, C.H., and Lu, T.K. (2016). Continuous genetic recording with self-targeting CRISPR-Cas in human cells. Science *353*, aag0511. https://doi.org/10.1126/science.aag0511.
36. Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., and Church, G.M. (2018). Developmental barcoding of whole mouse via homing CRISPR. Science *361*, eaat9804. https://doi.org/10.1126/science.aat9804.
37. Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B., Jost, M., Quinn, J.J., Yang, D., Jones, M.G., et al. (2019). Molecular recording of mammalian embryogenesis. Nature *570*, 77–82. https://doi.org/10.1038/s41586-019-1184-5.
38. Biddy, B.A., Kong, W., Kamimoto, K., Guo, C., Waye, S.E., Sun, T., and Morris, S.A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature *564*, 219–224. https://doi.org/10.1038/s41586-018-0744-4.
39. You, Z., Wang, L., He, H., Wu, Z., Zhang, X., Xue, S., Xu, P., Hong, Y., Xiong, M., Wei, W., et al. (2023). Mapping of clonal lineages across developmental stages in human neural differentiation. Cell Stem Cell *30*, 473–487.e9. https://doi.org/10.1016/j.stem.2023.02.007.
40. Sena, G., Wang, X., Liu, H.Y., Hofhuis, H., and Birnbaum, K.D. (2009). Organ regeneration does not require a functional stem cell niche in plants. Nature *457*, 1150–1153. https://doi.org/10.1038/nature07597.
41. Wagner, D.E., Wang, I.E., and Reddien, P.W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. Science *332*, 811–816. https://doi.org/10.1126/science.1203983.
42. Ikeuchi, M., Favero, D.S., Sakamoto, Y., Iwase, A., Coleman, D., Rymen, B., and Sugimoto, K. (2019). Molecular mechanisms of plant regeneration. Annu. Rev. Plant Biol. *70*, 377–406. https://doi.org/10.1146/annurev-arplant-050718-100434.
43. Ishikawa, M., Murata, T., Sato, Y., Nishiyama, T., Hiwatashi, Y., Imai, A., Kimura, M., Sugimoto, N., Akita, A., Oguri, Y., et al. (2011). Physcomitrella cyclin-dependent kinase A links cell cycle reactivation to other cellular changes during reprogramming of leaf cells. Plant Cell *23*, 2924–2938. https://doi.org/10.1105/tpc.111.088005.
44. Fehér, A. (2015). Somatic embryogenesis - Stress-induced remodeling of plant cell fate. Biochim. Biophys. Acta *1849*, 385–402. https://doi.org/10.1016/j.bbagrm.2014.07.005.
45. Ikeuchi, M., Sugimoto, K., and Iwase, A. (2013). Plant callus: mechanisms of induction and repression. Plant Cell *25*, 3159–3173. https://doi.org/10.1105/tpc.113.116053.
46. Valvekens, D., Van Montagu, M., and Van Lijsebettens, M. (1988). Agrobacterium tumefaciens-mediated transformation of Arabidopsis thaliana root explants by using kanamycin selection. Proc. Natl. Acad. Sci. USA *85*, 5536–5540. https://doi.org/10.1073/pnas.85.15.5536.
47. Xu, C., Chang, P., Guo, S., Yang, X., Liu, X., Sui, B., Yu, D., Xin, W., and Hu, Y. (2023). Transcriptional activation by WRKY23 and derepression by removal of bHLH041 coordinately establish callus pluripotency in Arabidopsis regeneration. Plant Cell *36*, 158–173. https://doi.org/10.1093/plcell/koad255.
48. Zhao, X.Y., Su, Y.H., Cheng, Z.J., and Zhang, X.S. (2008). Cell fate switch during in vitro plant organogenesis. J. Integr. Plant Biol. *50*, 816–824. https://doi.org/10.1111/j.1744-7909.2008.00701.x.
49. Skoog, F., and Miller, C.O. (1957). Chemical regulation of growth and organ formation in plant tissues cultured in vitro. Symp. Soc. Exp. Biol. *11*, 118–130.
50. Zhao, Z., Andersen, S.U., Ljung, K., Dolezal, K., Miotk, A., Schultheiss, S.J., and Lohmann, J.U. (2010). Hormonal control of the shoot stem-cell niche. Nature *465*, 1089–1092. https://doi.org/10.1038/nature09126.
51. Efroni, I., Mello, A., Nawy, T., Ip, P.L., Rahni, R., DelRose, N., Powers, A., Satija, R., and Birnbaum, K.D. (2016). Root Regeneration Triggers an Embryo-like Sequence Guided by Hormonal Interactions. Cell *165*, 1721–1733. https://doi.org/10.1016/j.cell.2016.04.046.
52. Zhai, N., and Xu, L. (2020). CRE/LOX-based analysis of cell lineage during root formation and regeneration in Arabidopsis. aBIOTECH *1*, 153–156. https://doi.org/10.1007/s42994-020-00025-y.
53. Takebe, I., Labib, G., and Melchers, G. (1971). Regeneration of whole plants from isolated mesophyll protoplasts of tobacco. Naturwissenschaften *58*, 318–320. https://doi.org/10.1007/bf00624737.
54. Xu, M., Du, Q., Tian, C., Wang, Y., and Jiao, Y. (2021). Stochastic gene expression drives mesophyll protoplast regeneration. Sci. Adv. *7*, eabg8466. https://doi.org/10.1126/sciadv.abg8466.
55. Atta, R., Laurens, L., Boucheron-Dubuisson, E., Guivarc'h, A., Carnero, E., Giraudat-Pautot, V., Rech, P., and Chriqui, D. (2009). Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. Plant J. *57*, 626–644. https://doi.org/10.1111/j.1365-313X.2008.03715.x.
56. Stewart, R.N., and Dermen, H. (1970). Determination of number and mitotic activity of shoot apical initial cells by analysis of mericlinal chimeras. Am. J. Bot. *57*, 816–826. https://doi.org/10.1002/J.1537-2197.1970.TB09877.X.
57. Wang, G., Zhang, Y., Li, C., Wang, X., and Fletcher, J.C. (2022). Signaling peptides direct the art of rebirth. Trends Plant Sci. *27*, 516–519. https://doi.org/10.1016/j.tplants.2022.03.009.
58. Zong, Y., Song, Q., Li, C., Jin, S., Zhang, D., Wang, Y., Qiu, J.L., and Gao, C. (2018). Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. Nat. Biotechnol. *36*, 950–953. https://doi.org/10.1038/nbt.4261.
59. Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. Cell *159*, 635–646. https://doi.org/10.1016/j.cell.2014.09.039.
60. Papikian, A., Liu, W., Gallego-Bartolomé, J., and Jacobsen, S.E. (2019). Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. Nat. Commun. *10*, 729. https://doi.org/10.1038/s41467-019-08736-7.
61. Tang, S., Yang, C., Wang, D., Deng, X., Cao, X., and Song, X. (2022). Targeted DNA demethylation produces heritable epialleles in rice. Sci. China Life Sci. *65*, 753–756. https://doi.org/10.1007/s11427-021-1974-7.
62. Yan, X., Hoek, T.A., Vale, R.D., and Tanenbaum, M.E. (2016). Dynamics of translation of single mRNA molecules in vivo. Cell *165*, 976–989. https://doi.org/10.1016/j.cell.2016.04.034.
63. Zuo, J., Niu, Q.W., and Chua, N.H. (2000). Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. *24*, 265–273. https://doi.org/10.1046/j.1365-313x.2000.00868.x.

64. Sugimoto, K., Jiao, Y., and Meyerowitz, E.M. (2010). Arabidopsis regeneration from multiple tissues occurs via a root development pathway. Dev. Cell *18*, 463–471. https://doi.org/10.1016/j.devcel.2010.02.004.

65. Kurup, S., Runions, J., Köhler, U., Laplaze, L., Hodge, S., and Haseloff, J. (2005). Marking cell lineages in living tissues. Plant J. *42*, 444–453. https://doi.org/10.1111/j.1365-313X.2005.02386.x.

66. Van Norman, J.M., Xuan, W., Beeckman, T., and Benfey, P.N. (2013). To branch or not to branch: the role of pre-patterning in lateral root formation. Development *140*, 4301–4310. https://doi.org/10.1242/dev.090548.

67. Torres-Martínez, H.H., Hernández-Herrera, P., Corkidi, G., and Dubrovsky, J.G. (2020). From one cell to many: Morphogenetic field of lateral root founder cells in Arabidopsis thaliana is built by gradual recruitment. Proc. Natl. Acad. Sci. USA *117*, 20943–20949. https://doi.org/10.1073/pnas.2006387117.

68. Su, Y.H., Zhou, C., Li, Y.J., Yu, Y., Tang, L.P., Zhang, W.J., Yao, W.J., Huang, R., Laux, T., and Zhang, X.S. (2020). Integration of pluripotency pathways regulates stem cell maintenance in the Arabidopsis shoot meristem. Proc. Natl. Acad. Sci. USA *117*, 22561–22571. https://doi.org/10.1073/pnas.2015248117.

69. Zhang, T.-Q., Lian, H., Zhou, C.-M., Xu, L., Jiao, Y., and Wang, J.-W. (2017). A two-step model for de novo activation of WUSCHEL during plant shoot regeneration. Plant Cell *29*, 1073–1087. https://doi.org/10.1105/tpc.16.00863.

70. Zhai, N., and Xu, L. (2021). Pluripotency acquisition in the middle cell layer of callus is required for organ regeneration. Nat. Plants *7*, 1453–1460. https://doi.org/10.1038/s41477-021-01015-8.

71. van Neerven, S.M., and Vermeulen, L. (2023). Cell competition in development, homeostasis and cancer. Nat. Rev. Mol. Cell Biol. *24*, 221–236. https://doi.org/10.1038/s41580-022-00538-y.

72. Yusupova, M., and Fuchs, Y. (2023). To not love thy neighbor: mechanisms of cell competition in stem cells and beyond. Cell Death Differ. *30*, 979–991. https://doi.org/10.1038/s41418-023-01114-3.

73. Horstman, A., Bemer, M., and Boutilier, K. (2017). A transcriptional view on somatic embryogenesis. Regeneration (Oxf) *4*, 201–216. https://doi.org/10.1002/reg2.91.

74. Yang, X., and Zhang, X. (2010). Regulation of somatic embryogenesis in higher plants. Crit. Rev. Plant Sci. *29*, 36–57. https://doi.org/10.1080/07352680903436291.

75. Touraev, A., and Heberle-Bors, E. (1999). Microspore embryogenesis and in vitro pollen maturation in tobacco. Methods Mol. Biol. *111*, 281–291. https://doi.org/10.1385/1-59259-583-9:281.

76. Yin, R., Chen, R., Xia, K., and Xu, X. (2024). A single-cell transcriptome atlas reveals the trajectory of early cell fate transition during callus induction in Arabidopsis. Plant Commun. *5*, 100941. https://doi.org/10.1016/j.xplc.2024.100941.

77. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

78. Edgar, R.C. (2022). Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat. Commun. *13*, 6968. https://doi.org/10.1038/s41467-022-34630-w.

79. Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLoS One *11*, e0163962. https://doi.org/10.1371/journal.pone.0163962.

80. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

81. Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. *16*, 735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x.

82. Tsutsui, H., and Higashiyama, T. (2017). pKAMA-ITACHI Vectors for Highly Efficient CRISPR/Cas9-Mediated Gene Knockout in Arabidopsis thaliana. Plant Cell Physiol. *58*, 46–56. https://doi.org/10.1093/pcp/pcw191.

83. Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.-L., Wang, D., and Gao, C. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat. Biotechnol. *35*, 438–440. https://doi.org/10.1038/nbt.3811.

84. Zhang, Q., Wang, Z., Lu, X., Yan, H., Zhang, H., He, H., Bischof, S., Harris, C.J., and Liu, Q. (2023). DDT-RELATED PROTEIN4-IMITATION SWITCH alters nucleosome distribution to relieve transcriptional silencing in Arabidopsis. Plant Cell *35*, 3109–3126. https://doi.org/10.1093/plcell/koad143.

85. Čermák, T., Curtin, S.J., Gil-Humanes, J., Čegan, R., Kono, T.J.Y., Konečná, E., Belanto, J.J., Starker, C.G., Mathre, J.W., Greenstein, R.L., et al. (2017). A multipurpose toolkit to enable advanced genome engineering in plants. Plant Cell *29*, 1196–1217. https://doi.org/10.1105/tpc.16.00922.

86. Andersen, T.G., Naseer, S., Ursache, R., Wybouw, B., Smet, W., De Rybel, B., Vermeer, J.E.M., and Geldner, N. (2018). Diffusible repression of cytokinin signalling produces endodermal symmetry and passage cells. Nature *555*, 529–533. https://doi.org/10.1038/nature25976.

87. Liu, Q., Wang, C., Jiao, X., Zhang, H., Song, L., Li, Y., Gao, C., and Wang, K. (2019). Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. Sci. China Life Sci. *62*, 1–7. https://doi.org/10.1007/s11427-018-9402-9.

88. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402–408. https://doi.org/10.1006/meth.2001.1262.

89. Ursache, R., Andersen, T.G., Marhavý, P., and Geldner, N. (2018). A protocol for combining fluorescent proteins with histological stains for diverse cell wall components. Plant J. *93*, 399–412. https://doi.org/10.1111/tpj.13784.

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli DH5a	N/A	N/A
Escherichia coli JM109	N/A	N/A
Agrobacterium tumefaciens GV3101	N/A	N/A
Chemicals, peptides, and recombinant proteins		
Murashige & Skoog basal salts medium with vitamins	Phyto Tech	Cat#M519
Murashige & Skoog basal salts medium	Phyto Tech	Cat#M524
Gamborg vitamin solution (1000x)	Phyto Tech	Cat#G219
Sucrose	Sigma-Aldrich	Cat#V900116
Plant-agar	Dingguo	Cat#DH010-1.1
Phytagel	Solarbio	Cat#P8170
β-estradiol	Coolaber	Cat#CE5114
2,4-dicholorophenoxyacetic acid	Caisson	Cat#D001
N-6-Benzyladenine	Caisson	Cat#B001
3-Indole butyric acid	BBI	Cat#A600725-0025
D-Biotin	Aladdin	Cat#B105433
Zeatin	BBI	Cat#A600747-0100
Gentamycin sulfate	Meilunbio	Cat#MB1331
Rifampicin	Meilunbio	Cat#MB1769
Spectinomycin Hydrochloride	Meilunbio	Cat#MB1497
Hygromycin B	Phyto Tech	Cat#H370
Glufosinate ammonium	Aladdin	Cat#G114499
MES hydrate	Sigma-Aldrich	Cat#6120
Triton X-100	Sigma-Aldrich	Cat#T8787
EDTA-Na2	Aladdin	Cat#E116429
β-mercaptoethanol	Sigma-Aldrich	Cat#M3148
KCl	Sigma-Aldrich	Cat#P9541
NaCl	Sigma-Aldrich	Cat#S5886
Spermine tetrahydrochloride	Sigma-Aldrich	Cat#S1141
Tris base	Sigma-Aldrich	Cat#V900483
Propidium iodide	Sigma-Aldrich	Cat#P4864
PBS	Sangon Biotech	Cat#B548117-0500
4% paraformaldehyde	Beyotime	Cat#P0099
Urea	Sigma-Aldrich	Cat#51456
Sodium deoxycholate	Sigma-Aldrich	Cat#30970
Xylitol	Sigma-Aldrich	Cat#X3375
Critical commercial assays		
2 x Phanta Max Master Mix	Vazyme	Cat#P525
CloneExpress II One Step Cloning Kit	Vazyme	Cat#C112
Mut Express II Fast Mutagenesis Kit V2	Vazyme	Cat#C214
2 x Taq Master Mix	CWBIO	Cat#CW0682
Experimental models: Organisms/strains		
Arabidopsis: Col-0	N/A	N/A
Arabidopsis: *XVE-Cas9-1/2*	This paper	N/A

(*Continued on next page*)

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Arabidopsis: *pGWB505-GCN4*	This paper	N/A
Arabidopsis: *Cas9-1/2-GCN4*	This paper	N/A
Arabidopsis: *XVE-Cre*	This paper	N/A
Arabidopsis: *Brainbow*	This paper	N/A
Arabidopsis: *XVE-Brainbow*	This paper	N/A
Arabidopsis: *pGWB505-BFP*	This paper	N/A
Arabidopsis: *A3A-PBE*	This paper	N/A
Oligonucleotides		
See Table S1	This paper	N/A
Recombinant DNA		
Plasmid *XVE-Cas9-1/2*	This paper	N/A
Plasmid *pGWB505-GCN4*	This paper	N/A
Plasmid *Cas9-1/2-GCN4*	This paper	N/A
Plasmid *XVE-Cre*	This paper	N/A
Plasmid *Brainbow*	This paper	N/A
Plasmid *XVE-Brainbow*	This paper	N/A
Plasmid *pGWB505-BFP*	This paper	N/A
Plasmid *A3A-PBE*	This paper	N/A
Software and algorithms		
Adobe Illustrator CC 2018	Adobe Acrobat	N/A
ImageJ	Schindelin et al.[77]	https://imagej.nih.gov/ij/
ZEN Microscopy Software	ZEISS	N/A
Origin Pro 2020b	OriginLab Corporation	N/A
SPSS 27.0	IBM	N/A
muscle (v3.8)	Edgar.[78]	https://www.drive5.com/muscle5/
seqtk (v1.3-r106)	Shen et al.[79]	https://github.com/lh3/seqtk
BWA v0.7.17-r1188	Li and Durbin.[80]	https://github.com/lh3/bwa
Other		
Confocal microscope LSM900	Carl Zeiss	N/A
M205 FCA stereo fluorescence microscope	Leica	N/A
BD FACS Aria SORP	BD Biosciences	N/A
Miracloth	Merck Millipore	Cat#475855
Falcon Cell Strainers	Corning	Cat#352340
4 Chamber glass bottom dishes	Cellvis	Cat#D35C4-20-0-N

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Arabidopsis thaliana of the Col-0 ecotype background were used in this study. For GCN4 DNA barcode analysis, the GCN4 transgenic lines were crossed with the *XVE-Cas9* transgenic lines (*Cas9-GCN4*). For fluorescence-based analysis, the final plants were obtained by co-infiltrating *Brainbow* and *XVE-Cre* constructs (*XVE-Brainbow*) or by co-infiltrating GFP reporter and *A3A-PBE* constructs. *Agrobacterium tumefaciens* strain GV3101 carrying the constructs was used for plant transformation via the standard floral dipping method.[81] The seedlings were grown on 1/2 MS solid medium and later on soil under a 16-h-light/8-h-darkness cycle at 22 °C. Table S1 contains the lists of genotyping primers used in this study.

METHOD DETAILS

Tissue culture

For tissue culture, seeds were grown on 1/2 MS medium plates placed vertically. Root explants of approximately 5 mm in length were obtained from 12-day-old seedlings and cultured on CIM (MS basal medium with vitamins, 3% w/v sucrose, 1 mg/L 2,4-dichlorophenoxyacetic acid, 0.05 mg/L N-6-Benzyladenine) under dark conditions at 22 °C for 5 days. The calli were then transferred to SIM (MS

basal medium without vitamins, 1% w/v sucrose, 0.05% w/v MES, 0.4 mg/L 3-Indole butyric acid (IBA), 1 mg/L Biotin, 1 mg/L Zeatin) for shoot regeneration, or RIM (MS basal medium without vitamins, 1% w/v sucrose, 0.05% w/v MES, 1 mg/L IBA) for root regeneration, and incubated under 16-h-light/8-h-darkness cycle at 22 °C.

Plasmid construction

To construct the GCN4 DNA barcode plasmid, the 10×GCN4 tandem array from the *SunTag VP64* vector[60] was amplified and recombined into *pDONR207* to generate *pDONR207-GCN4* using Exnase II (Vazyme, C112-01). It was further introduced into *pGWB505* through LR reaction (LR clonase, Invitrogen, 11791-100). To construct the *XVE-Cas9* plasmid, the coding sequence of Cas9 from *pKI1.1R*[82] was amplified and recombined into the XhoI-digested *pER8* vector[63] to obtain *pER8-Cas9*. Two oligomers consisting of the 4 bp-overhang sequence and the 20-bp GCN4 peptide-specific sgRNA sequence were phosphorylated with T4 polynucleotide kinase (NEB, #M0201) and hybridized. The hybridized primer was diluted 250-fold and then inserted into the BsaI-digested *AtU6-sgRNA* vector using T4 ligase (NEB, #M0202) to create a single sgRNA expression cassette. The eight sgRNA cassette was independently created. The first four sgRNA cassettes were assembled using the BioBrick method, with SpeI/NheI-SacI sites, to generate the 4×*AtU6-sgRNA* tandem array. The cassettes were then recombined into *pER8-Cas9* (*XVE-Cas9-1*). *XVE-Cas9-2*, which targets the second four GCN4 repeats, was similarly constructed.

The *Brainbow* plasmid was developed based on the *Brainbow 3.1* construct.[21] In the original Brainbow system, XFPs were farnesylated to facilitate their retention on cell membranes of neuron axons and dendrites, enhancing their visualization.[21] We modified the construct to create nuclear-localized fluorescence signals, which improved visualization in plant tissues. Four XFPs were amplified from the *Brainbow 3.1* plasmid using specific primers. Restriction enzyme digestion sites and nuclear localization signals were incorporated into the primers. The four fluorescent proteins containing nuclear localization signals were then digested, ligated, and inserted into the *pGWB501* vector, modified to be driven by the 35S promoter, creating *35S-pGWB501-XFPs*. Site-directed mutagenesis was performed on phiYFP (A65Y) to obtain the *35S-pGWB501-XFPs-muta* using the Mut Express II Fast Mutagenesis Kit V2 (vazyme, C214-01). The DNA sequence of Cre recombinase was amplified and inserted into the *pER8* vector to create *XVE-Cre*.

To construct the GFP reporter plasmid, the BFP sequence from *pUbi-BFP*[83] was cloned downstream of the 35S promoter of *pGWB505* using Exnase II (Vazyme, C112-01) to generate *pGWB505-BFP*. To construct the *A3A-PBE* plasmid, the Basta resistance sequence blpR from *DDR4-ZF*[84] were introduced into *pER8* to replace the Hygromycin resistance gene. The csy4 coding sequence from *pDIRECT_22C*[85] and the csy4 binding site containing gRNA scaffold linked by poly(A) were recombined downstream of the mini35S promoter of *pER8*. Finally, the A3A-PBE sequence[58] driven by the 2.5 kb promoter sequence upstream of the start codon of *XPP* (*AT4G30450*)[86] was cloned and introduced into the above vector to generate *pER8-csy4-gRNA-pXPP-A3A-PBE* (*A3A-PBE*).

DNA sequences amplified using PCR were confirmed by Sanger sequencing. Table S1 contains the lists of sgRNA oligomers and primers used for plasmid construction.

β-estradiol treatment

The response to β-estradiol induction was investigated using the roots explants of 12-day-old *Cas9-GCN4* transgenic seedlings. The explants were cultured on CIM or SIM supplemented with different concentrations of β-estradiol or an equal volume of DMSO ('mock' in figure panels). For DNA barcode analysis, the roots of 12-day-old seedlings from the transgenic lines *Cas9-GCN4* were used as explants and cultured on CIM supplemented with 100 μM β-estradiol. For Brainbow analysis, β-estradiol at concentrations ranging from 1 μM to 50 μM was tested, and no discernible differences in the efficiency of fluorescence labeling were seen among different treatments. Therefore, the results of the β-estradiol treatment group were collected regardless of the concentration applied. For lineage-specific tracing, 7-day-old seedlings from the transgenic line harboring both the GFP reporter and *A3A-PBE* construct were initially transferred to 1/2 MS medium supplemented with 10 μM β-estradiol for 3 days. Subsequently, the roots of these seedlings were used as explants for the regeneration assay. β-estradiol (Coolaber, CE5114) was prepared as a 100 mM stock solution in DMSO and stored at -80°C.

Amplification of DNA barcodes and deep-amplicon sequencing

The first-round PCR was performed to amplify the DNA barcodes in a 25 μL reaction volume containing less than 200 ng of genomic DNA, 0.4 μM of outer primers, and 12.5 μL of 2x Taq MasterMix (CWBIO, CW0682). The PCR products were diluted 10-fold and 2 μL of the diluted products were used for a second round of nested PCR, incorporating forward and reverse sequencing barcodes into the products.

For bulk samples, the PCR cycle number was set to 18 to ensure that the reaction was terminated within the linear range of amplification. Three independent PCR reactions were conducted for each DNA sample, and the resulting products were pooled into one tube. The final amplicons were submitted to the Hi-TOM platform for library construction and sequenced twice using the Illumina NextSeq 500 platform with 150-bp paired-end reads.[87] For single-cell samples, an independent PCR reaction with 25 cycles was performed for each sample. Table S1 contains the lists of first-round and nested primers used to amplify the GCN4 DNA barcode.

Quantitative real-time PCR

For quantitative real-time PCR experiments, samples were collected for RNA extraction using the Direct-zol RNA MiniPrep kit (Promega, LS1040). 500 ng of total RNA was then used for cDNA synthesis using the SuperScript III First-Strand Synthesis Supermix (Vazyme, R312-01). Quantitative real-time PCR was carried out using ChamQ Universal SYBR qPCR Master Mix (Vazyme, Q711-02).

Please cite this article in press as: Lu et al., Development of an inducible DNA barcoding system to understand lineage changes in *Arabidopsis* regeneration, Developmental Cell (2024), https://doi.org/10.1016/j.devcel.2024.10.023

The relative expression levels were determined using the $2^{-\Delta\Delta CT}$ calculation method with *ACTIN 7* as the internal control.[88] The amplification efficiency of each primer set was calculated according to the standard curve generated from a template dilution series. The primers used for quantitative real-time PCR are listed in Table S1.

Single-nucleus isolation using Fluorescence-activated Cell Sorting (FACS)

Rosette leaves, cauline leaves, and flower buds of the regenerated plant were collected. The tissues were chopped in 500 μL of nuclei isolation buffer LB01 (15 mM Tris, 2 mM Na_2EDTA, 0.5 mM spermine tetrahydrochloride, 80 mM KCl, 20 mM NaCl, 0.1% (v/v) Triton X-100, 15 mM β-mercaptoethanol, and adjust to pH 7.5 with 1 M NaOH). The samples were filtered using one layer of Miracloth and then centrifuged for 5 min (500 *g*, 4°C). The supernatant was removed, and the pellet was resuspended in 200 μL of nuclei isolation buffer LB01 without Triton X-100. PI was added to achieve a final concentration of 50 μg/mL. The stained nuclei were filtered using a 40 μm cell strainer before sorting.

A BD FACS Aria SORP (BD Biosciences) in single-cell mode was used to sort the isolated nuclei. The PI signals was achieved using a 535 nm laser. Col-0 nuclei were used as a reference to adjust the scatter gates. Subsequently, single nuclei droplets were randomly sorted into 96-well plates, each containing 5 μL of single cell lysis buffer. Eight tissue samples from a single regenerated plant, each consisting of ten cells, were individually analyzed. The experiment was repeated at least three times to characterize the barcode distribution in regenerated plants across different tissues.

Single-cell DNA barcode amplification and sequencing

The sorted single nuclei were lysed at 50°C for 3 h, followed by incubation at 75°C for 30 min. The lysates were then centrifuged at 9,000 x *g* for 1 min at 4°C. The 20 μL PCR mix without a template was directly added to the 5 μL cell lysate. The first-round PCR and nested PCR were performed as previously described. Deep amplicon sequencing was used to identify individual PCR products.

ClearSee and microscopy

Confocal microscopy of *XVE-Brainbow* transgenic lines was performed on root explants and calli using the Zeiss LSM900 with Airyscan2 confocal microscopy. Multiple fluorophores were separated using different channels with the GaAsP-PMT detector. All channels were acquired sequentially rather than simultaneously to minimize fluorescence cross-talk and optimize color separation. The specific fluorescence signals were obtained sequentially using the following settings: mKate2, excitation wavelength (ex): 589 nm, emission wavelength (em): 620-700 nm; mOrange2, ex: 546 nm, em: 450-590 nm; phiYFP, ex: 524 nm, em: 530-550 nm; eGFP, ex: 488 nm, em: 490-520 nm. Images were obtained using 10x/20x (0.8 NA) objectives. The Z-stacks module was used to detect roots and calli samples along the Z-axis to obtain Maximum Intensity Projection (MIP).

The regenerated plants of *XVE-Brainbow* were initially observed using a M205 FCA stereo fluorescence microscope equipped with a GFP filter (ex: 470/40 nm; em: 525/50 nm) for eGFP, a YFP filter (ex: 500/20 nm; em: 535/30 nm) for phiYFP, and a mCherry filter (ex: 560/40 nm; em: 630/75 nm). To further distinguish between mOrange2 and mKate2 fluorescence the plants were further confirmed using confocal microscopy after plants were treated with ClearSee reagent to become transparent. The ClearSee assays were performed as previously described.[89] In brief, leaves were immersed in fixation buffer (4% w/v paraformaldehyde in 1 x PBS) under vacuum for 1 h. The fixed leaves were then washed twice for 1 min in 1 x PBS and subsequently cleared in ClearSee solution for 4∼7 days at room temperature.

To achieve continuous observation of cell divisions during the regeneration process, the labeled root explants were placed on a 3.5 cm Petri dish containing CIM. The fluorescence signals from the same explant were observed continuously every day for 5 days of culture on CIM using the Zeiss LSM900 with Airyscan2 confocal (PI, ex:590 nm, em: 576-700 nm; eGFP, ex: 488 nm, em: 490-570 nm). Subsequently, the callus was transferred to SIM and its development was observed every two days using the M205 FCA stereo fluorescence microscope (eGFP, ex: 470/40 nm; em: 525/50 nm).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was conducted for all experiments and quantification data, with details included in the respective figure legends. The Student's t test ($P < 0.05$) was used for quantitative real-time PCR (mean ± SEM). One-way ANOVA, followed by Duncan's multiple range tests ($P < 0.05$), was used to analyze shoot regeneration rates (mean ± SD). Statistical analysis was performed using Origin Pro 2020b and SPSS 27.0.

For DNA barcode profiling, following high-throughput sequencing on the Hi-TOM platform, the raw data was converted to FASTA format, and the sequencing barcodes from each read were extracted to match the paired-end reads. After obtaining the reverse complement, the reads sequenced from two ends were separately subjected to multiple sequence alignment using muscle (v3.8)[78] and then merged to obtain the complete DNA barcode information. The reads that contain adapter sequences from both ends were considered to have a large fragment deletion between the first and fourth GCN4 sites, and were not merged. All reads were trimmed using seqtk (v1.3-r106)[79] software to remove adaptors, primers, and other bases that are unrelated to the mutation site. The reads were aggregated and filtered using a 1% threshold to remove the low-quality reads caused by sequencing errors. For evaluating the diversity of DNA barcode in a single callus (Figures 2E–2G), the sequences were filtered at 0.1%. The indels detected within the target site are considered as evidence of mutagenesis.

The copy numbers of the *Cas9-1-GCN4* and *Brainbow* transgenes were determined using whole genome re-sequencing (WGRS). The total DNA was prepared using the CTAB method and then sequenced using the HiSeq-PE150 platform. The clean reads were mapped to the *Arabidopsis* reference genome using BWA v0.7.17-r1188.[80] The clipped alignment was used to identify the insertion site and copy number of the transgenes.

All image analyses were performed using the ZEN Blue picture processing system and ImageJ[77] software.