Image-based methods for phenotyping growth dynamics and fitness in large plant populations

François Vasseur[1,2], George Wang[1], Justine Bresson[3], Rebecca Schwab[1] and Detlef Weigel[1*]

[1]Max Planck Institute for Developmental Biology, D-72076 Tübingen, Germany

[2]Current address: CNRS, UMR5175, Centre d'Ecologie Fonctionnelle et Evolutive, F-34000 Montpellier, France

[3]Center for Plant Molecular Biology (ZMBP), General Genetics, University of Tübingen, D-72076 Tübingen, Germany.

Email addresses:

FV: franc.vasseur@gmail.com

GW: biogeek@gmail.com

JB: justine.bresson@zmbp.uni-tuebingen.de

RS: rebecca.schwab@tuebingen.mpg.de

DW*: weigel@weigelworld.org

*Corresponding author.

Key words: biomass accumulation, fruit production, growth modelling, high-throughput, hydroponic system, ImageJ, logistic function, R codes, relative growth rate.

Word count: Background (464), Results (1971), Discussion (656), Methods (826), Acknowledgements (28), **Total** (3917).

7 Figures (Fig. 1, 2, 4 and 5 in color)

1 Table

Additional files: 5 Tables, 3 Figures and 4 Additional Files

Abstract

Background. With the development of next-generation sequencing technologies, high-throughput phenotyping has become the new bottleneck of quantitative genetics analyses. The model species *Arabidopsis thaliana* offers extensive resources to investigate intraspecific trait variability and the genetic bases of ecologically relevant traits, such as growth dynamics and reproductive allocation. However, reproducible and cost-effective methods need to be developed for the measurement of growth and especially fitness related traits in large populations. Here we describe image-based methods that can be adapted to a wide range of laboratory conditions, and enable the reliable estimation of biomass accumulation and fruit production in thousands of *A. thaliana* individuals.

Results. We propose a semi-invasive approach, where part of a population is used to predict plant biomass from image analysis. The other part of the population is daily imaged during three weeks, then harvested at the end of the life cycle where rosette and inflorescence are separately imaged. We developed ImageJ macros and R codes for image segmentation, 2D skeletonization and subsequent statistical analysis. First, ontogenetic growth is modelled from estimated and measured dry mass for all individuals with non-linear regressions, from which the dynamics of absolute growth rate (GR) and relative growth rate (RGR) are calculated. Second, analysis of the 2D inflorescence skeleton allows the estimation of fruit production, an important component of plant fitness. Our method was evaluated across 451 natural accessions of *A. thaliana*. Cross-validation revealed that our image-based method allows predicting approximately 90% of biomass variation and 70% of fruit production. Furthermore, estimated traits - like measured traits - showed high heritabilities and inter-experiment reproducibility.

Conclusions. We propose a flexible toolkit for the measurement of growth and fitness related traits in large plant populations. It is based on simple imaging, making the method

reproducible at low cost in different facilities. However, as manual imaging of large plant populations can quickly become a limiting factor, we also describe an automated high-throughput imaging coupled with micro-computers that enables large phenotypic screening for genome-wide association studies and stress experiments.

Background

Plants become less efficient to accumulate biomass as they get larger and older, resulting in a decline of the relative growth rate (RGR) in the course of ontogeny [1–3]. This is due to developmental constraints such as increasing allocation of biomass to supporting structures like stems, and self-shading due overlapping leaves in larger individuals. Strong variation of growth has been reported across plant species [1,4–11], which has been assumed to reflect the inherent diversity of plant strategies to cope with contrasting levels of resource availability [1,5,8,12,13]. For instance, species from resource-scarce environments generally show a reduced speed of biomass accumulation compared to species from resource-rich environments, even when they are placed in the same, non-limiting conditions [2,14]. The same observation has been made within species [5,15]. However, the genetic bases of plant growth dynamics remain unclear, due to the complexity of the genetic architecture of underlying traits [16,17], and thus, to the necessity of phenotyping very large plant populations.

To compare growth dynamics between genotypes or treatments, several methods have been proposed to compute RGR from a small numbers of observations [7,18,19]. These methods are valuable to investigate changes in growth over a short period of time, for instance in response to environmental stress. However, sequential harvesting of many individuals during ontogeny is space and time consuming. Furthermore, it is problematic for evaluating standard deviation and measurement error across genotypes, as the same individual is not followed during ontogeny. Ecophysiological and ecological studies

ultimately seek to compare individual variation of growth dynamics with fitness components like fruit and seed production. The model species *A. thaliana* is a valuable tool to make such comparison [20]. Because of its selfing mating system and its small size, numerous individuals can be grown in relatively small facilities, and experiments are reproducible indefinitely on the same genotypes [21]. In 2016, the genetic resources available for this species have reached a milestone with the complete sequencing of 1,135 natural accessions with publicly available seeds to conduct comparative analysis (http://1001genomes.org/) [22].

Many high-throughput platforms have been developed in the last decades for plant phenotyping in crops [23–27], and in *A. thaliana* [10,28–33]. However, most of them are (i) either commercial or require the purchase of expensive materials [23,25,27–29,31] but see [34], (ii) not suitable for the analysis of growth phenotypes from germination to fitness parameters [30,32,34,35], or (iii) do not allow the quantification of complex traits like biomass accumulation and RGR [30,36]. Thus, we aimed at developing flexible and easily reproducible methods to measure individual growth dynamics and fruit production across large populations of *A. thaliana*. Our pipeline is based on manual or automated imaging, and subsequent analysis with ImageJ [37] and R [38]. The method was evaluated on a set of 451 worldwide accessions.

Results

Analysis of growth dynamics in large populations with simple imaging

Plant imaging and population sub-sampling

In our study, we used a hydroponic system where plants are cultivated in trays on inorganic solid media (rockwool cubes) and nutrients are daily provided through a solution [39] (Fig. 1a; Table S1). This system allows plants to have an unlimited source of nutrients for growth, as well as a highly controlled environment to avoid pests and herbivores.

Furthermore, the hydroponic system also enables the measurement of the weight of the root system.

The individual plant growth was followed using a code system: the genotype sowed in each pot was defined before the beginning of the experiment by the label of the tray (*e.g.*, T01, T02, etc.) and the coordinate of the pot in each tray (*e.g.*, A1, C5, D8, etc.). The label of the tray was placed at the top-left corner, which defined the coordinate of the first position (A1) on each tray (Fig. 1b). Four replicates of 480 genotypes were grown, with pots randomly distributed in 64 trays of 30 pots each (Fig. 1b). After germination at 23 °C, all plants, having two cotyledons, were vernalized at 4 °C during 41 days to maximize the flowering of all the different accessions. Due to mortality, only 451 accessions out of the 472 were completely phenotyped for growth and fitness. After shoot harvesting, rockwool cubes were dried at 65 °C and weighed to estimate root dry mass.

A growth chamber equipped with Raspberry Pi Automated Phenotyping Array (hereafter RAPA) was built at the Max Planck Institute (MPI) of Tübingen for large-scale phenotyping. 192 micro-cameras (two per tray) were mounted between fluorescent light tubes to simultaneously take top-view pictures of all plants (Fig. 2). The recording and storage of tray pictures were managed through embedded computers (Raspberry Pi rev. 1.2, Raspberry Pi Foundation, UK). All trays were daily imaged until fruit ripening with the automatic RAPA system. In addition, all trays were daily imaged with a manual setup (left side on Fig. 2a) during the first 25 days after vernalization. To show the general applicability of the method, here we used only images taken manually from the regular camera.

For modelling growth, we partitioned the whole population into two sub-populations: replicates 1 and 2 of every genotype represent the *focal* population on which growth dynamics and fruit production were measured. They were harvested at the end of reproduction when the first fruits (siliques) were yellowing (stage 8.00 according to Boyes

et al. [40]). Inflorescences and rosettes were separated and both photographed with a high-resolution camera (Fig. 1b). Replicates 3 and 4 represent the *training* population. They were harvested at 16 days after the end of vernalization for building a predictive model of rosette biomass from images.

Estimation of rosette biomass from image analysis and modelling in the training population

We developed an ImageJ [37] macro to extract rosette morphological traits of all growing individuals from a tray image (Fig. 2b, c; Additional File 1). After automatic segmentation of the tray image (Fig. 2b), the macro returns individual rosette area and perimeter values in pixels, which can be then converted into cm^2 and cm, respectively, by measuring the area and perimeter (in pixels) of a 4 cm^2 calibrator present on trays. Three other morphological traits were also automatically measured with the macro: the rosette circularity ($cir = 4\pi \times \frac{Area}{Perimeter^2}$), the aspect ratio ($AR = \frac{Major\ axis\ length}{Major\ axis\ length}$), and the rosette roundness ($round = \frac{4 \times Area}{\pi \times (Major\ axis\ lenght)^2}$).

The macro automatically processes images when plants are young and do not overlap with each other. When they get bigger (*i.e.* 20-25 days after vernalization), the macro offers the possibility to manually separate plants. We estimated that, on a regular computer, the macro takes approximately 20-25 s per tray (30 individuals) when running on automatic mode, and between one and two minutes in manual mode (depending on the number of corrections to make). All individual rosette images extracted after segmentation are saved in .jpg format (Fig. 2c), called with the corresponding date, tray label and pot coordinates (for instance “20150509_Tray12_C5.jpg”). The ImageJ code of the macro is available in Additional File 1, with an example of tray image in Fig. S1.

147 We then developed a predictive model of rosette dry mass from the five traits
148 extracted from rosette images with ImageJ. Generalized linear models (*glm*) were
149 evaluated with a cross-validation approach on the *training* population (Additional File 2).
150 Prediction accuracy was measured as the coefficient of correlation (r^2) between measured
151 and predicted rosette dry mass. Results showed that we can explain *ca*. 90% ($P < 0.001$) of
152 rosette dry mass with the model that included rosette area, perimeter, circularity, aspect
153 ratio and roundness. However, as removing rosette circularity, aspect ratio and roundness
154 did not impact the prediction accuracy ($r^2 = 0.91$ when replicate 3 is used to train the
155 model, Fig. 3a; and $r^2 = 0.90$ when replicate 4 is used to train the model, Fig. 3b), these
156 traits were therefore removed in subsequent analyses. Thus, the final equation we used to
157 estimate rosette dry mass from rosette pictures, in the 25 first days following vernalization,
158 is given by Eq. 1 of Table 1, where the three coefficients (given in Table S2) were all
159 significant ($P = 0.03$ for the intercept, $P < 10^{-4}$ for both the coefficients of rosette area and
160 rosette perimeter).

161 *Fitting non-linear model of individual growth dynamics in the* focal *population*
162 Since some plants germinated during or, for a few, after vernalization, we considered the
163 first day of growth (t_0) for each individual of the *focal* population (replicates 1 and 2) as the
164 day at which it had a minimum size. For convenience, we used the size of the biggest
165 measured plant across all individuals at the first day of growth following vernalization.
166 This corresponded to a plant with first true leaves just emerged (2-3 mm each, Additional
167 File 2). This procedure allowed for normalization of growth trajectories from the same
168 starting point between individuals that differ in germination speed. We used the predictive
169 model established above on the training population to estimate rosette dry mass on the
170 focal population during the 25 first days following vernalization (Fig. 4).

Previous studies have shown that leaf and whole-plant growth follow a sigmoid curve in *A. thaliana* [41–43]. We thus used a three-order logistic function *f* of rosette dry mass (*M*) in the course of time *t* (Eq. 2 in Table 1) [3] as:

$$M(t) = \frac{A}{1 + e^{\frac{t_i - t}{B}}}$$

where A, B and t_i are the parameters characterizing the shape of the curve, which are expected to differ between individuals depending on genotypes and/or environmental conditions. A is the upper asymptote of the sigmoid curve, which was measured as rosette dry mass (mg) at maturity (Fig. 4). B controls the steepness of the curve, as the inverse of the exponential growth coefficient r ($r = 1/B$). t_i is the inflection point that, by definition, corresponds to the time point where the rosette is at 50% of its final dry mass, and where the absolute growth rate (GR, mg d^{-1}) is maximal [3]. Both B and t_i were estimated by fitting a logistic growth function on every individual in R (Table S2; Additional File 2). We then estimated several dynamical variables from the fitted parameters (Fig. 4; Table 1; Table S2), such as absolute growth rate ($GR(t)$, the derivative of the logistic growth function, mg d^{-1}; Eq. 3 in Table 1) and $RGR(t)$ ($GR(t) / M(t)$, mg d^{-1} g^{-1}; Eq. 4 in Table 1).

Image-based estimation of fruit production

Fruits (siliques) were manually counted on the pictures of the inflorescence of 352 out of the 856 plants harvested in the *focal* population. In addition, we analysed the inflorescence features of all the 856 harvested plants with ImageJ [37] to automatically estimate the number of siliques through image segmentation and 2D skeletonization with a dedicated macro (code in Additional File 3, and example of inflorescence image in Fig. S2). The analysis of the inflorescence 2D skeletons with ImageJ (Fig. 5) returned seven vectors of variables for each plant, which were automatically saved in a .csv file by the macro. These skeleton variables are: number of branches (sk_1), junctions (sk_2), end-point voxels (sk_3),

194 junction voxels (sk_4), slab voxels (sk_5), triple points (sk_6), and maximum branch length
195 (sk_7). These seven vectors were averaged per individual, giving a set of seven variables
196 with unique value that were used as predictors of silique number.

197 Using the same approach as for estimating rosette dry mass from images, we
198 estimated silique number from a cross-validation approach with *glm* performed on the 352
199 plants for which we had both manual measurements and skeleton parameters (detailed R
200 code in Additional File 4). The cross-validation was tested across 1000 permutations of a
201 random training dataset of 264 individuals, *i.e.* 75% of the 352 individuals. It showed that
202 the predicted and measured silique number had a high correlation with r^2 close to 0.7 (Fig.
203 5b and Additional File 4). We then used all the 352 plants as training dataset to predict the
204 number of siliques in all harvested individuals of the focal population. The parameters of
205 the equation to estimate the number of siliques (S) from ImageJ-processed inflorescence
206 images (Eq. 5 in Table 1) are given in Table S2.

207 **Robustness of the method and reproducibility of trait measurement**

208 Our analysis of trait variation focuses on eight phenotypic traits: five were measured
209 (growth duration, final rosette and root dry mass, root and reproductive allocation), and
210 three were estimated (GR and RGR at the inflection point, number of siliques). Trait values
211 for all individuals are available in Table S3. They were all strongly variable across the 451
212 accessions (Fig. 6). The part of phenotypic variance accounted by genetic variation across
213 individuals (broad-sense heritabilities, H^2) for all traits were higher than 0.70, except for
214 root allocation, a measured trait, with a H^2 of 0.40 (Table S4). Importantly, H^2 of the traits
215 estimated with our method (GR and RGR at inflection, and number of siliques) were in the
216 same range (0.91, 0.75 and 0.76, respectively) than for the heritabilities of the measured
217 traits (0.89, 0.95, 0.71, 0.40 and 0.84 for growth duration, rosette final dry mass, root final
218 dry mass, root and reproductive allocation, respectively; see Table S4).

To evaluate the reproducibility of trait value estimated across genotypes with our method, we repeated a second experiment on 18 accessions selected for their highly contrasted phenotypes (Fig. S3). We used nine phenotypic groups, each containing two accessions, spanning the range of growth duration and RGR variation (three groups of growth duration and three groups of RGR, see Table S5 and Fig. S3). These 18 accessions were grown using the same protocol (n = 8). Three replicates were harvested at the estimated inflection point of the growth trajectory of each genotype, and five replicates were harvested at the end of reproduction. Using the inflection point of the growth trajectory (*i.e.* when the rosette is estimated to be 50% of its final dry mass) had two goals: (i) test whether predicted rosette dry mass at later vegetative growth is consistent with measurement, and (ii) test whether the genotypic trait values are correlated between two independent experiments. In additions, the number of siliques was manually counted on the plants at the end of reproduction (stage 8.00 [40]).

Results showed a good correlation between the rosette dry mass at the inflection point estimated in the first experiment and the dry mass measured in the second experiment (r^2 = 0.67; Fig. 7a). This suggests that (i) the method is suitable for estimating rosette dry mass in the course of the of the whole life cycle, and (ii) model predictions are conserved across experiments. Furthermore, the number of siliques estimated from image analysis in the first experiment was well correlated with the number of siliques manually counted in the second experiment (r^2 = 0.70; Fig. 7b).

Discussion

Biomass accumulation is a key physiological parameter related to individual performance and stress response [1,44]. Here, we have developed a set of methods based on high-throughput image analysis for a rapid and accurate estimation of growth variation, including plant size, growth rate and RGR. The method uses low-cost techniques robustly

applicable in any facilities (e.g., growth chambers, greenhouses). Our method developed for *A. thaliana* has the advantage of being cost-effective and flexible to any facilities. From simple imaging, it allows a single experimenter to simultaneously measure rosette biomass accumulation, RGR and fruit production over thousands of plants. Importantly, all the growth parameters measured in this study varied strongly between individuals, notably between early and late flowering plants (Fig. 4, Table S3). *A. thaliana* is the favorite model in plant genetics and molecular biology, and it is also becoming a model in evolutionary biology and ecology [9,21,45]. Our approach therefore offers new avenues for dissecting the links between growth and fitness, and the genetic and evolutionary bases of phenotypic variation. The same method could be applicable to other species, but the 3D architecture in non-rosette species requires more sophisticated image analysis. A recent study in maize offers a nice example of high-throughput 3D reconstruction of plants and biomass prediction [25].

Technical considerations about high-throughput imaging and analysis

In this study, we took advantage of the RAPA system to collect plant images at high-throughput during growth. Additionally, we took pictures of the trays with a regular camera every day when we watered the plants. Measurements of rosette morphology (area, perimeter and shape descriptors) were the same on manual images and RAPA images. Thus, tray imaging can be done easily in any laboratory or facilities, and analyzed with the ImageJ macro developed in this study (Additional File 1). The time lapse and frequency of tray imaging is important for proper fitting of the growth curves. We used daily imaging during the 25 first days of growth after vernalization, but a logistic growth function could be fitted with only one picture every 2-3 days. Although less accurate, reducing the frequency of imaging would help the experimenter to save time during both image acquisition and analysis, notably if no automatic system like RAPA is available. Here we

showed that our method is robust and reproducible across experiments as long as we grow the plants in the same environment. However, we recommend making a new predictive model of plant biomass with cross-validation for each experimental design, specifically if growth conditions change, as the relationship between rosette morphology and rosette biomass may differ depending on genotypes and environments.

Reproducible and low-cost methods for genetic and ecophysiological studies

The semi-invasive approach based on statistical modelling and cross-validation to convert rosette morphology into biomass drastically reduces the number of replicates necessary to measure the dynamics of biomass accumulation. Here, we presented the analysis of eight growth trait across 451 accessions, for which we measured high heritabilities using only four replicates per genotype (Table S4). The reproducibility of our method is supported by the second experiment where 18 contrasted accessions were phenotyped in the same condition (Fig. 7). Importantly, we voluntarily chose 18 accessions that maximize differences in growth duration and RGR to reproduce the most extreme phenotypes (Fig. S3).

Importantly, our methods offer the possibility to evaluate the relationship between growth dynamics and fitness components on the same individuals. This enables high-throughput investigations about the mechanisms of plant acclimation to different conditions, as well as its genetic underpinnings (*e.g.* genotypes-by-environment interactions). Measuring growth dynamics and fitness components on large plant populations is a strong bottleneck for our understanding of the mechanisms of plant adaptation and evolution. The same approach we have developed here can easily be reproduced in any laboratory working with *A. thaliana*, and potentially with other species. Taking advantage of its flexibility, we hope that our method can help future studies to dissect the relationships between genetic diversity, environment and plant performance.

Methods

Plant material

472 natural accessions of *A. thaliana* were selected from the germplasm of the 1001 Genomes project [22] (http://1001genomes.org/; a list of the accessions used in this study is available in Table S4) for monitoring vegetative growth and fitness in the RAPA system at MPI-Tübingen. Seeds used in this study were obtained from parental plants propagated under similar conditions in greenhouse. All seeds were stored overnight at -80 °C and surface-sterilized with 100% ethanol before sowing.

Growth conditions

Plants were cultivated on a semi-hydroponic system. Seeds were sowed on 3.6 cm x 3.6 cm x 4 cm depth rockwool cubes (Grodan cubes, Rockwool International A/S, Denmark) fitted in 4.6 cm (diameter) x 5 cm (depth) circular pots (Pöppelmann GmbH and Co., Germany). The pots were covered with a black circle pierced in the centre (5-10 mm hole manually made with a puncher). Before sowing, the dry rockwool cubes were weighed with a microbalance. Then, they were watered by dipping them completely into 75% strength nutrient solution in order to achieve full humidification and fertilization. The chemical composition of the nutrient solution was obtained from Conn *et al*. [39] (also available in Table S2).

After sowing, trays were incubated for two days in the dark at 4 °C for seed stratification, and then transferred for six days to 23 °C for germination. After six days, when most seedlings had two cotyledons, trays were transferred to 4 °C for 41 days for vernalization. During germination and vernalization, all trays were watered once a week by putting them in 4-6 cm (*ca.* three-quarter of the pot height) of a 75% strength nutrient solution for 5-10 s. After 41 days of vernalization, when true leaves had emerged on most individuals, plants were thinned to one plant per pot, and trays were moved to the RAPA

room, set to 16 °C with a temperature variability of close to ± 0.1 °C, air humidity at 65 %, and 12 h day length, with a PPFD of 125 to 175 µmol m^{-2} s^{-1} provided by a 1:1 mixture of Cool White and Gro-Lux Wide Spectrum fluorescent lights (Luxline plus F36W/840, Sylvania, Germany). Light, temperature, and humidity were continuously monitored online and logged data were stored in a Structured Query Language (SQL) database. All trays were randomly positioned in the room, and watered every day with 100% strength nutrient solution.

The second experiment for validating the method was performed on a set of 18 selected accessions (Table S6, Fig. S1), grown in the same conditions. Plants were harvested at the estimated inflection point for rosette dry mass measurement (inflection point estimated from the first experiment, $n = 3$), and six replicates were harvested at the end of the life cycle for manual silique counting ($n = 5$).

Plant imaging and harvesting

All plants were daily imaged with 192 micro-cameras (5 Megapixel, OmniVision OV5647), two per tray, mounted between fluorescent light tubes in the RAPA room at MPI Tübingen. The recording and storage of tray pictures were managed through embedded computers (Raspberry Pi rev. 1.2, Raspberry Pi Foundation, UK). In addition, all trays were manually imaged every day during watering with a high-resolution camera (16.6 Megapixels, Fig. 1; Canon EOS-1, Canon Inc., Japan).

For the focal population harvested at fruit ripening, inflorescences and rosettes were separated and both separately photographed with a high-resolution camera (16.6 Megapixels, Canon EOS-1). The duration of growth (days) was estimated as the time in days between t_0 and harvesting. The rosettes, inflorescences and rockwool cubes that contain the roots were dried for at least three days at 65 °C, and weighed separately with a microbalance (XA52/2X, A. Rauch GmbH, Graz, Austria). Root dry mass of each

344 individual was estimated as the difference in the weight of the dry rockwool cube before
345 and after plant growth. Root and reproductive allocation were measured as the ratio of root
346 and inflorescence dry mass, respectively, over total (rosette + roots + inflorescence) plant
347 dry mass. For the training population harvested at 16 days after vernalization, rosette were
348 dried at 65 °C for three days, and separately weighed with a microbalance (XA52/2X).

349 **Statistical analyses**

350 All coefficients of correlation (r^2) used to estimate the prediction accuracy of our cross-
351 validation approaches were obtained from Pearson's coefficients between estimated and
352 measured data, using the *cor.test* function in R. Non-linear fitting of the logistic growth
353 functions (Eq. 2 in Table 1) were performed on every individual with the *nls* function in R
354 (detailed code in Additional File 2). Broad-sense heritabilities (H^2) were calculated with a
355 Bayesian approach implemented in a *MCMCglmm* model performed in R, considering the
356 accession as a random factor, as:

357 $$y_{ik} = G_i + e_{ik}$$

358 where y is any trait of interest after logarithmic transformation for the individual k of
359 genotype i, G_i is the identifier of the accession i, and e_{ik} is the residual error associated with
360 every individual. H^2 was calculated at the proportion of genotypic variance (σ^2_G) over
361 total variance ($\sigma^2_G+\sigma^2_e$):

$$H^2 = \frac{\sigma^2_G}{\sigma^2_G+\sigma^2_e}$$

Declarations

List of abbreviations

- t_0: first day of growth used for logistic fitting, when plants have two leaves of 3 mm^2 (see Additional File 2)
- t_i: inflection point (in days) of the logistic growth curve (corresponds to the point where the rosette is half it final dry mass, and where absolute growth rate is maximum)
- A: upper asymptote of the logistic growth curve (et as the rosette dry mass measured at the end of reproduction)
- B: inverse of the exponential constant of the logistic growth curve
- M: rosette dry mass (in mg)
- GR: absolute growth rate (in mg d^{-1}) measured by deriving the logistic equation of growth (Eq. 1 in Table 1)
- RGR: relative growth rate (mg d^{-1} g^{-1}) measured as the ratio between GR and M for each time point throughout ontogeny.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The datasets supporting the conclusions of this article are included within the article and its additional files. Correspondence and requests for materials should be addressed to weigel@weigelworld.org or franc.vasseur@gmail.com.

Competing interests

The authors declare no competing financial interests.

Funding

This work was supported by an ERC grant (ERC-AdG-IMMUNEMESIS, DW), and the Max Planck Society (FV, MEA, GW, DW). JB was funded by Alexander von Humboldt Foundation.

Author Contributions

FV, GW, JB, RS and DW designed the study. FV, RS and GW performed the experiments and extracted the data. FV, GW and JB performed image and statistical analyses, and wrote codes. All authors interpreted the results and wrote the paper.

Acknowledgements

We thank Jane Devos for her help in performing the experiments, and Denis Vile, Dominik Grimm and Moises Exposito-Alonso for helpful comments about image analysis and statistical modelling.

References

1. Poorter H, Garnier E. Ecological significance of inherent variation in relative growth rate and its components. Handbook of functional plant ecology. New York: Marcel Dekker; 1999;20:81–120.

2. Poorter H, Remkes C. Leaf area ratio and net assimilation rate of 24 wild species differing in relative growth rate. Oecologia. Springer-Verlag; 1990;83:553–9.

3. Paine CE, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, et al. How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. Methods Ecol. Evol. 2012;3:245–56.

4. Niklas KJ, Enquist BJ. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proceedings of the National Academy of Sciences. 2001;98:2922–7.

5. Poorter H, Van Rijn CPE, Vanhala TK, Verhoeven KJF, De Jong YEM, Stam P, et al. A genetic analysis of relative growth rate and underlying components in Hordeum spontaneum. Oecologia. 2005;142:360–77.

6. Sadok W, Naudin P, Boussuge B, Muller B, Welcker C, Tardieu F. Leaf growth rate per unit thermal time follows QTL-dependent daily patterns in hundreds of maize lines under naturally fluctuating conditions. Plant Cell Environ. 2006;30:135–46.

7. Paul-Victor C, Züst T, Rees M, Kliebenstein DJ, Turnbull LA. A new method for measuring relative growth rate can uncover the costs of defensive compounds in *Arabidopsis thaliana*. New Phytol. 2010;187:1102–11.

8. Vile D, Pervent M, Belluau M, Vasseur F, Bresson J, Muller B, et al. Arabidopsis growth under prolonged high temperature and water deficit: independent or interactive effects? Plant Cell Environ. 2012;35:702–18.

9. Vasseur F, Violle C, Enquist BJ, Granier C, Vile D. A common genetic basis to the origin of the leaf economics spectrum and metabolic scaling allometry. Ecol. Lett. 2012;15:1149–57.

10. Vasseur F, Bontpart T, Dauzat M, Granier C, Vile D. Multivariate genetic analysis of plant responses to water deficit and high temperature revealed contrasting adaptive strategies. J. Exp. Bot. 2014;65:6457–69.

11. Nunes-Nesi A, Nascimento V de L, de Oliveira Silva FM, Zsögön A, Araújo WL, Sulpice R. Natural genetic variation for morphological and molecular determinants of plant growth and yield. J. Exp. Bot. [Internet]. Soc Experiment Biol; 2016; Available from: http://dx.doi.org/10.1093/jxb/erw124

12. Grime JP. Plant strategies and the dynamics and structure of plant-communities. Nature. 1988;336:630–630.

13. Vergeer P, van den Berg LL, Bulling MT, Ashmore MR, Kunin WE. Geographical variation in the response to nitrogen deposition in Arabidopsis lyrata petraea. New Phytol. 2008;179:129–41.

14. Grime JP, Hunt R. Relative Growth-Rate: Its Range and Adaptive Significance in a Local Flora. J. Ecol. [Wiley, British Ecological Society]; 1975;63:393–422.

15. Li B, Suzuki J-I, Hara T. Latitudinal variation in plant size and relative growth rate in Arabidopsis thaliana. Oecologia. Springer; 1998;115:293–301.

16. Kroymann J, Mitchell-Olds T. Epistasis and balanced polymorphism influencing complex trait variation. Nature. 2005;435:95–8.

17. Rockman MV. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution. 2012;66:1–17.

18. Hoffmann WA, Poorter H. Avoiding bias in calculations of relative growth rate. Ann. Bot. Annals Botany Co; 2002;90:37–42.

19. Rees M, Osborne CP, Woodward FI, Hulme SP, Turnbull LA, Taylor SH. Partitioning the components of relative growth rate: how important is plant size variation? Am. Nat. JSTOR; 2010;176:E152–61.

20. Colautti RI, Lee C-R, Mitchell-Olds T. Origin, fate, and architecture of ecologically relevant genetic variation. Curr. Opin. Plant Biol. 2012;15:199–204.

21. Weigel D. Natural Variation in Arabidopsis: From Molecular Genetics to Ecological Genomics. Plant Physiol. 2012;158:2–22.

22. The 1001 Genomes Consortium. 1135 sequenced natural inbred lines reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell. 2016; Available from: http://dx.doi.org/10.1016/j.cell.2016.05.063

23. Chen D, Neumann K, Friedel S, Kilian B, Chen M, Altmann T, et al. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. Plant Cell. Am Soc Plant Biol; 2014;26:4636–55.

24. Mielewczik M, Friedli M, Kirchgessner N, Walter A. Diel leaf growth of soybean: a novel method to analyze two-dimensional leaf expansion in high temporal resolution based on a marker tracking approach (Martrack Leaf). Plant Methods. plantmethods.biomedcentral.com; 2013;9:1–14.

25. Cabrera-Bosquet L, Fournier C, Brichet N, Welcker C, Suard B, Tardieu F. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. New Phytol. [Internet]. Wiley Online Library; 2016; Available from: http://onlinelibrary.wiley.com/doi/10.1111/nph.14027/full

26. Andrade-Sanchez P, Gore MA, Heun JT, Thorp KR, Carmo-Silva AE, French AN, et al. Development and evaluation of a field-based high-throughput phenotyping platform. Funct. Plant Biol. CSIRO; 2013;41:68–79.

27. Honsdorf N, March TJ, Berger B, Tester M, Pillen K. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. PLoS One. journals.plos.org; 2014;9:e97047.

28. Granier C, Aguirrezabal L, Chenu K, Cookson SJ, Dauzat M, Hamard P, et al. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. New Phytol. 2006;169:623–35.

29. Arvidsson S, Pérez-Rodríguez P, Mueller-Roeber B. A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. New Phytol. 2011;191:895–907.

30. Jansen M, Gilmer F, Biskup B, Nagel KA, Rascher U, Fischbach A, et al. Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in Arabidopsis thaliana and other rosette plants. Funct. Plant Biol. CSIRO; 2009;36:902–14.

31. Tisne S, Serrand Y, Bach L, Gilbault E, Ben Ameur R, Balasse H, et al. Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. Plant J. Wiley Online Library; 2013;74:534–44.

32. Dhondt S, Gonzalez N, Blomme J, De Milde L, Van Daele T, Van Akoleyen D, et al. High-resolution time-resolved imaging of in vitro Arabidopsis rosette growth. Plant J. Wiley Online Library; 2014;80:172–84.

33. Bresson J, Vasseur F, Dauzat M, Koch G, Granier C, Vile D. Quantifying spatial heterogeneity of chlorophyll fluorescence during plant growth and in response to water stress. Plant Methods. plantmethods.biomedcentral.com; 2015;11:23.

34. Minervini M, Giuffrida MV, Perata P, Tsaftaris SA. Phenotiki: an open software and hardware platform for affordable and easy image-based phenotyping of rosette-shaped plants. Plant J. Wiley Online Library; 2017;90:204–16.

35. Busch W, Moore BT, Martsberger B, Mace DL, Twigg RW, Jung J, et al. A microfluidic device and computational platform for high-throughput live imaging of gene expression. Nat. Methods. nature.com; 2012;9:1101–6.

36. Zhang X, Hause RJ Jr, Borevitz JO. Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in *Arabidopsis thaliana*. G3. g3journal.org; 2012;2:29–34.

37. Rasband WS. 1997–2011. ImageJ. US National Institutes of Health, Bethdesda, MD, USA: http://imagej. nih. gov/ij. 2011;

501 38. R Development Core Team. R: a language and environment for statistical computing. Vienna,
502 Austria: R Foundation for Statistical Computing; 2009.

503 39. Conn SJ, Hocking B, Dayod M, Xu B, Athman A, Henderson S, et al. Protocol: optimising
504 hydroponic growth systems for nutritional and physiological analysis of Arabidopsis thaliana and
505 other plants. Plant Methods. 2013;9:4.

506 40. Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, et al. Growth stage--
507 based phenotypic analysis of Arabidopsis a model for high throughput functional genomics in
508 plants. Plant Cell. Am Soc Plant Biol; 2001;13:1499–510.

509 41. Cookson SJ, Van Lijsebettens M, Granier C. Correlation between leaf growth variables
510 suggests intrinsic and early controls of leaf size in *Arabidopsis thaliana*. Plant Cell Environ.
511 Blackwell Science Ltd; 2005;28:1355–66.

512 42. Chenu K, Franck N, Dauzat J, Barczi J, Rey H, Lecoeur J. Integrated responses of rosette
513 organogenesis, morphogenesis and architecture to reduced incident light in Arabidopsis thaliana
514 results in higher efficiency of light interception. Funct. Plant Biol. 2005;32:1123–34.

515 43. Christophe A, Letort V, Hummel I, Cournède P, de Reffye P, Lecœur J. A model-based analysis
516 of the dynamics of carbon balance at the whole-plant level in *Arabidopsis thaliana*. Funct. Plant
517 Biol. 2008;35:1147–62.

518 44. Lièvre M, Wuyts N, Cookson SJ, Bresson J, Dapp M, Vasseur F, et al. Phenotyping the
519 kinematics of leaf development in flowering plants: recommendations and pitfalls. Wiley
520 Interdiscip. Rev. Dev. Biol. 2013;

521 45. Blonder B, Vasseur F, Violle C, Shipley B, Enquist BJ, Vile D. Testing models for the leaf
522 economics spectrum with leaf and whole-plant traits in *Arabidopsis thaliana*. AoB Plants.
523 aobpla.oxfordjournals.org; 2015;7:lv049.

Table

Table 1. Equations of the models used in the study.

Eq.	Variable predicted	Units	Model	Parameters fitted on
1	Rosette dry mass	mg	$M = Rosette\ area + rosette\ perimeter$	population
2	Rosette dry mass	mg	$M = f(t) = \frac{A}{1 + e^{\frac{t_i - t}{B}}}$	individual
3	Growth rate	mg d^{-1}	$GR = f'(t) = rM\left(1 - \frac{M}{A}\right)$ with $r = \frac{1}{B}$	individual
4	RGR	mg d^{-1} g^{-1}	$RGR = \frac{f'(t)}{f(t)} = r\left(1 - \frac{M}{A}\right)$	individual
5	Number of siliques		$S = sk1 + sk2 + sk3 + sk4 + sk5 + sk6 + sk7$	population

Figures legend

Figure 1. The hydroponic system for *A. thaliana*. (**a**) Rockwool cube used in combination with small pot and black circle lid. (**b**) Example of tray image at 25 day old (top panel), individual pot at 25 day old in an early-flowering accession (bottom-left panel), and harvested adult plant with rosette and inflorescence (bottom-right panel).

Figure 2. The RAPA facility. (**a**) Entrance view of the growth chamber with a zoom on the camera installed between light tubes (top-left panel). On the right is the setup to water the plants and take manual tray picture. (**b**) Segmentation of tray image with the macro. (**c**) Examples of individual rosette images saved with the ImageJ macro, from which rosette area, perimeter and shape descriptors were extracted.

Figure 3. Cross-validation of the predictive model of rosette dry mass from rosette image. Rosettes harvested at 16 days after vernalization. (**a**) First cross-validation performed with replicate 3 used as training dataset and replicate 4 as testing dataset. (**b**) Second cross-validation performed with replicate 4 used as training dataset and replicate 3 as testing dataset. r^2 measured from Pearson's coefficient of correlation.

Figure 4. Logistic growth fitting and subsequent estimation of absolute growth rate and RGR. Top panels: example of growth fitting on a single individual (from left to right: logistic fit of rosette dry mass M, absolute growth rate GR as the derivative of the logistic function, and RGR as the GR/M at every time points). Bottom panels: Representation of the logistic fits of M, GR and RGR across the 451 accessions, colored by growth duration (days after vernalization).

Figure 5. Estimation of the number of siliques from inflorescence images. (**a**) From left to right: raw inflorescence image, segmented image and skeletonized image processed with the macro. (**b**) Cross-validation performed on a subset of the data with Pearson's

coefficient of correlation (r^2) between estimated and measured number of siliques on 352 test plants.

Figure 6. Variation in eight traits measured across the population of 451 accessions. In each barplot, accessions (n = 2) are ranked by increasing growth duration. *M*: final rosette dry mass, GR: absolute growth rate, RGR: relative growth rate, Root alloc.: ratio of final root dry mass over total plant dry mass (including root, rosette and inflorescence). Repro. Alloc: ratio of inflorescence dry mass over total plant dry mass. Error bars represent standard errors.

Figure 7. Reproducibility of rosette dry mass and silique number estimation. Measured across 18 contrasted accessions following the same protocol. (**a**) Pearson's coefficient of correlation (r^2) between rosette dry mass M estimated at the inflection point t_i in the first experiment and rosette dry mass M measured at t_i in the second experiment. (**b**) Pearson's coefficient of correlation (r^2) between the number of siliques estimated in the first experiment and the number of siliques measured in the second experiment.

Additional Files

Figure S1 ("FigS1_ex_TrayImage.JPG"). Example image of tray during vegetative growth.

Figure S2 ("FigS2_ex_InfloImage.JPG"). Example image of inflorescence at harvesting.

Figure S3 ("FigS3.pdf"). Representation of the 18 accessions chosen for testing reproducibility. Nine phenotypic groups represented by the purple circles (three groups of RGR and three groups of growth duration) were selected, each containing two accessions.

Table S1 ("TableS1_NutrientSolutionComposition.xlsx"). Chemical composition of the nutrient solution used in the hydroponic system.

574 **Table S2 ("TableS2_ModelParameters.xlsx").** Parameters of the models fitted,
575 corresponding to equations 1 (tab 1 in file), 2 to 4 (tab 2 in file), and 5 (tab 3 in file) of
576 Table 1.
577 **Table S3 ("TableS3_phenotypes.xlsx").** List of the 451 accessions with the eight
578 measured and estimated phenotypic traits presented in this study.
579 **Table S4 ("TableS4_Heritabilities.docx").** Broad-sense heritabilities (H^2) on eight
580 phenotypic traits with 95% Confidence Intervals (CI) estimated with *MCMCglmm*.
581 **Table S5 ("TableS5_RepeatedLines.xlsx").** List of the 18 accessions used for testing
582 reproducibility.
583 **Additional File 1.** Code of the ImageJ macro used for extracting individual rosette area,
584 perimeter and shape descriptors from tray images.
585 **Additional File 2.** Detailed R code used for modelling rosette biomass and logistic growth
586 fitting.
587 **Additional File 3.** Code of the ImageJ macro used for segmentation and skeletonization of
588 inflorescence images.
589 **Additional File 4.** Detailed R code used for modelling and cross-validation of the number
590 of siliques.

Figure 1.

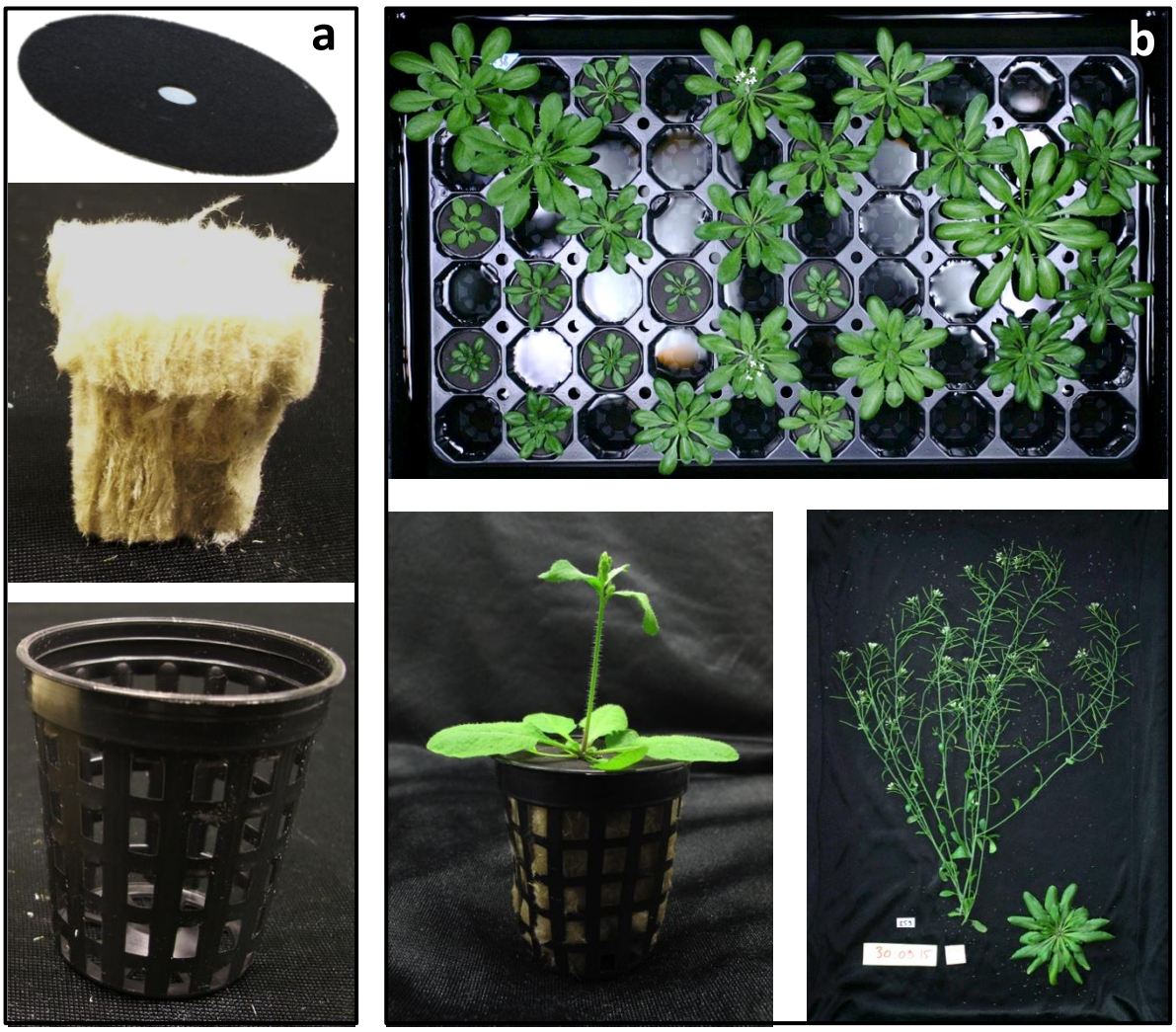


Figure 2.

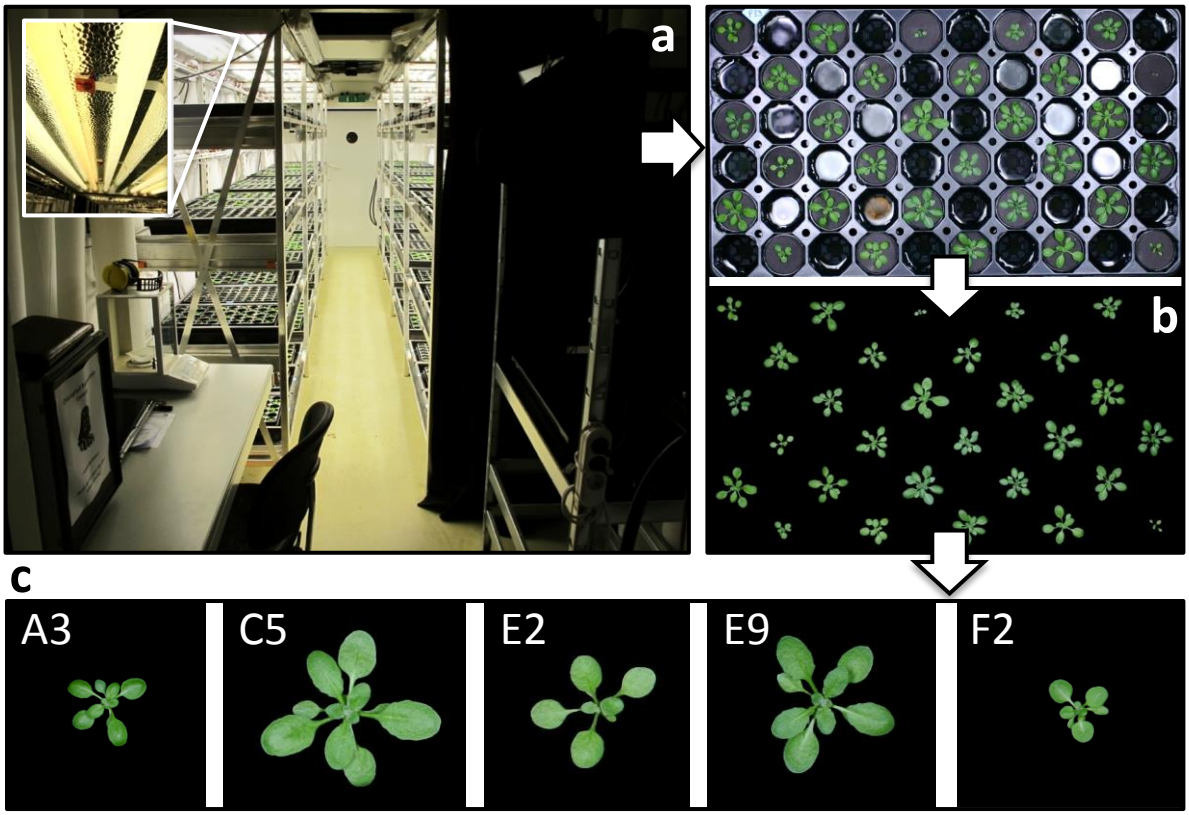


Figure 3.

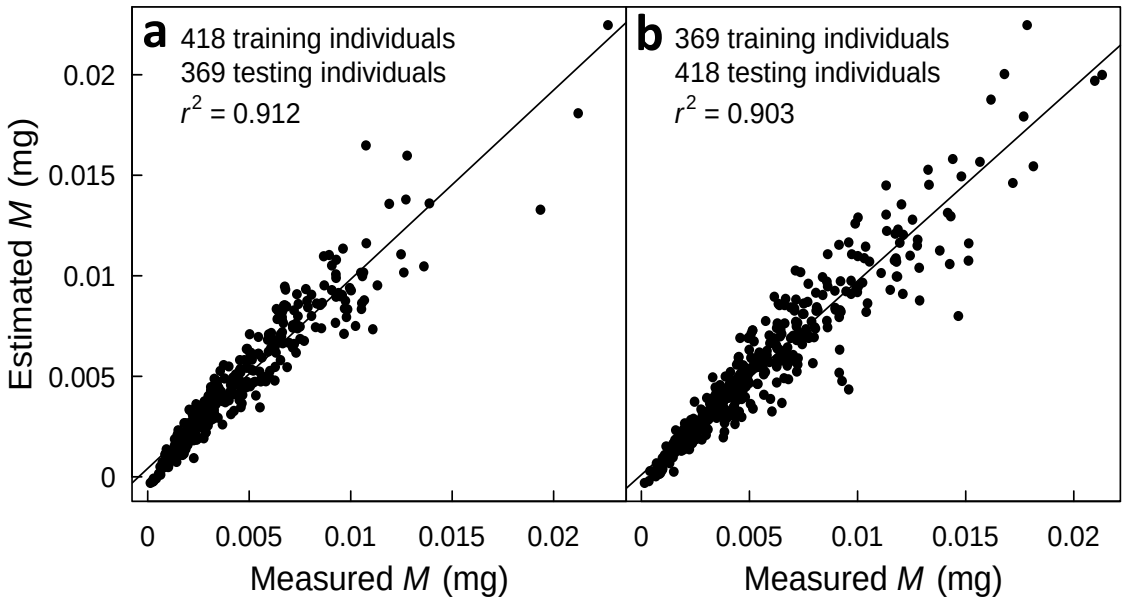

a
418 training individuals
369 testing individuals
r^2 = 0.912
b
369 training individuals
418 testing individuals
r^2 = 0.903
Estimated M (mg)
Measured M (mg)
Measured M (mg)
0
0.005
0.01
0.015
0.02

Figure 4.

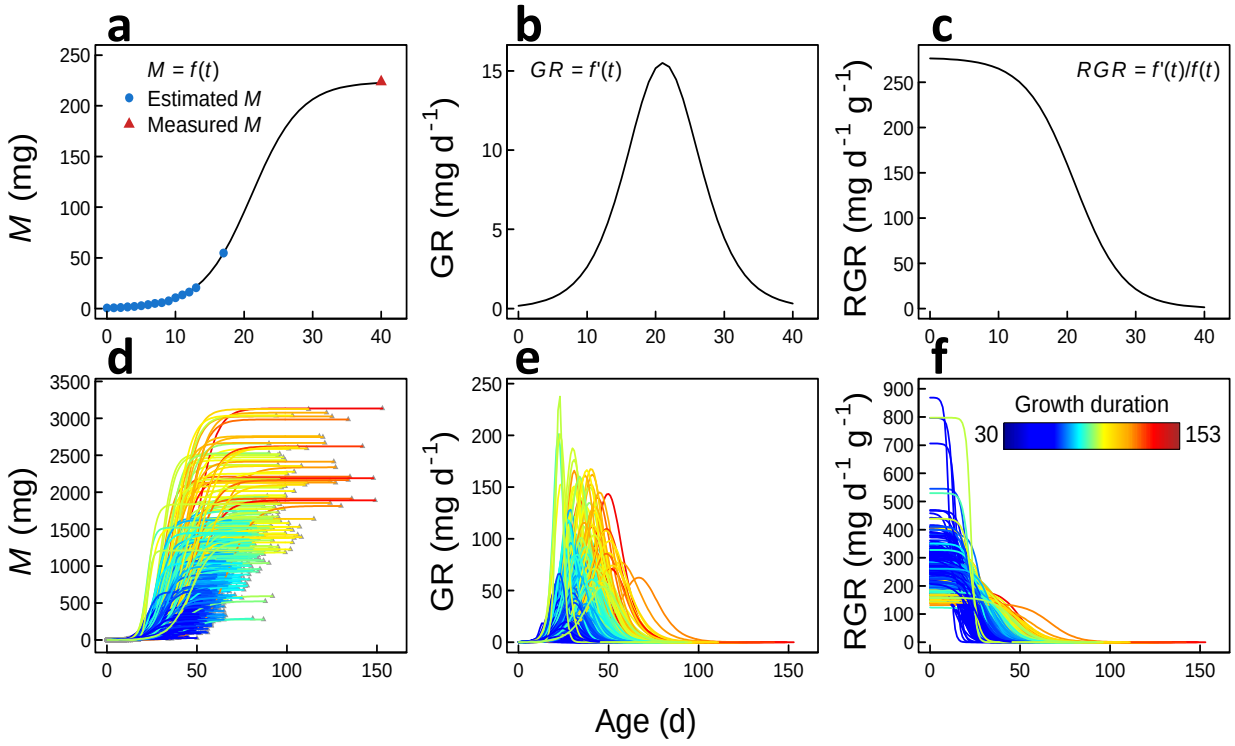

a
b
c
d
e
f
M = f(t)
Estimated M
Measured M
M (mg)
GR = f'(t)
GR (mg d-1)
RGR = f'(t)/f(t)
RGR (mg d-1 g-1)
Growth duration
30
153
Age (d)

Figure 5.

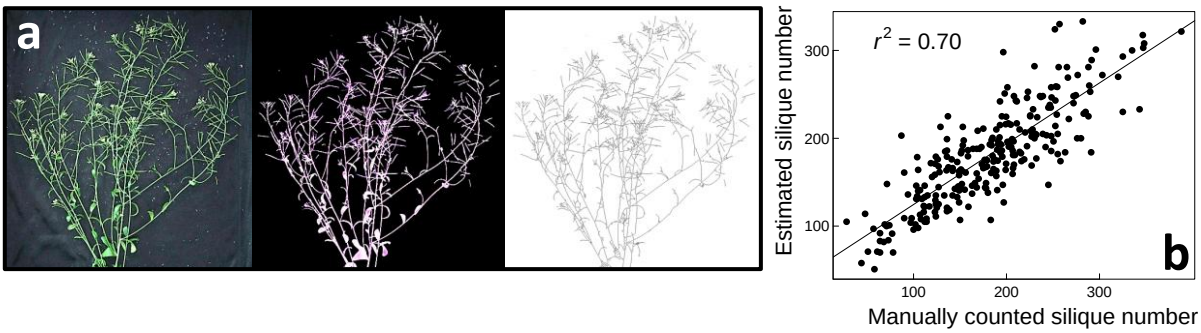

a
r^2 = 0.70
Estimated silique number
Manually counted silique number
100
200
300
b

Figure 6.

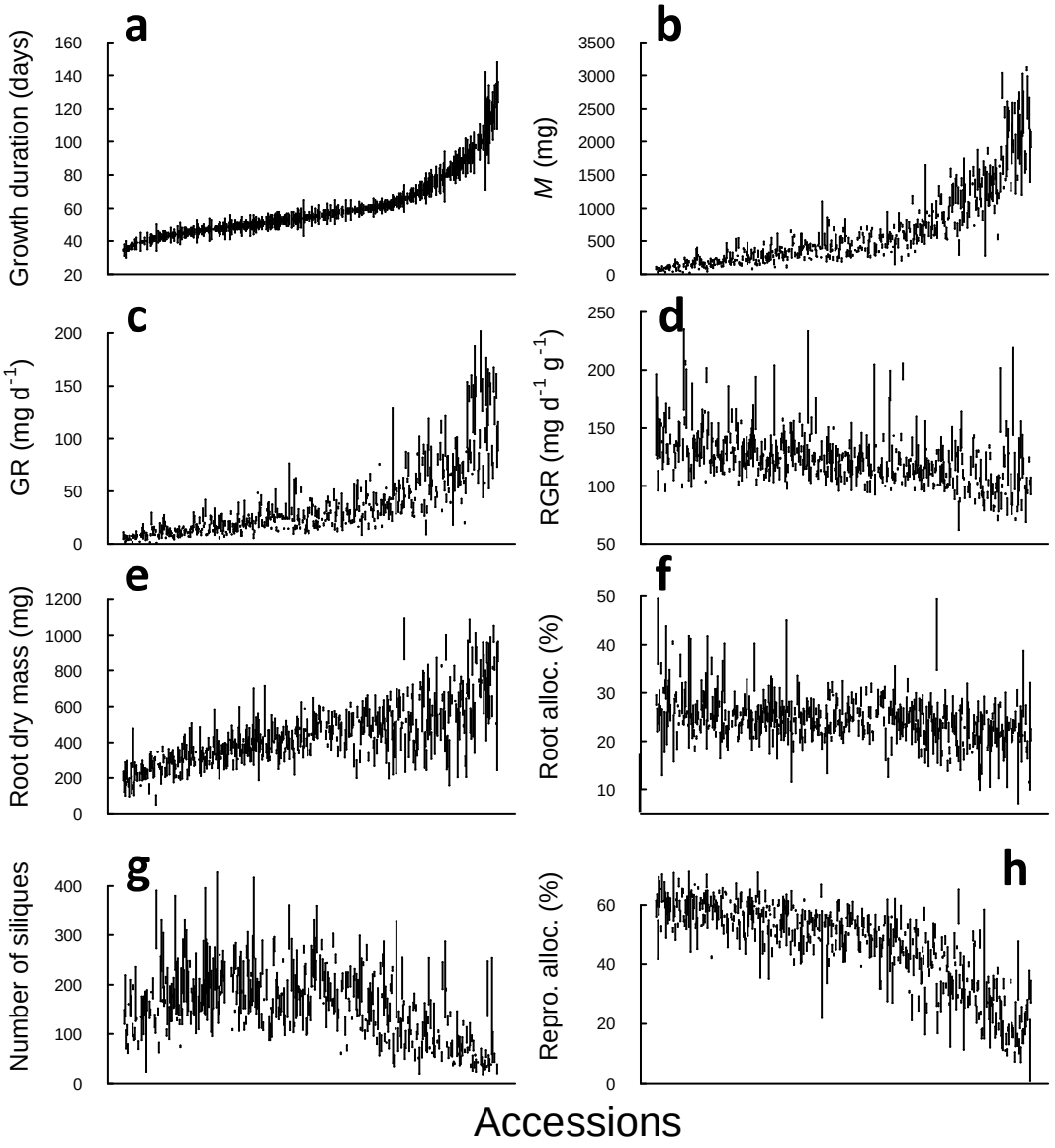
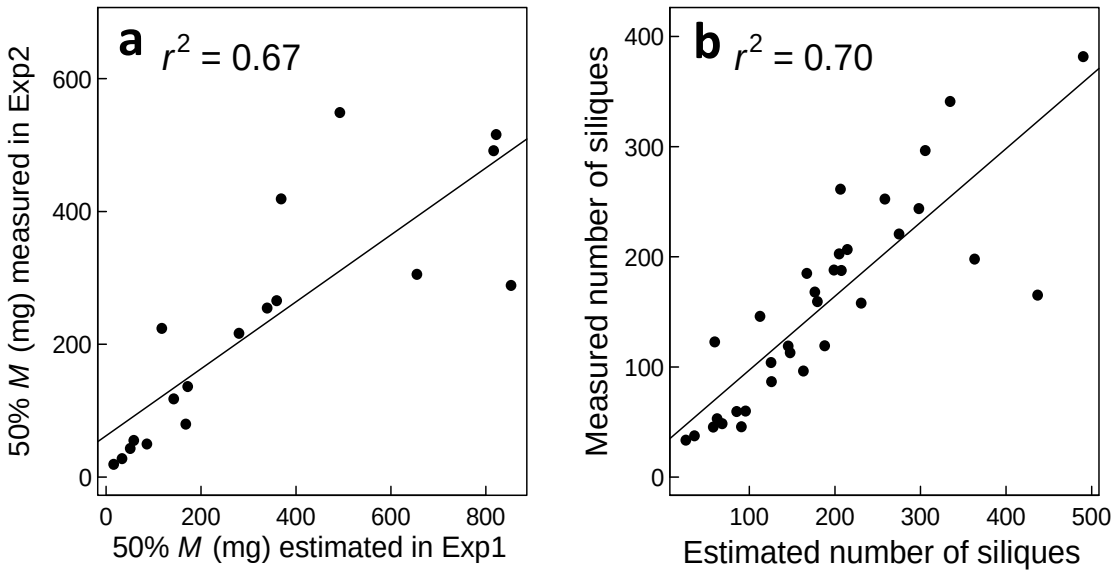

a
Growth duration (days)
160
140
120
100
80
60
40
20
b
M (mg)
3500
3000
2500
2000
1500
1000
500
0
c
GR (mg d-1)
200
150
100
50
0
d
RGR (mg d-1 g-1)
250
200
150
100
50
e
Root dry mass (mg)
1200
1000
800
600
400
200
0
f
Root alloc. (%)
50
40
30
20
10
g
Number of siliques
400
300
200
100
0
h
Repro. alloc. (%)
60
40
20
0
Accessions

Figure 7.


a
r^2 = 0.67
50% M (mg) measured in Exp2
50% M (mg) estimated in Exp1
0
200
400
600
800
b
r^2 = 0.70
Measured number of siliques
Estimated number of siliques
100
300
500