SOG1 activator and MYB3R repressors regulate a complex DNA damage network in Arabidopsis

Clara Bourbousses, Neeraja Vegenas, and Julie A. Law

Edited by Julia Bailey-Serres, University of California, Riverside, CA, and approved November 14, 2018 (received for review June 21, 2018)

To combat DNA damage, organisms mount a DNA damage response (DDR) that results in cell cycle regulation, DNA repair and, in severe cases, cell death. Underscoring the importance of gene regulation in this process, studies in Arabidopsis have demonstrated that all of the aforementioned processes rely on SUPPRESSOR OF GAMMA RESPONSE 1 (SOG1), a NAC family transcription factor (TF) that has been functionally equated to the mammalian tumor suppressor, p53. However, the expression networks connecting SOG1 to these processes remain largely unknown and, although the DDR spans from minutes to hours, most transcriptomic data correspond to single time-point snapshots. Here, we generated transcriptional models of the DDR from GAMMA (γ)-irradiated wild-type and sog1 seedlings during a 24-hour time course using DREM, the Dynamic Regulatory Events Miner, revealing 11 coexpressed gene groups with distinct biological functions and cis-regulatory features. Within these networks, additional chromatin immunoprecipitation and transcriptomic experiments revealed that SOG1 is the major activator, directly targeting the most strongly up-regulated genes, including TFs, repair factors, and early cell cycle regulators, while three MYB3R TFs are the major repressors, specifically targeting the most strongly down-regulated genes, which mainly correspond to G2/M cell cycle-regulated genes. Together these models reveal the temporal dynamics of the transcriptional events triggered by γ-irradiation and connects these events to TFs and biological processes over a time scale commensurate with key processes coordinated in response to DNA damage, greatly expanding our understanding of the DDR.

DNA damage response | DREM | SOG1 | transcriptional networks

The genomes of all organisms incur various types of DNA damage due to both endogenous processes and exposure to exogenous stresses or toxic compounds (1, 2). Of this damage, DNA double-strand breaks (DSBs) are particularly hazardous, as no intact strand remains to guide the DNA repair, potentially leading to chromosomal deletions and translocations (3, 4). To cope with such damage, mechanisms are in place to sense DNA lesions and initiate a DNA damage response (1, 5). This response involves the transcriptional and posttranscriptional regulation of diverse cellular pathways, ultimately leading to DNA repair, via the expression and/or targeting of repair factors to sites of damage, to cell cycle arrest, which provides additional time for DNA repair before replication, or to cell death, when the damage is too severe (5, 6). Given the importance of maintaining genome stability for proper cellular function and the faithful inheritance of genetic information (1–3, 5), it is critical to understand how the DNA damage response is initiated, coordinated, and executed.

Studies in yeast, plants, and mammals have revealed many highly conserved aspects of the DNA damage response (6–8). In the case of DSBs, conserved sensors, namely the MRN and Ku70/80 complexes, recognize the damaged DNA and transducers, including the ATAXIA-TELANGECTASIA MUTATED (ATM) and ATAXIA-TELANGECTASIA MUTATED AND RAD50-RELATED (ATR) kinases, initiate signaling cascades through the postranslational modification of target proteins (1, 8–10). These cascades modulate the activities of both shared and organism-specific effector proteins, culminating in the regulation of DSB repair, via homologous recombination and various nonhomologous end joining pathways, as well as the regulation of gene expression, cell cycle arrest, cell death, and endoreduplication (1, 6, 8, 11). To gain insight into the pathways and molecular interactions orchestrating these events, efforts in many organisms have focused on identifying and characterizing the key players, signaling cascades, and transcriptional programs that stem from the recognition of DNA damage.

In plants, the SUPPRESSOR OF GAMMA RESPONSE 1 (SOG1) transcription factor (TF) was identified from a DNA damage-suppressor screen (12) and was shown to be a major regulator of the DNA damage response (13). In the absence of SOG1, Arabidopsis plants exposed to DNA damaging agents display defects in gene regulation (13), cell cycle arrest (12), programmed cell death (14), endoreduplication (15), DNA repair, and genome stability (12, 13). These findings, along with those showing that SOG1 is regulated in an ATM-dependent manner via phosphorylation of conserved serine-glutamine motifs (16, 17), have led to SOG1 being functionally equated with p53 (8, 18), a mammalian tumor suppressor that coordinates the DNA damage response and is also phosphorylated in an ATM/ATR-dependent manner (19, 20). Despite the central role of SOG1 in the DNA damage response, and the numerous studies showing SOG1 is critical for coping with DNA damage (12–15, 21–26), global expression defects in sog1 mutants have only been assessed at single time points following γ-irradiation (γ-IR) (2 h) (13) or zeocin (1.5 h) (27) and, until recently (27), only a few SOG1 targets had been identified (22, 25, 26).

Significance

DNA damage triggers a highly conserved response that coordinates processes necessary to maintain genome integrity, including cell cycle arrest, DNA repair, and cell death. Despite the identification of primary transcription factors (TFs) that control these processes, knowledge regarding the downstream genes and regulatory networks controlled by these TFs remains poorly understood. Using Arabidopsis, we generated the first model of the DNA damage response transcriptional network, revealing 11 coexpressed gene groups with distinct biological functions and cis-regulatory features. Our characterization of this model demonstrates that SOG1 and three MYB3R TFs are, respectively, the major activator and repressors within this network, coordinating the rapid induction of DNA repair genes and TF cascades as well as the subsequent repression of cell cycle genes.

Author contributions: C.B., N.V., and J.A.L. designed research; C.B. and N.V. performed research; C.B. and J.A.L. analyzed data; and C.B. and J.A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The source data files and sequencing data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE112773).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810582115/-/DCSupplemental.

Published online December 12, 2018.
28). Furthermore, although the perception of DNA damage caused by exposure to γ-IR triggers events that occur on a time scale of minutes [e.g., the ATM/ATR-dependent phosphorylation of H2AX at DSBs (29–31)] to hours [e.g., cell cycle regulation (12, 29)], our understanding of the transcriptional changes coordinating these events is largely restricted to profiling experiments conducted at discrete time points (13, 32–39). Extending on these transcriptional snapshots, two previous studies profiled gene expression across several time points, but they utilized early array technology (40) or only included controls at a subset of time points (41). Thus, the expression dynamics of the DNA damage response, the full extent of SOG1’s role in gene regulation, and the transcriptional networks linking SOG1 to specific damage-associated processes remain to be determined.

To reveal the temporal features of the transcriptional response to DNA damage, and to further investigate the roles of SOG1 in executing this response, we performed transcriptomic analyses using γ-IR–treated wild-type and sog1 seedlings over a 24-h time course. These data, along with literature-curated gene–TF interactions, were then used to generate transcriptional network models of the Arabidopsis DNA damage response via DREM, the Dynamic Regulator Events Miner (42, 43). In total, ~2,400 differentially expressed (DE) genes were identified, greatly expanding upon the previously identified DNA damage-responsive genes. In the wild-type γ-IR treatment model, these genes were organized into 11 coexpressed groups with distinct expression patterns, promoter motifs, and gene ontology (GO) enrichments. Using this DREM model as a guide, additional analyses revealed both SOG1-dependent and -independent aspects of the DNA damage response and demonstrated that in addition to controlling the induction of many γ-IR responsive genes, SOG1 is also required for the repression of hundreds of genes. Furthermore, despite this dual effect in gene regulation, we found that SOG1 acts exclusively as a transcriptional activator, directly targeting ~300 genes, including many DNA repair and cell cycle factors, as well as a large subset of TFs, placing it at the top of a complex gene regulatory network. Finally, gene expression analysis of the myb3r1,3,5 mutant revealed that these TFs repress a large subset of G2/M-specific genes in response to DNA damage. Taken together, our findings not only shed light on the DNA damage response, but also provide a framework to begin connecting specific expression subnetworks to the diverse biological processes coordinated during this response.

Results and Discussion

Temporal Characterization of the DNA Damage Response Reveals Coexpressed Gene Sets with Distinct Biological Functions and Regulatory Features

To identify the transcriptional events underpinning the DNA damage response in Arabidopsis, mRNA sequencing (mRNA-seq) experiments were conducted at six time points from 20 min to 24 h after either mock or γ-IR treatments in wild-type plants (SI Appendix, Fig. S1A and Dataset S1). Furthermore, as the SOG1 TF is known to regulate many genes induced by DNA damage (13), a γ-IR time course experiment was also conducted in the sog1 mutant (SI Appendix, Fig. S1A and Dataset S1). Consistent with having selected a suitable time scale to capture the dynamics of the DNA damage response, identification of DE genes during the wild-type γ-IR time course (Dataset S2 A and C) [fold-change (FC) ≥ 2, false-discovery rate (FDR) ≤ 0.01] revealed that of the ~2,400 DE genes identified, 320 were already observed after just 20 min, while only 141 remained by 24 h (SI Appendix, Fig. S1B).

To determine whether the gene-expression changes observed in our γ-IR time course are consistent with previous assessments of the DNA damage response, the DE genes identified in at least one time point from the wild-type time course were compared with four previously published DE gene sets identified using similar γ-IR treatments and seedling stages, but at single time points (13, 32, 33, 38) (SI Appendix, Fig. S1 C and D). These analyses revealed that our γ-IR time course captures the vast majority of previously identified up-regulated genes, 65.2–91.4%, with the three datasets based on the most similar conditions and stages (13, 32, 33) showing the greatest overlaps (SI Appendix, Fig. S1C). For down-regulated genes, the overlap was lower but many common genes were still identified (SI Appendix, Fig. S1D). Based on these overlaps, we conclude that our γ-IR time course represents a stereotypical DNA damage response. However, in addition to the previously identified DE genes, over half of the up- and down-regulated genes identified in the γ-IR time course (655 and 519, respectively) were unique to this study (SI Appendix, Fig. S1 C and D).

Approximately 60% of these uniquely identified genes were regulated at commonly assessed time points (i.e., 1 h 30 min or 3 h) (SI Appendix, Fig. S1E), which likely reflects the increased sensitivity imparted by the use of mRNA sequencing rather than microarrays. The remaining 40% were specific to the additional time points covered in the γ-IR time course (SI Appendix, Fig. S1E). Collectively, these analyses identified an expanded set of DNA damage-responsive genes over a time span that corresponds to major events occurring in response to DNA damage, thereby providing a more comprehensive view of the DNA damage response.

To gain insight into how the thousands of γ-IR–regulated genes coordinate the DNA damage response, we used the DREM software (42, 43) to generate a temporal model of the underlying gene-regulatory network. This software groups genes with similar expression patterns together and assigns putative TFs to specific gene sets based on previously identified gene–TF interactions, enabling predictions to be made regarding specific features of the network. To facilitate subsequent comparisons with data generated in the sog1 mutant, the wild-type DREM model was constructed based on the log2 fold-change in expression between mock- and γ-IR-treated samples. Gene sets associated with the most similar conditions and stages (13, 32, 33) showing the broad peaks of repressing between 3 and 6 h. The remaining paths correspond to early up-regulated genes (W5 at 1 h 30 min and W6 at 20 min), early down-regulated genes (W8 at 1 h 30 min), and late mildly up-regulated genes (W7 at 3–12 h). Notably, ~half of the genes in the early and late responsive 11 groups of coexpressed genes with distinct biological functions and regulatory features (Fig. 1). At the level of gene expression, investigation of the DREM network confirmed that the 11 coexpressed groups, designated as paths W1–W11 (Fig. L4 and Dataset S3A), display highly correlated expression profiles across the entire γ-IR time course (SI Appendix, Fig. S24) and capture a wide range of expression dynamics, as exemplified by the representative gene-expression patterns shown in Fig. 1B. The two most prominent peaks revealed by the DREM analysis are a subset of paths (W1–W4) identified from the mock-treated γ-IR time course (Fig. S2A) showing broad peaks of repression between 3 and 6 h. The remaining paths correspond to early up-regulated genes (W5 at 1 h 30 min and W6 at 20 min), early down-regulated genes (W8 at 1 h 30 min), and late mildly up-regulated genes (W7 at 3–12 h). Notably, ~half of the genes in the early and late responsive paths (W5–W8), as well as many of the genes in the other paths, were uniquely identified in our γ-IR time course (SI Appendix, Fig. S2B).


to the remaining 40% were specific to the additional time points covered in the γ-IR time course (SI Appendix, Fig. S1E). Collectively, these analyses identified an expanded set of DNA damage-responsive genes over a time span that corresponds to major events occurring in response to DNA damage, thereby providing a more comprehensive view of the DNA damage response.

To gain insight into how the thousands of γ-IR–regulated genes coordinate the DNA damage response, we used the DREM software (42, 43) to generate a temporal model of the underlying gene-regulatory network. This software groups genes with similar expression patterns together and assigns putative TFs to specific gene sets based on previously identified gene–TF interactions, enabling predictions to be made regarding specific features of the network. To facilitate subsequent comparisons with data generated in the sog1 mutant, the wild-type DREM model was constructed based on the log2 fold-change in expression between mock- and γ-IR-treated samples. Gene sets associated with the most similar conditions and stages (13, 32, 33) showing the broad peaks of repressing between 3 and 6 h. The remaining paths correspond to early up-regulated genes (W5 at 1 h 30 min and W6 at 20 min), early down-regulated genes (W8 at 1 h 30 min), and late mildly up-regulated genes (W7 at 3–12 h). Notably, ~half of the genes in the early and late responsive paths (W5–W8), as well as many of the genes in the other paths, were uniquely identified in our γ-IR time course (SI Appendix, Fig. S2B).

Thus, while many genes show peaks of induction or repression at later, providing insights into the DNA damage response. To shed light on the biological functions of these gene sets, GO analyses were performed, revealing largely distinct enrichment terms for the DREM paths that capture the major processes previously connected with the DNA damage response [Fig. 1C. SI Appendix, Fig. S3, and Source Data 2 (44)]. These include DNA repair and DNA metabolism terms for the genes up-regulated in paths W1, W2, and—to a lesser extent—W3, cell cycle and associated terms for genes down-regulated in paths W9–W11, cell death terms for genes in paths W4 and W6, and respiratory burst along with other reactive oxygen species–associated terms for paths W6 and W7 (Fig. 1 C and D).
Fig. 1. DNA damage response DREM analysis reveals coexpressed genes with distinct biological functions and regulatory features. (A) DREM model (see Source Data 1 (44)) showing 11 groups of coexpressed genes, termed wild-type paths W1–W11. Here, and in all other DREM models, the y axis indicates the log₂ FC in expression in response to γ-IR, the x axis indicates the time in minutes (') and/or hours (h), and the number (N) of genes per path is indicated. All genes are listed in Dataset S3A. Comparisons with previously published DE gene sets are presented in SI Appendix, Fig. S1, expression patterns of the individual genes in each DREM path are shown in SI Appendix, Fig. S2A, and the TF families (i.e., NAC, TCP, HB, WRKY, and MYB) assigned to the DREM paths are indicated, with the lists of all of the TFs assigned to each path shown in SI Appendix, Fig. S4. (B) Screenshots showing the expression levels of representative genes from each DREM path. The gene indicated above is shown in blue and the neighboring genes are shown in gray. The difference between the mock and γ-IR–treated samples [+(+γ-IR av) − (−γ-IR av)] is shown for each time point on a scale of ±125. (C) Heatmap showing the significance of all of the GO terms with a log₁₀ P < −1.7 in at least one path, across all of the DREM paths. Gray indicates a P > 0.5. See SI Appendix, Fig. S3 and Source Data 2 (44) for all enriched GO terms remaining after the REVIGO similarity filter (100). (D) Table of select genes associated with the processes indicated above each column based on GO category enrichments and/or manual curation. The gene names and/or the number of genes per category (in parentheses) are colored based on the path in which they reside. (E) Select motifs enriched in the indicated DREM paths. The full set of enriched motifs identified are presented in SI Appendix, Fig. S5 [Source Data 3 (44)]. Below each motif, E-values calculated against all Arabidopsis promoter sequences, and similar TF families, identified via comparisons with the motifs in the DAP-seq database (48) using Tomtom (98), are indicated. Three similar, but independently identified motifs that correspond to the previously described mitosis-specific activator (MSA) element (AACGG) (55) are marked with an asterisk.
and and and and and and and and and and and and and and and and and γ

DREM model

www.pnas.org/cgi/doi/10.1073/pnas.1810582115

SI Appendix

2 and SI Appendix

0.01), most of the genes included in the sog1 mutant (771; FC 

SI Appendix

γ

C

Bourbousse et al.

SOG1 mutant (771; FC 

SI Appendix

γ

DREM model (Fig. S6 C). Thus, nearly two-thirds (1,451 of 2,395) of the transcriptional DNA damage response is completely dependent on SOG1. As detailed below, assessment of the features remaining in the sog1 DREM network (Fig. 2 A and B and SI Appendix, Fig. S6D) and visualization of the sog1 expression pattern specified for the context of the wild-type DREM model (Fig. 2C) reveal both SOG1-dependent and SOG1-independent aspects of the DNA damage response.

Analysis of the path S1 genes from the sog1 DREM model revealed substantial overlaps with the genes present in paths W5 and W6 of the wild-type DREM model (Fig. 2B). Furthermore, the expression levels of many of the W5 and W6 genes were comparable between the wild-type and sog1 datasets at the 20-min time point (Fig. 2C), demonstrating that a large portion of the initial burst of gene expression observed in the wild-type DNA damage response is SOG1-independent. However, for the path W5 genes, the sustained induction observed in the wild-type γ-IR time course was dependent on SOG1 (Fig. 2C). Importantly, the up-regulation of path W5 and W6 genes 20-min post-irradiation and the SOG1-independent nature of this early response were independently verified by a second set of wild-type and sog1 γ-IR time courses (SI Appendix, Fig. S7). Because many of the genes in path S1, like those in paths W5 and W6, are associated with diverse stress-response terms (Fig. 1C and SI Appendix, Figs. S6D and S8A), and because these expression profiles show similarities to profiles observed after various biotic and abiotic stresses (SI Appendix, Fig. S8B), rather than being true SOG1-dependent, we posit that the DNA damage response likely represents a more general stress response that may not be directly coupled to the detection of damaged DNA.

Analysis of the path S5 genes from the sog1 DREM model revealed that a specific subset of genes are repressed in a partially SOG1-independent manner (Fig. 2D). These genes correspond almost exclusively (98%) to paths W10 and W11 of the wild-type DREM model (Fig. 2B) and, consistent with this high degree of overlap, they are enriched for cell cycle–related GO terms associated with the promoters of G2/M-specific genes (55–57), identified in the promoters of the path W10 and W11 genes (Fig. 1E and SI Appendix, Fig. S5). Based on these TF predictions, additional genetic and genomic analyses were conducted to assign SOG1 and the Rep-MYB3R TFs to the regulation of specific genes and reveal mechanistic insights into how these TFs coordinate the DNA damage response.

SGO1 Controls Nearly All Transcriptional Aspects of the DNA Damage Response. Previously, Sog1 was shown to control the expression of ~300 up-regulated, and a few down-regulated genes, 1.5 h after exposure to γ-IR, demonstrating it plays a major role in gene regulation during the DNA damage response (13). Here, many additional DNA damage-responsive genes were identified as part of the wild-type γ-IR time course. To determine the role of Sog1 in regulating this network, a new DREM model was computed based on the same 2,395 genes included in the wild-type γ-IR time course and 218 sog1-specific DE genes), but using the RNA-seq data from the sog1 γ-IR time course. In this model, only five paths were identified, and they displayed a striking reduction in their expression changes compared with the wild-type DREM paths (Fig. 2B, SI Appendix, Fig. S6B, and Dataset S3B). Consistent with the identification of fewer DE genes in the sog1 mutant (771; FC ≥ 2 and FDR ≤ 0.01), most of the genes included in the sog1 DREM model based solely on their regulation in the wild-type γ-IR time course (1,451 of 1,624) reside in paths S3 and S4 and showed little to no concerted changes in expression across the entire sog1 γ-IR time course (SI Appendix, Fig. S6 B and C). The remaining 173 genes, which did not pass the thresholds for differential expression in the sog1 γ-IR time course, displayed subtle, but more concerted, expression changes and were assigned to paths S1, S2, or S5 in the sog1 DREM model (SI Appendix, Fig. S6C). Thus, nearly two-thirds (1,451 of 2,395) of the transcriptional DNA damage response is completely dependent on SOG1. As detailed below, assessment of the features remaining in the sog1 DREM network (Fig. 2 A and B and SI Appendix, Fig. S6D) and visualization of the sog1 expression pattern specified for the context of the wild-type DREM model (Fig. 2C) reveal both SOG1-dependent and SOG1-independent aspects of the DNA damage response.

Analysis of the path S1 genes from the sog1 DREM model revealed substantial overlaps with the genes present in paths W5 and W6 of the wild-type DREM model (Fig. 2B). Furthermore, the expression levels of many of the W5 and W6 genes were comparable between the wild-type and sog1 datasets at the 20-min time point (Fig. 2C), demonstrating that a large portion of the initial burst of gene expression observed in the wild-type DNA damage response is SOG1-independent. However, for the path W5 genes, the sustained induction observed in the wild-type γ-IR time course was dependent on SOG1 (Fig. 2C). Importantly, the up-regulation of path W5 and W6 genes 20-min post-irradiation and the SOG1-independent nature of this early response were independently verified by a second set of wild-type and sog1 γ-IR time courses (SI Appendix, Fig. S7). Because many of the genes in path S1, like those in paths W5 and W6, are associated with diverse stress-response terms (Fig. 1C and SI Appendix, Figs. S6D and S8A), and because these expression profiles show similarities to profiles observed after various biotic and abiotic stresses (SI Appendix, Fig. S8B), rather than being true SOG1-dependent, we posit that the DNA damage response likely represents a more general stress response that may not be directly coupled to the detection of damaged DNA.

Analysis of the path S5 genes from the sog1 DREM model revealed that a specific subset of genes are repressed in a partially SOG1-independent manner (Fig. 2D). These genes correspond almost exclusively (98%) to paths W10 and W11 of the wild-type DREM model (Fig. 2B) and, consistent with this high degree of overlap, they are enriched for cell cycle–related GO terms associated with the promoters of G2/M-specific genes (55–57), identified in the promoters of the path W10 and W11 genes (Fig. 1E and SI Appendix, Fig. S5). Based on these TF predictions, additional genetic and genomic analyses were conducted to assign SOG1 and the Rep-MYB3R TFs to the regulation of specific genes and reveal mechanistic insights into how these TFs coordinate the DNA damage response.

SGO1 Controls Nearly All Transcriptional Aspects of the DNA Damage Response. Previously, Sog1 was shown to control the expression of ~300 up-regulated, and a few down-regulated genes, 1.5 h after exposure to γ-IR, demonstrating it plays a major role in gene regulation during the DNA damage response (13). Here, many additional DNA damage-responsive genes were identified as part of the wild-type γ-IR time course. To determine the role of Sog1 in regulating this network, a new DREM model was computed based on the same 2,395 genes included in the wild-type γ-IR time course and 218 sog1-specific DE genes), but using the RNA-seq data from the sog1 γ-IR time course. In this model, only five paths were identified, and they displayed a striking reduction in their expression changes compared with the wild-type DREM paths (Fig. 2B, SI Appendix, Fig. S6B, and Dataset S3B). Consistent with the identifi
SOG1 Is a Transcriptional Activator that Directly Targets Nearly Half of the Genes Strongly Induced by DNA Damage. As described above, SOG1 is required for the induction or repression of thousands of genes in response to γ-IR. Until recently (27), only a few SOG1 target genes had been identified, namely SIAMESE-RELATED 5 (SMR5) and SIAMESE-RELATED 7 (SMR7) (26), FLAVIN-DEPENDENT MONOOXYGENASE 1 (FMO1) (22), CYCLIN B1.1 (CYCB1.1) (28), and BREAST CANCER SUSCEPTIBILITY 1 (BRCA1) (25). To determine which of the other genes regulated by SOG1 are direct targets, chromatin immunoprecipitation and sequencing (ChiP-seq) experiments were conducted (Dataset S4) using transgenic lines in which the sog1 mutation was complemented by a SOG1-3xFLAG construct driven by the endogenous SOG1 promoter (SI Appendix, Fig. S10 A and B). As SOG1 binding is expected to precede the transcriptional regulation of its targets, the ChiP-seq experiments were performed at two early time points after irradiation, 20 min and 1 h. Analysis of the resulting ChiP-seq profiles identified 307 SOG1 peaks (Dataset S4B) that are mainly located in promoters and transcribed regions (SI Appendix, Fig. S10C) and were assigned to 310 immediately adjacent gene targets (Dataset S4 C and D). These peaks were enriched relative to both input and wild-type ChIP samples (SI Appendix, Fig. S10D) and, attesting the reproducibility and quality of the ChiP-seq experiments, similar enrichment patterns were observed for both the 20-min and 1-h datasets (Fig. 3A). Furthermore, four of the five initially identified SOG1 targets (SI Appendix, Fig. S10E), and partially dependent on SOG1. This includes the 1,233 genes present in paths W1–W5 and W7 that require SOG1 for their induction between 20 min and 12 h, the 712 genes present in paths W8 and W9 that require SOG1 for their repression, and the 178 genes present in paths W10 and W11 that show a partial dependence on SOG1. These findings greatly expand the set of genes known to be induced in a SOG1-dependent manner and demonstrate that SOG1 is a master regulator not only for genes induced by DNA damage, but also for those that are repressed.

Fig. 2. SOG1 controls nearly all aspects of the transcriptional response to γ-IR. (A) sog1 DREM model [see Source Data 1 (44)] showing five sets of coexpressed genes, termed sog1 paths S1–S5. The expression profiles, enriched GO terms, and motifs are presented in SI Appendix, Figs. S6 and S9, respectively. All genes are listed in Dataset S3B. For comparison, the wild-type (wt) DREM model is shown as an inlay. (B) Scaled Venn diagrams showing the overlap of genes in DREM paths with similar trends in the wild-type and sog1 models. (C) Heatmaps showing the log FC in expression (γ-IR vs. mock) of the genes present in paths W1–W11 (Fig. 1A) using either the wild-type or the sog1 expression data. For each path, the heatmaps were ranked based on the wild-type expression level. See also SI Appendix, Fig. S7.

Fig. 3. SOG1 is a transcriptional activator that directly regulates nearly half of the genes strongly induced by γ-IR. (A) Heatmap showing SOG1 enrichment [log2(SOG1/wt ChIP)] at the 307 peaks (±3 kb) identified from both ChiP assays (20 min and 1 h) (SI Appendix, Fig. S10 and Dataset S4B). (B) Heatmaps showing the expression of SOG1 target genes (Datasets S4 C and D) during the wild-type (wt) and sog1 γ-IR time courses ranked based on the wild-type expression values at 1 h 30 min. (C) Expression profiles of the SOG1 target genes colored by gene paths. The fraction of genes in each path is indicated, with the percentage in parentheses. (D) Motif identified under the SOG1 ChiP peaks. The E-value, as reported by MEME (97, 102), and the fraction of peaks with the motif are indicated below (see also Dataset S4E).

6 h (Figs. 1A and 2C). However, essentially none of these genes are induced at the proper time (between 1 h 30 min and 6 h) in the sog1 mutant (Fig. 2C). As with the other features revealed by the sog1 DREM model, these expression profiles were independently verified (SI Appendix, Fig. S7). While many of the late-induced genes in the sog1 mutant correspond to general stress-response genes, several well-known DNA repair genes, including GMI1, BRCA1, PARP2, and PARG2 are also present [Dataset S3B and Source Data 2 (44)]. Thus, we posit that this latent response may be triggered by defects in DNA repair and the persistence of unpaired DNA lesions that may be accentuated by the loss of SOG1 function.

In sum, our analysis of the DNA damage response in sog1 mutants demonstrate that, with the exception of an initial burst in the expression of general stress-response genes, essentially all other aspects of the wild-type DNA damage response are either fully or
104 of the 146 recently identified SOG1 ChIP-seq targets (27), were among the 310 SOG1 targets identified after γ-IR treatment (SI Appendix, Fig. S10F). Thus, the SOG1 ChIP-seq results presented here are consistent with, but greatly expand upon, the previously identified SOG1 targets.

To determine whether genes targeted by SOG1 are induced, repressed, or both in response to DNA damage, their expression patterns after γ-IR treatment were determined from the wild-type and sog1 γ-IR time-course datasets (Fig. 3B). Over 80% of these genes were induced by γ-IR in a SOG1-dependent manner, while none were significantly repressed (Fig. 3B), demonstrating that SOG1 acts, either by itself or in association with other factors, as a transcriptional activator. In agreement with this finding, 206 of the 310 target genes were identified as part of the wild-type DREM model. Not only do all these genes map to up-regulated gene paths (Fig. 3C and SI Appendix, Fig. S10G), they correspond to 43% of all the strongly up-regulated genes (e.g., W1–W3) and 85% of the most highly induced genes (e.g., W1) (Fig. 3C). Strikingly, the same motif identified from the path W1–W3 genes of the DREM network (CTT[N][AAG]) (Fig. 1E) was found directly under nearly all (289 of 307) of the SOG1 peaks (Fig. 3C), underlining the importance of this SOG1 binding motif observed in vitro and in vivo assays.

Together, these analyses reveal several hundred direct targets of SOG1 and demonstrate that although SOG1 is required for both the induction and repression of DNA damage-responsive genes, it acts strictly as a transcriptional activator, directly controlling ~8% of the DNA damage transcriptional response.

Further investigation into the direct targets of SOG1 demonstrated that it regulates genes involved in many DNA damage-associated processes, including transcriptional and posttranscriptional regulation, DNA repair, and is also consistent with the findings of Ogita et al. (27), as this motif is similar to motifs observed for other NAC TFs (48, 51, 52) and is consistent with the findings of Ogita et al. (27), as this motif was identified in the identified regions of approximately half of the previously identified SOG1 target genes and was shown to be essential and sufficient for SOG1 binding based on in vitro and in vivo assays.

In agreement with SOG1 being required for cell cycle arrest in response to DNA damage (12), four cyclin-dependent kinase inhibitor (CKI) genes, belonging to two classes, [KIP-RELATED PROTEINS (KRP) and SIAMESE/SIAMESE-RELATED (SMR)] (80, 81), were identified as SOG1 targets here (Fig. 4) and in Ogita et al. (27). Current models posit that the up-regulation of CKIs by SOG1 plays an important role in stabilizing the Rep-MYB3R family of TFs, which are required for the repression of G2/M-regulated genes in response to DNA damage (53). As such, these SOG1 ChIP-seq findings expand the pool of CKIs that may be acting upstream of the Rep-MYB3Rs to include not just the initially identified SMR5 and SMR7 targets (26), but also SMR4 and KRP6 (SI Appendix, Fig. S12). Interestingly, SOG1 was also found to directly target WEE1, which encodes a kinase that is required for proper cell cycle arrest in response to replicative stress (82), but not in response to DSBs (15), suggesting that SOG1 activates cell cycle arrest programs associated with different types of damage. Finally, in addition to cell cycle arrest, SOG1 is also required for the promotion of endoreduplication in response to DNA damage (15). Although SOG1 does not directly regulate genes with known roles in the endoreplication cycle, it does target and induce the expression of WRKY25. Because this TF was shown by DAP-seq to bind the promoter of DP-E2F-LIKE 1 (DEL1) (48), a gene important for the repression of endoreduplication (83) (SI Appendix, Fig. S12), it can be hypothesized that WRKY25, or other related WRKY TFs, represses DEL1 expression in response to DNA damage to promote endoreduplication. Based on these findings it is clear that,

Fig. 4. Functional categorization of SOG1 target genes. Of the 206 SOG1 target genes in the DREM model, 141 were assigned to functional categories and colored based on the log2 FC for γ-IR at the 3-h time point from the wild-type DREM model. Genes underlined in blue have a human and/or mouse ortholog shown to be targeted and up-regulated by p53. See also Materials and Methods, SI Appendix, Figs. S11 and S12, and Source Data 4 (44).

 regulative TCP20, which encodes a TF that binds the CYCB1;1 promoter (79) and may help reinforce its up-regulation. Further extrapolation of the SOG1 TF network based on previously annotated gene–TF interactions from the AGRIS (45–47) and DAP-seq databases (48) allows the inference of a putative second tier within the network (SI Appendix, Fig. S11). Although this second tier remains to be experimentally validated, it provides a framework for the generation and examination of additional hypotheses regarding the transcriptional response to DNA damage.
although much of the molecular details remain to be elucidated, SOG1 plays critical, but indirect roles in controlling the suppression of cell cycle genes, which is consistent with the DREM model where the induction of DNA repair factors and other direct targets of SOG1 precede the suppression of cell cycle genes.

Finally, in addition to previously drawn parallels between SOG1 and the mammalian p53 protein, which focused on the activation of SOG1 by ATM and the common DNA damage-associated processes dependent on these two TFs (cell cycle arrest, cell death, overall genome stability, and the induction of damage-response genes), the identification and analysis of SOG1 target genes has revealed additional parallels. First, both proteins act as transcriptional activators (84, 85). Second, they target genes related to similar biological processes (19). And third, many of the SOG1 target genes have human and/or mouse orthologs identified as p53 targets (Fig. 4), including the RNR subunit, TSO2, for dNTP balance maintenance (86); the DNA polymerase kappa, POLK, for translesion DNA synthesis (87); the histone variant H3.1, which is deposited in a DNA-synthesis–dependent manner and is incorporated at damaged chromatin (88); and KRP6, which contains a cyclin-dependent kinase inhibitor domain similar to that of p21, a mammalian gene that mediates the p53-dependent down-regulation of cell cycle genes (89). However, SOG1 is unique in its selective targeting of numerous genes required for repair by HR (Fig. (27)). This is despite the fact that there is a sequence conservation between p53 and SOG1, they share a subset of conserved target genes, suggesting that they have been coopted to mediate both shared and unique aspects of the DNA damage response in plants versus mammals.

The Rep-MYB3R Family Is Required to Suppress Cell Cycle Genes after DNA Damage. Although the direct targets of SOG1 are activated in response to DNA damage, hundreds of repressed genes also depend on SOG1. Thus, events set into motion by the expression of SOG1 targets must play key roles in repressing genes in response to DNA damage. Of particular interest are those that regulate cell cycle genes, which are strongly repressed in the wild-type DREM model (e.g., W10 and W11). Consistent with the TF predictions from our DREM model (Fig. L4 and SI Appendix, Fig. S4), recent studies have revealed connections between the Rep-MYB3R family, cell cycle regulation, and DNA damage. First, all three family members—MYB3R1, MYB3R3, and MYB3R5—were found to act redundantly to suppress the expression of 34 genes associated with the G2/M phase of the cell cycle (67, 69) of which fall into paths W9, W10, and W11 (SI Appendix, Fig. S13A). Second, the myb3r3 and myb3r5 mutants display enhanced tolerance to DNA damage agents, including γ-IR, and show defects in cell cycle regulation and cell death (53). Finally, although the Rep-MYB3Rs are not transcriptionally regulated in response to DNA damage (SI Appendix, Fig. S13B) (53), they were placed in a SOG1-dependent pathway based on epistasis experiments (53). Together, these findings demonstrate that the Rep-MYB3Rs are essential for inhibiting cell division during the DNA damage response in connection with SOG1, yet only a few genes repressed in a MYB3R1/3/5-dependent manner after DNA damage have been identified (53).

To identify genes regulated by the Rep-MYB3R family in response to DNA damage, mRNA-seq experiments were conducted in wild-type and myb3r1,3,5 triple mutants 3 h after either mock or γ-IR treatments (SI Appendix, Fig. S13B and Dataset S1), a time when hundreds of genes are strongly down-regulated in the wild-type DREM model (Fig. L4). In agreement with previous studies showing minimal expression changes between wild-type and myb3r1,3,5 triple mutants in early seedling stages (90), comparison of the 6-old mock-treated seedlings (wild-type vs. myb3r1,3,5) revealed only 24 up-regulated genes, including just two G2/M phase genes and a single down-regulated gene (Dataset S5A). However, after γ-IR treatment, the DNA damage response was clearly altered in the myb3r1,3,5 triple mutant compared with the wild-type control. On a global level, the γ-IR response observed in the wild-type dataset was similar to the 3-h time point from the wild-type DREM model, but in the myb3r1,3,5 triple dataset, the genes in two of the three cell cycle-enriched paths (W10 and W11) were less repressed overall (Fig. 5A). At the level of individual genes, 80 loci significantly less repressed in the myb3r1,3,5 mutants after DNA damage (Dataset S5B and C) (FC ≥ 2 and FDR ≤ 0.05) were determined by considering both the experimental conditions (γ-IR vs. mock treatments) and the genotypes (wild-type vs. myb3r1,3,5). Nearly all of these genes (78 of 80) are present in the wild-type DREM model, constituting 72.3% of the path W11 genes (47 of 65), 24.8% of the path W10 genes (28 of 113), and 0.5% of the path W9 genes (3 of 751) (Fig. 5B and SI Appendix, Fig. S13C).

Functionally, 71 of these 80 genes are associated with the G2/M phase of the cell cycle (54, 57) (Fig. 5C and Dataset SSB). Approximately one-third of these genes (28 of 70) were previously shown to be repressed in a Rep-MYB3R–dependent manner either under normal growth conditions (90) or after exposure to DNA damage (53) (Dataset SSB). However, the remaining two-thirds (42 of 70) represent newly identified Rep-MYB3R–regulated genes (Dataset SSB). Finally, these 80 genes are likely direct targets of the Rep-MYB family, as they almost all (72 of 80) possess MSA motifs in their promoters and/or are associated with previously defined MYB3R3 peaks by ChIP-seq (q-value > 25 under nondamaged conditions) (90) or by ChIP-qPCR after DNA damage (Fig. 5D and SI Appendix, Fig. S13D). Furthermore, the association of MYB3R3 with these...
genes is quite specific, as the same DREM paths affected in the myb3r1,3,5 triple mutants displayed the highest levels of promoter-proximal MYB3R3 enrichment (SI Appendix, Fig. S13E). In contrast to the genes less repressed in the myb3r1,3,5 mutants after DNA damage, the genes significantly less induced showed poor overlap with genes in the wild-type DREM model (SI Appendix, Fig. S13C) and none overlapped with G2/M-expressed genes or MYB3R3 peaks (Dataset SS5), suggesting that these genes may not be directly regulated by the Rep-MYB3R3 family. Taken together, these findings more than double the number of cell cycle–regulated genes known to be regulated by the Rep-MYB3R3 family and demonstrate that these TFs control a large portion of the most strongly repressed genes in the wild-type DREM network in response to DNA damage.

Conclusions

Despite the identification of TFs that control most of the cellular outcomes of the DNA damage response, namely p53 in mammals and SOG1 in plants (8), knowledge regarding the gene expression networks controlled by these TFs, and the dynamics of the DNA damage response as a whole, has remained poorly understood. Here, DREM analysis of the transcriptional changes occurring in response to γ-IR resulted in the classification of the Arabidopsis DNA damage response into 11 groups of coexpressed genes with distinct expression profiles, biological functions, and cis-regulatory features. Attesting to the power of this approach, this network captured the major biological processes associated with the DNA damage response, including DNA metabolism, DNA repair, cell cycle regulation, and cell death, while at the same time providing new insight into early responses that occur in a largely SOG1-independent manner and late responses that occur specifically in the absence of SOG1. In addition, the DREM model revealed the order of key transcriptional events, including an initial burst in the expression of general stress genes, a rapid and sustained induction of DNA repair and DNA metabolism genes, and a slightly delayed suppression of cell cycle-associated genes. Finally, analysis of additional genetic and genomic experiments within the context of this model enabled the assignment of SOG1 and the Rep-MYB3R TFs to the most strongly up- and down-regulated gene clusters within the network, demonstrating that they act as the major activator and repressors of the DNA damage response, respectively (SI Appendix, Fig. S14).

The gene-expression profiles and TFs identified based on this DREM analysis represent a transcriptional roadmap of the DSB response in Arabidopsis that serves as a framework to continue exploring the expression of both known and unknown genes. These studies may prove especially useful for genes present in pathophysically enriched GO terms, as uncharacterized genes coexpressed with, for example, DNA repair factors or cell cycle regulators, may have similar functions. Similarly, this map will facilitate the identification and characterization of additional TFs and expression programs acting downstream of, or independent from, SOG1 to coordinate the diverse processes associated with the DNA damage response. In moving forward, it will also be important to continue integrating information into the DREM model for factors that, like the Rep-MYB3R3s and SOG1, are regulated at the posttranscriptional rather than the transcriptional level (16, 17, 53). Indeed, analysis of the Arabidopsis phospho-proteome in response to ionizing radiation identified hundreds of proteins phosphorylated upon DNA damage, including many TFs (91), but the functions of most of these posttranslational modifications remain unknown. Finally, it will be interesting to extend upon previous comparisons between the transcriptional responses elicited by various genotoxic stresses (41), via similar network approaches with matched experimental designs, to gain a higher-resolution view of how the profiles differ in response to specific types of DNA damage. Continued efforts toward addressing the aforementioned aspects of the DNA damage response will shed additional light on the inner workings of this process, which can ultimately be leveraged for crop improvement, because many environmental factors—including UV-B exposure, toxic levels of aluminum or boron in the soil, or endogenous processes like seed aging and germination—are associated with various types of DNA damage and show genetic requirements and transcriptional responses with links to the DNA damage response induced by DSBs (5, 41, 92).

Materials and Methods

Plant Materials. Arabidopsis lines include the Columbia (Col) wild-type strain, the sog1-1 EMS mutant provided by A. Britt, Department of Plant Biology, University of California, Davis, CA, which is in a mixed Landsberg erecta (Ler)/Col background containing a CYCB1;1 promoter::GUS fusion (13), and the myb3r1,3,5 triple mutant provided by M. Ito, Graduate School of Biological Sciences, Nagoya University, Chikusa, Nagoya, Japan (90).

mRNA-seq Experiments. Growth conditions and library preparation. For all mRNA-seq experiments, seeds were sterilized, sown on half MS media with 0.6% plant agar, stratified at 4 °C for 3 d, and transferred into a growth chamber [constant light (100 μmol·m⁻²·s⁻¹) and temperature (23 °C)]. For the wild-type, sog1-1, and myb3r1,3,5, datasets, 6-d-old seedlings were γ-irradiated at a dose of 100 Gy using a Co60 radioactive source. See SI Appendix, Materials and Methods for details regarding dose rates and experimental designs. Total RNA was extracted from pooled seedlings using the Quick-RNA MiniPrep kit (Zymo Research R1055) and mRNAs were purified from 1 to 2 μg of total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs E7490) before library preparation with the NEBNext Ultra or Ultral RNA Library Prep Kit (New England Biolabs E7530 or E7770). All sequencing was performed on an Illumina HiSeq 2500 in 150-bp single-end mode.

Data processing. For all mRNA-seq experiments, Illumina reads were mapped to the TAIR10 genome using STAR (93) (Dataset S1). Expression values for all genes and samples were retrieved using the analyzeRepeat.pl script from HOMER (94) with the -noad option. These values were used to generate the principle component analysis plots shown in SI Appendix, Fig. S1A with ggplot2 in R studio and to identify DE genes (FC ≥ 2 and an FDR ≤ 0.01) based on the experimental conditions (i.e., γ-IR vs. mock) (Dataset S2) or genotype (i.e., wild-type vs. mutant) (Dataset S5A) using the getDiffExpression.pl script from HOMER (94) with the -DEseq2 and -reps options. To identify DE genes based on both the experimental conditions and genotypes for the "wt vs myb3r1,3,5" experiment (Dataset S5B and C), the analysis was performed in R studio using DESeq2 with design = batch + Condition + Group + Condition:Group. To visualize the gene-expression data, University of California, Santa Cruz (UCSC) browser tracks were generated using the makeUCSCfile script from HOMER (94), normalizing to 10 million reads. Tracks showing the difference between the average mock- and γ-IR-treated samples (Fig. 18) were generated using deepTools (95) as detailed in SI Appendix, Materials and Methods. Heatmaps showing the expression of individual genes during the γ-IR time course were generated using the normalized logFC values generated by DESeq2 and plotted in R studio using pheatmap. Venn diagrams comparing DE gene sets were generated using VennMaster (96).

DREM Analysis. For data inputs and models, see Source Data 1 (44). Three types of files were generated as inputs for the DREM models. The first is a list of the 33,323 TAIR10 gene IDs, the second is a set of gene–TF interactions, and the third contains the expression values from the RNA-seq experiments. See SI Appendix, Materials and Methods for details regarding the content and generation of these files. As detailed in SI Appendix, Materials and Methods, the input files described above were used to generate the wild-type and sog1 DREM models by a two-step process using the DREM v2.0.3 (42, 43).

Motif Analysis. As detailed in SI Appendix, Materials and Methods, the MIME suite (97) was used to identify motifs enriched in the promoters of genes present in the individual DREM paths. All found motifs were then compared with the DAP-seq database of motifs (48) using Tomtom (98) to search for motifs overlapping by at least 5 bp with distance coefficient calculated by the Pearson method showing an E-value < 0.1. The matches with an E-value < 0.01 were kept.

GO Analysis. As detailed in SI Appendix, Materials and Methods, the enrichment for each DREM path were retrieved using GO-TermFinder (99) via the Princeton GO-TermFinder interface (https://go.princeton.edu/cgi-bin/GOTermFinder), simplified using the REVIGO (100) (revigo.irb.hr), and visualized as unclustered heatmaps in Treeview.
SOG1 ChIP-Seq Experiments.

ChIP: For ChIP, 1.3 g of 10-d-old seedlings from wild-type and SOG1-3xFLAG transgenic lines were harvested after 20 min and 1 h of γ-ir (75 Gy at a dose-rate of 3.75 Gy/min), ground in liquid nitrogen, and resuspended in 25 mL of nuclear isolation buffer. Proteins were cross-linked first with 1.5 mM ethyleneglycol bis(succinimidyI succinate) (Thermo Fisher Scientific 21565) for 20 min at room temperature, and then with formaldehyde at a final concentration of 1% for 10 min at room temperature. Cross-linking was stopped by adding 1.7 mL of 2 M glycine and incubating for 10 min at room temperature. The cross-linked samples were filtered through Miracloth and centrifuged at 4,500 × g for 20 min at 4 °C. Pellets were resuspended in 1-mL extraction buffer 2 and centrifuged at 12,500 × g for 10 min at 4 °C. Pellets were resuspended in 300 μL of nuclease lysis buffer. The chromatin was sonicated using a Bioruptor (Diagenode) for 20 cycles (30-s on/30-s off, on high intensity) at 4 °C. The debris was pelleted by centrifugation at 9,600 × g for 10 min at room temperature; 30 μL were kept as an input. The remaining sample was diluted 10 times using ChIP dilution buffer. Next, 25 μL of M2-FLAG magnetic bead slurry (Sigma M8823) was added and the sample was rotated for 1 h at 4 °C. The beads were washed sequentially in 1 mL of low-salt buffer, 1 mL of high-salt buffer, 1 mL UCi buffer, and 1 mL of TE. The chromatin elution was performed in 300 μL SDS elution buffer at 65 °C and the cross-linking was reversed by adding 12 μL of 5 M NaCl and 30 μL of 1 M DTT and incubating the sample overnight at 65 °C. Digestion with Proteinase K was performed at 45 °C for 1 h. The DNA was purified by phenol-chloroform-IAA extraction and precipitated with ethanol. Pellets of both inputs and IPs were resuspended in 30 μL of TE, pH 8.0. All buffers are described in Dataset S6.

Library preparation, sequencing, and analysis. Libraries were prepared from 25 μL of the IP samples or 1 μg of the input samples using the NEBNext Ultra II DNA Library Prep Set for Illumina (New England Biolabs E7645S). Sequencing was performed on an Illumina HiSeq 2500 in 50-bp single-end mode. Reads were mapped to the TAIR10 genome using bowtie2 (101) (Dataset S4A). UCSC browser tracks normalized to 10 million reads were generated using HOMER (94), and, as detailed in SI Appendix, Materials and Methods, heatmaps and metatranscripts showing SOG1 ChIP enrichments were generated using the deepTools suite (95). SOG1 peaks at 20 min or 1 h were called independently relative to two datasets (the wt/IP and input samples) using the HOMER findPeaks tool (94) and only peaks identified relative to both controls were kept. For details regarding peak calling and gene assignments, see SI Appendix, Materials and Methods. Heatmaps showing the expression of SOG1 target genes (Fig. 38) were generated using the log2 FC values generated by DESeq2 and plotted in R studio using heatmap and the overlaps between SOG1 target genes identified here or in ref. 27 were generated using VennMaster (96). Peak locations relative to genome features (promoters, introns, exons, and so forth) were determined using the HOMER annotatePeaks.pl script (94), where the promoter regions are defined as −1 kb and +100 bp relative to the transcription start site. Motifs under the SOG1 peaks were determined using the MEME tool (102) from the MEME suite (97) with the following options (-nostatus -maxsize 7500000 -nomots 10 -minmot 6 -maxw 18 -recomp -pdp -bfie).

MYBR3 ChIP-Seq Analysis. Using data from GSE60554 (90), the closest TAIR10 gene to each peak was determined using annotatepeaks.pl from HOMER (94). Mapping of the ChIP-seq reads and analysis of ChIP enrichment profiles using deepTools were as described for the SOG1 ChIP-seq. Venn diagrams were generated using VennMaster (96) based on known G2/M-expressed genes (54, 57) or previously identified MYBR3 ChIP peaks (90).

Cytoscape. Using Cytoscape 3.4.0, networks for either the 141 SOG1 target genes that were assigned to functional categories based on the GO analysis and their TAIR10 annotations [Source Data 4 (44)] or the DREM network targets of the 33 Ts downstream of SOG1 identified from the AGR15 (45–47) and DAP-seq (48), databases [SI Appendix, Fig. S1 and Source Data 4 (44)] were constructed and colored based on the log2 FC in expression 3 h after γ-ir. For Fig. 4, genes underlined in blue represent those that have a human or mouse ortholog, identified using the PANTHER (103) and Thalemine tools (104), that were shown to be targeted and up-regulated by p53 based on 13 genome-wide studies in humans (105) or a single study in primary mouse embryo fibroblasts study (106).

ACKNOWLEDGMENTS. We thank the J.A.L. laboratory and colleagues at the Salk Institute for helpful discussions; and Dr. C. Huang, Dr. L. Song, and Dr. Y. He for bioinformatics support. Work in the J.A.L. laboratory was supported by the Rita Allen and Heart Foundations (to J.A.L.). C.B. was supported by the Catharina Foundation Fellowship. N.V. was supported by the Jesse and Caryl Philips Foundation. This work was also supported by the Next-Generation Sequencing Core Facility and the Integrative Genomics and Bioinformatics Core Facility at the Salk Institute, with funding from NIH-NIH RdC Cancer Institute Cancer Center Support Grant P30 014195, the Glenn Foundation for Medical Research, the Chapman Foundation, and the Helmsley Charitable Trust.

1. Ciccia A, Elledge SJ (2010) The DNA damage response: Making it safe to play with knives. Mol Cell 40:179–204.
2. Manova V, Gruszka D (2015) DNA damage and repair in plants—From models to crops. Front Plant Sci 6:685.
3. Aguilera A, Garcia-Muse T (2013) Causes of genome instability. Annu Rev Genet 47:1–32.
4. Rodgers K, McVey M (2016) Error-prone repair of DNA double-strand breaks. J Cell Physiol 231:15–24.
5. Hu Z, Cools T, De Veylder L (2016) Mechanisms used by plants to cope with DNA damage. Annu Rev Plant Biol 67:439–462.
6. Su TT (2006) Cellular responses to DNA damage: One signal, multiple choices. Annu Rev Genet 40:37–61.
7. Amiard S, Gallego ME, White CI (2013) Signaling of double strand breaks and damage response in plants. Trends Plant Sci 18:377–385.
8. Yoshiyama KO, Sakaguchi K, Kimura S (2013) DNA damage response in plants: Annu Rev Genet 47:173–197.
9. Bleuyard JY, Gallego ME, White CI (2006) Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA Repair (Amst) 5:1–12.
10. Shihol Y, Ziv Y (2013) The ATM protein kinase: Regulating the cellular response to genotoxic stress, and more. Nat Rev Mol Cell Biol 14:197–210.
11. Knoll A, Fauser B, Putha H (2014) DNA repair pathways in somatic plant cells: Mechanisms and evolutionary consequences. Chromosoma Res 22:191–201.
12. Preuss SB, Britt AB (2003) A DNA-damage-induced cell cycle checkpoint in Arabidopsis. Genetics 164:323–334.
13. Vida M, Renkonen PA, Huener DF, Britt AB (2009) Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. Proc Natl Acad Sci USA 106:12843–12848.
14. Furukawa T, et al. (2010) A shared DNA-damage-response pathway for induction of stomata death in Arabidopsis. Science 328:62–64.
15. Adachi S, et al. (2011) Programmed induction of endoreduplication by DNA double-strand breaks in Arabidopsis. Plant Cell 23:2501–2515.
16. Yi D, et al. (2016) The Arabidopsis SIAMESE-RELATED cyclin-dependent kinase inhibitor SMR5 and SMR7 regulate the DNA damage checkpoint in response to reactive oxygen species. Plant Cell 26:296–309.
17. Ogita N, et al. (2018) Identifying the target genes of SUPPRESSOR OF GAMMA RESPONSE 1, a master transcription factor controlling DNA damage response in Arabidopsis. Plant J 94:439–453.
18. Weimer AK, et al. (2016) The plant-specific CDK81-CYCB1 complex mediates homologous recombination repair in Arabidopsis. EMBO J 35:2068–2086.
19. Charbonnel C, Gallego ME, White CI (2010) Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in Arabidopsis plants. Plant J 64:280–290.
20. Charbonnel C, Allain E, Gallego ME, White CI (2011) Kinetic analysis of DNA double-strand break repair pathways in Arabidopsis. DNA Repair (Amst) 10:611–619.
21. Friesner R, Li S, Liu B, Culligan K, Britt AB (2005) Inhibitors SMR5 and SMR7 selectivity of the DNA damage response in Arabidopsis. Cell 121:646–657.
22. Bourbousse et al. 2017 Nature 550:216–220.
55. Kim JH, et al. (2007) Transcriptomic profile of Arabidopsis rosette leaves during the reproductive stage after exposure to ionizing radiation. Radiat Res 167:267–280.

56. Kim JLG, et al. (2014) Differentially expressed genes in response to gamma-irradiation during the vegetative stage in Arabidopsis thaliana. Mol Biol Rep 41:2229–2241.

57. Nagata T, Yamada H, Du Z, Todokoro S, Kikuchi S (2005) Microarray analysis of genes that respond to gamma-irradiation in Arabidopsis. J Agric Food Chem 53:1022–1030.

58. Wang Z, Schwartz R, Kunze R (2016) DNA damage-induced transcription of transposable elements and long non-coding RNAs in Arabidopsis is rare and ATM-dependent. Mol Plant 9:1142–1155.

59. Kim DS, et al. (2011) Antioxidant response of tobacco to gamma-irradiation: Expression profiling of the ROS scavenging and signal transduction pathways. J Plant Physiol 168:1960–1970.

60. Ricaud L, et al. (2007) ATM-mediated transcriptional and developmental responses to gamma-rays in Arabidopsis. PLoS One 2:4240.

61. Missirian V, Conklin PA, Culligan KM, Hughen ND, Britt AB (2014) High atomic weight, high-energy radiation (HZE) induces transcriptional responses shared with conventional stresses in addition to a core “DSB” response specific to clastogenic treatments. Front Plant Sci 5:2364.

62. Chen J, Vainas Q, Harbison CT, Simon I, Bar-Joseph Z (2007) Reconstructing dynamic regulatory maps. Mol Syst Biol 3:74.

63. Schulz MH, et al. (2012) DREM 2.0: Improved reconstruction of dynamic regulatory interaction: Genome-wide expression profiling of the ROS scavenging and signal transduction pathways. J Plant Physiol 169:123–133.

64. Thomas JF, et al. (2011) AGROVOC: The Arabidopsis Gene Regulatory Information Server, an open source software for accessing Gene Regulatory Information Server, is rare and ATM-sensitive. Mol Syst Biol 7:154.

65. R1R2R3-Myb proteins are essential for inhibiting cell division in response to DNA damage. Nat Rev Mol Cell Biol 12:341–351.

66. Kim JB, et al. (2014) Differentially expressed genes in response to gamma-irradiation in Arabidopsis thaliana reveal functional differences relative to other Arabidopsis accessions. PLoS One 9:e94122.

67. Schommer C, Beven A, Lawrenson T, Shaw P, Sablowski R (2003) AHP2 is required for sister kinetochore movement during the vegetative stage in Arabidopsis. Plant J 35:269–280.

68. Shultz RW, Tatineni VM, Hanley-Bowdoin L, Thompson WF (2007) Genome-wide analysis of DNA double-strand break repair using a template rescue machinery in the higher plants Arabidopsis and rice. Plant Physiol 144:1697–1714.

69. Garcia-Ortiz MV, Ariza RR, Hoffman PD, Hays JB, Roldán-Arjona T (2004) Arabidopsis thaliana AtPOLK encodes a D1-like DNA polymerase that extends mispaired termini and is highly expressed in a variety of tissues. Plant J 39:98–144.

70. Yin H, et al. (2009) Epigenetic regulation, somatic homologous recombination, and array-ssic acid signaling are influenced by DNA polymerase epsilon mutation in Arabidopsis. Plant Cell 21:386–402.

71. Yilmaz A, et al. (2011) AGRIS: The Arabidopsis Gene Regulatory Information Server, an open source software for accessing Gene Regulatory Information Server, is rare and ATM-sensitive. Mol Syst Biol 7:154.

72. Yilmaz A, et al. (2011) AGRIS: The Arabidopsis Gene Regulatory Information Server, an open source software for accessing Gene Regulatory Information Server, is rare and ATM-sensitive. Mol Syst Biol 7:154.

73. Pedrosa-García JA, Nájera-Martínez M, de la Paz Sanchez M, Plasencia J (2015) Arabidopsis thaliana thymidin kinase 1 is ubiquitously expressed during development and contributes to confer tolerance to genotoxic stress. Plant Mol Biol 87:303–315.

74. Jacobs AL, Schäp P (2012) DNA glycosylases: In DNA repair and beyond. Chromosoma 121:1–20.

75. Laura S, et al. (2016) DREEM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. BMC Syst Biol 6:104.

76. St. Bourgeisse C, Vegenas N, Law JA (2018) The SOG1 transcriptional activator and the MYB3R repressors control a complex gene network in response to DNA damage in Arabidopsis. Gene Expression Omnibus (GEO). Available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112773. Deposited April 5, 2018.

77. Daoulas RA, et al. (2013) Systemic protection of the ROS scavenging and signal transduction pathways. J Plant Physiol 168:1960–1970.

78. O’Malley RC, et al. (2016) Cistrome and epicistrome features shape the regulary DNA landscape. Celf 165:1209–1292.

79. Fu ZQ, Dong X (2015) Systemic acquired resistance: Turning local infection into global defense. Annu Rev Plant Biol 64:839–863.

80. Song J, Bent AF (2014) Microbial pathogens trigger host DNA double-strand breaks whose abundance is reduced by plant defense responses. PLoS Pathog 10:e1004030.

81. Shao H, Wang H, Yang X (2015) NAC transcription factors in plant multiple abiotic stress responses: Progress and prospects. Front Plant Sci 6:902.

82. Zhang S, et al. (2015) C-terminal domains of a histone demethylase interact with a pair of transcription factors and mediate specific chromatin association. Cell Discov 1:15003.

83. Chen P, et al. (2017) Arabidopsis R1R2R3-Myb proteins are essential for inhibiting cell division in response to DNA damage. Nat Commun 8:365.

84. Haga N, et al. (2011) Mutations in MYB3R1 and MYB3R4 cause pleiotropic developmental defects and preferential down-regulation of multiple G2M-specific genes in Arabidopsis. Plant Physiol 157:706–717.

85. Ito M, et al. (1998) A novel cis-acting element in promoters of plant B-type cyclin genes activates M phase-specific transcription. Plant Cell 10:331–341.

86. Ito M, et al. (2001) G2M-phase-specific transcription during the plant cell cycle is mediated by Myb-like transcription factors. Plant Cell 13:1891–1905.

87. Meewisse M, et al. (2007) AGRIS: The Arabidopsis Gene Regulatory Information Server, an open source software for accessing Gene Regulatory Information Server, is rare and ATM-sensitive. Mol Syst Biol 3:994–996.

88. Waterworth WM, Bray CM, West CE (2015) The importance of safeguarding genome integrity in germination and seedling growth and development. J Exp Bot 66:3549–3558.

89. Dobin A, et al. (2013) STAR: Ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21.

90. Heinz S, et al. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576–589.

91. Ramírez F, et al. (2016) deepTools2: A next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44:W160-W165.

92. Heerler WV, et al. (2016) VennMaster: Area-proportional euler diagrams for functional GO analysis of microarrays. BMC Bioinformatics 9:67.

93. Bailey TL, et al. (2009) MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208.

94. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS (2007) Quantifying similarity between motifs. Genome Biol 8:R24.

95. Boyle EL, et al. (2004) GO-TermFinder—Open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics 20:3130–3137.

96. Supé F, Bosnjak M, Šáčka N, Smuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6:e21800.

97. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

98. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36.

99. Mi H, et al. (2016) PANTHER version 11: Expanded annotation data from gene ontology and pathway databases, and pathway data tool enhancements. Nucleic Acids Res 45:D183–D189.

100. Krizhanskaya V, et al. (2015) Arabidopsis Information portal. Nucleic Acids Res 43:D1003–D1009.

101. Fischer M (2017) Censure and evaluation of p53 target genes. Oncogene 36:3943–3956.

102. Kenzelmann Brod D, et al. (2013) Global genomic profiling reveals an extensive p53-regulated autophagy program contributing to key p53 responses. Genes Dev 27:1016–1031.