

RESEARCH ARTICLE

Open Access



Stress-responsive pathways and small RNA changes distinguish variable developmental phenotypes caused by *MSH1* loss

Mon-Ray Shao¹, Sunil Kumar Kenchanmane Raju¹, John D. Laurie^{1,2}, Robersy Sanchez¹ and Sally A. Mackenzie^{1*}

Abstract

Background: Proper regulation of nuclear-encoded, organelle-targeted genes is crucial for plastid and mitochondrial function. Among these genes, *MutS Homolog 1 (MSH1)* is notable for generating an assortment of mutant phenotypes with varying degrees of penetrance and pleiotropy. Stronger phenotypes have been connected to stress tolerance and epigenetic changes, and in *Arabidopsis* T-DNA mutants, two generations of homozygosity with the *msh1* insertion are required before severe phenotypes begin to emerge. These observations prompted us to examine how *msh1* mutants contrast according to generation and phenotype by profiling their respective transcriptomes and small RNA populations.

Results: Using RNA-seq, we analyze pathways that are associated with *MSH1* loss, including abiotic stresses such as cold response, pathogen defense and immune response, salicylic acid, MAPK signaling, and circadian rhythm. Subtle redox and environment-responsive changes also begin in the first generation, in the absence of strong phenotypes. Using small RNA-seq we further identify miRNA changes, and uncover siRNA trends that indicate modifications at the chromatin organization level. In all cases, the magnitude of changes among protein-coding genes, transposable elements, and small RNAs increases according to generation and phenotypic severity.

Conclusion: Loss of *MSH1* is sufficient to cause large-scale regulatory changes in pathways that have been individually linked to one another, but rarely described all together within a single mutant background. This study enforces the recognition of organelles as critical integrators of both internal and external cues, and highlights the relationship between organelle and nuclear regulation in fundamental aspects of plant development and stress signaling. Our findings also encourage further investigation into potential connections between organelle state and genome regulation vis-à-vis small RNA feedback.

Keywords: *MSH1*, Transcriptome, Plastid, Organelle, Stress, Small RNA

Background

As sessile organisms, plants must be able to perceive and adapt to changing environmental cues such as stress conditions. Stresses can be present in combinations, such as the simultaneous heat and drought, or a combination of abiotic and biotic stresses, and can be highly detrimental to plant growth [1, 2]. Under such circumstances, plant

responses are often non-additive compared to the individual stresses, suggesting a complex regulatory network with significant crosstalk [2, 3]. Understanding these regulatory networks, their signaling components, and the genes that influence them are therefore important topics in plant biology.

Because plastids are required for photosynthesis, the production of major phytohormones, and other metabolic processes, they are in a key position to sense alterations in the environment and communicate accordingly to the nucleus [4]. For example, light quality affects the

* Correspondence: sally.mackenzie@unl.edu

¹Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA

Full list of author information is available at the end of the article



plastid transcriptome while light quantity affects plastid signaling, and both influence photomorphogenesis [5, 6]. Plastids also play a role in tolerance and signaling against drought, freezing, heat, and oxidative stress [4, 7, 8]. Furthermore, disruption of regulators of plastid function, such as pentatricopeptide repeat (PPR) proteins, can lead to defects in growth, embryo development, photosynthesis, and leaf pigmentation, among others [9]. As such, the status of plastids has a profound level of control over the entire plant, including its responses to environmental and cellular stress.

MSH1, previously known as *CHM1*, is a plant-specific, nuclear-encoded *MutS* homolog that is targeted to both plastids and mitochondria. Loss of *MSH1* causes an array of phenotypes, including variegation, dwarfism, altered leaf morphology, delayed flowering, and male sterility [10–13]. Additional phenotypes are environmentally-dependent, such as secondary stem growth and aerial rosette formation under short-day conditions [14]. From *MSH1*-suppressed RNAi lines in sorghum, pearl millet, tomato, tobacco, and soybean, it is apparent that many of these phenotypes are conserved between monocots and eudicots [14–17]. In addition, manipulation of *MSH1* is associated with tolerance to heat, high light, and drought [15, 18, 19], particularly in *MSH1*-depleted plants showing strong developmental phenotypes. Indeed, *MSH1* transcript levels are endogenously down-regulated during stress [15], leading to the possibility that in *msh1* mutants, stress responses are triggered to cause growth suppression and other phenotypes.

A further consequence of *MSH1* loss is epigenetic, with evidence first appearing from the segregation of *MSH1* RNAi plants maintained for multiple generations as hemizygotes. A proportion of subsequent wild-type segregant progeny lacking the RNAi transgene still retained altered growth and delayed flowering phenotypes, which were not cytoplasmically heritable [14]. Furthermore, whole genome bisulfite-sequencing of *msh1* T-DNA mutants revealed numerous changes in DNA methylation over both gene bodies and transposable elements [20]. In plants, one role of DNA methylation is used to silence transposable elements, which can become activated in during stress conditions [21]. In some cases, changes in DNA methylation have also been associated with stress-induced gene regulation, such as during phosphate starvation or *Pseudomonas syringae* infection [22, 23], and may also provide the mechanisms basis for stress priming and memory [23, 24].

During propagation of the *msh1* T-DNA materials, we observed that first-generation homozygous *msh1* mutants (S1) had either no phenotype or only slight variegation, whereas second-generation homozygous *msh1* mutant (S2) plants displayed the full range of *msh1*-associated phenotypes [20]. Compared to S1 generation

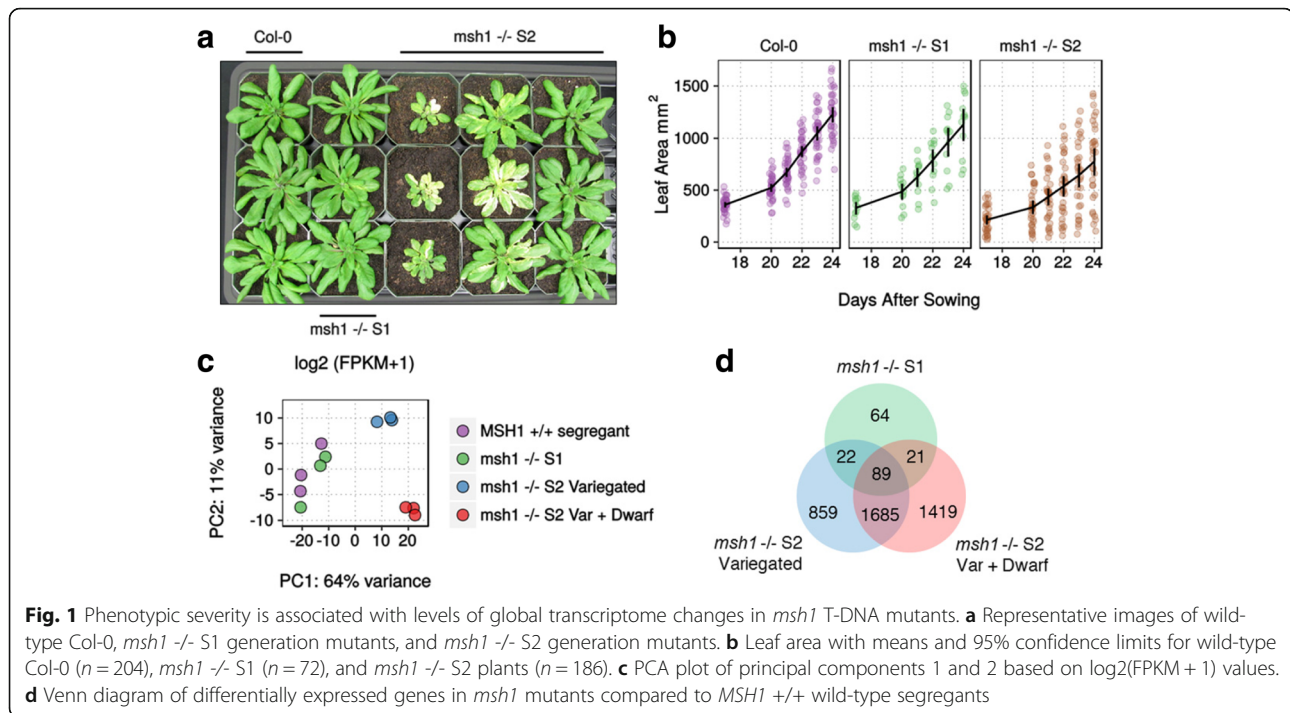
msh1 plants, S2 generation *msh1* plants with mutant phenotypes also had markedly increased amounts of methylation changes in the non-CG context [20]. This contrast raises questions as to what transcriptional changes begin to occur in the *msh1* *-/-* S1 plants, as opposed to *msh1* *-/-* S2 plants or later generations. We hypothesized that the degree of gene expression changes would parallel phenotype and methylome state, distinguishing the transition between the S1 and S2 generations. In this study, we performed RNA-seq to identify genes that are altered in the first sporophytic generation of *MSH1* loss, as well as those that are induced with the onset of strong phenotypes in the subsequent generation. Using small RNA-seq, we also show that miRNA profiles and repeat-associated siRNA levels change according to *msh1* generation and phenotype. Together, these data indicate that the variable phenotypes resulting from *msh1* loss are caused by the triggering of large gene expression networks associated with stress and other pathways, which also ultimately influence genome-wide changes in chromatin organization.

Results

Phenotypic and transcriptomic changes from *MSH1* loss build over two generations

We propagated a T-DNA insertion line for the *MSH1* locus that contained a mixture of seeds hemizygous and homozygous for the exon insertion (Additional file 1: Figure S1A-B). By self-pollinating a hemizygous *MSH1* +/- plant, we could observe the phenotype and changes occurring in the immediate progeny generation (S1) lacking *MSH1*. As noted in a previous study [20], these *msh1* *-/-* S1 plants were mild in phenotype (Fig. 1a), with occasional, slight leaf variegation being the only phenotype observed under normal growth conditions. In contrast, homozygous *msh1* *-/-* plants from generation S2 show the range of phenotype seen in *chm1-1* mutants, such as strong variegation and a wide variation in leaf area (Fig. 1a-b), growth rate and flowering time. From two S2 lines, variegation was present in approximately 70% of plants (Additional file 1: Figure S1C). As such, mutants showing dwarfism or stunted growth typically, though not necessarily always, also display some form of variegation or chlorotic phenotype.

To determine what transcriptome changes were associated with particular phenotypes, we performed RNA-seq on *msh1* *-/-* S1 plants, *msh1* *-/-* S2 variegated plants, and *msh1* *-/-* S2 variegated & dwarf plants chosen from a single lineage. *MSH1* *+/+* wild-type segregants also derived from the hemizygous T-DNA plants were used as the control group. Although there is a possibility that these wild-type segregants may contain subtle dosage-related effects from the previous hemizygous generation, they have no apparent phenotype. Furthermore, they are



very closely related in terms of lineage to the *msh1* mutant individuals, an important advantage for the purposes of this study where small RNA changes are also of interest and epigenetics changes may be present [20].

Because of the large overall number of gene expression changes observed in *msh1* mutants, and to minimize the number of potential false positives, we retained only differentially expressed genes (DEGs) identified by both Cuffdiff2 and DESeq2 (see Methods). Compared against the *MSH1* +/+ wild-type segregant control, we found 196 DEGs in *msh1* S1 plants, 2655 DEGs in *msh1* S2 variegated plants, and 3214 DEGs in *msh1* S2 variegated & dwarf plants (Additional file 2), suggesting that increasing phenotype severity is associated with increased gene expression changes. From principal component analysis, the *msh1* S2 groups are also more distinct from the wild-type segregants compared to the *msh1* S1 group (Fig. 1c), and a majority (66%) of the DEGs in *msh1* S2 variegated plants are also shared with *msh1* S2 variegated & dwarf plants (Fig. 1d).

Despite the relatively small total number of DEGs in *msh1* S1 plants, over two-thirds of these are also differentially expressed in at least one of the *msh1* S2 groups. Furthermore, of the 89 genes that are differentially expressed in all *msh1* mutants (S1, S2 variegated, and S2 variegated & dwarf), 38 genes show an intensifying trend between the S1 and S2 generation (Additional file 3: Table S1), defined by at least a four-fold change. Among these, several are transcription factors, such as *SCARECROW-LIKE 13*, *STZ*, *NAC036*, and *DEAR3*. Many are

also defense-related, including *WRKY40*, *WRR4*, *LURP1*, *PUB23*, *SDR3*, *CAF1a*, and *CNII* while others are responsive to environmental conditions such as cold (*CCR1*) or light (*PRIN2*). Therefore, despite the lack of a dramatic outward phenotype, several key pathways already appear to be modulated in the *msh1* -/- S1 plants. On the other hand, 7 genes showed a reversal in differential expression direction (i.e., down-regulation in S1 and up-regulation in S2, or vice versa), including *AOX1D*, the low carbon-induced *PHI-1/EXL1*, senescence-associated *SAG13*, the chloroplast-targeted aldo-keto reductase *AKR4C9*, the calcium transporter *CAX3*, and a cold acclimation *WCOR413*-family gene.

An increasing number of studies also indicate that alternative splicing is an important form of regulation in plant development and environmental responses [25, 26]. From our transcriptome data, a gene was considered alternatively spliced if it was identified by Cuffdiff2/spliceR [27, 28] as having differential isoform expression and by JunctionSeq [29] as having differential exon or splice junction usage. Using this methodology, there were 30, 139, and 470 genes with differential isoform expression in the *msh1* S1, *msh1* S2 variegated, and *msh1* S2 variegated & dwarf plants, respectively (Additional file 4). The majority of isoform-specific events pertained to alternative 5' or 3' splice sites, followed closely by alternative transcription start or termination sites (Additional file 5: Figure S2D). Roughly half of the isoforms (43 to 54%, depending on sample group) that are differentially expressed in the *msh1* mutants are also predicted to be sensitive to

nonsense-mediated decay (Additional file 5: Figure S2E), suggesting that these alternative isoforms may be a means to regulate functional transcript abundance [30]. Nevertheless, similar to gene expression, there is a trend of subtler isoform-specific changes in the *msh1* S1 generation followed by greater changes in the *msh1* S2 generation.

Expansion of differential expression and splicing in specific pathways reflect upon phenotype

We performed enrichment analysis using MapMan [31] functional categories on the differentially expressed genes. Several categories were shared between the *msh1* S2 mutants, such as auxin metabolism, stress, photosystem II components, regulation of WRKY transcription factors, and receptor kinase signaling (Fig. 2a). Cell organization, cell wall modifications, and regulation of bHLH, MYB-related, and Aux/IAA transcription factor families were particularly enriched in the *msh1* S2 variegated plants, while jasmonate metabolism, calcium signaling, and regulation of *CONSTANS*-like transcription factor families were particularly enriched in the *msh1* S2 variegated and dwarf plants. *CONSTANS*-like transcription factors are involved in photoperiod, flowering, and circadian rhythm [32, 33], which may be directly related with delayed growth and development in the dwarfed plants.

Among genes with differential isoform expression in *msh1* S2 variegated & dwarf plants, gene ontology (GO) enrichment analysis [34] found 26 categories that were enriched, including *response to salt stress* and *response to fructose* (Additional file 6). No such enrichment was statistically detectable in *msh1* S1 plants or *msh1* S2 variegated plants, although this is likely affected by the

lower overall number of differentially regulated isoforms in those groups. However, nearly all (22/26) of the enriched GO categories from analysis of differentially regulated isoforms in *msh1* S2 variegated & dwarf plants were also found in GO enrichment analysis of differentially expressed genes in *msh1* S2 mutants (Additional file 5: Figure S2F, red arrow), including *MAPK cascade*, *regulation of defense response*, *response to ethylene*, and *response to cold*. Therefore, in addition to differential overall gene expression, alternative isoform regulation and differential splicing likely play a meaningful role in the organismal response to *MSH1* loss.

Together, multiple enrichment analyses indicate that abiotic stress and several related pathways are significantly affected in *msh1* S2 mutants of both phenotypic classes. Among 41 multiple stress regulatory genes [35], 28 showed differential gene or isoform expression in *msh1* S2 mutants (Fig. 2b). Additionally, 95 drought-responsive genes [36] were differentially expressed in *msh1* mutants, with 7 differentially expressed as early as the *msh1* S1 generation (Additional file 7: Figure S3A). Differential expression of drought-responsive genes here is consistent with findings of heat and drought tolerance from *MSH1* perturbation [19]. However, 22 cold-responsive transcription factors [37] were also differentially expressed in *msh1* mutants, with 3 (*STZ*, *ERF2*, *ERF6*) up-regulated in the *msh1* S1 generation (Additional file 7: Figure S3B). The observed changes across multiple abiotic stress pathways indicate that loss of *MSH1* elicits a general stress response, rather than targeted changes to a specific stress. Triggering of these stress pathways in *msh1* mutants likely confers the previously observed stress tolerances, but may also be linked to growth rate.

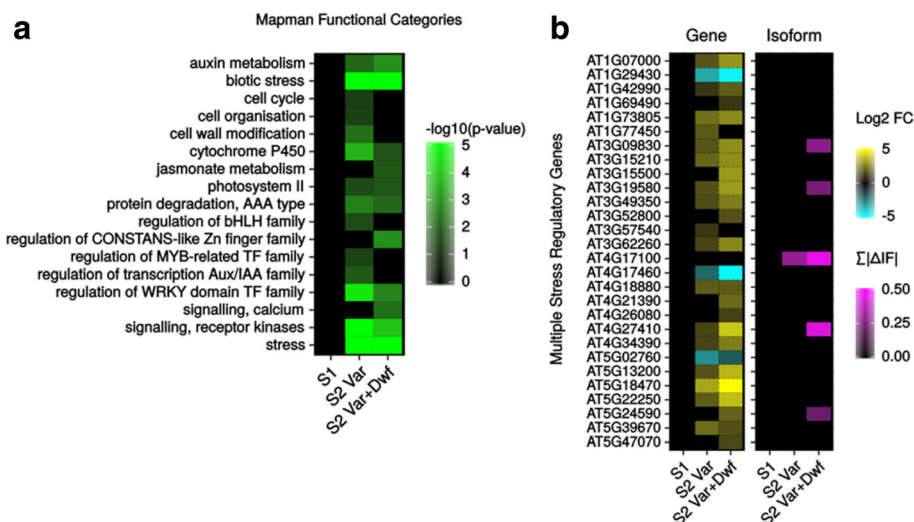


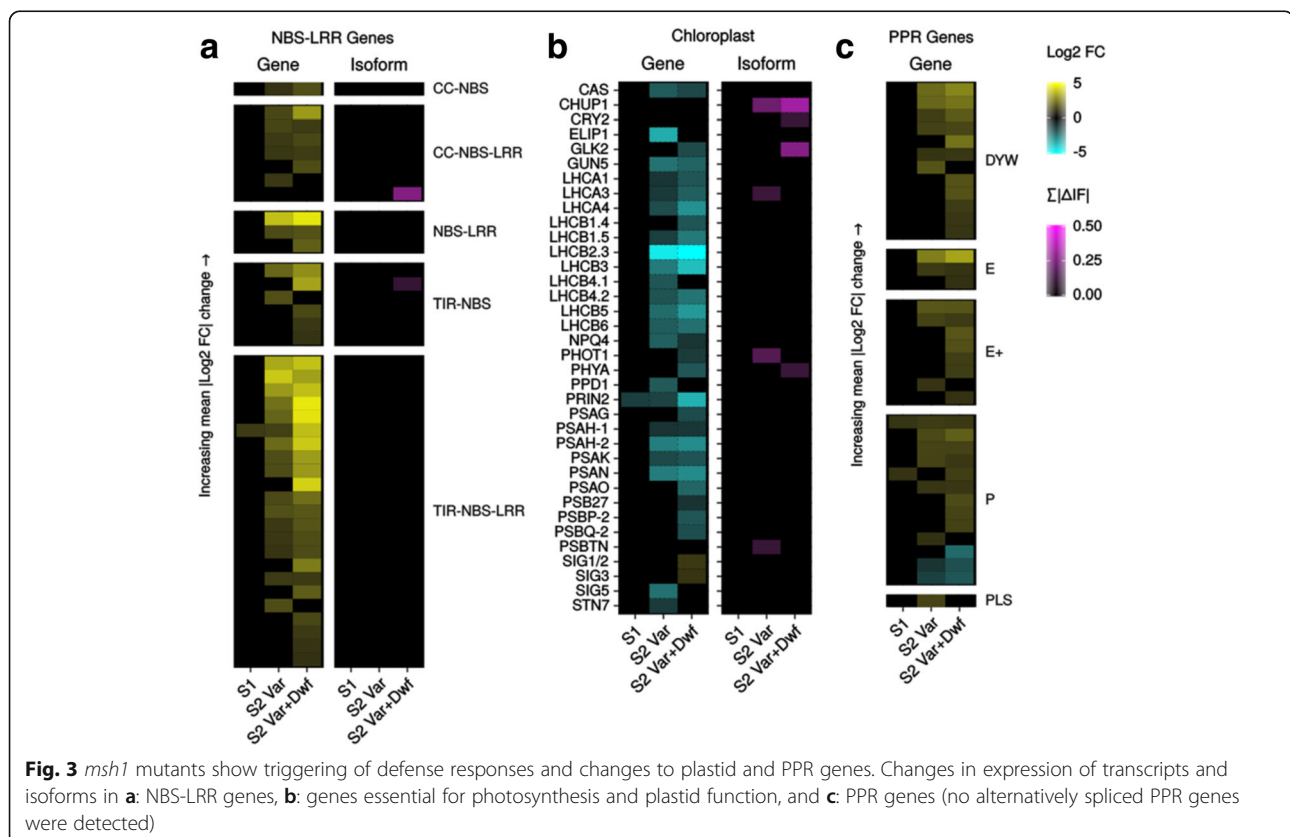
Fig. 2 *msh1* S2 mutants share enriched functional categories and changes in general stress gene. **a** Enrichment of MapMan functional categories among differentially expressed genes. **b** Changes in expression of transcripts and isoforms in multiple stress regulatory genes. Gene-level changes are measured by log₂ fold-change, while isoform-level changes are measured by the total change in isoform fractions for each gene

In addition to abiotic stress, the repeated observation of enriched GO categories related to biotic stress, such as *systemic acquired resistance*, *regulation of defense response*, and *regulation of plant-type hypersensitive response* is likely related to accumulating evidence of crosstalk between ROS signaling and defense responses [38–40]. This would be consistent with the enrichment of *respiratory burst involved in defense response* as early as the S1 generation, and sufficient activation of the immune response triggers a decrease in plant growth [41]. As suggested from enrichment analysis, we found that many genes involved in salicylic acid and jasmonate pathways were differentially expressed (Additional file 8: Figure S4A-B), as well as other phytohormones such as auxin (Additional file 8: Figure S4C). Further inspection revealed that a significant number of *NBS-LRR* genes, the largest family of disease resistance (*R*) genes, were also up-regulated (Fig. 3a). *R* genes can be triggered by a variety of environmental changes [42], and may also be particularly labile to changes in DNA methylation [43].

As expected, we also found down-regulation of many genes related to photosynthesis and plastid function, such as light harvesting complex and photosystem I and II components, as well as up-regulation of several *PPR* genes of various families (Fig. 3b-c). This includes down-regulation of *PRIN2*, which is involved in redox-mediated

retrograde signaling and expression of genes transcribed by the plastid-encoded RNA polymerase [44], beginning in the *msh1* S1 generation. *GLN5*, the ChlH subunit of Mg-chelatase involved in the tetrapyrrole-based retrograde signaling [45], was also down-regulated in *msh1 -/-* S2 plants. Previous studies found altered redox levels in *msh1* mutants, with plastoquinone and phylloquinone biased towards a reduced state [15]. Here we also see differential expression of genes related to redox, including the H₂O₂ scavenger *APX5*, and *MDAR3*, involved in the regeneration of reduced ascorbate (Fig. 4a). Among nuclear-encoded plastid or mitochondria-associated genes [46], *msh1* S2 variegated plants had a bias towards down-regulation, while in *msh1* S2 variegated & dwarf plants these genes are more evenly split between up-regulation and down-regulation (Additional file 9: Figure S5A-B). However, 46 and 40% of the same plastid-associated and mitochondrial-associated DEGs, respectively, were shared between the two groups (Additional file 9: Figure S5C-D).

Therefore, while one of the initial responses to *MSH1* loss is transcriptional changes in genes directly associated with the organelles, many other changes occur in the nucleus as an additional consequence. In addition to abiotic and biotic stress pathways previously mentioned, we observed differential expression of several core components of the circadian clock in *msh1* S2 mutants,



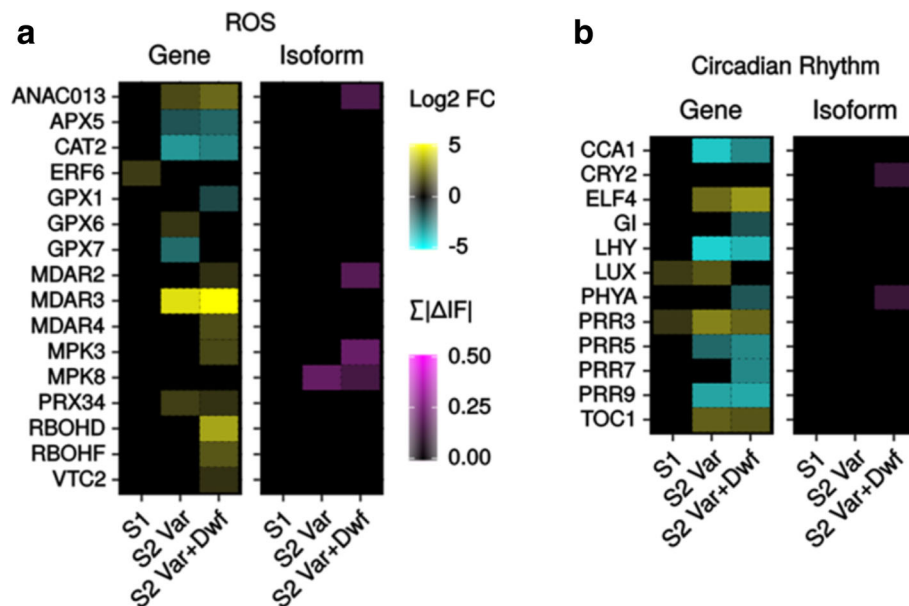


Fig. 4 ROS and circadian rhythm genes are altered in *msh1* mutants. Changes in expression of transcripts and isoforms in **a**: ROS related genes, and **b**: circadian rhythm genes

including *LHY* and *TOC1* (Fig. 4b). In *msh1* S2 variegated & dwarf plants, expression of *GI* was also altered. In addition, *PRR3*, a vascular regulator of *TOC1* [47], is differentially expressed beginning in *msh1* S1 generation mutants. Since circadian rhythm is linked to many other gene networks [48, 49], including ROS and phytohormone pathways that are altered in *msh1* mutants, these changes may be an important contributor to the broad pleiotropy in *msh1* mutants.

Rather than an unrelated assortment of pathways, there is accumulating literature indicating that many of those affected in these *msh1* mutants are interconnected as part of a large signaling network. For example, organelle perturbation is known to affect ROS levels [50], which impacts circadian rhythm and vice versa [49, 51]. In turn, both influence phytohormones [38, 52–54], calcium signaling [55, 56], and biotic and abiotic stress responses [57–60]. Therefore, many of the pathways affected in *msh1* can be considered as a series of direct consequences from the initial organelle perturbation, with signs of this beginning in the *msh1* S1 generation. In support of this, comparison of the *msh1* transcriptome revealed that a number of changes in such pathways were also found in the transcriptomes of plants with other plastid or chloroplast mutations, or chemical disruption of organelle function (Additional file 10: Figure S6) [61–70]. In *msh1* mutants, the number of genes within these pathways that are differentially expressed or spliced, and the intensity of their modulation, are associated with increasing phenotypic severity and comprise a distinguishing

feature between *msh1* S1, *msh1* S2 variegated, and *msh1* S2 variegated & dwarf plants.

Small RNA changes and differential TE expression suggest global chromatin alterations in *msh1* mutants

Because microRNAs are known to cause large effects in gene regulation and phenotype [71], we also performed small RNA-sequencing and aligned 20 to 24-nt reads to the genomic sequences of all high-confidence *A. thaliana* microRNAs annotated by miRBase [72]. In *msh1* S1 plants, only miR156 was up-regulated compared to the wild-type segregant control, and miR156 was also up-regulated in both the *msh1* S2 variegated and *msh1* S2 variegated & dwarf plants (Additional file 11: Table S2). The miR156 regulates the transition from vegetative to reproductive phase, and was previously found to be differentially expressed in *chm1* mutants [14]. miR163, involved in root architecture and secondary metabolites [73, 74], was also found to be up-regulated in both the *msh1* S2 variegated and *msh1* S2 variegated & dwarf plants, as was miR391, whose function is still unclear. miR169, which is involved in stress response and nutrient signaling [75–77], was down-regulated in both the *msh1* S2 variegated and *msh1* S2 variegated & dwarf plants.

A previous study indicated that the loss of *MSH1* leads to epigenetic changes [20], particularly in phenotypically dwarf plants and with respect to DNA methylation in the CHG and CHH sequence contexts. CHG/CHH methylation is primarily associated with heterochromatic regions

and transposable elements, and with short interfering RNA (siRNA) production. Small RNAs are intimately linked with chromatin organization [78], and believed to play a role in some stress responses, both biotic and abiotic [79–81]. We analyzed siRNA levels by mapping small RNA-sequencing reads to the entire genome using Short-Stack [82]. Most siRNA clusters identified belonged to the 24-nt class, particularly those that were differentially expressed (Additional file 12: Figure S7A), and so were the focus of further analysis. Analogous to what was observed with gene expression changes, *msh1* S1 plants only had 276 differentially expressed 24-nt siRNA clusters. In contrast, *msh1* S2 variegated and *msh1* S2 variegated & dwarf plants had 2646 and 2810 differentially expressed 24-nt siRNA clusters, respectively. In the *msh1* S2 variegated plants, the majority of these differentially expressed clusters were down-regulated, whereas in the *msh1* S2 variegated & dwarf plants, the majority were up-regulated (Fig. 5a).

In the *msh1* S2 plants, down-regulated 24-nt siRNA clusters are more localized to the pericentromere, whereas up-regulated 24-nt siRNA clusters are more spread throughout the chromosome and encompass euchromatic regions (Additional file 12: Figure S7B). Therefore, down-regulated 24-nt siRNA clusters more frequently overlapped transposable elements (TEs) and were generally closer to TEs than up-regulated 24-nt siRNAs were (Fig. 5b-c, Additional file 12: Figure S7C). However, up-regulated 24-nt siRNAs still tended to be located near TEs, but were also much closer to genes

than down-regulated 24-nt siRNAs were. Consistent with this observation, down-regulated 24-nt siRNA clusters were mostly associated with *Gypsy* elements, particularly in *msh1* S2 variegated plants, while up-regulated siRNA clusters were mostly associated with *Helitron* elements, particularly in *msh1* S2 variegated & dwarf plants (Fig. 5d). To a lesser degree, both up-regulated and down-regulated 24-nt siRNA clusters were associated with *MuDR* elements across all *msh1* mutants.

We then evaluated whether any transposable element families showed differential expression in *msh1* mutants using RNA-seq data. *Copia* and *MuDR* elements showed greatest tendency toward differential expression, regardless of direction of change (Additional file 13: Figure S8; Additional file 14: Table S3). Expression of TEs did not appear strongly correlated to siRNA changes, although this may be masked when evaluating at the superfamily or family level rather than individual elements. However, in a similar trend to DEGs and differential siRNA clusters, *msh1* *-/-* S2 variegated & dwarf plants showed the most changes (71 TE families), followed by *msh1* *-/-* S2 variegated plants (53 TE families), with *msh1* *-/-* S1 plants having the fewest (4 TE families). Overall, there were more instances of increased TE expression than decreased TE expression, and two families, *ATHILA4C* and *ATMU8*, had increased expression in all *msh1* mutant groups. Furthermore, in *msh1* S2 variegated and dwarf plants, DEGs were 1.7-fold enriched within 1 kb of *RathE1_cons* elements (chi-square test $p = 0.003$), which were also differentially expressed (Additional file 14: Table S3).

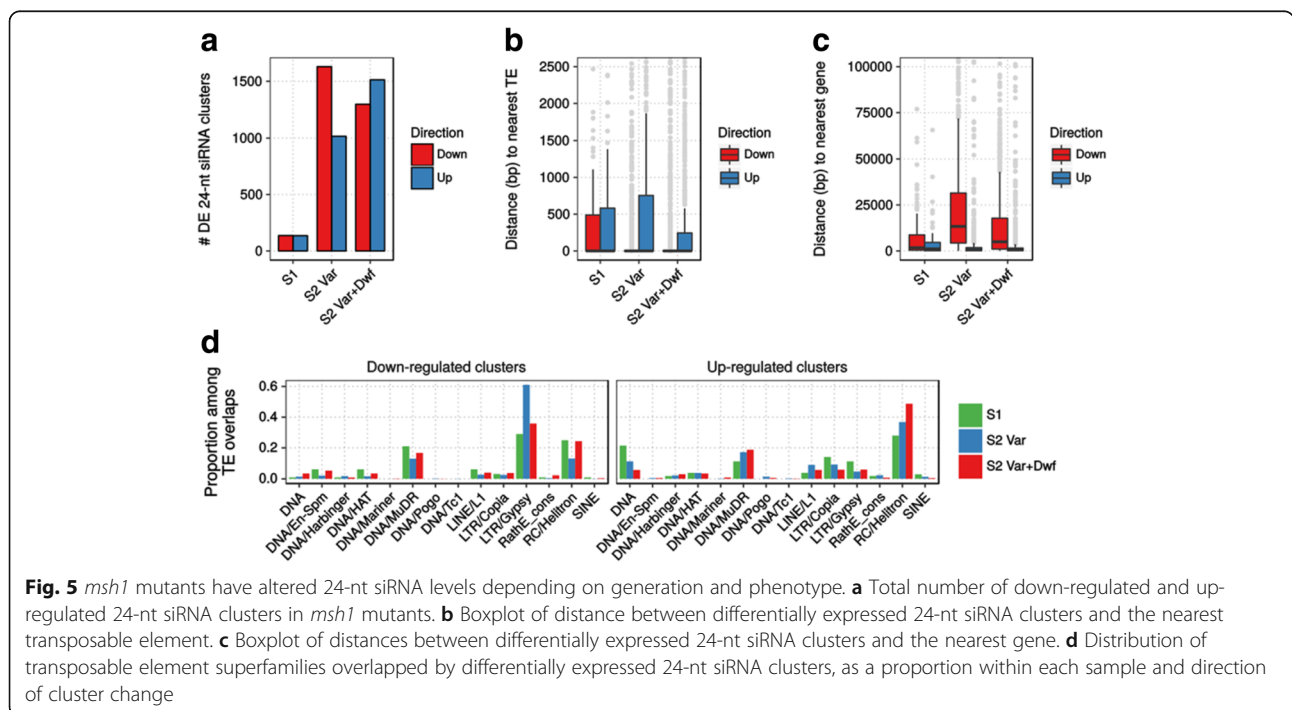


Fig. 5 *msh1* mutants have altered 24-nt siRNA levels depending on generation and phenotype. **a** Total number of down-regulated and up-regulated 24-nt siRNA clusters in *msh1* mutants. **b** Boxplot of distance between differentially expressed 24-nt siRNA clusters and the nearest transposable element. **c** Boxplot of distances between differentially expressed 24-nt siRNA clusters and the nearest gene. **d** Distribution of transposable element superfamilies overlapped by differentially expressed 24-nt siRNA clusters, as a proportion within each sample and direction of cluster change

Therefore, although not apparently widespread, differential expression of certain transposable elements could be affecting nearby gene expression [21].

We did not find any statistically significant overall enrichment of DEGs proximal to differentially expressed 24-nt siRNA clusters (using maximum distances varying between 50 bp and 10 kb), suggesting that siRNA changes here do not impact gene expression at the genome-wide level in an obvious manner. However, two large gene families, *R* genes and *PPR* genes, tend to be associated with repeats and thus may be susceptible to regulation via siRNAs [83, 84]. Although no enrichment was found for *R* genes, we observed a 2.2-fold enrichment for differential expression of *PPR* genes within 5 kb of a differentially expressed siRNA cluster (chi-square test $p = 0.02$) in *msh1* S2 variegated & dwarf plants, suggesting a possible link between *PPR* gene expression and proximal siRNAs. These findings of differential siRNA and TE expression in *msh1* mutants, coupled with previous studies showing DNA methylation changes, indicate that loss of *msh1* results in broad, global alterations in chromatin organization.

Discussion

Organelles are essential centers of regulation for multiple cellular signals; seedling, leaf, and flower development; energy and carbon metabolism; and light and temperature adaptation in plants [4, 85–87]. Loss of the dual organelle-targeted gene *MSH1* triggers changes involving plastid and mitochondria states, resulting in altered ROS levels and organellar genome instability, especially within white leaf sectors, and a large reduction in sucrose metabolism [12–15]. The resulting variable phenotypes, which requires two generations of homozygosity for complete elaboration but is not amenable to fixation, complicates studies into its function. Here, we performed RNA-seq and sRNA-seq on mutants newly homozygous for *msh1* T-DNA insertion to expand upon the global and pathway-specific changes associated with *MSH1* loss of function, and in particular to each phenotype.

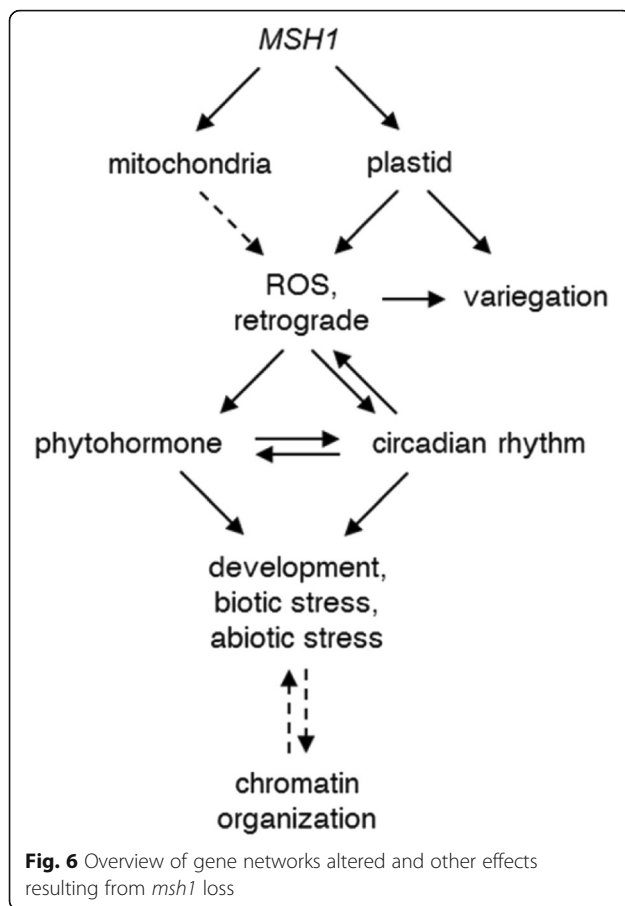
Using this segregating *msh1* T-DNA material, we found subtle gene expression changes in the first homozygous generation related to plastid redox regulation, plant defense, temperature response, and circadian rhythm. Despite this, the first homozygous generation shows little alteration in phenotype, and the total number of changes and the magnitude of the changes are still relatively small (196 DEGs). Since *MSH1* expression is particularly high within carpels [88], *MSH1* function may be particularly important during reproduction and the gametophyte stage, where disruption of organelle functions during this stage could have larger consequences on the cell [89]. This would be consistent with the much stronger phenotypes

observed by the second sporophytic generation, where gene expression changes expand into additional pathways, including auxin and other phytohormones influencing plant development, along with high induction of stress response pathways such as salt, cold, and oxidative stress.

Consistent with enrichment for abiotic and biotic stress transcriptional changes, *msh1* mutants have been found to be tolerant to high light and heat stress [15, 19]. In plants, repeated stress experience from drought, heat, or pathogens can lead to priming that then facilitates faster future responses within the same generation, and in some cases the underlying mechanism of memory appears to be epigenetic [24, 90]. Whole genome bisulfite-sequencing of *msh1* mutants previously revealed numerous changes in DNA methylation over both gene bodies and transposable elements [20], leading to the possibility of epigenetic feedback as a response to *MSH1* loss and heritable methylation changes at stress-responsive loci [24]. However, identification of causative DNA methylation changes is hampered by spontaneous differences that are a function of generational distance between individual plants or other stochastic events [91, 92]. This has complicated efforts to determine essential locus-specific methylation signals from *MSH1* loss, but the application of new methods in methylome analysis may aid future investigation [93, 94].

As another approach to epigenetic changes, we examined small RNA populations within *msh1* mutants. Much greater 24-nt siRNA changes occurred by the second sporophytic generation of *MSH1* loss than within the first generation, with different preferences in chromosome context (euchromatic or heterochromatic) between up-regulated and down-regulated clusters. However, with the exception of some *PPR* genes, we did not observe an association between differential siRNA levels and gene expression. Thus, although changes in siRNAs, TE expression, and DNA methylation strongly suggest changes to chromatin organization in *msh1* mutants, these could be broad but not locus-specific effects associated with induction of stress responses. Alternatively, the impact of specific chromatin-level changes upon gene expression may simply be uncommon or difficult to conclusively detect, but are important occurrences [95]. A summary schematic of the downstream effects of *MSH1* perturbation, from organelle perturbation to eventual chromatin-level alterations, is depicted in Fig. 6.

Variegation, chlorosis, or stunted growth phenotypes are also observed in several other plastid or chloroplast mutations [96]. In addition, chemical inhibitors of chloroplast biogenesis or function have also been identified. We reanalyzed public data from 12 such mutants or chemical treatments and compared enriched pathways from differentially expressed genes. We found that most mutants or chemical treatments also induced some degree of abiotic and biotic stress pathways that were also enriched in the



msh1 mutant. This supports the broader concept that organelle function is directly linked to stress responses. However, the *MSH1* gene has an unusual combination of characteristics, including a plastid nucleoid localization pattern within the epidermis and vascular parenchyma tissue, and a large number of secondary phenotypes [14, 15, 19]. Furthermore, to our knowledge, no other organellar gene mutant has been directly assayed for changes in DNA methylation [20]. Identification of protein interactors with *MSH1* would help place the functional relationship of *MSH1* with other chloroplast proteins and variegation mutants. Future studies will also be important for elucidating the direct importance of stress responses, circadian rhythm, and plastid function in *msh1* phenotypes, and could reveal previously undescribed loci under epigenetic regulation.

Conclusion

T-DNA mutants of the organellar-targeted gene *MSH1* have a wide variety of phenotypes – with variegation and stunted growth being the most obvious among them – but this is only apparent after two generations of homozygosity. Multiple abiotic stress response pathways, pathogen defense, photosynthesis, cell cycle, circadian rhythm,

MAPK signaling, and phytohormone regulation are all significantly impacted with far-reaching consequences to development, as the level of their induction differs according with phenotypic severity. Additionally, changes in small RNA populations and transposable element activity provide further evidence of alterations at the chromatin organization level. As *msh1* mutants have been shown to be tolerant to abiotic stress, these results highlight the connection between environment, organelle function, and nuclear regulation, an interplay that is important for understanding stress biology in plants.

Methods

Plant growth conditions and RNA extraction

Arabidopsis seeds were sown into pots containing Fafard Germination Mix soil added with Turface MVP. After 2–3 days of cold stratification at 4 °C, pots were moved to growing conditions set at 22 °C. Light regimes were set on a 12:12 h light–dark cycle with a light intensity of 150 $\mu\text{E m}^{-2} \text{s}^{-1}$.

For RNA-seq, individual *msh1* *-/-* S1, *msh1* *-/-* S2 variegated, *msh1* *-/-* S2 variegated & dwarf, and *MSH1* *+/+* wild-type segregant plants were used as biological replicates, with 3 replicates per group. Leaf tissue was harvested from plants near bolting developmental phase. For each sample, approximately 100 mg of frozen tissue was ground and extracted for total RNA using a standard TRIzol reagent protocol. RNA samples were then treated with DNaseI (Qiagen catalog #79254).

Seed stock and PCR genotyping

The *msh1* T-DNA seeds used in this study are part of the Syngenta Arabidopsis Insertion Library (SAIL) collection and ordered from the Arabidopsis Biological Resource Center (germplasm name SAIL_877_F01, stock number CS877617). Plants were genotyped using the following primers:

LP: 5' – ACGGAAAAAGTTCTTTCCAGG – 3'
 RP: 5' – GCTTCCATCGGCTAGGTTAG – 3'
 LB3: 5' – TAGCATCTGAATTCATAACCAATCT
 CGAT ACAC – 3'

with LP + RP to amplify the wild-type allele and LB3 + RP to amplify the T-DNA insertion. Polymerase chain reaction (PCR) based genotyping was performed using a temperature protocol of 95 °C for 3 min, followed by 41 cycles of 95 °C for 30 s, 50 °C for 30 s, 72 °C for 90 s, and finally 72 °C for 5 min. Leaf area was measured using a LemnaTec Scanalyzer HTS.

RNA-sequencing and analysis

Qiagen RNeasy Plant Mini Kit (Qiagen catalog #74904) was used to clean total RNA samples from *msh1* T-DNA

and wild-type segregants prior to RNA-sequencing (RNA-seq). Poly(A)-enriched RNA-seq was performed by CoFactor Genomics, generating at least 30.9 M single-end 50 bp reads per sample. To reduce false positives, we used two software pipelines and only retained concordant results. In both cases, reads were aligned to the Arabidopsis TAIR10 reference genome sequence with annotation from Araport11 PreRelease3.

In the first approach, alignment was performed using TopHat 2.1 (options *-b2-very-sensitive*, default mismatch and edit distance parameters) [97], resulting in ≥ 30.2 M mapped reads per sample. Cufflinks was used for pre-processing and masking reads derived from rRNAs, tRNAs, organellar-encoded transcripts, and transposable elements. Cuffdiff2 (options *-dispersion-method per-condition -library-norm-method geometric -u -b*) was then used to quantify and detect differentially expressed protein-coding genes ($FDR < 0.05$, $|\log_2(\text{fold change})| \geq 1$, and mean FPKM of control or test group ≥ 1). SpliceR was applied to the Cuffdiff2 output to quantify and characterize differentially expressed isoforms ($FDR < 0.05$, $|\Delta\text{isoform fraction}| \geq 0.1$, and mean FPKM of control or test group ≥ 1), and predict nonsense-mediated decay (NMD) sensitive transcripts.

In the second approach, alignment was performed using RUM 2.0.4 (default parameters) [98] keeping only uniquely mapped reads. DESeq2 [99] was used for gene count normalization and to identify differentially expressed genes ($FDR < 0.05$, $|\log_2(\text{fold change})| \geq 1$, and mean FPKM of control or test group ≥ 1). JunctionSeq was used for count normalization of known exons and splice junctions from Araport11 annotation and to then identify genes with at least one differentially used exon or splice junction (gene-wise $FDR < 0.05$, at least one exon or splice junction with $|\log_2(\text{fold-change})| \geq 0.5$).

For subsequent gene-level analysis, only those genes which were detected as differentially expressed in the same direction by both Cuffdiff2 and DESeq2 were kept (reported $\log_2(\text{fold-change})$ values in Results are from Cuffdiff2 estimates). Principle component analysis was based on FPKM (fragments per kilobase million) estimates from Cufflinks, using the 500 most variable genes across samples. For isoform-level analysis, only known isoform variants detected as differentially expressed by Cuffdiff2/spliceR and corresponding to genes with at least one differentially used exon or known splice junction according to JunctionSeq were kept (Additional file 5: Figure S2A-C), reported as the sum of absolute isoform fraction changes from Cuffdiff2/spliceR estimates, or $\sum|\Delta\text{IF}|$. Isoforms that are sensitive to nonsense mediated decay were identified by spliceR using default parameters, and are indicated in Additional file 2. All gene ontology enrichment analyses were performed using DAVID 6.8b, using the Benjamini method for multiple testing adjustment of p -values.

For TE family expression analysis, reads were aligned using the STAR 2-pass method [100] allowing up to 100 multi-mapped locations (per recommendation of TETranscripts). Quantification and testing for differential expression of TEs were performed using TETranscripts [101] with the developer-provided Arabidopsis TE family annotation.

Small RNA-sequencing and analysis

RNA samples of *msh1* T-DNA and wild-type segregants were extracted from the same plants and tissue used in RNA-seq. Small RNA-sequencing (sRNA-seq) was then performed by CoFactor Genomics, generating at least 11.1 M single-end 49 bp reads per sample, which were subsequently trimmed to remove adapters. One replicate of *Msh1* *+/+* wild-type segregant was dropped from analysis due to outlier concerns.

ShortStack v3.4 was used to align 20–24 nt sRNA-seq reads to the TAIR10 reference genome sequence with 0 mismatches allowed, resulting in ≥ 6.7 M mapped reads per sample. ShortStack-identified clusters were combined across all samples (stratified by size class), further merging any clusters within 75 bp of each other and removing those corresponding to miRNA, tRNA, or rRNA loci. This yielded a total of 44,348 clusters across the nuclear genome, with a median cluster length of 208 bp. For each sample, sRNA counts within each cluster by size class were tabulated. DESeq2 was used to identify differential sRNA clusters ($FDR < 0.05$, $|\log_2 \text{fold-change}| \geq 0.5$).

For miRNA analysis, 20 to 24-nt sRNA-seq reads were aligned using ShortStack 3.4 to the hairpin precursor sequences of all *A. thaliana* miRNAs from miRBase with 0 mismatches allowed, resulting in ≥ 161 k mapped reads per sample. After filtering these to only include miRNAs annotated as high confidence by miRBase 21, DESeq2 was used to identify differentially expressed miRNAs ($FDR < 0.05$, $|\log_2 \text{foldchange}| \geq 1$).

Microarray analysis

Publicly available microarray data were obtained from NCBI Gene Expression Omnibus and are described in Additional file 15: Table S4. For each data set, normalization and analysis was conducted using the R using the package “limma” [102]. For Agilent arrays, data was normalized using the *backgroundCorrect()* and *normalizeBetweenArrays()* functions. For Affymetrix arrays, data was normalized using the *rma()* function. Each data set was subsequently fit to a linear model and calculated for expression statistics using the empirical Bayes approach provided by “limma”. Differentially expressed genes against each study’s included wild-type control were selected using a cutoff of $FDR < 0.05$ and a $|\log_2 \text{foldchange}| \geq 0.5$. Gene ontology enrichment of differentially expressed genes was performed with DAVID 6.8b, using the Benjamini method for multiple testing adjustment of p -values.

Additional files

Additional file 1: Figure S1. a: Schematic of the T-DNA insertion at the *MSH1* locus. b: Diagram of pedigree relationship between T-DNA materials used in this study. Hemizygous *MSH1* T-DNA insertion individuals were self-pollinated to create first-generation homozygous *msh1* *-/-* individuals (S1) as well as homozygous *MSH1* *+/+* wild-type segregants. *msh1* *-/-* S1 plants were then self-pollinated to generate second-generation homozygous *msh1* *-/-* plants, which showed a range of phenotypes. c: Phenotypic scoring of 260 plants from two *msh1* *-/-* S2 lines gave an estimated variegation frequency of roughly 70%. (PDF 573 kb)

Additional file 2: Spreadsheet of differentially expressed genes in *msh1* mutants. Only concordant results (with Cuffdiff2 chosen as output) are included. (XLSX 940 kb)

Additional file 3: Table S1. List of genes with switching (from up-regulated to down-regulated or vice versa) or intensifying trends (≥ 4 -fold change in same direction) between S1 and S2 generations. (PDF 460 kb)

Additional file 4: Spreadsheet of differentially expressed isoforms in *msh1* mutants. Concordant results (with Cuffdiff2 chosen as output) and JunctionSeq results are included. (PDF 632 kb) (XLSX 15417 kb)

Additional file 5: Figure S2. Overlap of genes with differential isoform expression from TopHat2 + Cuffdiff2 and genes with at least one differentially used exon or splice junction from RUM + JunctionSeq2, in a: *msh1* S1 plants, b: *msh1* S2 variegated plants, and c: *msh1* S2 variegated & dwarf plants. d: Alternative splicing events and nonsense-mediated decay among differentially expressed isoforms. Each isoform may contain more than one type of alternative splicing event. e: Proportion of differentially expressed isoforms that are predicted to be sensitive to nonsense-mediated decay. f: Overlap of enriched GO terms from differentially expressed genes in *msh1* mutants, and genes with differentially expressed isoforms in *msh1* S2 variegated and dwarf plants. (PDF 632 kb)

Additional file 6: Spreadsheet of GO enrichment results for differentially expressed genes in *msh1* mutants and differentially expressed isoforms in *msh1* S2 variegated & dwarf plants. (XLSX 32 kb)

Additional file 7: Figure S3. Changes in expression of a: transcripts and isoforms of drought-responsive genes, and b: cold-responsive transcription factors. (PDF 879 kb)

Additional file 8: Figure S4. Changes in expression of transcripts and isoforms in a: salicylic acid, b: jasmonate, and c: auxin. (PDF 553 kb)

Additional file 9: Figure S5. Changes in expression of transcripts and isoforms among a: 2163 plastid-associated genes, and b: 1089 mitochondria-associated genes. Overlap of differentially expressed c: plastid-associated genes, and d: mitochondrial-associated genes, between *msh1* mutants. (PDF 532 kb)

Additional file 10: Figure S6. Comparison of transcriptome changes of several organelle mutants and chemical treatments, against biological processes enriched in *msh1* *-/-* S2 variegated & dwarf plants. Samples are arranged from left to right according to assay type (microarray or RNA-seq) and number of shared enriched categories with *msh1*. For RNA-seq of public data sets, differentially expressed genes were called using DESeq2. To reduce the otherwise very large number of GO enriched categories, only mappings directly annotated by the source database were used ("GO DIRECT"). Only mutants or treatments with at least 4 enriched GO DIRECT categories in common with *msh1* kept retained for analysis. Data were obtained from sources and studies listed in Additional file 7. (PDF 697 kb)

Additional file 11: Table S2. List of differentially expressed miRNA in *msh1* mutants. (PDF 470 kb)

Additional file 12: Figure S7. a: Total number of differentially-expressed siRNA clusters for each size class, by sample. b: Genomic distribution of differentially expressed 24-nt siRNA clusters. c: Proportion of differentially expressed 24-nt siRNA clusters overlapping genes and transposable elements. (PDF 616 kb)

Additional file 13: Figure S8. Number of differentially expressed transposable element families in *msh1* mutants, grouped by superfamily. (PDF 374 kb)

Additional file 14: Table S3. List of differentially expressed TE families in *msh1* mutants. (PDF 491 kb)

Additional file 15: Table S4. List of plastid or chloroplast mutants or chemical treatments used for comparative analysis, and their sources. (PDF 475 kb)

Abbreviations

DEGs: Differentially expressed gene; FPKM: Fragments per kilobase million; MSH1: MutS Homolog 1; ROS: Reactive oxygen species; S1: First generation homozygous self-pollinated plants; S2: Second generation homozygous self-pollinated plants; sRNA: Small RNA; TE: Transposable element

Acknowledgements

Not applicable.

Funding

We gratefully acknowledge support from The Bill and Melinda Gates Foundation (OPP1088661) to SM. The funding body had no role in the design, collection, analysis, or interpretation of the data, or in the writing of the manuscript or the decision to submit for publication.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available in the NCBI BioProject repository under the accession PRJNA342475 and can be obtained at <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA342475>. Information on public data sets used for analysis and their sources are listed in Additional file 15: Table S4.

Authors' contributions

M-RS, SK-KR, JDL, RS, and SM contributed to the design of the research. JDL generated the biological materials and sequencing data. SK-KR performed phenotypic scoring and measurements. M-RS, SK-KR, and RS contributed to analysis of RNA-seq data. M-RS performed analysis of small RNA-seq data. M-RS and SM wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

We have read and understood BMC policy on declaration of interests and declare the following interests: SM serves as a founder and member of the Scientific Advisory Board to EpiCrop Technologies.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA. ²Sainsbury Laboratory, University of Cambridge, Cambridge, UK.

Received: 19 September 2016 Accepted: 8 February 2017

Published online: 20 February 2017

References

- Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R. When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. *Plant Physiol.* 2004;134(4):1683–96.
- Atkinson NJ, Urwin PE. The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot.* 2012;63(10):3523–43.
- Mittler R. Abiotic stress, the field environment and stress combination. *Trends Plant Sci.* 2006;11(1):15–9.
- Chan KX, Phua SY, Crisp P, McQuinn R, Pogson BJ. Learning the languages of the chloroplast: retrograde signaling and beyond. *Annu Rev Plant Biol.* 2016;67:25–53.
- Dhingra A, Bies DH, Lehner KR, Folta KM. Green light adjusts the plastid transcriptome during early photomorphogenic development. *Plant Physiol.* 2006;142(3):1256–66.

6. Martín G, Leivar P, Ludevid D, Tepperman JM, Quail PH, Monte E. Phytochrome and retrograde signalling pathways converge to antagonistically regulate a light-induced transcriptional network. *Nat Commun.* 2016;7:11431.
7. Sun AZ, Guo FQ. Chloroplast retrograde regulation of heat stress responses in plants. *Front Plant Sci.* 2016;7:398.
8. Chan KX, Mabbitt PD, Phua SY, Mueller JW, Nisar N, Gigolashvili T, Stroehrer E, Grassl J, Arlt W, Estavillo GM, Jackson CJ, Pogson BJ. Sensing and signaling of oxidative stress in chloroplasts by inactivation of the SAL1 phosphoadenosine phosphatase. *Proc Natl Acad Sci U S A.* 2016;113(31):E4567–76.
9. Barkan A, Small I. Pentatricopeptide repeat proteins in plants. *Annu Rev Plant Biol.* 2014;65:415–42.
10. Rédei GP. Extra-chromosomal mutability determined by a nuclear gene locus in *Arabidopsis*. *Mut Res.* 1973;18:149–62.
11. Sakamoto W, Kondo H, Murata M, Motoyoshi F. Altered mitochondrial gene expression in a maternal distorted leaf mutant of *Arabidopsis* induced by chloroplast mutator. *Plant Cell.* 1996;8(8):1377–90.
12. Sandhu AP, Abdelnoor RV, Mackenzie SA. Transgenic induction of mitochondrial rearrangements for cytoplasmic male sterility in crop plants. *Proc Natl Acad Sci U S A.* 2007;104(6):1766–70.
13. Zhao N, Xu X, Wamboldt Y, Mackenzie SA, Yang X, Hu Z, Yang J, Zhang M. MutS HOMOLOG1 silencing mediates ORF220 substoichiometric shifting and causes male sterility in *Brassica juncea*. *J Exp Bot.* 2016;67(1):435–44.
14. Xu YZ, Santamaria Rde L, Virdi KS, Arrieta-Montiel MP, Razvi F, Li S, Ren G, Yu B, Alexander D, Guo L, Feng X, Dweikat IM, Clemente TE, Mackenzie SA. The chloroplast triggers developmental reprogramming when *MutS HOMOLOG1* is suppressed in plants. *Plant Physiol.* 2012;159(2):710–20.
15. Xu YZ, Arrieta-Montiel MP, Virdi KS, de Paula WB, Widhalm JR, Basset GJ, Davila JJ, Elthon TE, Elowsky CG, Sato SJ, Clemente TE, Mackenzie SA. MutS HOMOLOG1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light. *Plant Cell.* 2011;23(9):3428–41.
16. de la Rosa Santamaria R, Shao MR, Wang G, Nino-Liu DO, Kundariya H, Wamboldt Y, Dweikat I, Mackenzie SA. MSH1-induced non-genetic variation provides a source of phenotypic diversity in *Sorghum bicolor*. *PLoS One.* 2014;9(10):e108407.
17. Yang X, Kundariya H, Xu YZ, Sandhu A, Yu J, Hutton SF, Zhang M, Mackenzie SA. MutS HOMOLOG1-derived epigenetic breeding potential in tomato. *Plant Physiol.* 2015;168(1):222–32.
18. Shedge V, Davila J, Arrieta-Montiel MP, Mohammed S, Mackenzie SA. Extensive rearrangement of the *Arabidopsis* mitochondrial genome elicits cellular conditions for thermotolerance. *Plant Physiol.* 2010;152(4):1960–70.
19. Virdi KS, Wamboldt Y, Kundariya H, Laurie JD, Keren I, Kumar KR, Block A, Basset G, Luebker S, Elowsky C, Day PM, Roose JL, Bricker TM, Elthon T, Mackenzie SA. MSH1 is a plant organellar DNA binding and thylakoid protein under precise spatial regulation to alter development. *Mol Plant.* 2016;9(2):245–60.
20. Virdi KS, Laurie JD, Xu YZ, Yu J, Shao MR, Sanchez R, Kundariya H, Wang D, Riethoven JJ, Wamboldt Y, Arrieta-Montiel MP, Shedge V, Mackenzie SA. *Arabidopsis* MSH1 mutation alters the epigenome and produces heritable changes in plant growth. *Nat Commun.* 2015;6:6386.
21. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet.* 2007;8(4):272–85.
22. Yong-Villalobos L, González-Morales SI, Wrobel K, Gutiérrez-Alanis D, Cervantes-Peréz SA, Hayano-Kanashiro C, Oropeza-Aburto A, Cruz-Ramírez A, Martínez O, Herrera-Estrella L. Methylome analysis reveals an important role for epigenetic changes in the regulation of the *Arabidopsis* response to phosphate starvation. *Proc Natl Acad Sci U S A.* 2015;112(52):E7293–302.
23. Downen RH, Pelizzola M, Schmitz RJ, Lister R, Downen JM, Nery JR, Dixon JE, Ecker JR. Widespread dynamic DNA methylation in response to biotic stress. *Proc Natl Acad Sci U S A.* 2012;109(32):E2183–91.
24. Kinoshita T, Seki M. Epigenetic memory for stress response and adaptation in plants. *Plant Cell Physiol.* 2014;55(11):1859–63.
25. James AB, Syed NH, Bordage S, Marshall J, Nimmo GA, Jenkins GI, Herzyk P, Brown JW, Nimmo HG. Alternative splicing mediates responses of the *Arabidopsis* circadian clock to temperature changes. *Plant Cell.* 2012;24(3):961–81.
26. Shikata H, Hanada K, Ushijima T, Nakashima M, Suzuki Y, Matsushita T. Phytochrome controls alternative splicing to mediate light responses in *Arabidopsis*. *Proc Natl Acad Sci U S A.* 2014;111(52):18781–6.
27. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol.* 2013;31(1):46–53.
28. Vitting-Seerup K, Porse BT, Sandelin A, Waage J. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics.* 2014;15:81.
29. Hartley SW, Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. *Nucleic Acids Res.* 2016;44(15):e127.
30. Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol.* 2004;16(3):293–9.
31. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 2004;37(6):914–39.
32. Suárez-López P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G. CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature.* 2001;410(6832):1116–20.
33. Valverde F. CONSTANS and the evolutionary origin of photoperiodic timing of flowering. *J Exp Bot.* 2011;62(8):2453–63.
34. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
35. Ransbotyn V, Yeger-Lotem E, Basha O, Acuna T, Verduyn C, Gordon M, Chalifa-Caspi V, Hannah MA, Barak S. A combination of gene expression ranking and co-expression network analysis increases discovery rate in large-scale mutant screens for novel *Arabidopsis thaliana* abiotic stress genes. *Plant Biotechnol J.* 2015;13(4):501–13.
36. Sham A, Moustafa K, Al-Ameri S, Al-Azzawi A, Itrat R, AbuQamar S. Identification of *Arabidopsis* candidate genes in response to biotic and abiotic stresses using comparative microarrays. *PLoS One.* 2015;10(5):e0125666.
37. Leviatan N, Alkan N, Leshkowitz D, Fluhr R. Genome-wide survey of cold stress regulated alternative splicing in *Arabidopsis thaliana* with tiling microarray. *PLoS One.* 2013;8(6):e66511.
38. Rivas-San Vicente M, Plasencia J. Salicylic acid beyond defense: its role in plant growth and development. *J Exp Bot.* 2011;62(10):3321–38.
39. Wrzaczek M, Brosché M, Kangasjärvi J. ROS signaling loops - production, perception, regulation. *Curr Opin Plant Biol.* 2013;16(5):575–82.
40. Herrera-Vásquez A, Salinas P, Holuigue L. Salicylic acid and reactive oxygen species interplay in the transcriptional control of defense genes expression. *Front Plant Sci.* 2015;6:171.
41. Zhu Z, Xu F, Zhang Y, Cheng YT, Wiermer M, Li X, Zhang Y. *Arabidopsis* resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *Proc Natl Acad Sci U S A.* 2010;107(31):13960–5.
42. MacQueen A, Bergelson J. Modulation of *R*-gene expression across environments. *J Exp Bot.* 2016;67(7):2093–105.
43. Kawakatsu T, Huang SS, Jupe F, Sasaki E, Schmitz RJ, Ulrich MA, Castanon R, Nery JR, Barragan C, He Y, Chen H, Dubin M, Lee CR, Wang C, Bemm F, Becker C, O'Neil R, O'Malley RC, Quarles DX, 1001 Genomes Consortium, Schork NJ, Weigel D, Nordborg M, Ecker JR. Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell.* 2016;166(2):492–505.
44. Kindgren P, Kremnev D, Blanco NE, de Dios Barajas López J, Fernández AP, Tellgren-Roth C, Kleine T, Small I, Strand A. The plastid redox insensitive 2 mutant of *Arabidopsis* is impaired in PEP activity and high light-dependent plastid redox signaling to the nucleus. *Plant J.* 2012;70(2):279–91.
45. Mochizuki N, Brusslan JA, Larkin R, Nagatani A, Chory J. *Arabidopsis* genes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc Natl Acad Sci U S A.* 2001;98(4):2053–8.
46. Van Aken O, Whelan J. Comparison of transcriptional changes to chloroplast and mitochondrial perturbations reveals common and specific responses in *Arabidopsis*. *Front Plant Sci.* 2012;3:281.
47. Para A, Farré EM, Imaizumi T, Pruneda-Paz JL, Harmon FG, Kay SA. PRR3 is a vascular regulator of TOC1 stability in the *Arabidopsis* circadian clock. *Plant Cell.* 2007;19(11):3462–73.
48. Greenham K, McClung CR. Integrating circadian dynamics with physiological processes in plants. *Nat Rev Genet.* 2015;16(10):598–610.
49. Lai AG, Doherty CJ, Mueller-Roebber B, Kay SA, Schippers JH, Dijkwel PP. *CIRCADIAN CLOCK-ASSOCIATED 1* regulates ROS homeostasis and oxidative stress responses. *Proc Natl Acad Sci U S A.* 2012;109(42):17129–34.
50. Lepage É, Zampini É, Brisson N. Plastid genome instability leads to reactive oxygen species production and plastid-to-nucleus retrograde signaling in *Arabidopsis*. *Plant Physiol.* 2013;163(2):867–81.

51. Zhou M, Wang W, Karapetyan S, Mwimba M, Marqués J, Buchler NE, Dong X. Redox rhythm reinforces the circadian clock to gate immune response. *Nature*. 2015;523(7561):472–6.
52. Covington MF, Harmer SL. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS Biol*. 2007;5(8):e222.
53. Arana MV, Marín-de la Rosa N, Maloof JN, Blázquez MA, Alabadi D. Circadian oscillation of gibberellin signaling in *Arabidopsis*. *Proc Natl Acad Sci U S A*. 2011;108(22):9292–7.
54. Atamian HS, Harmer SL. Circadian regulation of hormone signaling and plant physiology. *Plant Mol Biol*. 2016;91(6):691–702.
55. Wood NT, Haley A, Viry-Moussaid M, Johnson CH, van der Luit AH, Trewavas AJ. The calcium rhythms of different cell types oscillate with different circadian phases. *Plant Physiol*. 2001;125(2):787–96.
56. Mazars C, Thuleau P, Lamotte O, Bourque S. Cross-talk between ROS and calcium in regulation of nuclear activities. *Mol Plant*. 2010;3(4):706–18.
57. Mühlenbock P, Szechynska-Hebda M, Plaszczyc M, Baudo M, Mateo A, Mullineaux PM, Parker JE, Karpinska B, Karpinski S. Chloroplast signaling and *LESION SIMULATING DISEASE1* regulate crosstalk between light acclimation and immunity in *Arabidopsis*. *Plant Cell*. 2008;20(9):2339–56.
58. Baxter A, Mittler R, Suzuki N. ROS as key players in plant stress signaling. *J Exp Bot*. 2014;65(5):1229–40.
59. Zhang C, Xie Q, Anderson RG, Ng G, Seitz NC, Peterson T, McClung CR, McDowell JM, Kong D, Kwak JM, Lu H. Crosstalk between the circadian clock and innate immunity in *Arabidopsis*. *PLoS Pathog*. 2013;9(6):e1003370.
60. Sanchez A, Shin J, Davis SJ. Abiotic stress and the plant circadian clock. *Plant Signal Behav*. 2011;6(2):223–31.
61. Luesse DR, Wilson ME, Haswell ES. RNA sequencing analysis of the *msl2ms3*, *cri*, and *gpps1* mutants indicates that diverse sources of plastid dysfunction do not alter leaf morphology through a common signaling pathway. *Front Plant Sci*. 2015;6:1148.
62. Hotto AM, Schmitz RJ, Fei Z, Ecker JR, Stern DB. Unexpected diversity of chloroplast noncoding RNAs as revealed by deep sequencing of the *Arabidopsis* transcriptome. *G3 (Bethesda)*. 2011;1(7):559–70.
63. Woodson JD, Perez-Ruiz JM, Schmitz RJ, Ecker JR, Chory J. Sigma factor-mediated plastid retrograde signals control nuclear gene expression. *Plant J*. 2013;73(1):1–13.
64. Miura E, Kato Y, Sakamoto W. Comparative transcriptome analysis of green/white variegated sectors in *Arabidopsis yellow variegated2*: responses to oxidative and other stresses in white sectors. *J Exp Bot*. 2010;61(9):2433–45.
65. Bermúdez MA, Páez-Ochoa MA, Gotor C, Romero LC. *Arabidopsis* S-sulfocysteine synthase activity is essential for chloroplast function and long-day light-dependent redox control. *Plant Cell*. 2010;22(2):403–16.
66. Laloi C, Stachowiak M, Pers-Kamczyc E, Warzych E, Murgia I, Apel K. Crosstalk between singlet oxygen- and hydrogen peroxide-dependent signaling of stress responses in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A*. 2007;104(2):672–7.
67. Dietzel L, Gläßer C, Liebers M, Hiekel S, Courtois F, Czarniecki O, Schlicke H, Zubo Y, Börner T, Mayer K, Grimm B, Pfanschmidt T. Identification of early nuclear target genes of plastidial redox signals that trigger the long-term response of *Arabidopsis* to light quality shifts. *Mol Plant*. 2015;8(8):1237–52.
68. Ng S, Ivanova A, Duncan O, Law SR, Van Aken O, De Clercq I, Wang Y, Carrie C, Xu L, Kmiec B, Walker H, Van Breusegem F, Whelan J, Giraud E. A membrane-bound NAC transcription factor, ANAC017, mediates mitochondrial retrograde signaling in *Arabidopsis*. *Plant Cell*. 2013;25(9):3450–71.
69. Ng S, Giraud E, Duncan O, Law SR, Wang Y, Xu L, Narsai R, Carrie C, Walker H, Day DA, Blanco NE, Strand Å, Whelan J, Ivanova A. Cyclin-dependent kinase E1 (CDKE1) provides a cellular switch in plants between growth and stress responses. *J Biol Chem*. 2013;288(5):3449–59.
70. Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J. Signals from chloroplasts converge to regulate nuclear gene expression. *Science*. 2007;316(5825):715–9.
71. Rubio-Somoza I, Weigel D. MicroRNA networks and developmental plasticity in plants. *Trends Plant Sci*. 2011;16(5):258–64.
72. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42(Database issue):D68–73.
73. Chung PJ, Park BS, Wang H, Liu J, Jang IC, Chua NH. Light-inducible miR163 targets *PXMT1* transcripts to promote seed germination and primary root elongation in *Arabidopsis*. *Plant Physiol*. 2016;170(3):1772–82.
74. Ng DW, Zhang C, Miller M, Palmer G, Whiteley M, Tholl D, Chen ZJ. *cis*- and *trans*-Regulation of miR163 and target genes confers natural variation of secondary metabolites in two *Arabidopsis* species and their allopolyploids. *Plant Cell*. 2011;23(5):1729–40.
75. Zhao B, Ge L, Liang R, Li W, Ruan K, Lin H, Jin Y. Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. *BMC Mol Biol*. 2009;10:29.
76. Zhao M, Ding H, Zhu JK, Zhang F, Li WX. Involvement of miR169 in the nitrogen-starvation responses in *Arabidopsis*. *New Phytol*. 2011;190(4):906–15.
77. Xu MY, Zhang L, Li WW, Hu XL, Wang MB, Fan YL, Zhang CY, Wang L. Stress-induced early flowering is mediated by miR169 in *Arabidopsis thaliana*. *J Exp Bot*. 2014;65(1):89–101.
78. Olovnikov I, Aravin AA, Fejes Toth K. Small RNA in the nucleus: the RNA-chromatin ping-pong. *Curr Opin Genet Dev*. 2012;22(2):164–71.
79. Sunkar R, Chinnusamy V, Zhu J, Zhu JK. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci*. 2007;12(7):301–9.
80. Phillips JR, Dalmay T, Bartels D. The role of small RNAs in abiotic stress. *FEBS Lett*. 2007;581(19):3592–7.
81. Huang J, Yang M, Zhang X. The function of small RNAs in plant biotic stress response. *J Integr Plant Biol*. 2016;58(4):312–27.
82. Johnson NR, Yeoh JM, Coruh C, Axtell MJ. Improved placement of multi-mapping small RNAs. *G3 (Bethesda)*. 2016;6(7):2103–11.
83. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell*. 2003;15(4):809–34.
84. Oliver KR, McComb JA, Greene WK. Transposable elements: powerful contributors to angiosperm evolution and diversity. *Genome Biol Evol*. 2013;5(10):1886–901.
85. Jarvis P, López-Juez E. Biogenesis and homeostasis of chloroplasts and other plastids. *Nat Rev Mol Cell Biol*. 2013;14(12):787–802.
86. Millar AH, Whelan J, Soole KL, Day DA. Organization and regulation of mitochondrial respiration in plants. *Annu Rev Plant Biol*. 2011;62:79–104.
87. Jacoby RP, Li L, Huang S, Pong Lee C, Millar AH, Taylor NL. Mitochondrial composition, function and stress response in plants. *J Integr Plant Biol*. 2012;54(11):887–906.
88. Shedje V, Arrieta-Montiel M, Christensen AC, Mackenzie SA. Plant mitochondrial recombination surveillance requires unusual *RecA* and *MutS* homologs. *Plant Cell*. 2007;19(4):1251–64.
89. Kelliher T, Walbot V. Hypoxia triggers meiotic fate acquisition in maize. *Science*. 2012;337(6092):345–8.
90. Boyko A, Kovalchuk I. Genome instability and epigenetic modification—heritable responses to environmental stress? *Curr Opin Plant Biol*. 2011;14(3):260–6.
91. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*. 2011;480(7376):245–9.
92. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urlich MA, Libiger O, Schork NJ, Ecker JR. Transgenerational epigenetic instability is a source of novel methylation variants. *Science*. 2011;334(6054):369–73.
93. Sanchez R, Mackenzie SA. Information thermodynamics of cytosine DNA methylation. *PLoS One*. 2016;11(3):e0150427.
94. Sanchez R, Mackenzie SA. Genome-wide discriminatory information patterns of cytosine DNA methylation. *Int J Mol Sci*. 2016;17(6):E938.
95. Probst AV, Mittelsten Scheid O. Stress-induced structural changes in plant chromatin. *Curr Opin Plant Biol*. 2015;27:8–16.
96. Rodermel S. *Arabidopsis* variegation mutants. *Arabidopsis Book*. 2002;1:e0079.
97. Kim D, Perteu G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
98. Grant GR, Farkas MH, Pizarro AD, Lahens NF, Schug J, Brunk BP, Stoeckert CJ, Hogenesch JB, Pierce EA. Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics*. 2011;27(18):2518–28.
99. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
100. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.

101. Jin Y, Tam OH, Paniagua E, Hammell M. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*. 2015;31(22):3593–9.
102. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

