Heat-induced transposition increases drought tolerance in *Arabidopsis*

Michael Thieme[1]*, Arthur Brêchet[2], Bettina Keller[1], Etienne Bucher[3] and Anne C. Roulin[1]

[1] Department of Plant and Microbial Biology, University of Zurich, Zollikerstrasse 107, 8008 Zürich, Switzerland

[2] Department of Environmental Sciences - Botany, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

[3] Crop Genome Dynamics Group, Agroscope, 1260 Nyon, Switzerland

* author for correspondence: michael.thieme@botinst.uzh.ch

Abstract

Eukaryotic genomes contain a vast diversity of transposable elements (TEs) [1]. Formerly often described as selfish and parasitic DNA sequences, TEs are now recognized as a source of genetic diversity and powerful drivers of evolution [2]. Yet, because their mobility is tightly controlled by the host, studies experimentally assessing how fast TEs may mediate the emergence of adaptive traits are scare. Here, we show that the heat-induced transposition of a low-copy TE increases phenotypic diversity and leads to the emergence of drought-tolerant individuals in *Arabidopsis thaliana*. We exposed high-copy TE lines (hcLines) [3] with up to ~8 fold increased copy numbers of the heat-responsive *ONSEN* TE (*AtCOPIA78*) [4–6] to drought as a straightforward and ecologically highly relevant selection pressure [7]. We provide evidence for increased drought tolerance in five out of the 23 tested hcLines and further pinpoint one of the causative mutations to an exonic insertion of *ONSEN* in the *ribose-5-phosphate-isomerase 2* gene. The resulting loss-of-function mutation caused a decreased rate of photosynthesis and water consumption. This is one of the rare empirical examples [8] substantiating the adaptive potential of mobilized stress-responsive TEs in eukaryotes. Our work sheds light on the relationship between TEs and their hosts and demonstrates the importance of TE-mediated loss-of-function mutations in stress adaptation, particularly in the face of global warming.

Keywords: *Arabidopsis thaliana*, *ONSEN,* transposable element, adaptation, evolution, loss-of-function mutation, drought stress, climate change

Main

Plants are constantly exposed to fluctuating environments. To successfully reproduce, they rely on mechanisms that allow them to react and adapt to suboptimal growth conditions. Genetic variation, whether as a result of natural processes or artificially induced, is a prerequisite for adaptation and the evolution of new traits. There is evidence that severe stresses can not only trigger the formation of small-scale mutations [9] but also increase genetic diversity through the stress-induced activation of transposable elements (TEs) [10]. Mobile, highly mutagenic [11,12] and especially abundant in eukaryotic genomes [1], TEs are believed to facilitate rapid adaptation to challenging environments [2,13,14]. Yet, to ensure a limited mutation rate and to safeguard genome stability, TE mobility is usually restricted by epigenetic silencing mechanisms, which in plants involves the RNA-directed DNA methylation (RdDM) pathway [15]. As a consequence, only few TE families have been observed transposing *in planta* and therefore the immediate evolutionary consequences of stress-induced TE bursts are largely unknown. By combining a transient chemical inhibition of TE silencing with a heat-shock, we have previously generated genetically stable high-copy TE lines (hcLines) [3] carrying novel insertions of the heat-responsive retrotransposon *ONSEN* (*AtCOPIA78*) [4–6] in *Arabidopsis thaliana*. Due to climate change and global temperature increase, drought is predicted to constitute one of the most severe environmental constraints to which plants will have to adapt in the near future [7,16]. In this context, we tested here whether the heat-induced transposition of *ONSEN* may help individuals to survive in warmer and hence water-limited environments.

Results and discussion

We grew the S4 generation of 23 hcLines originating from 13 independent heat stressed and transiently de-methylated plants [3] and validated by quantitative PCR (qPCR) an up to ~8 fold copy number increase (thus up to ~64 stably inserted *ONSEN* copies; Fig 1a) compared to Col-0 wild type (wt). However, in some hcLines (e.g. hcLine9), no additional *ONSEN* copies were detectable by qPCR even though some were previously validated by sequencing [17]. As the hcLines were originally exposed to a combination of heat and the drugs zebularine (Z) and alpha-amanitin (A) (HS+AZ) [3], we controlled for a potential *ONSEN*-independent phenotypic variation caused by epi/genetic changes induced by the heat-stress or the chemical demethylation. To do so, we included a Col-0 wt plant that was propagated on soil and five independent controls, i.e. lines that originated from plants that were also grown *in vitro* but only exposed to control-stress (CS; two lines), heat stress (HS; two lines) or CS plus chemical demethylation (CS+AZ; one line) [3]. In accordance with previous observations [3,17], we did not detect an increase in *ONSEN* copy numbers in these control lines when compared to the wt (Fig 1a).

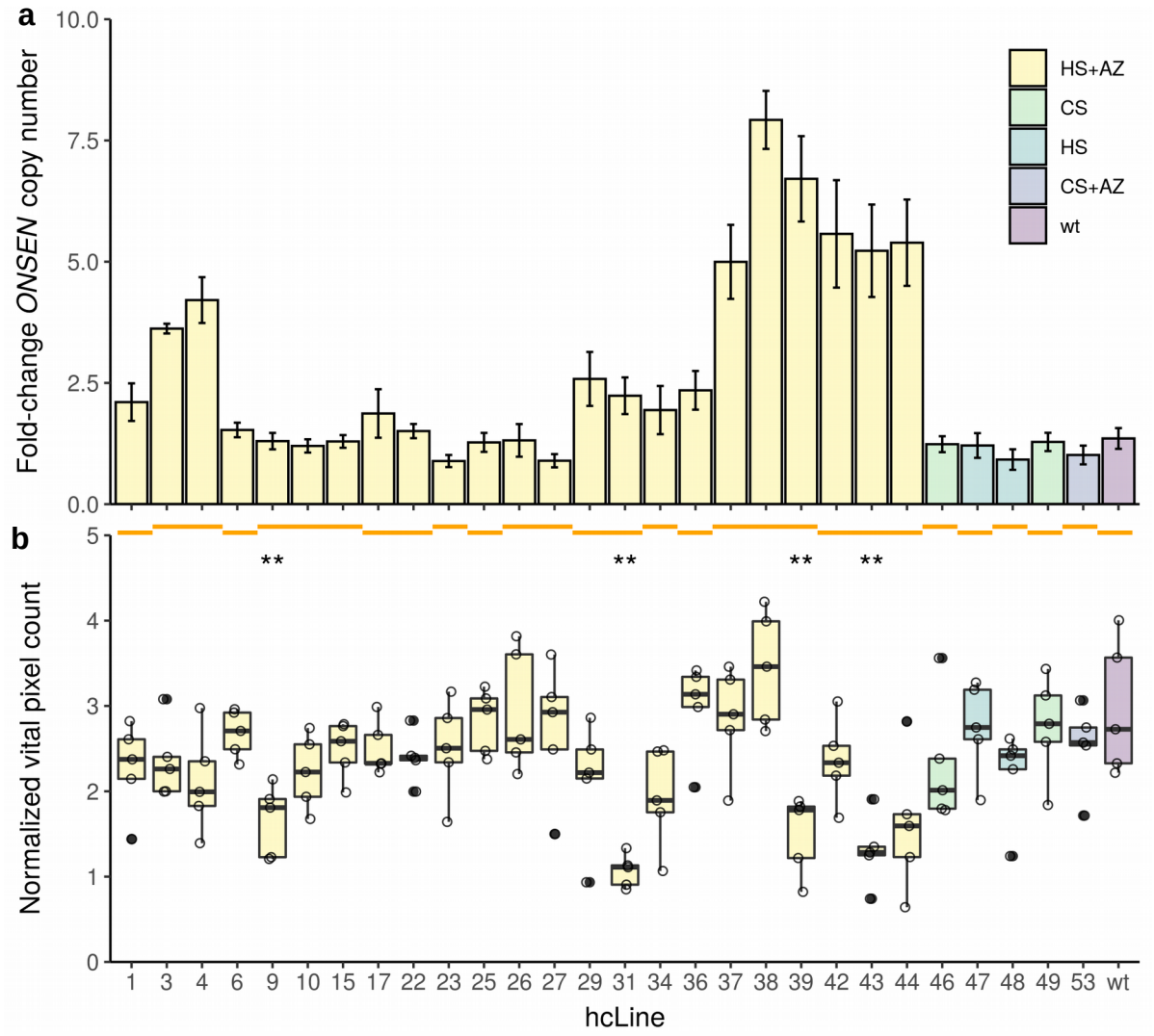

Fig. 1 ***ONSEN* copy numbers and size of hcLines.** Box colors indicate the history of the lines: control-stress (CS), heat-stress (HS), chemical de-methylation (AZ) and wild type propagated on soil (wt). Orange lines spanning multiple hcLines indicate their origin from a common parent. **a** Fold-change of *ONSEN* copy numbers measured by qPCR compared to the wt (n=3 technical replicates +/- SD). **b** Vital pixel count of hcLines and controls before the occurrence of necrotic leaves (day 8). Significant differences to the wt are indicated. Horizontal line defines median, hinges represent 25th and 75th percentiles, whiskers extend to 1.5 * IQR and outliers are shown as filled dots. n= 5 biological replicates. Wilcoxon-test P< 0.01 (**).

We suspended watering after 36 days of growth and recorded plant development and water loss of the pots every week by taking top view pictures and by weighting the pots. To quantify the growth and the degree of drought-induced leaf senescence, we trained the image-based interactive learning and segmentation toolkit (ilastik) [18] to specifically detect living (hereafter vital) and necrotic leaf segments. We first tested the reliability of the prediction by placing one to three punched leaf discs of necrotic or vital segments onto a pot that did not contain a plant. After processing the images with ilastik, we obtained a linear increase of vital and necrotic pixel-counts according to the number of segments placed onto the pots (Fig S1), confirming the reliability of the method. For the rest of the study, we therefore use vital pixel counts as a proxy for plant size. To assess size variations between the hcLines, we analyzed the pixel counts of predicted vital areas before the appearance of necrotic leaves eight days after watering was suspended (Fig 1b, Fig 2a). Although heat stress and the AZ drug treatment have been shown to induce epi/genetic mutations [9,17,19], we did not observe differences in growth among our five control lines. Yet, in accordance with the fact that transposition is predominantly associated with fitness loss of the host [17,20–22], we found a significantly reduced number of vital pixels compared to the wt for four hcLines (Wilcoxon-test $P<0.01$, Fig 1b). Notably, in some cases we also observed a strong size variation between hcLines originating from the same parent (e.g., hcLine37 and hcLine38 vs hcLine39) indicating genetic segregation of the lines (Fig 1b). We did not find any significant global correlation between *ONSEN* copy numbers and plant size ($P>0.05$, Fig S2). Taken together, these observations suggest that single TE insertions rather than the overall *ONSEN* load were responsible for the observed phenotypic variation.

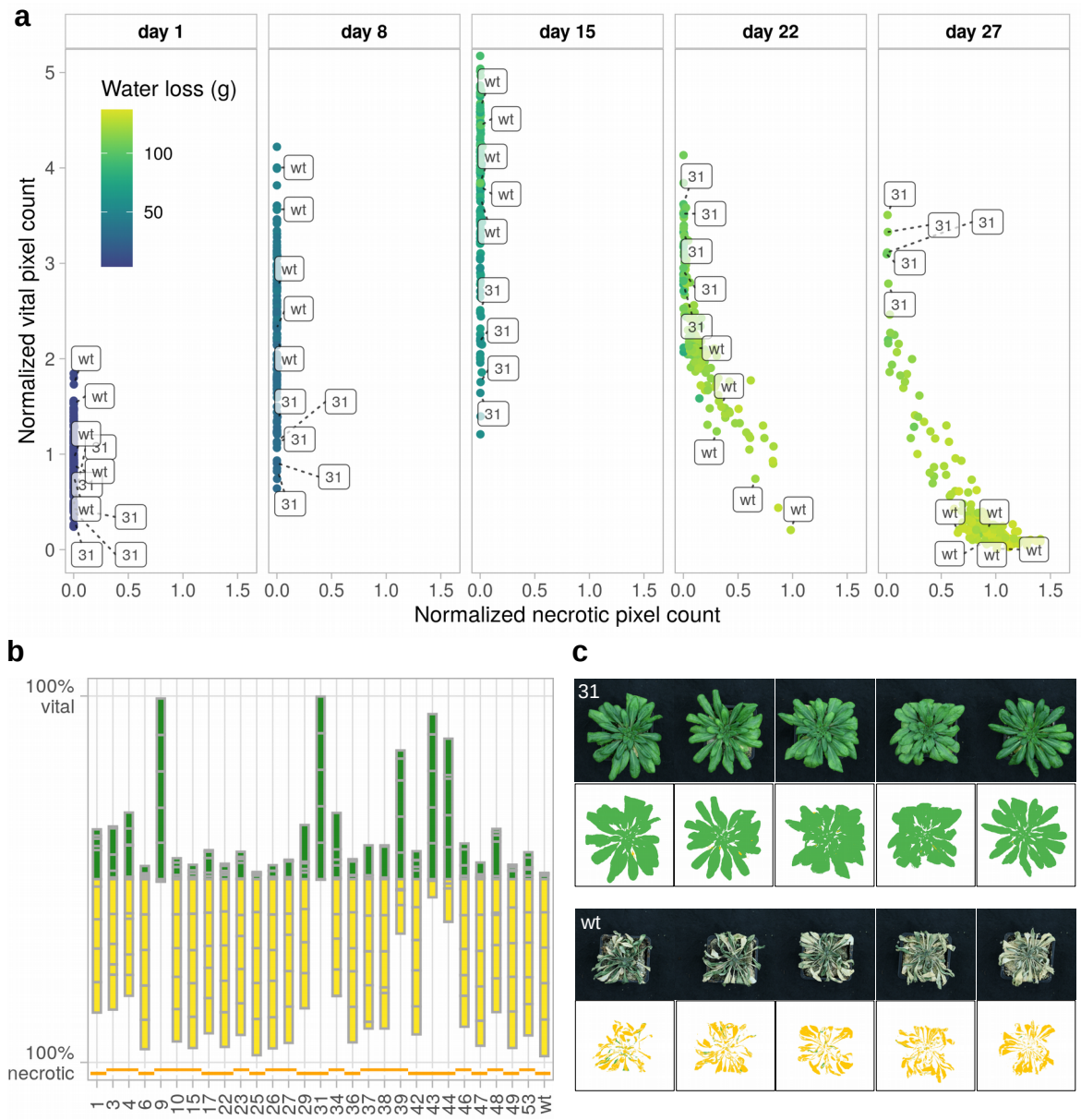

Fig. 2 **Drought tolerance of *ONSEN* hcLines. a** Pixel counts of living and necrotic tissues during the drought stress. Each dot represents one individual plant and the color code indicates the cumulative water loss over time. n=3-5 biological replicates per line are shown. **b** Percentage of vital (green) and necrotic (yellow) tissues of all five replicates per line two days after recovery (day 29). The contribution of each replicate to the total amount of vital and necrotic pixels is indicated with gray bars. Orange lines spanning multiple hcLines indicate their origin from a common parent. **c** Original (top) and processed (ilastik, bottom) images of the five replicates of hcLine31 (31) (top) and wt (bottom) on day 29. Predicted vital leaf segments are depicted in green, necrotic segments in yellow.

The machine learning approach allowed us to further track the dynamics of growth and drought-induced necrosis of the individual lines. None of the plants had necrotic leaves until 15 days after watering was suspended, (Fig 2a). After 27 days, we resumed watering and recorded leaf areas after two days of recovery (day 29). In contrast to the high drought-induced mortality of the five control lines and the wt, five out of the 23 hcLines originating from four independent parental lines were more stress tolerant (mean vital leaf area > 50%) (Fig 2b). While the wt was most susceptible to drought (0.03 % mean vital area), hcLine31 did not show any signs of necrosis (97.9 % mean vital area, Fig 2a, b and c). The consistent vitality of all five replicates of hcLine31 indicated that a homozygous mutation was underlying the observed drought tolerance of this line. Because hcLine31 displayed the most drought-tolerant phenotype, we selected it to characterize the functional link between the heat-induced insertion of novel *ONSEN* copies and the observed increase in drought tolerance.

Table 1 **Novel *ONSEN* insertions in hcLine31**. Location, description (Araport11) and zygosity (-/- homozygous *ONSEN*, +/- heterozygous) of predicted *ONSEN* insertion sites adapted from [17].

chr	***coordinates***	**context**	***ID***	***description***	***zygosity***
1	21761950– 55	exon	*AT1G58602*	LRR and NB-ARC domains-containing disease resistance protein	(-/-)
2	149387- 91	exon	*AT2G01290*	cytosolic ribose-5-phosphate isomerase	(-/-)
3	19300993– 97	exon	*AT3G52020*	serine carboxypeptidase-like 39	(+/-)
3	22631755- 59	exon	*AT3G61150*	homeodomain GLABROUS 1; HD-ZIP IV family.	(+/-)
5	853776- 80	exon	*AT5G03435*	Ca2+dependent plant phosphoribosyltransferase family protein	(-/-)
5	10632816– 20	TE	*AT5G28626*	AT5TE38720; SADHU; Sadhu non-coding retroTE family	(-/-)
5	18850327- 31	promoter	*AT5G46490*	disease resistance protein (TIR-NBS-LRR class) family	(-/-)
5	21050239- 43	intron	*AT5G51800*	protein kinase superfamily protein	(-/-)
5	21602030- 34	exon	*AT5G53240*	hypothetical protein (DUF295)	(+/-)
5	22846432- 36	intron	*AT5G56400*	FBD, F-box, Skp2-like and Leucine Rich Repeat domains containing protein	(+/-)

We first used whole-genome re-sequencing data to locate all transposon-insertion polymorphisms (TIPs) of hcLine31 [17]. In accordance with its insertion bias towards actively transcribed regions in the *A. thaliana* genome [17,23], six out of the ten *ONSEN* TIPs detected in hcLine31, were located in exons (Table 1). Notably, we detected a homozygous TIP in the

ribose-5-phosphate-isomerase 2 (*RPI2, At2G01290*), a gene involved in chloroplast photosynthetic capacity [24]. Analysis of previously published RNA-seq data of hcLine31 [17] revealed a premature transcriptional stop coinciding with the detected exonic *ONSEN*-insertion in *RPI2* (Fig 3a).

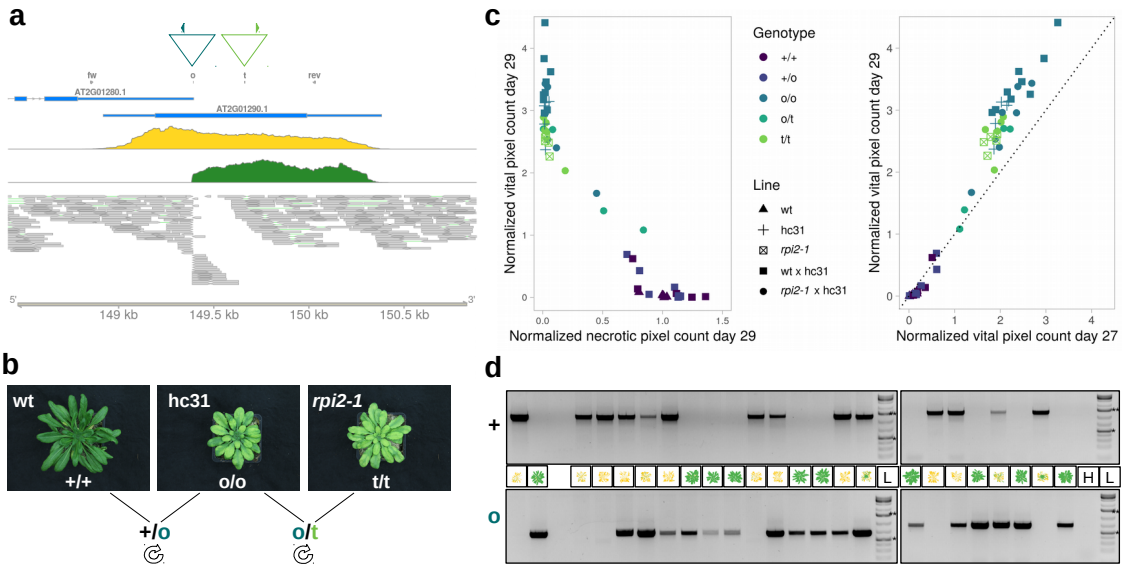

Fig. 3 **An *ONSEN*-insertion in *RPI2* leads to an increased drought tolerance of hcLine31 (hc31). a** Triangles indicate the insertion site of *ONSEN* (turquoise) and the location of the T-DNA in the *rpi2-1* mutant (green) on chromosome 2. Primer locations used for the genotyping in **c** and **d** are depicted as filled triangles. Annotation track (blue), RNA-seq coverage of the wt (yellow) and hcLine31 (green) and aligned genomic reads from hcLine31 are shown. **b** Images and crossing scheme of the wt, hcLine31, and *rpi2-1*. Genotypes are depicted with + (wt), o (*ONSEN*), and t (T-DNA). **c** Pixel counts of living tissue after two days of recovery (day 29) in relation to necrotic tissues on the same day (left panel) and compared to vital pixel counts before recovery (day 27) (right panel). Shapes indicate the plant line (parental or segregating F2 individuals of the crosses from panel b and colors indicate the genotype of *RPI2*. **d** Geno- and phenotypes (day 29) of the segregating F2 population of hcLine31 x wt. A wt and a hcLine31, (first two lanes) are shown as references. Primers specific for the wt (top gel, +) and the *ONSEN* insertion (bottom gel, o) (see **a**) were used. H=water control, L=GeneRuler 1 kb Plus ladder. 0.5 kb (*) and 1.5 kb (**) bands are marked.

To test whether the *ONSEN* insertion in *RPI2* was the causative mutation for the increased drought tolerance of hcLine31, we used the mutant *rpi2-1* (SALK_022117) as it harbors a homozygous T-DNA insertion in the exon of *RPI2* which also leads to a knock-out of the gene [24]. We crossed hcLine31 to the wt and to the *rpi2-1*-mutant (Fig 3b) and assessed the drought response in the segregating F2 generations obtained from self-fertilization. All F2 individuals that were either homozygous for the *ONSEN* or the T-DNA insertion or that carried both the *ONSEN* and the T-DNA-insertions in *RPI2* survived the drought stress and showed continued growth two days after recovery. In contrast, plants that carried at least one wt allele showed a high degree of drought-induced necrosis and no growth increase two days after recovery from water-limitation (Fig 3c,d). These results confirmed that the *ONSEN* insertion in *RPI2* results in a recessive loss-of-function mutation causing the increased drought tolerance of hcLine31.

A previous study indicated that the loss of *RPI2* leads to chloroplast dysfunction and reduced chlorophyll content [24]. Therefore, we assessed carbon/nitrogen (C/N)-ratios and Soil Plant Analysis Development (SPAD) values, which are directly linked to photosynthetic activity [25,26], following growth under well-watered conditions and before the emergence of necrotic leaves. SPAD and C/N ratios were significantly reduced (Wilcoxon-test, $P<0.05$) in hcLine31 compared to the wt (Fig 4a). As noted earlier, the onset of water-limitation symptoms in hcLine31 occurred significantly later than that in the wt (Fig 2a), therefore we looked for a general link between plant size and drought tolerance. We first fitted a linear model where water loss (before the occurrence of necrotic leaves at day 8 of the experiment) was entered as the response variable and the plant lines, vital pixel count and their interaction as explanatory variables. While the overall model was significant and explained a large part of the variance (R^2 =0.72, $P < 2.2e-16$), we did not detect a significant contribution of the plant line nor of the interaction between the line and vital pixel count, indicating that all hcLines, control lines and the wt had a similar efficiency regarding water use. However, we found that vital pixel count had a significant effect on water loss ($P = 5.19e-11$). This was further confirmed by a reduced model where the line and interaction effects were dropped

(Fig 4b, R^2 0.71, P<0.001). These experiments suggested that the *ONSEN* insertion in *RPI2* resulted in reduced photosynthetic capacity leading to slower growth and a reduced water consumption, allowing hcLine31 to escape severe drought stress. Accordingly, we also found that *RPI2* alleles in natural *A. thaliana* ecotypes were quantitatively associated with aridity levels in their local habitats (4.3% of the variance explained; Fig S3) and concluded that our candidate *ONSEN* insertion in hcLine31 boosts the native function of *RPI2* in respect of adaptation to drought.

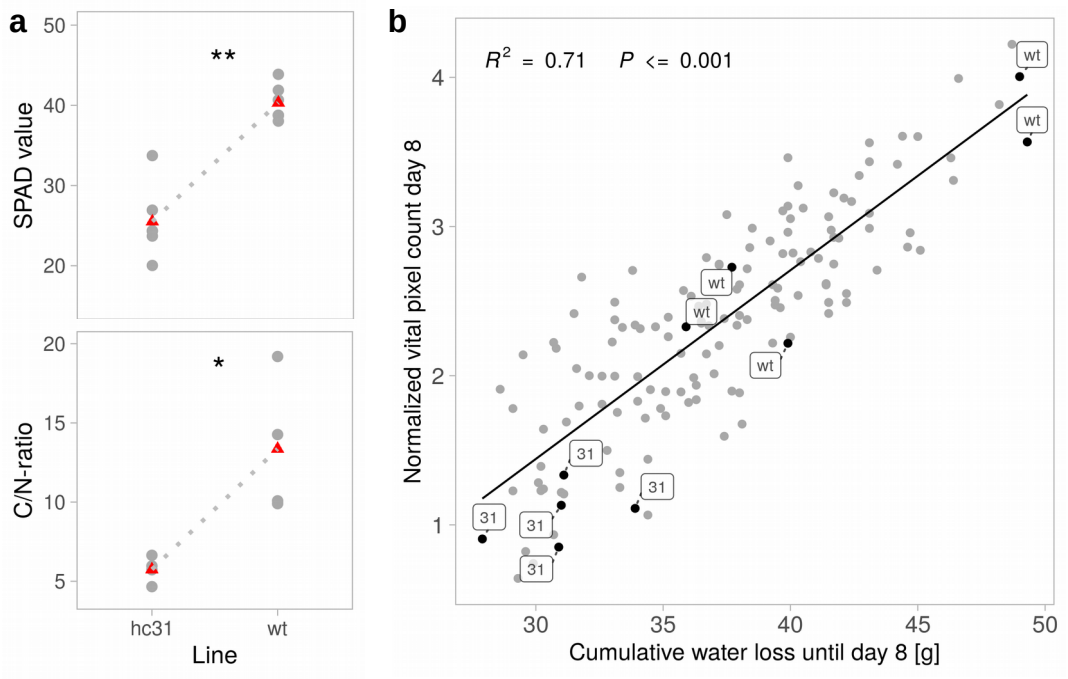

Fig. 4 **Physiological differences between hcLine31 and wt plants. a** SPAD meter values (top) and C/N-ratio (bottom) of the two lines. n= 4-6 biological replicates. Means are marked with red triangles, Wilcoxon-test P<0.05 (*), 0.01 (**). **b** Relationship between pixel counts of living tissues and cumulative water loss until day 8 before the occurrence of necrotic leaves. Linear regression model, adjusted R-squared: 0.71. P<0.001.

Such an insertion of *ONSEN* or other heat-induced TEs could thus provide a selective advantage in the face of global warming [2,23,27,28]. Yet, because the knock-out of *RPI2* leads to a reduced photosynthetic activity, one may expect large-scale mutations providing such a growth penalty under control conditions to be purged by natural selection. In fact, we do did find any natural TE insertions in *RPI2* [2]. This could suggest that such distinct, TE-mediated adaptive effects might be transient and thus difficult to capture with population genomics data. On the other hand, positively selected loss-of-function mutations leading to semi-dwarfness have been reported earlier in *A. thaliana* [29] and the absence of natural TE insertions in *RPI2* could also indicate that this transposition event did not take place in the wild or only occurred in marginal populations not yet sampled. Indeed, TE-mediated mechanisms leading to large-effect mutations might be especially important in less-adapted, frequently stressed populations where drastically altered phenotypes could be advantageous. In agreement with this hypothesis, Baduel et al. (2021) recently pointed towards a link between positive selection of weak alleles of the largest subunit of RNA polymerase V, a key component of RNA-directed DNA methylation, and a globally relaxed silencing of TEs in *A. thaliana* ecotypes growing under extreme conditions.

Conclusion

In the context of climate change, assessing how fast plants will adapt to longer periods of severe heat and drought stress is crucial both for conservation biology and food security [7,16,30]. Selection from standing variation, as opposed to the emergence of new mutations [31], is classically expected to lead to faster evolution [32]. TEs however, may challenge this prediction due to the combined effect of their stress-inducible activity and large mutagenic properties [2]. Here, we experimentally confirmed with a real-time setup that a single novel insertion of *ONSEN* can indeed rapidly lead to drought tolerance. In *A. thaliana,* several studies have established a functional link between loss-of-function alleles and adaptive traits (e.g. [33–35]) also in the context of adaptation to drought [36,37]. Indeed, variations in

the growth scaling of natural accessions of *A. thaliana* have been shown to be linked to abiotic parameters such as temperature and precipitation [38]. Together with this large body of evidence, our study substantiates the “less is more” hypothesis [39] i.e., that in contrast to intuitive expectations, gene loss may fuel evolution. Whether the observed high frequency of five independent drought-tolerant hcLines could be explained by the insertion preference of *ONSEN* towards actively transcribed H2A.Z-rich regions [17,23], remains to be elucidated. Selection experiments with populations of hcLines whose TE-composition has not yet been shaped by natural selection will further allow us to extrapolate the overall gain or loss of plant fitness following a heat-induced transposition. In conclusion, our study demonstrates that stress-induced TE mobility can lead to an increased stress tolerance of the host. These findings also confirm that TEs, by rapidly modulating gene expression and traits, may provide a powerful mean for plants to keep pace with rapidly changing environmental conditions.

Methods

4.1 Plant material

ONSEN hcLines were generated by treating *Arabidopsis thaliana* Col-0 plants with a combination of a heat-shock and drugs that inhibit TE-silencing, as described previously [3]. Briefly, Col-0 seeds were germinated and grown under long-day conditions (16h light) at 24 °C (day) 22 °C (night) on ½ MS medium with 1% sucrose and 0.5% Phytagel, pH 5.8. To reduce TE silencing and increase the rate of *ONSEN* transposition, seedlings were grown analogously on ½ MS medium supplied with a combination sterile filtered zebularine (Z, 40 µM) and α-amanitin (A, 5 mg/ml). After seven days of growth on control ½ MS or medium supplied with A and Z, seedlings were ether exposed to control stress (CS, 24 h at 6 °C followed by 24 h at normal conditions) or heat stress (HS, 24 h at 6°C followed by 24h at 37 °C), then transferred to soil and selfed to obtain the S1 generation. Individual S1 plants originating from plants that were either only exposed to CS or HS or additionally treated with A and Z (AZ) were separated and repeatedly self-fertilized until we obtained the S4 generation. In this study we used 23 *ONSEN* hc-lines originating from 13 plants that were treated with HS+AZ, five independent control lines that were either only exposed to CS (two lines), HS (two lines) or CS+AZ (one line) and the Col-0 wild type that was propagated on soil. The *rpi2-1* mutant (SALK_022117) [24] was obtained from the Nottingham Arabidopsis Stock Centre [40].

4.2 qPCR for *ONSEN* copy numbers

To determine the average *ONSEN* copy numbers of the hcLines and controls used in this study, we extracted DNA of the aboveground parts of at least 24 pooled individuals per line of the S4 generation grown for eight days under sterile conditions on ½ MS medium (1% sucrose, 0.5% Phytagel, pH 5.8) under long day condtions (16h light) at 24 °C (day) 22 °C (night) using the DNeasy Plant Kit (Qiagen). *ONSEN* copy numbers were determined by qPCR using 12 ng total DNA using the KAPA SYBR FAST master mix universal on a C1000

Touch (Bio-Rad) machine. *ACTIN2* (*At*3g18780) was used to normalize DNA levels and DNA of Col-0 served as a control. Three technical replicates were used and data were analyzed with the Bio-Rad CFX Manager 3.1 software. Sequences of oligos are listed in Table S2.

4.3 Identification and visualization of *ONSEN* and T-DNA insertions

Novel *ONSEN* insertions of hcLine31 were identified and characterized recently [17] by whole-genome sequencing and using the Transposable Insertion Finder v1.6 [41] and the TAIR10 version of the *A. thaliana* Col-0 reference genome [42]. To validate the presence and zygosity of the *ONSEN* and T-DNA insertions in *RPI2* in the segregating F2 populations, we designed primers (mto_007 and mto_067) spanning the predicted insertion sites of *ONSEN* and the T-DNA (based on SIGnAL) and combined them with primers specific to *ONSEN* (mto_196) or the T-DNA (LBb1.3 mto_063) (Fig 3a, Table S2). For the PCRs we used a standard Taq DNA polymerase (Sigma Aldrich) and limited the elongation time to 90 seconds so that an homozygous insertion of the 5-kb *ONSEN* TE or the T-DNA would prevent the formation of a PCR-product. For the genotyping of the F2 populations, we used DNA of homozygous parental plants (wt, hcLine31 and *rpi2-1* (SALK_022117)).

RNA-seq data of one representative biological replicate of the Col-0 wt and hcLine31 exposed to CS and whole-genome sequencing data of hcLine31 were obtained from and analyzed according to a previous report [17]. Genomic reads were mapped to the TAIR10 version of the *A. thaliana* genome using bwa mem (v. 0.7.17-r1188) [43] with the -M parameter set. The insertion site of *ONSEN* was then visualized using the packages Gviz (v. 1.28.3) [44], rtracklayer (v. 1.44.4) [45] and the annotation package TxDb.Athaliana.BioMart.plantsmart28 (v. 3.2.2) [46] using R (v. 3.6.3) [47] in Rstudio (v. 1.1.456) [48].

4.4 Drought assay

To obtain comparable and robust results, we ran one comprehensive drought experiment where we tested the S4 generation of hcLines, the control lines and the segregating F2 generations of crosses between hcLine31 and Col-0, and hcLine31 and *rpi2-1* (SALK_022117) in parallel. We included five replicates for each high-copy and control line and for the parents of the cross between *rpi2-1* and hcLine31, and tested 22 F_2 individuals of the cross of hcLine31 with the wt and 16 F_2 individuals of the cross of hcLine31 and *rpi2-1* (SALK_022117). Seeds were sown in pots filled with Einheitserde that was incubated with a solution (75 mg/L) of the insecticide Kohinor (Leu+Gygax AG) and kept at 4°C for three days. After stratification, pots were moved into a Hiros climate chamber (Clitec) set to short-day conditions with 10 h light (LED Valoya Ns12 C75/65, ~120 µmol*m-2*s-1) at 22°C (day) and 19°C (night), with 60% humidity. After ten days of growth, seedlings were piqued into pots filled with equal amounts of Kohinor-treated soil and grown under well-watered conditions for 36 days. The position of pots was frequently shifted to ensure similar growth conditions. Before watering was suspended, pots were again saturated with water and weighted to obtain the maximal water content. One day later (day 1 of the experiment), top-view pictures were taken with a Canon EOS 70D camera on a tripod at the following settings: 5.6 s shutter opening, 1/60 shutter speed, ISO 200. This procedure was repeated three times at an interval of seven days until day 22 of the experiment. Due to technical issues nine out of 193 images (affecting one to two biological replicates of seven different lines and one biological replicate of the F2 of the cross of hcLine31 and *rpi2*-1) from day 15 are missing in the analysis. On day 27, pots were again weighted, top-view pictures were taken, and drought stress was stopped by filling trays with water and allowing the pots to absorb water over night. After two days of regeneration under well-watered conditions, final pictures of the plants were taken (day 29). To account for different zoom levels during the course of the experiment, we took pictures of a white label that later served as a calibrator to normalize predicted vital and necrotic leaf areas. One day after the last pictures were taken, one leaf of each plant from the segregating F2 populations was sampled for DNA extractions and

genotyping. Pots were removed from the climate chamber and dried for 8 weeks at room temperature to obtain the dry weight in order to calculated the water content of each pot. Pictures and weight were determined on two successive days (except for day 27); therefore, we extrapolated the weight measurements to determine the water content on the exact day pictures were taken.

4.5 Machine-learning-based prediction of necrotic and vital leaf areas

We used the pixel classification tool of ilastik (v. 1.3.3post3) [18] with all 13 features for color/intensity, edge, texture of sigma 0.3, 0.7 and 10.00 selected. We defined three different pixel classes: “background”, “necrotic” and “vital”. We performed an iterative manual training to gradually improve the accuracy of the prediction and finally used 24 images of plants at different stages of the experiment to train the model. For the disc assay, we combined one to three leaf-discs that were punched from vital and necrotic leaves onto soil of a single pot that did not contain a plant. We took three pictures of each combination. To cover a broad spectrum of possibilities, leaf discs were shuffled and/or moved on the pot between pictures. Hence some discs were photographed multiple times but in different combinations and/or positions. Similar to the model training with living and necrotic tissues, we used five images to train ilastik to detect background and the white labels that served as a scale to normalize pixel counts between different days of the experiment. We then processed and exported all image files in ilastik with the following settings: source, “simple segmentation”; convert to datatype, “floating 32 bit”; format, “tif”, and used the getHistogram function of ImageJ (v. 1.53g, Java 1.8) [49] to extract the pixel counts of the areas predicted by ilastik. Pixel counts of vital and necrotic leaves were then normalized using the predicted areas of the size references for each time point of the drought experiment using R.

4.6 SPAD-value and C/N ratio

We grew S4 generation plants under the same conditions and watering regime as for the drought experiment and used a chlorophyll meter SPAD-502 chlorophyll meter [25] to determine the Soil Plant Analysis Development (SPAD) values of three or six leaves of six wt and hcLine31 plants that were grown as triplicates in two independent experiments. SPAD values were measured before the occurrence of necrotic leaves 15 days after watering was suspended.

To determine the carbon/nitrogen ratios of the wt and hcLine31 we grew S3 generation plants under well-watered conditions on Einheitserde under short day conditions in a Sanyo MLR-350 growth chamber with 8 h light at 20 °C (day) and 18 °C (night) for 14 weeks. We sampled one leaf from each of four plants of the wt and hcLine31, inactivated them for 30 seconds in a microwave and dried them for eight days at 60°C. Plant material was then ground for 3 minutes at 30 hz with an oscillating mill (M M 400, Retsch, Germany). Then, 2 mg of plant material were put in tin capsule and C/N-ratios were analyzed with a thermal conductivity detector by the Basel Stable Isotope Lab.

4.7. *RPI2* analysis in natural accessions

We used the vcf file produced by The 1001 Genomes Consortium [50] to extract SNPs for 1135 sequenced accessions of *A. thaliana*. To limit the effect of the phylogenetic relationship in further analyses, we used the function –relatedness2 from vcftools [51] to keep only ecotypes with a kinship coefficient k <0.5. For the remaining ecotypes, bioclimatic variables (https://www.worldclim.org/data/worldclim21.html) and Global Aridity Index (https://cgiarcsi.community/data/global-aridity-and-pet-database/) were extracted using the Rpackages raster (v. 3.5-2; [52]) and rgdal (v. 1.5-27; [53]). We fitted a linear-mixed model with the R package lme4 (v. 1.1-27; [54]) to test the association between aridity levels averaged over May to August and SNPs in *RPI2* harboring a minor allele frequency (maf) > 0.3. We added the admixture groups defined by The 1001 Genomes Consortium [50] as a random

effect to account for population structure. We used QGIS (v. 3.16; https://www.qgis.org/en/site/) to display *RPI2* alleles in Eurasia. For the rest of the manuscript, all statistical analyses were performed in R and parametric or non-parametric tests were used according to the sample size and data distribution.

Data availability

Raw and segmented images and ilastik classifications were uploaded to figshare.

Acknowledgements

We would like to thank Anja Schmutz for her assistance with R, Mahendra Mariadassou for his kind support with statistics and Yann Bourgeois and Christoph Stritt for their comments on the manuscripts. We further want to thank Ansgar Kahmen and Thomas Boller for their involvement at the beginning of the project.

Author contributions

Experimental design: M.T.; data acquisition: M.T., B.K. and A.B.; data analysis: M.T. and A.R.; writing of the manuscript: M.T. with contributions from A.R., E.B., A.B. and B.K.

Funding

University of Zurich Research Priority Programs (URPP) *Evolution in Action* (to M.T., B.K. and A.R.);

European Commission [PITN-GA-2013-608422-IDP BRIDGES] and Freiwillige Akademische Gesellschaft Basel (to M.T.);

[SNSF 31003A_182785] (to A.R. and B.K.)

European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [725701, BUNGEE] (to E.B.);

Competing interests

The authors declare no competing interests.

Material & Correspondence

Correspondence and material requests should be addressed to Michael Thieme (michael.thieme@botinst.uzh.ch)

References

1. Wells, J. N. & Feschotte, C. A Field Guide to Eukaryotic Transposable Elements. *Annual Review of Genetics* **54**, 539–561 (2020).

2. Baduel, P. *et al.* Genetic and environmental modulation of transposition shapes the evolutionary potential of Arabidopsis thaliana. *Genome Biol.* **22**, 138 (2021).

3. Thieme, M. *et al.* Inhibition of RNA polymerase II allows controlled mobilisation of retrotransposons for plant breeding. *Genome Biol* **18**, 134 (2017).

4. Tittel-Elmer, M. *et al.* Stress-induced activation of heterochromatic transcription. *PLoS Genet* **6**, e1001175 (2010).

5. Ito, H. *et al.* An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* **472**, 115–119 (2011).

6. Cavrak, V. V *et al.* How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet* **10**, e1004115 (2014).

7. Exposito-Alonso, M. *et al.* Natural selection on the Arabidopsis thaliana genome in present and future climates. *Nature* **573**, 126–129 (2019).

8. Esnault, C., Lee, M., Ham, C. & Levin, H. L. Transposable element insertions in fission yeast drive adaptation to environmental stress. *Genome Res.* **29**, 85–95 (2019).

9. Belfield, E. J. *et al.* Thermal stress accelerates arabidopsis thaliana mutation rate. *Genome Res.* **31**, 40–50 (2021).

10. Negi, P., Rai, A. N. & Suprasanna, P. Moving through the stressed genome: Emerging regulatory roles for transposons in plant stress response. *Front. Plant Sci.* **7**, 1448 (2016).

11. McClintock, B. THE ORIGIN AND BEHAVIOR OF MUTABLE LOCI IN MAIZE. *Proc Natl Acad Sci U S A.* **36**, 344–355 (1950).

12. Lisch, D. How important are transposons for plant evolution? *Nature Reviews Genetics* **14**, 49–61 (2013).

13. Naito, K. *et al.* Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. *Nature* **461**, 1130–1134 (2009).

14. Li, Z.-W. *et al.* Transposable Elements Contribute to the Adaptation of Arabidopsis thaliana. *Genome Biol. Evol.* **10**, 2140–2150 (2018).

15. Matzke, M. A. & Mosher, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* **15**, 394–408 (2014).

16. Brás, T. A., Seixas, J., Carvalhais, N. & Jagermeyr, J. Severity of drought and heatwave crop losses tripled over the last five decades in Europe. *Environ. Res. Lett.* **16**, 065012 (2021).

17. Roquis, D. *et al.* Genomic impact of stress-induced transposable element mobility in Arabidopsis. *Nucleic Acids Res.* (2021). doi:10.1093/nar/gkab828

18. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226–1232 (2019).

19. Liu, C. H. *et al.* Repair of DNA Damage Induced by the Cytidine Analog Zebularine Requires ATR and ATM in Arabidopsis. *Plant Cell* **27**, 1788–1800 (2015).

20. Wilke, C. M. & Adams, J. Fitness effects of Ty transposition in Saccharomyces cerevisiae. *Genetics* **131**, 31–42 (1992).

21. Boissinot, S., Davis, J., Entezam, A., Petrov, D. & Furano, A. V. Fitness cost of LINE-1 (L1) activity in humans. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9590–9594 (2006).

22. Chuong, E. B., Elde, N. C. & Feschotte, C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet* **18**, 71–86 (2017).

23. Quadrana, L. *et al.* Transposition favors the generation of large effect mutations that may facilitate rapid adaption. *Nat. Commun.* **10**, 1–10 (2019).

24. Xiong, Y., Defraia, C., Williams, D., Zhang, X. & Mou, Z. Deficiency in a cytosolic ribose-5-phosphate isomerase causes chloroplast dysfunction, late flowering and premature cell death in Arabidopsis. *Physiol. Plant.* **137**, 249–263 (2009).

25. Ling, Q., Huang, W. & Jarvis, P. Use of a SPAD-502 meter to measure leaf chlorophyll concentration in Arabidopsis thaliana. *Photosynth. Res.* **107**, 209–214 (2011).

26. Otori, K. *et al.* Enhanced photosynthetic capacity increases nitrogen metabolism through the coordinated regulation of carbon and nitrogen assimilation in Arabidopsis thaliana. *J. Plant Res.* **130**, 909–927 (2017).

27. Ito, H. *et al.* A Stress-Activated Transposon in Arabidopsis Induces Transgenerational Abscisic Acid Insensitivity OPEN. *Nat. Publ. Gr.* (2016). doi:10.1038/srep23181

28. Quadrana, L. *et al.* The Arabidopsis thaliana mobilome and its impact at the species level. *Elife* **5**, (2016).

29. Barboza, L. *et al.* Arabidopsis semidwarfs evolved from independent mutations in GA20ox1, ortholog to green revolution dwarf alleles in rice and barley. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15818–15823 (2013).

30. Loarie, S. R. *et al.* The velocity of climate change. *Nature* **462**, 1052–1055 (2009).

31. Hermisson, J. & Pennings, P. S. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* **8**, 700–716 (2017).

32. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic variation. *Trends in Ecology and Evolution* **23**, 38–44 (2008).

33. Johanson, U. *et al.* Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science (80-.).* **290**, 344–347 (2000).

34. Gujas, B., Alonso-Blanco, C. & Hardtke, C. S. Natural arabidopsis brx loss-of-function alleles confer root adaptation to acidic soil. *Curr. Biol.* **22**, 1962–1968 (2012).

35. Kroymann, J., Donnerhacke, S., Schnabelrauch, D. & Mitchell-Olds, T. Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14587–14592 (2003).

36. Monroe, J. G. *et al.* Drought adaptation in arabidopsis thaliana by extensive genetic loss-of-function. *Elife* **7**, (2018).

37. Xu, Y.-C. *et al.* Adaptation and Phenotypic Diversification in Arabidopsis through Loss-of-Function Mutations in Protein-Coding Genes. *Plant Cell* **31**, 1012–1025 (2019).

38. Vasseur, F. *et al.* Adaptive diversification of growth allometry in the plant Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3416–3421 (2018).

39. Olson, M. V. *MOLECULAR EVOLUTION '99 When Less Is More: Gene Loss as an Engine of Evolutionary Change*. *Am. J. Hum. Genet* **64**, (1999).

40. Alonso, J. M. *et al.* Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science (80-.).* **301**, 653–657 (2003).

41. Nakagome, M. *et al.* Transposon Insertion Finder (TIF): A novel program for detection of de novo transpositions of transposable elements. *BMC Bioinformatics* **15**, 71 (2014).

42. Berardini, T. Z. *et al.* The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome. *genesis* **53**, 474–485 (2015).

43. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

44. Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and bioconductor. in *Methods in Molecular Biology* **1418**, 335–351 (Humana Press Inc., 2016).

45. Lawrence, M., Gentleman, R. & Carey, V. rtracklayer: An R package for interfacing with genome browsers. *Bioinformatics* **25**, 1841–1842 (2009).

46. Carlson, M. & Maintainer, B. P. TxDb.Athaliana.BioMart.plantsmart28: Annotation package for TxDb object(s). (2015).

47. R Core Team. R: A Language and Environment for Statistical Computing. (2020).

48. RStudio Team. RStudio: Integrated Development Environment for R. (2016).

49. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).

50. Alonso-Blanco, C. *et al.* 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166**, 481–491 (2016).

51. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

52. Hijmans, R. J. & van Etten, J. raster: Geographic analysis and modeling with raster data. (2012).

53. Keitt, T. H., Bivand, R., Pebesma, E. & Rowlingson, B. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. (2010).

54. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

Supplementary tables

Table S1 SNPs in *RIP2* and bioclimatic variables for the subset of 596 ecotypes.

Table S2 Names and sequences of oligos used for the qPCR to determine *ONSEN* copy numbers and for genotyping of *RPI2*

Table S3 Raw data of pixel counts and weight, SPAD values and C/N-measurements

Extended data figures

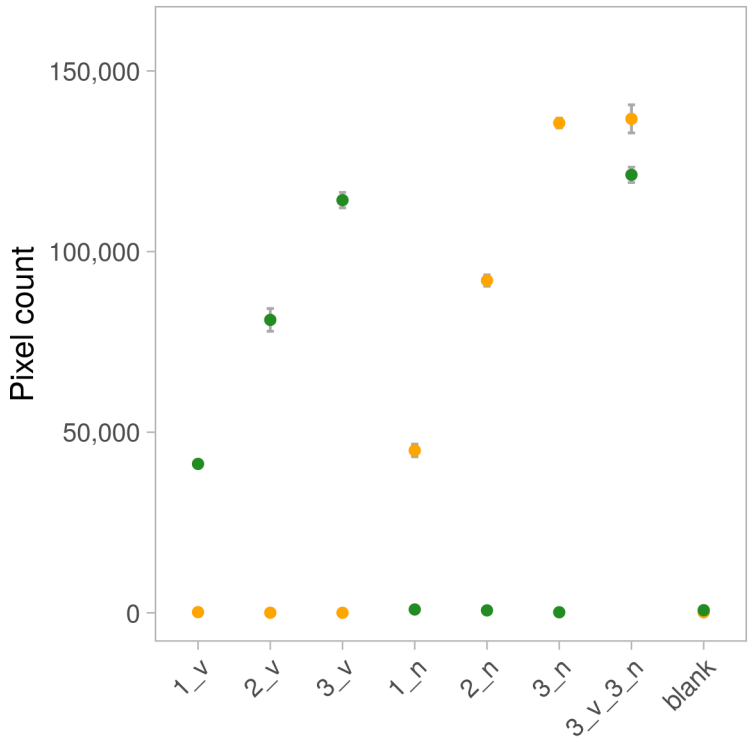

Fig S1 Leaf disc assay to validate the accuracy of the pixel counts for the machine learning-based prediction of vital (green) and necrotic (yellow) segments using ilastik. One, two or three discs were punched from vital (v) or necrotic (n) leaf tissues and placed onto soil of an empty pot (blank). n=3 technical replicates.

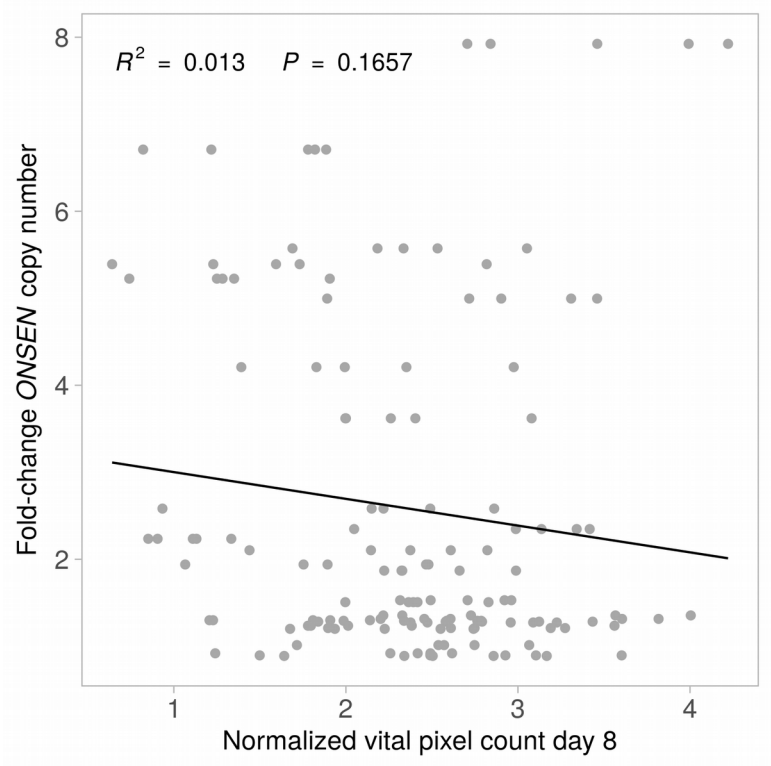

Fig S2 Pixel count of vital tissues in all tested lines before the occurrence of necrotic leaves (day 8) (n=5 biological replicates) and the fold change of *ONSEN* copy numbers measured by qPCR relative to the wt. n=3 technical replicates. Linear regression model, adjusted R-squared: 0.013, P=0.1657.

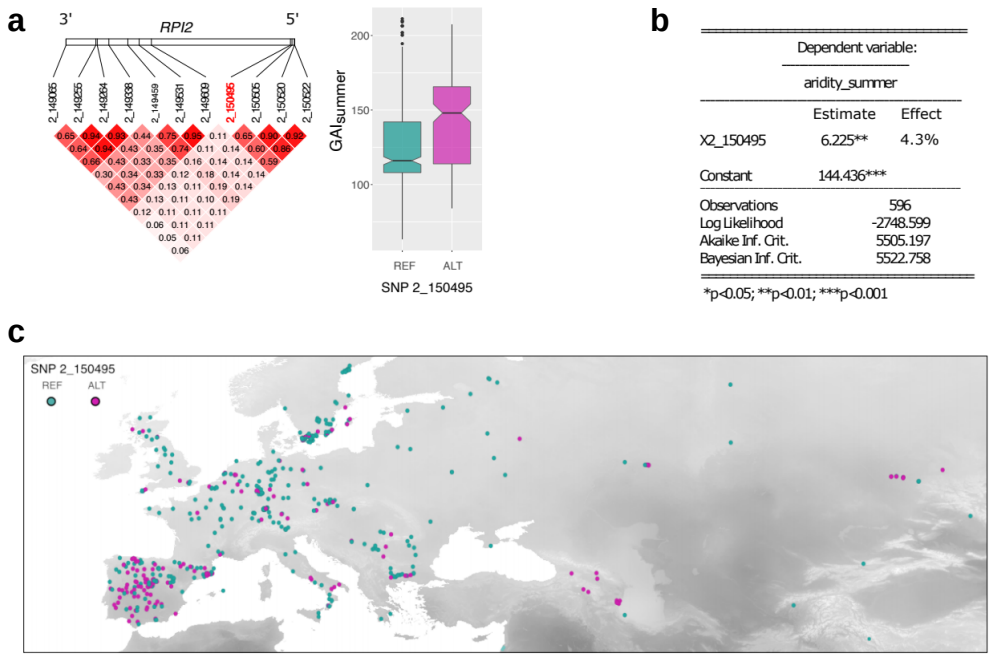

Fig S3 The association between aridity and *RPI2* variations in natural populations of *A. thaliana.* **a** Linkage disequilibrium among the 11 SNPs located in *RPI2* (left panel) and the association between SNP 2_150495 and Global Aridity Index in summer (GAIsummer; right panel). **b** Linear mixed-model results based on 596 accessions. **c** Occurrence of the reference and alternative alleles of SNP 2_150495 in Eurasia. The intensity of the background map displays aridity levels in July from low (light grey) to high (dark grey)