

Trade-Off Evaluator enables dissecting the genetic architecture of trade-offs between two traits

Shang Liu

University of Potsdam

Amna Sarmad

University of Helsinki

Roosa Anna Emilia Laitinen

University of Helsinki

Zoran Nikoloski

`zoran.nikoloski@uni-potsdam.de`

University of Potsdam

Mikael Brosche

University of Helsinki

Short Report

Keywords: Trade-off, GWAS, Pareto frontier, protein content, growth, Arabidopsis

Posted Date: October 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10999364/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Trade-Off Evaluator enables dissecting the genetic architecture of trade-offs between two traits

Shang Liu¹, Amna Sarmad³, Mikael Brosché^{3,4}, Roosa A. E. Laitinen^{3*}, Zoran Nikoloski^{1,2*}

¹Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam, Germany

²Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

³Department of Organismal and Evolutionary Biology, Viikki Plant Science Centre, University of Helsinki, 00014, Helsinki, Finland

⁴ Department of Molecular and Integrative Biosciences, Viikki Plant Science Centre, University of Helsinki, 00014, Helsinki, Finland

Official email address:

Zoran Nikoloski: zoran.nikoloski@uni-potsdam.de

Shang Liu: shang.liu@uni-potsdam.de

Roosa Anna Emilia Laitinen: roosa.laitinen@helsinki.fi

Amna Sarmad: amna.sarmad@helsinki.fi

Mikael Brosche: mikael.brosche@helsinki.fi

Full contact details of the corresponding authors: zoran.nikoloski@uni-potsdam.de; roosa.laitinen@helsinki.fi

Abstract

Biological trade-offs between quantitative traits are pervasive in plants. Yet, there is no framework to dissect their genetic basis due to the lack of measures to quantify trade-offs in individuals. Here, we developed the Trade-Off Evaluator, a framework based on Pareto Frontier that quantifies trade-off between two traits at the level of an individual. To validate the Trade-Off Evaluator, we employed the established fresh weight–protein content trade-off in 326 *Arabidopsis thaliana* accessions. Using cross-validation GWAS with the fresh weight–protein content trade-off scores, resulting from the Trade-Off Evaluator, we identified a locus harboring a cluster of seven membrane bound O-acyl transferase (*MBOAT*) encoding genes. We demonstrated that only the disruption of *MBOAT4* (AT5G55350) showed higher fresh weight compared to the wild type at no penalty to protein content, thus, breaking the fresh weight–protein content trade-off. Furthermore, genomic prediction with the trade-off score demonstrated that the ability to decouple traits in trade-offs can be predicted from genetic markers. Collectively, the Trade-off Evaluator provides a framework for trade-off quantification that enables genetic dissection, with implications to multi-trait crop improvement.

Keywords: Trade-off, GWAS, Pareto frontier, protein content, growth, *Arabidopsis*

Background

Trade-offs between quantitative traits are pervasive in plants (e.g., yield–drought tolerance, biotic stress–abiotic stress resistance, and growth–immunity) as limited resources cannot be simultaneously used to fulfill multiple biological tasks^{1–3}. Trade-offs constrain the simultaneous improvement of multiple agronomic traits and represent a major challenge for breeding. Despite the relevance of trade-offs in plant growth and development as well as their role in shaping evolutionary outcomes, genetic loci underlying trade-offs remain largely elusive^{4–7}.

Trade-offs between two traits have been quantified by negative phenotypic or genetic correlations at a population level⁸. This approach does not allow us to quantify trade-offs in an individual and, therefore, precludes application of quantitative genetics approaches to dissect their genetic architecture. Another approach based on antagonistic pleiotropy identifies loci with contrasting effects on two traits by aggregating their genetic architectures⁹. This strategy does not allow the quantification of the strength of trade-off and variance explained by the putative trade-off loci. A framework that quantifies trade-offs at

the individual level would resolve these issues and would allow genetic mapping of trade-offs.

The Pareto frontier defines the theoretical performance boundary for two negatively correlated traits¹⁰. Individuals located on the Pareto frontier have reached the optimal balance between two competing traits, since an increase of one trait inevitably requires decrease of the other. By contrast, individuals away from the Pareto frontier have the potential to break the trade-off. Here, we developed the Trade-Off Evaluator (TOE) that quantifies the extent of a trade-off between two traits in an individual by the distance of the individual to the Pareto frontier. Using the fresh weight–protein content trade-off in *Arabidopsis* as a proof-of-concept, we demonstrated that TOE enables both identification of loci underlying a trade-off and genomic prediction of the extent of trade-off in a given genotype¹¹.

Results and Discussion

TOE is a computational framework that scores a trade-off as a quantitative trait amenable for quantitative genetics analyses. To illustrate the design of TOE, we focus on two negatively correlated traits, fresh weight and protein content, in a panel of 326 *Arabidopsis thaliana* accessions¹¹ (Pearson correlation coefficient, $r=-0.65$, $p=2\times 10^{-41}$, Fig. 1A). After the min-max scaling on fresh weight and protein content, a convex hull was identified from the panel of accessions, and the vertices of the convex hull above the $y = -x$ line were selected (Fig. 1A). These selected vertices represent accessions for which an increase in fresh weight inevitably requires decrease in protein content and *vice versa*. The Pareto frontier, comprising a subset of accessions, was then fitted using a quadratic function to define the performance boundary for fresh weight and protein content for the 326 accessions (Fig. 1A). We defined the trade-off between fresh weight and protein content for an accession as its distance to the fitted Pareto frontier (Fig. 1A).

With the fresh weight-protein content trade-off score, we performed five-folds cross validation-GWAS (CV-GWAS) based on over 200,000 single nucleotide polymorphisms (SNPs) to detect the stable genetic architecture of this trade-off (Fig. 1B, Methods). Among all 1,917 SNPs significantly associated with this trade-off ($\text{LOD}\geq 3$), only the SNP m206146 (Chr5:22445118) was detected across the five batches of the five-folds CV-GWAS (Fig. 1C). According to m206146, the 326 accessions were divided into two haplotypes (hap1, trade-off constrained; hap2, trade-off breaking). Whilst we did not observe divergence in fresh weight between the two haplotypes, there was a slight difference in protein content ($p\text{-value} = 0.02$, indicating that the effect of identified trade-off locus on individual trait means is negligible

(Fig. 1D). Seven clustered genes encoding for membrane bound O-acyl transferase (MBOAT1-7) proteins were located around the strongest SNP (Fig. 1E), and three of them (i.e., *MBOAT4*, *MBOAT6*, and *MBOAT7*) were expressed in roots according to a previous reported single-cell transcriptomic atlas covering root, stem, and rosette (Additional file1: Fig. S1A)^{12,13}.

To experimentally validate whether these three candidate *MBOATs* break the fresh weight-protein content trade-off, we performed three rounds of sequential screening of the candidate T-DNA insertion lines (Materials and Methods, Additional file 1: Fig. S1). Based on the first two sequential validations, *MBOAT4* was identified to break the fresh weight-protein content trade-off by increasing fresh weight without affecting the protein content (Additional file1: Fig. S1B and S1C). A further validation of this line in an independent soil-grown experiment, the disruption of *MBOAT4* again led to an increase in fresh weight without compromising protein content (Fig. 1F; Additional file 1: Fig. S1D), identifying *MBOAT4* as a gene underlying the fresh weight-protein content trade-off.

The trade-off score between fresh weight and protein content is suitable for use not only in GWAS, but also genomic prediction. This enables the prediction of lines in which traits in trade-offs are decoupled. To this end, we inspected the predicted accuracies of fresh weight, protein content, and their trade-off, assessed by the Pearson correlation coefficient (PCC), between measured and predicted values with the cross-validation splits from the CV-GWAS. The mean accuracy of the genomic prediction model for the fresh weight-protein content trade-off was 0.22, while that of fresh weight and protein content was 0.32 and 0.15, respectively (Fig. 1G). These results suggest that trade-offs can be directly incorporated into downstream breeding applications.

Existing strategies identify trade-off genes indirectly by analysing traits in trade-off separately. TOE overcomes this limitation by directly quantifying trade-off as a genetically tractable trait. TOE can be extended to diverse trade-offs among multiple biological layers, providing a general framework for studying resource allocation and facilitating multi-trait molecular breeding.

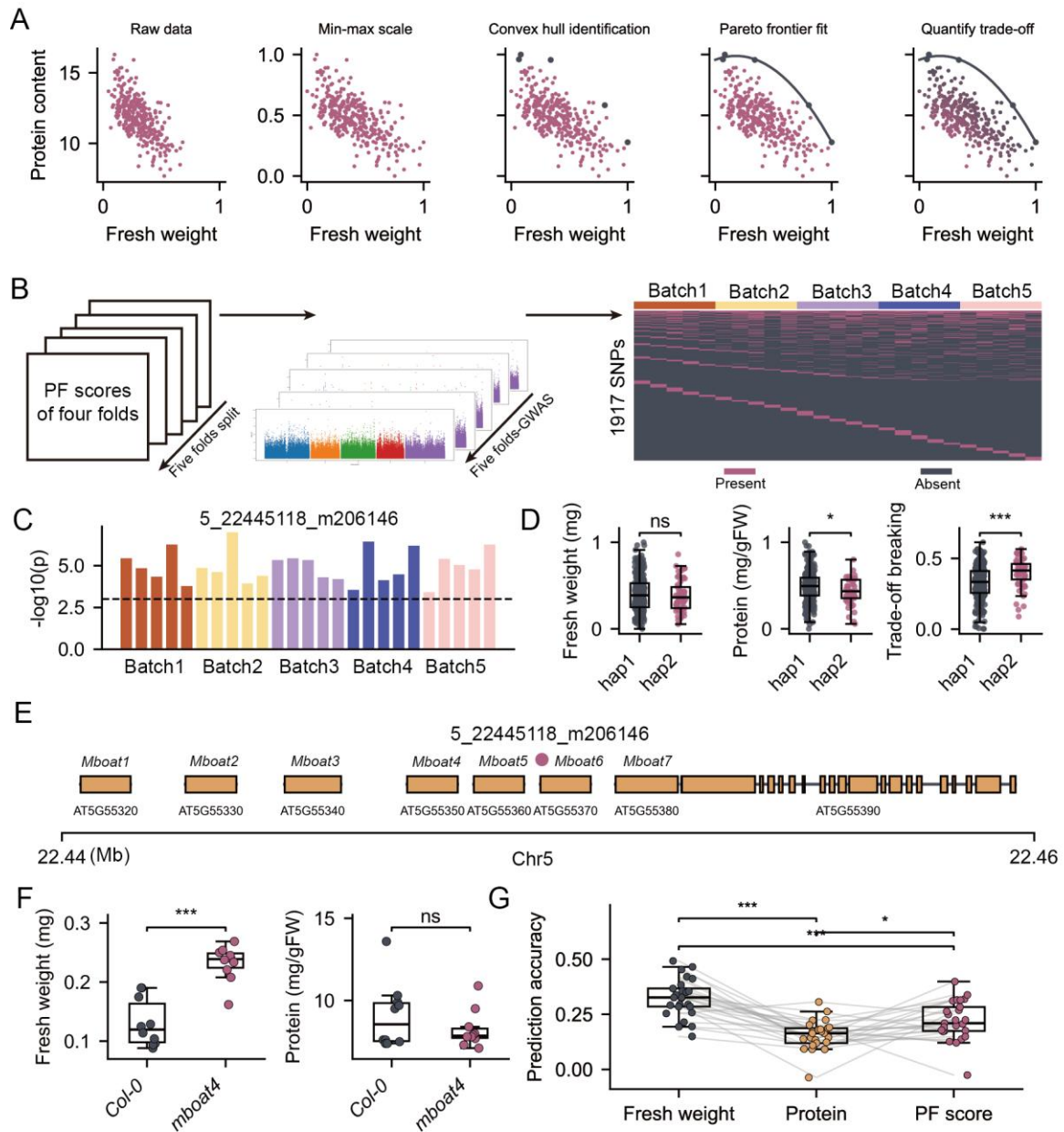


Figure1. Genetic mapping on trade-off between fresh weight and protein content in *Arabidopsis thaliana*. **A.** Development of the Trade-Off Evaluator (TOE) for quantification of trade-off between two traits in a population of genotypes. With fresh weight (mg) and protein content (mg/gFW) for 326 *Arabidopsis thaliana* accessions, TOE comprises four steps: (1) the two traits are transformed to [0, 1] range by min-max scaling, (2) a convex hull for all individuals is determined, and the vertices of the convex hull above the $y = -x$ line, marked in black, are selected to yield the Pareto frontier of the genotypes for the two traits, (3) a function is then fitted to the points at the Pareto frontier, (4) the distance of each point, representing a genotype, to the Pareto frontier is calculated as a quantified trade-off breaking ability used in association analysis. **B.** Workflow of 5-fold cross validation GWAS (CV-GWAS) on the quantified trade-off breaking ability for 326 *Arabidopsis thaliana* accessions. The 326 accessions were split into five folds of which four were used for GWAS; five GWAS results were generated from the 5-folds CV-GWAS, and loci associated with

phenotype ($p \leq 0.001$) in each GWAS result were retained. The present/absent map of significant loci in each CV-GWAS were constructed for target loci selection. The x-axis represents five batches of five-folds CV-GWAS, marked in different colors. The y-axis represents 1,917 SNPs associated with trade-off breaking among five batches of the CV-GWAS. **C.** The log-transformed p-values of 5_22445118_m206146 among 25 CV-GWAS results. **D.** Divergence of fresh weight (mg), protein content (mg/gFW), and quantified trade-off breaking ability between two haplotypes classified by 5_22445118_m206146. Two-side t-test (***, $p \leq 0.001$; ns, $p > 0.05$). **E.** Seven *MBOAT* genes within the 10 kb upstream and downstream region of 5_22445118_m206146. **F.** Fresh weight (mg) and protein content (mg/gFW) of soil-grown Col-0 and T-DNA mutant lines of *mboat4*. Two-side t-test (***, $p \leq 0.001$; ns, $p > 0.05$). **G.** Accuracies of genomic predictions on fresh weight, protein content and fresh weight-protein trade-off breaking. Pair t-test (***, $p \leq 0.001$; *, $p \leq 0.05$)

Conclusions

We developed TOE as a general framework to quantify trade-offs between two traits at the level of an individual relying on the Pareto frontier concept. Using the fresh weight–protein content trade-off in *A. thaliana* as a proof-of-concept, we demonstrated the utility of TOE in trade-off mapping and genomic prediction. TOE is broadly applicable to trade-offs within multiple biological layers and provides a practical strategy for genetic dissection on resource allocation.

Method

The phenotype data and SNP markers of 326 *Arabidopsis thaliana* accessions

The fresh weight, protein content, and genetic markers of 326 *A. thaliana* accessions were downloaded from a previous publication¹¹.

Development of the Trade-Off Evaluator (TOE) based on Pareto-Frontier

The TOE was developed by Python code, and the script as well as sample data used in this study are available at <https://github.com/LiuShang-777/TOE>

Five-folds cross-validation genome wide association study (GWAS) and gene association

To identify robust genetic loci associated with fresh weight-protein content trade-off, we performed a five-folds cross-validation GWAS. Phenotype and genetic markers of 326

accessions were randomly and equally divided into five subsets. In each round, we excluded one subset and used the remained four subsets to perform GWAS by EMMAX¹⁴. This leave-one-fold-out GWAS was repeated five times to get 25 GWAS results. Loci detected across five repeated five-fold GWAS ($\text{LOD} \geq 3$) were considered to be of high confidence. To associate significant loci identified from five-folds CV-GWAS, we annotated the significant fresh weight-protein content trade-off (FPT) locus by annovar.

The phenotype validation of T-DNA mutant lines for *MBOATs*

Seven candidate genes for the fresh weight-protein trade-off were investigated using *Arabidopsis* T-DNA insertion mutants in three independent trials. In the first trial, a total of 16 T-DNA insertion lines for seven candidate genes (Additional file 2: Supplementary Table 1) were grown and measured for fresh weight and protein content. All lines were grown on 0.5xMS plates, which were first stratified at dark at 4 °C for five days whereafter they were moved to LD growth conditions cabinet (16 h light/8 h dark, 21°C/18°C, $\sim 120 \mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity). After 15-17 days, plates were imaged and fresh weight (g) of 5 pooled plants was measured. Total protein content (mg/gFW) Total protein content (mg/gFW) was determined from the pellet remaining after ethanol extraction. The pellet was resuspended in 0.1 M NaOH following a procedure adapted from previous research¹⁵, and protein concentration was determined using the Bradford assay.

To verify the function of *At5g55350* encoding for *MBOAT4* protein, underlying the fresh weight- protein content trade-off, two additional trials were conducted (Additional file 2: Supplementary Table 2). The second trial was done like the first, with four biological replicates, each consisting of five pooled plants. The third trial was performed by growing plants on 1:1 peat and vermiculite mix. Prior to sowing, seeds were stratified in 0.1 % agarose at 4°C for 5 days before grown under 12 h light/12 h dark ($\sim 240 \mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity, 23°C/19°C, 50–60% humidity) (Additional file 3: Supplementary Table S3). After 14 days, fresh weight and protein content was measured from 10 replicate plants as described above.

To confirm homozygosity of T-DNA insertion line for *At5g55350*, genomic DNA was extracted from leaves by gene-specific primers (LP and RP) alongside a T-DNA left-border primer (LB) (Additional file2: Supplementary Table 4)¹⁶. To confirm the loss of expression of mutant lines three replicates of leaf and root samples were ground and RNA was isolated using TRIzol (Invitrogen, 15596026) according to the manufacturer's guidelines. The extracted nucleic acid was treated with DNaseI (Thermo fisher, 18047019), and 1 μg of RNA was used for cDNA synthesis using RevM Revertase (Klever lab, E-RT) according to the manufacturer's guidelines. Real-time qRT-PCR reactions were performed using 2X Fast Dye

qPCR Mastermix, (Klever labs, M2-FD), on a Bio-Rad Opus CFX-384. YLS8, SAND and TIP41 were used as housekeeping genes for normalisation, considering a standard deviation of ≤ 0.5 Ct as quality control (primer list at the end of the document). *MBOAT* expression levels were calculated using the $2^{-\Delta\Delta C_T}$ method¹⁷.

Genomic prediction of fresh weight, protein content and their trade-off

With the data splits following CV-GWAS, we performed genomic prediction on fresh weight, protein content and trade-off using the genomic best linear unbiased prediction (GBLUP)¹⁸. The genetic markers used for GWAS were also applied to calculate genomic relationship matrix (GRM). Genomic prediction was conducted using the following mixed linear model:

$$y = 1\mu + Zu + e,$$

where y denotes the phenotype, μ is the overall mean, Z is the random effect incidence matrix, u is additive genetic effect, and e is the residual error. The random genetic effects and residual error were assumed to follow: $u \sim N(0, K\sigma_u^2)$ and $e \sim N(0, I\sigma_e^2)$.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Availability of data and materials

All data during this study are included in this published article Fusari, C. M. *et al.* Genome-Wide Association Mapping Reveals That Specific and Pleiotropic Regulatory Mechanisms Fine-Tune Central Metabolism and Growth in Arabidopsis. *Plant Cell* **29**, 2349–2373 (2017).

Funding

Part of this research was funded by the Horizon Europe research and innovation programme, project BOLERO (breeding for coffee and cocoa root resilience in low-input farming systems based on improved rootstock), HORIZON-CL6-2021-BIODIV-01-13, under grant agreement no. 101060393 (to Z.N.).

Authors' contributions

S.L., Z.N., M.B. and R.L. designed and supervised the study. A.S. performed wet-lab experiments. S.L. performed the data analysis and prepared the manuscript. Z.N., R.L., A.S. and M.B. revised manuscript. All authors have read and approved the manuscript.

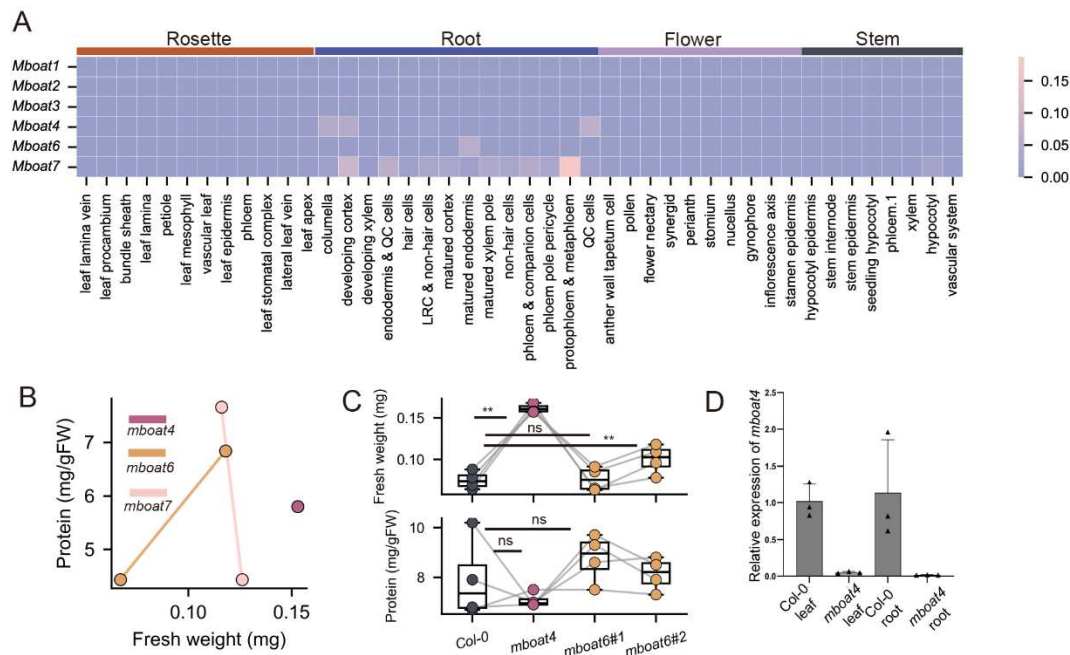
Acknowledgements

Hirofumi Ishihara is thanked for his guidance in the protein measurements.

References:

1. Sun, X. *et al.* The role of transposon inverted repeats in balancing drought tolerance and yield-related traits in maize. *Nature Biotechnology* **41**, 120–127 (2023).
2. Gu, X. *et al.* The OsSGS3-tasiRNA-OsARF3 module orchestrates abiotic-biotic stress response trade-off in rice. *Nature Communications* **14**, 4441 (2023).
3. Qi, Y. *et al.* Chloroplast elongation factors break the growth–immunity trade-off by simultaneously promoting yield and defence. *Nature Plants* **10**, 1576–1591 (2024).
4. Liu, S. *et al.* Cotton transposon-related variome reveals roles of transposon-related variations in modern cotton cultivation. *Journal of Advanced Research* <https://doi.org/10.1016/j.jare.2024.05.019> (2024) doi:10.1016/j.jare.2024.05.019.
5. Liu, S. *et al.* Cotton pedigree genome reveals restriction of cultivar-driven strategy in cotton breeding. *Genome Biol* **24**, 282 (2023).
6. Liu, Y. *et al.* Pan-Genome of Wild and Cultivated Soybeans. *Cell* **182**, 162-176.e13 (2020).
7. Guo, D. *et al.* A pangenome reference of wild and cultivated rice. *Nature* <https://doi.org/10.1038/s41586-025-08883-6> (2025) doi:10.1038/s41586-025-08883-6.
8. Laitinen, R. A. E. & Nikoloski, Z. Strategies to identify and dissect trade-offs in plants. *Molecular Ecology* **33**, e16780 (2024).
9. Brigos-Barril, E. *et al.* Genetic trade-offs in fertility and longevity explain the maintenance of disease-associated alleles in humans. *Nat Ecol Evol* **10**, 1774–1785 (2026).
10. Kleessen, S. *et al.* Metabolic efficiency underpins performance trade-offs in growth of *Arabidopsis thaliana*. *Nature Communications* **5**, 3537 (2014).

11. Fusari, C. M. *et al.* Genome-Wide Association Mapping Reveals That Specific and Pleiotropic Regulatory Mechanisms Fine-Tune Central Metabolism and Growth in Arabidopsis. *Plant Cell* **29**, 2349–2373 (2017).
12. Lee, T. A. *et al.* A single-cell, spatial transcriptomic atlas of the Arabidopsis life cycle. *Nat Plants* **11**, 1960–1975 (2025).
13. Shahan, R. *et al.* A single-cell Arabidopsis root atlas reveals developmental trajectories in wild-type and cell identity mutants. *Dev Cell* **57**, 543–560.e9 (2022).
14. Kang, H. M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* **42**, 348–354 (2010).
15. Hendriks, J. H. M., Kolbe, A., Gibon, Y., Stitt, M. & Geigenberger, P. ADP-Glucose Pyrophosphorylase Is Activated by Posttranslational Redox-Modification in Response to Light and to Sugars in Leaves of Arabidopsis and Other Plant Species. *Plant Physiology* **133**, 838–849 (2003).
16. Edwards, K., Johnstone, C. & Thompson, C. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucl Acids Res* **19**, 1349–1349 (1991).
17. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–408 (2001).
18. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome* **4**, 250–255 (2011).



Supplementary Figure. S1 Expression pattern of Mboats and phenotypic validation on the fresh weight-protein content trade-off by T-DNA insertion lines.

A. The single cell transcriptomic map of six MBOAT genes (normalized read counts) in rosette (GSE226097), root (GSM4626009), flower (GSE226097) and stem (GSE226097) of Col-0. Expression level was exhibited as the normalized read count after removing batch effects. **B.** The fresh weight and protein content of agar-grown T-DNA mutant lines, *mboat6* and *mboat7*. Each dot represents an independent T-DNA mutant line. **C.** Fresh weight (mg) and protein content (mg/gFW) of agar-grown Col-0, one T-DNA mutant lines of *Mboat4*, and two T-DNA mutant lines of *Mboat6*. Two-side t-test (**, $p \leq 0.01$; ns, $p > 0.05$). **D.** Expression level of *MBOAT4* in Col-0 and T-DNA insertion lines of *mboat4*, in both leaf and root.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ST.xlsx](#)