PNAS
www.pnas.org

Supplementary Information for:

Transcriptional profiling reveals signatures of latent developmental potential in Arabidopsis stomatal lineage ground cells

Chin-Min Kimmy Ho[1,4*], Martin Bringmann[1], Yoshimi Oshima[3], Nobutaka Mitsuda[3], Dominique C. Bergmann[1, 2*]
[1] Department of Biology, Stanford University, Stanford, CA, 94305-5020, USA
[2] Howard Hughes Medical Institute
[3] Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Higashi 1-1-1, Tsukuba 305-8562, Japan
[4] Present address, Institute of Plant and Microbial Biology, Academia Sinica, Taipei, 115, Taiwan

*Correspondence chmho@gate.sinica.edu.tw, dbergmann@stanford.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Datasets S1 to S6

SUPPLEMENTARY METHODS

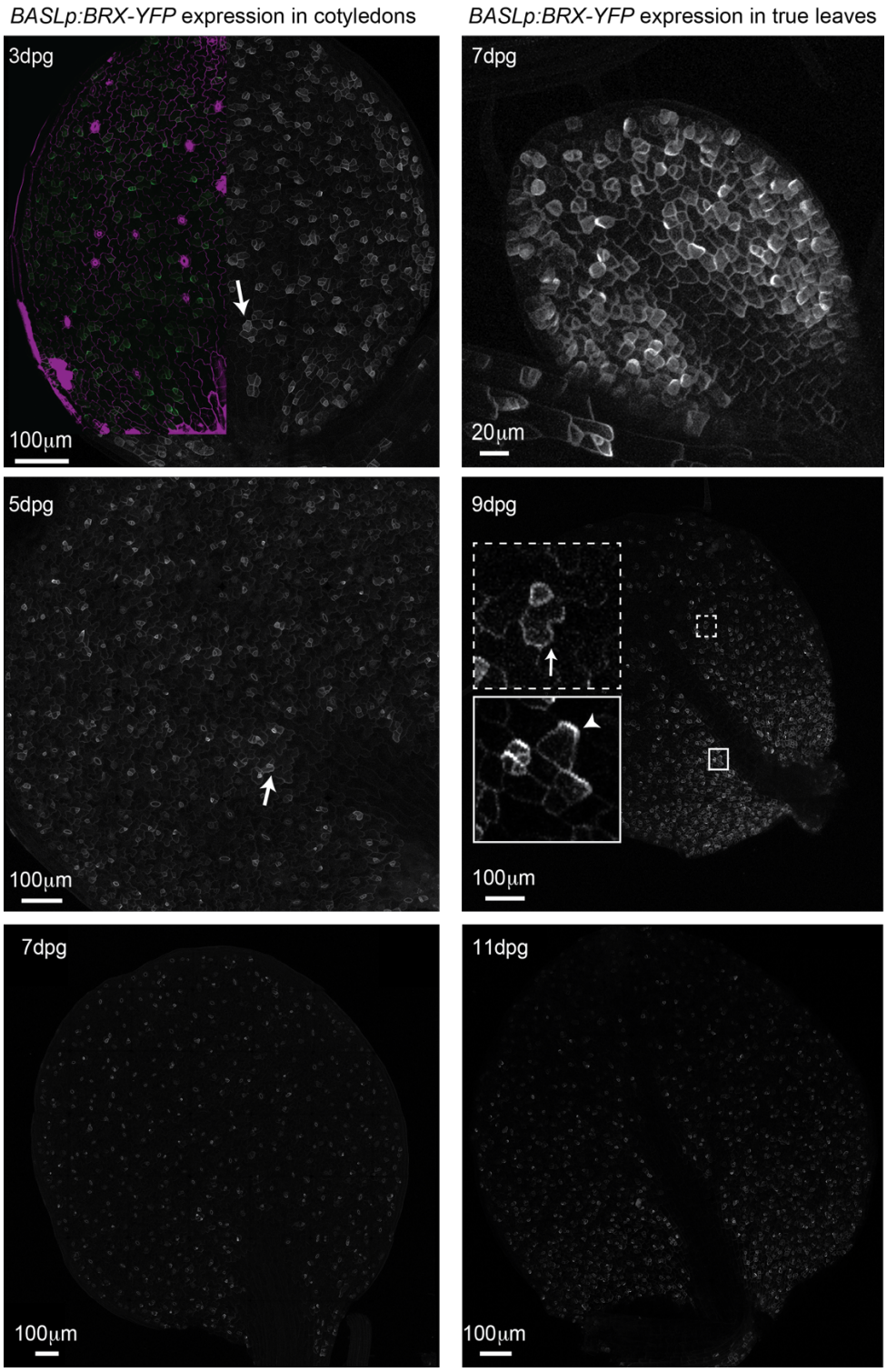
Plant reporter lines and mutants

Col-0 served as the WT in all experiments and all transgenics were made in this accession. DEKp:DEK-YFP and DEKlikep:DEKlike-YFP was created through Gateway (Thermo Fisher Scientific) cloning using pENTR and pHGY (1) with primers DEKproF/DEKR and DEKlikepro2_F_NotI/DEKlike_R_AscI. Mutant alleles of *DEK-LIKE* (Salk_097030) and *DEK* (Salk_014079c) were obtained from ABRC and confirmed by PCR using primer sets of DEK_Salk_014079_LP / DEK_Salk_014079_RP and DEKlike_Salk_097030 LP / DEKlike_Salk_097030 RP. *dek* and *deklike CRISPR* lines were made using YAO promoter-driven CRISPR/Cas9 system and guide RNA were designed as shown in Figure 4E-F (2). *DEK* and *DEK-like* gene expression was detected by dek-qPCR-F/dek-qPCR-R and deklike-qPCR-F/deklike-qPCR-R. Fibrillarin-RFP plasmid (CD3-795059) (3) was obtained from ABRC and transformed into Arabidopsis Col-0. To construct MYB16pro:MYB16-GFP, the protein-coding region of MYB16 was amplified by PCR using primers: AT5G15310N and AT5G15310C and was cloned into SmaI site of p35SGFPG (which was generated by insertion of a GFP fragment using primers: GFP-N and GFP-C into SmaI and SalI site of p35SG (4)). The MYB16-GFP fragment was amplified using a primer set: attB1-AT5G15310F/attB2-GFP and cloned into pDONR207 by Gateway BP reaction (Thermo Fisher Scientific). The contents of pDONR207_MYB16GFP and pDONRG-P4P1R_MYB16pro (5) were transferred using LR reaction (Thermo Fisher Scientific) into R4pGWB4_stop_HSP T-DNA vector (4). MYB16-SRDX was reported previously (5). Primers are listed in Table S1. **Phylogeny:** Protein sequences of SLIDEs and DEKs were obtained from TAIR and analyzed by MEGA6 (6). All sequences were aligned by ClustalW with default settings and the neighbor-joining method with 1000 bootstraps was used for constructing the phylogenetic tree. **Transient Expression:** DEKp:DEK-YFP and DEKlikep:DEKlike-YFP were transformed into *Arabidopsis* mesophyll protoplasts with Arabidopsis Fibrillarin-RFP (CD3-795095) following the protocol in (7) or into *N. benthamiana* with Nb-Fibrillarin-RFP (8).

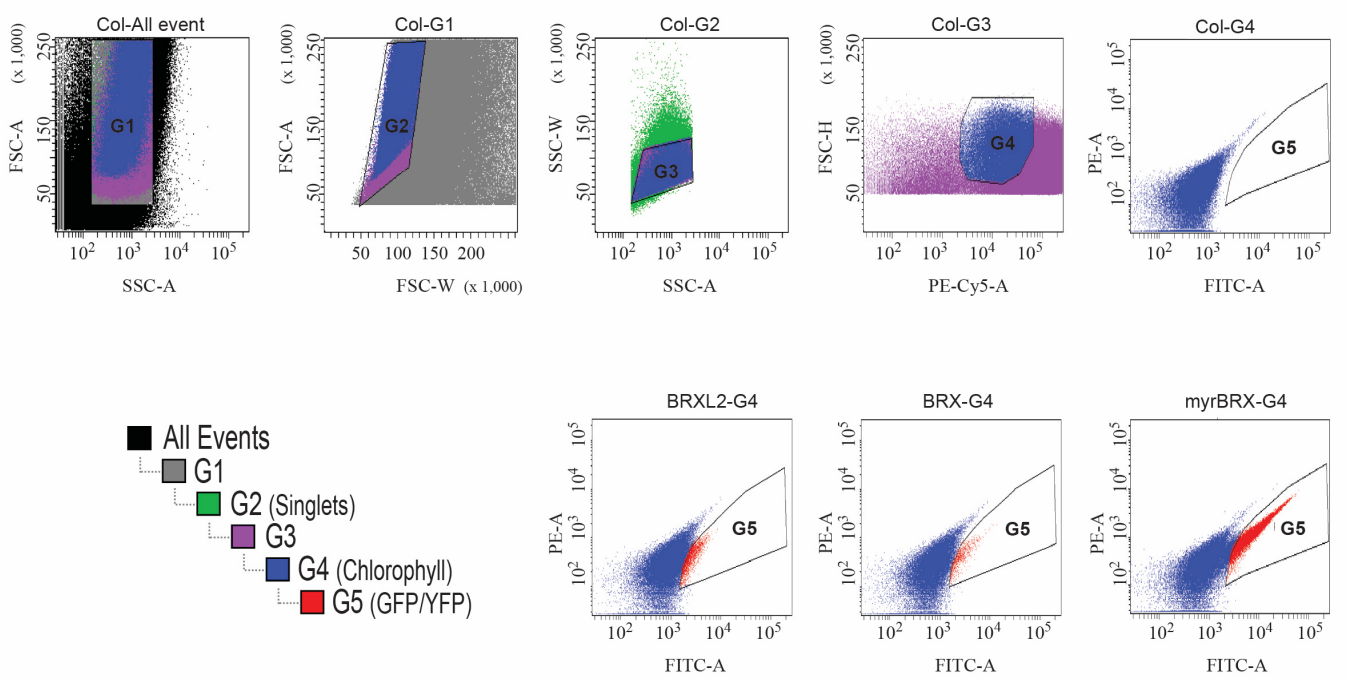
Generation of SLGC transcriptomes:

9 dpg seedings of reporter lines were used for protoplast isolation and FACS as described in (9, 10). FACS was performed on Aria II (BD biosciences) with an 80 mm nozzle. Approximately 20,000 YFP positive cells/sample were sorted into 350 µl RNA extraction buffer and total RNA were extracted using RNeasy Micro Kit (QIAGEN). Cell commitment stages A and B correspond to MUTEpro:nucYFP (10) and MUTEpro:MUTE-YFP, respectively. The MUTEpro:MUTE-YFP transcriptome was generated via the same protocols as in Adrian, 2015 (10)**.** 20 ng total RNA with RIN (Agilent bioanalyzer) greater than 8 from each sample was submitted to Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley for mRNA isolation (NuGen-Ovation® RNA-Seq System V2) and cDNA library construction. Fifty-bp single end reads were generated from a HiSeq2000 sequencer (Illumina) in high-output mode. All sequencing and data analysis were done using Illumina's standard protocols. Reads were mapped to TAIR10.18 via Bowtie2 (11) (read statistics in Dataset S1). TPM normalized counts were analyzed by ICI (scheme in Figure S2B). To generate an SLGC cluster (Figure S2C), counts were normalized via via DESeq2 (12) using default settings and DE genes (19707 genes) were obtained from all possible pairwise comparisons with FDR < 0.05, then clustered via FANNY (13) with $k = 5$ and a probability cutoff of 0.6.

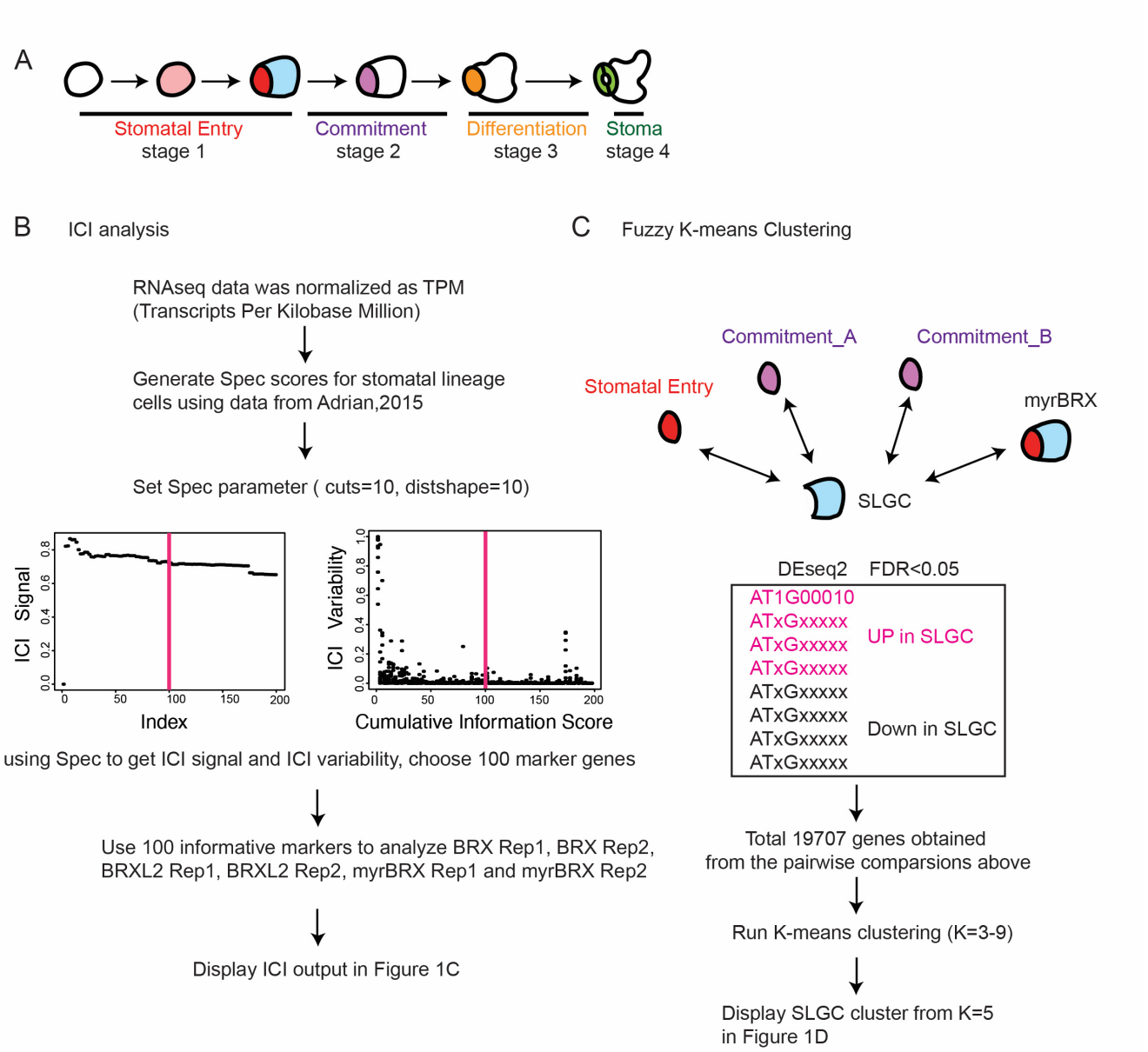
RNA-seq datasets are in GEO with accession number GSE129938.



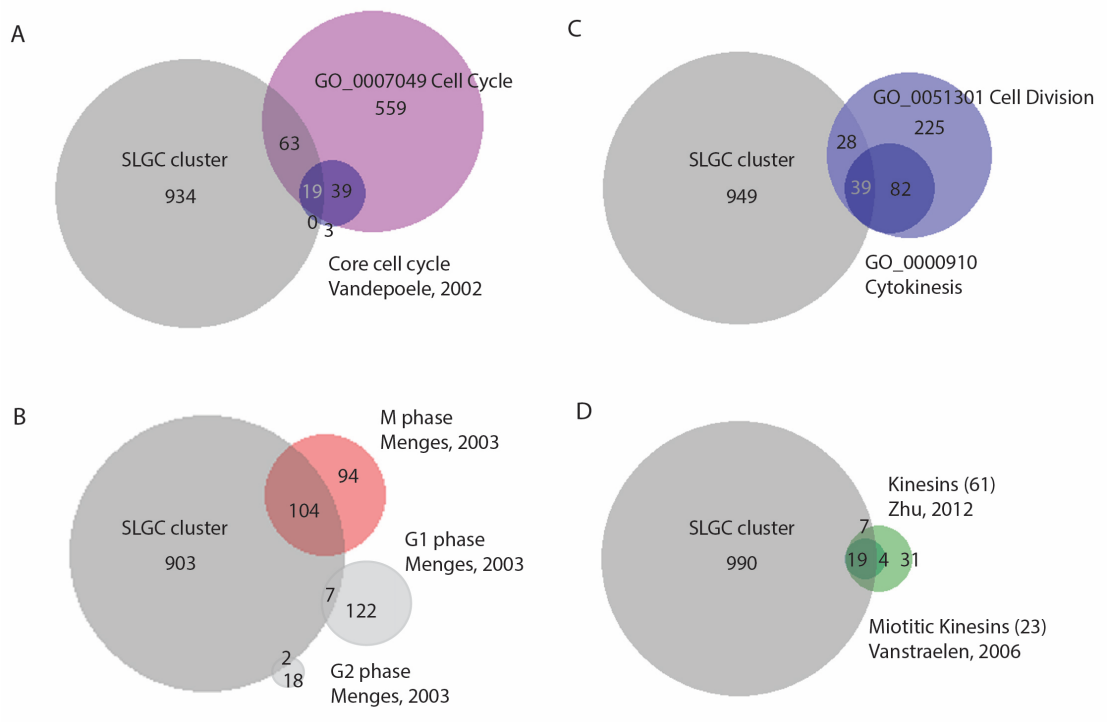
Supplemental Figure 1. Confocal imaging timecourse of *BASLp:BRX-YFP* expression pattern in cotyledons and leaves. Top left panel includes overlay of cell outline marker (magenta) for general orientation relative to BRX-YFP (green). Signal in all other panels is only BRX-YFP. At all stages, BRX-YFP is primarily in meristemoids and young SLGCs (arrowhead in inset at 9dpg), and occasionally in small cells that had begun to lobe (arrows). BRX-YFP was never observed in large, lobed pavement cells. All images were acquired at the same laser intensity.



Supplemental Figure 2. Parameters used for FACS isolation of stomatal lineage cells. Forward scatter (FCS) and side scatter (SSC) were used to identify singlets (Gates (G)1-G3). To discriminate against mesophyll cells, protoplasts with relatively lower amounts of chlorophyll (G4) were selected. YFP positive cells (G5) were identified by higher GFP (FITC) and YFP (PE) spectrums when compared with wild-type cells (YFP negative cells).



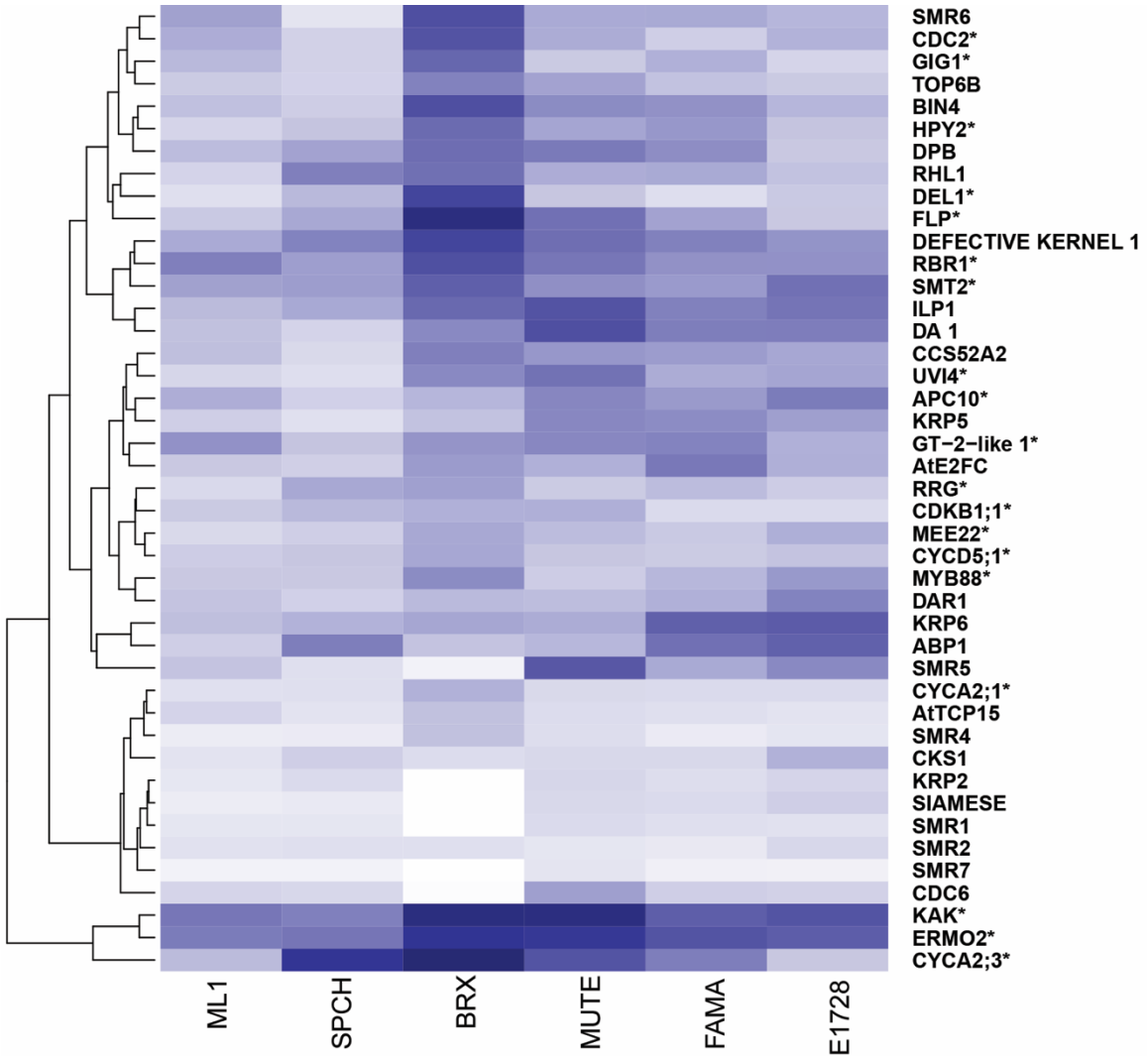
Supplemental Figure 3. Flowcharts of ICI analysis and Fuzzy K-means Clustering. (A) Illustration of stomatal development from entry, commitment, differentiation to stoma stages. (B) General scheme of ICI analysis (14). Stomatal lineage (10) and SLGC RNAseq data generated from this study were TPM (transcripts per kilobase million) normalized for further comparisons. Spec parameter, ICI signal and ICI variability were obtained using stomatal lineage profiles (10) and Dataset S3. 6 SLGC samples: BRX rep1/2, BRXL2 rep1/2 and myrBRX rep1/2 with marker number (100) were used for ICI approach and output was shown in Figure 1C. (C) Generation of SLGC enriched cluster by fuzzy K-means clustering. Pairwise comparisons via DEseq2 were made with BASLp:BRX-YFP sample to stomatal entry, commitment_A (MUTEp:nucGFP), commitment_B (MUTEp:MUTE-YFP) and BASLp:myrBRX-YFP samples. The sum of differentially expressed (DE) genes with FDR<0.05 were clustered via FANNY (13). The program was run with cluster numbers K=3 through K=9, with K=5 empirically determined to best recapitulate known biology. The SLGC cluster in Figure 1D was obtained with K=5 and a probability cutoff=0.6.



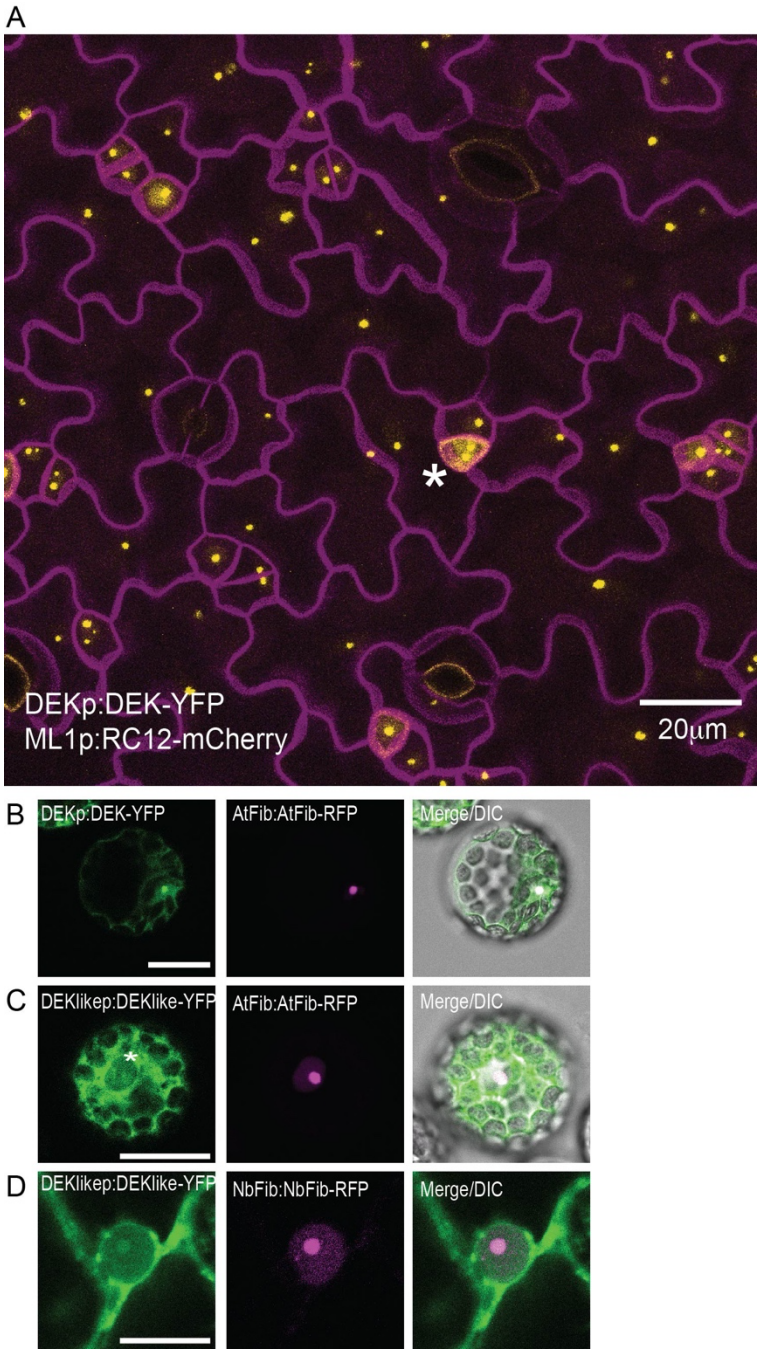
Supplemental Figure 4. Comparison of genes in the SLGC cluster with previously published cell cycle-related datasets.
Genes in the SLGC cluster comparison with (A) cell cycle Gene Ontology Term (GO:0007049) and core cell cycle genes (15), (B) G1, G2 and M phase associated genes (16) (C) cell division (GO:0051301) and cytokinesis (GO:0000910) GO terms, (D) sixty-one Arabidopsis kinesins (17, 18) and mitotic kinesin (19).

Overlap SLGC cluster- Cell cycle (GO_0007049): P(X>= 82) = 1.14932e-26;
Overlap SLGC cluster- Core cell cycle (Vandepoele): P(X>= 19) = 1.023581e-14;
Overlap SLGC cluster- M phase (Menges): P(X>= 104) = 6.357579e-104;
Overlap SLGC cluster- G1 phase (Menges): P(X>= 7) = 0.09696155;
Overlap SLGC cluster- G2 phase (Menges): P(X>= 1) = 0.1213557;
Overlap SLGC cluster- Cell division (GO_0051301): P(X>= 67) = 3.627694e-32;
Overlap SLGC cluster- Cytokinesis (GO_0000910): P(X>= 28) = 2.042789e-18;
Overlap SLGC cluster- Sixty-one Kinesins : P(X>= 19) = 1.023581e-14;
Overlap SLGC cluster- Mitotic Kinesins : P(X>= 19) = 9.040754e-26.

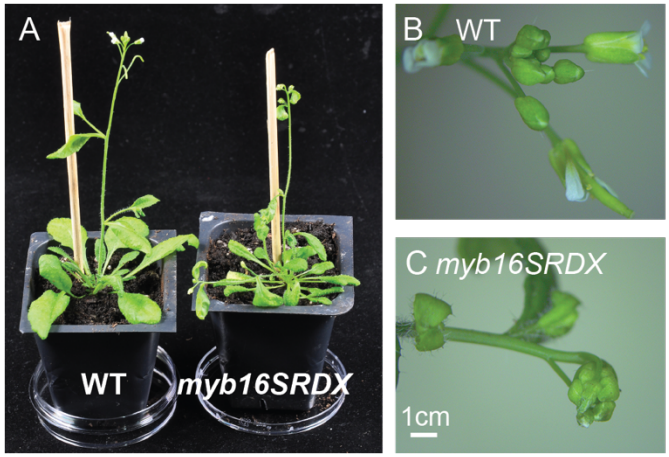
X = Sample size for overlap of genes between two categories using Hypergeometric Test



Supplemental Figure 5. Expression of DNA endoreduplication genes in stomatal lineage stages. Heatmap showing the expression pattern of genes associated with DNA endoreduplication (GO:0042023) in transcriptional profiles of stomatal lineage cells. X-axis is labeled with names of reporters used to isolate cells. Negative regulators are noted with asterisks (*). The BRX (presumptive SLGC cluster) is defined by expression of both positive and negative cell cycle regulators. Two clusters at the bottom of the map related to endoreplication. Repressors of endoreduplication CYCA2;3 (20), KAKTUS (KAK) (21) and ERMO2 (22) peak in BRX/SLGC whereas positive regulators such as KRP2, SIAMESE and SMRs are depleted.



Supplemental Figure 6. Tissue-wise and subcellular expression patterns of DEK and DEK-like (A) Confocal image of DEK translational reporter (yellow) in leaf cells from a larger field of view than Figure 4B. Asterisk marks peak expression in stomatal precursor cell. Cell outlines (magenta) are marked by ML1p:RC12A-mcherry. (B-C) Expression of DEK and DEK-like (green) in *Arabidopsis* protoplasts. Both are in the nucleolus, co-localized with *Arabidopsis* Fibrillarin (magenta), but DEK-like is also abundant in the endoplasmic reticulum (ER). (D) Expression of DEK-like (green) in *N. benthamina* leaf cells, co-localized with *Nicotiana* NbFibrillarin (magenta) and also in the ER. Scale bars represent 20μm.



Supplemental Figure 7. Whole plant phenotypes of *myb16SRDX* lines (A) Whole plant images at bolting for WT (Col-0) and *myb16SRDX lines*, the latter exhibiting slightly smaller leaves. (B) Magnified view of WT flower buds. (C) Magnified view of *myb16SRDX* flower buds showing organ fusions. B-C are at the same magnification.

Supplemental Table 1: Primers used in this study

DEKproF	CACCTATATAGCCCAGTTTTTCAAGTCCAAG
DEKR	AATACTCGTCGACTCTATTCTCTTGGG
DEK_Salk_014079_LP	TCATCAAAGCATCATCCTCATC
DEK_Salk_014079_RP	TCTTCTTTATTTTCCTCCGCAG
DEKlike_Salk_097030 LP	TCGATTGTCATAATCGCCTTC
DEKlike_Salk_097030 RP	TTCTTGTATTGTGGGGACTCC
AT5G15310N	gATGGGTAGATCACCGTGTTGTGACAAATT
AT5G15310C	GAACATCGGTGAATCCGACGGTGAAGGATC
GFP-N	GGGATGGTGAGCAAGGGCGAGGAGCTGTTC
GFP-C	GTCGACTTACTTGTACAGCTCGTCCATG
attB1-AT5G15310F	ggggacaagtttgtacaaaaaagcaggcttcATGGGTAGATCACCGTGTTGTGACAAATT
attB2-GFP	ggggaccactttgtacaagaaagctgggtcTTACTTGTACAGCTCGTCCATGCCGTGAGT
dek-qPCR-F	TCTGAGGAATTTGAGCTCTTGC
dek-qPCR-R	AAGTTTGAAGAGGTGTCCTGCA
deklike-qPCR-F	CAAAATCATCGTCAGAGAAGCC
deklike-qPCR-R	GATTGTCATAATCGCCTTCTCGA
DEK_CRISPRgRNA9_F	ATTGAAGAAGCAGTAAAGTTGAT
DEK_CRISPRgRNA9_R	AAACATCAACTTTACTGCTTCTT
DEKlike_CRISPRgRNA7_F	AATGGAAGAAGAAGCGAGACAAG
DEKlike_CRISPRgRNA7_R	AAACCTTGTCTCGCTTCTTCTTC
DEKlikeProF2_NotI	CACCGCGGCCGCcttgtattgtggggactccttatgtc
DEKlike_R_AscI	CGGggcgcgcccAGCAGGTGATGACTCTTCTTCTTCCAC

Legends for Supplementary Datasets

Dataset S1 Genes differentially expressed between BRX and myrBRX sorted populations

Dataset S2 Genes differentially expressed between BRX and BRXL2 sorted populations

Dataset S3 Informative markers for each cell type used in ICI analysis (10)

Dataset S4 Genes in the SLGC cluster, including cell cycle, cell division and cytokinesis comparisons (data source for Supplemental Figure 4) (10, 23-26)

Dataset S5 Kinesins found in the SLGC cluster related to Arabidopsis kinesins (data source for Figure 2A) (18, 19)

Dataset S6 Genes related to endoreduplication (data source for Supplemental Figure 5)

SI References

1. Kubo M*, et al.* (2005) Transcription switches for protoxylem and metaxylem vessel formation. *Genes & development* 19(16):1855-1860.
2. Yan L*, et al.* (2015) High-Efficiency Genome Editing in Arabidopsis Using YAO Promoter-Driven CRISPR/Cas9 System. *Mol Plant* 8(12):1820-1823.
3. Pendle AF*, et al.* (2005) Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. *Molecular biology of the cell* 16(1):260-269.
4. Oshima Y*, et al.* (2011) Novel vector systems to accelerate functional analysis of transcription factors using chimeric repressor gene-silencing technology (CRES-T). *Plant Biotechnology* 28(2):201-210.
5. Oshima Y*, et al.* (2013) MIXTA-like transcription factors and WAX INDUCER1/SHINE1 coordinately regulate cuticle development in Arabidopsis and Torenia fournieri. *The Plant cell* 25(5):1609-1624.
6. Tamura K, Stecher G, Peterson D, Filipski A, & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution* 30(12):2725-2729.
7. Yoo S-D, Cho Y-H, & Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature protocols* 2(7):1565-1572.
8. Chang CH*, et al.* (2016) The Nucleolar Fibrillarin Protein Is Required for Helper Virus-Independent Long-Distance Trafficking of a Subviral Satellite RNA in Plants. *The Plant cell* 28(10):2586-2602.
9. Bargmann BO & Birnbaum KD (2010) Fluorescence activated cell sorting of plant protoplasts. *J Vis Exp* (36):1673.
10. Adrian J*, et al.* (2015) Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Developmental cell* 33(1):107-118.
11. Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* 9(4):357-359.
12. Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome biology* 11(10):R106.
13. Mächler M, Rousseeuw P, Struyf A, Hubert M, & Hornik K (2012) *Cluster: Cluster Analysis Basics and Extensions*.(R package version 1, 2012)
14. Efroni I, Ip P-L, Nawy T, Mello A, & Birnbaum KD (2015) Quantification of cell identity from single-cell gene expression profiles. *Genome biology* 16(1):9-9.
15. Vandepoele K*, et al.* (2002) Genome-wide analysis of core cell cycle genes in Arabidopsis. *The Plant cell* 14(4):903-916.
16. Menges M, Hennig L, Gruissem W, & Murray JA (2003) Genome-wide gene expression in an Arabidopsis cell suspension. *Plant molecular biology* 53(4):423-442.
17. Lee Y-RJ & Liu B (2004) Cytoskeletal Motors in Arabidopsis. Sixty-One Kinesins and Seventeen Myosins. 136(4):3877-3883.
18. Zhu C & Dixit R (2012) Functions of the Arabidopsis kinesin superfamily of microtubule-based motor proteins. *Protoplasma* 249(4):887-899.
19. Vanstraelen M, Inze D, & Geelen D (2006) Mitosis-specific kinesins in Arabidopsis. *Trends in plant science* 11(4):167-175.
20. Imai KK*, et al.* (2006) The A-Type Cyclin CYCA2;3 Is a Key Regulator of Ploidy Levels in Arabidopsis Endoreduplication. *The Plant cell* 18(2):382.
21. El Refy A*, et al.* (2004) The Arabidopsis KAKTUS gene encodes a HECT protein and controls the number of endoreduplication cycles. *Molecular Genetics and Genomics* 270(5):403-414.
22. Nakano RT*, et al.* (2009) GNOM-LIKE1/ERMO1 and SEC24a/ERMO2 Are Required for Maintenance of Endoplasmic Reticulum Morphology in Arabidopsis thaliana. *The Plant cell* 21(11):3672.
23. Lau OS*, et al.* (2018) Direct Control of SPEECHLESS by PIF4 in the High-Temperature Response of Stomatal Development. *Current biology : CB* 28(8):1273-1280 e1273.
24. Shpak ED, McAbee JM, Pillitteri LJ, & Torii KU (2005) Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science* 309(5732):290-293.

25. MacAlister CA, Ohashi-Ito K, & Bergmann DC (2007) Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature* 445(7127):537-540.
26. Dong J, MacAlister CA, & Bergmann DC (2009) BASL controls asymmetric cell division in Arabidopsis. *Cell* 137(7):1320-1330.