Genetics: Early Online, published on August 26, 2013 as 10.1534/genetics.113.153346

High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response

Candace R. Moore[*], Logan S. Johnson[*], Il-youp Kwak[§], Miron Livny[†]

Karl W. Broman[‡], Edgar P. Spalding[*]

[*]Department of Botany

[§]Department of Statistics

[†]Department of Computer Sciences

[‡]Department of Biostatistics and Medical Informatics

University of Wisconsin-Madison
Madison, Wisconsin 53706
USA

Short title: Adding developmental time to a quantitative trait map

Key words: root gravitropism, QTL, natural variation, Arabidopsis

Corresponding author:

Edgar P. Spalding

University of Wisconsin-Madison

430 Lincoln Dr.

Madison, WI 53706

ABSTRACT

Automated image acquisition, a custom analysis algorithm, and a distributed computing resource were used to add time as a third dimension to a quantitative trait locus (QTL) map for plant root gravitropism, a model growth response to an environmental cue. Digital images of *Arabidopsis thaliana* seedling roots from two independently reared sets of 162 recombinant inbred lines (RILs) and one set of 92 near isogenic lines (NILs) derived from a Cape Verde Islands (Cvi) x Landsberg *erecta* (L*er*) cross were collected automatically every two minutes for eight hours following induction of gravitropism by 90 degree reorientation of the sample. High Throughput Computing (HTC) was used to measure root tip angle in each of the 1.1 million images acquired and to perform statistical regression of tip angle against the genotype at each of the 234 (RIL) or 102 (NIL) DNA markers independently at each time point using a standard stepwise procedure. Time-dependent QTL were detected on chromosomes 1, 3, and 4 by this mapping method and by an approach developed to treat the phenotype time course as a function-valued trait. The QTL on chromosome 4 was earliest, appearing at 0.5 h and remaining significant for 5 h, while the QTL on chromosome 1 appeared at 3 h and thereafter remained significant. The Cvi allele generally had a negative effect of 2.6–4.0%. Heritability due to the QTL approached 25%. This study shows how computer vision and statistical genetic analysis by HTC can characterize the developmental timing of genetic architectures.

INTRODUCTION

Methodologies for characterizing genomes and genotypes are more advanced in terms of their degree of automation and throughput than their phenotype counterparts. This imbalance hinders progress in mapping genotype to phenotype in model systems such as the *Arabidopsis thaliana* plant, the subject of the present study. Computer vision methodologies can help address this imbalance by enabling objective, accurate, and potentially automated measurements of size and shape captured in digital images (SPALDING 2009). The Arabidopsis seedling root is well suited for this approach because it grows and responds well to stimuli in experimental scenarios that permit acquisition of high-quality images from which shape and size descriptors such as midlines, local curvature, and angles can be computationally extracted (CHAVARRIA-KRAUSER *et al.* 2008; FRENCH *et al.* 2009; MILLER *et al.* 2007). For example, Durham Brooks *et al.* (2010) used automated analysis of high-resolution images to measure the influence of seed size, seedling age, and media composition on the time course of root gravitropism, a dynamic growth response to a change in orientation with respect to the gravity vector (BLANCAFLOR and MASSON 2003; MORITA 2010). Miller *et al.* (2010) used the same approach plus extended computational analysis to quantify a transient gravitropism phenotype and its precise time of onset in young roots of *glr3.3* glutamate receptor mutants. Other researchers have taken advantage of automated image acquisition and analysis to measure shoot growth features over time (ZHANG *et al.* 2012; TISNÉ *et al.* 2013), or to acquire rotational series of branching root systems growing in a transparent gel to capture a third spatial dimension instead of time (CLARK *et al.* 2011; IYER-PASCUZZI *et al.* 2010).

The degree of automation made possible by computational image analysis increases the feasibility of systematically acquiring high-resolution phenotype data from genetically-structured populations for the purpose of large-scale statistical genetic studies. The benefits of accuracy and objectivity attending this approach increase the quality of the resulting phenotype data. Accordingly, automated or semi-automated image analysis is beginning to impact the mapping of quantitative trait loci in plants (HERRIDGE *et al.* 2011; MOORE *et al.* 2013; TISNÉ *et al.* 2013; TOPP *et al.* 2013). The project presented here combines an ability to capture the time course of gravitropism in high resolution with the throughput needed to make mapping populations feasible subjects. Gravitropism was selected for study because it is a model that embodies perception and transduction of an external cue, hormonal control of growth, and adaptation to a new environmental state (BOONSIRICHAI *et al.* 2002). These are fundamental physiological and developmental processes, now made accessible to statistical genetic analysis by novel application of automated computer vision and computation methodologies.

MATERIALS AND METHODS

Germplasm

The Arabidopsis seed stocks used here were described in a recent machine-vision study of seed size and shape QTL (MOORE *et al.* 2013). As in the previous report, RIL1 refers to a set of 162 F_{10} RILs derived from a cross between the Landsberg *erecta* (L*er*) and Cape Verde Islands (Cvi) accessions of *Arabidopsis thaliana* (ALONSO-BLANCO *et al.* 1998) donated to us by Professor Patrick Masson, University of Wisconsin, and to the phenotype data set obtained with those seeds. RIL2 refers to progeny of RIL1 produced in a random-designed and carefully controlled

seed-bulking exercise in order to duplicate the population in a different maternal environment as described in Moore *et al.* (2013). NIL refers to 92 near isogenic lines created by introgression of various short regions of Cvi DNA into the L*er* background (KEURENTJES *et al.* 2007) and raised in the same manner as the RIL2 seeds, again as described in (MOORE *et al.* 2013). The plant populations or the phenotype data sets derived from them are referred to as RIL1, RIL2, or NIL.

Plant culture

In a row on a Petri plate containing 1 mM KCl, 1 mM $CaCl_2$, 5 mM MES, pH 5.7 with BTP and gelled with 1% agar, three seeds were sown approximately 0.5 cm apart. Plates with seeds were placed at 4° for 2-4 d of stratification before being cultured vertically in a 22° growth chamber under constant white light for 3 d, at which point the primary root was 2-8 mm in length.

Root imaging

An imaging system consisting of 11 CCD cameras (Marlin F146B; Allied Vision Technologies) equipped with a macro-zoom lens (model NT59-157; Edmund Optics) was used to capture digital images of roots on as many Petri plates, each backlit with 880 nm radiation in a room maintained at approximately 22°, essentially as previously described (DURHAM BROOKS *et al.* 2010; MILLER *et al.* 2010). Instructions on how to create a similar apparatus are posted at http://phytomorph.wisc.edu/hardware/fixed-cameras.php. To initiate the experiment, each Petri dish was held in a custom plate holder in front of a camera and rotated by 90°. Framing, focusing, and starting image acquisition, in that order, were performed within 20 s of sample rotation. Images were acquired automatically every 2 min for 8 h at a resolution of 100

pixels/mm, resulting in a 241-frame movie of the root gravitropic response. The resulting images showed the dark roots of the three seedlings on a plate, against a light background. A video of an example experiment is provided as Figure S1. Because each response recording lasted 8 h, two successive trials per camera were acquired within a given day to increase the throughput of phenotype acquisition. To ensure that the roots being compared were at similar growth stages, only those with initial lengths between 3 and 8 mm were used for QTL analysis. RIL1, RIL2, and NIL were phenotyped consecutively. Within each population, genotypes were selected for measurement according to a randomized order.

Tip angle measurement

The camera firmware that controlled acquisition also automatically saved the images to a local disk array. A file synchronization protocol nightly copied the images to disks housed in the Department of Computer Sciences from where they were automatically read and processed for tip angle measurements on the University of Wisconsin Center for High Throughput Computing (CHTC) grid, which is managed and scheduled by the HTCondor software tool (THAIN *et al.* 2005). The custom algorithm that measured tip angle at each time point in a time series of frames was implemented in Matlab. The algorithm first binarized the grayscale images to isolate the roots from the background (Figure 1A-B). Erosion and pruning produced a continuous, branchless skeleton of the root. The set of pixels forming the boundary of the root apex was located within an 81 x 81 pixel patch of the binarized root image centered at the terminus of the skeleton. The local curvature at each position of the boundary was determined by continuous wavelet transformation of a parameterized trace of the boundary pixels. The point of maximum curvature was labeled the root tip (Figure 1C). The largest circle fit within a 60 x 60

pixel patch centered at the root tip was used to mask (set to zero) a portion of the root apex large enough that the tip angle could be equated with the first principal component determined by principal components analysis (Figure 1D). The angle of each root tip in each frame was thus determined, stored in comma-separated value files, and automatically copied to the local server one day after the experiment was performed. The large majority of trials were successfully processed automatically, resulting in 8 - 20 independent time course measurements for each genotype. The most common problems encountered were blurry images caused by excessive condensation forming on the lid of the Petri plate, and two roots colliding during the experiment. In all, 2123, 2325, and 2536 roots were successfully imaged and analyzed in the RIL1, RIL2, and NIL populations, respectively.

Stepwise QTL analysis using HTCondor

To add a high-resolution time axis to a QTL map, stepwise QTL analyses (MANICHAIKUL *et al.* 2009) were performed at each of the 241 time points using tip angle as the phenotype. The approach uses a penalized LOD score, with the LOD score for a multiple-QTL model (the $\log_{10}$ likelihood ratio comparing the given model to the null model with no QTL) penalized by the complexity of the model with separate penalties for main effects and pairwise interactions derived by a permutation test with a nominal significance level of 5% in the context of a two-dimensional, two-QTL genome scan. The approach uses a penalty on the main effects, T_m, and two different penalties on interactions, T_i^H (heavy) and T_i^L (light), and considers the structure of a QTL model to assign these penalties; see Manichaikul *et al.* (2009).

Performing 241 separate multiple-QTL analyses presented computational challenges especially because, to establish the penalties for each time point, 25,000 permutations of the

phenotype against the genotype were performed in the context of a two-dimensional, two-QTL genome scan, using Haley-Knott regression (HALEY and KNOTT 1992). The calculations were performed with R/qtl (BROMAN *et al.* 2003), an add-on package for the R statistical software (R CORE TEAM 2013). A directed acyclic graph (DAG) created in DAGman (COUVARES *et al.* 2007) structured the permutation tests into 5 per job and the HTCondor software tool was used to execute the 1.2 million separate compute jobs (5000 jobs per time point x 241 time points) in parallel on the CHTC or the Open Science Grid (OSG) distributed computing resources. After the results of each permutation test were automatically collated, the DAG triggered automatic model selection and QTL processing, again performed using R/qtl (BROMAN *et al.* 2003). Model fitting was performed in the manner described in Moore *et al.* (2013).

Heat maps of profile LOD scores were used to visualize the evidence for QTL in the context of a multiple-QTL model. This is analogous to a display technique used in Zeng *et al.* (2000): Each QTL was considered separately, and its position was allowed to vary, keeping all other QTL fixed at their estimated positions. For each possible position of the QTL under test, the LOD score comparing that multiple-QTL model to the model with the given QTL omitted was calculated.

Source code, raw and processed data

All of the raw images, custom image analysis computer code, quantified phenotype data, R/qtl input files, and reduced QTL results are available for download at http://phytomorph.wisc.edu/download/. These resources enable repetition of the analyses presented here but, probably more importantly, they enable computer vision scientists to

extract additional traits from the image sets and statisticians to develop new methods for mapping phenotypes quantified with high time resolution.

RESULTS

The gravitropism dataset

The biological material studied here was a set of 162 recombinant inbred lines of *Arabidopsis thaliana* derived from a cross between the Cape Verde Islands (Cvi) and Landsberg *erecta* (L*er*) ecotypes (ALONSO-BLANCO *et al.* 1998) and a set of 92 near isogenic lines containing introgressions of Cvi DNA in the L*er* background. The seed stocks used were the same previously studied by image analysis to determine the size distributions and genetic architectures of length, width, and area traits (MOORE *et al.* 2013). The phenotype data are time series of root tip angles extracted from sequences of grayscale digital images automatically acquired at 2 min intervals during gravitropism and batch-processed on a HTCondor-managed high-throughput computing grid (THAIN *et al.* 2005). Figure 2A shows five images of a representative root undergoing gravitropism for 8 h. Each of the 6,984 separate trials in this study consists of 241 frames that form a time series. The average tip angle responses of the two parental ecotypes and three example members of the mapping population are shown in Figure 2B. Regardless of genotype, all roots reoriented by approximately 90° within 8 h though the time courses differed significantly. The set of phenotype data obtained for this set of recombinant inbred lines is referred to as RIL1. A separate rearing and harvest of the recombinant inbred lines including the parental accessions and a repeat of the phenotype measurements from the next generation of seedlings produced

the RIL2 dataset. In general, RIL2 responses were slower than RIL1 responses despite the genetic identity of the plants (Figure 2C). The same measurements were performed on the near isogenic lines, the seeds of which were generated in the manner of the second RIL generation, to produce the NIL dataset (Figure 2D).

Analysis of variance

The variance attributed to additive genetic variation and environmental variance in RIL1, RIL2, and NIL are plotted in Figure 3. Additive genetic variance was initially low in all populations but increased as the gravitropic response progressed, peaking 4 to 6 h after gravistimulation (Figure 3A). Environmental contribution to total phenotypic variance exceeded the variance attributed to additive genetic variation, developing mostly during the first few hours of the response, especially in NIL (Figure 3B). These analyses of variance are based on the differences between the average responses of genetically distinct lines.

Time-resolved QTL map

The mean tip angle at each of the 241 time points for each genotype was used as the phenotype in 241 separate and independent multiple-QTL modeling analyses. The result was a highly time-resolved QTL map for each of the datasets (Figure 4A-C). Logarithm of the odds (LOD) score is shown on a color-intensity scale as a function of time (abscissa) and genome position (ordinate). In RIL1 (Figure 4A), a total of 8 loci, with a maximum of 6 loci at any given time, were detected during the 8 h response. Two loci in particular were present during a large portion of the response. The locus on chromosome 4 at 40.3 cM affected variation in the response from approximately 0.5 to 6 h past gravistimulation, and the other on chromosome 1

at 64 cM contributed from approximately 3 to 8 h. Some QTL were more limited in duration. For example, shortly after 3 h a QTL appeared on chromosome 3 at 44 cM and lost influence at 5 h. This particular QTL illustrates a problematic feature of treating the phenotype at one time point independent of the next, namely that a QTL can abruptly disappear or shift position due to data at adjacent time points being best fit by different QTL models.

RIL2 supported the selection of 3 loci as genetic contributors to variation in the gravitropism response (Figure 4B). Two loci contributed from 2 to 8 h after gravistimulation, with the third locus on chromosome 3 at 17 cM arising later to influence the response between 4 and 8 h after the start of the experiment. A locus on chromosome 4 at 40 cM was identified and displayed a similar time course in RIL1 and RIL2 (Figure 4B).

Figure 4C shows that 3 or possibly 4 loci were statistically significant in NIL during the early and middle parts of the response. The locus on chromosome 3 at 17 cM detected throughout the first 6 h of the response was also featured prominently in RIL2, and in the 'alternate' model centered around 4 h in RIL1. A second locus on chromosome 5 at 55 cM, influential from 3.5 to 5.5 h after gravistimulation, was also apparent over the same time period in RIL1. NIL was the only dataset to support a QTL on chromosome 2, located at 40 cM, which contributed to the variation in the response during the first hour.

A function-valued approach

The analysis described above adjusted for the search across QTL models but not for the search across time points. Another shortcoming is that the selected model occasionally abruptly changed between adjacent time points despite the phenotype following a smooth curve across

time. A simple but effective approach to integrating information across time points was devised to accommodate the high time resolution in the phenotype data generated by the automated image analysis. The method used a model comparison criterion analogous to that of Manichaikul *et al.* (2009). It was initiated by performing a genome scan, or single-QTL analysis, at each time point to generate a 2D matrix of LOD scores (genomic positions x time points). Figure 4D-F shows the results of these single-QTL analyses in the three populations. For RIL1, strong support was obtained for loci on chromosomes 1 and 4 acting at different times (Figure 4D), while for RIL2, loci affecting this response were predicted on chromosomes 3 and 4 (Figure 4E). The locus on chromosome 4 was predicted to be at approximately 40.3 cM in both RIL1 and RIL2. The NIL population showed some support for a locus on the proximal end of chromosome 4 and some support for loci on chromosomes 3 and 5 (Figure 4F). The next step in the method was to calculate the maximum LOD (MLOD) and the average LOD (SLOD) at each genomic position across time. Significance thresholds were obtained by a permutation test (CHURCHILL and DOERGE 1994), using the corresponding SLOD or MLOD statistic. SLOD exhibits higher power to detect QTL with effects across a large time interval, while MLOD exhibits higher power to detect QTL with large effects over a narrow interval. The MLOD or SLOD statistics were used to derive a penalized LOD criterion, which multiple-QTL analyses sought to maximize in a stepwise search, starting with the first QTL at the position with the highest MLOD or SLOD score. Additional QTL were added to the model if the significance threshold was met. After selecting the QTL model with the highest penalized LOD score, profile LOD scores were derived to evaluate the evidence and localization of each QTL, evaluating each time point in the 8 h response, individually: The position of each QTL was varied, one at a time, and at each

location for a given QTL we derived a LOD score comparing the multiple-QTL model to the model with the given QTL omitted. Figure 4G-I present these profile LOD scores, based on the model identified with the stepwise QTL analysis using the SLOD statistic. In RIL1, a 2-QTL model was identified using the SLOD statistic, with loci on chromosomes 1 and 4 (Figure 4G), and a 2-QTL model with evidence of epistasis was predicted using the MLOD statistic, with the same QTL on chromosome 4 and an additional one on the distal end of chromosome 3 (data not shown but presented for download). For RIL2, the SLOD statistic predicted a 3-QTL model, with loci on chromosomes 1, 3, and 4 (Figure 4H), while the MLOD statistic gave a 4 QTL model, with two loci on chromosome 1 and one locus on chromosomes 3 and 4 (data not shown but presented for download). In the NIL dataset, the SLOD statistic resulted in a 2-QTL model with both loci on chromosome 3, at 16.4 and 39.7 cM (Figure 4I). Analysis of this population using the MLOD statistic found both the two QTL on chromosome 3, as well as an additional locus on chromosome 2 at 81.3 cM (data not shown but presented for download). None of the loci identified by the above methods were found in a QTL study of Arabidopsis root skewing (the degree to which a seedling root axis deviates from the gravity vector when grown on a vertically maintained agar surface) that utilized the same germplasm (VAUGHN and MASSON 2011).

Heritability

Analysis of variance was performed to calculate the heritability (h^2) due to the identified QTL across time in each population, using both the stepwise QTL analysis results at individual time points as well as the results achieved via stepwise QTL analysis using the SLOD statistic. For

RIL1, h^2 ranged from a low of 10%, occurring at the initiation of the gravitropic stimulus, to a high of 27% around 6 h post-gravistimulation using trait data from each time point (Figure 5A). Analysis of RIL1 with the SLOD statistic showed a similar range in h^2, from 5% to a high of 33% (Figure 5B). The heritability calculated from the RIL2 population was somewhat lower than RIL1, ranging from 8% – 24% and 5% – 23% for analysis at each time point and via the SLOD statistic, respectively (Figure 5A-B). For the NIL population, the measured heritability was less than that of the RIL1 and RIL2 populations, with estimates of 3% – 15% from individual time point data (Figure 6A) and 12% – 22% from the SLOD statistic analysis (Figure 5B). The lower heritability of NIL can probably be attributed to the lesser amount of genetic variation present in these lines.

A four QTL model

Independent analysis of the three data sets identified some shared and some unique QTL. In an effort to consolidate the results in a single analysis, loci at chromosome 1 at 64 cM (1@64 cM), 3@17 cM, 4@40.3 cM, and 5@61 cM, identified as contributors by stepwise QTL analysis at each time point (Figure 4A-C), were used to create a four-QTL model that was evaluated in each of the three data sets. Figure 6 shows plots of the effect on the tip angle resulting from replacement of the L*er* allele with the Cvi allele at the indicated positions, i.e. the Cvi allele effect, as inferred from the four-QTL model fit separately to each of the data sets. Figure 6A shows that the Cvi allele at the 1@64 cM locus was estimated to have a negative time-dependent effect on the root tip angle in all three populations. The LOD scores for this position (Figure 4A-C) indicate strong support for the negative effect of this allele in RIL1 and RIL2,

peaking around 5 h. Locus 3@17 cM was also estimated to have a negative effect on root tip angle, this time with NIL providing strong support, peaking about 4 h into the response (Figure 6B). The other two loci, 4@40.3 cM and 5@61 cM, also showed a negative Cvi allele effect but with inconsistent (Figure 6C) or little (Figure 6D) time dependence.

DISCUSSION

Gravitropic signaling initiates in specific cells of the root cap (HASHIGUCHI *et al.* 2013), resulting in a redistribution of indole-3-acetic acid (IAA) flowing back shootward such that a growth-inhibiting level of this auxin accumulates on the lower side of the root (SPALDING 2013). Slowing of cell expansion on the lower side drives downward bending, which is first detectable within approximately 15 minutes (LEWIS *et al.* 2007; MILLER *et al.* 2007). The response time course in Arabidopsis seedling roots depends on factors such as seedling age, the size of the seed from which the seedling emerged, nutrient conditions (DURHAM BROOKS *et al.* 2010), and even on the previous generation's culture environment (ELWELL *et al.* 2011) but, generally speaking, the bending speed increases during the first two hours, peaks, then gradually slows as the tip angle approaches the new vertical. The present study captured the time course of this process across genetically structured populations of Arabidopsis seedlings so that the dynamics of its genetic architecture could be determined.

HTC advantages matched technical needs of temporal QTL study

Automated image acquisition and feature extraction were important contributors to the feasibility of this study but were not sufficient for success. Also necessary was the

establishment of an automated HTC workflow. HTCondor software enabled batch processing of raw images and statistical genetic modeling of the quantified feature by matching jobs and available resources in a distributed computing environment. Obtaining the results presented in Figure 4A-C alone entailed launching 21,700 jobs that consumed approximately 5000 central processor unit hours. Such work would probably not have been feasible in the absence of such a computing infrastructure. Data quality was also increased by adopting HTC because the degree of throughput and automation increased the feasibility of larger sample sizes.

Comparison of the two methods of QTL analysis

The first of the two methods of searching for significant QTL used in this study relied on automated stepwise QTL selection of a model based on tip angle data at a single time point, while the second method combined information from all time points to choose a multiple-QTL model. The first method is conceptually simpler, treating each time point independently and separately as a new trait to map. However, especially in the case of RIL1, the independent treatment of time points produced occasional discontinuities along the time axis as the selected QTL model jumped between alternatives. The alternate QTL positions, for example around 4 h on chromosomes 3 and 5 (Figure 4A), could reflect genes that contribute to the response but without sufficient strength to force their selection in each independent modeling exercise. In the second method, the function-valued approach, one model is chosen to represent the entire response, leading to improved power to detect QTL and better separation of linked QTL. The second approach is also more conservative as it seeks to control for the search across time points as well as the search across QTL models. When applied to the two RIL data sets, both methods detected the QTLs on chromosomes 1, 3, and 4 (Figure 4). Replacement of L*er* with

Cvi alleles at these loci had a negative effect on the root tip angle (Figure 6) as did the Cvi allele at the weakly supported QTL on chromosome 5, which was only detected with the first method and only in RIL1 (Figure 4A).

Similarities and differences between RIL1 and RIL2

Independent rearings of the same set of recombinant inbred lines created the seed stocks used to produce the RIL1 and RIL2 data sets. Therefore, differences between the two data sets cannot be due to genetic differences. Instead, environmental differences during the production of the seed lots are the most probable causes of the differences between RIL1 and RIL2 root responses. Variation in seed size and shape within the two stocks were previously measured by image analysis and subjected to statistical genetic analysis. The seeds used to generate RIL2 were found to be significantly larger than those which generated RIL1. Such maternal environment effects on Arabidopsis seed size are known to affect next-generation growth of seedlings including root gravitropism (DURHAM BROOKS *et al.* 2010; ELWELL *et al.* 2011). Therefore, the QTL on chromosome 3 that is apparent in RIL2 (Figure 4H) but only weakly and briefly or not at all in RIL1 (Figure 4A,G) may be a locus that affects gravitropism only in certain seed-size contexts. However, none of the seed size or shape QTL previously identified (MOORE *et al.* 2013) using the same seed stocks overlapped with the root tip angle QTL mapped here.

Candidate genes

The 1.5-LOD support intervals calculated for each QTL permitted consideration of whether or not any corresponded to genes known to encode components of the gravitropism mechanism, such as PIN2 auxin transporters (CHEN *et al.* 1998; MÜLLER *et al.* 1998), determinants of starch

content in the sedimenting statoliths (CASPAR and PICKARD 1989; KISS *et al.* 1989), and plastid outer envelope proteins (STANGA *et al.* 2009), among others (BALDWIN *et al.* 2013; MORITA 2010).

The confidence interval of QTL 1@61 spans a region approximately from gene AT1G21945 to gene AT1G22420. One gene within this region is *AUXIN UP-REGULATED F-BOX PROTEIN2* (*AUF2*), which functions with the related AUF1 protein to regulate root growth responses to the hormones cytokinin and auxin (ZHENG *et al.* 2011). Mutating *AUF1* and *AUF2* impairs both of the major auxin transport streams in the root and makes root growth more sensitive to auxin transport inhibitors (ZHENG *et al.* 2011). Natural variation at this genetic locus could potentially affect the gravitropism time course.

The confidence interval for QTL 3@15 spans from AT3G23980 to AT3G24400. No gene within this region is known to function in gravitropism, hormonal control of root growth, or other processes that could support a hypothesis about how natural variation within it could affect the phenotype measured here.

The confidence interval for QTL4@40.3 extends from approximately AT4G15080 to AT4G15396 and includes *ABCG30*, a gene predicted to encode a membrane protein belonging to the ATP-binding cassette (ABC) superfamily of transporters (VERRIER *et al.* 2008). Members of the B subfamily of ABC transporters have been linked to IAA transport (NOH *et al.* 2001) and subsequently shown to affect gravitropism (LEWIS *et al.* 2007; NOH *et al.* 2003), but only relatively recently have members of the G subfamily been connected to auxin transport. Specifically, ABCG36 and ABCG37 have a subcellular localization and an apparent transport activity that leads to efflux from roots of indole-3-butyric acid (RŮŽIČKA *et al.* 2010; STRADER and BARTEL 2009), a compound that has intrinsic auxinic activity and is believed to be a precursor of

IAA (STRADER and BARTEL 2011). *ABCG30*, formerly known as *PLEIOTROPIC DRUG RESISTANCE 2* (*PDR2*), resides near *ABCG36* and *ABCG37* in the *ABCG* phylogenetic tree, and it isexpressed specifically in roots (VAN DEN BRULE and SMART 2002). Expression in roots and a close sequence relationship with membrane proteins now believed to function in auxin homeostasis makes *ABCG30* an interesting candidate for the cause of phenotypic variation associated with QTL 4@40.3.

The genes within the confidence interval surrounding QTL 5@61, covering loci AT5G16770 to AT5G17330, include *MERISTEM-DEFECTIVE* (*MDF*), which encodes a protein in the SART-1 family. MDF is expressed throughout the plant and mediates root, shoot, and flower development (CASSON *et al.* 2009). Messenger RNA levels of *PIN2* and *PIN4* are lower in *mdf* mutants, affecting the ability of the root to achieve a normal auxin maximum in the meristem, which may be how MDF functions to maintain meristematic activity (CASSON *et al.* 2009). The influence of MDF on meristem activity does not make it a strong candidate for the causative gene at this QTL because cell division is not expected to play an important part during an 8 h gravitropic response. However, the effect of MDF on PIN2 expression makes it a very interesting candidate of causation. PIN2 mediates the shootward flow of auxin that is an important component of the mechanism that creates differential growth between the upper and lower sides of the root during gravitropism (SPALDING 2013). Natural variation in MDF may cause natural variation in PIN2 expression, and thereby affect the strength of the shootward auxin transport stream and the tropic responses that depend on it.

Adding a highly-resolved time dimension to an analysis of quantitative trait loci posed some significant technical challenges met here by an emphasis on automated data acquisition

and HTC for data analysis and modeling, but it also generates significant new insights and perspectives on the genetic architecture of the process that gives rise to a trait, rather than a trait itself. With image analysis becoming more common in studies of plant biology, in scenarios ranging from confocal microscopes to whole plants (SPALDING and MILLER 2013), time may be an axis on many more QTL profiles in the future.

REFERENCES

ALONSO-BLANCO, C., A. J. M. PEETERS, M. KOORNNEEF, C. LISTER, C. DEAN *et al.*, 1998 Development of an AFLP based linkage map of L*er*, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a L*er*/Cvi recombinant inbred line population. Plant J. **14:** 259-271.

BALDWIN, K. L., A. K. STROHM and P. H. MASSON, 2013 Gravity sensing and signal transduction in vascular plant primary roots. Am. J. Bot. **100:** 126-142.

BLANCAFLOR, E. B., and P. H. MASSON, 2003 Plant gravitropism. Unraveling the ups and downs of a complex process. Plant Physiol. **133:** 1677-1690.

BOONSIRICHAI, K., C. GUAN, R. CHEN and P. H. MASSON, 2002 Root gravitropism: an experimental tool to investigate basic cellular and molecular processes underlying mechanosensing and signal transmission in plants. Annu. Rev. Plant Biol. **53:** 421-447.

BROMAN, K. W., H. WU, S. SEN and G. A. CHURCHILL, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics **19:** 889-890.

CASPAR, T., and B. G. PICKARD, 1989 Gravitropism in a starchless mutant of *Arabidopsis*. Implications for the starch-statolith theory of gravity sensing. Planta **177:** 185-197.

CASSON, S. A., J. F. TOPPING and K. LINDSEY, 2009 MERISTEM-DEFECTIVE, an RS domain protein, is required for the correct meristem patterning and function in Arabidopsis. Plant J. **57:** 857-869.

CHAVARRIA-KRAUSER, A., K. A. NAGEL, K. PALME, U. SCHURR, A. WALTER *et al.*, 2008 Spatio-temporal quantification of differential growth processes in root growth zones based on a novel combination of image sequence processing and refined concepts describing curvature production. New Phytol. **177:** 811-821.

CHEN, R., P. HILSON, J. SEDBROOK, E. ROSEN, T. CASPAR *et al.*, 1998 The *Arabidopsis thaliana AGRAVITROPIC 1* gene encodes a component of the polar-auxin-transport efflux carrier. Proc. Natl. Acad. Sci. USA **95:** 15112-15117.

CHURCHILL, G. A., and R. W. DOERGE, 1994 Empirical threshold values for quantitative trait mapping. Genetics **138:** 963-971.

CLARK, R. T., R. B. MACCURDY, J. K. JUNG, J. E. SHAFF, S. R. MCCOUCH *et al.*, 2011 Three-dimensional root phenotyping with a novel imaging and software platform. Plant Physiol. **156:** 455-465.

COUVARES, P., T. KOSAR, A. ROY, J. WEBER and K. WENGER, 2007 *Workflow in Condor*. Springer Press.

DURHAM BROOKS, T. L., N. D. MILLER and E. P. SPALDING, 2010 Plasticity of Arabidopsis root gravitropism throughout a multidimensional condition space quantified by automated image analysis. Plant Physiol. **152:** 206-216.

ELWELL, A. L., D. S. GRONWALL, N. D. MILLER, E. P. SPALDING and T. L. BROOKS, 2011 Separating parental environment from seed size effects on next generation growth and development in Arabidopsis. Plant Cell Environ. **34:** 291-301.

FRENCH, A., S. UBEDA-TOMAS, T. J. HOLMAN, M. J. BENNETT and T. PRIDMORE, 2009 High-throughput quantification of root growth using a novel image-analysis tool. Plant Physiol. **150:** 1784-1795.

HALEY, C. S., and S. A. KNOTT, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity **69:** 315-324.

HASHIGUCHI, Y., M. TASAKA and M. T. MORITA, 2013 Mechanism of higher plant gravity sensing. Am. J. Bot. **100:** 91-100.

HERRIDGE, R. P., R. C. DAY, S. BALDWIN and R. C. MACKNIGHT, 2011 Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. Plant Methods **7:** 3.

IYER-PASCUZZI, A. S., O. SYMONOVA, Y. MILEYKO, Y. HAO, H. BELCHER *et al.*, 2010 Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. Plant Physiol. **152:** 1148-1157.

KEURENTJES, J. J., L. BENTSINK, C. ALONSO-BLANCO, C. J. HANHART, H. BLANKESTIJN-DE VRIES *et al.*, 2007 Development of a near-isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a recombinant inbred line population. Genetics **175:** 891-905.

KISS, J. Z., R. HERTEL and F. D. SACK, 1989 Amyloplasts are necessary for full gravitropic sensitivity in roots of *Arabidopsis thaliana*. Planta **177:** 198-206.

LEWIS, D. R., N. D. MILLER, B. L. SPLITT, G. WU and E. P. SPALDING, 2007 Separating the roles of acropetal and basipetal auxin transport on gravitropism with mutations in two Arabidopsis *Multidrug Resistance-Like* ABC transporter genes. Plant Cell **19:** 1838-1850.

MANICHAIKUL, A., J. Y. MOON, S. SEN, B. S. YANDELL and K. W. BROMAN, 2009 A model selection approach for the identification of quantitative trait loci in experimental crosses, allowing epistasis. Genetics **181:** 1077-1086.

MILLER, N. D., T. L. DURHAM BROOKS, A. H. ASSADI and E. P. SPALDING, 2010 Detection of a gravitropism phenotype in *glutamate receptor-like 3.3* mutants of *Arabidopsis thaliana* using machine vision and computation. Genetics **186:** 585-593.

MILLER, N. D., B. M. PARKS and E. P. SPALDING, 2007 Computer-vision analysis of seedling responses to light and gravity. Plant J. **52:** 374-381.

MOORE, C. R., D. S. GRONWALL, N. D. MILLER and E. P. SPALDING, 2013 Mapping quantitative trait loci affecting *Arabidopsis thaliana* seed morphology features extracted computationally from images. G3: Genes, Genomes, Genetics **3:** 109-118.

MORITA, M. T., 2010 Directional gravity sensing in gravitropism. Annu. Rev. Plant Biol. **61:** 705-720.

MÜLLER, A., C. GUAN, L. GALWEILER, P. TANZLER, P. HUIJSER *et al.*, 1998 *AtPIN2* defines a locus of Arabidopsis for root gravitropism control. EMBO J. **17:** 6903-6911.

NOH, B., A. BANDYOPADHYAY, W. A. PEER, E. P. SPALDING and A. S. MURPHY, 2003 Enhanced gravi- and phototropism in plant *mdr* mutants mislocalizing the auxin efflux protein PIN1. Nature **423:** 999-1002.

NOH, B., A. S. MURPHY and E. P. SPALDING, 2001 Multidrug resistance-like genes of Arabidopsis required for auxin transport and auxin-mediated development. Plant Cell **13:** 2441-2454.

R CORE TEAM, 2013 R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, pp., Vienna, Austria.

RŮŽIČKA, K., L. C. STRADER, A. BAILLY, H. YANG, J. BLAKESLEE *et al.*, 2010 Arabidopsis *PIS1* encodes the ABCG37 transporter of auxinic compounds including the auxin precursor indole-3-butyric acid. Proc. Natl. Acad. Sci. USA.

SPALDING, E. P., 2009 Computer Vision as a Tool to Study Plant Development, pp. 317-326 in *Plant Systems Biology*, edited by D. A. BELOSTOTSKY. Totowa: Humana Press Inc.

SPALDING, E. P., 2013 Diverting the downhill flow of auxin to steer growth during tropisms. Am. J. Bot. **100:** 203-214.

SPALDING, E. P., and N. D. MILLER, 2013 Image analysis is driving a renaissance in growth measurement. Curr. Op. Plant Biol. **16:** 100-104.

STANGA, J. P., K. BOONSIRICHAI, J. C. SEDBROOK, M. S. OTEGUI and P. H. MASSON, 2009 A role for the TOC complex in Arabidopsis root gravitropism. Plant Physiol. **149:** 1896-1905.

STRADER, L. C., and B. BARTEL, 2009 The Arabidopsis PLEIOTROPIC DRUG RESISTANCE8/ABCG36 ATP Binding Cassette transporter modulates sensitivity to the auxin precursor indole-3-butyric acid. Plant Cell **21:** 1992-2007.

STRADER, L. C., and B. BARTEL, 2011 Transport and metabolism of the endogenous auxin precursor indole-3-butyric acid. Mol. Plant **4:** 477-486.

THAIN, D., T. TANNENBAUM and M. LIVNY, 2005 Distributed computing in practice: the Condor experience. Concurrency and Computation: Practice & Experience **17:** 323-356.

TISNÉ, S., Y. SERRAND, L. BACH, E. GILBAULT, R. BEN AMEUR *et al.*, 2013 Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. Plant J. **74:** 534-544.

TOPP, C. N., A. S. IYER-PASCUZZI, J. T. ANDERSON, C. R. LEE, P. R. ZUREK *et al.*, 2013 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. Proc. Natl. Acad. Sci. USA **110:** E1695-1704.

VAN DEN BRULE, S., and C. C. SMART, 2002 The plant PDR family of ABC transporters. Planta **216:** 95-106.

VAUGHN, L. M., and P. H. MASSON, 2011 A QTL study for regions contributing to *Arabidopsis thaliana* root skewing on tilted surfaces. G3: Genes, Genomes, Genetics **1:** 105-115.

VERRIER, P. J., D. BIRD, B. BURLA, E. DASSA, C. FORESTIER *et al.*, 2008 Plant ABC proteins – a unified nomenclature and updated inventory. Trends Plant Sci. **13:** 151-159.

ZENG, Z. B., J. J. LIU, L. F. STAM, C. H. KAO, J. M. MERCER *et al.*, 2000 Genetic architecture of a morphological shape difference between two Drosophila species. Genetics **154:** 299-310.

ZHANG, X., R. J. HAUSE, and J. O. BOREVITZ, 2012 Natural genetic variation for growth and development revealed by high-throughput phenotyping in *Arabidopsis thaliana*. G3: Genes, Genomes, Genetics. **2:** 29-34.

ZHENG, X. H., N. D. MILLER, D. R. LEWIS, M. J. CHRISTIANS, K. H. LEE *et al.*, 2011 *AUXIN UP-REGULATED F-BOX PROTEIN1* regulates the cross talk between auxin transport and cytokinin signaling during plant root growth. Plant Physiol. **156:** 1878-1893.

ACKNOWLEDGEMENTS

This work was supported by grant IOS-1031416 from the National Science Foundation Plant Genome Research Program to E.P.S. and by National Institutes of Health grant GM074244 to K.W.B.

FIGURE LEGENDS

Figure 1. Root tip angle determined by image analysis. **A.** Grayscale image of a representative root undergoing a growth response to a change in the gravity vector, or gravitropism. Scale bar = 1mm. **B.** Binarized image of the responding root. **C.** Curvature values are calculated at each boundary point along the root apex. The point of highest curvature is taken to be the root tip. Color legend indicates curvature values in mm^{-1}. **D.** A patch centered at the root tip is subjected to principal components analysis. The first eigenvector (black line) determines the tip angle relative to the horizon.

Figure 2. Quantifying root tip angle during gravitropism generates a time-course phenotype. **A.** A representative series of images of a root responding to gravity at 2 h intervals. At the beginning of the experiment, seedlings were rotated 90° such that the root tip was approximately horizontal to the gravity vector. By 8 h after rotation, the root had grown to reorient its tip parallel to the direction of gravity. Scale bar = 0.5 mm. **B-D.** Tip angle development during gravitropism in the RIL1, RIL2, and NIL datasets. The response of the Cvi parental line is shown by a black line, L*er* by an orange line, and the other lines indicate the responses of representatives of the population of inbred lines (labeled A, B, and C). Vertical bars indicate the standard error of the mean. For RIL1, n = 27, 28, 10, 10, and 18 for Cvi, L*er*, RIL1-A, RIL1-B, and RIL1-C. For RIL2, n= 14, 18, 18, 11, and 12 for Cvi, L*er*, RIL2-A, RIL2-B, and RIL2-C. For NIL, n = 32, 33, and 20 for NIL-A, NIL-B, and NIL-C.

Figure 3. Genetic and environmental contributions to variance. **A.** The proportion of the phenotypic variance attributed to the additive genetic variation. **B.** The environmental

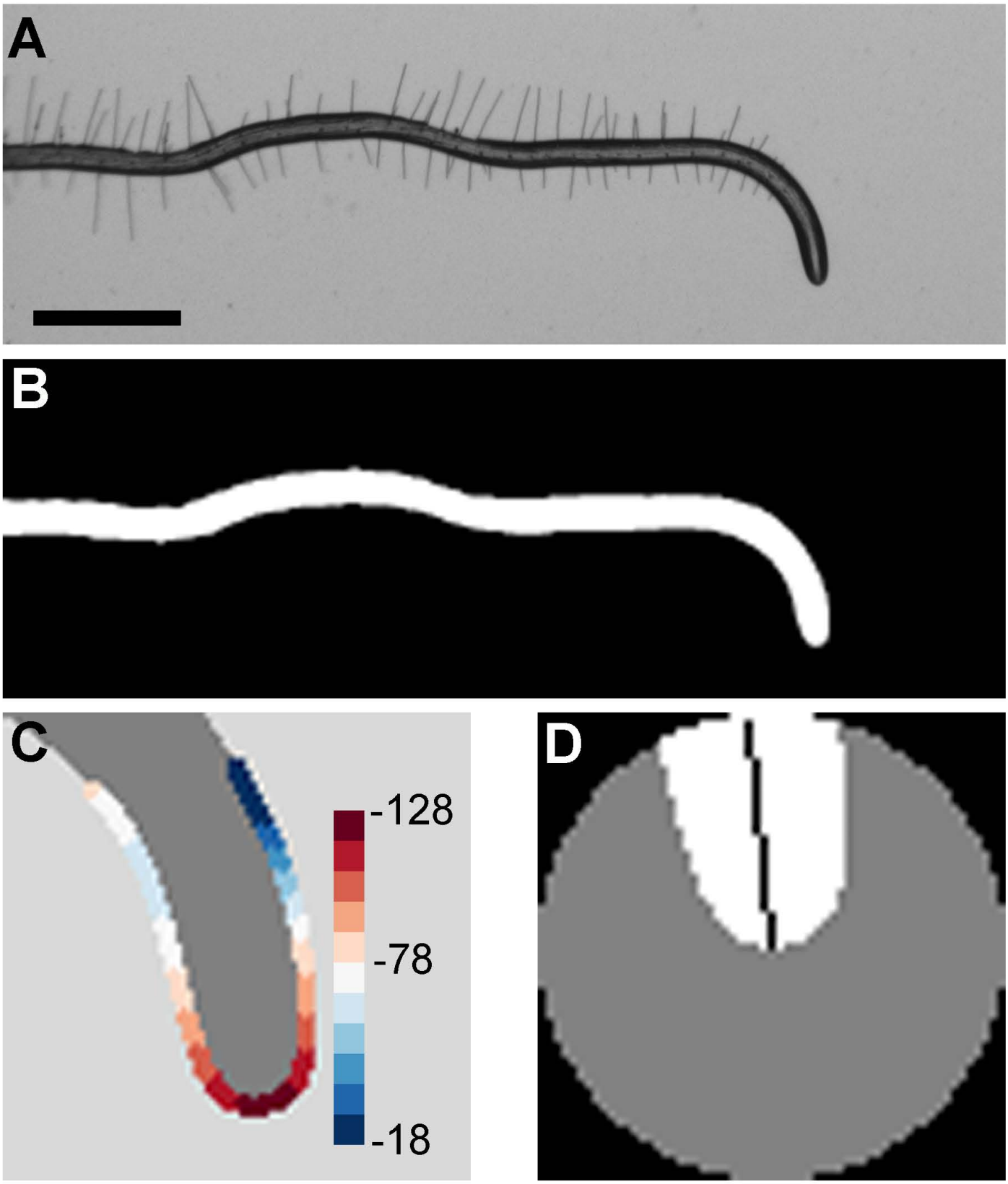
contribution to the phenotypic variance. Orange, blue, and green lines indicate the RIL1, RIL2, and NIL populations, respectively.

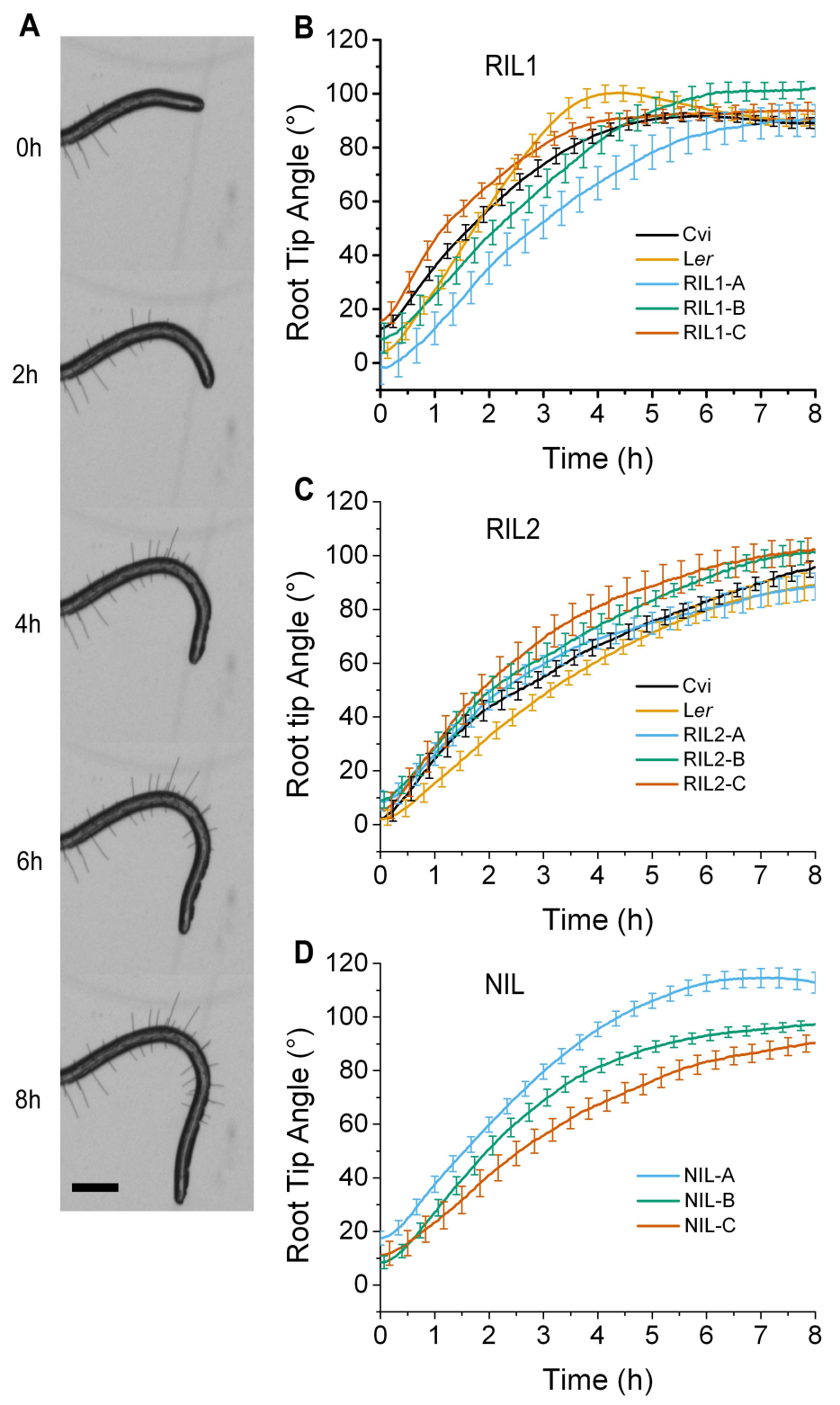
Figure 4. Time course of QTL development. Magnitude of LOD score is displayed as color intensity as a function of time. **A-C.** Profile LOD scores for models selected by stepwise QTL analyses, considering each time point individually. For RIL1, (T_m, T_i^H, T_i^L) = (2.58, 3.41, 1.53), for RIL2, (T_m, T_i^H, T_i^L) = (2.56, 3.44, 1.63), and for NIL, (T_m, T_i^H, T_i^L) = (2.75, 2.40, 0.67). **D-F.** Single-QTL analysis results from RIL1, RIL2, and NIL populations. Threshold for significance of single QTL is 1.91 for RIL1, 1.99 for RIL2, and 1.91 for NIL. **G-I.** Profile LOD scores for the model selected by stepwise QTL analyses using the SLOD criterion. For RIL1, (T_m, T_i^H, T_i^L) = (1.91, 2.41, 1.10), for RIL2, (T_m, T_i^H, T_i^L) = (1.99, 2.62, 1.51), and for NIL, (T_m, T_i^H, T_i^L) = (1.91, 2.40, 1.66). The position of the loci on the y-axis is shown as the cumulative position of the five Arabidopsis chromosomes. Horizontal lines indicate chromosome breaks. The x-axis shows time in hours since onset of gravistimulation. All coordinates whose LOD scores were not significant are shown in white. Darker blue colors indicate higher support for the loci.

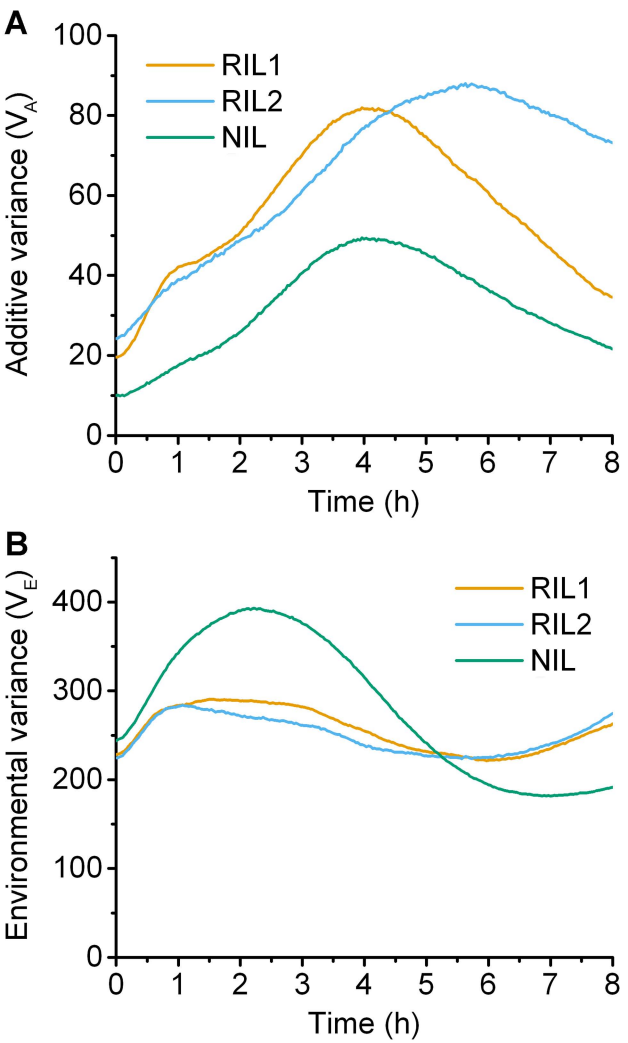
Figure 5. A. Heritability calculated via results from stepwise QTL analysis at individual time points. **B.** Heritability based on the chosen SLOD model from each population. Orange, blue, and green lines indicate the RIL1, RIL2, and NIL populations, respectively.

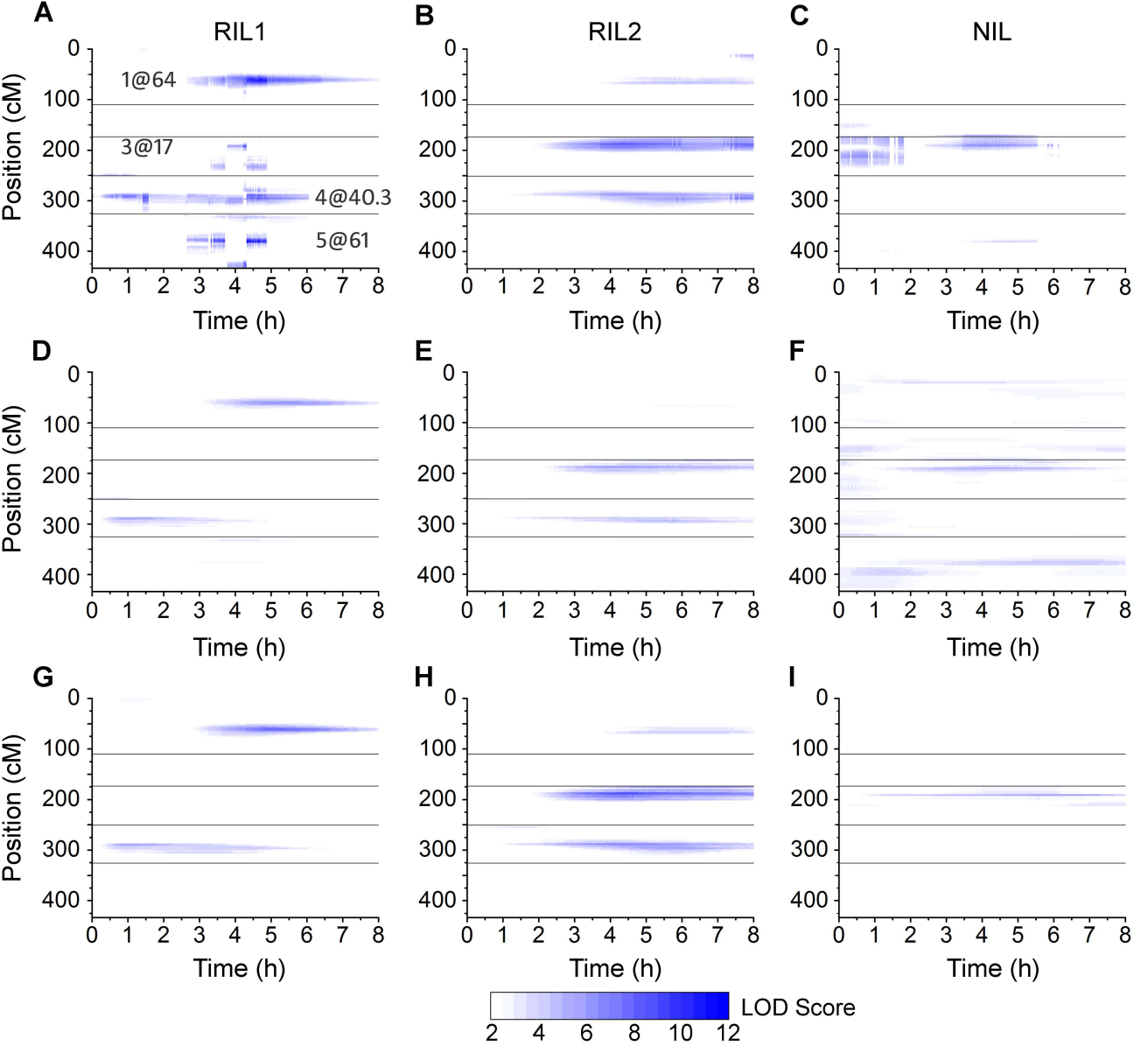
Figure 6. Time-dependent allele effects on root tip angle during gravitropism. The four QTL found by stepwise QTL analyses at each time point that were common to at least two populations were added to a model and the contribution of each locus was estimated. Positive values indicate that substitution of a Cvi allele at the indicated locus increases the tip angle

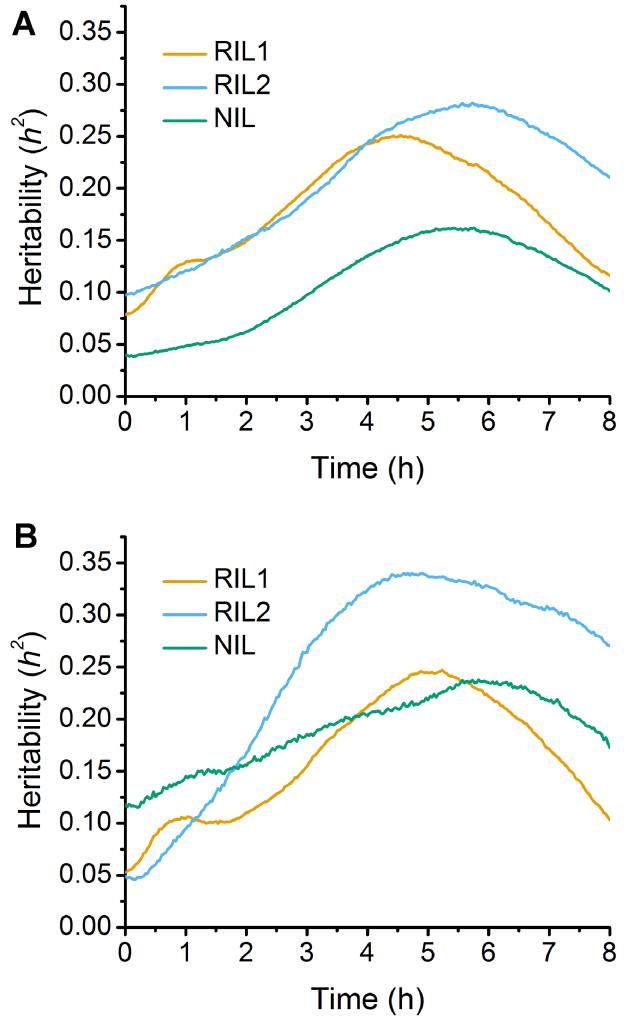
trait, while a negative value corresponds to the L*er* allele increasing the trait value. The 95 percent confidence interval for the degree of the effect is displayed as error bars. **A.** Chromosome 1 at 64 cM. **B.** Chromosome 3 at 17 cM. **C.** Chromosome 4 at 40.3 cM. **D.** Chromosome 5 at 61 cM.

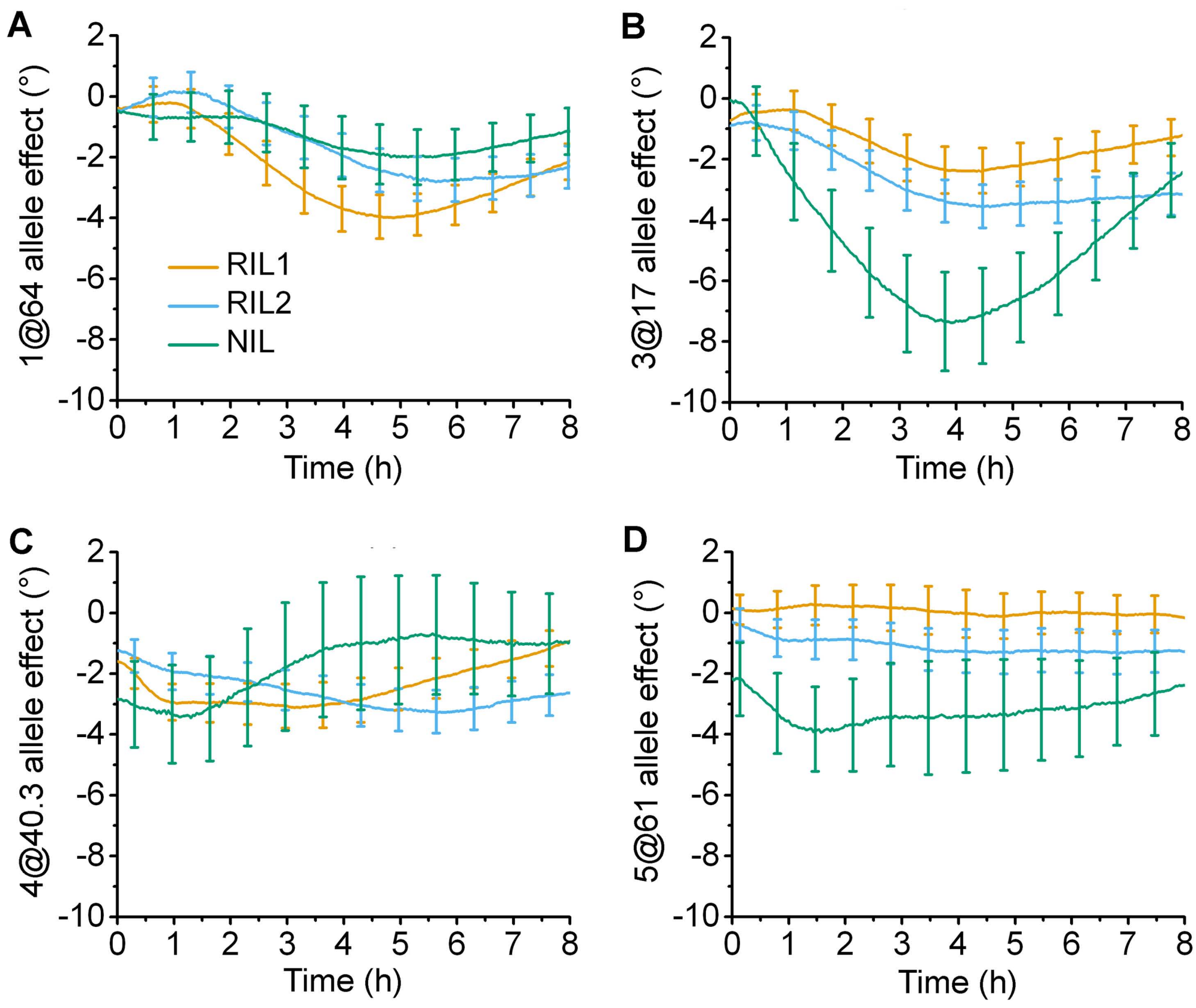

A
B
C
-128
-78
-18
D


A
0h
2h
4h
6h
8h
B
RIL1
Root Tip Angle (°)
Time (h)
Cvi
Ler
RIL1-A
RIL1-B
RIL1-C
C
RIL2
Root tip Angle (°)
Time (h)
Cvi
Ler
RIL2-A
RIL2-B
RIL2-C
D
NIL
Root Tip Angle (°)
Time (h)
NIL-A
NIL-B
NIL-C


A
Additive variance (V_A)
RIL1
RIL2
NIL
Time (h)
B
Environmental variance (V_E)
RIL1
RIL2
NIL
Time (h)


A
RIL1
B
RIL2
C
NIL
D
E
F
G
H
I
Position (cM)
Time (h)
1@64
3@17
4@40.3
5@61
0
100
200
300
400
0 1 2 3 4 5 6 7 8
LOD Score
2 4 6 8 10 12


A
RIL1
RIL2
NIL
Heritability (h^2)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3
4
5
6
7
8
Time (h)
B
RIL1
RIL2
NIL
Heritability (h^2)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3
4
5
6
7
8
Time (h)


A
1@64 allele effect (°)
RIL1
RIL2
NIL
Time (h)
B
3@17 allele effect (°)
Time (h)
C
4@40.3 allele effect (°)
Time (h)
D
5@61 allele effect (°)
Time (h)