

Small RNAs reflect grandparental environments in apomictic dandelion

Lionel Morgado^{1*}, Veronica Preite^{2*}, Carla Oplaat², Sarit Anava³, Julie Ferreira de Carvalho², Oded Rechavi³, Frank Johannes^{4,5} and Koen J.F. Verhoeven²

* Equal contributions

- 1 Groningen Bioinformatics Centre, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands
- 2 Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands
- 3 Department of Neurobiology, Wise Faculty of Life Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel
- 4 Population epigenetics and epigenomics, Technical University of Munich, Liesel-Beckmann-Strasse 2, Freising 85354, Germany.
- 5 Institute for Advanced Study, Technical University of Munich, Lichtenbergstrasse 2a, Garching 85748, Germany

Corresponding author: Koen Verhoeven (k.verhoeven@nioo.knaw.nl)

Abstract

Plants can show long-term effects of environmental stresses and in some cases a stress ‘memory’ has been reported to persist across generations, potentially mediated by epigenetic mechanisms. However, few documented cases exist of transgenerational effects that persist for multiple generations and it remains unclear if or how epigenetic mechanisms are involved. Here we show that the composition of small regulatory RNAs in apomictic dandelion lineages reveals a footprint of drought stress and salicylic acid treatment experienced two generations ago. Overall proportions of 21nt and 24nt RNA pools were shifted due to grandparental treatments. While individual genes did not show strong up- or downregulation of associated sRNAs, the subset of genes that showed the strongest shifts in sRNA abundance was significantly enriched for several GO terms including stress-specific functions. This indicates that a stress-induced signal was transmitted across multiple unexposed generations leading to persistent and functional changes in epigenetic gene regulation.

Stress exposure triggers responses that are mediated by changes in gene regulation (Cramer et al. 2011; Heil 2002; Shao et al. 2008). In plants, some responses to environmental stresses are long-lived. For instance, upon mild pathogen infection, plants can enter a “primed” state which is expressed as a quicker or more vigorous defense response upon a second infection later in life (Conrath et al. 2002). Defense-related priming effects, and also responses to other environmental triggers, have been demonstrated to persist into offspring generations in some cases (Agrawal 2002; Mandal et al. 2012; Slaughter et al. 2012; Wang et al. 2016).

Although several different mechanisms can underlie inherited environmental effects in plants (Crisp et al. 2016), epigenetic mechanisms are considered prime candidates because of their potential for environmental sensitivity (Downen et al. 2012) and transgenerational stability (Cortijo et al. 2014). Especially DNA methylation can be transgenerationally stable in plants and this mechanism is often proposed to mediate environmental effects that persist for multiple generations (Bilichak et al. 2015; Boyko et al. 2010; Boyko et al. 2007; Ou et al. 2012; Verhoeven et al. 2010) although empirical support for this hypothesis remains scarce (Pecinka and Scheid 2012).

Accumulating evidence indicates that regulatory small RNAs (sRNAs) also have a role in plant transgenerational effects. Indeed, sRNA biogenesis mutants in *A. thaliana* show compromised transgenerational herbivore resistance (Rasmann et al. 2012), suggesting that sRNAs are required to sustain priming responses across generations. Changes in sRNA composition have also been demonstrated in a number of species in response to heat (Bilichak et al. 2015; Ito et al. 2011; Song et al. 2016), drought (Matsui et al. 2008; Tricker et al. 2012), salinity (Borsani et al. 2005; Ding et al. 2009; Matsui et al. 2008; Song et al. 2016), cold and osmotic stress (Song et al. 2016). In some cases, these sRNA alternations have been shown to persist in the offspring of stressed plants. The mechanisms that maintain changes of sRNAs across

generations remain largely unclear but may involve feed-forward interactions with (transiently) heritable DNA methylation changes (Wibowo et al. 2016).

Here, we used apomictic dandelions (*Taraxacum officinale*) to test the impact of environmental stress on sRNA composition in unexposed offspring two generations after stress treatment. Due to apomictic reproduction dandelion offspring are considered clonal copies (Bicknell and Koltunow 2004) allowing for multi-generation experiments without confounding effects of genetic differences between samples. In this species, first-generation offspring of stress-exposed plants have previously been demonstrated to show modified phenotypes and DNA methylation patterns, suggesting potential for environment-induced transgenerational epigenetic inheritance (Verhoeven et al. 2010; Verhoeven and van Gurp 2012).

We grew first-generation (G1) plants under either drought stress, salicylic acid exposure (SA; a plant hormone that is involved in several processes including defense signaling in response to pathogens; Vicente and Plasencia 2011), or under control conditions. Second (G2) and third (G3) generation apomictic progenies were obtained by single-seed descent for four replicate lineages per experimental group and were grown under common control conditions (Supplementary text: **S1**). sRNAs were sequenced at generation G3 in four individual plants per experimental group (Supplementary text: **S1-S2**). As no reference assembly currently exists for dandelion, we first assessed differences in sRNA composition between treatment groups and control using the total sRNA libraries. We found significant shifts in all size classes of sRNAs ranging from 21 to 24 nt (**Fig. 1A and 1D**, Supplementary text: **S3**). The most pronounced changes occurred for sRNAs of size 24nt whose relative abundance in the total sRNA population was markedly reduced in both of the stress conditions compared to the control. Relative loss of TE-associated 24nt sRNA has been reported for a variety of biotic and abiotic stressors (Downen et al. 2012; Lunardon et al. 2016; McCue et al. 2012). These changes appear to be mainly the result of hypomethylation and loss of RdDM targeting of transposable

element (TE) sequences (McCue et al. 2013; Tran et al. 2005). Loss of 24 nt sRNAs is typically accompanied by gains in 21 nt sRNAs due to an increase in the transcription of precursors for this class of sRNA (Downen et al. 2012; McCue et al. 2012). Consistent with this, the relative loss in 24 nt sRNAs after drought stress co-occurred with gains in 21 nt sRNAs, although a similar trend could not be observed for the SA treatment (**Fig. 2A** and **2D**).

The fact that the libraries show signatures of known TE-associated stress-induced changes in sRNA (at least for the drought treatment) suggests that the observed effects are dominated by changes within TE sequences. To test this more directly we took advantage of a recent TE database that was generated based on de novo clustering of repetitive sequences from the *T. officinale* genome (Ferreira de Carvalho et al. 2016a)(SI text: **S2**). We aligned sRNAs to these TE sequences, and compared the relative size abundance across conditions. A loss of 24 nt sRNAs was also observed in these TE-annotated sequences (**Fig. 2B** and **2E**). Moreover, the gain in 21 nt sRNAs was now clearer and occurred in both stress treatments, suggesting that stress-induced increases in 21 nt sRNA are subtle and may have been obscured by the complexity of the total sRNA libraries.

It is unclear how and if such sRNA shifts impact gene expression. Previous studies have reported that TEs proximal to or overlapping genes can affect transcription under stress, probably as a result of RdDM-mediated DNA methylation loss (Downen et al. 2012; Hollister and Gaut 2009; Lister et al. 2008; Quadrana et al. 2016; Wang et al. 2013). Another mechanism by which sRNAs can affect gene expression is through *trans*-acting post-transcriptional modifications (Borsani et al. 2005). We recently assembled the complete transcriptome of dandelion (Ferreira de Carvalho et al. 2016b). Over 13500 genes could be annotated by homology with *A. thaliana*. We aligned our sRNA libraries to these transcriptomes (SI text: **S2**). On average about 53% of the reads from each library met our quality control criteria and could be successfully aligned. In contrast to TE-aligned sRNAs, the relative abundance of

transcriptome-aligned sRNAs was reduced after grandparental stress exposure not only for 24nt but also for 21nt sRNAs (**Fig. 2C** and **2F**).

The observed loss of transcript-associated 24 nt sRNAs was enigmatic, as 24nt sRNA are typically depleted in genic sequences. To explore this issue in more detail, we studied the density distribution of 24nt sRNA along our annotated transcripts. Our analysis shows that 24 nt sRNAs mapped more frequently towards the 5' and 3' flanks of the genes, suggesting vestiges of a sRNA signal that originates from sequences outside of gene bodies, such as promoters or intergenic regions (**Fig. 3**). In contrast, 21 nt sRNA showed a peak density toward the center of gene bodies. The distributional patterns reported here resemble previously reported genic signatures of sRNA abundance in well-annotated genomes such as maize (Gent et al. 2013; Lunardon et al. 2016), *A. thaliana* (Downen et al. 2012) and rice (Li et al. 2012). The relative decrease in transcript-associated 24 nt sRNA after stress exposure may suggest a loss of methylation in gene flanking regions (and possibly also in TE sequences within genes) and consequent gene expression upregulation. However, a quality genome reference assembly will be required to test this hypothesis. Together, the specific changes in sRNA profiles that we observed are in line with previous observations in stress-exposed plants, but our results indicate the stress-associated sRNA footprint is maintained transgenerationally for at least two unexposed generations after the stress treatment.

In order to identify specific genes that show different sRNA abundance comparing control and stress treatments, we performed differential analysis using DESeq2 (Love et al. 2014). We applied DESeq2 to different sRNA lengths: 21nt, 24nt sRNAs and for all length classes combined (18 to 30nt). After adjusting for multiple testing (FDR=0.10), our results showed virtually no significant sRNA enrichment or depletion at specific genes (Supplementary table: **S4**). We argue that the induced and transgenerationally inherited sRNA effects are subtle and may not be readily detectable using our approach that involved sRNA sequencing of individual

plants, not pooled samples. We therefore focused, instead, on sets of genes that were either most depleted or enriched for sRNA in the stress groups and we tested if these gene sets were overrepresented for specific GO-terms. Our analysis revealed a strong overrepresentation for hundreds of GO categories, depending on grandparental treatment and sRNA length class (Supplementary text: **S5**). A large fraction of these GO-categories overlapped between the two stress treatments, suggesting a generalized response. Genes that were depleted or enriched for 21 nt sRNAs were significantly enriched for about 400-500 GO categories in both the control-drought comparison and in the control-SA comparison. For 24 nt sRNAs the downregulated genes were enriched for many more GO terms than the upregulated genes, suggesting a strong biological signal in the relaxation of 24 nt-based gene silencing after grandparental stress.

We searched the list of enriched GO terms for specific keywords that are associated with the grandparental stresses: “water” and “drought” for drought treatment, “salicylic” and “hormone” for SA treatment and “response to stress”, “abiotic stimulus” and “wounding” for stress treatments in general (see **Fig. 3**). For instance, the GO term 0006950 (“response to stress”) was specially enriched, pointing to an active stress memory. For all other key words, except “drought”, significantly enriched GO terms were found in both stress treatments, suggesting that these GO terms reflect a general stress response rather than a treatment-specific response. However, two SA-related GO categories (GO term 0009862: systemic acquired resistance, salicylic mediated signaling pathway; and GO term 0009914: hormone transport) were affected only in the SA set, indicating a more treatment-specific pattern.

In summary, it is well known that stress responses can be mediated by changes in sRNA-associated gene silencing. Our results suggest that this regulation may persist for several generations after stress. sRNA-based multi-generational inheritance of environmental stress has been previously demonstrated in some animal systems (e.g. Gapp et al. 2014; Rechavi et al. 2014) where underlying mechanisms of sRNA inheritance are at least partly different from

plants. Although effects on gene expression remain to be evaluated, our study is to our knowledge the first demonstration in plants of modified sRNAs two generations removed from the stress trigger. Our results show no clear statistically significant effects on individual genes, which may be due to low sequencing depth of the libraries, or lack of sensitivity of our differential analysis. However, we were able to uncover a sRNA signal among genes involved in stress-related functions. This illustrates that an epigenetic signal travelled between generations preserving footprints of grandparental stress, and that this memory implicates genes that are known to be involved in stress responses. Although we did not explore the nature of the transgenerationally inherited epigenetic signal, this signal could be a stress-induced change in TE-associated DNA methylation, which in plants can be stably inherited and can trigger RNA-mediated gene expression changes in offspring (Wibowo et al. 2016).

Acknowledgments and funding information

Keygene N.V., Wageningen, kindly provided *T. officinale* BAC sequences that we used for preliminary analysis of TE-mapping sRNAs. Seeds from the *T. officinale hemicyclum* lineage were kindly provided by Jan Kirschner from the CAS Institute of Botany, Pruhonice. This work was supported by the Netherlands Organisation for Scientific Research (grant numbers 864.10.008 and 884.10.003); and an ERC starting grant to O.R. (grant number 335624). sRNA data generated for this study are deposited in SRA, submission number SUB2306247.

Literature Cited

- Agrawal AA. 2002. Herbivory and maternal effects: Mechanisms and consequences of transgenerational induced plant resistance. *Ecology* 83(12):3408-3415.
- Bicknell RA, Koltunow AM. 2004. Understanding apomixis: Recent advances and remaining conundrums. *Plant Cell* 16:S228-S245.

- Bilichak A, Ilnytskyy Y, Woycicki R, Kepeshchuk N, Fogen D, Kovalchuk I. 2015. The elucidation of stress memory inheritance in *Brassica rapa* plants. *Frontiers in Plant Science* 6.
- Borsani O, Zhu JH, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* 123(7):1279-1291.
- Boyko A, Blevins T, Yao YL, Golubov A, Bilichak A, Ilnytskyy Y, Hollander J, Meins F, Kovalchuk I. 2010. Transgenerational Adaptation of *Arabidopsis* to Stress Requires DNA Methylation and the Function of Dicer-Like Proteins. *Plos One* 5(3).
- Boyko A, Kathiria P, Zemp FJ, Yao YL, Pogribny I, Kovalchuk I. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants (Virus-induced plant genome instability). *Nucleic Acids Research* 35(5):1714-1725.
- Conrath U, Pieterse CMJ, Mauch-Mani B. 2002. Priming in plant-pathogen interactions. *Trends in Plant Science* 7(5):210-216.
- Cortijo S, Wardenaar R, Colome-Tatche M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury J-M, Wincker P et al. . 2014. Mapping the Epigenetic Basis of Complex Traits. *Science* 343(6175):1145-1148.
- Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. 2011. Effects of abiotic stress on plants: a systems biology perspective. *Bmc Plant Biology* 11.
- Crisp PA, Ganguly D, Eichten SR, Borevitz JO, Pogson BJ. 2016. Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. *Science Advances* 2(2).
- Ding D, Zhang LF, Wang H, Liu ZJ, Zhang ZX, Zheng YL. 2009. Differential expression of miRNAs in response to salt stress in maize roots. *Annals of Botany* 103(1):29-38.
- Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America* 109(32):E2183-E2191.
- Ferreira de Carvalho J, de Jager VCL, Van Gurp TP, Wagenmaker N, Verhoeven KJF. 2016a. Recent and dynamic transposable elements contribute to genomic divergence under asexuality. *BMC Genomics* 17(884):1-12.
- Ferreira de Carvalho J, Oplaat C, Pappas N, Derks M, de Ridder D, Verhoeven KJF. 2016b. Heritable gene expression differences between apomictic clone members in *Taraxacum officinale*: Insights into early stages of evolutionary divergence in asexual plants. *BMC Genomics* 17(203):1-13.
- Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska E, Mansuy IM. 2014. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nature Neuroscience* 17(5):667-+.
- Gent JI, Ellis NA, Guo L, Harkess AE, Yao YY, Zhang XY, Dawe RK. 2013. CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. *Genome Research* 23(4):628-637.
- Heil M. 2002. Ecological costs of induced resistance. *Current Opinion in Plant Biology* 5(4):345-350.

- Hollister JD, Gaut BS. 2009. Epigenetic silencing of transposable elements: A trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Research* 19(8):1419-1428.
- Ito H, Gaubert H, Bucher E, Mirouze M, Vaillant I, Paszkowski J. 2011. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472(7341):115-U151.
- Li X, Zhu JD, Hu FY, Ge S, Ye MZ, Xiang H, Zhang GJ, Zheng XM, Zhang HY, Zhang SL et al. . 2012. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *Bmc Genomics* 13.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 133(3):523-536.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12).
- Lunardon A, Forestan C, Farinati S, Axtell MJ, Varotto S. 2016. Genome-Wide Characterization of Maize Small RNA Loci and Their Regulation in the required to maintain repression6-1 (rnr6-1) Mutant and Long-Term Abiotic Stresses. *Plant Physiology* 170(3):1535-1548.
- Mandal R, Kathiria P, Psychogios N, Bouatra S, Krishnumarthy R, Wishart D, Kovalchuk I. 2012. Progeny of tobacco mosaic virus-infected *Nicotiana tabacum* plants exhibit trans-generational changes in metabolic profiles. *Biocatalysis and Agricultural Biotechnology* 1(2):115-123.
- Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, Okamoto M, Nambara E, Nakajima M, Kawashima M et al. . 2008. Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant and Cell Physiology* 49(8):1135-1149.
- McCue AD, Nuthikattu S, Reeder SH, Slotkin RK. 2012. Gene Expression and Stress Response Mediated by the Epigenetic Regulation of a Transposable Element Small RNA. *Plos Genetics* 8(2).
- McCue AD, Nuthikattu S, Slotkin RK. 2013. Genome-wide identification of genes regulated in trans by transposable element small interfering RNAs. *Rna Biology* 10(8):1379-1395.
- Ou XF, Zhang YH, Xu CM, Lin XY, Zang Q, Zhuang TT, Jiang LL, von Wettstein D, Liu B. 2012. Transgenerational Inheritance of Modified DNA Methylation Patterns and Enhanced Tolerance Induced by Heavy Metal Stress in Rice (*Oryza sativa* L.). *Plos One* 7(9).
- Pecinka A, Scheid OM. 2012. Stress-Induced Chromatin Changes: A Critical View on Their Heritability. *Plant and Cell Physiology* 53(5):801-808.
- Quadrana L, Silveira AB, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V. 2016. The Arabidopsis thaliana mobilome and its impact at the species level. *Elife* 5.
- Rasmann S, De Vos M, Casteel CL, Tian DL, Halitschke R, Sun JY, Agrawal AA, Felton GW, Jander G. 2012. Herbivory in the Previous Generation Primes Plants for Enhanced Insect Resistance. *Plant Physiology* 158(2):854-863.

- Rechavi O, Hourri-Ze'evi L, Anava S, Goh WSS, Kerk SY, Hannon GJ, Hobert O. 2014. Starvation-Induced Transgenerational Inheritance of Small RNAs in *C. elegans*. *Cell* 158(2):277-287.
- Shao HB, Chu LY, Jaleel CA, Zhao CX. 2008. Water-deficit stress-induced anatomical changes in higher plants. *Comptes Rendus Biologies* 331(3):215-225.
- Slaughter A, Daniel X, Flors V, Luna E, Hohn B, Mauch-Mani B. 2012. Descendants of Primed Arabidopsis Plants Exhibit Resistance to Biotic Stress. *Plant Physiology* 158(2):835-843.
- Song YP, Ci D, Tian M, Zhang DQ. 2016. Stable methylation of a non-coding RNA gene regulates gene expression in response to abiotic stress in *Populus simonii*. *Journal of Experimental Botany* 67(5):1477-1492.
- Tran RK, Zilberman D, de Bustos C, Ditt RF, Henikoff JG, Lindroth AM, Delrow J, Boyle T, Kwong S, Bryson TD et al. . 2005. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in Arabidopsis. *Genome Biology* 6(11).
- Tricker PJ, Gibbins JG, Lopez CMR, Hadley P, Wilkinson MJ. 2012. Low relative humidity triggers RNA-directed de novo DNA methylation and suppression of genes controlling stomatal development. *Journal of Experimental Botany* 63(10):3799-3813.
- Verhoeven KJF, Jansen JJ, van Dijk PJ, Biere A. 2010. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist* 185(4):1108-1118.
- Verhoeven KJF, van Gorp TP. 2012. Transgenerational Effects of Stress Exposure on Offspring Phenotypes in Apomictic Dandelion. *Plos One* 7(6).
- Vicente MRS, Plasencia J. 2011. Salicylic acid beyond defence: its role in plant growth and development. *Journal of Experimental Botany* 62(10):3321-3338.
- Wang X, Weigel D, Smith LM. 2013. Transposon Variants and Their Effects on Gene Expression in Arabidopsis. *Plos Genetics* 9(2).
- Wang X, Xin CY, Cai J, Zhou Q, Dai TB, Cao WX, Jiang D. 2016. Heat Priming Induces Trans-generational Tolerance to High Temperature Stress in Wheat. *Frontiers in Plant Science* 7.
- Wibowo A, Becker C, Marconi G, Durr J, Price J, Hagemann J, Papareddy R, Putra H, Kageyama J, Becker J et al. . 2016. Hyperosmotic stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. *Elife* 5.

Figure legends

Figure 1. Length composition for the read libraries: all sRNAs (A and D), mapped to annotated TEs (B and E) and mapped to gene-annotated transcripts (C and F). The plots in the bottom (D, E and F) show the p-value for the bootstrap analysis performed for 21nt and 24 nt sRNA. Error bars, 95% bootstrap confidence intervals. ** p value = 10^{-4} . Treatment groups (C: control; D: drought; S: salicylic acid) refer to grandparental treatments.

Figure 2. Distribution for sRNA fold change in gene-mapping transcripts after grandparental drought stress (A) and salicylic acid (B) against control. Bar plots show GO terms enrichment compared to a bootstrap analysis in the case of the drought (C) and the salicylic acid (D) sets. Error bars, 95% bootstrap confidence intervals.

Figure 3. Spatial accumulation of 21 and 24 nt sRNA reads in gene-mapping transcripts. Lines represent density distributions of sRNA mapping location along the transcript. Each gene-mapping transcript was scaled to a length of 1000 bp and sRNA mapping positions (pooled replicates) inside each transcript were transformed accordingly. The counts for each transcript were afterwards collapsed into a single transcript model by calculating the averaged number of sRNA hits for each transcript coordinate across all length-normalized transcripts. Color code for treatment groups: control (green), drought (blue) and SA (red).

Figure 1

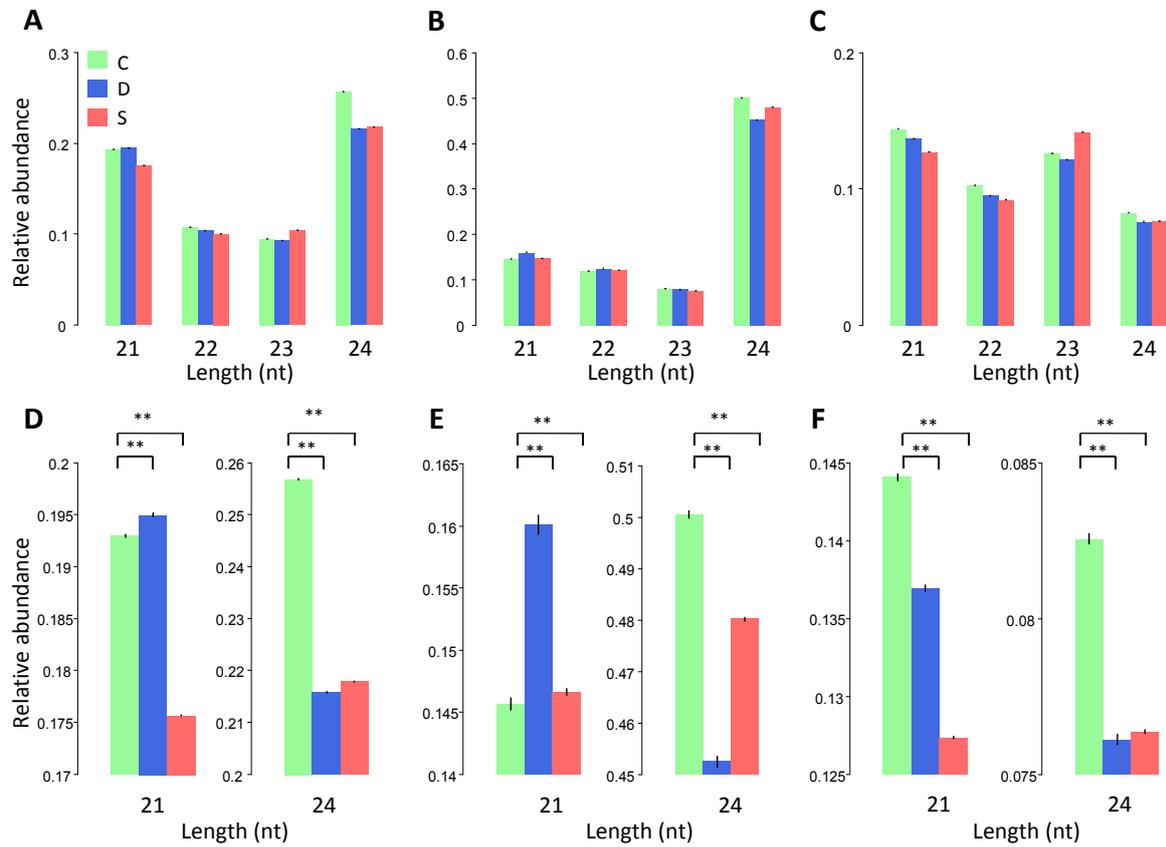


Figure 2

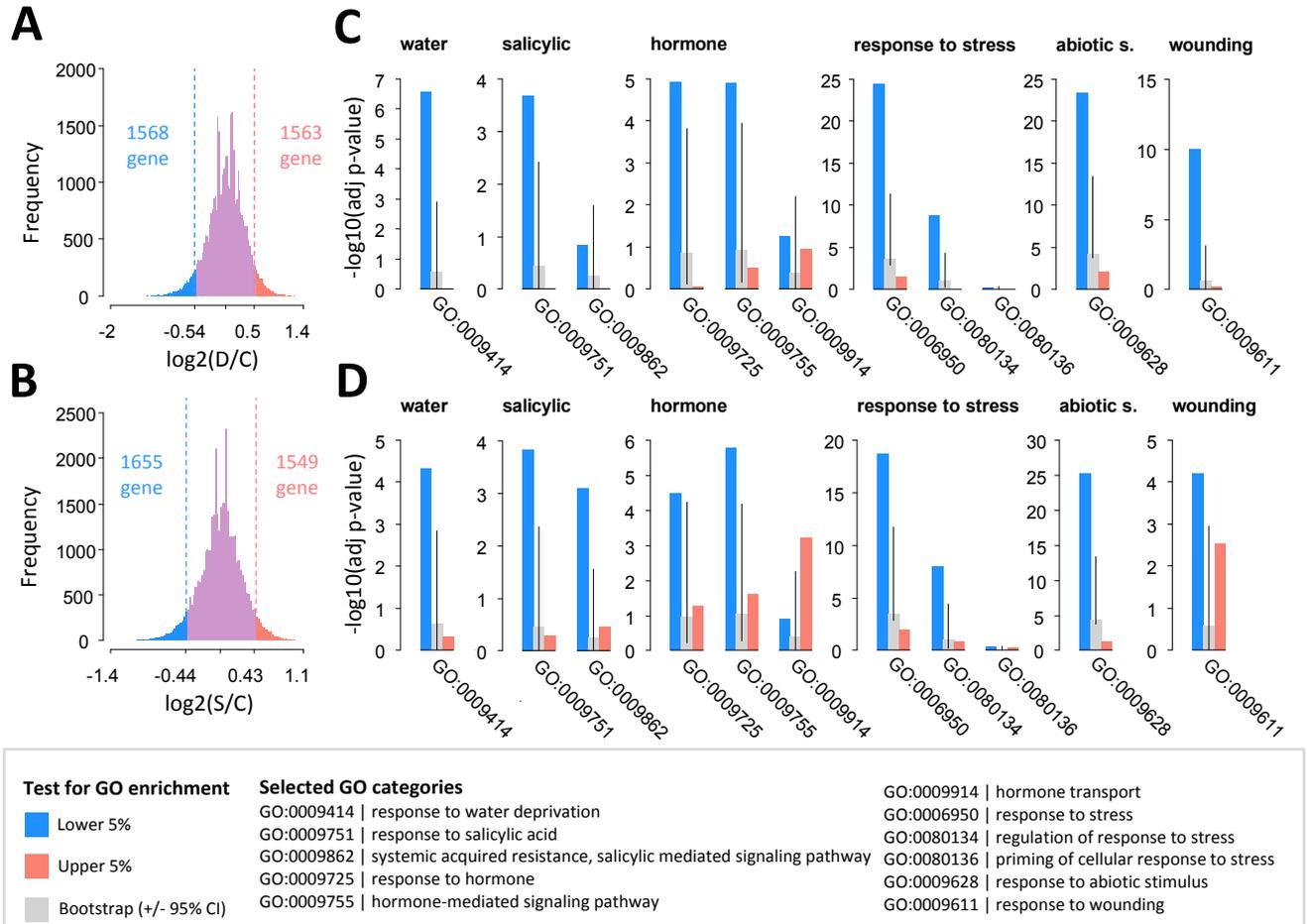
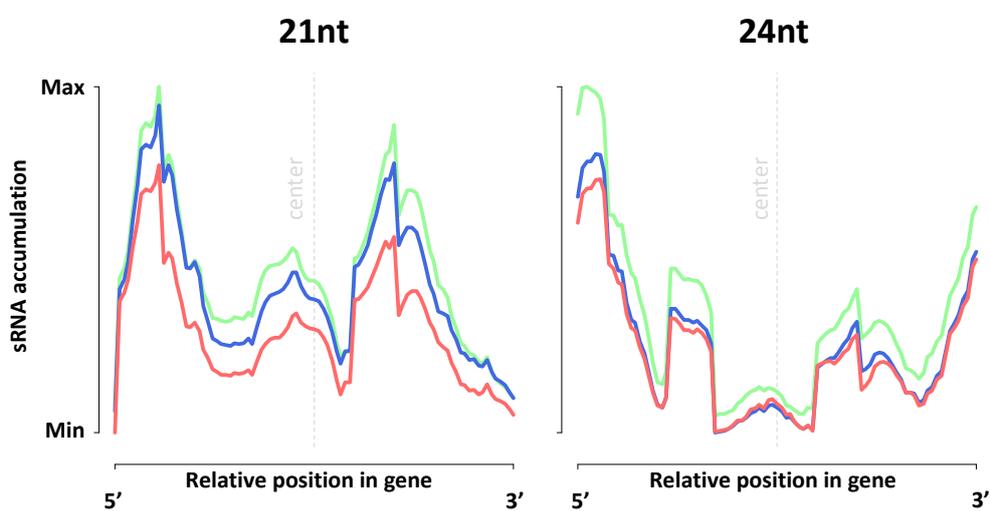


Figure 3



Supplementary text S1. Experimental methods

The experimental design is visualized in Figure S1. Apomictic dandelion plants from a single lineage (*T. officinale hemicyclum*) were used (Kirschner et al., 2016). G1 plants in the experiment derived through apomictic reproduction from a single G0 individual. Seedlings were distributed over three treatment groups: control, drought and salicylic acid. Plants within a group were randomized and placed in rows to ensure non-touching between the treatment groups. After four weeks of growth in a climate chamber (14 h light / 10 h dark, 18 °C / 14 °C, 60%) drought treatment started: during ten brief drought episodes over a period of four weeks, water was withheld until at least 80% of all individuals showed wilted leaves. At five weeks after seedling growth, SA treatment was applied once by spreading 0.5 ml of 10 mM SA solution over the surface of three leaves. The control group received no treatment. At eight weeks, plants were moved to a cold room for a 5-week vernalization period (continuous 4°C, 16 h light / 8 h dark), after which plants were placed on greenhouse benches and seeds were collected six weeks later. Using single-seed descent the subsequent two generations, G2 and G3, were grown in a common control environment under the same conditions as described for G1.

Leaf tissue was sampled from five weeks old G3 plants. Sixteen leaf discs of 8 mm in diameter were punched from one young and fully-developed leaf. Leaf discs were snap frozen in liquid nitrogen and stored in -80°C until usage. From liquid nitrogen-ground leaf tissue total RNA was extracted using 1 ml of Trizol (Ambion, Life technologies, The Netherlands) according to the manufacturer protocol but with an additional precipitation step with isopropanol and 3M sodium acetate (pH 5.2) and subsequent washing steps with ethanol. The final pellet was dissolved in 50 µl DNase/RNase-free water. Quality was checked on agarose gel electrophoresis and concentration on a NanoDrop 1000 spectrophotometer. For RNA sequencing library preparation the New England Biolabs NEBNext multiplex small RNA kite7300 kit (Ornat, Rechovot, Israel) was used with an initial 1 µg of RNA according to the manufacturer protocol, using barcoded primers and using T4 ligase truncated for 3' ligation. To select for sRNA with a length of 20-30nt, cuttings from a E-Gel EX 4% Agarose gel (Invitrogen, Life technologies, Israel) were taken between 140-150nt size fraction (accounting for two adapter sequences of 60nt each). Bands were cleaned using MiniElute Gel Extraction kit (QIAGEN, Eldan, Israel). After clean-up and a final Bioanalyzer quality check (Agilent Technologies, Eldan, Israel) using Agilent High Sensitivity DNA Kit, the 12 samples were pooled into a single sequencing library which was sequenced on two Illumina HiSeq2500 lanes (50bp, single-end).

References

- Kirschner, J., Oplaat, C., Verhoeven, K. J. F., Zeisek, V., Uhlemann, I., Travnicek, B., Rasanen, J., Wilschut, R. A. and Stepanek, J. (2016) Identification of oligoclonal agamosperous microspecies: taxonomic specialists versus microsatellites, *Preslia*, 88(1), pp. 1-17.

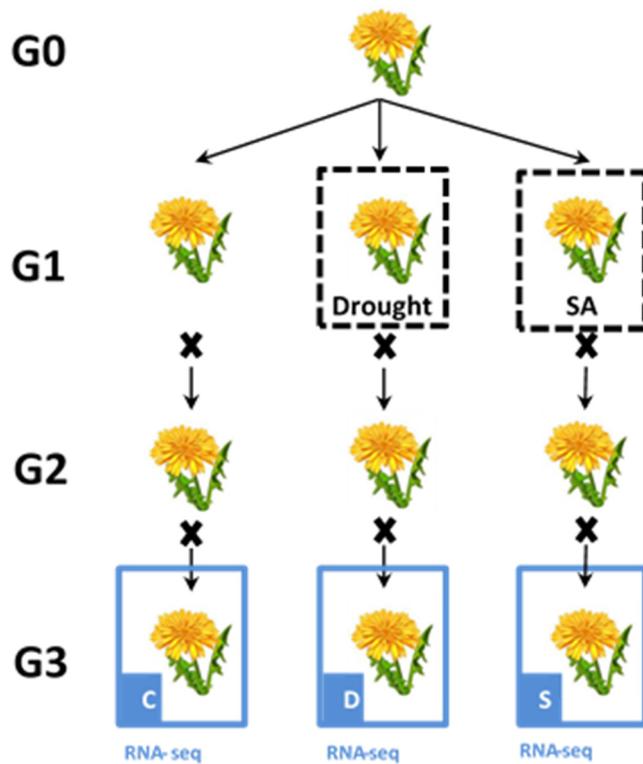


Figure S1. Apomictic Dandelion lines generated for the current study. Generations 1 to 3 are marked by G1, G2 and G3. Each plant cartoon accounts for 4 replicates, except for the single initial common ancestor individual (G0). Lines subjected to sRNA sequencing are indicated with a blue box. Treatment sets: control (C), drought (D); salicylic acid (S).

Supplementary text S2. Data processing and mapping

FastQC v0.11.3 software (www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used for preliminary quality check. Reads were processed using adapter trimming Cutadapt (Martin, 2011), filtered for quality (Phred score>33), and length (18nt to 30nt). Read counts per sRNA showed overall good correlations between replicate samples within experimental groups (Fig. S2). Reads were aligned with BWA (Li and Durbin, 2009; perfect matches) to:

1. A *Taraxacum officinale* reference transcriptome (Ferreira de Carvalho et al., 2016b). The transcriptome was assembled *de novo* from RNAseq data, which resulted in a total of 123,232 transcripts of which 39,685 transcripts were annotated to TAIR genes (BLASTn, e-value< 1e-05). The transcriptome is available in the Dryad repository under identifier number doi:10.5061/dryad.6p2n6.

2. A *Taraxacum officinale* database of transposable element sequences. This TE database was based on *de novo* clustering of genomic sequencing reads of the apomict *T. officinale* 'macranthoides accession 12' using RepeatExplorer, as presented in Ferreira de Carvalho et al. (2016a) and available in the Dryad repository under identifier number doi:10.5061/dryad.f2j00. From the list of clusters assembled by RepeatExplorer, the most abundant clusters (i.e. containing at least 0.20% of the inputs reads) were selected. Thus, a total of 120 clusters were used for sRNA mapping and represent highly repetitive annotated TE sequences present in the genome of *T. officinale*.

References

- Ferreira de Carvalho J, de Jager VCL, Van Gurp TP, Wagenmaker N, Verhoeven KJF. 2016a. Recent and dynamic transposable elements contribute to genomic divergence under asexuality. *BMC Genomics* 17(884):1-12.
- Ferreira de Carvalho J, Oplaat C, Pappas N, Derks M, de Ridder D, Verhoeven KJF. 2016b. Heritable gene expression differences between apomictic clone members in *Taraxacum officinale*: Insights into early stages of evolutionary divergence in asexual plants. *BMC Genomics* 17(203):1-13.
- Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754-1760.
- Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), pp. 10-12.

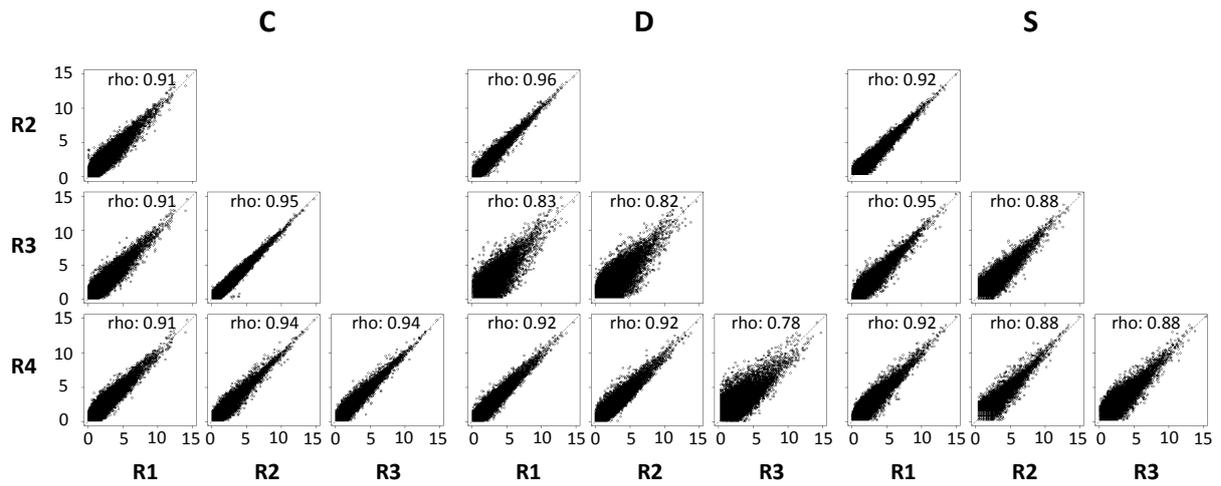


Figure S2. Correlation between sRNA read counts among replicate libraries (R1, R2, R3 and R4) in each treatment group (C, D and S). Axes in each plot represent the log₂(RPM+1). Rho: Pearson correlation coefficient based on read count of sRNAs that are shared between the two libraries.

Supplementary text S3. Bootstrap procedure

The bootstrap procedure to determine significance of global shifts in sRNA composition among treatment sets compared a pooled control with a pooled treatment sample (drought or salicylic acid); in both groups the 4 replicate plants were pooled in silico into a single sample. Generically, the bootstrap procedure for hypothesis testing implemented considers two observed independent samples:

$$x = (x_{obs,1}, x_{obs,2}, \dots, x_{obs,n})$$

and

$$y = (y_{obs,1}, y_{obs,2}, \dots, y_{obs,m});$$

with n and m the respective sizes, and for which no assumptions of normality must be made. From these observed samples the mean difference

$$t_{obs}^* = \bar{x}_{obs}^* - \bar{y}_{obs}^*$$

can be calculated, for which the null hypothesis

H_0 : both samples come from the same population is tested against the alternative

H_1 : both samples are NOT from the same population.

The method encompasses 5 steps:

1. Merging of the two observed samples into one sample of $(n + m)$ observations
2. A bootstrap of $(n + m)$ observations with replacement is drawn from the merged sample
3. The mean \bar{x}^* for the first n observations is calculated and then the mean \bar{y}^* for the remaining m observations, afterward which the test statistic $t^* = \bar{x}^* - \bar{y}^*$ is calculated
4. Steps 2 and 3 must be repeats B (e.g. 4000) times to obtain B values of the test statistic
5. A p-value is then estimated according to:

$$p - value = \frac{\text{number of times}(t^* > t_{obs}^*)}{B}$$

If $p - value < \alpha$, being α the significance level (e.g.: 0.05) then H_0 is rejected, and retained otherwise.

Supplementary text S4. DESeq2 results

Table S4. Per-transcript analysis of differential sRNA abundance. DESeq2 results for: all consensus sRNA (sRNA shared by all replicates from at least one of the treatments), TE-mapping sRNA and gene-mapping sRNA. Up- or downregulated sRNAs are significant after multiple testing correction in DESeq2 at FDR = 0.1 threshold.

Group	Comparison	sRNA length (nt)	#instances tested	#Up-regulated	#Down-regulated
Consensus set	D/C	All	569268	3	0
		21	102759	2	0
		24	235625	0	0
	S/C	All	545459	15	12
		21	99072	5	8
		24	227383	0	6
TEs	D/C	All	98	0	0
		21	98	0	0
		24	98	0	0
	S/C	All	98	0	0
		21	98	0	0
		24	98	0	0
Gene-mapping transcripts	D/C	All	35880	0	5
		21	27153	0	0
		24	27730	0	1
	S/C	All	36202	0	0
		21	27975	0	0
		24	28683	0	0

Supplementary text S5. Gene Ontology analysis

Gene Ontology (GO) analysis was performed with the R package topGO (Alexa and Rahnenfuhrer, 2004) and annotations from the Arabidopsis genome (Carlson *GO.db* v.3.1.2). To control for potential biases due to the use of Arabidopsis as a reference in the enrichment analysis, a baseline enrichment level was determined via random subsampling from the dandelion transcripts. This subsampling was repeated 1000 times picking a number of random transcripts equal to the top 5% of the respective treatment-control set (control versus drought and control versus SA). We considered the enrichment to be significant if the FDR-adjusted p-values from the topGO enrichment analyses exceeded the 95% confidence interval of the bootstrapped subsamples.

References

Alexa, A. and Rahnenfuhrer, J. 2004. topGO:Enrichment analysis for gene ontology.

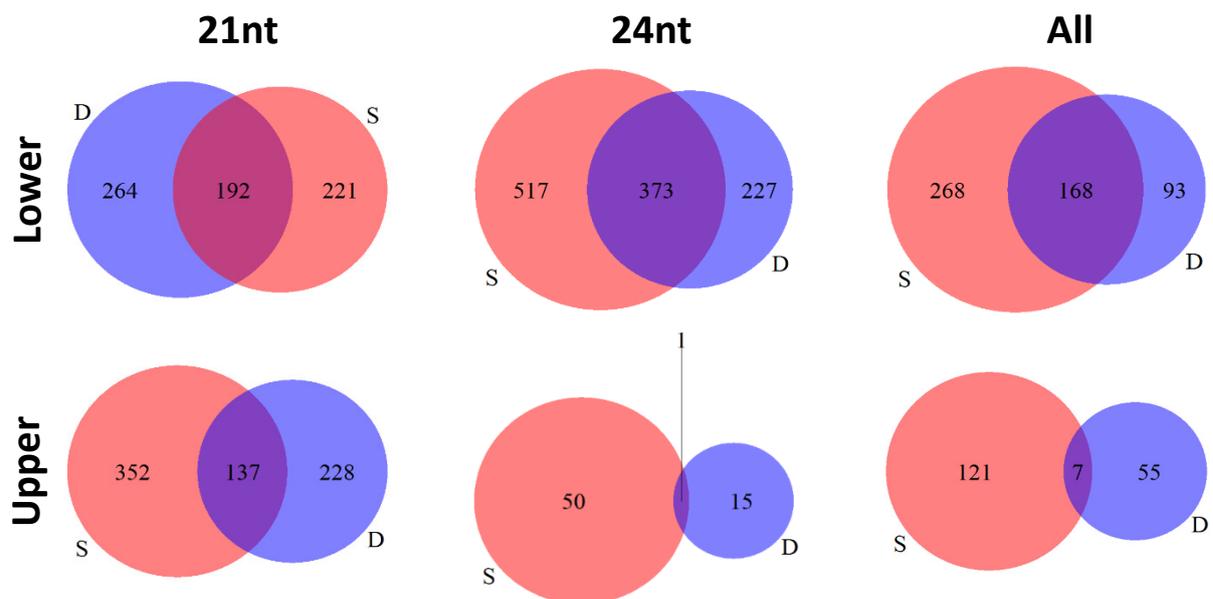


Figure S5. Overlap between GO terms that are significantly enriched in the list of most downregulated genes (lower: 5% of transcript with strongest reduction in mapped sRNAs due to grandparental stress compared to control treatment) and in the list of most upregulated genes (upper: 5% of transcripts with strongest increase in mapped sRNA due to grandparental stress treatment). Numbers for significantly enriched GO terms are shown as Venn diagrams for effects of grandparental drought and grandparental SA treatment for 21nt and 24nt sRNAs and for all length classes between 18 and 30nt.

Table S5. Number of genes for each of the 11 terms selected in the GO analysis.

GO id	Term	#Annotated genes	#Significant genes			
			CvsD		CvsS	
			Lower 5%	Upper 5%	Lower 5%	Upper 5%
GO:0009414	response to water deprivation	419	57	28	52	31
GO:0009751	response to salicylic acid	471	53	25	55	34
GO:0009862	systemic acquired resistance, salicylic ...	251	24	13	33	21
GO:0009725	response to hormone	1601	145	100	148	120
GO:0009755	hormone-mediated signaling pathway	773	82	58	88	66
GO:0009914	hormone transport	107	14	13	13	19
GO:0006950	response to stress	4030	406	280	398	286
GO:0080134	regulation of response to stress	569	77	34	77	46
GO:0080136	priming of cellular response to stress	2	1	0	1	1
GO:0009628	response to abiotic stimulus	2796	307	208	326	197
GO:0009611	response to wounding	335	57	25	44	38