A rainfall-manipulation experiment with 517 *Arabidopsis thaliana* accessions

Moises Exposito-Alonso[1], Rocío Gómez Rodríguez[2], Cristina Barragán[1], Giovanna Capovilla[1], Eunyoung Chae[1], Jane Devos[1], Ezgi S. Dogan[1], Claudia Friedemann[1], Caspar Gross[1], Patricia Lang[1], Derek Lundberg[1], Vera Middendorf[1], Jorge Kageyama[1], Talia Karasov[1], Sonja Kersten[1], Sebastian Petersen[1], Leily Rabbani[1], Julian Regalado[1], Lukas Reinelt[1], Beth Rowan[1], Danelle K. Seymour[1], Efthymia Symeonidi[1], Rebecca Schwab[1], Diep Thi Ngoc Tran[1], Kavita Venkataramani[1], Anna-Lena Van de Weyer[1], François Vasseur[1], George Wang[1], Ronja Wedegärtner[1], Frank Weiss[1], Rui Wu[1], Wanyan Xi[1], Maricris Zaidem[1], Wangsheng Zhu[1], Fernando García-Arenal[2], Hernán A. Burbano[1], Oliver Bossdorf[3], Detlef Weigel[1].

[1] Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany
[2] Center for Plant Biotechnology and Genomics, Technical University of Madrid, Pozuelo de Alarcón, Spain
[3] Institute of Ecology and Evolution, University of Tübingen, Tübingen, Germany

Abstract

The gold standard for studying natural selection is to quantify lifetime fitness in individuals from natural populations that have been grown together under different field conditions. This has been widely done in ecology to measure phenotypic selection in nature for a wide range of organisms — an evolutionary force that seems to be most determined by local precipitation patterns. Studies that include whole-genome data would enable the translation of coefficients of selection to the genetic level, but such studies are still scarce, even though this type of genetic knowledge will be critical to predict the effect of climate change in natural populations. Here we present such an experiment including rainfall-manipulation with the plant *Arabidopsis thaliana*. The experiment was carried out in a Mediterranean and a Central European field station with rainout shelters to simulate a high and low rainfall treatment within each location. For each treatment combination, we planted 7 pots with one individual and 5 pots with 30 counted seeds of 517 whole-genome sequenced natural accessions covering the global species distribution. Survival, germination, flowering time, and final seed output were measured for ca. 25,000 pots, which contained ca. 14,500 individual plants and over 310,00 plants growing in small populations. This high-throughput phenotyping was only possible thanks to image analysis techniques using custom-made scripts. To make the data and processing code available, we created an R package "dryAR" (http://github.com/MoisesExpositoAlonso/dryAR).

Running title: A Climate Change Experiment with *A. thaliana*
Keywords: Field experiment, Climate Change, *Arabidopsis thaliana*

Field experiment design

The accessions from the 1001 Genomes Project

The 1001 Genomes (1001G) project (1001 Genomes Consortium 2016) comprises 1,135 sequenced genetic lines called accessions (Fig. 1). To select the most genetically and geographically informative and least biased 1001G lines, we set some quality criteria. These consisted of several filters: (1) First we removed the accessions with the lowest genome quality. We discarded those with < 10X genome coverage of Illumina sequencing reads and < 90% congruence of SNPs called from MPI and GMI pipelines (1001 Genomes Consortium 2016), which resulted in 959 accessions. (2) We removed almost-identical individuals. Using Plink (Purcell et al. 2007) we computed identity by state genome-wide across the 1,135 accessions. For pairs of accessions with < 0.01 differences per SNP, we randomly picked one. This resulted in 889 accessions. After applying criteria (1) and (2) sequentially, 762 accessions remained. (3) Finally, we reduced geographic ascertainment. Sampling for 1001G was not performed in either a random nor regularly structured scheme. Some laboratories provided several lines per location whereas others provided lines that were collected at least several hundred kilometres apart. Employing latitude and longitude degrees, we computed Euclidean distances across the 1,135 accessions and identified all pairs that were < 0.0001 distance apart, that is, accessions from the same population (<< 100 meters). From such pairs, we randomly picked one. Applying this filter independently of (1) and (2) resulted in 682 accessions. We intersected the resulting lists of accessions after the quality filtering procedures and obtained a final set of 523 accessions. We propagated accessions in controlled conditions. We stratified the seeds one week at 4°C, then we sowed them in trays with industrial soil (CL-P, Einheitserde Werkverband e.V., Sinntal-Altengronau Germany) and placed them in a growth room at 16h light and 23°C for a week. Trays were then vernalized for 60 days at 4°C vernalization and 8h light. After vernalization, trays were moved back to 23°C and 16 h light for final growth and reproduction. This generated sufficient seeds for 517 accessions that were then grown in the field (Fig. 1). Seeds descendant from the same parents can be ordered from the 1001G seed stock at the Arabidopsis Biological Resource Center (CS78942).

Field settings and watering

Rainout shelter design

We built two 30 m x 6 m tunnels of PVC plastic foil to fully exclude rainfall in Madrid and in Tübingen (Fig. 2). The foil tunnels are different from a regular greenhouse in that they are completely open on two sides. Thus, ambient temperatures vary almost as much as in an outdoor experiment (see Environmental sensors section). In each location, we supplied artificial watering at two contrasting regimes: abundant watering and reduced watering. Inside each tunnel, we created an approximate 4% slope and set up four flooding tables on the ground (1 m x 25 m, Hellmuth Bahrs GmbH & Co KG, Brüggen, Germany) covered with soaking mats (4 L/m^2, Gärtnereinkauf Münchingen GmbH, Münchingen, Germany). The lower end of the flooding table was used to

drain the water provided from the other, higher end of the table (Fig. 2). To imitate rainfall, watering was also supplied using a watering gun.

We used potting trays of 8x5 cells (5.5 cm x 5.5 cm x 10 cm size) and industrial soil (CL-P, Einheitserde Werkverband e.V., Sinntal-Altengronau Germany). One genotype was planted per cell, excluding corner cells, to avoid edge effects. We grew a total of 12 replicates per genotype per treatment. Five replicates were planted at a density of 30 counted seeds per cell and grew without further intervention ("population replicate"). Seven were planted at low density (ca. 10 seeds) and once germinated, one seedling was selected at random and all others were removed ("individual replicate"). While the population replicates should more faithfully reflect survival from seed to reproductive adult, the individual replicates were useful because they could be more accurately (individually) monitored for flowering time and seed set.

We used a randomized incomplete block design (Fig. 2). Because 36 pots were used per tray, a total of 14.36 trays amounted to one replicate of all 517 genotypes. A total of 16 treatment blocks were established. For each watering treatment there were two intercalated blocks. Within each flooding table there were four also intercalated blocks, two of individual replicates and two of population replicates. The genotypes were randomized within replicate block and were distributed along the treatment block. The design was identical in Madrid and Tübingen (Fig. 2).

Environmental sensors

Environmental variables — air temperature, photosynthetic active radiation and soil water content — were monitored in real time (one record every 15 minutes) throughout the experiment using multi-purpose sensors (Flower Power sensor, Parrot SA, Paris, France). This enabled us to adjust watering depending on the degree of local evapotranspiration during the course the experiment. The sensors outside of the tunnel in Madrid (i.e. only natural rainfall) showed a interquartile range between 1 and 17% soil water content. This overlapped which the range of 10 to 22% water content of the drought treatment we artificially imposed inside of the tunnels both in Madrid and in Tübingen. The relatively lower measurements by the outside sensor in Madrid is due to a complete lack of natural rainfall during the first two months of the experiment. On the other hand, the sensor outside of the tunnel in Tübingen recorded an interquartile range of soil water content percentage, 22 to 27%, that was comparable to the high watering treatments in Tübingen and Madrid, from 20 to 33%. These values confirmed that our low and high watering treatment not only were different, but also that they mimicked natural watering of two contrasting locations. Air temperatures were overall higher in Madrid (5-6°C) than in Tübingen (8-10°C), as expected, and the difference in tempreature between the sensors inside and outside of the tunnel was only of one degree on average (Table 1). The photosynthetic active radiation (PAR, wavelengths from 400 to 700 nm) had a median of 0.1 mole m^{-2} day^{-1} at night for all experiments. At mid-day (11:00-13.00 hrs), the median PAR outside the tunnel in Madrid was 57.81 mole m^{-2} day^{-1} and 45.24 and 46.24 for the low and high treatments inside the tunnel. In Tübingen, the median values were 29.02 outside, and 34.36 and 27.50 inside the tunnel.

Sowing and quality control

During sowing, contamination of neighboring pots with adjacent genotypes can occur for multiple reasons. In order to avoid such contamination, we chose a day with no wind and we sowed the seeds at only 1-2 cm height from the soil. Additionally, watering during the first days was gentle to avoid seed-carryover. We also tried to remove human error during sowing by preparing and curating 2 mL plastic tubes containing the seeds to be sown in cardboard boxes with the same cells (5x8) as in the target trays and arranged them in their corresponding (randomised) locations. During sowing, each experimenter took a box at random and went to the corresponding previously labeled and arranged tray in the field (Fig. 2). This reduced the possibility of sowing errors.

Later, during the vegetative growth, we could identify germinated seedlings that looked like neighbour contamination and removed such plants. Although this meant the loss of a number of plants, the high replication of the experiment allowed this sacrifice. During the recording of flowering time, we used the homogeneity of flowering within a pot as a further indicator for contamination. When a plant had a completely different flowering timing and leaf phenotypes did not coincide with the majority of the pot, this plant was removed.

After sowing and removal of errors, the total number of pots was 24,747 instead of the original 24,816 pots.

Monitoring of plants

Image analysis of vegetative rosettes

Top-view images were taken every four to five days (median) with a Panasonic DMC-TZ61 digital camera and a customized closed dark box (Fig. 3) at a distance of 40 cm from each tray. After testing different camera parameters, we used an exposure of -2/3 and an ISO of 100. White balance was set for flashlight. As we used a dark box from all sides closed, this was the only source of illumination, which ensured that the white balance and illumination were consistent from picture to picture. Photos were saved both in .jpeg and .raw to allow for *a posteriori* adjustments if needed. Using a calibration board with white and dark squares of 1.3 cm x 1.3 cm, we studied the error in retrieving the true area across the tray. This provided us with a median resolution estimate of 101.5 pixels mm^{-2}. The deviations from the true area were minimal, typically from 0 to 5%, with a maximum of 8-9% deviation in area in the extreme corners of the tray (where we did not sow any seeds). We are confident that such small variation in retrieved area are more than compensated by the randomized locations of genotypes within the trays.

In total, we imaged each tray at 20 timepoints throughout the vegetative growth. All images are deposited at http://datadryad.org/[updatehere] and the Python module to process and

analyse them is available at http://github.com/MoisesExpositoAlonso/hippo. The implemented segmentation was virtually the same as in (Exposito-Alonso et al. 2017), which relies on the Open CV Python library (Itseez 2015). We began by transforming images from RGB to HSV channels. We applied a hard segmentation threshold of HSV values as (H=30-65, S=65-255, V=20-220). The threshold was defined after manually screening 10 different plants in order to capture the full spectrum of greens from different accessions and of different developmental stages. This was followed by several iterations of morphology transformations based on erosion and dilation. Then, for the resulting binary image we counted the number of green pixels.

During field monitoring we noticed that seeds in some pots had not germinated. Sometimes this was due to lack of seeds or improper soil compaction. In these cases, we left a red mark in those pots, which we could detect in the same way as the existence of green pixels (with threshold H=150-179, S=100-255, V=100-255). These pots were excluded from survival analysis as they did not contain any plants. An example of transformed images is shown in Fig. 3.

The resulting raw data consist of green and red area (pixel counts) per pot (Fig. 3). Some trays were photographed twice on the same day by mistake. We took advantage of this as a blind control to verify whether our camera settings and segmentation pipeline would recover the same area, i.e. to what degree the images were consistent and the pipeline was replicable. In total there were 1,508 pots whose area was estimated twice, distributed across 11 timepoints and different trays. The Spearman's rank correlation was $r=0.97$, $n=1508$, $p<10^{-16}$. This confirmed that replicability was high.

In order to remove pots that did not contain germinated seeds from the analyses, we performed an analysis of variance between pots above and below a moving threshold of red pixels to determine the number of red pixels. This provided us with the threshold at which which a pot was highly likely to have a red mark (indicating an empty pot). As expected, the distribution of pixels was bimodal (Fig. 3), what made this process straightforward and reliable.

Then we estimated germination timing. One approach to do this was to model growth trajectories (Fig. 3) of green pixels per pot as a sigmoidal curve, fitting the function:

$$y = \frac{a}{1+e^{-(b\times(x-c))}}$$

, starting on the sowing day and until the apparent peak of green pixels per pot. The sigmoidal curve could be fitted for 12,636 pots. The three parameters a, b, and c, inform about the different shapes of growth curves. We also computed less complex indicators of growth: an analogous linear model that was used to determine the intersection with 1,000 pixels, i.e., the day that over 1,000 green pixels were observed (~ 10 mm^2, Fig. 4), the day that a fitted spline passed over 1,000 green pixels, and a total count of green and red pixels through all timepoints. A detailed R markdown document of data loading and cleaning can be found at http://github.com/MoisesExpositoAlonso/field/data-cleaning/gen_vegetative.html. The final dataset

contained data for 22,779 pots — after the removal of pots with red labels — for which we had a time series of green areas.

Manual recording of flowering time

We visited the experiment every 1-2 days and manually recorded the pots with flowering plants. To keep track of previous visits and avoid errors, we labeled the pots where flowering had already been recorded with blue pins. This removed another potential source of human error. To calculate flowering time, we counted the number of days from the date of sowing to the recorded flowering date. Fig. 4 shows the raw flowering time data per pot in the original spatial distribution (Fig. 2) and the distribution of flowering time per treatment combination. Note that grey boxes in Fig. 4 are pots with plants that did not survive until flowering. For more visualizations of flowering time see http://github.com/MoisesExpositoAlonso/dryAR/analyses/flowering_exploration.html. In total, we gathered data for 16,858 flowering pots.

Image analysis of reproductive plants

Once the first dry fruits were observed, we harvested them and took a final 'studio photograph' of the rosette and the inflorescence (Fig. 5). In total, we took 13,849 photographs. The camera settings were the same as for the vegetative monitoring, but here we included an 18% grey card approximately in the same location in case *a posteriori* adjustments would be needed. The Python module to analyse the inflorescence pictures (Fig. 5) is available at http://github.com/MoisesExpositoAlonso/hitfruit. We first used a cycle of morphological transformations of erode-and-dilate to produce the segmented image (Fig. 5). This generated a segmented white/black image without white noise. Then, we used the thin (erode cycles) algorithm from the Mahotas library (Coelho 2013) to generate a binary picture reduced to single-pixel paths — a process called skeletonisation (Fig. 5). Finally, to detect the branching points in the skeletonised image we used a hit or miss algorithm from Mahotas. We used customized structural elements to maximize the branch (Fig. 5) and end point detection (Fig. 5). This resulted in four variables per image: total segmented inflorescence area, total length of the skeleton path, number of branching points and number of end points (Fig. 5).

Because we ran the same segmentation and skeletonization software on rosette images, we could leverage the different image patterns that rosettes and inflorescences have to identify labeling errors (i.e. mistakes in inputting sample information of the pictures). To do this, we first trained a random forest model to predict the manually labeled organ by the four image variables. From this exercise, a total of 92.1% were correctly predicted from image analysis, and ca. 2,000 images were incorrectly predicted. This could be either because there might be ranges of organ morphology that are relatively similar (for instance, we noticed that very small inflorescences and rosettes were confounded), or could be due to real mislabeling. Manually re-labeling about 500 pictures, we discovered that only the 2.5% of them had been incorrectly labeled. As the mislabeled

subset of 2,000 images picked by the machine learning algorithm must contain an overrepresentation of errors, we are confident that the labeling error in the dataset of 13,848 imaged plants must be below 2.5%.

A detailed R markdown document of data loading and cleaning can be found at http://github.com/ MoisesExpositoAlonso/field/data-cleaning/gen_harvesting.html.

Prediction of number of fruits and seeds

Although the study of natural selection is based on studying relative fitness, sometimes it is useful to have at least an approximation of the absolute fitness. In order to provide an approximate number of how many seeds each plant had produced, we generated two allometric relationships by manual counting of fruits per plant and seeds per fruit. In order to be sure that the counts corresponded to single plants, we counted fruits and seeds of only individual replicates of accessions, not the population replicates (see Field experiment design section). The first allometric relationship was built by manually counting the number of fruits per inflorescence of three sets of inflorescences, very small, intermediate, and very large ones (n=11). The variance explained by the carefully counted number of fruits and the four image analysis variables was high (R^2=0.97, $p=4\times10^{-4}$). We believe that the prediction of the number of fruits is appropriate for this type of data, as we had already shown a similar relationship with 350 manually counted fruits (Vasseur et al. 2017). The second relationship was the average number of seeds per fruit. To do this, in the same samples as before, we counted all seeds inside one fruit (n=11). We tried to sample fruits capturing the entire range of fruit size variation. The mean was 28.3 seeds per fruit and the standard deviation was 11.2 seeds. The two aforementioned allometric relationships were used to predict, first, the number of fruits per inflorescence using the four image analysis variables, and second, the number of seeds corresponding to the number of fruits per inflorescence.

Conclusions

This high-throughput field experiment has generated an invaluable dataset to study natural selection and adaptation in the context of global climate change — at the genetic level.

Author contributions

MEA conceived and designed the project. MEA carried out the experiment in Tübingen. MEA and RGR carried out the experiment in Madrid. All authors contributed to specific tasks in the experiments (see detailed description below). OB provided the field site in Tübingen and FGA provided the site in Madrid. DW secured funding for the project. MEA carried out the analyses and wrote the first draft of the manuscript. All authors commented and approved the manuscript.

AUTHOR	Conceived_idea	Funding	Advice	Coordination	Materials	Bulking_seeds	Seed_aliquoting	Field_setup	Pictures_plants	Sowing_Madrid	Sowing_Tuebingen	Thinning_seedlings	Field_care	Image_processing	Foil_tunnel_reparation	Fresh_harvesting_Madrid	Fresh_harvesting_Tuebingen	Dry_imaging_Madrid	Dry_imaging_Tuebingen	Flowering_monitoring	Image_processing	Data analysis/processing	Writing
Moises Exposito-Alonso	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Rocio Gomez Rodriguez							x	x	x	x			x			x				x			
Detlef Weigel		x	x		x												x						
Hernán A Burbano			x							x													
Oliver Bossdorf			x		x																		
Rebecca Schwab			x	x	x												x						
Fernando García Arenal			x		x																		
George Wang			x																				
François Vasseur			x								x												
Julian Regalado							x																
Derek Lundberg											x						x						
Ronja Wedegärtner							x	x	x		x		x				x						
Frank Weiss									x														
Danelle Seymour											x												
Beth Rowan											x				x		x						
Patricia Lang									x		x	x			x	x	x						
Jorge Kagayema											x												
Rui Wu											x				x		x						
Wanyan Xi											x												
Kavita Venkataramani											x				x	x	x						
Giovanna Capovilla												x			x		x						
Efthymia Symeonidi									x			x			x		x						
Vera Middendorf												x							x	x			
Anna-Lena Van de Weyer												x											
Jane Devos												x											
Diep Thi Ngoc Tran												x											
Sonja Kersten					x						x				x								
Wangsheng Zhu															x								
Maricris Zaidem															x								
Sebastian Petersen																	x						
Ezgi Dogan																	x						
Claudia Friedemann																x	x						
Talia Karasov																x							
Cristina Barragán																x							
Leily Rabbani																	x						
Caspar Gross																	x			x			
Lukas Reinelt													x							x			
Eunyoung Chae																	x						

Acknowledgements

We are thankful to Belen Mendez-Vigo, Carlos Alonso-Blanco and the technical service at CBGP-UMP, Antolín López Quirós, Marisa López Herránz and Miguel Ángel Mora Plaza, for assistance during sowing in Madrid, and Xavi Picó for experimental design advice.

References

1001 Genomes Consortium. 2016. "1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis Thaliana." *Cell* 166 (2). Elsevier: 481–91. doi:10.1016/j.cell.2016.05.063.

Coelho, Luis Pedro. 2013. "Mahotas: Open Source Software for Scriptable Computer Vision." *Journal of Open Research Software* 1 (1): e3. doi:10.5334/jors.ac.

Exposito-Alonso, Moises, Francois Vasseur, Wei Ding, George Wang, Hernan A. A. Burbano, and Detlef Weigel. 2017. "Genomic Basis and Evolutionary Potential for Extreme Drought Adaptation in Arabidopsis Thaliana." *bioRxiv*. doi:10.1101/118067.

Itseez. 2015. "Open Source Computer Vision Library." https://github.com/itseez/opencv.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81 (3): 559–75. doi:10.1086/519795.

Vasseur, Francois, Moises Exposito-Alonso, Oscar Ayala-Garay, George Wang, Brian J. Enquist, Cirile Violle, Denis Ville, and Detlef Weigel. 2017. "Scaling Irregularities Explained by Local Adaptation in Arabidopsis Thaliana." *Submitted*.

262

Tables

263

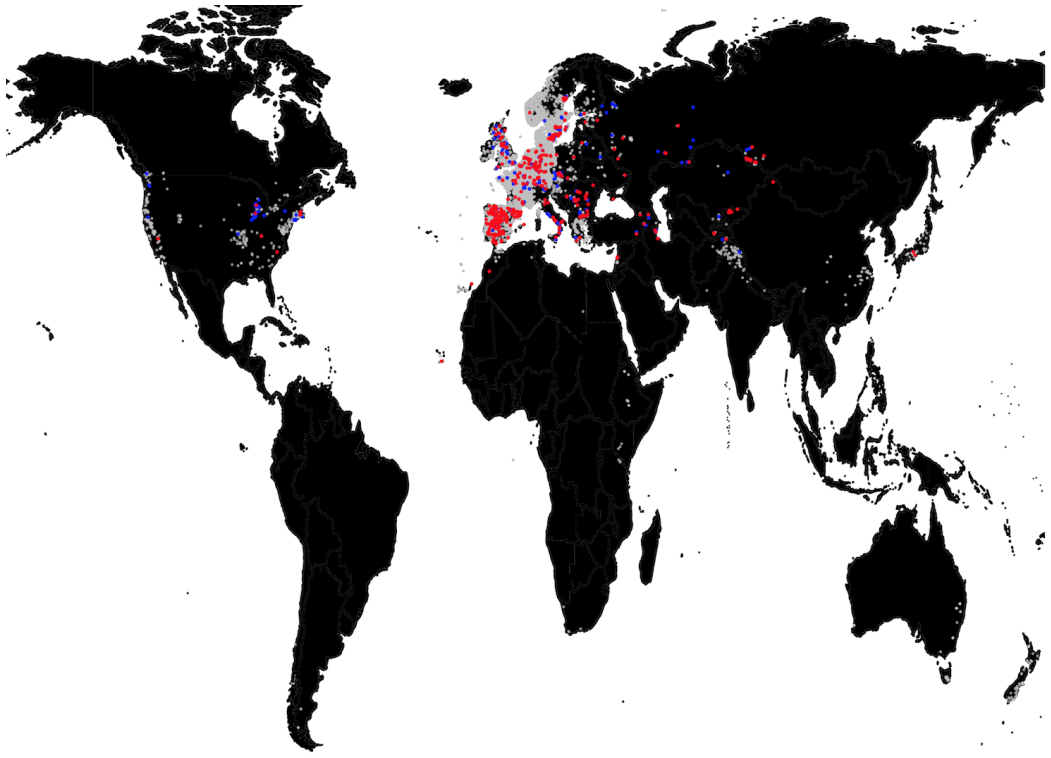
Table 1 Summaries of measurements from environmental sensor

Site	Rainfall	Soil water content (%)	Air temperature (°C)
Madrid	out	14.53 (1.09, 17.46)	8.45 (5.34, 12.39)
Madrid	low	16.07 (11.38, 22.51)	9.96 (6.95, 15.13)
Tübingen	low	14.74 (10.76, 20.09)	6.57 (3.27, 10.78)
Tübingen	out	27.67 (22.82, 30.50)	5.60 (2.44, 9.54)
Tübingen	high	24.62 (20.73, 29.02)	6.57 (3.27, 10.78)
Madrid	high	27.77 (22.62, 33.00)	9.84 (6.82, 15.13)

292 A total of 34 sensors were placed in the different treatment blocks as well as outside (out) of the foil
293 tunnels (Fig. 2). The median (interquartile) values of all sensors per treatment and location are shown.

Figures

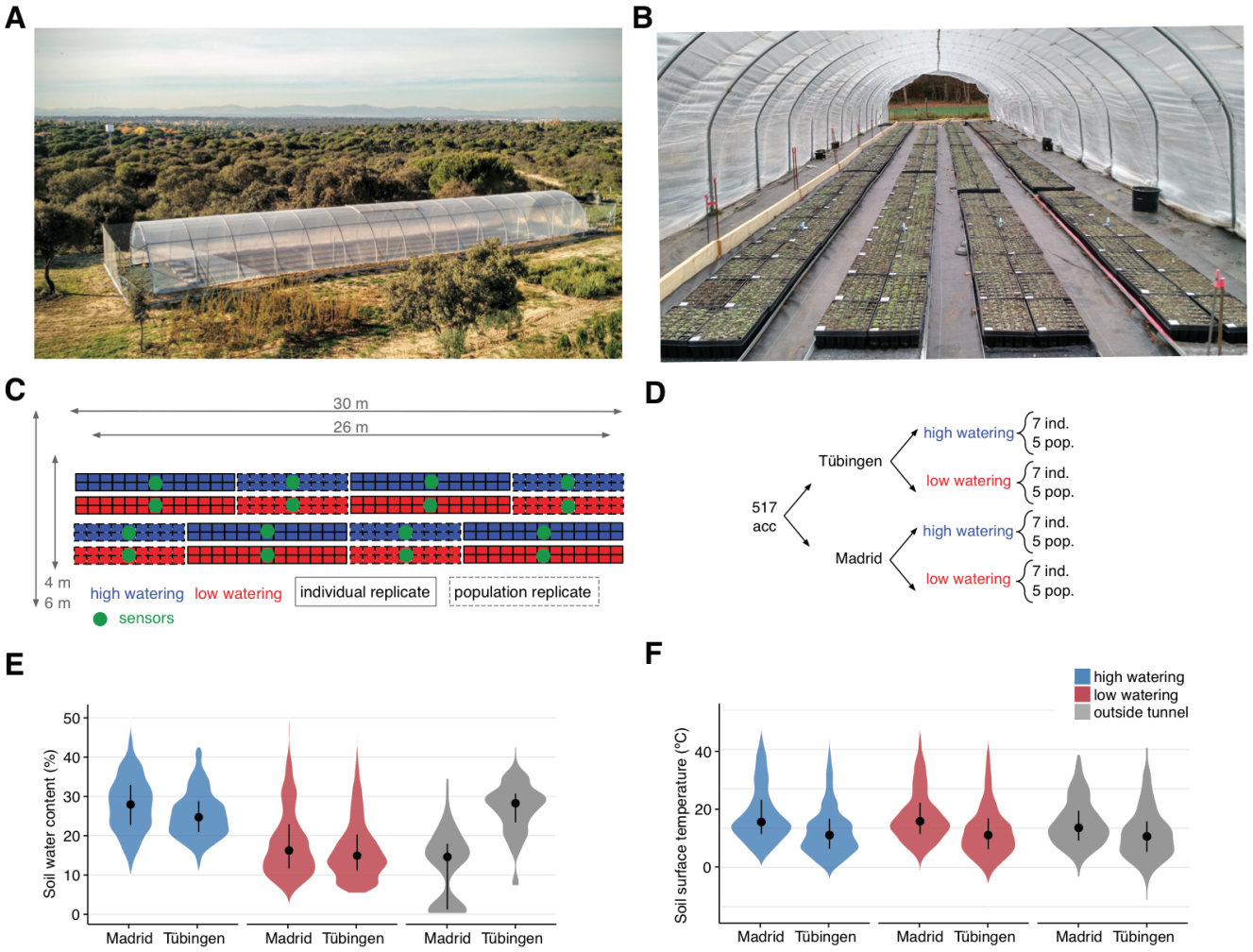
Figure 1. Geographic distribution of accessions



Locations of *Arabidopsis thaliana* accessions used in this experiment (red), 1001G accessions (blue), and all observations of the species in gbif.org (grey).

298

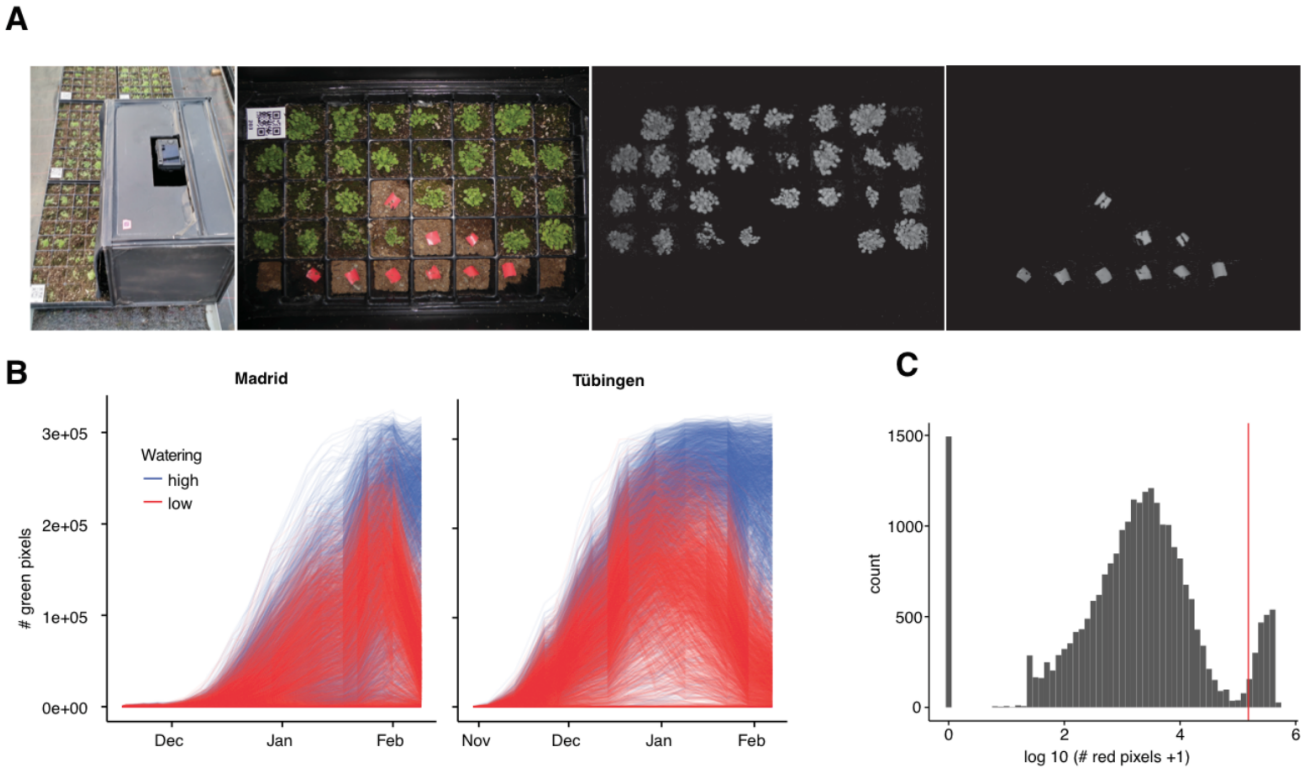
Figure 2. Field experiment design



299 (A) Aerial picture of foil tunnel settings in Madrid and (B) photo inside the foil tunnel in Tübingen. (C)
300 Spatial distribution of blocks and replicates and (D) experimental design. (E) Soil water content from the 34
301 sensors monitoring each experimental block and conditions outside the tunnel.

302

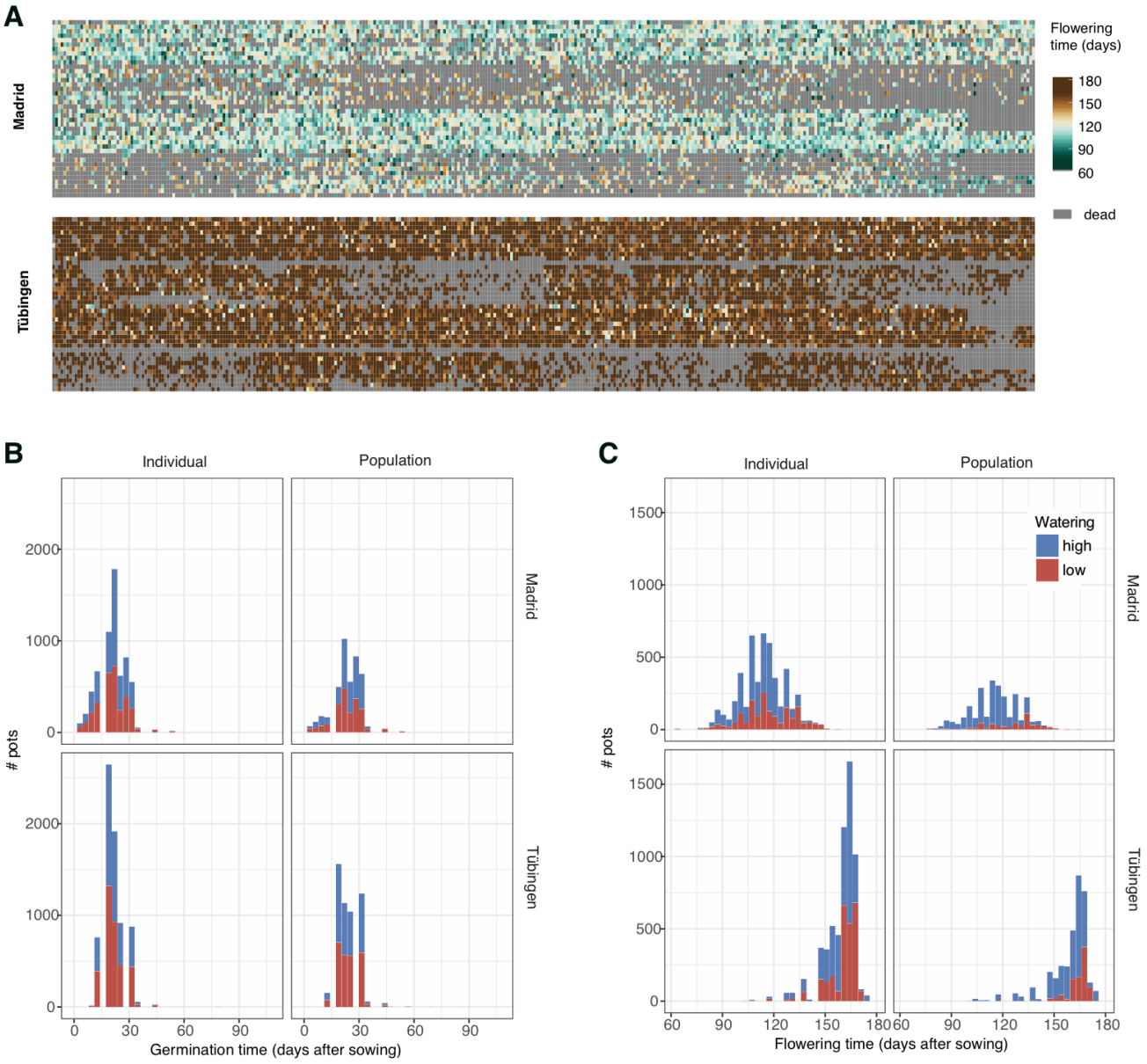
Figure 3. Rosette monitoring



303 (A) Customized dark box for image acquisition and example tray with the corresponding green and red
304 segmentation. (B) Trajectories of number of green pixels per pot, indicating rosette area, for Madrid and
305 Tübingen. (C) Distribution of the sum of red pixels per pot over all time frames. The red vertical line
306 indicates the heuristically chosen threshold to define whether the pot actually had a red label.

307

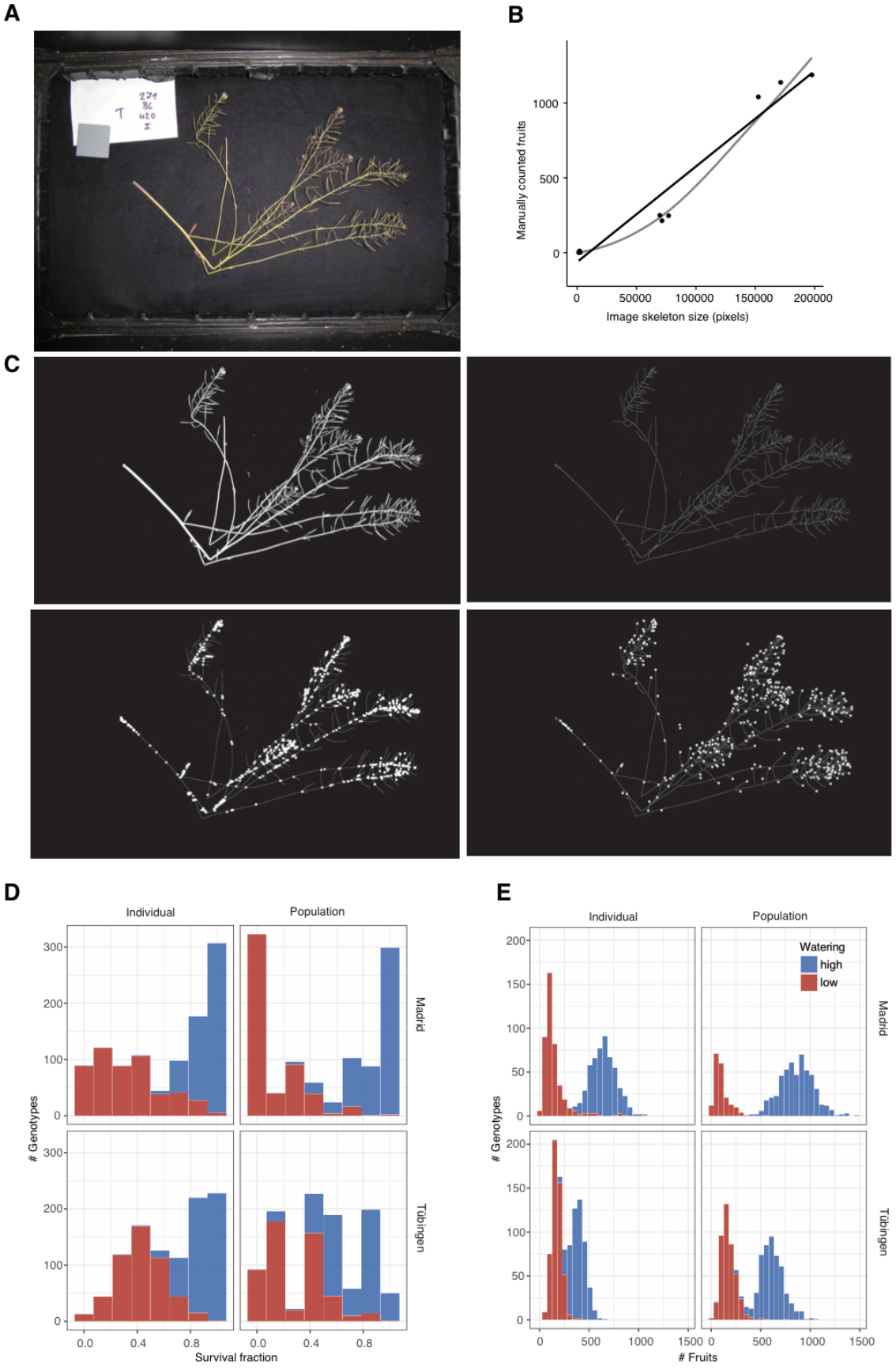
Figure 4. Flowering time distributions



308 (A) Flowering times per pot in the same spatial arrangement as in each tunnel (see Fig. 3). (B) Distribution
309 of germination times. (C) Distribution of flowering times.

310

Figure 5. Inflorescence and seed set estimation



(A) Representative inflorescence picture and the resulting variables from image processing (B): total segmented area (upper-left), skeletonized inflorescence (upper-right), branching points (lower-left), and endpoints (lower-right). (C) Regression between the fruits of few manually counted inflorescences and the inflorescence size. The four variables inferred in (B) accurately predicted the manually counted inflorescences as example (R^2=0.97, n=11, $p=10^{-4}$). Distribution of survival to reproduction (D) and fruits per plant (E) in the four environment treatments.