

RESEARCH ARTICLE

Predictable and stable epimutations induced during clonal plant propagation with embryonic transcription factor

Anjar Tri Wibowo^{1,2,3}, Javier Antunez-Sanchez¹, Alexander Dawson¹, Jonathan Price¹, Cathal Meehan¹, Travis Wrightsman², Maximilian Collenberg², Ilya Bezrukov², Claude Becker^{2,4}, Moussa Benhamed⁵, Detlef Weigel^{2*}, Jose Gutierrez-Marcos^{1*}

1 School of Life Science, University of Warwick, Coventry, United Kingdom, **2** Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany, **3** Department of Biology, Faculty of Science and Technology, Airlangga University, Surabaya City, East Java, Indonesia, **4** Ludwig-Maximilians-University of Munich, Faculty of Biology, Biocenter, Martinsried, Germany, **5** Université Paris-Saclay, Centre National de la Recherche Scientifique, Institut National De La Recherche Agronomique, University of Évry, Institute of Plant Sciences Paris-Saclay (IPS2), Orsay, France

☞ These authors contributed equally to this work.

* detlef.weigel@tuebingen.mpg.de (DW); j.f.gutierrez-marcos@warwick.ac.uk (JGM)



OPEN ACCESS

Citation: Wibowo AT, Antunez-Sanchez J, Dawson A, Price J, Meehan C, Wrightsman T, et al. (2022) Predictable and stable epimutations induced during clonal plant propagation with embryonic transcription factor. *PLoS Genet* 18(11): e1010479. <https://doi.org/10.1371/journal.pgen.1010479>

Editor: Nathan M. Springer, University of Minnesota, UNITED STATES

Received: April 27, 2022

Accepted: October 15, 2022

Published: November 16, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1010479>

Copyright: © 2022 Wibowo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequence data (BS-seq, RNA-seq and DNA-seq) and *Ws-2* genome assembly have been deposited at the European

Abstract

Clonal propagation is frequently used in commercial plant breeding and biotechnology programs because it minimizes genetic variation, yet it is not uncommon to observe clonal plants with stable phenotypic changes, a phenomenon known as somaclonal variation. Several studies have linked epigenetic modifications induced during regeneration with this newly acquired phenotypic variation. However, the factors that determine the extent of somaclonal variation and the molecular changes underpinning this process remain poorly understood. To address this gap in our knowledge, we compared clonally propagated *Arabidopsis thaliana* plants derived from somatic embryogenesis using two different embryonic transcription factors- *RWP-RK DOMAIN-CONTAINING 4 (RKD4)* or *LEAFY COTYLEDON2 (LEC2)* and from two epigenetically distinct founder tissues. We found that both the epigenetic (genetic) status of the explant and the regeneration protocol employed play critical roles in shaping the molecular and phenotypic landscape of clonal plants. Phenotypic variation in regenerated plants can be largely explained by the inheritance of tissue-specific DNA methylation imprints, which are associated with specific transcriptional and metabolic changes in sexual progeny of clonal plants. For instance, regenerants were particularly affected by the inheritance of root-specific epigenetic imprints, which were associated with an increased accumulation of salicylic acid in leaves and accelerated plant senescence. Collectively, our data reveal specific pathways underpinning the phenotypic and molecular variation that arise and accumulate in clonal plant populations.

Nucleotide Archive (ENA) under BioProject PRJNA819891

Funding: This work was supported by ERA-CAPS AUREATE (DFG) and the Max Planck Society to (DW), an European Research Council (ERC) Marie Skłodowska-Curie Fellowship (751204-H2020-MSCA-IF-2016) and Airlangga University Hibah Riset Mandat Grant (387/UN3.14/PT/2020) to (ATW), and BBSRC grants (BB/L003023/1, BB/N005279/1, BB/N00194X/1 and BB/P02601X/1) to (JGM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Clonal propagation is commonly used in plants to propagate selected genotypes and to aid genetic/genomic manipulation. Although cloning minimizes genetic variation, clonal plants commonly display phenotypic changes, a phenomenon known as somaclonal variation, that are at times stable after multiple cycles of sexual reproduction. The newly acquired phenotypic variation exhibited by clonal plants has been linked to non-genetic modifications induced during the regeneration process, though the precise nature of these modifications and the factors implicated remain poorly understood. By generating clonal plants through two different methods and using two distinct tissues—roots and leaves—, we show that the phenotypic variation found in clonal plants is linked to heritable changes in DNA methylation marks already present in the founder tissues. Further, these novel phenotypes are associated with specific patterns of gene expression and metabolic changes. Our findings unveil specific pathways underpinning new phenotypic and molecular variation that arises and accumulates in clonal plant populations.

Introduction

Many multicellular organisms can reproduce both sexually and asexually, with clonal reproduction being especially common in plants, where it is prevalent in ferns, mosses and many angiosperms [1, 2]. Most plant species can switch to asexual reproductive programs in response to environmental stimuli by forming specialized structures such as rhizomes, stolons and bulbils [2]. Such flexibility in reproductive strategies is thought to be possible due to the high degree of plasticity and totipotency of plant cells. This property has been traditionally exploited for the clonal propagation and genomic manipulation of many economically important species [3]. Clonal propagation can be achieved in different ways, from simple methods such as grafting and cuttings to more advanced methods such as tissue culture-induced regeneration using phytohormones or embryonic/meristematic transcription factors. In *Arabidopsis thaliana* (hereafter, *Arabidopsis*), somatic embryogenesis (SE) and plant regeneration can be initiated through the ectopic overexpression of embryonic transcription factors such as *AGAMOUS-LIKE15 (AGL15)* [4], *BABY BOOM (BBM)* [5], *LEAFY COTYLEDON2 (LEC2)* [6], *WUSCHEL (WUS)* [7–9] and *RWP-RK DOMAIN-CONTAINING 4 (RKD4)* [10].

A key characteristic of asexual propagation is that the individuals produced are genetically identical to their parents, except for a few spontaneous *de novo* mutations [1]. However, clonal plants often display heritable phenotypic variation [11, 12], known as somaclonal variation, which is linked to genetic [13, 14] and epigenetic changes [12, 15, 16]. Moreover, this phenotypic diversity could contribute to the adaptation to recurrent environmental challenges. In a previous study we found that *Arabidopsis* plants regenerated through the ectopic activation of an embryonic transcription factor, resulted in clonal plants that inherited molecular signatures, DNA methylation and gene expression profiles, characteristic of the tissue of origin [12]. Whether this molecular and phenotypic variation is primarily caused by the direct regeneration of clonal individuals from different above and belowground organs, or driven by the organ-specific activity of the embryonic transcription factor employed, has been unclear. To address this gap in our knowledge, we generated clonal progeny from *Arabidopsis* roots and leaves through the ectopic activation of two different embryonic transcription factors. We found that the founder organ employed for cloning significantly contributes to clonal progeny, with root regenerants inheriting tissue-specific DNA methylation imprints and transcriptional states, independent of the regeneration method used. Results from our study reveals avenues

for the manipulation of phenotypic traits through cloning for the adoption in crop improvement programs.

Results

Phenotypic variation present in clonal plants is underpinned by tissue of origin

In a previous study, we identified heritable molecular and phenotypic changes in *Arabidopsis* Col-0 plants propagated through the induction of the embryonic transcription factor *RKD4* [12]. To better understand how this somaclonal variation is created, in particular the importance of the genetic background and/or the transcription factor employed, we generated clonal progeny from *Arabidopsis* Ws-2 leaf and root tissues using transgenic lines that allow the controlled expression of either *RKD4* or *LEC2*, two transcription factor gene necessary for the initiation of different stages of embryo development [6, 10]. To determine the stability of the phenotypes arising during cloning, we propagated seeds from each regenerated individual (G0) by selfing over three consecutive generations (G1-G3) under similar growth conditions, thus limiting possible effects caused by tissue culture. When we grew G2 and G3 clonal progeny, we noticed that the morphology of plants regenerated from roots using *LEC2* ectopic expression (RO-*LEC2*) was noticeably different to control plants, while leaf regenerants (LO-*LEC2*) were indistinguishable from the control plants (Fig 1A). Notably, most RO-*LEC2* lines (6/10) had reduced biomass, smaller leaf area and senesced early (Fig 1C and 1D), with different temperature growth conditions having no discernible effects on these phenotypes (S1 Fig). These phenotypes were only found in RO-*LEC2* lines and plants regenerated using *RKD4*, either from root or leaf, did not produce any observable phenotypes. Collectively, our data suggest that the tissue of origin employed for regeneration and the regeneration protocol could contribute to phenotypic diversity in the clonal plants and their progenies, thus pointing to the long-term inheritance of somatic epigenetic states in the sexual progeny of clonal plants.

Tissue of origin affects the activity of defense-related genes

To determine what molecular changes might underpin the observed morphological variation, we performed whole-genome transcriptome analysis in leaves and roots from randomly selected, independent G3 lines. Using stringent thresholds [false discovery rate (FDR) <0.05; absolute log-two fold change >1.5], we obtained lists of Differentially Expressed Genes (DEGs) by comparing leaves and roots of regenerants and control plants (S1 and S2 Tables). There were notable differences between the regenerants induced by the two embryonic transcription factors; leaves of *RKD4*-induced regenerants were the most affected independent of the tissue of origin (401 and 253 DEGs in leaves for LO and RO respectively, compared to 147 and 48 in roots). Notably, we found more DEGs for LO than for RO plants (Figs 2A and S2), which resembled the effects found in Col-0 plants cloned using *RKD4* [12]. *LEC2* regenerants had the highest numbers of DEGs in the opposite tissue from which the plant was regenerated, i.e., in roots of LO-*LEC2* plants (282 DEGs versus 75 in leaves) and in leaves of RO-*LEC2* plants (595 DEGs versus 61 in roots) (Figs 2A and S3). Since gene expression was strongly dysregulated in the leaves of both RO-*RKD4* and RO-*LEC2* regenerants, we looked for an overlap between the DEGs. We found more shared dysregulated genes than expected by chance (p -value = 2.1×10^{-17}), suggesting common effects independent of the regeneration strategy.

To obtain further insight into the possible function of these DEGs, we performed Gene Ontology (GO) analyses, finding that DEGs in leaves from both RO-*LEC2* and RO-*RKD4*

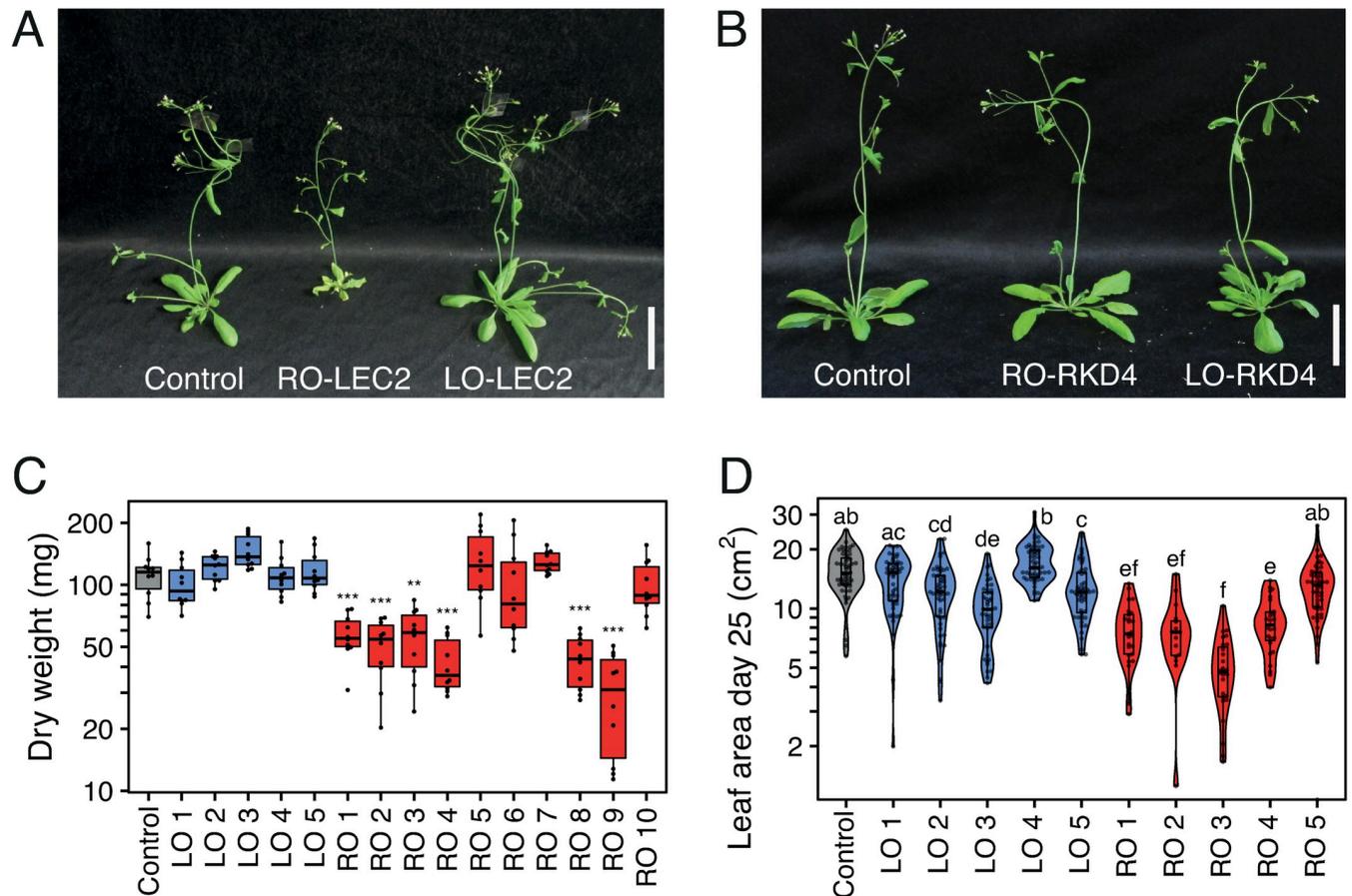


Fig 1. Growth phenotypes of Arabidopsis plants cloned from roots (RO) and leaves (LO) through LEC2/RKD4-induced regeneration and propagated by self-fertilization over three generations. (A) Representative images of 28-day old Arabidopsis *Ws-2* plants. Control (*Ws-2/indLEC2*), and clonal RO-LEC2 and LO-LEC2 G3 progeny. Scale bars, 5 cm. (B) Representative images of 28-day old Arabidopsis *Ws-2* plants. Control (*Ws-2/indRKD4*), and clonal RO-RKD4 and LO-RKD4 G3 progeny. Scale bars, 5 cm. (C) Boxplot showing variation in plant biomass for control (*indLEC2/Ws-2*) and independent G3 lines. Asterisks represent significant differences, determined by Student's *t*-test, ** $p < 0.01$, *** $p < 0.001$, sample size $n = 10$. (D) Violin plot and boxplot of leaf area for control (*Ws-2/indLEC2*) and independent G3 lines. Letters represent groups of statistically significantly different samples ($p < 0.05$), determined by Tukey's test, sample size $n = 20$.

<https://doi.org/10.1371/journal.pgen.1010479.g001>

lines were enriched for stress and defense responses (FDR < 0.05, Fig 2C). In contrast, DEGs in roots from either LO-LEC2 or LO-RKD4 lines were primarily associated with metabolic processes (Fig 2C). Network analysis revealed that genes related to pathogen responses and senescence were upregulated in RO plants, independent of the embryonic transcription factor used for clonal propagation (Fig 2D and 2E). As several of the genes upregulated in RO-LEC2 leaves were part of the salicylic acid (SA) biosynthesis pathway (Fig 2F), we measured the concentration of total SA in G2 progenies of LEC2 regenerated plants. Reduced biomass and early senescence in RO-LEC2 lines were strongly associated with high levels of SA, indicating a causal relationship between transcriptional changes, differences in production of a defense hormone known to negatively affect growth and to induce early senescence [17], and morphological changes in the regenerants (Fig 2G). The senescence phenotypes observed are likely caused by the disturbance of free SA and conjugated SA. Collectively, our analyses revealed that the clonal propagation of plants using embryonic transcription factors results in the transmission of organ-specific transcriptional states.

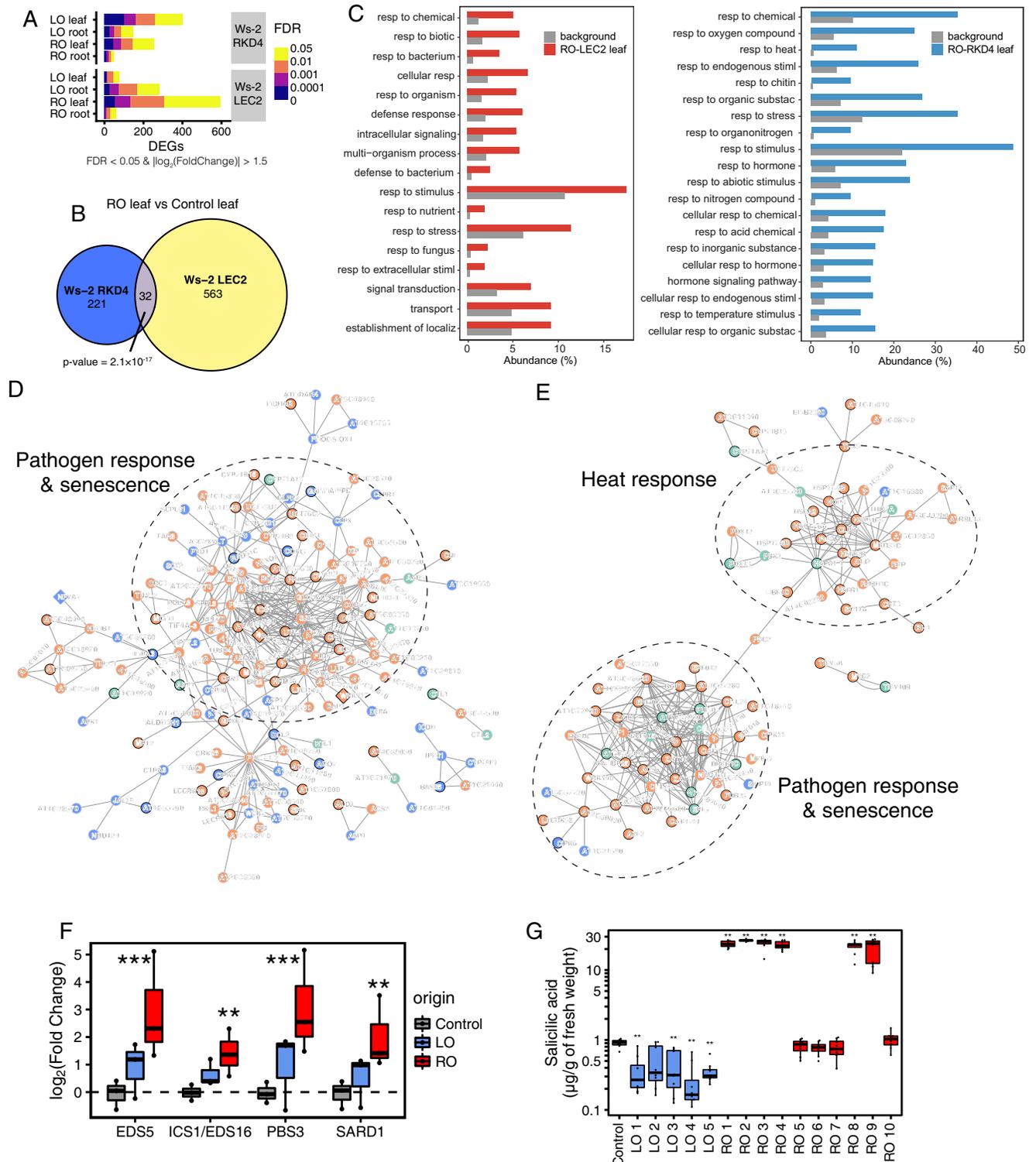


Fig 2. Differential expression of pathogen response and senescence genes in Arabidopsis Ws-2 plants regenerated from roots or leaves using RKD4 or LEC2. (A) Number of DEGs for different comparisons split by FDR levels. Data from three randomly selected regenerants for each tissue. (B) Euler diagram of common DEGs found in RO plants regenerated using two different transcription factors. p-value indicates results from Fisher’s exact test. (C) Gene Ontology analysis showing significant enrichment for stress-related genes among DEGs in RO-LEC2 (left) and RO-RKD4 (right) leaves. Categories are sorted by decreasing FDR. (D) Interaction network of DEGs found in leaves of RO-LEC2 plants. Orange, upregulated genes; blue, downregulated genes; green, non-

DEGs with significant interactions with the network. Diamond shaped nodes, transcription factors. Main GO terms associated with clusters are highlighted. (E) Interaction network of DEGs found in leaves of RO-RKD4 plants. Orange, upregulated genes; blue, downregulated genes; green, non-DEGs with significant interactions with the network. Diamond shaped nodes, transcription factors. Main GO terms associated with clusters are highlighted. (F) Expression levels from RNA-seq of genes associated with SA biosynthesis in LO-LEC2 and RO-LEC2 plants compared to control (Ws-2/indLEC2) parent. Asterisks represent significant differences, determined by Wald test and adjusted with the Bonferroni–Holm method: ** $p < 0.01$, *** $p < 0.001$, (G) SA concentrations in leaves of LEC2 regenerants at the G3 generation. Asterisks represent significant differences, determined by Student's *t*-test, ** $p < 0.01$, sample size of $n = 10$. Control = Ws-2/indLEC2 parent.

<https://doi.org/10.1371/journal.pgen.1010479.g002>

Heritable DNA methylation imprints in LEC2-induced clonal progeny

We postulated that the phenotypic and transcriptional changes found in plants regenerated with embryonic transcription factors may be caused by the inheritance of heritable tissue-specific epigenetic imprints. To test this hypothesis, we studied the methylome of LEC2 regenerants that had been propagated for two successive generations (G2 and G3). To map methylome data we generated a *de novo* assembly of the Ws-2 genome (see [methods](#)). In an all-against-all comparison of RO-LEC2 and LO-LEC2 plants, we found 226, 120 and 386 Differentially Methylation Regions (DMRs) in CG, CHG and CHH contexts, respectively ([S3 Table](#)). Principal component analysis (PCA) of the 226 CG-DMRs showed that the main source of variance in our dataset was the tissue analyzed (69% of variance), with methylome profile of leaves of RO-LEC2 plants being slightly skewed towards root samples ([Fig 3A](#)). Another 6% of variance can be explained by the tissue of origin, separating RO-LEC2 and LO-LEC2 samples from each other and from the controls, independent of the tissue sampled. Some of the remaining variance (3%) comes from the regeneration process independent of tissue sampled or founder tissue used for cloning ([Fig 3A](#)). These data suggest that even though the methylome of regenerated plants mainly corresponds to the characteristic methylome of the tissue sampled, a number of epigenetic imprints are maintained from the tissue employed for cloning, even after plants had undergone three cycles of sexual reproduction. Our data also show that regeneration by itself contributes to discrete methylation imprints that are stably inherited in clonal plants ([S4 Fig](#)). Similar, albeit weaker patterns were observed for non-CG DMRs ([S4 Fig](#)). To define the potential impact of these epigenetic changes, we mapped DMRs to genome features, which revealed that regions flanking protein coding genes were particularly likely to be associated with methylation changes ([S5 Fig](#)).

When only considering DMRs distinguishing leaves and roots of non-regenerated plants, RO-LEC2 leaves, although distinct from other samples, resembled the methylation of root samples ([Fig 3B](#)). Most DMRs in leaves of RO-LEC2 regenerants were primarily hypomethylated compared to leaves of control plants, resembling the methylation profile of these regions in control roots, and they were stably inherited ([Figs 3C and S4](#)). Similar to roots of control plants, these DMRs were also hypomethylated in roots of LO-LEC2 regenerants, thus pointing to active DNA demethylation in *Arabidopsis* roots. Collectively, our analyses demonstrate that induction of somatic embryogenesis and subsequent regeneration results in tissue-specific epigenetic imprints that are stable during sexual reproduction.

Precocious leaf senescence in root-regenerants caused by epigenetic deregulation of the salicylic acid pathway

To determine the molecular basis for the early senescence phenotype observed in RO regenerants, we performed reciprocal crosses between Ws-2 and RO-LEC2 plants. Five reciprocal F_1 hybrids were self-pollinated to generate F_2 seed progeny ([Fig 4A](#)). In the F_2 generation, some plants were early senescing, characteristic of RO-LEC2 plants, while others were not. The fraction of early-senescing plants was around one quarter, with 16% (13 in 80) early-senescing F_2

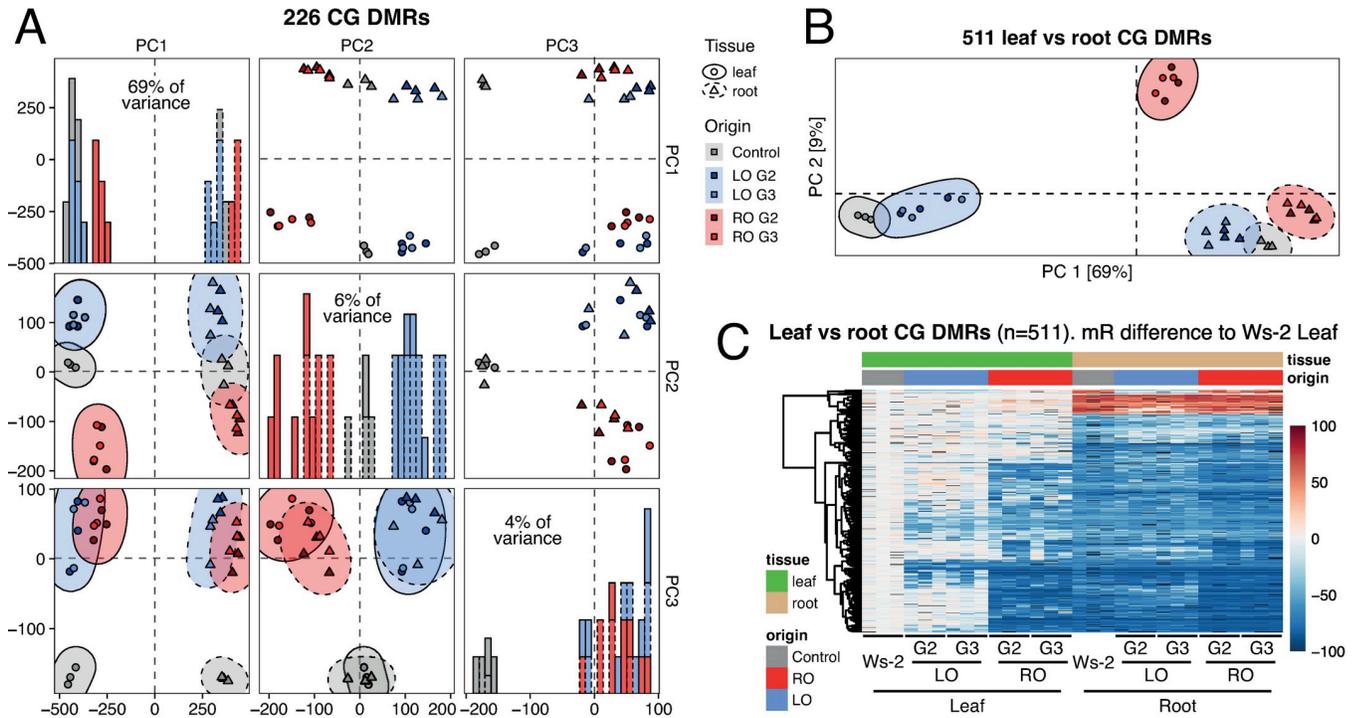


Fig 3. DNA methylomes of LEC2 regenerants. (A) PCAs of CG DNA methylation levels at 226 CG-DMRs identified in all-against-all comparison and histograms of the distribution of samples across each principal component. Probability ellipses (95% distribution) for each group are shown. Gray, control; Blue, LO-LEC2; Red, RO-LEC2. Data from three randomly selected regenerants for each tissue. (B) PCA of CG methylation levels at 511 CG-DMRs identified in the comparison of control leaves and roots. Probability ellipses (95% distribution) for each group are shown. Gray, control; Blue, LO-LEC2; Red, RO-LEC2. Data from three randomly selected regenerants for each tissue. (C) Heatmap of CG methylation levels in leaves and roots of control plants and LEC2 regenerants. Each column corresponds to an independently regenerated line. Ws-2 = Ws-2/indLEC2. Data from three randomly selected regenerants for each tissue.

<https://doi.org/10.1371/journal.pgen.1010479.g003>

progeny derived from Ws-2/indLEC2 x RO-LEC2 hybrids, and 23% (28 in 120) early-senescent F₂ progeny derived from RO-LEC2 x Ws-2/indLEC2 hybrids ($p > 0.05$, Fisher's Exact test). These data suggest that early senescence in RO-LEC2 regenerants may be under the control of a single recessively acting locus. To determine whether this was due to a genetic mutation, we whole-genome shotgun sequenced 10 phenotypically normal and 10 early-senescent individuals from five different segregating F₂ populations (Fig 4A). However, we could not identify genetic variants linked to this phenotype (S6 Fig), suggesting that the phenotypes were caused by an epigenetic change.

We investigated the mode of inheritance of RO imprints by generating DNA methylation profiles of leaves from F₂ progeny derived from Ws-2/indLEC2 x RO-LEC2 hybrids. If differences in DNA methylation were responsible for the growth phenotypes characteristic of RO plants, we reasoned the methylation imprints must be already present in the parents. We therefore focused on DMRs between leaves of Ws-2/indLEC2 and RO-LEC2 plants. Most regions that are hypomethylated in RO-LEC2 regenerants were stably inherited both mitotically and meiotically (Fig 4B). On the other hand, most hypermethylated DMRs had disappeared after two generations of paternal transmission, although they were faithfully inherited maternally (Fig 4B). This was observed for both CG- and CHG-hypermethylated DMRs, while CHH-hypermethylation was reset both paternally and maternally (S7 Fig). These data suggest that the early senescence phenotype observed in RO-LEC2 plants may be caused by an epimutation associated with a reduction in DNA methylation, as plants with early senescence

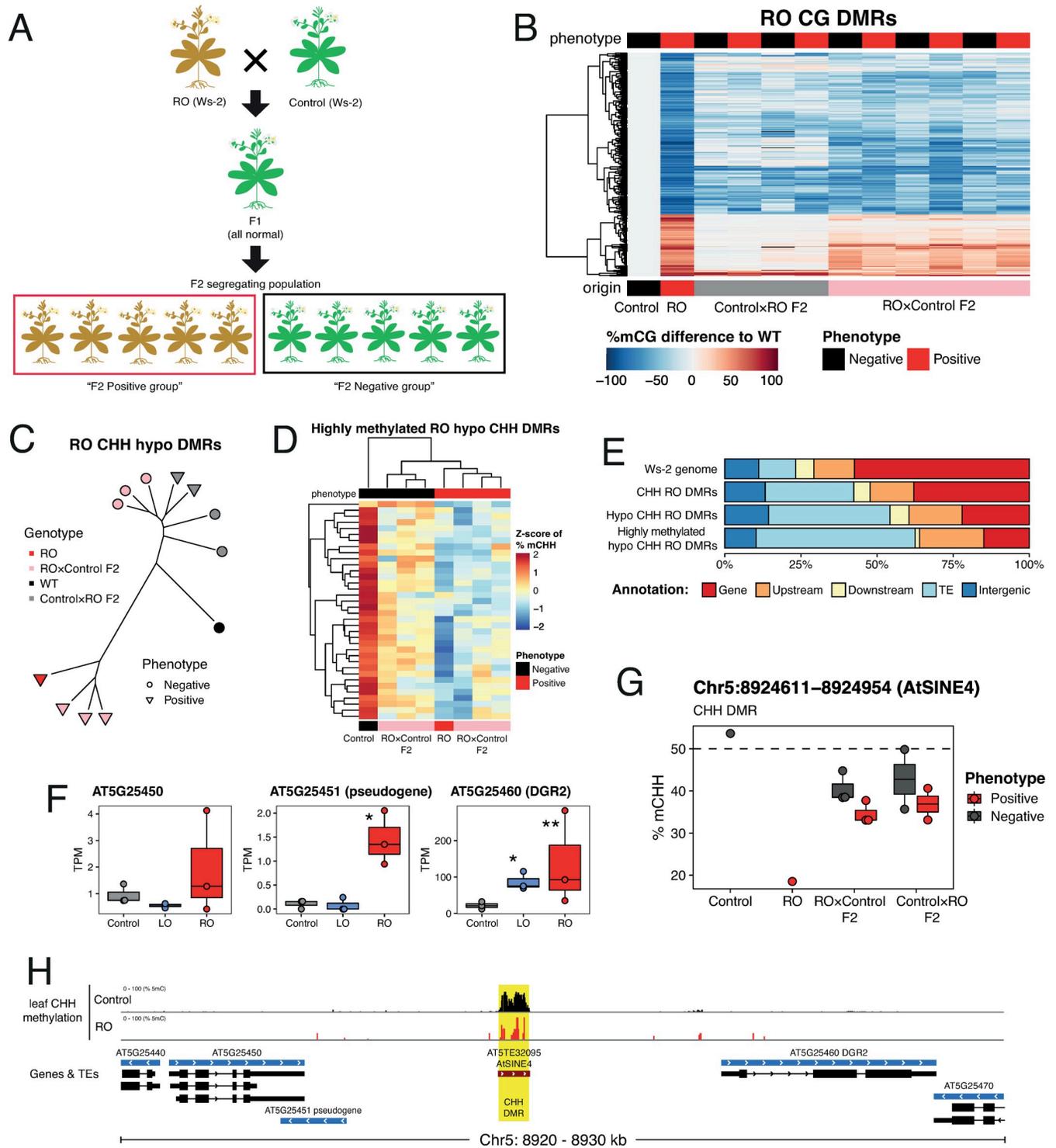


Fig 4. Leaf senescence and TE hypomethylation in RO-LEC2 root regenerants. (A) Experimental design to identify epi- and DNA mutations linked to RO-LEC2 phenotypes. Green indicates phenotypically normal plants (n = 10) and brown early senescing ones (n = 10). Analysis was carried out using five independent F₂ progeny. (B) Heatmap showing gains and losses of CG methylation for RO-LEC2-leaf DMRs in F₂ progeny from reciprocal backcrosses between control (Ws-2/indLEC2) and RO-LEC2 plants. Each column represents an independent F₂ progeny individual. (C) Dendrogram of F₂ progeny and parents at parental CHH-hypo-DMRs. (D) Heatmap of Z-scores in CHH methylation for highly methylated (Ws-2/indLEC2 %mC > 50) parental CHH-DMRs, showing clustering of parents and RO x Ws-2/indLEC2 F₂ progeny. (E) Annotation of several subsets of parental CHH DMRs to different genetic features compared to the distribution of features in the whole genome. Upstream and downstream are defined as 2 kb away from a gene and Intergenic as all sequences

more than 2 kb away from any gene. The priority of annotations is Gene, TE, Upstream, Downstream and Intergenic in that order. (F) Expression levels from RNA-seq for genes proximal to Chr5:892411–894943 (AtSINE4) DMR in leaves of control (Ws-2/indLEC2) and regenerants. Asterisks represent significant differences, determined by Wald test and adjusted with the Bonferroni–Holm method: * $p < 0.05$, ** $p < 0.01$ (G) CHH methylation levels for Chr5:892411–894943 (AtSINE4) DMR in Ws-2/indLEC2, RO-LEC2, and F₂ reciprocal crosses, showing association between demethylation and morphological phenotype. (H) Genomic view around Chr5:892411–894943 (AtSINE4) DMR. Yellow box represents RO-LEC2 CHH-DMR.

<https://doi.org/10.1371/journal.pgen.1010479.g004>

appeared in both directions of the cross. After clustering samples based on the levels of methylation for DMRs, we found that CHH-hypomethylated DMRs as a class predicted phenotype, (Fig 4C), and this was clearest for RO-LEC2 x Ws-2/indLEC2 progeny. Thirty-six DMRs co-segregated with the growth phenotype. Normal plants were highly methylated, as in the Ws-2/indLEC2 control, and early senescing plants had low methylation at these 36 DMRs (Fig 4D). These phenotype-associated DMRs were strongly enriched in transposable elements (TEs) (Fig 4E). Collectively, these data suggest that demethylation of TEs is associated with early senescence of RO-LEC2 plants.

To understand how losses of DNA methylation at specific loci could cause the phenotype, we searched for deregulated genes in the proximity of the 36 DMRs. We found two genomic regions where methylation status co-segregated with the phenotype for both directions of the crosses (Fig 4G). One region included three genes in its proximity that are upregulated in leaves of RO-LEC2 regenerants (Fig 4F). The DMR is located in a TE of the AtSINE4 family, between two of the genes found to be misregulated in RO-LEC2 plants when compared to Ws-2/indLEC2 control (Fig 4H). We investigated the chromatin conformation of these regions in leaves and roots of wild-type plants, and found notable differences, thus suggesting it may act as a long-distance regulatory element for flanking genes (S8 Fig). Another DMR associated with early senescence in RO-LEC2 plants was located in an AtMUNX1 TE upstream of the *PRE1/BNQ1* locus (S9 Fig), in a region flanking a CarG box known to be a target of MADS box containing transcription factors [18] and negatively regulated by the transcription factor pair APETALA3 (AP3) and PISTILLATA (PI) [19].

Discussion

A common concern surrounding clonal propagation in plants by tissue culture is the appearance of new undesirable phenotypes. This outcome may be in part attributed to the type of regeneration employed ie. organogenesis or somatic embryogenesis. For instance, using phytohormones is known to cause stochastic changes in DNA methylation, primarily hypomethylation, that are associated with changes in gene expression [11, 13, 16, 20, 21]. On the other hand, regeneration via somatic embryogenesis using embryonic transcription factors results in nonrandom changes in DNA methylation, which are heritable and associated with transcriptional and phenotypic variation [12]. To define the factor(s) that contribute to the molecular and phenotypic diversity found in plants propagated through somatic embryogenesis, we conducted a detailed analysis of regenerants created either by the ectopic expression of LEC2 or RKD4, two transcription factors that are part of distinct embryonic transcriptional networks [10, 20, 22]. Through this approach, we were able to identify heritable tissue-specific DNA methylation imprints in these regenerants that were associated with specific patterns of gene expression, leading to significant phenotypic variation. Notably, the molecular and phenotypic variation observed in clonal plants was independent of the transcription factor used for regeneration and instead strongly linked to the founder tissue employed.

The molecular mechanism(s) underpinning the inheritance of tissue-specific DNA methylation imprints that accumulate in these clonal plants are yet unknown, but they may be linked to the capacity of these two transcription factors to reprogram somatic cells such that they

adopt an embryonic program without altering their epigenomic configuration. One explanation for this phenomenon could be that tissue-specific imprints arising during cloning are developmentally programmed, since leaves and roots have distinct epigenetic characteristics [23–27]. The tissue-specific imprints found in clonal plants may be linked to the distinct chromatin architecture found in root and shoot cells, which is largely influenced by chromatin loops associated with repressive histone marks [28]. Chromatin interactions likely contribute to the developmental differentiation of specific cell types in plants, as is the case in animals [29, 30], and it would be interesting to test if these interactions are maintained or reset in clonal plants. Nevertheless, our data show that some epigenetic imprints can be fixed independently of the pathway recruited to achieve somatic embryogenesis and regeneration. This suggests that genetic and epigenetic programs can be uncoupled in clonal plants.

Our data further show that irrespective of the transcription factor employed, plants regenerated from roots senesced early—a phenotype that was inherited by the progeny either after selfing or outcrossing. The observed early leaf senescence strongly correlated with DNA hypomethylation of specific TEs flanking genes associated with the SA pathway. Our data suggest that epimutations at TEs could accumulate in populations of species that reproduce asexually, and that these epimutations could confer an adaptive advantage under fluctuating environments. Several studies have shown that the long-distance regulatory action of TEs in Arabidopsis is associated with changes in DNA methylation [31]. Moreover, transposon mobilization and variation in TE methylation have been linked to environmental adaptation in natural plant populations [32] and that this epigenetic variation often occurs near genes associated with environmental responses [33, 34]. As such, the methylation of transposons has been selected to provide new regulatory functions in most eukaryotic genomes [35].

The SA pathway may be an important target for the accumulation of epimutations in plants, as it is critical for defense [36], and is strongly correlated with biomass, a characteristic that has traditionally been selected by breeders to increase yield in various crops, including maize [37], rice [38], and wheat [39]. Notably, mutants that cause a global change in chromatin organization, such as *CROWDED NUCLEI* (*CRWN*), display constitutive activity of the salicylic acid-dependent pathway of immunity, with *crwn* plants displaying dwarfism and premature senescence [40, 41].

Epigenetic imprints generated during cloning can be inherited over multiple cycles of sexual reproduction, suggesting that the epigenetic machinery in plants is unable to repair these epimutations [11, 12, 16]. The stable accumulation of epimutations has also been documented in crosses between wild-type plants and mutants defective in DNA methylation [42–46] and histone demethylation [47]. Notably, while the tissue-specific DNA hypomethylation imprints of clonal plants are transmitted through both the maternal and paternal germline to the offspring. However, DNA hypermethylation imprints are only transmitted maternally. The molecular mechanisms implicated in the resetting of these DNA hypermethylated imprints are not yet known but may be linked to the active epigenomic reprogramming that takes place during male gametogenesis [48, 49] and/or embryo development [50]. Nevertheless, our study reveals that the parental transmission of the epigenetic imprints generated during clonal propagation could be used as a means of controlling their maintenance or erasure in plant breeding.

In summary, our study has revealed that clonal propagation via somatic embryogenesis using embryonic transcription factors enables the generation of specific epigenomic changes leading to regenerated plants with novel phenotypes. Moreover, because these epigenomic changes are stable albeit exhibit selection/bias during sexual reproduction, stable epimutations could arise and be maintained in populations of plants that alternate between asexual and sexual programs.

Methods

Plant growth and clonal propagation

For all experiments, *A. thaliana* Ws-2 plants were grown in long day conditions (16-h light/8-h dark, 22°C). To facilitate direct plant regeneration from different type of organs we used transgenic lines harboring an chemically-inducible transgenes to overexpress the GRANDE (GRD)/RWPRK motif-containing (RKD4) transcription factor [10] (indRKD4) or the B3 domain transcription factor LEAFY COTYLEDON2 (LEC2) [51] (indLEC2). Seeds from transgenic indRKD4 and indLEC2 lines were germinated on Murashige and Skoog (MS) medium containing 30 μM dexamethasone or 50 μM estradiol, respectively, and incubated for 14 days. Plants were transferred to MS media without dexamethasone or estradiol for 7 days to allow the formation of somatic embryos, which were then isolated and transferred to fresh MS plates using tungsten needles. After 2–3 weeks, the somatic embryos from leaves (LO) or roots (RO) were fully regenerated into new plants. We generated 10 independent regenerants from each organ (G0 generation), which were grown in soil to produce seeds. We grew 24 plants from each line (G1 generation) and selected 10 individuals at random for each regenerated line to produce seed. We continued the sexual propagation of each regenerant population for two additional generations (G2 and G3) following the same scheme.

Plasmid constructs and transgenic lines

To generate estradiol-inducible RKD4 lines (indRKD4) we cloned using Gateway recombination the coding region of GRD/RKD4 (AT5G53040) in pMDC7 [52] modified by the insertion after PmeI digestion of a pOLE1:OLE1-RFP fragment from pFAST-R [53]. The estradiol-inducible LEC2 lines (indLEC2) were generated by amplifying the LEC2 coding sequence by PCR and cloning in pBI-ΔGR vector. The LEC2-GR fragment was then cloned in pCGN18 [54] to create p35S::LEC2-GR. Plants were transformed by floral dipping, transgenic plants were selected in Kanamycin (indLEC2) or RFP seed selection (indRKD4), and propagated for three generations to isolate homozygous lines.

Phenotypic analysis

For biomass analysis, plants were grown under long days for 4 weeks (16-h light/8-h dark, 22°C). Dry weight was recorded using aerial parts excised and dried at 80°C for 48 hours. For image-based phenotyping, images were acquired once per day in top view using two cameras per tray. The cameras were equipped with OmniVision OV5647 sensors with a resolution of 5 megapixels. Each camera was controlled by a Raspberry Pi computer (Revision 1.2, Raspberry Pi Foundation, UK). Stitched whole-tray images were separated into individual pots using a predefined mask for each position. Segmentation of plant tissue against background was performed by removing the background pixels and applying GrabCut-based automatic post processing. The leaf area of each plant was calculated based on the segmented plant images.

SA measurement

The total levels of SA and SA glucoside (SAG) were measured from each sample using the *Acinetobacter* sp. ADPWH-*lux* biosensor [55]. In brief, 175 mg of leaves material were harvested from four-week-old plants and homogenized in 250 μl 0.1 M sodium acetate buffer pH 5.5. Samples were then centrifuged for 15 min at 16,000 g. For total SA measurement, 200 μl of supernatant was transferred to a new tube and incubated with 20 μl of 0.5 U/μl β-glucosidase (Sigma-Aldrich) for 90 min at 37°C. Following β-glucosidase treatment, 60 μl of LB, 30 μl of plant extract, and 50 μl of *Acinetobacter* sp. ADPWH-*lux* (OD = 0.4) were incubated together

at 37°C for 150 min. Luminescence was read with a Tecan luminometer. SA levels were calculated based on the SA standard curve, which was constructed using known amounts of SA stock (from 0 to 2500 ng/μl) diluted in *NahG* plant extract [56].

Genome assembly and annotation

Plants were grown in long day conditions (16 h of light) at 23°C and a relative humidity of 65%. Philips GreenPower TLED modules (Philips Lighting GmbH, Hamburg, Germany) with 110–140 μmol m⁻² s⁻¹ light were used. Plants were grown for 3 weeks. In order to reduce the accumulation of starch all plants were placed in darkness for 30 h before harvesting young leaves. Fresh plant material was immediately flash-frozen after harvesting. HMW DNA extraction was performed following a custom protocol. Approximately thirty grams of leaf tissue were ground in liquid nitrogen and subsequently transferred into a 1 L flask containing 500 mL of ice-cold nuclei isolation buffer. The NIB was composed of 10 mM Tris (pH 8), 0.1 M KCl, 10 mM EDTA (pH 8), 0.5 M sucrose, 4 mM spermidine, and 1 mM spermine-4HCl. Flasks with ground plant material and NIB were kept on ice for fifteen minutes with occasional gentle swirling. The homogenate was filtered through four layers of miracloth (Merck, Germany). Subsequently, NIB solution with 20% Triton X-100 was added. The solution was kept on ice for fifteen minutes with occasional gentle swirling. Afterwards, the solution was distributed into 50 mL Falcon tubes and centrifuged for 15 min at 3,250 rpm at four degrees using a tabletop centrifuge. Nuclei containing pellets were washed by adding a 40 mL NIB solution with 1% Triton X-100 prior to centrifugation for 15 min at 4°C and 3,250 rpm. Subsequently, the supernatant was discarded before resuspending the nuclei pellet in 30 mL prewarmed (37°C) G2 lysis buffer (Qiagen, Germany). Resuspended pellets were incubated with RNase A (50 μg/ml) for 30 minutes while gently inverting the tube every five minutes. Proteinase K was added (250 μg/ml) and samples were incubated overnight at 50°C. Subsequently, the lysate was centrifuged at 8,000 rpm at 4°C for 15 min. The supernatant was poured onto an equilibrated Genomic Tip 100 (Qiagen, Germany). Afterwards the Genomic Tip 100 was washed twice with 7.5 mL QC Buffer (Qiagen, Germany) per wash. Genomic DNA was eluted from the column using 5 mL of prewarmed (50°C) QF Buffer (Qiagen, Germany). DNA was precipitated by adding 0.7 volumes of isopropanol. HMW DNA was obtained by spooling DNA strings onto a glass hook. The spooled DNA pellet was washed with 80% EtOH prior to resuspending it in 300 μL EB. DNA concentration was measured using Qubit HS DNA kit.

For library construction, high molecular weight DNA was sheared with a Megaruptor (Diagenode) set to a 40 kb fragment size. The DNA library was prepared using the Accel-NGS XL Library Kit (Swift Biosciences, United States). The library was size-selected with a minimum fragment size of 20 kb using a BluePippin (Sage Science, Beverly, United States and a High-Pass v3 cassette. The library was sequenced on the Pacific Biosciences (Menlo Park, United States) Sequel platform (ICS v5.0.0) with the v5.0.1 chemistry, which produced 5.09 Gb of long read sequences with a subread N50 of 15,750 bp.

The genome was assembled into 120 contigs of total length of 123.5 Mb with an N50 of 6.57 Mb using a custom Snakemake pipeline, “auto-asm” (<https://github.com/weigelworld/auto-asm>). Briefly, long reads were assembled into contigs using canu [57] and scaffolded without polishing steps into chromosomes against TAIR10 using reveal [58]. Whole Genome Annotation was lifted over using the CAT [59] pipeline with TAIR10 as the reference genome and annotation.

Transcriptome analysis

For mRNA-seq analysis, total RNA (10 μg) for each sample was used to purify polyA+ mRNA that was used for synthesis and amplification of cDNA. The RNA-seq libraries were prepared

using TruSeq RNA Sample Preparation Kit from Illumina (San Diego, CA). RNA libraries were sequenced on an Illumina HiSeq2000 instrument (100 bp single-end). Read quality was assessed using FastQC and trimming of low-quality bases at the 3' end of reads and adapter removal was done using Trimmomatic. Reads were mapped to the TAIR10 reference genome using Tophat (parameters -i 20 -I 30,000) with on average 85.4% unique mappings (min = 82.2, max = 88.8). The read counts for these libraries are given in Dataset S1. We used the R package DESEQ2 (version 1.10.1) ⁸. RNA-seq libraries from samples Control-LEC2 root 2 and Control-RKD4 root 3 were discarded for further analysis due to bad quality and low replication fidelity. Genes were classified as significantly differentially expressed at either an $FDR < 0.01$ or < 0.05 and $|\log_2 \text{fold-change}| > 1.5$. We used AgriGO v2.0 [60]. The background used for identifying enrichment was the suggested background provided by AgriGo, which contains all GO annotated genes in Arabidopsis. Enrichment was performed using a hypergeometric test, p-values were corrected by multiple correction (Benjamini Hochberg FDR; $\alpha = 0.05$).

Identification of polymorphisms in RO-LEC2 lines

For library construction, high molecular weight DNA from individual plants was isolated using DNeasy Plant mini kit (Qiagen). The DNA library was prepared using a NxSeq Amp-FREE Low DNA Library kit (Lucigen) and sequenced on a Illumina HiSeq2000 with a 100bp paired-end mode. Sequencing reads were pre-processed using fastp with default settings and the overrepresentation analysis parameter to remove reads with low quality score, irregular GC content, short length, and sequencing adapters present [61]. Trimmed reads were then quality checked with outputs from fastp to ensure trimming had been successful. Trimmed reads were mapped to the TAIR10 reference genome using bowtie2 [62]. Variant calling was performed on aligned and sorted BAM files using samtools mpileup and piped to bcftools to produce VCF files [63]. VCF files were converted to SHOREmap format and analysed using SHOREmap backcross and then annotated to output tables and plots of backcross SNPs [64].

Bisulfite sequencing

Bisulfite libraries were prepared from 100 ng genomic DNA using the TruSeq Nano kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions, with the following modification. After adapter ligation and cleanup, samples were bisulfite-treated using the Epi-ect Plus kit (Qiagen, Hilden, Germany). Thermal cycle incubation was repeated once before clean-up. DNA fragments were amplified using the Kapa HiFi Hotstart Uracil+ PCR mix in 14 cycles of PCR (Kapa Biosystems, Wilmington, MA, USA). Bisulfite libraries were sequenced on an Illumina HiSeq3000 instrument (150 bp paired-end reads).

Computational analysis of paired end BS-seq

Paired-end quality was assessed using FASTQC [65]. Trimmomatic [66] was used for quality trimming. Parameters for read quality filtering were set as follows: Minimum length of 40 bp; sliding window trimming of 4 bp with required Phred quality score of 20. Trimmed reads were mapped to the *Arabidopsis thaliana* TAIR10 genome assembly using bwa-meth [67] with default parameters. Mapped reads were deduplicated using picardtools [68], and numbers of methylated/unmethylated reads per position were retrieved using MethylExtract [69]. We calculate differentially methylated regions (DMRs) using MethylScore [70].

Data visualization

For visualizing BS-seq genomic data we used Integrative Genomic Viewer (IGV) [71]. For figures, we used R version 3.5.1 (www.r-project.org) with packages ggplot2 [72], eulerr [73] and pheatmap [74].

Supporting information

S1 Fig. Arabidopsis Ws-2 plants from clonal RO-LEC2 and LO-LEC2 G3 progeny grown under different growth conditions.

(EPS)

S2 Fig. DEGs in LO-RKD4 and RO-RKD4 regenerants compared to Ws-2/indRKD4 control plants. Data from three randomly selected regenerants for each tissue.

(EPS)

S3 Fig. DEGs in LO-LEC2 and RO-LEC2 regenerants compared to Ws-2/indLECs control plants. Data from three randomly selected regenerants for each tissue.

(EPS)

S4 Fig. DNA methylation variation in LEC2 regenerants. (A). CG methylation levels at 226 CG-DMRs identified in an all-against-all comparison. (B). CHG methylation levels at 120 CHG-DMRs identified in an all-against-all comparison. (C). CHH methylation levels at 386 CHH-DMRs identified in an all-against-all comparison.

(EPS)

S5 Fig. DNA methylation variation in LEC2 regenerants affects all contexts and appears mainly in proximity to genes. (A). PCAs of CHG DNA methylation levels at 120 CHG-DMRs identified in all-against-all comparison (left two panels) and PCAs of CHH methylation levels at 386 CHH-DMRs identified in all-against-all comparison (right two panels). Percentages in brackets indicate the proportion of variance explained by each principal component (PC). (B) Distribution of different genetic features in DMRs identified in all-against-all comparison for all three contexts, compared to regions identified as methylated in each sample and to the distribution of features across the whole genome. Intergenic is defined as all sequences more than 2 kb away from a gene.

(EPS)

S6 Fig. Distribution of SNPs associated with early senescence of RO-LEC2 plants. Allelic frequencies across chromosomes of foreground markers derived from comparing phenotypically positive (foreground) against phenotypically negative (background) plants.

(EPS)

S7 Fig. Hypermethylated RO-LEC2 DMRs revert to normal methylation levels after backcrossing. Heatmaps showing gains and losses of CHG (A) and CHH (B) DNA methylation levels for RO-LEC2 leaf DMRs in F₂ progeny from reciprocal backcrosses between control (Ws-2/indLEC2) and RO-LEC2 plants. Each column represents plants from independent F₂ progenies.

(EPS)

S8 Fig. Epigenome browser view of AtSIN4 genomic region showing DNA methylation, ATAC-seq, H3K27me3 and Hi-C data from roots and shoots.

(EPS)

S9 Fig. Expression of *PRE1/BNQ1* is associated with the hypomethylation of a flanking *AtMUNX1* transposon.

(EPS)

S1 Table. Differentially Expressed Genes in Ws-2/LEC2 regenerated lines.

(XLSX)

S2 Table. Differentially Expressed Genes in Ws-2/RKD4 regenerated lines.

(XLSX)

S3 Table. List of Differentially Methylated Regions at Differentially Expressed Genes in Ws-2/LEC2 regenerated lines.

(XLSX)

Acknowledgments

We thank Gary Grant for help with plant husbandry, and Liliana M. Costa for discussions and comments on the manuscript.

Author Contributions

Conceptualization: Anjar Tri Wibowo, Detlef Weigel, Jose Gutierrez-Marcos.

Data curation: Anjar Tri Wibowo, Javier Antunez-Sanchez, Alexander Dawson, Jonathan Price, Cathal Meehan, Travis Wrightsman, Maximillian Collenberg, Ilja Bezrukov, Claude Becker.

Formal analysis: Anjar Tri Wibowo, Javier Antunez-Sanchez, Alexander Dawson, Jonathan Price, Cathal Meehan, Travis Wrightsman, Maximillian Collenberg, Ilja Bezrukov, Claude Becker, Moussa Benhamed.

Funding acquisition: Detlef Weigel, Jose Gutierrez-Marcos.

Investigation: Anjar Tri Wibowo, Javier Antunez-Sanchez, Claude Becker.

Project administration: Detlef Weigel, Jose Gutierrez-Marcos.

Writing – original draft: Anjar Tri Wibowo, Javier Antunez-Sanchez, Detlef Weigel, Jose Gutierrez-Marcos.

Writing – review & editing: Anjar Tri Wibowo, Javier Antunez-Sanchez, Detlef Weigel, Jose Gutierrez-Marcos.

References

1. de Meeus T, Prugnolle F, Agnew P. Asexual reproduction: genetics and evolutionary aspects. *Cell Mol Life Sci.* 2007; 64(11):1355–72. <https://doi.org/10.1007/s00018-007-6515-2> PMID: 17396223.
2. Zhang Y, Zhang D. Asexual and sexual reproductive strategies in clonal plants. *Frontiers of Biology in China.* 2007; 2(3):256–62. <https://doi.org/10.1007/s11515-007-0036-0>
3. Ikeuchi M, Ogawa Y, Iwase A, Sugimoto K. Plant regeneration: cellular origins and molecular mechanisms. *Development.* 2016; 143(9):1442–51. <https://doi.org/10.1242/dev.134668> PMID: 27143753.
4. Zheng Q, Zheng Y, Perry SE. AGAMOUS-Like15 promotes somatic embryogenesis in Arabidopsis and soybean in part by the control of ethylene biosynthesis and response. *Plant Physiol.* 2013; 161(4):2113–27. Epub 20130301. <https://doi.org/10.1104/pp.113.216275> PMID: 23457229; PubMed Central PMCID: PMC3613480.
5. Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, et al. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell.* 2002; 14(8):1737–49. <https://doi.org/10.1105/tpc.001941> PMID: 12172019; PubMed Central PMCID: PMC151462.

6. Stone SL, Braybrook SA, Paula SL, Kwong LW, Meuser J, Pelletier J, et al. Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proc Natl Acad Sci U S A*. 2008; 105(8):3151–6. Epub 20080219. <https://doi.org/10.1073/pnas.0712364105> PMID: 18287041; PubMed Central PMCID: PMC2268600.
7. Bouchabke-Coussa O, Obellianne M, Linderme D, Montes E, Maia-Grondard A, Vilaine F, et al. Wuschel overexpression promotes somatic embryogenesis and induces organogenesis in cotton (*Gossypium hirsutum* L.) tissues cultured in vitro. *Plant Cell Rep*. 2013; 32(5):675–86. Epub 20130330. <https://doi.org/10.1007/s00299-013-1402-9> PMID: 23543366.
8. Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho MJ, et al. Morphogenic Regulators Baby boom and Wuschel Improve Monocot Transformation. *Plant Cell*. 2016; 28(9):1998–2015. Epub 20160906. <https://doi.org/10.1105/tpc.16.00124> PMID: 27600536; PubMed Central PMCID: PMC5059793.
9. Wang K, Shi L, Liang X, Zhao P, Wang W, Liu J, et al. The gene TaWOX5 overcomes genotype dependency in wheat genetic transformation. *Nat Plants*. 2022; 8(2):110–7. Epub 20220113. <https://doi.org/10.1038/s41477-021-01085-8> PMID: 35027699.
10. Waki T, Hiki T, Watanabe R, Hashimoto T, Nakajima K. The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. *Curr Biol*. 2011; 21(15):1277–81. Epub 20110728. <https://doi.org/10.1016/j.cub.2011.07.001> PMID: 21802301.
11. Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, et al. Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*. 2015; 525(7570):533–7. Epub 20150909. <https://doi.org/10.1038/nature15365> PMID: 26352475; PubMed Central PMCID: PMC4857894.
12. Wibowo A, Becker C, Durr J, Price J, Spaepen S, Hilton S, et al. Partial maintenance of organ-specific epigenetic marks during plant asexual reproduction leads to heritable phenotypic variation. *Proc Natl Acad Sci U S A*. 2018; 115(39):E9145–E52. Epub 20180910. <https://doi.org/10.1073/pnas.1805371115> PMID: 30201727; PubMed Central PMCID: PMC6166847.
13. Jiang C, Mithani A, Gan X, Belfield EJ, Klingler JP, Zhu JK, et al. Regenerant Arabidopsis lineages display a distinct genome-wide spectrum of mutations conferring variant phenotypes. *Curr Biol*. 2011; 21(16):1385–90. Epub 20110728. <https://doi.org/10.1016/j.cub.2011.07.002> PMID: 21802297; PubMed Central PMCID: PMC3162137.
14. Miyao A, Nakagome M, Ohnuma T, Yamagata H, Kanamori H, Katayose Y, et al. Molecular spectrum of somaclonal variation in regenerated rice revealed by whole-genome sequencing. *Plant Cell Physiol*. 2012; 53(1):256–64. Epub 20111207. <https://doi.org/10.1093/pcp/pcr172> PMID: 22156226.
15. Han Z, Crisp PA, Stelpflug S, Kaeppler SM, Li Q, Springer NM. Heritable Epigenomic Changes to the Maize Methylome Resulting from Tissue Culture. *Genetics*. 2018; 209(4):983–95. Epub 20180530. <https://doi.org/10.1534/genetics.118.300987> PMID: 29848487; PubMed Central PMCID: PMC6063226.
16. Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, et al. Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife*. 2013; 2:e00354. Epub 20130319. <https://doi.org/10.7554/eLife.00354> PMID: 23539454; PubMed Central PMCID: PMC3601819.
17. Rivas-San Vicente M, Plasencia J. Salicylic acid beyond defence: its role in plant growth and development. *J Exp Bot*. 2011; 62(10):3321–38. Epub 20110228. <https://doi.org/10.1093/jxb/err031> PMID: 21357767.
18. Aerts N, de Bruijn S, van Mourik H, Angenent GC, van Dijk ADJ. Comparative analysis of binding patterns of MADS-domain proteins in Arabidopsis thaliana. *BMC Plant Biol*. 2018; 18(1):131. Epub 20180625. <https://doi.org/10.1186/s12870-018-1348-8> PMID: 29940855; PubMed Central PMCID: PMC6019531.
19. Mara CD, Huang T, Irish VF. The Arabidopsis floral homeotic proteins APETALA3 and PISTILLATA negatively regulate the BANQUO genes implicated in light signaling. *Plant Cell*. 2010; 22(3):690–702. Epub 20100319. <https://doi.org/10.1105/tpc.109.065946> PMID: 20305124; PubMed Central PMCID: PMC2861465.
20. Barwale UB, Widholm JM. Somaclonal variation in plants regenerated from cultures of soybean. *Plant Cell Rep*. 1987; 6(5):365–8. <https://doi.org/10.1007/BF00269562> PMID: 24248847.
21. Etienne H, Bertrand B. Somaclonal variation in Coffea arabica: effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. *Tree Physiol*. 2003; 23(6):419–26. <https://doi.org/10.1093/treephys/23.6.419> PMID: 12642244.
22. Tedeschi F, Rizzo P, Rutten T, Altschmied L, Baumlein H. RWP-RK domain-containing transcription factors control cell differentiation during female gametophyte development in Arabidopsis. *New Phytol*. 2017; 213(4):1909–24. Epub 20161121. <https://doi.org/10.1111/nph.14293> PMID: 27870062.
23. Chwialkowska K, Nowakowska U, Mroziewicz A, Szarejko I, Kwasniewski M. Water-deficiency conditions differently modulate the methylome of roots and leaves in barley (*Hordeum vulgare* L.). *J Exp Bot*.

- 2016; 67(4):1109–21. Epub 20160105. <https://doi.org/10.1093/jxb/erv552> PMID: 26739862; PubMed Central PMCID: PMC4753852.
24. Ferreira LJ, Azevedo V, Maroco J, Oliveira MM, Santos AP. Salt Tolerant and Sensitive Rice Varieties Display Differential Methylome Flexibility under Salt Stress. *PLoS One*. 2015; 10(5):e0124060. Epub 20150501. <https://doi.org/10.1371/journal.pone.0124060> PMID: 25932633; PubMed Central PMCID: PMC4416925.
 25. Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D. Evolution of DNA methylation patterns in the Brassicaceae is driven by differences in genome organization. *PLoS Genet*. 2014; 10(11):e1004785. Epub 20141113. <https://doi.org/10.1371/journal.pgen.1004785> PMID: 25393550; PubMed Central PMCID: PMC4230842.
 26. Vining KJ, Pomraning KR, Wilhelm LJ, Priest HD, Pellegrini M, Mockler TC, et al. Dynamic DNA cytosine methylation in the *Populus trichocarpa* genome: tissue-level variation and relationship to gene expression. *BMC Genomics*. 2012; 13:27. Epub 20120117. <https://doi.org/10.1186/1471-2164-13-27> PMID: 22251412; PubMed Central PMCID: PMC3298464.
 27. Widman N, Feng S, Jacobsen SE, Pellegrini M. Epigenetic differences between shoots and roots in *Arabidopsis* reveals tissue-specific regulation. *Epigenetics*. 2014; 9(2):236–42. Epub 20131029. <https://doi.org/10.4161/epi.26869> PMID: 24169618; PubMed Central PMCID: PMC3962533.
 28. Huang Y, Sicar S, Ramirez-Prado JS, Manza-Mianza D, Antunez-Sanchez J, Brik-Chaouche R, et al. Polycomb-dependent differential chromatin compartmentalization determines gene coregulation in *Arabidopsis*. *Genome Res*. 2021. Epub 20210603. <https://doi.org/10.1101/gr.273771.120> PMID: 34083408; PubMed Central PMCID: PMC8256866.
 29. Lupianez DG, Spielmann M, Mundlos S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet*. 2016; 32(4):225–37. Epub 20160207. <https://doi.org/10.1016/j.tig.2016.01.003> PMID: 26862051.
 30. Zheng H, Xie W. The role of 3D genome organization in development and cell differentiation. *Nat Rev Mol Cell Biol*. 2019; 20(9):535–50. <https://doi.org/10.1038/s41580-019-0132-4> PMID: 31197269.
 31. Kawakatsu T, Huang SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, et al. Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions. *Cell*. 2016; 166(2):492–505. <https://doi.org/10.1016/j.cell.2016.06.044> PMID: 27419873; PubMed Central PMCID: PMC5172462.
 32. Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Paolo Casale F, et al. DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *Elife*. 2015; 4:e05255. Epub 20150505. <https://doi.org/10.7554/eLife.05255> PMID: 25939354; PubMed Central PMCID: PMC4413256.
 33. Wibowo A, Becker C, Marconi G, Durr J, Price J, Hagmann J, et al. Hyperosmotic stress memory in *Arabidopsis* is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *Elife*. 2016; 5. Epub 20160531. <https://doi.org/10.7554/eLife.13546> PMID: 27242129; PubMed Central PMCID: PMC4887212.
 34. Zhang B, Tieman DM, Jiao C, Xu Y, Chen K, Fei Z, et al. Chilling-induced tomato flavor loss is associated with altered volatile synthesis and transient changes in DNA methylation. *Proc Natl Acad Sci U S A*. 2016; 113(44):12580–5. Epub 20161017. <https://doi.org/10.1073/pnas.1613910113> PMID: 27791156; PubMed Central PMCID: PMC5098663.
 35. Zhou W, Liang G, Molloy PL, Jones PA. DNA methylation enables transposable element-driven genome expansion. *Proc Natl Acad Sci U S A*. 2020; 117(32):19359–66. Epub 20200727. <https://doi.org/10.1073/pnas.1921719117> PMID: 32719115; PubMed Central PMCID: PMC7431005.
 36. Peng Y, Yang J, Li X, Zhang Y. Salicylic Acid: Biosynthesis and Signaling. *Annu Rev Plant Biol*. 2021; 72:761–91. Epub 20210323. <https://doi.org/10.1146/annurev-arplant-081320-092855> PMID: 33756096.
 37. Thomas H, Howarth CJ. Five ways to stay green. *J Exp Bot*. 2000; 51 Spec No:329–37. https://doi.org/10.1093/jexbot/51.suppl_1.329 PMID: 10938840.
 38. Yoo SC, Cho SH, Zhang H, Paik HC, Lee CH, Li J, et al. Quantitative trait loci associated with functional stay-green SNU-SG1 in rice. *Mol Cells*. 2007; 24(1):83–94. PMID: 17846502.
 39. Spano G, Di Fonzo N, Perrotta C, Platani C, Ronga G, Lawlor DW, et al. Physiological characterization of 'stay green' mutants in durum wheat. *J Exp Bot*. 2003; 54(386):1415–20. <https://doi.org/10.1093/jxb/erg150> PMID: 12709488.
 40. Choi J, Strickler SR, Richards EJ. Loss of CRWN Nuclear Proteins Induces Cell Death and Salicylic Acid Defense Signaling. *Plant Physiol*. 2019; 179(4):1315–29. Epub 20190129. <https://doi.org/10.1104/pp.18.01020> PMID: 30696746; PubMed Central PMCID: PMC6446779.
 41. Hu B, Wang N, Bi X, Karaaslan ES, Weber AL, Zhu W, et al. Plant lamin-like proteins mediate chromatin tethering at the nuclear periphery. *Genome Biol*. 2019; 20(1):87. Epub 20190430. <https://doi.org/10.1186/s13059-019-1694-3> PMID: 31039799; PubMed Central PMCID: PMC6492433.

42. Cortijo S, Wardenaar R, Colome-Tatche M, Johannes F, Colot V. Genome-wide analysis of DNA methylation in Arabidopsis using MeDIP-chip. *Methods Mol Biol.* 2014; 1112:125–49. https://doi.org/10.1007/978-1-62703-773-0_9 PMID: 24478012.
43. Gallego-Bartolome J, Gardiner J, Liu W, Papikian A, Ghoshal B, Kuo HY, et al. Targeted DNA demethylation of the Arabidopsis genome using the human TET1 catalytic domain. *Proc Natl Acad Sci U S A.* 2018; 115(9):E2125–E34. Epub 20180214. <https://doi.org/10.1073/pnas.1716945115> PMID: 29444862; PubMed Central PMCID: PMC5834696.
44. Ji L, Jordan WT, Shi X, Hu L, He C, Schmitz RJ. TET-mediated epimutagenesis of the Arabidopsis thaliana methylome. *Nature Communications.* 2018; 9(1):895. <https://doi.org/10.1038/s41467-018-03289-7> PMID: 29497035
45. Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, et al. Assessing the Impact of Transgenerational Epigenetic Variation on Complex Traits. *PLoS Genetics.* 2009; 5(6): e1000530–e. <https://doi.org/10.1371/journal.pgen.1000530> PMID: 19557164
46. Reinders J, Wulff BBH, Mirouze M, Mari-Ordonez A, Dapp M, Rozhon W, et al. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & Development.* 2009; 23(8):939–50. <https://doi.org/10.1101/gad.524609> PMID: 19390088
47. Antunez-Sanchez J, Naish M, Ramirez-Prado JS, Ohno S, Huang Y, Dawson A, et al. A new role for histone demethylases in the maintenance of plant genome integrity. *Elife.* 2020; 9. Epub 20201027. <https://doi.org/10.7554/eLife.58533> PMID: 33107825; PubMed Central PMCID: PMC7671693.
48. Borges F, Calarco JP, Martienssen RA. Reprogramming the epigenome in Arabidopsis pollen. *Cold Spring Harb Symp Quant Biol.* 2012; 77:1–5. Epub 20130425. <https://doi.org/10.1101/sqb.2013.77.014969> PMID: 23619015; PubMed Central PMCID: PMC5556929.
49. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T, et al. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell.* 2012; 151(1):194–205. Epub 20120920. <https://doi.org/10.1016/j.cell.2012.09.001> PMID: 23000270; PubMed Central PMCID: PMC3697483.
50. Papareddy RK, Paldi K, Paulraj S, Kao P, Lutzmayr S, Nodine MD. Chromatin regulates expression of small RNAs to help maintain transposon methylome homeostasis in Arabidopsis. *Genome Biol.* 2020; 21(1):251. Epub 20200917. <https://doi.org/10.1186/s13059-020-02163-4> PMID: 32943088; PubMed Central PMCID: PMC7499886.
51. Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, et al. Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci U S A.* 2006; 103(9):3468–73. Epub 20060221. <https://doi.org/10.1073/pnas.0511331103> PMID: 16492731; PubMed Central PMCID: PMC1413938.
52. Curtis MD, Grossniklaus U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* 2003; 133(2):462–9. <https://doi.org/10.1104/pp.103.027979> PMID: 14555774; PubMed Central PMCID: PMC523872.
53. Shimada TL, Shimada T, Hara-Nishimura I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *Plant J.* 2010; 61(3):519–28. Epub 20091125. <https://doi.org/10.1111/j.1365-3113X.2009.04060.x> PMID: 19891705.
54. Krizek BA, Meyerowitz EM. The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development.* 1996; 122(1):11–22. <https://doi.org/10.1242/dev.122.1.11> PMID: 8565821.
55. Defraia CT, Schmelz EA, Mou Z. A rapid biosensor-based method for quantification of free and glucose-conjugated salicylic acid. *Plant Methods.* 2008; 4:28. Epub 20081231. <https://doi.org/10.1186/1746-4811-4-28> PMID: 19117519; PubMed Central PMCID: PMC2654556.
56. Abreu ME, Munne-Bosch S. Salicylic acid deficiency in NahG transgenic lines and sid2 mutants increases seed yield in the annual plant Arabidopsis thaliana. *J Exp Bot.* 2009; 60(4):1261–71. Epub 20090202. <https://doi.org/10.1093/jxb/ern363> PMID: 19188277; PubMed Central PMCID: PMC2657544.
57. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 2017; 27(5):722–36. Epub 20170315. <https://doi.org/10.1101/gr.215087.116> PMID: 28298431; PubMed Central PMCID: PMC5411767.
58. Linthorst J, Hulsman M, Holstege H, Reinders M. Scalable multi whole-genome alignment using recursive exact matching. *bioRxiv.* 2015. <https://doi.org/10.1101/022715>
59. Fiddes IT, Armstrong J, Diekhans M, Nachtweide S, Kronenberg ZN, Underwood JG, et al. Comparative Annotation Toolkit (CAT)-simultaneous clade and personal genome annotation. *Genome Res.* 2018; 28(7):1029–38. Epub 20180608. <https://doi.org/10.1101/gr.233460.117> PMID: 29884752; PubMed Central PMCID: PMC6028123.

60. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* 2010; 38(Web Server issue):W64–70. Epub 20100430. <https://doi.org/10.1093/nar/gkq310> PMID: 20435677; PubMed Central PMCID: PMC2896167.
61. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018; 34(17):i884–i90. <https://doi.org/10.1093/bioinformatics/bty560> PMID: 30423086; PubMed Central PMCID: PMC6129281.
62. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9(4):357–9. Epub 20120304. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286; PubMed Central PMCID: PMC3322381.
63. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9. Epub 20090608. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943; PubMed Central PMCID: PMC2723002.
64. Sun H, Schneeberger K. SHOREmap v3.0: fast and accurate identification of causal mutations from forward genetic screens. *Methods Mol Biol.* 2015; 1284:381–95. https://doi.org/10.1007/978-1-4939-2444-8_19 PMID: 25757783.
65. Andrews S, Krueger F. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc2010>.
66. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30(15):2114–20. Epub 20140401. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404; PubMed Central PMCID: PMC4103590.
67. Pedersen BS, Eyring K, D S., Yang IV, Schwartz DA. Fast and accurate alignment of long bisulfite-seq reads 2014.
68. toolkit P. Picard toolkit. Broad Institute, GitHub repository. 2019.
69. Barturen G, Rueda A, Oliver JL, Hackenberg M. MethylExtract: High-Quality methylation maps and SNV calling from whole genome bisulfite sequencing data. *F1000Res.* 2013; 2:217. Epub 20131015. <https://doi.org/10.12688/f1000research.2-217.v2> PMID: 24627790; PubMed Central PMCID: PMC3938178.
70. Hüther P, Hagmann J, Nunn A, Kakoulidou I, Pisupati R, Langenberger D, et al. MethylScore, a pipeline for accurate and context-aware identification of differentially methylated regions from population-scale plant WGBS data. *bioRxiv.* 2022:2022.01.06.475031. <https://doi.org/10.1101/2022.01.06.475031>
71. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 2013; 14(2):178–92. Epub 20120419. <https://doi.org/10.1093/bib/bbs017> PMID: 22517427; PubMed Central PMCID: PMC3603213.
72. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.
73. Larsson J. eulerr: Area-proportional Euler and Venn diagrams with ellipses. 2019.
74. Kolde R. pheatmap: Pretty heatmaps [Software]. 2015.