The complex genetic architecture of shoot growth natural variation in Arabidopsis thaliana

Elodie Marchadier, Mathieu Hanemian, Sebastien S. Tisne, Lien Bach, Christos Bazakos, Elodie Gilbault, Parham Haddadi Haddadi, Laetitia Virlouvet, Olivier Loudet

HAL Id: hal-02629200

https://hal.inrae.fr/hal-02629200

Submitted on 27 May 2020

The complex genetic architecture of shoot growth natural variation in *Arabidopsis thaliana*

Elodie Marchadier[¤a☯], **Mathieu Hanemian**[¤b☯], **Sébastien Tisné**[¤c☯], **Liên Bach,**
Christos Bazakos[¤d], **Elodie Gilbault, Parham Haddadi**[¤e], **Laetitia Virlouvet,**
Olivier Loudet *

Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay, 78000, Versailles,
France

☯ These authors contributed equally to this work.
¤a Current address: GQE Le Moulon, INRA, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-
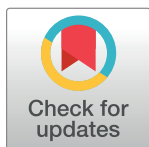Saclay, Gif-sur-Yvette, France
¤b Current address: Institute of Plant Science, Bern University, Bern, Switzerland
¤c Current address: CIRAD, UMR AGAP, Montpellier, France
¤d Current address: Institute of Plant Breeding and Genetic Resources ELGO-DEMETER, Thermi,
Thessaloniki, Greece / Department of Comparative Development and Genetics, Max Planck Institute for Plant
Breeding Research, Cologne, Germany
¤e Current address: Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada,
Saskatoon, SK, Canada
* Olivier.Loudet@inra.fr

Abstract

One of the main outcomes of quantitative genetics approaches to natural variation is to
reveal the genetic architecture underlying the phenotypic space. Complex genetic architec-
tures are described as including numerous loci (or alleles) with small-effect and/or low-fre-
quency in the populations, interactions with the genetic background, environment or age.
Linkage or association mapping strategies will be more or less sensitive to this complexity,
so that we still have an unclear picture of its extent. By combining high-throughput phenotyp-
ing under two environmental conditions with classical QTL mapping approaches in multiple
Arabidopsis thaliana segregating populations as well as advanced near isogenic lines con-
struction and survey, we have attempted to improve our understanding of quantitative phe-
notypic variation. Integrative traits such as those related to vegetative growth used in this
work (highlighting either cumulative growth, growth rate or morphology) all showed complex
and dynamic genetic architecture with respect to the segregating population and condition.
The more resolutive our mapping approach, the more complexity we uncover, with several
instances of QTLs visible in near isogenic lines but not detected with the initial QTL map-
ping, indicating that our phenotyping accuracy was less limiting than the mapping resolution
with respect to the underlying genetic architecture. In an ultimate approach to resolve this
complexity, we intensified our phenotyping effort to target specifically a 3Mb-region known
to segregate for a major quantitative trait gene, using a series of selected lines recombined
every 100kb. We discovered that at least 3 other independent QTLs had remained hidden in
this region, some with trait- or condition-specific effects, or opposite allelic effects. If we
were to extrapolate the figures obtained on this specific region in this particular cross to the

genome- and species-scale, we would predict hundreds of causative loci of detectable phenotypic effect controlling these growth-related phenotypes.

Author summary

The question of the complexity of the genetic variants underlying diversity in plant size and shape is central in evolutionary biology to better understand the impacts of selection and adaptation. In this work, we have combined the high resolution of a robotized platform designed to grow Arabidopsis plants under strictly-controlled conditions and the power of quantitative genetics approaches to map the individual genetic components (the 'QTLs') controlling diverse phenotypes, and hence reveal the so-called 'genetic architecture' of these traits. We show that the more we increase our resolution to map QTLs, the more complex of a genetic architecture we reveal. For instance, by focusing all of our mapping power on a small region representing 2.5% of the genome in an unprecedented phenotyping effort, we reveal that several independent QTLs had remained hidden in this region beyond a major-effect QTL that is always clearly visible. If this region is representative of the genome, this means that our current understanding misses potentially hundreds of variants finely controlling traits of evolutionary or agronomical interest.

Introduction

Fine-tuning plant growth throughout development and in response to environmental limitations is a decisive process to optimize fitness and population survival in the wild. As a sessile organism, plants have to cope with environmental fluctuations and evolved a wide range of responses. This is well illustrated by their great phenotypic plasticity and their ability to colonize very diverse habitats, through intraspecific genetic diversity as revealed in most pathways [1]. Aerial and below-ground growth represent a balance between resource investment in the structures and resource acquisition (respectively photosynthesis and water / nutrient uptake). Thus, growth is a highly complex trait controlled by many genes with constitutive or more specific roles depending on developmental stage, tissue, timing, environment [2–7]. In this context, plant growth can be considered as a model complex trait to increase our knowledge in the genetics of evolution, as well as to improve plant performance.

Forward mutant analysis plays a central role in plant biology to blindly identify gene functions associated with a phenotype [8], but sometimes remains limiting to reveal genes with modest phenotypic effect, or when addressing genes from redundant families. With regard to growth and stress tolerance, these limitations are likely to be relevant given the multigenic nature of growth phenotypes, the low mean effect at each locus and/or epistatic interaction they involve [9, 10]. Thus, the use of naturally-occurring variation through quantitative genetic approaches designed to map quantitative trait loci (QTLs) is interesting notably to complement the search for alleles selected during evolution which may not be brought out with classical loss-of-function approaches. Linkage mapping and genome-wide association lead to the identification of large amount of alleles involved in intraspecific phenotype variation from different plant species [1, 11].

With the drop of sequencing and genotyping costs, phenotyping clearly is the limiting factor for quantitative genetics approaches [12]. However, the complexity of the genetic architecture of a given trait, which depends on the contribution and the number of loci controlling a

trait and their interactions with the genetic background and the environment, has direct consequences on how much phenotyping remains limiting. Highly heritable traits with a limited number of contributing loci (in a given segregating material, or at the species scale) are more likely to be well understood than more complex traits. For instance, a large part of the phenotypic variation for flowering time in *Arabidopsis thaliana* maps to a limited number of loci [13–16], including *FRIGIDA* and *FLC* genes [17–19] and thus has a relatively simple genetic architecture, although many more loci make smaller contributions -at least in some environments- and allelic heterogeneity also interferes [15, 16, 20, 21]. By contrast, traits like fitness or growth can be expected to have a more complex basis as they integrate many upstream traits, and consequently many genes, each prone to residual variation and heterogeneity. Smaller contributions from individual loci means that, although one can still estimate total heritabilities, the accuracy and throughput of phenotyping will be limiting to confirm individual QTLs' contributions. Heritabilities for flowering time-related traits will often be above 80%, while biomass accumulation or fitness' heritabilities are essentially found in the range 20–60% [22–27].

Another factor that will influence the genetic complexity of a trait is its response to the environment through phenotypic plasticity [28]. Part of the environmental fluctuations may be controlled in an experimental design, while another part may contribute to the residuals. Whether the sensitivity of a pathway or trait to the environment depends on the number and architecture of the contributing loci remains an open question, however the relationships between higher plasticity and lower heritability are described [26]. Water availability is an environmental factor that varies through space and time and shows great heterogeneity which certainly constrains plant growth and shapes plant distribution in nature and in agricultural systems. Prevalence of drought episode is expected to increase with global climate change making the understanding of plant response to drought one of the major challenge of the next decades [29, 30]; this includes deciphering the genetic basis for variation in mechanisms such as drought escape, avoidance and tolerance [31]. Hence, this environmental parameter is definitely a good candidate to understand the genetic architecture of GxE. However, drought is both difficult to control and hard to predict, because of interactions with almost all other factors in the environment (temperature, air flow, light) and interplay with other constraints (especially nutrient-related or osmotic). The development of robotic phenotyping tools throughout the community makes it now feasible to acquire traits on hundreds or thousands of plants in precisely controlled and reproducible conditions [32–35], pushing a bit further one of the main limitation for a better decomposition of the genetic architecture of these complex traits.

Still, regarding plant growth variation in nature, mainly genomic regions with relatively large effect were identified in Arabidopsis and were often related to development, immunity or major hormones (for instance [36–43]). A limited number of non-theoretical studies seem to confirm that many genes with smaller effect–potentially involving epistasis and linked loci–would be responsible for part of the phenotypic variation of such complex traits (for instance [44–46]).

Here, we undertook a precise analysis of plant growth genetic architecture under both optimal watering condition and mild drought stress (as revealed under artificial conditions of reduced and constant soil water content), using a classical linkage mapping approach on 4 biparental segregating populations [47]. As a dynamic trait, we chose to follow growth during the vegetative phase using a high-throughput phenotyping robot (the *Phenoscope* [34]; https://phenoscope.versailles.inra.fr/) to map major- to small- effect QTLs as well as their interaction with drought stress. Zooming in on the loci, we use near-isogenic lines to validate these QTLs and reveal in more detail the genetics behind a single QTL peak. We then focus more precisely

on a region where a major Quantitative Trait Gene (QTG) is segregating (= *CRY2*, a known polymorphic actor with major pleiotropic phenotypic consequences), and show that other loci with additive or opposite effect are also present in its vicinity, illustrating the complexity of growth genetic architecture.

Results

Characterization of shoot growth-related phenotypes

The four RIL sets used in this study (BurxCol, CvixCol, BlaxCol, YoxCol) were conducted under well-watered (WW) and moderate water deficit (WD) conditions on our high-through-put phenotyping platform. Ensuring that growth occurs in a highly controlled and homogeneous environment, the Phenoscope records a number of image-based quantitative traits describing shoot development (Fig 1). Taking daily pictures gave access to cumulative (Projected Rosette Area, PRA) and dynamics (Relative Expansion Rate, RER) growth parameters for individual plants (Fig 1C & 1D) as well as other descriptive or derived traits (rosette
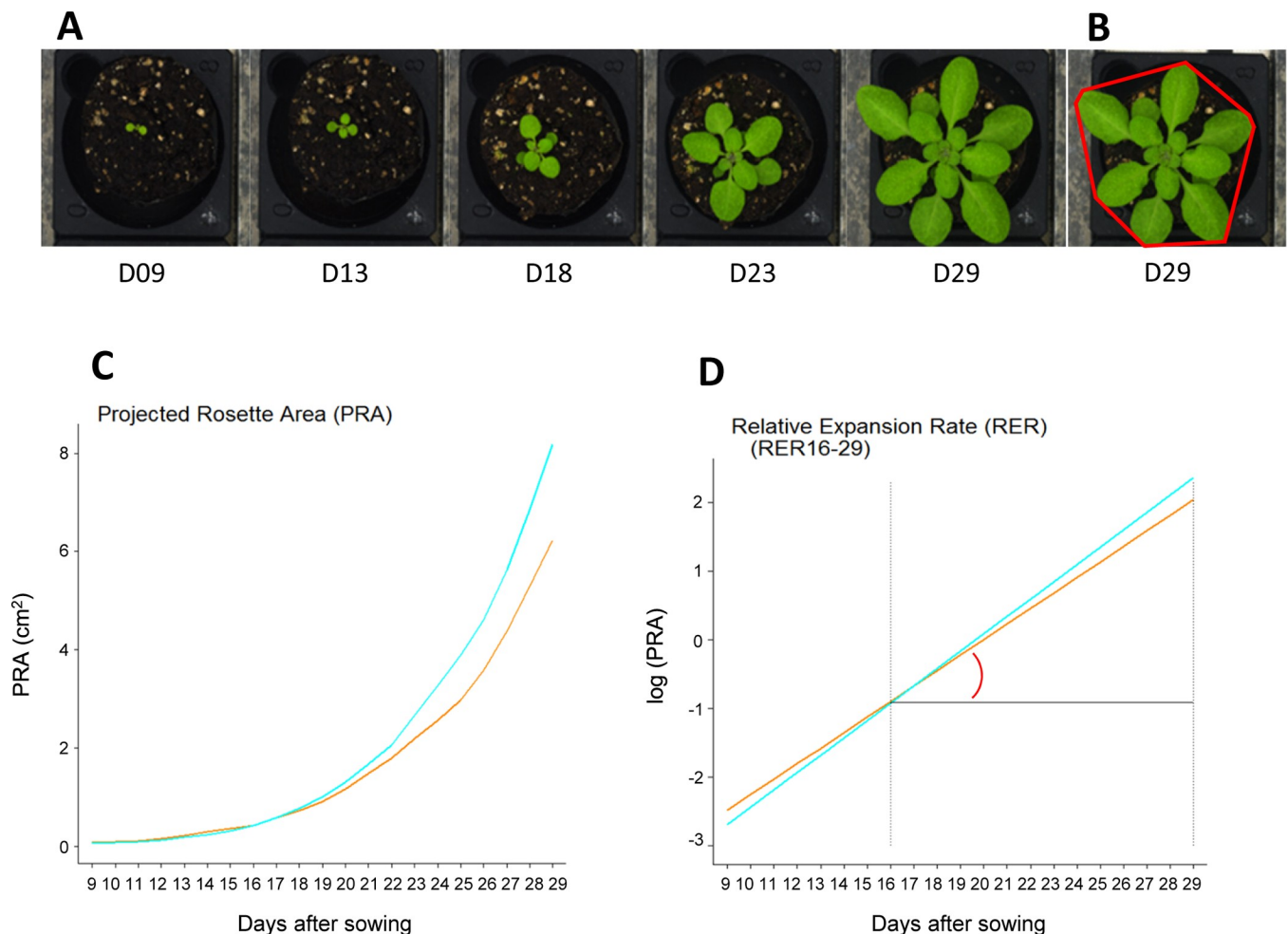
Fig 1. Extracting phenotypic information from zenithal images. A: Pictures of plants on the Phenoscope at different selected days after sowing. B: Representation of the rosette-encompassing convex hull (red shape) used to calculate the compactness (here illustrated at day 29). C: Growth kinetics obtained by extracting the projected rosette area from the daily images. D: Relative Expansion Rate calculated between two dates (for instance RER16-29 was integrated from day 16 to 29). C and D: Blue: 'WW' = Well Watered (soil water content at 60%); Orange: 'WD' = Water Deficit (soil water content at 30%).
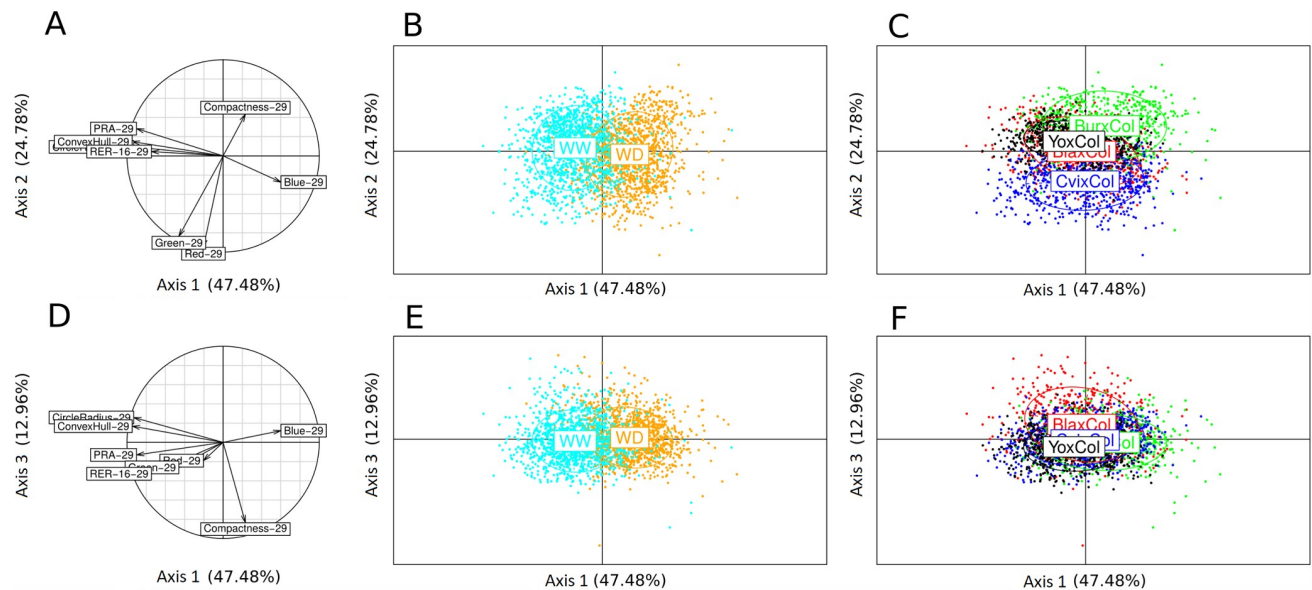
Fig 2. Principal component analysis of phenotypic traits. A to C show the results of the PCA based on the 2D plane corresponding to the first 2 axes. D to F show the results of the PCA based on the 2D plane corresponding to axes 1 and 3. A and D: circles of correlations with projected variables on the 2D plane. B and E: individual lines (RILs) projected on the 2D planes, clustered and colored according to the watering conditions as indicated. C and F: individual lines (RILs) projected on the 2D planes, clustered and colored according to the RIL set as indicated. B, C, E and F: Ellipses are centered on the gravity center of the individuals and encompass 67% of the individuals.

https://doi.org/10.1371/journal.pgen.1007954.g002

morphology and RGB colour components) [48]. A principal component analysis (PCA) was performed using all picture-based phenotypes at the final day of the experiment, 29 Days After Sowing (DAS; hence 'PRA29' etc) and relative expansion rate calculated between 16 and 29 DAS (RER16-29; Fig 2). The first axis explained a major part of the total variance, essentially through final rosette size (PRA29) and expansion rate (RER16-29). However, PRA29 and RER16-29 variables were not perfectly correlated, with genotypes exhibiting moderate PRA29 despite high RER16-29. The red (Red29) and green (Green29) components colour phenotypes mainly contributed to the second axis and were positively correlated, and both were negatively correlated with rosette compactness at 29DAS (Compactness29) which was the main trait contributing to the third axis. Individual projection showed that the first axis strongly structured the individuals according to the watering treatment (WW versus WD) while the axes 2 and 3 represented cross (RIL set) effects, differentiating CvixCol and BurxCol (axis 2) and BlaxCol (axis 3) from the other RIL sets. PRA29, RER16-29 and Compactness29 are complementary growth phenotypes that were investigated further in this study to quantify different aspects of shoot development variation: final projected rosette area is a cumulative proxy for biomass and photosynthetically-active surface, rosette compactness is an informative parameter describing the rosette morphology, and relative expansion rate highlights the dynamics of growth.

Phenotypic distribution among the RILs compared to their parents (S1 Fig) revealed extensive transgressive segregation for most of the traits and the crosses studied. As expected, mild drought stress (WD) condition impacted the distribution of the RILs for PRA29 and RER16-29 with generally reduced values for both traits. Interestingly, compactness distribution was much more robust to stress, which indicated that overall the morphology of the rosette is less

affected by mild drought. In order to estimate the part of phenotypic variation that is explained by genetic factors for each traits, heritabilities were calculated (S1 Table) and were essentially below 0.2 for RER (except in BurxCol where they were higher), essentially around 0.5 for cumulative PRA and generally above 0.5 for Compactness (this trait is certainly less sensitive to shifts in developmental stage that could be induced, for instance, by small changes in germination time). Overall, heritabilities were also lower in stress conditions than in control, as if the stress was inducing noisier phenotypic variations.

According to ANOVA analyses (S2 Table), all traits and RIL sets showed significant variation according to both the watering condition (WD versus WW) and the genotype (within each RIL set), except for Compactness29 response to stress in BlaxCol. There were also weaker (compared to the genotype and watering condition effects) but still significant biological replicates' effects (i.e. independent Phenoscope experiments), but less genotype x experiment interactions (with the exception of PRA and RER traits in BlaxCol for instance). PRA is more prone to genotype x experiment interactions than other traits, especially in CvixCol and YoxCol. Genotype x condition interactions are often milder than genotype or condition effects, and overall compactness–or YoxCol–show much less genotype x condition interactions than other traits/sets.

The phenotypic values of each RIL were then corrected for inter-experiment differences (indicated by the significant biological replicates' effect).

QTL mapping-based shoot growth genetic architecture

Our experimental design allowed the identification of many QTLs for all combinations of traits, conditions and RIL sets, and also for the GxE interaction term using genotype x condition effects from the ANOVA model for each trait (Fig 3 and S3 Table). Globally 112 QTLs were identified all along the genome when conditions are studied independently (62 under WW + 50 under WD) likely corresponding to at least 18 independent loci (genetically distant-enough from each other to be most likely considered as independent), yielding a median of 4 QTLs per modality of cross x trait x condition (ranging from 1 to 8 QTLs). QTL hotspots across RIL sets and traits were identified for instance at the beginning of chromosome 1, at the bottom of chromosome 2 and 5. These hotspots include very highly significant QTLs with LOD scores above 10, and up to 32. Chromosome 3 appeared to show less significant QTLs in all crosses, especially for PRA29 and RER16-29. Individual QTL contributions to phenotypic variance (R^2) ranged from 1 up to 30%, and showed a L-shaped distribution of effect (S2 Fig). Using empirical significance boundaries according to the observed distribution of QTLs effects, ~10% of the QTLs could be considered as showing major effects and significance ($R2>10\%$ and/or $LOD>15$); ~25% of the QTLs could be considered as showing intermediate effects and significance ($5\%<R2<10\%$ and/or $7<LOD<15$); the remaining 2/3rd of the QTLs could be considered as showing minor effects and significance $R2<5\%$ and/or $LOD<7$). Many more potential QTLs not listed here were only suggestive with LOD score peaks just below our permutation-based threshold (<2.4 LOD).

Most of the QTL profiles are stable across the 2 watering conditions, especially for the major-effect loci. However, QTLs specific for one condition were detected, e.g for RER16-29 under WD in YoxCol on chromosome 4, and at the top of chromosome 1 under WW. We also mapped QTLs for the interaction term with the drought treatment, yielding 19 QTLs (Fig 3 and S3 Table). These essentially emphasize large effect QTLs showing a modulation of their effect in response to stress (especially for PRA, see chromosome 1 and 5), with no new loci revealed. For RER16-29 in YoxCol, this would confirm the interaction of the above-mentioned locus on chromosome 4 (although its exact position is questionable), but not for the one at the
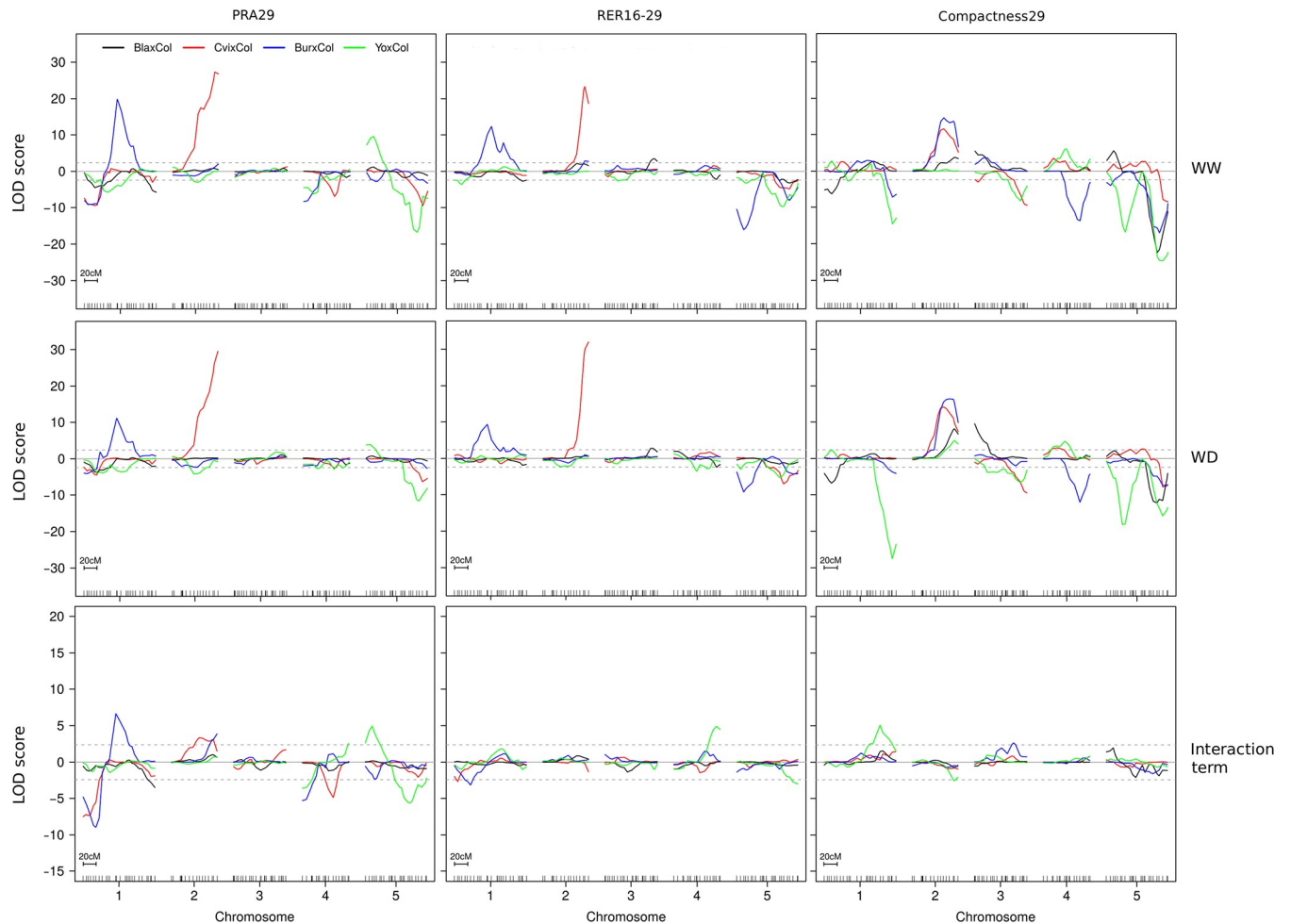
Fig 3. QTL mapping for 4 RIL sets, 3 traits and 2 conditions. QTL maps obtained from four RIL sets (BlaxCol in black; CvixCol in red; BurxCol in blue; YoxCol in green) using three phenotypes (PRA29, RER16-29 and Compactness29), generated independently in two conditions (top panels: WW; middle panels: WD) and from the genotype x environment interaction term in the model (bottom panels: Interaction term). LOD score values above (below) zero indicate that the Col allele increases (decreases) trait value with respect to the other parental allele. Significance threshold (LOD = 2.4) is shown as dotted lines. The x-axis indicates consensus markers along the genetic map (scale is shown for 20cM). The parameters of the significant QTLs detected are listed in S3 Table.

https://doi.org/10.1371/journal.pgen.1007954.g003

top of chromosome 1. There may be some power issues when comparing across conditions due to lower heritabilities of traits under WD.

Although derived from PRA29, Compactness29 showed an independent genetic basis, as exemplified by the major peaks on the bottom of chromosome 4 in BurxCol or 5 in BlaxCol. Although contributing to RER16-29, PRA29 does not always share the same contributing loci, for instance on the top of chromosome 1 in CvixCol and BurxCol. Other more complex cross x trait patterns are apparent, like at the bottom of chromosome 2 where a major QTL for Compactness29 in three of the four RIL sets seems to colocalize with a significant PRA29 and RER16-29 QTL (in the same direction), but only in one cross. It may be that these Compactness29 QTLs are actually independent in each RIL set.

A two-dimensions search for epistatic interactions was performed across all traits, conditions and RIL sets (S3 Fig). Overall, the BurxCol and CvixCol RIL sets showed more significant epistasis compared to the two other sets. Interestingly, pairwise interactions controlling growth phenotypes are overall quite different depending on RIL set and growth phenotypes. Shared

Fig 4. A significant epistatic interaction for PRA29 in BlaxCol. A significant epistatic effect on PRA29 between the bottom of chromosomes 4 and 5 segregating in the BlaxCol RIL set. Results shown in WW condition. Left panel: heatmaps representin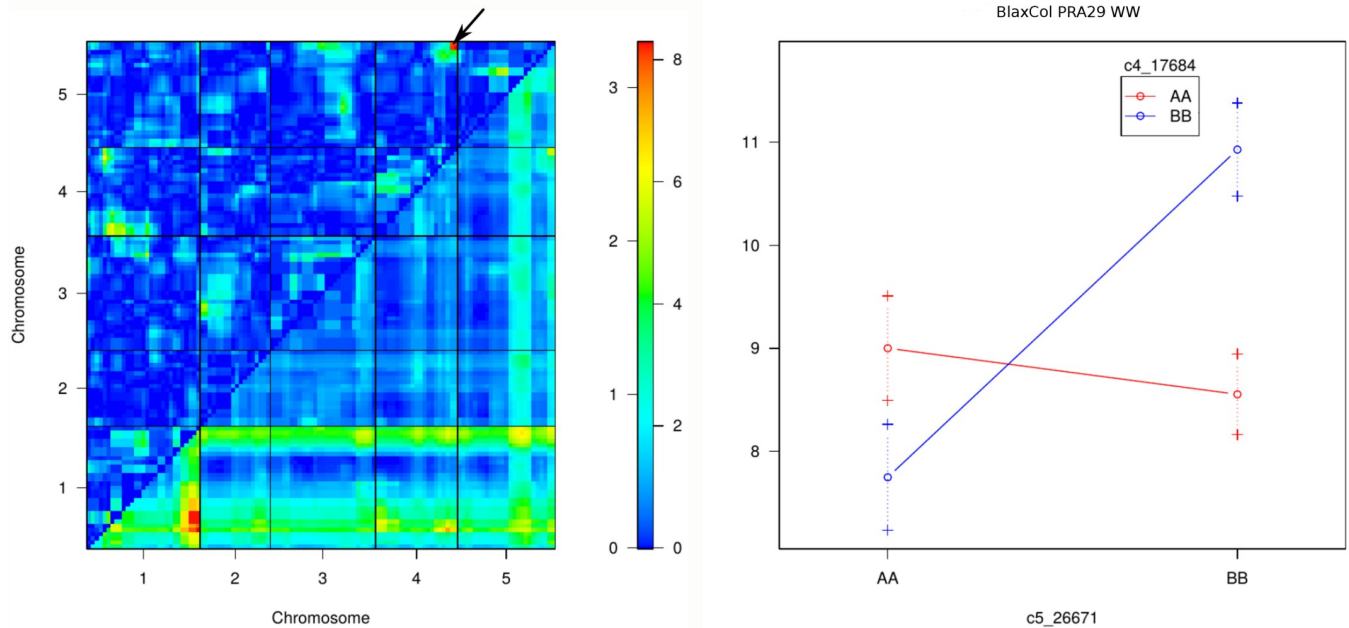g the LOD score of the full model with additive and interaction effects (triangle below diagonal; 0–8 LOD scale) and that of the interaction effects only (triangle above diagonal; 0–3 LOD scale) for PRA29 in BlaxCol and WW condition. The LOD scales for significance are indicated on the right of the heatmaps. An arrow indicates the most significant interaction, which phenotypic consequences are detailed on the right panel: the interaction plot represents the average RIL phenotype (PRA29 in cm2; +/-1 SE) for each allelic combination ('AA' = homozygous Col; 'BB' = homozygous Bla) at the 2 closest markers.

https://doi.org/10.1371/journal.pgen.1007954.g004

epistasis effects are potentially detected in both watering conditions: they appear as symmetrical across the diagonal on S3 Fig. One of the most significant epistatic interaction was observed between 2 loci on the top and bottom of chromosome 4 in the BurxCol cross (S3 Fig; interaction component highly significant for RER16-29 in WW: 5.1 LOD). Positions and directions of effect match perfectly with the previously published SG3 x SG3i interaction known to segregate in this cross [49]. Based on its significance, another relevant interaction (3.34 LOD in WW and 3.12 LOD in WD) was observed for PRA29 in the BlaxCol cross between the bottom of chromosomes 4 and 5 (Fig 4). The effect of the bottom of chromosome 5 QTL on PRA29 is observed only when RILs carry Bla alleles at the bottom of chromosome 4. As a consequence of this epistasis, these QTLs appear barely significant in single QTL scans (Fig 3).

We also performed dynamic QTL detection on daily-recorded traits (PRA and Compactness) to reveal the evolution of QTL effect throughout the experiment on the Phenoscope: these interactive QTL profiles can be accessed at http://www.inra.fr/vast/PhenoDynamic.htm

Most of the PRA QTLs observed after 29 days of growth correspond to locus that become gradually significant across the experiment and are essentially not time-specific. These are most likely contrasted allelic effect on growth that cumulate their effect over time. There are only a few exceptions for PRA, like the bottom of chromosome 5 locus segregating in BlaxCol which remains significant only until 13DAS and thereafter is canceled out. For Compactness, the picture is rather different with numerous examples of QTLs that are essentially significant around specific time-points, even sometimes in successive waves of significance (BlaxCol, bottom of chromosome 5: QTL peaking at Days 11, 17 and 27 –providing that this is a unique locus).
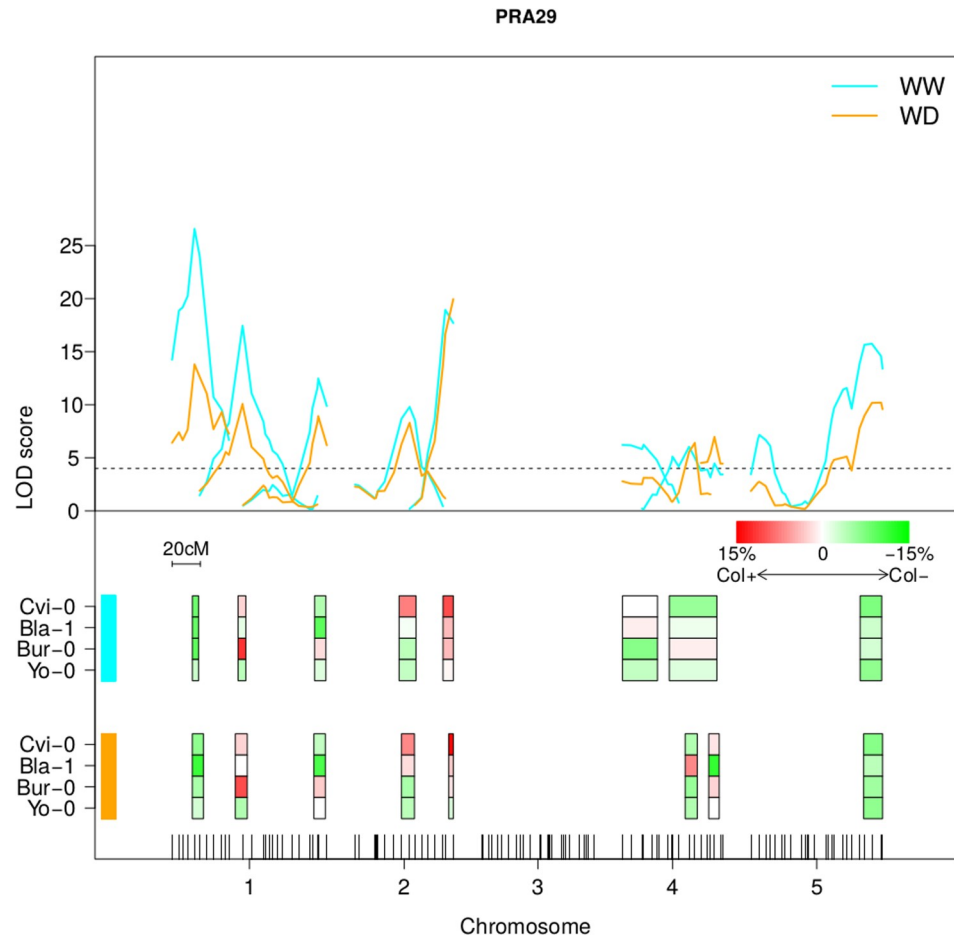
Fig 5. Multi-cross QTL analysis for PRA29. The upper panel of the figure represents the LOD scores calculated by the MCQTL method along the genetic map in WW (blue) and WD (orange) conditions. Each QTL is depicted by a LOD score profile which peak is represented by colored boxes on the lower panel of the figure (WW conditions above and WD conditions below). The color of the box indicates the deviation of the trait value for the Cvi, Bla, Bur or Yo alleles relative to the Col allele as indicated. Chromosome 3 shows no significant combined QTLs for PRA29. This figure shows results obtained for trait PRA29 only; other traits are in S4 and S5 Figs.

https://doi.org/10.1371/journal.pgen.1007954.g005

In order to take advantage of the different crosses to Col-0 in our experimental setup, QTL mapping was performed on the whole dataset using MCQTL (Multi-Cross QTL) tool to compare allelic effects in a multicross design and potentially reveal shared QTLs. The combined QTL maps obtained (Fig 5 for PRA29, S4 Fig for RER16-29, S5 Fig for Compactness29) highlight 10 independent loci, including at least 4 regions with contrasting allelic effect on PRA29: for instance, the middle of chromosome 1 region shows contrasted phenotypic consequences in different crosses, particularly when comparing Bur and Yo alleles (with respect to Col). At this scale, it can be difficult to distinguish between different alleles at the same QTL and different QTLs. Conversely, combining the information of multiple crosses sometimes allows to predict narrower QTL intervals than with the initial QTL mapping, enabling the detection of distinct linked loci, for instance for PRA on the bottom half of chromosome 2; a location where the dynamic QTL analysis for PRA in CvixCol was already showing signs of 2 different segregating loci with slightly distinct dynamics over time. Another striking example is for Compactness on chromosome 5, with neighbouring QTLs predicted to show opposite allelic effects (e.g. BlaxCol).

Near isogenic lines-based shoot growth genetic architecture

To confirm and investigate further the complexity of the genetic architecture of these traits, we used near-isogenic lines to mendelize QTLs and assess in more details the role of smaller chromosomal regions. Using 81 independent Heterogeneous Inbred Families (HIFs) scattered across the genome or chosen to decompose candidate regions in specific crosses [50], we tested a total of 79 QTL effects from the 24 modalities of RIL set x trait x condition (Fig 6 for PRA29; S6 Fig for RER16-29 and Compactness29). Globally, 60% of the HIFs with a segregating region covering a candidate region previously identified showed significant effect with consistent direction, thus validating the QTL. Larger effect QTL were more often validated in HIF, with ~75% of the tested major or intermediate effect QTLs confirmed, compared to ~50% of the minor effect QTL. Specifically, for PRA29 trait in the two conditions studied (Fig 6), 28 QTLs were assessed with at least one HIF, among which 17 (60%) were confirmed. Some (minor-effect) QTL x condition interactions detected in the RIL set were also significantly confirmed using HIF, such as the top of chromosome 5 locus in BurxCol which has significance (PRA29) only under WW.

We are also interested in positive HIF results that do not match with the results of the initial QTL mapping. This was particularly possible in regions with high HIF coverage and shows that several LOD score peaks were explained by more than one underlying QTL, as also suggested by the MCQTL analysis. A good example lies in the CvixCol cross, where the bottom
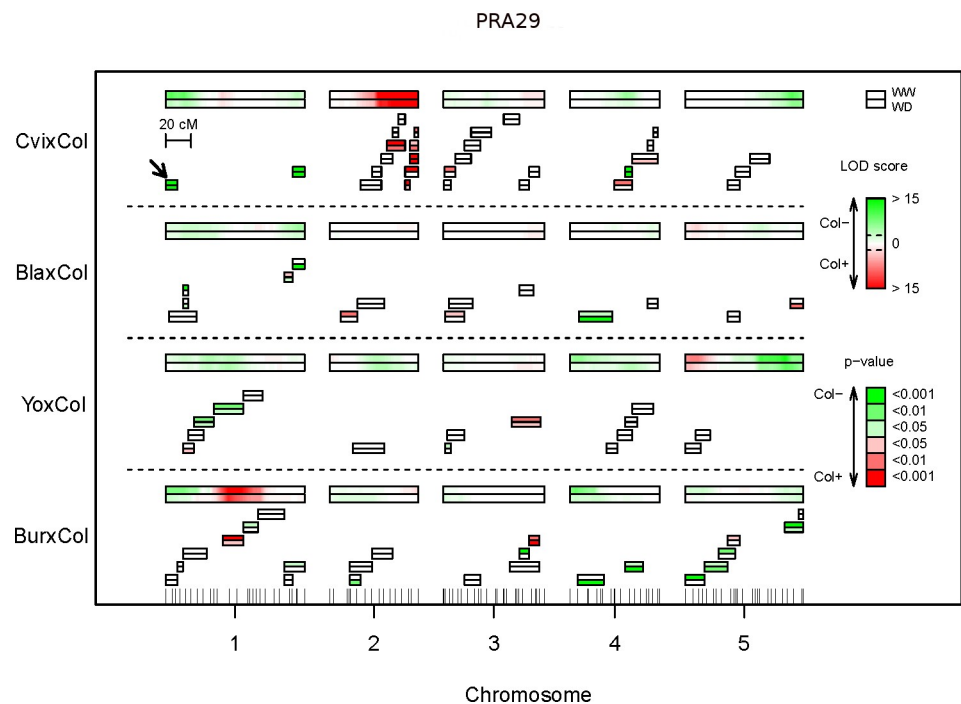
Fig 6. Near isogenic lines-based validation of QTLs for PRA29. QTL mapping results (same LOD score profiles as Fig 3) are shown as LOD score heatmaps (see color scale on the right) for each RIL set, across long twin boxes representing the 5 chromosomes (the upper frame is for WW, the lower frame is for WD conditions). Below each chromosome are series of short twin boxes representing individual HIFs segregating region, used to test phenotypes associated with specific regions (for each twin box the upper frame is for WW, the lower frame is for WD). The significance and the direction of the difference between Col-fixed HIF lines and alternate parental allele-fixed HIF lines are indicated by the p-value color scale as indicated. Details of each HIF used here are gathered in a deposited dataset [50]. An arrow indicates the HIF exploited further in the 'microStairs' approach. This figure shows results obtained for trait PRA29 only; other traits are in S6 Fig.

https://doi.org/10.1371/journal.pgen.1007954.g006

half of chromosome 4 seemed to control PRA29 as a single locus in the initial QTL analysis (Fig 3) and is now subdivided in at least 2 independent loci with opposite allelic effects after the HIF analysis (Fig 6). The 2 adjacent HIFs do not only show opposite allelic effect, but also different allelic effect amplitude on PRA29 ($-1.4cm^2$ versus $+0.9cm^2$), potentially explaining why only a Col-negative allelic effect was detected in the initial QTL mapping in this region, since the effect in this direction is stronger. Another example segregates in YoxCol at the top of chromosome 3 where 2 QTLs with opposite-effect on Compactness29 under WW conditions seem to localize closely according to the HIFs (S6 Fig) but wasn't detected at all in the QTL mapping at 29DAS (Fig 3) and only remained significant at intermediate stages around 16DAS according to the dynamic analysis. Here, the 2 adjacent HIFs show very similar allelic effect amplitude on Compactness29 ($+0.0301$ versus -0.0295), likely explaining why no QTL was detected in this region for this trait. These could be examples of a lack of power of the RIL design to detect complex patterns due to the confusing effect of linked loci; alternatively, the confusing effect of epistasis could also interfere when comparing QTL mapping results from RILs and HIFs, because of the specific genetic background of each HIF.

Finally, we can occasionally exploit the localization of the segregating region(s) of the tested HIFs to narrow down the candidate QTL region, like for a PRA29 QTL in CvixCol at the bottom of chromosome 2 which is narrowed down to the extremity of the chromosome and shown to be distinct (= confirmed in non-overlapping HIFs) from another nearby QTL segregating in the same cross (with allelic effect in the same direction), as predicted by the multi-cross analysis above. Here again the dynamic analysis also helped to distinguish these loci based on their effect through time. Still, the 'precision' remains approximately at the Mb level at this stage.

To tackle further the question of the complexity of the genetic architecture at a higher resolution than with simple HIFs, we decided to systematically dissect a region of 3Mb at the very beginning of chromosome 1 in CvixCol: in this region all previous approaches have predicted a single QTL with intermediate to major effect significance on PRA29 (Fig 3), which was confirmed in an HIF (Fig 6). This HIF was used to zoom in on the region across 30 bins (stairs) of ~100kb, defined by successive recombination breakpoints (S4 Table) and phenotypically evaluated individually in the same conditions as above (an approach coined 'microStairs'). We selected this interval to test if a region harbouring a large-effect growth QTL may typically also include independent loci, maybe of smaller effect or hidden by the main QTL. A pairwise comparison of their growth phenotypes allows to test the impact of Cvi versus Col alleles over relatively short physical intervals of expected average size 100kb, with a maximum ~200kb, depending on recombination breakpoints location and marker interval. We either compared only the 2 successive recombinants ('stair by stair'), or we took advantage of the support from all pairwise comparisons ('staircase') to increase our power to detect QTL (Fig 7 for PRA29 and S7 Fig for RER16-29 and Compactness29).

The strongest phenotypic effect for PRA29 in WW and WD (but also–to a lesser extent–for RER16-29 in WD) was fine-mapped down to the stair between recombinants 111 and 112, covering potentially a physical interval (bin) extending to the maximum between 1.070 and 1.291 Mb (an interval including 54 genes). In the middle of this region lies the obvious candidate gene *CRYPTOCHROME 2* (AT1G04400; *CRY2*, a blue-light photoreceptor; S5 Table) that is known to harbour a functional variant in Cvi (a single amino-acid change) and impact plant photomorphogenic development especially in short days [51, 52]. This is a good test case of our approach, as it is very likely that *CRY2* is primarily responsible for the growth difference overall observed in this HIF.

Still, on either side of this locus, a few other QTLs also affect growth, underlying a much more complex genetic architecture than expected after the initial HIF results: in both watering
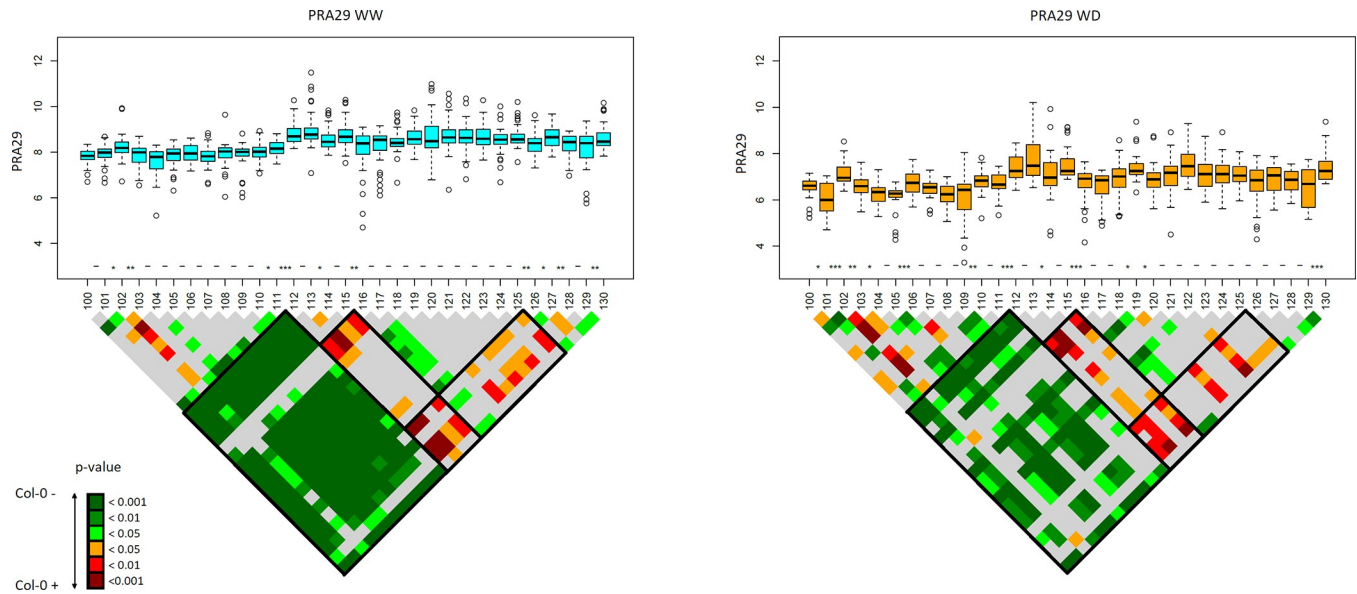
Fig 7. Dissection of a genomic region in CvixCol for PRA29 ('microStairs' approach). The upper panels' are quartile boxplots indicating phenotypic values observed for each of the 31 successive recombinant lines along the 3Mb region studied (details of each line used here are in S4 Table). Significant differences between successive lines (stairs) are indicated along the X axis with '*' (p-value < 0.05), '**' (p-value < 0.01), '***' (p-value < 0.001) or '-' for no significant difference. The lower panel shows a triangular matrix gathering the significance of all pairwise comparisons between lines (see the p-value color scale also indicating the direction of the allelic effect). Blocks of significance ('staircase') are interpreted and highlighted in black-framed boxes. This figure shows results obtained for trait PRA29 only (in 2 conditions), PRA is expressed in cm2; other traits are in S7 Fig.

https://doi.org/10.1371/journal.pgen.1007954.g007

conditions, a milder PRA29 QTL, with opposite allelic effect than *CRY2*, was detected in the stair defined by recombinants 115 and 116 (1.495 to 1.703 Mb positions). Another PRA29 QTL was predicted in the bin 125/126 (2.509 to 2.710 Mb). For Compactness29 specifically under WD treatment, *CRY2* did not seem to be causal and the causative locus for the observed HIF phenotype would most likely be in the bin 118/119 (1.801 to 2.027 Mb); this QTL was not clear from the initial QTL analysis for this trait on Day 29 (S6 Fig), but appeared significant earlier in the experiment (cf dynamic analyses for CvixCol around 14DAS and later). There are probably even more loci, especially at the very beginning of the region for PRA29 (stairs between recombinant 101 and 104 at least), but we reach the limits of our experimental design and phenotyping precision to be able to conclude accurately, with either too complex genetic architecture in this region or not enough recombinant lines to robustly support each intervals' effect.

We then looked for high impact polymorphisms (premature stop codon, frameshift or non-synonymous mutations) likely affecting gene function between Col-0 and Cvi-0 within the most promising bins to identify candidate genes. Because of the numerous non-synonymous changes between these accessions (S5 Table), we decided to arbitrarily filter the genes with the criteria of at least 3 non-synonymous mutations to increase our chance to detect genes with altered function or degenerated sequences after loss-of-function mutations. Here, we discuss some of those polymorphic genes where previous publications have indicated a putative function or effect that could relate to our phenotypes.

Within the bin 115/116 lie at least 2 interesting candidates:

URIDINE DIPHOSPHATE GLYCOSYLTRANSFERASE 74E2 (AT1G05680; *UGT74E2*) is an auxin glycosyltransferase whose overexpression was shown to modify plant morphology and the size of the petioles, to delay flowering, and to increase drought tolerance [53]. The

gene is also known for ample natural variation in expression, including potentially cis-acting variants [54]: http://www.bioinformatics.nl/AraQTL/multiplot/?query=AT1G05680.

GLUTAMATE RECEPTOR 3.4 (AT1G05200; *GLR3.4*) is a calcium-dependent abiotic stimuli-responsive gene [55] expressed throughout the plant and impacting at least lateral root initiation [56]. It harbours several non-synonymous variants in Cvi-0 (compared to Col-0), but also higher expression in Col-0 as shown by several local-eQTLs in different crosses (http://www.bioinformatics.nl/AraQTL/multiplot/?query=AT1G05200) confirmed to be cis-regulated through Allele-Specific Expression assay [57] (reported in S6 Table).

Within the bin 125/126 lie at least 2 interesting candidates:

AT1G08130 encodes *DNA LIGASE 1* (*LIG1*), a ligase involved in DNA repair, which mutation causes severe growth defects [58]. Its expression is also known to be controlled by a local-eQTL (most likely cis-acting) in LerxCvi:

http://www.bioinformatics.nl/AraQTL/multiplot/?query=AT1G08130

AT1G08410 is *DROUGHT INHIBITED GROWTH OF LATERAL ROOTS* 6 (*DIG6*), encoding a large 60S subunit nuclear export GTPase 1 that impacts several developmental processes regulated by auxin, including growth [59].

Further work is required to prove any link between the observed phenotypic variation and these candidate genes.

Discussion

Owing to its fine regulation throughout development and in interaction with the environment, plant growth represents a highly complex trait potentially controlled by numerous factors and interactions. Little is known on the actual genetic architecture of plant growth natural variation, with essentially a few genes of major effect being identified until now and only a few exceptions of more complex genetics revealed [44]. Association genetics hold great promise to dissect the underlying molecular bases of complex traits [12], however one can wonder if the genetic architecture of highly integrative traits like growth or fitness is amenable to genome-wide association studies (GWAS) at the species scale: GWAS especially lacks power to decompose traits controlled by many loci of small effects when the underlying alleles have low frequency in the mapping population. For instance when exploiting worldwide collections of accessions, very little (if any) significant associations for growth parameters were found [60], even when using very anti-conservative thresholds [2, 61] or using morphological parameters with higher heritabilities [62]. Even with more targeted growth traits like root cell length, Meijon *et al.* did not detect any significant signal above the threshold, although the first peak just below the threshold identified a causal gene [63]. Overall, it is argued that linkage and GWA studies are complementary in the loci that they are able to reveal, depending on the genetic architecture of the trait in the population considered (e.g. [21, 64]). Here, by studying four different crosses to the reference Col-0, we find many cross-specific loci, especially of mild effects, several of which might correspond to low frequency-alleles that would not likely be pictured in GWAS (not enough power due to low frequency x effect size).

Whether linkage or association mapping, these approaches are both similarly phenotyping-intensive and prone to interaction with uncontrolled environmental parameters (increasingly so with the scale of the experiments). Using our high-throughput phenotyping robots to grow individual plants under tightly controlled conditions, we intended to dissect the genetic basis of plant growth under optimal and limiting watering conditions, to a level of accuracy rarely reached so far. We focused on vegetative growth from days 8 to 29 (after sowing) and selected three non-fully-correlated variables allowing to characterize plant growth dynamically: rosette area (PRA), relative growth rate (RER) and compactness. PRA is a typical cumulative trait:

what is observed on day n is not independent of what has occurred from day 1 to n-1. Compactness, which basically represents a measure of PRA normalized by rosette width, is rather independent of cumulative phenomenons and hence shows much more age-specific QTLs; this is particularly striking when comparing the dynamic analyses for these traits. Finally, RER integrates growth rates over a specific period of time, and is independent of plant size (= relative). Because estimated during the exponential phase of growth, RER is very much similar to what is obtained by fitting exponential models to the PRA data and exploiting model parameters [2].

Hence, growth-related phenotypes, depending on how they are exploited, will present different genetic architectures throughout age, with individual loci making different contributions to cumulative or age-specific traits. It has previously been shown that heritabilities for growth-related traits change over time [6, 65]; one possible explanation is that QTLs are more or less likely to act at specific time points. For instance, studying growth dynamics in maize [66, 67] and root tip growth in Arabidopsis [68] allowed to identify marker-trait associations that would not be detectable by considering the cumulative trait only at a single (final) time point. At the other end of time-resolution for growth, going into much more details of the dynamics (several images per day) may result in noisy raw data requiring further treatment before exploiting, for example due to projected growth estimates interfering with circadian leaf movements [65].

Our work has been performed under two environmental conditions, a control condition and one that moderately limits growth due to water (but not nutrient) availability [34], i.e. a mild drought treatment. The QTL profiles obtained at the genomic scale are very robust to mild drought with most of the large effect QTL showing no clear signs of interaction with water availability in our conditions. Some of them still change their level of significance with conditions, but it is difficult to know if this is a real interaction with drought, a change in trait variance under stress, or a change in the rest of the genetic architecture of the trait (which will impact the significance of individual QTL). Condition-specific QTLs detected here always are of small effect, which also raises the question of the power of these comparisons across different conditions/experiments due to false negative in mapping QTL. Still, this result (a relatively smaller part of the phenotypic variation is plastic rather than constitutive) is similar to what was found previously for instance in linkage [69] and association mapping [61], showing an overlap in the network of genes that regulate plant size under control and mild drought conditions [60]. Drought stress might have pleiotropic effects on different tightly interrelated phenotypic traits and impose strong constraints on them, reducing trait variability [31]. Of course, this may highly depend on the type of stress (intensity, stage of application, stability . . .) that is applied. Here, we chose a mild stress intensity, to remain physiologically relevant and avoid the squeezing of trait variation concomitant to strong stress levels. Also, our robot is compensating (twice a day) the individual plant size effect on transpiration that may otherwise artificially increase stress intensity according to intrinsic plant size difference.

One advantage of using a star-like cross design, increasingly used in nested association mapping (NAM) populations, is to combine the power of individual crosses and take advantage of the comparison of multiple alleles, with respect to the reference allele in order to identify allelic variants more efficiently [70–73]. These could correspond to variants originating from Col-0 or to shared allelic variants among the other parents (same direction of the allelic effect in all crosses), or to allelic series (other parents could have divergent direction/intensity of effect with respect to the reference parent). This multi-population study has confirmed the effect of several loci across traits and environments, with particular power for compactness, and in some instances already allows to predict that independent loci actually underlie major

peaks. However, this approach is still limited by the mapping resolution which makes it difficult to distinguish shared variants from linked loci.

Considering the precision of our phenotyping (which has an impact on the part of phenotypic variation that is amenable to genetic dissection), the need for higher-resolution approaches to better describe the trait's genetic architecture is obvious here. If the architecture of variation in our crosses is more complex than just a limited number of loci independently segregating for intermediate or major effects, then the density of recombination observed in a simple RIL set will not allow to decipher the full architecture [74, 75]. We first went deeper in resolution by phenotyping numerous pairs of near-isogenic lines (actually, HIFs) that each interrogate a specific portion of a chromosome (2-3Mb on average) in an otherwise fixed genetic background. This nicely confirms a majority of loci but also shows some effects that were unexpected after the initial QTL mapping results, already indicating more complex genetic architecture than anticipated; this includes single peaks splitting up in independent loci or complex patterns of linkage versus pleiotropy (when comparing different traits) and linkage versus GxE (when comparing traits in different conditions). It seems that QTL colocalization among crosses (linkage versus shared variants) is also often questioned, although this requires to be able to compare multiple HIFs showing positive and/or negative results, which can be difficult for several reasons. Indeed, many factors can explain that a QTL is not validated in an HIF: the QTL could be under epistatic interaction with another locus (thus, a specific HIF may not represent the adequate genetic background), the QTL could be mislocalised by QTL mapping (thus, out of the HIF segregating region) or the HIF harbours a more complex genotype than expected at the segregating region, such as a double recombined region (thus, the HIF doesn't actually allow to test the whole region). This makes the comparison between HIFs difficult, even in the same cross, and negative HIF results should particularly be interpreted carefully. In this context, the rate of QTL validation obtained here is rather satisfying. The difficulties to identify genes responsible for complex phenotypes also depends on their involvement in epistatic interactions. HIFs are particularly sensitive to epistasis (compared to traditional NILs) as they each harbour a different genetic background, which means that they allow epistasis to be interrogated providing that one can test enough independent HIFs, otherwise they have to be compared with care.

Epistatic interactions are detected in all crosses / traits (although not always with very high significance) even when the trait heritability or variation is not so high in a cross, illustrating another factor of the complex genetic architecture. Some interactions seem to be condition-specific, but power issues are likely to be limiting in understanding these patterns of GxGxE. Furthermore, HIF have an expected candidate segregating region of several Mb usually, so our observations are likely just a glimpse at the real complexity of growth as it is known that linkage and epistasis is also active at a very local scale [44].

Still, the sensitivity of our approach here is validated by the detection of QTL colocalizing with several already-known QTG expected to segregate in our crosses, like *CRY2* as discussed above [52, 76], *MPK12* which would explain nicely the bottom of chromosome 2 QTL in Cvix-Col [77, 78], or SG3 detected here through its epistasis with SG3i in BurxCol [49, 79].

To avoid genome-wide epistasis and better reveal local-scale architecture, we have investigated in further details a single HIF background for a specific 3Mb region containing a known QTG of large and pleiotropic effect (*CRY2*) in an original approach. Our analysis reveals that there are at least 3 other QTGs in this interval controlling one of the traits in at least one condition. For PRA29, the picture seems to be even more complex with traces of at least one more locus; here, it seems that phenotyping accuracy becomes again limiting after all. Obviously these loci with opposite allelic effects, different patterns of pleiotropy and interaction with the environment, and just a few hundreds of kb from each other, remain cryptic in simple QTL

mapping. This major result of local-scale independent complex genetic architecture for different traits and conditions should lead us to a lot of caution when interpreting colocalizing QTLs from different traits / conditions / age, as these may very well be independent loci rather than a single pleiotropic locus, as shown here for PRA and Compactness. If we were to extrapolate the figures obtained on this specific region in this particular cross to the genome- and species-scale, we would expect hundreds of causative loci of detectable phenotypic effect controlling these growth-related phenotypes. One way to approach these individual loci would be to decompose their independent signature based on different dynamics or underlying traits (transcriptomics, metabolomics . . .) in a 'systems genetics' strategy [54, 60, 80].

Complex genetic architecture as revealed in this study has consequences on quantitative genetics experimental design and interpretation, arguing in favor of linkage mapping or GWAS depending on the balance between genetic complexity due to linked loci (where association is expected to behave better than linkage mapping) and genetic complexity due to small effect/rare alleles (where association will behave poorly). Other intermediate experimental designs like multiparental populations or nested-association mapping should bring more power [12, 81]. Resolution is improved by pushing recombination densities to its limits and it was shown to help resolve more complex genetic architecture in yeast [82]. In plants, using 'hyper-recombinant' mutations to generate new segregating populations could also be a strategy in the future [83].

Materials and methods

Genetic material. Recombinant Inbred Lines (RILs)

The 4 RIL sets used for this work were generated at the Versailles Arabidopsis Stock Center, France (http://publiclines.versailles.inra.fr/) and were either described previously [47] or on the Publiclines website where all relevant information (description and associated data) is gathered. Versailles stock center ID are indicated as 'xxxAV' and 'xxRV' for Accessions and RILs respectively). They are derived from crosses between the following pairs of accessions, chosen to maximize genetic and phenotypic diversity [84]:

RIL set 'BlaxCol' (ID = 2RV): Bla-1 (76AV) x Col-0 (186AV) / 259 RILs

RIL set 'CvixCol' (ID = 8RV): Cvi-0 (166AV) x Col-0 / 358 RILs

RIL set 'BurxCol' (ID = 20RV): Bur-0 (172AV) x Col-0 / 283 RILs

RIL set 'YoxCol' (ID = 23RV): Yo-0 (250AV) x Col-0 / 358 RILs

Heterogeneous Inbred Families (HIFs)

HIFs have been selected in each RIL set to cover several regions of the genome, some of which are expected to segregate for QTLs while others were chosen at random locations. As described previously [85] they are derived from the progeny of one RIL which is heterozygous only at the locus of interest. Hence, one HIF family is composed of 2 or 3 lines fixed for one parental allele at the segregating region and 2 or 3 lines fixed for the alternate parental allele at the segregating region, in an otherwise identical genetic background. Each HIF (ID = 'xxHVyyy') is named after the RIL ID ('xxRVyyy') from which it has been generated ('xx' is the ID of the RIL set, 'yyy' is the ID of the RIL): for example, the family 2HV142 is fixed from the RIL 2RV142. The tentative QTL validation is based on the phenotypic comparison of these fixed lines within the HIF family. We have used 81 HIF families. The complete dataset gathering genotypic information on the RILs used to generate HIFs has been submitted to INRA institutional data repository (https://data.inra.fr) [50], with the genotypic conventions and ID from the Versailles Arabidopsis Stock Center (http://publiclines.versailles.inra.fr/).

microStairs

The progeny of CvixCol RIL 8RV294 (also used to generate HIF family 8HV294), segregating for the first 3Mb of chromosome 1, has been screened to detect recombined individuals. 4000 individuals were genotyped with markers at the edge of the heterozygous region and then 29 evenly distributed recombinants were selected using markers spaced every ~100kb. These recombinants were genotyped and fixed for the remaining segregating region in such a way that they each differ genotypically from the next recombinant by a ~100kb bin on average (S4 Table). Similarly, as for the HIF, the descendance of 3 similarly-fixed lines are saved and phenotyped to account for possible maternal environmental effect. This is similar the approach taken by Koumproglou *et al.* [86], but at a much finer scale, hence the name '*microStairs*'.

Phenotyping

Phenotyping was performed on the Phenoscope robots as previously described [34] (https://phenoscope.versailles.inra.fr/). Every RIL set and their respective parental accessions have been phenotyped in 2 independent Phenoscope experiments (= biological replicates), except for Cvix-Col (3 biological replicates), with 1 individual (plant) per RIL per condition. In short, the peat-moss plugs' soil water content (SWC) is gradually adjusted for each plant individually as a fraction of the initially-saturated plug weight. We worked at 2 watering conditions: 60% SWC for non-limiting conditions (called 'WW' for well-watered) and 30% SWC for mildly growth-limiting watering conditions (called 'WD' for water deficit). The growth room is set at a 8 hours short-days photoperiod (230 μmol m-2 sec-1) with days at 21˚C/65%RH and nights at 17˚C/65%RH. A picture of each individual plant is taken every day at the same day-time and a semi-automatic segmentation process (with some manual corrections when required) is performed to extract leaf pixels. From this we exploit different traits: Projected Rosette Area (PRA), circle radius, convex hull area, average Red, Green and Blue components (leaf pixels, RGB colour scale), and derived phenotypic traits are calculated, such as the compactness (ratio PRA / convex hull area) and the Relative Expansion Rate (RER) over specific time windows (Fig 1), as previously described [34]. The complete raw phenotypic dataset has been submitted to INRA institutional data repository (https://data.inra.fr) [48]. Principal component analysis (PCA) was performed using the ade4 R package based on phenotypic data from all the RIL sets in WW and WD conditions.

Heritabilities were calculated based on the mean squares (MS) of the following ANOVA model, used as estimators of genetic and residual variances:

$$Y_{ij} \sim \mu + \alpha_i + \varepsilon_{ij}$$

where Y_{ij} is the phenotypic value, μ is the mean, α_i is the Genotype factor and ε_{ij} is the Residuals. The genetic variance VarG was estimated by $(MS_\alpha - MS_\varepsilon)/n$ (where n is the number of replicates). The residual variance VarR was estimated by MS_ε. $h^2 = VarG/(VarG+VarR)$.

Then, for QTL detection and further analyses, the phenotypic values were corrected for experiment effects. Corrected phenotypic values were calculated using the intercept (μ), the condition (α), genotype (γ), and genotype*condition (σ) effects of the following linear model:

$$Y_{ijkl} \sim \mu + \alpha_i + \beta_j + \gamma_k + \delta_{ij} + \lambda_{jk} + \sigma_{ik} + \varepsilon_{ijkl}$$

where Y_{ijkl}: phenotype; μ: mean; α_i: effect of the condition; β_j: effect of the experiment; γ_k: effect of the genotype; δ_{ij}: effect of the interaction condition*experiment; λ_{jk}: effect of the interaction experiment*genotype; σ_{ik}: effect of the interaction condition*genotype; ε_{ijkl}: residuals. The interaction term (independently used for QTL mapping) refers to the condition*genotype interaction component.

Note that one of the replicates of the CvixCol phenotypic data was already analyzed for QTL mapping -with a different statistical model- in Tisné *et al.* [34].

Similarly, the set of near isogenic lines (microStairs) were phenotyped and analyzed from 3 full biological replicates in independent Phenoscope experiments.

QTL MQM

QTL detections were performed using Multiple QTL Mapping algorithm (MQM) implemented in the R/qtl package [87, 88] using a backward selection of cofactors. At first, genotype missing data were augmented, then one marker every three markers were selected and used as cofactors. Important markers were selected through backward elimination. Finally, a QTL was moved along the genome using these pre-selected markers as cofactors, except for the markers in the 25.0 cM window around the region of interest. QTL were identified based on the most informative model through maximum likelihood. According to permutation results (computed with the mqmpermutation() function in R/qtl), a general LOD threshold of 2.4 was chosen for all QTL maps to ensure a FDR below 0.05 and remain conservative. Interactive QTL maps for time-course series were generated using the R/qtl charts package [89]. All QTL positions were projected on the consensus genetic map of the 4 crosses built with R/qtl (this is the map shown along the x axis on Figs 3, 5 and 6). A joint genotype dataset was constructed with 'A' alleles coding Col alleles (the common parent), 'B' alleles for non-Col alleles, and monomorphic markers in a cross coded as missing. The linkage groups were considered known from the individual maps and the physical position of markers, and a first marker order was calculated using orderMarkers() function. In case of conflicting marker order between individual, physical and consensus maps, the function switch.order() was used to retain the most probable order (i.e with the lowest number of recombination). The consensus genetic map markers' identities and positions are available together with the HIF dataset in Dataverse [50].

Epistatic interactions

Epistatic interactions were identified using the scantwo() function of the R/qtl package. LOD scores were calculated for additive, interaction and full models for all pairwise combination of markers, except for adjacent markers and a general conservative LOD threshold (3 LOD) was determined from permutations. Effect plots for the pairs of markers were drawn using the R/qtl package.

QTL mapping in the multi-cross design

QTL mapping in the multi-cross design was performed with the MCQTL package [90]. The model was described as additive (no dominance effect) and connected (Col-0 centered design) and the following 3 steps process was applied. Step 1: thresholds were calculated by trait on the whole genome using 1000 resampling replicates (PRA29_WW = 3.73; PRA29_WD = 3.84; RER16-29_WW = 4.22; RER16-29_WD = 3.95; Compactness29_WW = 3.64 Compactness29_WD = 3.30). Step 2: QTL detection was performed using iQTLm method with a general and conservative threshold of 4 LOD. To perform this detection, cofactors were automatically chosen by backward selection with a threshold of 2.8 LOD among a skeleton with a minimal inter distance of 10cM. Search for QTL was not allowed within +/-10cM window surrounding the QTL to avoid linked genetic regions. Step 3: model estimations were performed for each trait and condition using the shared QTL positions identified at step 2.

Fine dissection of a genomic region 'microStairs'

The phenotypes of the recombined HIFs lines were modeled using the following linear equation:

$$Y_{ij} \sim \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

where Y_{ij} is the value of the phenotype; μ is the mean of the phenotype; α_i is the effect of the stair (bin) i; β_j is the effect of the line j (maternal replicate within each stair) and ε_{ij} is the residuals. An anova was performed with this model and the p-value of the stair effects were adjusted by a Benjamini-Hochberg correction.

Polymorphic candidate genes (Cvi versus Col) were listed for each PRA29 'microStairs' significant interval according to variants listed on the 1001Genomes website (http://1001genomes.org/), through the Polymorph1001 tool. Differentially cis-regulated variants were extracted from Cvi/Col Allele-Specific Expression (ASE) data [57] and from CvixCol local-eQTLs data [91] across the whole region.

Datasets

Two datasets have been submitted to INRA Dataverse repository (https://data.inra.fr).

Raw phenotypic data obtained on the Arabidopsis RILs with the Phenoscope robots [48]

This dataset gathers the main raw phenotypic data obtained and exploited in Marchadier, Hanemian, Tisné *et al.* (2018). It contains data from 4 RIL sets across 9 Phenoscope experiments.

For each Phenoscope experiment, Recombinant Inbred Line (RIL) and Condition ('WW' = Well Watered / 'WD' = Water Deficit), the data set indicates the phenotypic value for 6 traits at 21 successive time points. 'Trait.XX' = Trait at XX days after sowing, with 'XX' = 09 to 29 and 'Trait' = PRA (Projected Rosette Area; in cm2), GreenMean / RedMean / BlueMean (rosette pixels' colour components; arbitrary unit), ConvexHullArea (area of the convex hull encompassing the rosette; in cm2) and CircleRadius (radius of the smallest circle encompassing the rosette; in cm). RIL set IDs and RIL IDs are according to Publiclines http://publiclines.versailles.inra.fr/rils/index

Genotypic description of the near isogenic lines (HIFs) used for QTL validation and significance of the observed segregating phenotypes [50]

Each row represents a single HIF and the genotype of the F7 RIL it originates from is indicated along the chromosomes with RIL ID, markers and genotypic conventions from Publiclines http://publiclines.versailles.inra.fr/rils/index (i.e. 'A' = Col allele; 'B' = alternate parental allele; 'C' = heterozygous). The region highlighted in yellow is the segregating region that is tested in the HIF family through several fixed lines for each parental allele. For each of the 3 growth traits (Compactness29 = rosette compactness 29 days after sowing; PRA29 = Projected Rosette Area 29 days after sowing; RER16-29 = Relative rosette Expansion Rate between days 16 and 29 after sowing) in 2 conditions ('WW' = Well Watered / 'WD' = Water Deficit), whenever significant, the p-value of the comparison between allelic lines ('Pval') and the direction of the allelic effect ('sign' calculated as [Col-Xxx] where Xxx is the alternate parental allele) are indicated in the last columns of the table.

A specific webpage is associated with this work to display interactive graphes for dynamic QTL analyses at http://www.inra.fr/vast/PhenoDynamic.htm

Supporting information

S1 Fig. Distributions of the mean phenotypic values across RILs. Frequency histograms showing the distributions of the PRA29, RER16-29 and Compactness29 traits within each of the four RIL sets in well-watered (WW: blue) and water deficit (WD: orange) conditions. Phenotypic values for parental accessions from these specific experiments are indicated by blue (WW) and red (WD) ticks (see inset for legend) just above the x axes.
(PDF)

S2 Fig. Distribution of QTL effects. Distribution of the QTL effects (R2%) in the four RILs populations (CvixCol, BurxCol, BlaxCol, YoxCol), two conditions (WW and WD) and for the 3 traits considered in the analysis (PRA29, RER16-29, Compactness29).
(PDF)

S3 Fig. 2D scans for epistasis. Heatmaps representing the LOD score of the interaction effects between all pairs of loci for 4 RIL sets, 3 traits, in 2 conditions. Each heatmap shows the pairwise interaction effects obtained in control condition WW (triangle above diagonal) and in water deficit condition WD (triangle below diagonal). The color scale (LOD score values) shared among all heatmaps is indicated on the right (note that it is different from the scale of Fig 4). Diagonal values are canceled and enlarged to exclude pairs of adjacent markers from the test.
(PDF)

S4 Fig. Multi-cross QTL analysis for RER16-29. Same legend as Fig 5. Chromosome 3 shows no significant combined QTLs for RER16-29.
(PDF)

S5 Fig. Multi-cross QTL analysis for Compactness29. Same legend as Fig 5.
(PDF)

S6 Fig. Near isogenic lines-based validation of QTLs for RER16-29 and Compactness29. Same legend as Fig 6.
(PDF)

S7 Fig. Dissection of a genomic region in CvixCol for RER16-29 and Compactness29 ('microStairs' approach). Same legend as Fig 7.
(PDF)

S1 Table. Heritabilities of the observed phenotypes. Broad-sense heritabilities (h2) in the 4 RIL sets (BlaxCol, CvixCol, BurxCol and YoxCol) for PRA29, RER16-29 and Compactness29 in WW and WD conditions. H2 > 0.4 are highlighted in bold. Broad-sense heritabilities were calculated with the following equation h2 = Var(G)/Var(P) with Var(P) = Var(G)+Var(E)
(PDF)

S2 Table. Analyses of variance for genotype, experiment and condition factors, and their interactions. For each RIL set, -log(p values) of main and interaction effects calculated by the analysis of variance (ANOVA) from the following model:
$Y_{ijkl} \sim \mu + \alpha_i + \beta_j + \gamma_k + \delta_{ij} + \lambda_{jk} + \sigma_{ik} + \varepsilon_{ijkl}$.
where Y_{ijkl}: phenotype; μ: mean; α_i: effect of the condition; β_j: effect of the experiment; γ_k: effect of the genotype; δ_{ij}: effect of the interaction condition*experiment; λ_{jk}: effect of the interaction experiment*genotype; σ_{ik}: effect of the interaction condition*genotype; ε_{ijkl}: residuals.

NS = Not Significant.
(PDF)

S3 Table. Mapped QTL parameters across RIL sets, traits and conditions. For each RIL set, trait and condition (or GxE interaction term: 'inter'): every independent QTL peak is a row in the table, indicating its localisation (chromosome and position on the genetic map), its significance (maximum LOD score) and its effect (direction of the allelic effect: 'sign' indicates the sign of the allelic effect estimated as [Col–Xxx], where Xxx is the alternate parental allele; R2%). Independent QTLs peaks are considered when the LOD score curve returns below threshold between 2 peaks, to ensure the loci are not too genetically linked.
(PDF)

S4 Table. microStairs; genotypic description of recombinant lines. The genotypes of the 31 lines selected for the microStairs experiment are indicated along this 3Mb region of the top of chromosome 1 with markers localized on the physical map (kb).
(PDF)

S5 Table. microStairs; polymorphisms segregating in different bins. This table lists all known polymorphisms between Col-0 and Cvi-0 that can potentially affect gene function (according to 1001Genomes data) within PRA29-significant bins only: in the selected physical interval, genes with more than 3 non-synonymous SNP or a high-impact polymorphism are highlighted in yellow and orange respectively. *CRY2* candidate gene is highlighted in red.
(PDF)

S6 Table. microStairs; differentially expressed genes controlled in cis. List of Cvi/Col differentially expressed genes in an allele-specific fashion ([log2]>0.8 in at least one condition; from Cubillos *et al*., 2014) across the whole region.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Elodie Marchadier, Mathieu Hanemian, Sébastien Tisné, Olivier Loudet.

Data curation: Elodie Marchadier, Mathieu Hanemian, Sébastien Tisné, Liên Bach, Christos Bazakos, Elodie Gilbault, Parham Haddadi, Laetitia Virlouvet.

Formal analysis: Elodie Marchadier, Mathieu Hanemian, Sébastien Tisné, Liên Bach, Christos Bazakos, Elodie Gilbault, Parham Haddadi, Laetitia Virlouvet, Olivier Loudet.

Funding acquisition: Olivier Loudet.

Investigation: Olivier Loudet.

Methodology: Olivier Loudet.

Project administration: Olivier Loudet.

Supervision: Olivier Loudet.

Validation: Elodie Marchadier, Mathieu Hanemian, Sébastien Tisné.

Writing – original draft: Elodie Marchadier, Mathieu Hanemian, Sébastien Tisné, Olivier Loudet.

Writing – review & editing: Liên Bach, Christos Bazakos, Elodie Gilbault, Parham Haddadi, Laetitia Virlouvet, Olivier Loudet.

References

1. Alonso-Blanco C, Mendez-Vigo B. Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. Curr Opin Plant Biol. 2014; 18:37–43. Epub 2014/02/26. https://doi.org/10.1016/j.pbi.2014.01.002 PMID: 24565952.

2. Bac-Molenaar JA, Vreugdenhil D, Granier C, Keurentjes JJ. Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. J Exp Bot. 2015; 66(18):5567–80. Epub 2015/04/30. https://doi.org/10.1093/jxb/erv176 PMID: 25922493; PubMed Central PMCID: PMC4585414.

3. Bouteille M, Rolland G, Balsera C, Loudet O, Muller B. Disentangling the intertwined genetic bases of root and shoot growth in Arabidopsis. PLoS ONE. 2012; 7(2):e32319. Epub 2012/03/03. https://doi.org/10.1371/journal.pone.0032319 PMID: 22384215; PubMed Central PMCID: PMC3286473.

4. Geng Y, Wu R, Wee CW, Xie F, Wei X, Chan PM, et al. A spatio-temporal understanding of growth regulation during the salt stress response in Arabidopsis. Plant Cell. 2013; 25(6):2132–54. Epub 2013/07/31. https://doi.org/10.1105/tpc.113.112896 PMID: 23898029; PubMed Central PMCID: PMC3723617.

5. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, et al. A gene expression map of Arabidopsis thaliana development. Nat Genet. 2005; 37(5):501–6. Epub 2005/04/05. https://doi.org/10.1038/ng1543 PMID: 15806101.

6. Zhang X, Hause RJ Jr., Borevitz JO. Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in *Arabidopsis thaliana*. G3 (Bethesda). 2012; 2(1):29–34. Epub 2012/03/03. https://doi.org/10.1534/g3.111.001487 PMID: 22384379; PubMed Central PMCID: PMC3276187.

7. Clauw P, Coppens F, De Beuf K, Dhondt S, Van Daele T, Maleux K, et al. Leaf responses to mild drought stress in natural variants of Arabidopsis. Plant Physiol. 2015; 167(3):800–16. Epub 2015/01/22. https://doi.org/10.1104/pp.114.254284 PMID: 25604532; PubMed Central PMCID: PMC4348775.

8. Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, et al. 50 years of Arabidopsis research: highlights and future directions. New Phytol. 2016; 209(3):921–44. Epub 2015/10/16. https://doi.org/10.1111/nph.13687 PMID: 26465351.

9. Gonzalez N, Inze D. Molecular systems governing leaf growth: from genes to networks. J Exp Bot. 2015; 66(4):1045–54. Epub 2015/01/21. https://doi.org/10.1093/jxb/eru541 PMID: 25601785.

10. Vanhaeren H, Gonzalez N, Coppens F, De Milde L, Van Daele T, Vermeersch M, et al. Combining growth-promoting genes leads to positive epistasis in Arabidopsis thaliana. Elife. 2014; 3:e02252. Epub 2014/05/21. https://doi.org/10.7554/eLife.02252 PMID: 24843021; PubMed Central PMCID: PMC4014012.

11. Trontin C, Tisné S, Bach L, Loudet O. What does Arabidopsis natural variation teach us (and does not teach us) about adaptation in plants? Curr Opin Plant Biol. 2011; 14(3):225–31. https://doi.org/10.1016/j.pbi.2011.03.024 PMID: 21536479

12. Bazakos C, Hanemian M, Trontin C, Jimenez-Gomez JM, Loudet O. New Strategies and Tools in Quantitative Genetics: How to Go from the Phenotype to the Genotype. Annu Rev Plant Biol. 2017; 68:435–55. Epub 2017/02/23. https://doi.org/10.1146/annurev-arplant-042916-040820 PMID: 28226236.

13. Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D, et al. What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell. 2009; 21(7):1877–96. https://doi.org/10.1105/tpc.109.068114 PMID: 19574434

14. Grillo MA, Li C, Hammond M, Wang L, Schemske DW. Genetic architecture of flowering time differentiation between locally adapted populations of Arabidopsis thaliana. New Phytol. 2013; 197(4):1321–31. Epub 2013/01/15. https://doi.org/10.1111/nph.12109 PMID: 23311994.

15. Salome PA, Bomblies K, Laitinen RA, Yant L, Mott R, Weigel D. Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. Genetics. 2011; 188(2):421–33. Epub 2011/03/17. https://doi.org/10.1534/genetics.111.126607 PMID: 21406681; PubMed Central PMCID: PMC3122318.

16. Tabas-Madrid D, Mendez-Vigo B, Arteaga N, Marcer A, Pascual-Montano A, Weigel D, et al. Genome-wide signatures of flowering adaptation to climate temperature: Regional analyses in a highly diverse

native range of Arabidopsis thaliana. Plant Cell Environ. 2018; 41(8):1806–20. Epub 2018/03/10. https://doi.org/10.1111/pce.13189 PMID: 29520809.

17. Johanson U, West J, Lister C, Michaels SD, Amasino RM, Dean C. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in Arabidopsis flowering time. Science. 2000; 290:344–7. PMID: 11030654

18. Michaels SD, Amasino RM. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell. 1999; 11:949–56. PMID: 10330478

19. Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, et al. Role of FRIGIDA and FLOW-ERING LOCUS C in determining variation in flowering time of Arabidopsis. Plant Physiol. 2005; 138 (2):1163–73. Epub 2005/05/24. https://doi.org/10.1104/pp.105.061309 PMID: 15908596; PubMed Central PMCID: PMC1150429.

20. Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, et al. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature. 2010; 465(7298):627–31. Epub 2010/ 03/26. https://doi.org/10.1038/nature08800 PMID: 20336072.

21. Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, et al. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet. 2010; 6:e1000940. Epub 2010/05/ 14. https://doi.org/10.1371/journal.pgen.1000940 PMID: 20463887; PubMed Central PMCID: PMC2865524.

22. El-Lithy ME, Clerkx EJ, Ruys GJ, Koornneef M, Vreugdenhil D. Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. Plant Physiol. 2004; 135 (1):444–58. Epub 2004/05/04. https://doi.org/10.1104/pp.103.036822 PMID: 15122039; PubMed Central PMCID: PMC429397.

23. Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F. Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. Theor Appl Genet. 2002; 104:1173–84. https://doi.org/10.1007/s00122-001-0825-9 PMID: 12582628

24. Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F. Quantitative Trait Loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiol. 2003; 131(1):345–58. https://doi.org/10.1104/pp.102.010785 PMID: 12529542

25. Tisné S, Schmalenbach I, Reymond M, Dauzat M, Pervent M, Vile D, et al. Keep on growing under drought: genetic and developmental bases of the response of rosette area using a recombinant inbred line population. Plant Cell Environ. 2010; 33(11):1875–87. Epub 2010/06/16. https://doi.org/10.1111/j.1365-3040.2010.02191.x PMID: 20545881.

26. Tonsor SJ, Elnaccash TW, Scheiner SM. Developmental instability is genetically correlated with phenotypic plasticity, constraining heritability, and fitness. Evolution. 2013; 67(10):2923–35. Epub 2013/10/08. https://doi.org/10.1111/evo.12175 PMID: 24094343.

27. Dittmar EL, Oakley CG, Agren J, Schemske DW. Flowering time QTL in natural populations of Arabidopsis thaliana and implications for their adaptive value. Mol Ecol. 2014; 23(17):4291–303. Epub 2014/07/22. https://doi.org/10.1111/mec.12857 PMID: 25039363.

28. El-Soda M, Malosetti M, Zwaan BJ, Koornneef M, Aarts MG. Genotype x environment interaction QTL mapping in plants: lessons from Arabidopsis. Trends Plant Sci. 2014; 19(6):390–8. Epub 2014/02/05. https://doi.org/10.1016/j.tplants.2014.01.001 PMID: 24491827.

29. Kooyers NJ. The evolution of drought escape and avoidance in natural herbaceous populations. Plant Sci. 2015; 234:155–62. Epub 2015/03/26. https://doi.org/10.1016/j.plantsci.2015.02.012 PMID: 25804818.

30. Exposito-Alonso M, Vasseur F, Ding W, Wang G, Burbano HA, Weigel D. Genomic basis and evolutionary potential for extreme drought adaptation in Arabidopsis thaliana. Nat Ecol Evol. 2018; 2(2):352–8. Epub 2017/12/20. https://doi.org/10.1038/s41559-017-0423-0 PMID: 29255303; PubMed Central PMCID: PMC5777624.

31. Juenger TE. Natural variation and genetic constraints on drought tolerance. Curr Opin Plant Biol. 2013; 16(3):274–81. Epub 2013/03/07. https://doi.org/10.1016/j.pbi.2013.02.001 PMID: 23462639.

32. 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(3):623–35. https://doi.org/10.1111/j.1469-8137.2005.01609.x PMID: 16411964

33. Skirycz A, Vandenbroucke K, Clauw P, Maleux K, De Meyer B, Dhondt S, et al. Survival and growth of Arabidopsis plants given limited water are not equal. Nat Biotechnol. 2011; 29(3):212–4. Epub 2011/03/11. https://doi.org/10.1038/nbt.1800 PMID: 21390020.

34. Tisné 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. 2013; 74:534–44. Epub 2013/03/05. https://doi.org/10.1111/tpj.12131 PMID: 23452317.

35. Walter A, Scharr H, Gilmer F, Zierer R, Nagel KA, Ernst M, et al. Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. New Phytol. 2007; 174(2):447–55. Epub 2007/03/29. https://doi.org/10.1111/j.1469-8137.2007.02002.x PMID: 17388907.

36. Barboza L, Effgen S, Alonso-Blanco C, Kooke R, Keurentjes JJ, Koornneef M, et al. Arabidopsis semi-dwarfs evolved from independent mutations in GA20ox1, ortholog to green revolution dwarf alleles in rice and barley. Proc Natl Acad Sci USA. 2013; 110(39):15818–23. Epub 2013/09/12. https://doi.org/10.1073/pnas.1314979110 PMID: 24023067; PubMed Central PMCID: PMC3785751.

37. Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, et al. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol. 2007; 5(9): e236. https://doi.org/10.1371/journal.pbio.0050236 PMID: 17803357

38. Loudet O, Michael TP, Burger BT, Le Mette C, Mockler TC, Weigel D, et al. A zinc knuckle protein that negatively controls morning-specific growth in *Arabidopsis thaliana*. Proc Natl Acad Sci USA. 2008; 105 (44):17193–8. Epub 2008/10/31. https://doi.org/10.1073/pnas.0807264105 PMID: 18971337; PubMed Central PMCID: PMC2579400.

39. Masle J, Gilmore SR, Farquhar GD. The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. Nature. 2005; 436(7052):866–70. Epub 2005/07/12. https://doi.org/10.1038/nature03835 PMID: 16007076.

40. Mouchel CF, Briggs GC, Hardtke CS. Natural genetic variation in Arabidopsis identifies *BREVIS RADIX*, a novel regulator of cell proliferation and elongation in the root. Genes Dev. 2004; 18(6):700–14. https://doi.org/10.1101/gad.1187704 PMID: 15031265

41. Sureshkumar S, Todesco M, Schneeberger K, Harilal R, Balasubramanian S, Weigel D. A genetic defect caused by a triplet repeat expansion in *Arabidopsis thaliana*. Science. 2009; 323(5917):1060–3. Epub 2009/01/20. 1164014 [pii] https://doi.org/10.1126/science.1164014 PMID: 19150812.

42. Todesco M, Balasubramanian S, Hu TT, Traw MB, Horton M, Epple P, et al. Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. Nature. 2010; 465(7298):632–6. Epub 2010/06/04. https://doi.org/10.1038/nature09083 PMID: 20520716.

43. Trontin C, Kiani S, Corwin JA, Hematy K, Yansouni J, Kliebenstein DJ, et al. A pair of receptor-like kinases is responsible for natural variation in shoot growth response to mannitol treatment in *Arabidopsis thaliana*. Plant J. 2014; 78(1):121–33. Epub 2014/02/01. https://doi.org/10.1111/tpj.12454 PMID: 24479634.

44. Kroymann J, Mitchell-Olds T. Epistasis and balanced polymorphism influencing complex trait variation. Nature. 2005; 435(7038):95–8. https://doi.org/10.1038/nature03480 PMID: 15875023

45. Prinzenberg AE, Barbier H, Salt DE, Stich B, Reymond M. Relationships between growth, growth response to nutrient supply, and ion content using a recombinant inbred line population in Arabidopsis. Plant Physiol. 2010; 154(3):1361–71. Epub 2010/09/10. https://doi.org/10.1104/pp.110.161398 PMID: 20826703.

46. Lemmon ZH, Doebley JF. Genetic dissection of a genomic region with pleiotropic effects on domestication traits in maize reveals multiple linked QTL. Genetics. 2014; 198(1):345–53. Epub 2014/06/22. https://doi.org/10.1534/genetics.114.165845 PMID: 24950893; PubMed Central PMCID: PMC4174946.

47. Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, et al. QTL mapping in five new large RIL populations of *Arabidopsis thaliana* genotyped with consensus SNP markers. Genetics. 2008; 178:2253–64. https://doi.org/10.1534/genetics.107.083899 PMID: 18430947

48. Loudet O. "Raw phenotypic data obtained on the Arabidopsis RILs with the Phenoscope robots (Marchadier, Hanemian, Tisné et al., 2018)", https://doi.org/10.15454/OCOP9B, Portail Data Inra, V1 2018.

49. Vlad D, Rappaport F, Simon M, Loudet O. Gene transposition causing natural variation for growth in *Arabidopsis thaliana*. PLoS Genet. 2010; 6(5):e1000945. https://doi.org/10.1371/journal.pgen.1000945 PMID: 20485571

50. Loudet O. "Genotypic description of the near isogenic lines (HIFs) used for QTL validation and significance of the observed segregating phenotypes (Marchadier, Hanemian, Tisné et al., 2018)", https://doi.org/10.15454/EORHL8, Portail Data Inra, V1. 2018.

51. Botto JF, Alonso-Blanco C, Garzaron I, Sanchez RA, Casal JJ. The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon unfolding in the absence of blue light in Arabidopsis. Plant Physiol. 2003; 133(4):1547–456. Epub 2003/11/08. https://doi.org/10.1104/pp.103.029546 PMID: 14605225; PubMed Central PMCID: PMC300712.

52. El-Din El-Assal S, Alonso-Blanco C, Peeters AJ, Raz V, Koornneef M. A QTL for flowering time in Arabidopsis reveals a novel allele of *CRY2*. Nat Genet. 2001; 29(4):435–40. https://doi.org/10.1038/ng767 PMID: 11726930.

53. Tognetti VB, Van Aken O, Morreel K, Vandenbroucke K, van de Cotte B, De Clercq I, et al. Perturbation of indole-3-butyric acid homeostasis by the UDP-glucosyltransferase UGT74E2 modulates Arabidopsis architecture and water stress tolerance. Plant Cell. 2010; 22(8):2660–79. Epub 2010/08/28. https://doi.org/10.1105/tpc.109.071316 PMID: 20798329; PubMed Central PMCID: PMC2947170.

54. Nijveen H, Ligterink W, Keurentjes JJ, Loudet O, Long J, Sterken MG, et al. AraQTL—workbench and archive for systems genetics in *Arabidopsis thaliana*. Plant J. 2017; 89(6):1225–35. Epub 2016/12/21. https://doi.org/10.1111/tpj.13457 PMID: 27995664.

55. Meyerhoff O, Muller K, Roelfsema MR, Latz A, Lacombe B, Hedrich R, et al. AtGLR3.4, a glutamate receptor channel-like gene is sensitive to touch and cold. Planta. 2005; 222(3):418–27. Epub 2005/05/03. https://doi.org/10.1007/s00425-005-1551-3 PMID: 15864638.

56. Vincill ED, Clarin AE, Molenda JN, Spalding EP. Interacting glutamate receptor-like proteins in Phloem regulate lateral root initiation in Arabidopsis. Plant Cell. 2013; 25(4):1304–13. Epub 2013/04/18. https://doi.org/10.1105/tpc.113.110668 PMID: 23590882; PubMed Central PMCID: PMC3663269.

57. Cubillos FA, Stegle O, Grondin C, Canut M, Tisne S, Gy I, et al. Extensive *cis*-regulatory variation robust to environmental perturbation in Arabidopsis. Plant Cell. 2014; 26(11):4298–310. Epub 2014/11/28. https://doi.org/10.1105/tpc.114.130310 PMID: 25428981.

58. Waterworth WM, Kozak J, Provost CM, Bray CM, Angelis KJ, West CE. DNA ligase 1 deficient plants display severe growth defects and delayed repair of both DNA single and double strand breaks. BMC Plant Biol. 2009; 9:79. Epub 2009/06/30. https://doi.org/10.1186/1471-2229-9-79 PMID: 19558640; PubMed Central PMCID: PMC2708163.

59. Zhao H, Lu S, Li R, Chen T, Zhang H, Cui P, et al. The Arabidopsis gene DIG6 encodes a large 60S sub-unit nuclear export GTPase 1 that is involved in ribosome biogenesis and affects multiple auxin-regulated development processes. J Exp Bot. 2015; 66(21):6863–75. Epub 2015/08/15. https://doi.org/10.1093/jxb/erv391 PMID: 26272902; PubMed Central PMCID: PMC4623693.

60. Clauw P, Coppens F, Korte A, Herman D, Slabbinck B, Dhondt S, et al. Leaf Growth Response to Mild Drought: Natural Variation in Arabidopsis Sheds Light on Trait Architecture. Plant Cell. 2016; 28:2417–34. Epub 2016/10/13. https://doi.org/10.1105/tpc.16.00483 PMID: 27729396.

61. Bac-Molenaar JA, Granier C, Keurentjes JJ, Vreugdenhil D. Genome-wide association mapping of time-dependent growth responses to moderate drought stress in Arabidopsis. Plant Cell Environ. 2016; 39(1):88–102. Epub 2015/07/04. https://doi.org/10.1111/pce.12595 PMID: 26138664.

62. Kooke R, Kruijer W, Bours R, Becker F, Kuhn A, van de Geest H, et al. Genome-Wide Association Mapping and Genomic Prediction Elucidate the Genetic Architecture of Morphological Traits in Arabidopsis. Plant Physiol. 2016; 170(4):2187–203. Epub 2016/02/13. https://doi.org/10.1104/pp.15.00997 PMID: 26869705; PubMed Central PMCID: PMC4825126.

63. Meijon M, Satbhai SB, Tsuchimatsu T, Busch W. Genome-wide association study using cellular traits identifies a new regulator of root development in Arabidopsis. Nat Genet. 2014; 46(1):77–81. Epub 2013/11/12. https://doi.org/10.1038/ng.2824 PMID: 24212884.

64. Rishmawi L, Buhler J, Jaegle B, Hulskamp M, Koornneef M. Quantitative trait loci controlling leaf venation in Arabidopsis. Plant Cell Environ. 2017; 40(8):1429–41. Epub 2017/03/03. https://doi.org/10.1111/pce.12938 PMID: 28252189.

65. Flood PJ, Kruijer W, Schnabel SK, van der Schoor R, Jalink H, Snel JF, et al. Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. Plant Methods. 2016; 12:14. Epub 2016/02/18. https://doi.org/10.1186/s13007-016-0113-y PMID: 26884806; PubMed Central PMCID: PMC4754911.

66. Zhang X, Huang C, Wu D, Qiao F, Li W, Duan L, et al. High-Throughput Phenotyping and QTL Mapping Reveals the Genetic Architecture of Maize Plant Growth. Plant Physiol. 2017; 173(3):1554–64. Epub 2017/02/06. https://doi.org/10.1104/pp.16.01516 PMID: 28153923; PubMed Central PMCID: PMC5338669.

67. Muraya MM, Chu J, Zhao Y, Junker A, Klukas C, Reif JC, et al. Genetic variation of growth dynamics in maize (Zea mays L.) revealed through automated non-invasive phenotyping. Plant J. 2017; 89(2):366–80. Epub 2016/10/08. https://doi.org/10.1111/tpj.13390 PMID: 27714888.

68. Moore CR, Johnson LS, Kwak IY, Livny M, Broman KW, Spalding EP. High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. Genetics. 2013; 195(3):1077–86. Epub 2013/08/28. https://doi.org/10.1534/genetics.113.153346 PMID: 23979570; PubMed Central PMCID: PMC3813838.

69. Mojica JP, Mullen J, Lovell JT, Monroe JG, Paul JR, Oakley CG, et al. Genetics of water use physiology in locally adapted Arabidopsis thaliana. Plant Sci. 2016; 251:12–22. Epub 2016/09/07. https://doi.org/10.1016/j.plantsci.2016.03.015 PMID: 27593459.

70. Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L. Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor Appl Genet. 2006; 113(2):206–24. Epub 2006/06/23. https://doi.org/10.1007/s00122-006-0287-1 PMID: 16791688.

71. Nice LM, Steffenson BJ, Brown-Guedira GL, Akhunov ED, Liu C, Kono TJ, et al. Development and Genetic Characterization of an Advanced Backcross-Nested Association Mapping (AB-NAM) Population of Wild x Cultivated Barley. Genetics. 2016; 203(3):1453–67. Epub 2016/05/18. https://doi.org/10.1534/genetics.116.190736 PMID: 27182953; PubMed Central PMCID: PMC4937491.

72. Yu J, Holland JB, McMullen MD, Buckler ES. Genetic design and statistical power of nested association mapping in maize. Genetics. 2008; 178(1):539–51. Epub 2008/01/19. https://doi.org/10.1534/genetics.107.074245 PMID: 18202393; PubMed Central PMCID: PMC2206100.

73. Brenton ZW, Cooper EA, Myers MT, Boyles RE, Shakoor N, Zielinski KJ, et al. A Genomic Resource for the Development, Improvement, and Exploitation of Sorghum for Bioenergy. Genetics. 2016; 204 (1):21–33. Epub 2016/07/01. https://doi.org/10.1534/genetics.115.183947 PMID: 27356613; PubMed Central PMCID: PMC5012387.

74. Huang YF, Madur D, Combes V, Ky CL, Coubriche D, Jamin P, et al. The genetic architecture of grain yield and related traits in Zea maize L. revealed by comparing intermated and conventional populations. Genetics. 2010; 186(1):395–404. Epub 2010/07/02. https://doi.org/10.1534/genetics.110.113878 PMID: 20592258; PubMed Central PMCID: PMC2940303.

75. Studer AJ, Doebley JF. Do large effect QTL fractionate? A case study at the maize domestication QTL teosinte branched1. Genetics. 2011; 188(3):673–81. Epub 2011/04/26. https://doi.org/10.1534/genetics.111.126508 PMID: 21515578; PubMed Central PMCID: PMC3176526.

76. Sanchez-Bermejo E, Zhu W, Tasset C, Eimer H, Sureshkumar S, Singh R, et al. Genetic Architecture of Natural Variation in Thermal Responses of Arabidopsis. Plant Physiol. 2015; 169(1):647–59. Epub 2015/07/22. https://doi.org/10.1104/pp.15.00942 PMID: 26195568; PubMed Central PMCID: PMC4577429.

77. Des Marais DL, Auchincloss LC, Sukamtoh E, McKay JK, Logan T, Richards JH, et al. Variation in MPK12 affects water use efficiency in Arabidopsis and reveals a pleiotropic link between guard cell size and ABA response. Proc Natl Acad Sci USA. 2014; 111(7):2836–41. Epub 2014/02/20. https://doi.org/10.1073/pnas.1321429111 PMID: 24550314; PubMed Central PMCID: PMC3932934.

78. Jakobson L, Vaahtera L, Toldsepp K, Nuhkat M, Wang C, Wang YS, et al. Natural Variation in Arabidopsis Cvi-0 Accession Reveals an Important Role of MPK12 in Guard Cell CO2 Signaling. PLoS Biol. 2016; 14(12):e2000322. Epub 2016/12/07. https://doi.org/10.1371/journal.pbio.2000322 PMID: 27923039; PubMed Central PMCID: PMC5147794.

79. Michael TP, Jupe F, Bemm F, Motley ST, Sandoval JP, Lanz C, et al. High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. Nat Commun. 2018; 9(1):541. Epub 2018/02/09. https://doi.org/10.1038/s41467-018-03016-2 PMID: 29416032; PubMed Central PMCID: PMC5803254.

80. Civelek M, Lusis AJ. Systems genetics approaches to understand complex traits. Nat Rev Genet. 2014; 15(1):34–48. Epub 2013/12/04. https://doi.org/10.1038/nrg3575 PMID: 24296534; PubMed Central PMCID: PMC3934510.

81. de Koning DJ, McIntyre LM. Back to the Future: Multiparent Populations Provide the Key to Unlocking the Genetic Basis of Complex Traits. Genetics. 2017; 206(2):527–9. Epub 2017/06/09. https://doi.org/10.1534/genetics.117.203265 PMID: 28592493; PubMed Central PMCID: PMC5494722.

82. She R, Jarosz DF. Mapping Causal Variants with Single-Nucleotide Resolution Reveals Biochemical Drivers of Phenotypic Change. Cell. 2018; 172(3):478–90. Epub 2018/01/27. https://doi.org/10.1016/j.cell.2017.12.015 PMID: 29373829; PubMed Central PMCID: PMC5788306.

83. Fernandes JB, Seguela-Arnaud M, Larcheveque C, Lloyd AH, Mercier R. Unleashing meiotic crossovers in hybrid plants. Proc Natl Acad Sci USA. 2018; 115(10):2431–6. Epub 2017/12/01. https://doi.org/10.1073/pnas.1713078114 PMID: 29183972; PubMed Central PMCID: PMC5877974.

84. McKhann HI, Camilleri C, Berard A, Bataillon T, David JL, Reboud X, et al. Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. Plant J. 2004; 38(1):193–202. https://doi.org/10.1111/j.1365-313X.2004.02034.x PMID: 15053772

85. Loudet O, Gaudon V, Trubuil A, Daniel-Vedele F. Quantitative trait loci controlling root growth and architecture in *Arabidopsis thaliana* confirmed by heterogeneous inbred family. Theor Appl Genet. 2005; 110 (4):742–53. https://doi.org/10.1007/s00122-004-1900-9 PMID: 15678326

86. Koumproglou R, Wilkes TM, Townson P, Wang XY, Beynon J, Pooni HS, et al. STAIRS: a new genetic resource for functional genomic studies of Arabidopsis. Plant J. 2002; 31(3):355–64. PMID: 12164814.

87. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 2003; 19(7):889–90. Epub 2003/05/02. PMID: 12724300.

88. Arends D, Prins P, Jansen RC, Broman KW. R/qtl: high-throughput multiple QTL mapping. Bioinformatics. 2010; 26(23):2990–2. Epub 2010/10/23. https://doi.org/10.1093/bioinformatics/btq565 PMID: 20966004; PubMed Central PMCID: PMC2982156.

89. Broman KW. R/qtlcharts: interactive graphics for quantitative trait locus mapping. Genetics. 2015; 199 (2):359–61. Epub 2014/12/21. https://doi.org/10.1534/genetics.114.172742 PMID: 25527287; PubMed Central PMCID: PMC4317647.

90. Jourjon MF, Jasson S, Marcel J, Ngom B, Mangin B. MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics. 2005; 21(1):128–30. Epub 2004/08/21. https://doi.org/10.1093/bioinformatics/bth481 PMID: 15319261.

91. Cubillos FA, Yansouni J, Khalili H, Balzergue S, Elftieh S, Martin-Magniette ML, et al. Expression variation in connected recombinant populations of *Arabidopsis thaliana* highlights distinct transcriptome architectures. BMC Genomics. 2012; 13(1):117. Epub 2012/03/29. https://doi.org/10.1186/1471-2164-13-117 PMID: 22453064.