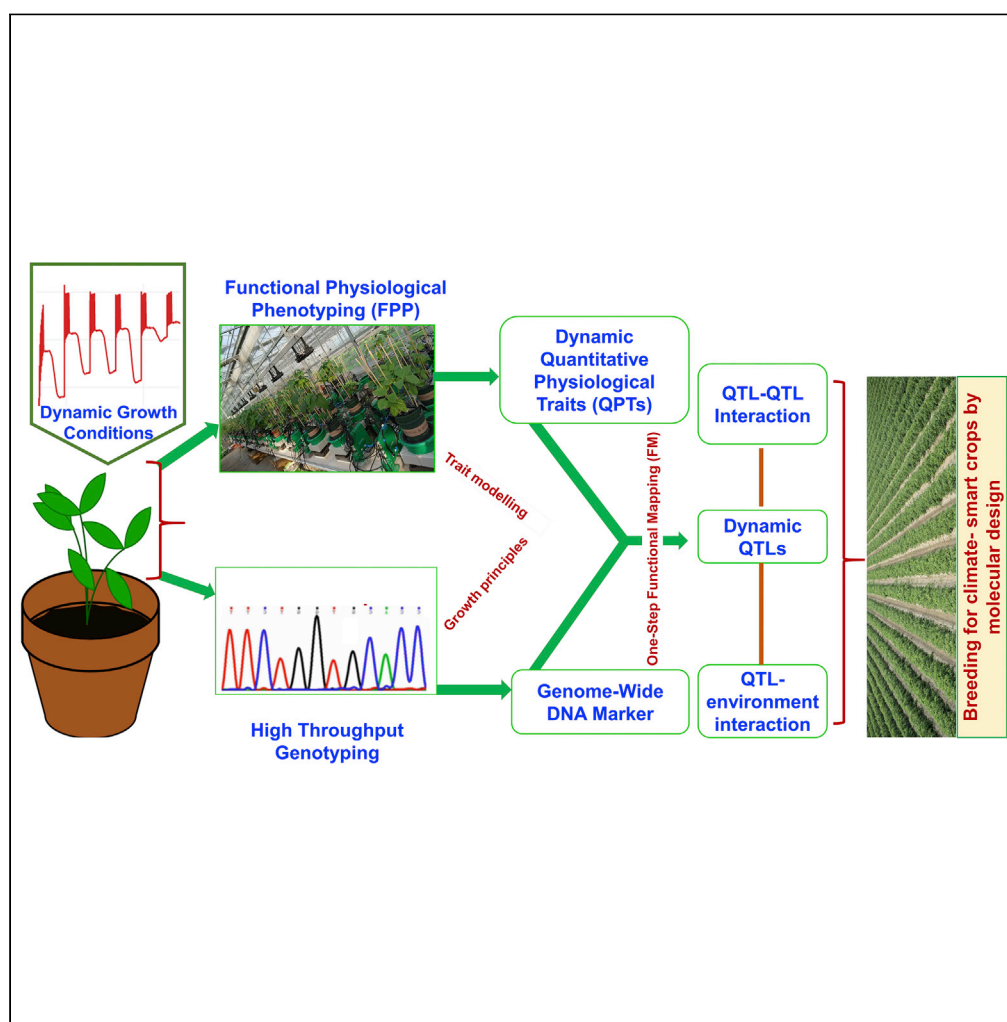**Article**

# Functional physiological phenotyping with functional mapping: A general framework to bridge the phenotype-genotype gap in plant physiology

Arun K. Pandey, Libo Jiang, Menachem Moshelion, ..., Qin Lin, Rongling Wu, Pei Xu

menachem.moshelion@mail.huji.ac.il (M.M.)
rwu@phs.psu.edu (R.W.)
peixu@cjlu.edu.cn (P.X.)

## Highlights

FPP-FM allows real-time depicting of genetic interactions in response to environment

It estimates genetic effects from physiological and developmental principles of growth

FPP-FM serves the translation of physiological data into genetic gain

## Article

# Functional physiological phenotyping with functional mapping: A general framework to bridge the phenotype-genotype gap in plant physiology

Arun K. Pandey,[1] Libo Jiang,[2] Menachem Moshelion,[3,*] Sanbon Chaka Gosa,[3] Ting Sun,[1] Qin Lin,[4] Rongling Wu,[5,*] and Pei Xu[1,6,*]

## SUMMARY

**The recent years have witnessed the emergence of high-throughput phenotyping techniques. In particular, these techniques can characterize a comprehensive landscape of physiological traits of plants responding to dynamic changes in the environment. These innovations, along with the next-generation genomic technologies, have brought plant science into the big-data era. However, a general framework that links multifaceted physiological traits to DNA variants is still lacking. Here, we developed a general framework that integrates functional physiological phenotyping (FPP) with functional mapping (FM). This integration, implemented with high-dimensional statistical reasoning, can aid in our understanding of how genotype is translated toward phenotype. As a demonstration of method, we implemented the transpiration and soil-plant-atmosphere measurements of a tomato introgression line population into the FPP-FM framework, facilitating the identification of quantitative trait loci (QTLs) that mediate the spatiotemporal change of transpiration rate and the test of how these QTLs control, through their interaction networks, phenotypic plasticity under drought stress.**

## INTRODUCTION

Plant phenotyping technologies have been described as the quantification of specific plant traits for research or breeding purposes that include growth, development, environmental responses, and other morphological and physiological changes (Araus et al., 2018). High-throughput phenotyping techniques built on automatic or semiautomatic platforms have emerged in recent years, providing the capacity to characterize structural and functional traits of plants from the cellular level to the whole plant or canopy level under controlled conditions (Tardieu et al., 2017). On the other hand, next-generation genotyping technologies have also progressed rapidly and provided great value in the field of genomic profiling and DNA sequencing (Fahlgren et al., 2015). Such advancement in technology is seeking to build an era of digital agriculture that is expected to significantly benefit plant researchers, breeders, and ultimately, consumers.

Drought is a severe agricultural problem facing plant scientists and farmers worldwide. From an agronomic perspective, drought resistance is defined by the capacity of a plant to sustain yield under water deficit during the life cycle (Negin and Moshelion, 2017). Physiological traits related to plant water relations are among the most relevant traits to agronomic drought resistance, whose fine-tuning would lead to optimized water management in response to drought stress (Blum, 2009; Sinclair et al., 2017). These traits include, for example, water use efficiency (Blum, 2009), transpiration-limited traits (Sinclair et al., 2017), midday water potential (Williams and Araujo, 2002), and stomatal conductance (Roche, 2015). Recently, the term "functional physiological traits" or "quantitative physiological traits" (QPTs) was coined to refer to these traits (Negin and Moshelion, 2017). Compared with morphological traits, which are defined as the changes to the outward appearance of a plant or the form and structure of internal parts (e.g. stem greenness, leaf wiltness), physiological traits are more sensitive to drought stress and thus provide a means of manipulating plant drought responses in an earlier stage before any morphological response is visible

[1]College of Life Sciences, China Jiliang University, Hangzhou 310018, China

[2]Center for Computational Biology, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100080, China

[3]The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel

[4]Biozeron Biotechnology Co., Ltd, Shanghai 201800, China

[5]Center for Statistical Genetics, Departments of Public Health Sciences and Statistics, The Pennsylvania State University, Hershey, PA 17033, USA

[6]Lead contact

*Correspondence:
menachem.moshelion@mail.huji.ac.il (M.M.),
rwu@phs.psu.edu (R.W.),
peixu@cjlu.edu.cn (P.X.)
https://doi.org/10.1016/j.isci.2021.102846

(Negin and Moshelion, 2017). In field conditions, plant drought responses are highly dynamic because the ambient environment is changing. Depending on the drought scenarios (mild or severe, transient, or long-term), a certain QPT may impact the yield positively or negatively. For example, low stomatal conductance is beneficial under severe drought conditions as it prevents excess water loss, but it may be undesired under mild drought as it constrains photosynthesis as well (Galkin et al., 2018). The property of an organism to produce distinct phenotypes in response to the environment is known as phenotypic plasticity (Davière and Achard, 2016; Ibañez et al., 2017). The degree of the phenotypic plasticity of quantitative physiological traits in plants can be determinant of their adaptive advantage to the environment (Nicotra et al., 2010; Pham and McConnaughay, 2014), thus regarded as being a valuable breeding trait related to food security (Gratani, 2014; Puglielli et al., 2017). However, it has never been easy to harness QPT plasticity. First, there are difficulties in precise phenotyping of yield-related QPTs due to their dynamic nature. A robust method requires a phenotyping system capable of continuously and simultaneously monitoring plant physiological traits responding to dynamic environmental fluctuations, in particular, soil-plant-atmosphere continuum (SPAC) parameters (Xu et al., 2015). Second, there is currently a huge phenotype-genotype knowledge gap to handle the massive quantitative physiological traits data acquired from pedigreed or natural populations to enable the identification of specific quantitative trait loci (QTLs) and their genetic interactions underlying phenotypic variations. Traditional genetic mapping methods deal with the time-series data each at a single time point (thus is "static"), usually failing to uncover the inherent dynamics of QTL effects (Li et al., 2015; Feldman et al., 2018; Guo et al., 2018). In addition, as the data volume increases radically in the phenomics era, traditional mapping may become technically intractable. There is hence an urgent need for more robust ways to integrate physiological knowledge, statistical genetics, and computational biology to dissect when (temporal) and where (spatial) a given genotype or allele has comparative advantages under stressful conditions.

Recently, several high-throughput phenotyping platforms tailored for physiological measurements have emerged to provide solutions to the aforementioned challenge in precise QPT phenotyping. Among them, platforms based on weighing lysimeter and gravimetric arrays are nondestructive, noninvasive systems enabling simultaneous and continuous monitoring of plant water relations under dynamic radiation, temperature, relative humidity, and vapor pressure deficit conditions (Vadez et al., 2015; Rymaszewski et al., 2017) Nevertheless, efficient genetic mapping of these longitudinal quantitative traits remains a major challenge, despite a few recent advancements in functional mapping (FM) using different modeling strategies such as Bayesian group lasso approach and estimating equations approach (Li et al., 2015; Li and Sillanpää, 2015). FM is a general statistical mapping framework proposed to characterize QTLs that underlie a complex dynamic trait (Li et al., 2015; Wu and Lin, 2006). It was developed based on the assumption of function-valued traits that phenotypic values expressed in the course of plant growth or treatment follow a continuous function/curve over time (Gomulkiewicz et al., 2018). By using mathematical equations to integrate the dynamic pathways behind the phenotypic formation, FM can unveil QTLs involved in rate-limiting processes and quantify the dynamic change of their genetic effects across a time or space scale (He et al., 2010; Li and Wu, 2010). From a statistical perspective, functional mapping estimates parameters that determine a genotype-specific growth or response curve and/or the covariance structure, thus strikingly increasing its statistical power for QTL detection (Sillanpää et al., 2012; Jiang et al., 2016). FM has shown remarkable power in associating QTLs with dynamic traits in plants and animals, such as growth trajectory, allometric scaling (Sillanpää et al., 2012; Jiang et al., 2016), biological shape (Sun et al., 2018), and drug or nutrient responses (Wang et al., 2013; Jiang et al., 2019).

Here, we developed a novel framework called FPP-FM in which phenotypic data from functional physiological phenotyping (FPP) under multiple treatments, together with genotypic (marker) information, are implemented into a functional mapping (FM) algorithm in order to reveal the QTLs behind the QPT and assess the interplay between QTLs or QTL-environment interactions over time (Figure 1). FPP is a physiology-based, high-throughput, nondestructive, and non-invasive phenotyping technique that continuously measures the plant and its ambient conditions including soil and atmosphere (Gosa et al., 2019). By showcasing the application of FPP-FM to an introgression line population of tomato under progressive drought stress followed by recovery, we demonstrate the efficiency and power of this method in dissecting the genetic architecture of the dynamic trait of weight-normalized transpiration rate (E). Because the ideas of FPP and FM both are universal, this joint framework can be applied to other longitudinal traits related to development and responses to broader types of abiotic and biotic stresses.
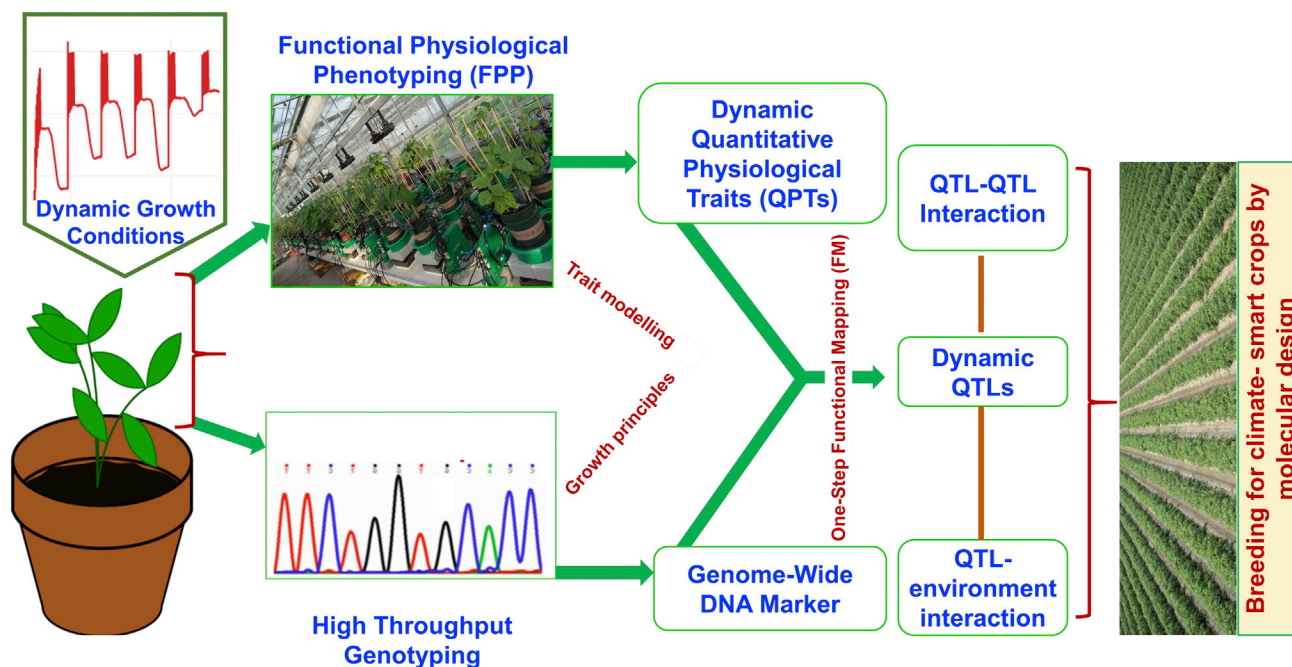
**Figure 1. Schematic display of the principle of joint FPP-FM analysis**
As high-throughput genotyping has no longer been a bottleneck in the post-genome era, the cornerstones of the joint FPP-FM framework should include: (1) a high-throughput FPP platform enabling acquisition of dynamic quantitative physiological traits (QPTs) data from a natural or pedigree population. Exemplified here is the PlantArray system that collects dynamic data of environmental parameters (e.g. temperature, radiation), stress strength (soil water content) and system weight from hundreds of plants simultaneously, which are translated into critical whole-plant water-relation parameters (e.g. E, Gs, WUE) by using proved, physiology-based algorithms. (2) a powerful genetic framework in which dynamic QPTs data, together with genome-wide DNA marker information, are implemented into a one-step FM algorithm to reveal the QTLs behind the QPT plasticity and assess the interplay between QTLs or QTL-environment interactions over time. Through joint FPP-FM analysis, the dynamic effects of the QTLs, their interaction with specific environmental scenarios and the QTL network information that highlights QTL topology, organization, evolution and hub QTLs in the course of treatment are elucidated.

## RESULTS

### Functional physiological trait analysis

The first step of FPP-FM is to acquire dynamic quantitative physiological traits data from a natural or pedigreed population. Here, the population used was an introgression line population of tomato, which were subjected to a treatment scenario comprising normal irrigation, progressive drought, and water recovery in a high-throughput, automatic phenotyping system (Figure 1). During the entire course of this experiment, the pheno-typic data of multiple drought-responsive traits such as plant weight and transcription rate were continually recorded alongside with the stress parameters. A detailed description of the FPP process for this population was already published by Halperin et al. (2017). For a reasonable computational intensity, here we used tran-spiration rate (Tr) as an example to show how our proposed framework translated the dynamic phenotypic in-formation into genetic knowledge. Detailed phenotypic and genotypic data used in this study is deposited in Table S1 (source data). In Figures 2A–2C, the dynamic patterns of transpiration rate from five randomly selected IL lines were displayed, which were recorded every three minutes spanning the before-drought (Fig-ure 2A), drought-stressed (Figure 2B), and the recovery phases (Figure 2C). Transpiration rate is known to associate with yield penalty under drought stress (De Wit, 1958; Sade et al., 2010). Upon water withholding, a profound reduction of transpiration rate was observed for all five lines (Figure 2B), whereas the genotypic differences in agility and magnitude of the response were apparent. Following water resumption, significant genotypic differences in the pattern of recovery were observed, which, however, were not correlated to the plant behaviors in response to water deficit. The line IL2-1-1, for example, maintained the highest transpiration rate under drought stress (Figure 2B), but its capacity to recover was weak following rewatering (Figure 2C).

### Fitting of multiphased phenotypic data and QTL mapping by functional mapping

We implemented the Legendre function and the structured antedependence model (Zhao et al., 2005) to the phenotypic data. We modeled the QTL genotypic mean vector of transpiration rate and approximated
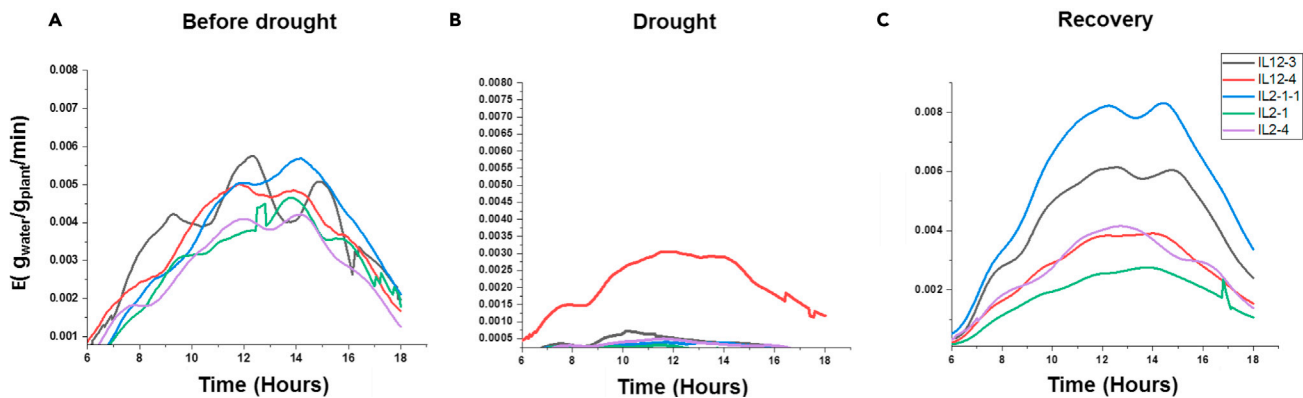
**Figure 2. Graphic display of transpiration rate for five representatives ILs**

Data from the before-drought (A), drought-stress (B), and recovery (C) phases are shwon. Measurements were taken using the FPP platform PlantArray 3.0 system (Plant-DiTech, Rehovot, Israel). Plants were grown in 3.9-L pots filled with sand, in a partially controlled greenhouse in which the temperature was moderated by fans cooling a wet mattress and the plants were exposed to natural day length and light. The weight output of the load cells was monitored every 3 min.

.

the structure of the covariance matrix, which was a key step to integrate the fitted parameters into the functional mapping framework (see "STAR methods" for details). The mean curves of the normalized transpiration rate for all the ILs in nonstressed (day 0 to day 12), drought-stressed (day 13 to day 23), and recovery (day 24 to day 30) phases were demonstrated in Figure 3A. It clearly showed that, as a whole, in the first phase, the transpiration rate increased gradually as the plants grew; in the second phase, its trend turned to a gradual decrease; and in the third phase, the transpiration rate increased rapidly, which reached the maximum level on the fifth day after rewatering and was followed by a decrease since then.

Following incorporating the fitted parameters into functional mapping, we identified, in a single step, 7 large-effect QTLs under the threshold of $p < 3.979 \times 10^{-6}$ calculated by permutation test (Figure 3B). The most significant QTL was located on chromosome 11 (97.30–103.4 cM, *Qtr.11* hereafter). The effects of the two alleles (QQ and qq) at this QTL were found to show a major difference only in the third phase, suggesting that *Qtr.11* is a conditional QTL depending on the specific treatment scenario of rehydration (Figure 3C). Indeed, by further testing hypothesis 11 (see "STAR methods"), we revealed that the expression of *Qtr.11* in response to water recovery was through QTL-environment interaction to affect the transpiration rate ($p < 0.01$).

## Integration of QTLs to develop a network

We further reconstructed networks of QTLs that interact with each other to show how they confer adaptation of plants to environmental changes. The limited number (1,077) of discrete SNPs in the IL population due to the cosegregation of SNPs (2,521 in total) in each chromosome bins allowed us to reconstruct reasonably sized, full-scale networks to reveal the interactions among all QTLs with major or minor effects. Through modeling analysis, networks for the three treatment scenarios were constructed that in total included 71 nodes, with each node denoting a QTL and each edge representing the interaction between two QTLs. The shifts of nodes in different networks led to changes of the network topology in various scenarios, reflecting dynamic responses of QTL interactions to the environment. We identified four hub nodes, Q11, Q42, Q46, and Q60, under the normal (well-watered) condition, which modulated the genetic network of transpiration rate predominately by activating other QTLs. Under drought condition, 8 hub QTLs (Q64, Q28, Q51, Q14, Q49, Q43, Q48, and Q66) were detected, which played a role either by activating or repressing the expressions of other QTLs and were under positive or negative regulations by other core or peripheral QTLs themselves. We detected two hub nodes (Q66 and Q42) that interacted with other nodes to maintain the stability and dynamics of the recovery network. Highly interestingly, Q66 and Q42 were previously found in the drought and normal networks, respectively, and thus were considered to be plasticity QTLs linking normal transpiration and stress responses for balanced plant growth and drought tolerance upon water recovery. Because episodes of severe stress are often followed by periods of alleviated tension in field conditions, our findings of environment-depend QTLs, in particular, the discovery of the hub
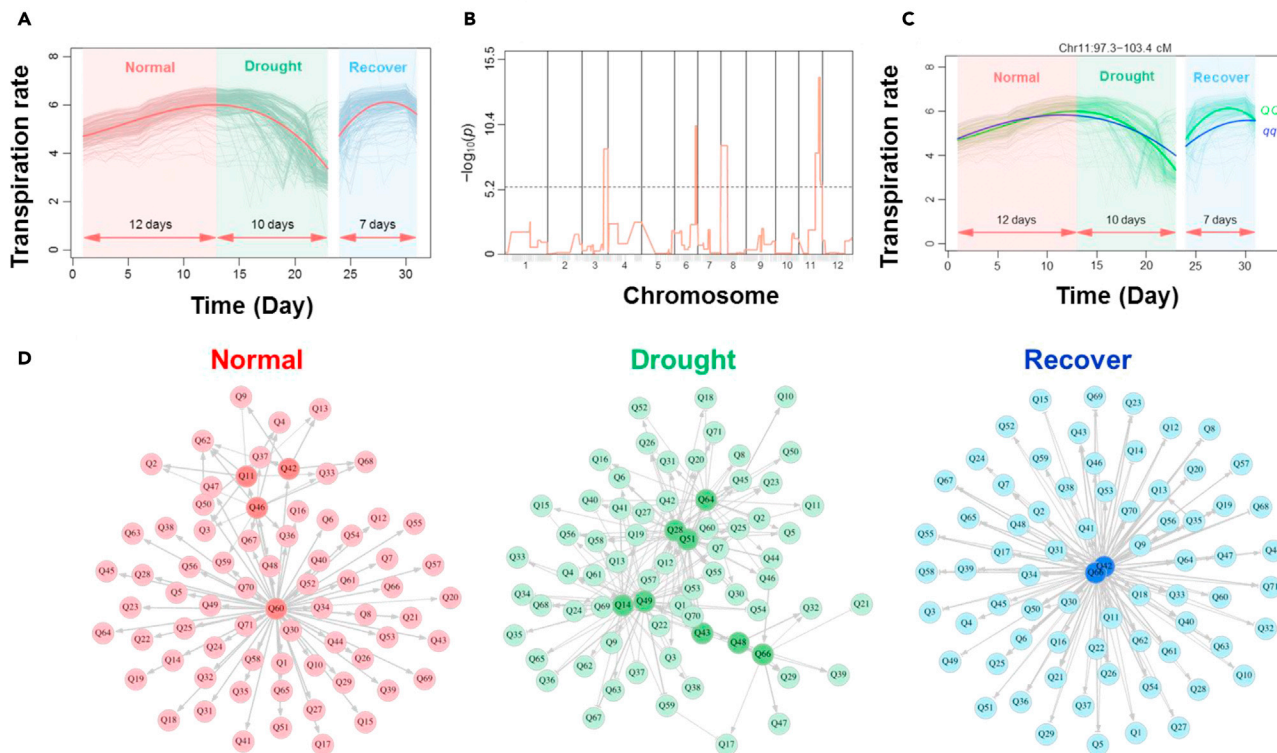
**Figure 3. Dynamic QTLs and QTL networks controlling transpiration rate**

(A), longitudinal transpiration rates of all ILs whose mean curve (red) was fitted by Legendre function; (B), Manhattan plot of significances for the association of markers with dynamic changes of E across the tomato genome by FPP-FM. The dashed horizontal line indicates the critical threshold of p value at the 5% significance level determined by the permutation test.

(C), dynamic effects of the two genotypes (*QQ* and *qq*) at the most significant QTL (chromosome 11) on transpiration rate.

(D), QTL interaction networks in the pre-treatment, drought-stressed, and recovery phases. Each node denotes a QTL and each edge represents the interaction between two QTLs. The hub QTLs were highlighted with darkened colors in each network.

plasticity QTLs in the recovery phase, have great agronomic implications. Leveraging this core genetic knowledge gained from FPP-FM analysis would facilitate the modulation of plant fitness and resiliency through molecular breeding approaches, to ultimately achieve the goal of minimum yield loss under various drought conditions.

## Comparison with existing methods and simulation analysis

To our knowledge, there is no currently available tool like ours that is capable of fitting multiphased physiological data and being integrated with functional mapping. We conducted a Bayesian group lasso analysis and ANOVA with the same genotypic and phenotypic data to determine the extent to which the QTLs mapping was affected by the choice of methods. As a result, the Bayesian method detected only one QTL with a relative high p value ($-\log_{10}P = 1.5$, Figure S1) whereas the ANOVA was able to detect only two QTLs, one for the fourth day ($-\log_{10}P = 10.2$, Figure S2, Table S2) and the other for the 22$^{nd}$ day ($-\log_{10}P = 6.3$, Figure S3, Table S3) of the experiment. The positions of the significant markers detected by different methods were consistent. These results demonstrate that FPP-FM was precise and more powerful in detecting QTLs from dynamic physiological data.

To further provide a theoretical estimate of the powers of FPP-FM and the traditional mapping method (ANOVA, for IL population), a simulation analysis was performed. As expected, ANOVA well detected significant QTLs from the uncorrelated data at a large sample size and heritability; however, its power for QTL detection in time-correlated data was substantially low (0.34–0.61; Table 1). In contrast, our new model showed higher power to detect QTLs underlying the time-correlated data, whereas its performance in handling time-uncorrelated data was similar to that of ANOVA (Table 1). The new model also showed

**Table 1. Power of QTL detection by ANOVA and FPP-FM from the data simulated under different heritabilities (0.05 and 0.10), for the sample sizes 500 and 200**

| Method | | Simulation | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Time-uncorrelated | | | | time-correlated | | | |
| | | N = 200 | | N = 500 | | N = 200 | | N = 500 | |
| | | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 |
| | ANOVA | 74% | 84% | 86% | 92% | 34% | 57% | 42% | 61% |
| | FPP-FM | 72% | 83% | 88% | 93% | 78% | 92% | 91% | 95% |

much lower false-positive rate in QTLs detection in the simulation analysis, especially for the time-correlated data (1/2–1/7 of that from ANOVA, Table 2). To conclude, FPP-FM was proved to be a powerful new tool to identify specific QTLs that govern processes and patterns of physiological traits over time.

## DISCUSSION

The phenomics era is witnessing a burst of multiphased phenotypic data derived from various treatments. In order to leverage this advantage in breeding programs, the genetic architecture behind the traits must be elucidated to enable breeders to manipulate these traits with DNA markers for a large number of genotypes. To date, the contribution of the physiological trait-based selection of optimal drought response is minor (Ghanem et al., 2015) as compared with morphological traits (Sadras and Richards, 2014) largely because of the big challenges in genetic dissection of these traits (Easlon et al., 2014; Feldman et al., 2018). This has become a rate-limiting step in reaping the fruit of phenomics. For example, to increase carbon assimilation toward a higher yield, one may consider manipulating Gs or the biosynthesis of its regulator, ABA (Sinclair et al., 1984; Blum, 2009). However, as stomata is the gateway of both $H_2O$ and $CO_2$, selection for stomata behavior needs to be very conditional and dynamic to balance water evaporation and carbon intake according to specific ambient conditions that typically related to geographic locations, cultivation seasons, and plant growth stages. Clearly, the success of harnessing quantitative physiological trait plasticity to improve agronomic drought resistance relies heavily on prior knowledge on QTLs governing the individual traits, and hence, it is imperative to map and identify a comprehensive set of QTLs that are involved in the ecophysiological dynamics of drought response during the plant's life cycle.

In this article, we propose that the principle of functional physiological phenotyping and that of functional mapping can be unified to provide a powerful tool to draw real-time pictures that illustrate the causal networks of genetic interactions along the short-term axis of physiological response to the environment and the long-term axis of development. By taking transpiration rate as an example, we demonstrated how the proposed FPP-FM framework help to better unravel and comprehend the genetic architecture of phenotypic formation and plasticity. As bridging the phenotype-genotype gap is at the core of advancing our mechanistic understanding of phenotypic complexities, alterations, and novelties, our conceptual framework potentially provides an engine for this aim and specifically offers a novel physiological-genetic-environmental functional strategy to dissect the main contributors to yield stableness. We implement the idea of FM (Wu and Lin, 2006) to map multitreatment physiological data in time-series by combining multistage nonparametric equation and high-dimensional covariance matrix. Our approach detects and estimates genetic effects due to allelic/nonallelic actions and QTL interactions from physiological and developmental principles of growth. The simulation study clearly showed an over performance of our approach to the traditional static mapping for longitudinal phenotypic data. By choosing an optimal order of LOP via information criteria, our approach overcomes the overfitting or underfitting of multiphased physiological data. The approach can reconstruct QTL networks, providing a powerful tool to understand how QTL-QTL and QTL-environment interactions affect trait formation in a set of specific environmental scenarios and thereby allows for identifying the hub QTLs in network topology, organization, and function.

In the demonstrative case study here, a tomato pedigreed population was analyzed; if a mapping population is constructed using a natural population, the theory of linkage disequilibrium mapping should be integrated within the functional mapping framework to provide a powerful tool for fine-scale mapping of

**Table 2. False positive rate of QTL detection by ANOVA and FPP-FM from the data simulated under different heritabilities (0.05 and 0.10), for the sample sizes 500 and 200**

| | | Simulation | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Time-uncorrelated time-correlated | | | | | | | |
| Method | | N = 200 | | N = 500 | | N = 200 | | N = 500 | |
| | | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 | $H^2 =$ 0.05 | $H^2 =$ 0.1 |
| | ANOVA | 22% | 14% | 11% | 8% | 45% | 34% | 31% | 28% |
| | FPP-FM | 19% | 15% | 9% | 6% | 23% | 17% | 11% | 4% |

complex traits (Lou et al., 2003). Therefore, the extension of our model to tackle genotype-environment interactions using genome-wide association study is straightforward. It is also noteworthy that the limited capacity of physiological phenotyping and the challenge in acquiring difficult-to-measure traits like photosynthesis still exists, which restricts the immediate use of our method to larger and richer types of populations. However, the highly uniform environments (hence less noises) in this kind of experiments would reduce the size requirement of population in genetic mapping. More importantly, with the rapid deployment of phenotyping facilities all over the world along with their growing phenotyping capacity (e.g. more number of units, smaller size) and the successful integration of lysimeters and imaging system (Weksler et al., 2020), FPP-FM will show its usefulness in serving the translation of physiological data into genetic gain. In principle, our method is extendable to morphological traits showing plasticity and dependence on treatment because the concept of functional mapping is general and the idea of integrating dynamic phenotypes into functional mapping is also general. Equation (1) is a redundant joint likelihood, thus can be extended to consider multiple phenotypes at once. Nevertheless, because different types of traits are usually captured by different phenotyping systems, a case-by-case modification of the fitting algorithm for the phenotype will be needed.

The proposed FPP-FM framework can further be integrated with molecular biology approaches such as transcriptomic, proteomic and epigenetic profiling as well as gene/protein coexpression networking, to more deeply and systematically reveal how the plant water status and yield are affected by the interplay between genetic, epigenetic, and environmental factors. In such an effort, greater biological relevance will be achieved if the molecular analysis is guided by the principle of FPP because it allows for sampling the plant tissues based on physiological status rather than time of treatment. It is therefore expected to offer unprecedented opportunities to reveal the molecular mechanisms underlying genotypic differences of plant responses to the environment. From the perspective of breeding, FPP-FM is also a useful tool. First, it provides markers linked to yield-related physiological traits, which will allow for earlier stage, more direct, and efficient selection of the genes regulating transpiration and photosynthesis under drought stresses. Second, the dissection of the dynamic expression of the QTLs, i.e. when they are expressed as the stress scenario evolves, is valuable for the molecular design of crops adaptive to a certain ecoclimatic condition. For example, in the current case, the QTL *Q66* may represent a good selection target in areas featured by intermittent drought, provided that this QTL plays a role in balancing drought response and plant growth; however, in long-term, severe drought-prone areas, QTLs that confer stably conservative water use may be preferred.

### Limitations of the study

As aforementioned, the limited capacity of physiological phenotyping in many current platforms and the still-existing challenge in acquiring difficult-to-measure traits like photosynthesis are the major factors restricting the immediate use of our method to larger populations. Also, despite the principle of our method is general, the application of it to traits that are captured by other phenotyping systems may need a case-by-case modification of the fitting algorithm.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102846.

## AUTHOR CONTRIBUTIONS

P.X, M.M, and R.W conceived the idea. P.X, L.J, A.K.P, M.M., and R.W designed the experiments and developed statistical algorithms. A.K.P, L.J, S.C.G, and Q.L performed the experiment. L.J, P.X, S. C.G and M.M analyzed and interpreted the data. A.K.P and P.X. wrote the manuscript. All authors read and approved the final manuscript.

## DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

## REFERENCES

Araus, J.L., Kefauver, S.C., Zaman-Allah, M., Olsen, M.S., and Cairns, J.E. (2018). Translating high-throughput phenotyping into genetic gain. Trends Plant Sci. *23*, 451–466.

Blum, A. (2009). Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. Field Crops Res. *112*, 119–123.

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. InterJournal, Complex Systems *1695*, 1–9.

Dalal, A., Shenhar, I., Bourstein, R., Mayo, A., Grunwald, Y., Averbuch, N., Attia, Z., Wallach, R., and Moshelion, M. (2020). A telemetric, gravimetric platform for real-time physiological phenotyping of plant–environment interactions. J. Vis. Exp. *162*, e61280.

Davière, J.M., and Achard, P. (2016). A pivotal role of DELLAs in regulating multiple hormone signals. Mol. Plant *9*, 10–20.

De Wit, C.T. (1958). Transpiration and crop yields. Versl. Landbouwk. Onderz. *64*, 6, Inst. Biol. Chem. Res. On Field Crops and Herbage, Wageningen,The Netherlands.

Easlon, H.M., Nemali, K.S., Richards, J.H., Hanson, D.T., Juenger, T.E., and McKay, J.K. (2014). The physiological basis for genetic variation in water use efficiency and carbon isotope composition in Arabidopsis thaliana. Photosynth. Res. *119*, 119–129.

Eshed, Y., and Zamir, D. (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics *141*, 1147–1162.

Fahlgren, N., Feldman, M., Gehan, M.A., Wilson, M.S., Shyu, C., Bryant, D.W., Hill, S.T., McEntee, C.J., Warnasooriya, S.N., Kumar, I., and Ficor, T. (2015). A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in Setaria. Mol. Plant *8*, 1520–1535.

Feldman, M.J., Ellsworth, P.Z., Fahlgren, N., Gehan, M.A., Cousins, A.B., and Baxter, I. (2018). Components of water use efficiency have unique genetic signatures in the model $C_4$ grass Setaria. Plant Physiol. *178*, 699–715.

Galkin, E., Dalal, A., Evenko, A., Fridman, E., Kan, I., Wallach, R., and Moshelion, M. (2018). Risk-management strategies and transpiration rates of wild barley in uncertain environments. Physiol. Plant *164*, 412–428.

Ghanem, M.E., Marrou, H., and Sinclair, T.R. (2015). Physiological phenotyping of plants for crop improvement. Trends Plant Sci. *20*, 139–144.

Gomulkiewicz, R., Kingsolver, J.G., Carter, P.A., and Heckman, N. (2018). Variation and evolution of function-valued traits. Annu. Rev. Ecol. Evol. Syst. *49*, 139–164.

Gosa, S.C., Lupo, Y., and Moshelion, M. (2019). Quantitative and comparative analysis of whole-plant performance for functional physiological traits phenotyping: new tools to support pre-breeding and plant stress physiology studies. Plant Sci. *282*, 49–59.

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. Adv. Bot. 1–17.

Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., Jiang, N., Feng, H., Huang, C., Yang, P., and Zhao, H. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. Mol. Plant *11*, 789–805.

Halperin, O., Gebremedhin, A., Wallach, R., and Moshelion, M. (2017). High-throughput physiological phenotyping and screening system for the characterization of plant–environment interactions. Plant J. *89*, 839–850.

He, Q., Berg, A., Li, Y., Vallejos, C.E., and Wu, R. (2010). Mapping genes for plant structure, development and evolution: functional mapping meets ontology. Trends Genet. *26*, 39–46.

Ibañez, C., Poeschl, Y., Peterson, T., Bellstädt, J., Denk, K., Gogol-Döring, A., Quint, M., and Delker, C. (2017). Ambient temperature and genotype differentially affect developmental and phenotypic plasticity in Arabidopsis thaliana. BMC Plant Biol. *17*, 1–14.

Jiang, L., Sun, L., Ye, M., Wang, J., Wang, Y., Bogard, M., Lacaze, X., Fournier, A., Beauchêne, K., Gouache, D., and Wu, R. (2019). Functional mapping of N deficiency-induced response in wheat yield-component traits by implementing high-throughput phenotyping. Plant J. *97*, 1105–1119.

Jiang, Q., Huang, B., and Yan, X. (2016). GMM and optimal principal components-based Bayesian method for multimode fault diagnosis. Comput. Chem. Eng. *84*, 338–349.

Li, J., Wang, Z., Li, R., and Wu, R. (2015). Bayesian group lasso for nonparametric varying-coefficient models with application to functional genome-wide association studies. Ann. Appl. Stat. *9*, 640–664.

Li, Y., and Wu, R. (2010). Functional mapping of growth and development. Biol. Rev. *85*, 207–216.

Li, Z., and Sillanpää, M.J. (2015). Dynamic quantitative trait locus analysis of plant phenomic data. Trends Plant Sci. *20*, 822–833.

Lin, M., and Wu, R. (2006). A joint model for nonparametric functional mapping of longitudinal trajectory and time-to-event. BMC Bioinformatics *7*, 1–2.

Lou, X.Y., Casella, G., Littell, R.C., Yang, M.C., Johnson, J.A., and Wu, R. (2003). A haplotype-based algorithm for multi-locus linkage disequilibrium mapping of quantitative trait loci with epistasis. Genetics *163*, 1533–1548.

Negin, B., and Moshelion, M. (2017). The advantages of functional phenotyping in pre-field screening for drought-tolerant crops. Funct. Plant Biol. *44*, 107–118.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., Davidson, A.M., Finnegan, E.J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F., and van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. Trends Plant Sci. *15*, 684–692.

Pham, B., and McConnaughay, K. (2014). Plant phenotypic expression in variable environments. In Ecology and the Environment. The Plant Sciences 8, R. Monson, ed. (Springer), pp. 119–141.

Puglielli, G., Catoni, R., Spoletini, A., Varone, L., and Gratani, L. (2017). Short-term physiological plasticity: Trade-off between drought and recovery responses in three Mediterranean Cistus species. Ecol. Evol. *7*, 10880–10889.

Roche, D. (2015). Stomatal conductance is essential for higher yield potential of C3 crops. Crit. Rev. Plant Sci. *34*, 429–453.

Rymaszewski, W., Vile, D., Bediee, A., Dauzat, M., Luchaire, N., Kamrowska, D., Granier, C., and Hennig, J. (2017). Stress-related gene expression reflects morphophysiological responses to water deficit. Plant Physiol. *174*, 1913–1930.

Sade, N., Gebretsadik, M., Seligmann, R., Schwartz, A., Wallach, R., and Moshelion, M. (2010). The role of tobacco Aquaporin1 in improving water use efficiency, hydraulic conductivity, and yield production under salt stress. Plant Physiol. *152*, 245–254.

Sadras, V.O., and Richards, R.A. (2014). Improvement of crop yield in dry environments: benchmarks, levels of organisation and the role of nitrogen. J. Expt. Bot. *65*, 1981–1995.

Sillanpää, M.J., Pikkuhookana, P., Abrahamsson, S., Knürr, T., Fries, A., Lerceteau, E., Waldmann, P., and García-Gil, M.R. (2012). Simultaneous estimation of multiple quantitative trait loci and growth curve parameters through hierarchical Bayesian modeling. Heredity *108*, 134–146.

Sim, S.C., Durstewitz, G., Plieske, J., Wieseke, R., Ganal, M.W., Van Deynze, A., Hamilton, J.P., Buell, C.R., Causse, M., Wijeratne, S., and Francis, D.M. (2012). Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. PLoS One *7*. https://doi.org/10.1371/journal.pone.0040563.

Sinclair, T.R., Devi, J., Shekoofa, A., Choudhary, S., Sadok, W., Vadez, V., Riar, M., and Rufty, T. (2017). Limited-transpiration response to high vapor pressure deficit in crop species. Plant Sci. *260*, 109–118.

Sinclair, T.R., Tanner, C.B., and Bennett, J.M. (1984). Water-use efficiency in crop production. Bioscience *34*, 36–40.

Sun, L., Wang, J., Zhu, X., Jiang, L., Gosik, K., Sang, M., Sun, F., Cheng, T., Zhang, Q., and Wu, R. (2018). Hp QTL: a geometric morphometric platform to compute the genetic architecture of heterophylly. Brief. Bioinformatics *19*, 603–612.

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., and Bennett, M. (2017). Plant phenomics, from sensors to knowledge. Curr. Biol. *27*, 770–783.

Vadez, V., Kholová, J., Hummel, G., Zhokhavets, U., Gupta, S.K., and Hash, C.T. (2015). LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. J. Exp. Bot. *66*, 5581–5593.

Wang, Z., Luo, J., Fu, G., Wang, Z., and Wu, R. (2013). Stochastic modeling of systems mapping in pharmacogenomics. Adv. Drug Deliv. Rev. *65*, 912–917.

Weksler, S., Rozenstein, O., Haish, N., Moshelion, M., Walach, R., and Ben-Dor, E. (2020). A hyperspectral-physiological phenomics system: measuring diurnal transpiration rates and diurnal reflectance. Remote Sens. *12*, 1493.

Williams, L.E., and Araujo, F.J. (2002). Correlations among predawn leaf, midday leaf, and midday stem water potential and their correlations with other measures of soil and plant water status in Vitis vinifera. J. Am. Soc. Hortic. Sci. *127*, 448–454.

Wu, R., and Lin, M. (2006). Functional mapping—how to map and study the genetic architecture of dynamic complex traits. Nat. Rev. Genet. *7*, 229–237.

Xu, P., Moshelion, M., Wu, X., Halperin, O., Wang, B., Luo, J., Wallach, R., Wu, X., Lu, Z., and Li, G. (2015). Natural variation and gene regulatory basis for the responses of asparagus beans to soil drought. Front. Plant Sci. *6*, 891.

Zhao, W., Chen, Y.Q., Casella, G., Cheverud, J.M., and Wu, R. (2005). A non-stationary model for functional mapping of complex traits. Bioinformatics *21*, 2469–2477.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT OR RESOURCE | SOURCE | IDENTIFIER |
| --- | --- | --- |
| R software | | https://github.com/FFP-FM/Version1 |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact: Prof. Pei Xu (peixu@cjlu.edu.cn)

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

All statistical algorithms that build our framework into a user-friendly R package for public use are available on https://github.com/FFP-FM/Version1.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS SECTION

An introgression line (IL) population of tomato composed of 62 lines were used to determine whole-plant transpiration rates using an array of lysimeters placed in the greenhouse (Plantarray 3.0 system; Plant-DiTech) in the ''iCORE Center for Functional Phenotyping'' (http://departments.agri.huji.ac.il/plantscience/icore.phpon). We followed the Bayesian group lasso method with associated MCMC algorithms as developed by Li et al. (2015). ANOVA is a common (and ''static'') method used to detect associations between molecular markers and traits in IL populations.

## METHOD DETAILS

### Plant materials and growth conditions

In 2018, an introgression line (IL) population of tomato composed of 62 lines (Eshed and Zamir, 1995) were used to determine whole-plant transpiration rates (Table S1) using an array of lysimeters placed in the greenhouse (PlantArray 3.0 system; Plant-DiTech) in the ''iCORE Center for Functional Phenotyping'' (http://departments.agri.huji.ac.il/plantscience/icore.phpon), as described in details by Halperin et al. (2017). PlantArray combines gravimetric system, soil and atmospheric probes, controller and irrigation valves in a unit, which can measure or calculate plant water relation traits as well as the soil-atmosphere parameters at the same time, during the whole plant growth process. Briefly, plants were grown in 3.9L-pots filled with an inert substance (quartz sand) under semi-controlled temperature conditions (20-34°C day and 18-24°C night temperature) with natural light. Pots were placed on a temperature-compensated load cell with digital output (**Vishay Tedea Hunt Leigh**). Plants were grown in a circle cut out of white plates that is used to prevent evaporation. We irrigated pots at night in multiple irrigations (3 minutes irrigation, every 15 minutes for 6 hours) during the pre-drought treatment period with dripper assemblies (**Netafim, Israel**) pushed into the sand to ensure the medium was evenly wet. Once the plants reach high daily transpiration (above 400g/day), drought treatment was started by gradually irrigating plants at 60% of their daily transpiration while allowing all plants to gradually experience water stress. After each plant reached its $\theta_{crit}$, the point at which the availability of soil water becomes a limiting factor to transpiration rate, the irrigation was resumed to evaluate the recovery of each genotype. The SNPs genotypic information (Table S1) of this population was retrieved from Sim et al. (2012) and used during analysis.

### Whole-plant transpiration and soil-atmospheric measurements

The PlantArray 3.0 platform is a multi-sensor telemetric platform that enables simultaneous and continuous monitoring of key plants' physiological traits. A detailed description of the system was recently published in Dalal et al. (2020). Briefly, the system consists of highly sensitive, temperature compensated load cells that

are used as a weighing lysimeter. A lysimeter is a measuring device that can be used to measure the amount of actual transpiration which is released by plants from a monotonic weight reduction throughout the day. The amount of water loss by transpiration can be worked out by calculating the difference between the weight before and after irrigation. The water loss is the phase of daily transpiration which was calculated from the weight difference between the two data points for individual plants. The plant weight was calculated as the difference between the total system weight and the sum of the tare weight of pot plus drainage container, the weight of soil at pot capacity, and weight of water in the drainage container at the end of the free drainage (Halperin et al., 2017; Gosa et al., 2019). By using soil weight and initial plant weight, a set of phenotypic data can be extracted including data on growth rates, transpiration, canopy conductance, and soil water content. Soil-atmosphere parameters including temperatures, relative humidity, radiation, and soil moisture were monitored by using the LI-COR 190 Quantum Sensor (**Lincoln, NE, USA**) and the soil moisture sensor (**5TE; Decagon Devices, Pullman, WA, USA**), respectively. The measurement was made every 3 minutes. The data of soil-atmosphere parameters and plant growth and physiological parameters including plant weight, daily transpiration, transpiration rate (Tr) and normalized transpiration (E) were simultaneously saved on the online web-based software SPAC analytics (**Plant-Ditech, Israel**). The data of whole plant transpiration rate (Tr) was used in the following genetic analysis.

## Statistical model for functional mapping

*Joint likelihood.* Let $y_{il} = (y_{il}(t^1_{1il}), ...., y_{il}(t^1_{Til}), y_i(t^2_{1il}), ...., y_i(t^2_{Til}), ..., y_i(t^s_{1il}), ...., y_i(t^s_{Til}))$ denotes the phenotypic value of a trait $l$ ($l=1, ..., L$) for individual $i$ at different time points under treatment $s$ ($s=1, ..., S$) from a natural or controlled population. Note that $(t^s_{1il}, ..., t^s_{Til})$ represents a series $T$ time points which reflects the treatment stage such as drought or recovery. At a given QTL with the $j$th genotype having $n_j$ individuals, the joint likelihood of phenotypic values at this marker is expressed as

$$L = \prod_{l=1}^{L} \prod_{j=1}^{J} \prod_{i=1}^{n_j} f_{lj}(y_{li}; \mu_{lj}, \Sigma_i) \qquad \text{(Equation 1)}$$

where $f_{lj}(y_{li}; \mu_{lj}, \Sigma_i)$ is a multivariate normal distribution with genotype-specific mean vector over different traits, expressed as

$$\mu_{lj} = \left(\mu^1_{lj}, \mu^2_{lj}, .., \mu^S_{lj}\right) = \left(\mu_{lj}(t^1_{1l}), ..., \mu_{lj}(t^1_{Tl}), \mu_{lj}(t^2_{1l}), ..., \mu_{lj}(t^2_{Tl}), ..., \mu_{lj}(t^S_{1l}), ..., \mu_{lj}(t^S_{Tl})\right) \qquad \text{(Equation 2)}$$

and covariance matrix is expressed as

$$\Sigma = \begin{bmatrix} \Sigma_1 & \cdots & \Sigma_{1l} \\ \vdots & \ddots & \vdots \\ \Sigma_{l1} & \cdots & \Sigma_l \end{bmatrix} \qquad \text{(Equation 3)}$$

where $\Sigma_l$ is the residual covariance matrix for trait $l$, $\Sigma_{1l} = \Sigma_{l1}$ is the longitudinal covariance matrix between the two traits.

*Modeling the genotype-specific mean vector.* Functional mapping models the genotypic means for an assumed QTL at different time points by Legendre Orthogonal Polynomials (LOP). This function can fit any kind of curve (Lin and Wu, 2006). The general form of the Legendre polynomial of order $k$ implemented to model the mean value of $j$th genotype for a marker is generally expressed as

$$P_{lsk}(t^s_{Tl}) = \left(P_{ls0}(\tau^s_{Tl}), P_{ls1}(\tau^s_{Tl}), ..., P_{lsk}(\tau^s_{Tl})\right) \qquad \text{(Equation 4)}$$

where

$$P_{lsk}(\tau^s_{Tl}) = \frac{1}{2^k} \sum_{r=0}^{k/2} \frac{(-1)^r (2k-2r)!}{r!(k-r)!(k-2r)!} (\tau^s_{Tl})^{k-2r} \qquad \text{(Equation 5)}$$

where $k = 0, 1, ..., K$ and $\tau^s_{Tl} = -1 + \frac{2(t^s_{Tl} - \min(t^s_{Tl}))}{\max(t^s_{Tl}) - \min(t^s_{Tl})}$ with $\min(t^s_{Tl})$ and $\max(t^s_{Tl})$ being, respectively, the first and last time for the trait $l$ under stage $s$. We rewrite the genotype-specific mean vectors as

$$\mu_{lj} = \left(\mu^1_{lj}, \mu^2_{lj}, .., \mu^S_{lj}\right) = \left(P_{l1k}(t^1_{Tl}) v'_{l1j}, P_{l2k}(t^2_{Tl}) v'_{l2j}, ..., P_{lsk}(t^s_{Tl}) v'_{lsj}\right) \qquad \text{(Equation 6)}$$

where $v_{lsj} = (v_{ls0j}, v_{ls1j}, ...v_{lskj})$ is a vector of genotypic values. By maximizing likelihood (1), time-invariant genotypic values $v_{lsj}$ for the longitudinal trait are estimated.

*Modeling the covariance matrix.* The covariance matrix is modeled by structured antedependence (SAD). SAD models autocorrelate structure of the covariance without relying on a stationary assumption of time-dependent variance and correlation (Zhao et al., 2005). The first-order SAD (SAD(1)) only considers the dependence of errors at the immediate time point, which can be generally applied to model the longitudinal changes of variances and covariances. The residual variance at time $t$ and residual covariance between two times $t_1$ and $t_2$ can be derived as

$$\sigma^2(t) = \left[\frac{1 - \rho^{2t}}{1 - \rho^2}\right]\varphi^2$$

$$\sigma(t_1, t_2) = \rho^{t_2 - t_1}\left[\frac{1 - \rho^{2t_1}}{1 - \rho^2}\right]\varphi^2$$

(Equation 7)

where $\varphi^2$ as the innovation variance at a different time is the constant, and $\rho$ is the degree of antedependence.

*Model selection.* To obtain the best fit of the data, the optimal order of the Legendre polynomial to model the genotypic mean vector should be determined. One can use the AIC information criterion as the model selection criterion for the polynomial order. The AIC is defined as

$$\text{AIC} = -2\ln L(\hat{\mu}, \hat{\Sigma}|k) + 2\text{dimension }(\hat{\mu}, \hat{\Sigma}|k)$$

(Equation 8)

where $(\hat{\mu}, \hat{\Sigma}|k)$ is the MLE of parameters for the Legendre polynomial of order $k$ and dimension $(\hat{\mu}, \hat{\Sigma}|k)$ represents the number of independent parameters under order $k$. The optimal model displays the minimum AIC value.

*Computing algorithm.* It is difficult to derive a closed-form of genotype-specific and covariance structure parameters; due to this reason, Nelder-Mead Simplex or the Quasi-Newton algorithm is used to estimate curve equation as well as the variance-covariance structuring parameters. These two methods do not depend on explicit equations that are implemented.

**Hypothesis tests**

*Existence of a QTL.* Under likelihood (1), whether there exists a QTL for the longitudinal phenotypes can be tested by formulating the null and alternative hypotheses as follows

$$H_0: v_{lsj} \equiv v_{ls} \text{ for } l = 1, ..., L; \ s = 1, ..., S; \ j = 1, ..., J$$
$$H_1: \text{Not all equalities in the } H_0 \text{ hold}$$

(Equation 9)

under each of which, the likelihoods are used to calculate a log-likelihood ratio which is compared against a critical threshold. If the null hypothesis above is rejected, a significant QTL is claimed. By reshuffling the phenotypic data for 1000 times, the critical threshold can be determined from permutation tests. At each time, an LR value was calculated and the top 5% of LR values among 1000 times was then chosen as the critical threshold at the 5% significance level.

*Pleiotropic effect of a QTL.* If a significant QTL exists, then we can test whether this QTL has a pleiotropic effect for different traits. This test is formulated as

$$H_0: v_{lsj} \equiv v_{sj} \text{ for } l = 1, ..., L; \ s = 1, ..., S; \ j = 1, ..., J$$
$$H_1: \text{Not all equalities in the } H_0 \text{ hold}$$

(Equation 10)

If the null hypothesis is rejected, this implies that the significant QTL detected affects pleiotropically longitudinal traits.

*QTL by environment interaction.* Further, we can test whether the QTL effect on the longitudinal trait is due to QTL × Environment interaction. This can be tested for different cues using

$$H_0: v_{lsj} \equiv v_{lj} \text{ for } l = 1, ..., L; \ s = 1, ..., S; \ j = 1, ..., J$$
$$H_1: \text{Not all equalities in the } H_0 \text{ hold}$$

(Equation 11)

The rejection of the above null hypothesis means that a significant genotype by environment interaction exists due to allelic sensitivity to a varying environment.

The LR for these hypothesis tests (Equations 10 and 11) are thought of as being chi-square distributed with the degrees of freedom equal to the difference in the number of parameters to be estimated under the alternative and null hypotheses. *P*-value was calculated by the chi-square distribution function for each LR, in which less than 0.05 was considered statistically significant. When these tests (Equations 10 and 11) were undertaken, bonferroni correction was used to control false positives.

*QTL network reconstruction.* For a given QTL of quantitative trait, its time-varying genetic variance is expressed as:

$$V(t) = 2p_1 p_0 \left(a + (p_1 - p_0)d\right)^2 + 4p_1^2 p_0^2 d^2 \qquad \text{(Equation 12)}$$

where $p_1$ and $p_0$ are allele frequency for two alleles, respectively; a and d are the estimate of the additive and dominant effect for QTL, respectively. Here, we implement a procedure to reconstruct a genetic network among all significant QTLs based on their genetic variance. Let $s_j(t) = \sqrt{V_j(t)}$ denote the genetic standard deviation of SNP *j* (*j=1, ..., J*) at time *t*. We construct a *J*-dimensional ordinary differential equations:

$$\frac{ds_j(t)}{dt} = W_j\left(g_j(t) : \Theta_j\right) + \sum_{j'=1, j' \neq j}^{J} W_{jj'}\left(g_{j'}(t) : \Theta_{jj'}\right) \qquad \text{(Equation 13)}$$

where $W_j(g_j(t) : \Theta_j)$ describes the main effect of this QTL that is expressed independently from the effects of any other QTLs, and $\sum_{j'=1, j' \neq j}^{J} W_{jj'}(g_{j'}(t) : \Theta_{jj'})$ reflects the interactive effects of different QTLs on QTL j. $W_j(g_j(t) : \Theta_j)$ and $\sum_{j'=1, j' \neq j}^{J} W_{jj'}(g_{j'}(t) : \Theta_{jj'})$ are fitted by the nonparametric function such as B-spline or LOP. By estimate a set of ODE parameters $\Theta_j$, we can determine the pattern and magnitude of the independent effect of QTL. Meanwhile, the estimation of another set of ODE parameters $\Theta_{jj'}$ enables us to characterize whether and how the effect of a QTL depends jointly on all other QTLs. Networks were visualized by using the package *igraph* in R software (Csardi and Nepusz, 2006).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### QTL analysis with existing methods

The performance of FPP-FM in detecting QTL from multiple-treatment longitudinal data was compared with those of the Bayesian method and one-way ANOVA. We followed the Bayesian group lasso method with associated MCMC algorithms as developed by Li et al. (2015). ANOVA is a common (and "static") method used to detect associations between molecular markers and traits in IL populations. The same set of 2,521 SNPs markers (Sim et al., 2012) was used during analysis.

### Simulation analysis

We conducted a computer simulation to validate the statistical properties of FPP-FM based on the working example used in this study. Data were simulated by assuming that traits are either correlated (using FPP-FM) or uncorrelated (using ANOVA, which is suitable for IL population) at different time points. The phenotype was determined by a set of QTLs among 1,000 simulated markers, plus a residual error following a multivariate normal distribution. We adjusted the innovation variance to obtain curve heritability levels of 0.05 and 0.10. The simulation considered two sample sizes: 500 and 200. Data were analyzed reciprocally using ANOVA and FPP-FM. For each simulated QTL, the estimate for the QTL parameter was the average of the corresponding estimates from the counted samples. The number of QTLs detected in every 1,000 simulated datasets represents the detection power of the QTL.

### DATA AND SOFTWARE PACKAGE

The phenotypic and genotypic data of 62 introgression line (IL) population of tomato are available in Table S1. The genetic mapping data of E using Bayesian (Figure S1) and ANOVA (Tables S2 and S3, Figures S2 and S3) methods are available as supplementary files. We have coded all statistical algorithms that build our framework into a user-friendly R package for public use (https://github.com/FFP-FM/Version1).