
MR. HAITAO LI (Orcid ID : 0000-0001-5757-6375)

DR. LIZHONG XIONG (Orcid ID : 0000-0003-0490-1474)

PROF. KEDE LIU (Orcid ID : 0000-0002-1395-2049)

MR. WANNENG YANG (Orcid ID : 0000-0003-1095-1355)

Article type : Research Article

High-throughput phenotyping accelerates the dissection of the dynamic genetic architecture of plant growth and yield improvement in rapeseed

Haitao Li^{a,b,d}, Hui Feng^{a,d}, Chaocheng Guo^a, Shanjing Yang^a, Wan Huang^a, Xiong Xiong^c, Jianxiao Liu^a, Guoxing Chen^a, Qian Liu^c, Lizhong Xiong^a, Kede Liu^a, Wanneng Yang^{a,*}

^a National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research, and Hubei Key Laboratory of Agricultural Bioinformatics, Huazhong Agricultural University, Wuhan 430070, P. R. China.

^b State Key Laboratory of Biocatalysis and Enzyme Engineering, and Hubei Collaborative Innovation Center for Green Transformation of Bio-resources, School of Life Sciences, Hubei University, Wuhan 430062, P. R. China.

^c Britton Chance Center for Biomedical Photonics, Wuhan National Laboratory for Optoelectronics, and Key Laboratory of Ministry of Education for Biomedical Photonics, Department of Biomedical Engineering, Huazhong University of Science and Technology, Wuhan 430074, P. R. China.

^d These authors contributed equally to this work.

*Corresponding author:

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/PBI.13396](https://doi.org/10.1111/PBI.13396)

This article is protected by copyright. All rights reserved

Wanneng Yang, National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research, Huazhong Agricultural University, Wuhan 430070, P. R. China. Email: ywn@mail.hzau.edu.cn. Fax: +86-27-87287092. Tel: +86-27- 87282130)

Keywords: High-throughput phenotyping, i-trait, quantitative trait loci, dynamic genetic architecture, yield, rapeseed

Summary

Rapeseed is the second most important oil crop species and is widely cultivated worldwide. However, overcoming the “phenotyping bottleneck” has remained a significant challenge. A clear goal of high-throughput phenotyping is to bridge the gap between genomics and phenomics. In addition, it is important to explore the dynamic genetic architecture underlying rapeseed plant growth and its contribution to final yield. In this work, a high-throughput phenotyping facility was used to dynamically screen a rapeseed intervarietal substitution line population during two growing seasons. We developed an automatic image analysis pipeline to quantify 43 dynamic traits across multiple developmental stages, with 12 time points. The time-resolved i-traits could be extracted to reflect shoot growth and predict the final yield of rapeseed. Broad phenotypic variation and high heritability were observed for these i-traits across all developmental stages. A total of 337 and 599 QTLs were identified, with 33.5% and 36.1% consistent QTLs for each trait across all 12 time points in the two growing seasons, respectively. Moreover, the QTLs responsible for yield indicators colocalized with those of final yield, potentially providing a new mechanism of yield regulation. Our results indicate that high-throughput phenotyping can provide novel insights into the dynamic genetic architecture of rapeseed growth and final yield, which would be useful for future genetic improvements in rapeseed.

Introduction

Rapeseed (*Brassica napus*; canola) is a relatively recent allopolyploid species formed by interspecific hybridization between *Brassica rapa* and *Brassica oleracea* ~7500 years ago (Chalhoub et al., 2014). Currently, rapeseed is the second most important oil crop species and is widely cultivated worldwide due to several characteristics, such as its ability to be used as a forage

for cattle, as an oil plant for nutritional and industrial purposes and as a protein crop for producing animal feed (Friedt et al., 2018). With the increase in the global population and living standards, the demand for vegetable oils and proteins is expected to increase by more than 40% until 2030 (Delourme et al., 2018). Therefore, it is very important to study agronomic traits, including yield-related traits and seed oil content and quality, which could provide new foundations for breeding elite cultivars.

Quantitative trait loci (QTLs) responsible for complex traits in crop species are usually identified by linkage or association analysis using appropriate computational methods based on the genotype and phenotype of each individual in a mapping population. Recent advances in crop genomic sequencing and high-throughput genotyping methods have greatly facilitated the genotyping of numerous lines for detecting the genetic basis of phenotypic variations (Huang and Han, 2014). Twelve genomes of rapeseed, encompassing all three ecotypes, are available, providing useful resources for identifying genetic variations (Bayer et al., 2017; Chalhoub et al., 2014; Song et al., 2020; Sun et al., 2017; Zou et al., 2019). High-throughput genotyping approaches, including single nucleotide polymorphism (SNP) Illumina Infinium array, genotyping by sequencing (GBS) and resequencing-based genotyping, have been established in rapeseed (Chen et al., 2013; Clarke et al., 2016; Wang et al., 2018). Based on these resources, substantial amounts of SNPs, insertions/deletions (InDels) and presence and absence variations (PAVs) have been identified in various permanent populations, such as recombinant inbred lines (RILs), double haploids (DHs), intervarietal substitution lines (ISLs), and in nested association mapping (NAM) and natural populations (Chen et al., 2013; Hu et al., 2018; Luo et al., 2015; Song et al., 2020; Yang et al., 2018; Yang et al., 2016). In contrast, the development of phenotyping technology is very slow in rapeseed and is still dominated by traditional manual scoring. The investigated traits in rapeseed are usually based on general breeding targets and include seed germination, seedling root development, flowering time, yield-related traits, aerial plant architecture, pod shatter, seed oil content, quality and tolerance to biotic/abiotic stress (Delourme et al., 2018). These traits are evaluated manually at a single stage, which is often time-consuming and destructive. Fortunately, recently developed high-throughput phenotyping methods can potentially relieve this bottleneck.

Combined with genome-wide association analyses, high-throughput phenotyping has been proven to be an effective tool to dissect the genetic basis of early seedling growth in rapeseed (Knoch et al., 2020). However, the temporal dynamic genetic architecture of plant growth throughout development and its contribution to yield need to be further investigated. In a previous study, we developed a high-throughput rice phenotyping facility (HRPF) to phenotype large populations of rice plants (Yang et al., 2014). In this study, the HRPF (Figure S1) was used to study the dynamic architecture of plant growth within a rapeseed intervarietal substitution line (ISL) population at 12 time points of multiple stages throughout the growth phase. We developed a robust image analysis pipeline for rapeseed and quantified 43 image-based traits (i-traits) and 30 growth-related traits. It was found that a combination of several i-traits measured in the multiple growth stages could be used as good predictors of the final yield. Interestingly, the QTLs responsible for yield indicators colocalized with those of final yield, potentially providing a new mechanism of yield regulation.

Results and Discussion

Biomass estimation and performance evaluation of the phenotyping platform

To evaluate the measuring accuracy, an additional 120 rapeseed plants (recurrent parent ZY821 and 9 ISLs, 12 replicates per accession) were grown and tested. For each inspection (T1-T12, shown in Table S1), 10 plants were automatically (using the HRPF) and manually measured for their height, fresh weight, and dry weight. To predict the fresh weight and dry weight, 11 models (including linear, quadratic, exponential, and power models) were compared using total projected area in side (TPA_SV) and top view (TPA_TV). Compared with the other models, the power model ($\ln(FW) = a + b \times \ln(m)$) had a higher R^2 , and lower mean absolute percentage error (MAPE) and standard deviation of the absolute percentage error (SD_{APE}), for both dry weight and fresh weight in the two growing seasons (Tables S2 and S3), which was also validated by a 10-fold cross-validation approach (Tables S4 and S5). These results also suggested that the i-trait projected area exhibited good correlation with manually measured rapeseed shoot biomass. Similar results were observed in our previous studies on rice (Yang et al., 2014) and maize (Zhang

et al., 2017) and have also been proven by other groups (Chen et al., 2014; Knoch et al., 2020; Muraya et al., 2017). Scatterplots showed that the R^2 value between the manual and automatic measurements of these three traits was greater than 0.85 ($P < 0.001$) in both growing seasons (Figure S2), which suggests a high accuracy of our phenotyping platform for rapeseed.

Yield prediction using i-traits at multiple growth stages

It would be beneficial for crop breeding if we could use simple i-traits at the multiple growth stages, especially the early development stages, to predict the final rapeseed yield. By the use of the extracted i-traits of all investigated lines, the variance explained for yield with 43 traits at 12 growth stages was evaluated using linear stepwise regression, and the results showed that up to 75.5% of the phenotypic variance of yield could be explained by combining 10 i-traits across the 12 time points (the statistical details of the coefficients are shown in Table S6). If only 8 i-traits at 3 time points (the 1st, 7th, and 12th time point, which reflect the early seedling stage, early bolting stage, and flowering stage, respectively) were used, 68.2% of the phenotypic variance of the yield could be explained (Figure 1a), which indicated that the i-traits at these three growth stages could be used as early predictors of rapeseed yield. The statistical details of the 8 coefficients are shown in Table S7. The prediction results were also validated using a 5-fold cross-validation approach, which is listed in Table S8.

For example, the first selected i-trait of the model was FDNIC_TV_7 (fractal dimension without image cropping in top view at the 7th time point), which reflected the top-view projected shoot area and leaf shape and was positively correlated with yield ($R^2 = 0.369$, $p < 0.001$, Figure 1b). Interestingly, after comparing the correlation between FDNIC_TV and yield across 12 time points, we found that FDNIC_TV at the 7th time point (FDNIC_TV_7) had the best correlation with yield (Figure 1b). On the basis of the plant height change across 12 time points and time series images (Figure 1c and 1d), it was confirmed that the 7th time point was during the early bolting stage (DAS 95), which indicated that FDNIC_TV in the early bolting stage could be used as a good predictor of rapeseed yield. In the future, in conjunction with the use of unmanned aerial vehicles and other phenotyping tools in the field, the image analysis pipeline and these extracted

i-traits could be potentially adapted for field-based phenotyping.

Digital biomass accumulation modeling and extraction of growth-related trait

It would be interesting for rapeseed breeding if we could dissect the growth model and derive growth-related traits. The growth model could be used to predict future biomass using the biomass data in early stages, and the growth-related traits could be used to identify rapeseed populations that display rapid growth. Based on the shoot digital biomass (TPA_SV and TPA_TV) at the 12 time points, we compared 6 growth models (the linear, quadratic, exponential, power, logarithmic, and sinusoidal models) and generated 30 growth-related traits. By comparing the R^2 value, mean absolute percentage error (MAPE), and standard deviation of the absolute percentage error (SD_{APE}) for TPA_SV accumulation, five models (the logarithmic model was excluded) showed good prediction ability (the R^2 was greater than 0.93, and both the MAPE and SD_{APE} were below 19%). The digital biomass accumulation prediction of the exponential model is shown in Figure 2a, and the exponential formula is described as follows:

$$TPA_SV_t = a_Exp_SV \times Exp(b_Exp_SV \times t) \quad t = 1, \dots, 12 \quad (1)$$

In formula 1, TPA_SV is the total projected area in the side view, t is the number of time points, and a_Exp_SV reflects the biomass and growth speed and was positively correlated with yield (Figure 2b), which indicated that the faster growth of plants in the vegetative growth stage would lead to higher yield. For TPA_TV accumulation, the quadratic and sinusoidal models showed good prediction ability (the R^2 values was greater than 0.93, and both MAPE and SD_{APE} were below 12%). The detailed results of the digital biomass modeling are shown in Table S9. The correlations between yield and the 30 growth-related traits are illustrated in a heatmap (Figure 2c).

Phenotypic variation and heritability

Extensive phenotypic diversity was observed for all 43 i-traits in the ISL population at each time point (Figure 3a; Table S10). The average fold change for all traits was 1.93, ranging from 1.02 to 6.18 among the different time points. Of these, FDIC_TV/SV, FDNIC_TV/SV (reflecting the projected shoot area and leaf shape), and GCV_TV (reflecting the degree of leaf green) had the

lowest range of phenotypic variation, with <1.17-fold differences across all time points (Figure 3a). In contrast, DW, FW, and PC6_SV (reflecting the compactness of the whole plant; Yang et al., 2014) had a wide range of phenotypic variation across all time points, with the fold change of the traits ranging from 2.07 to 6.18 (Figure 3a).

The H^2 of all traits among the 12 time points ranged from 0 to 0.84, with an average of 0.49 (Figure 3b; Table S10). Thirty-one traits showed higher heritability (>0.50) at more than half of the time points, which was exemplified by TPA_SV and FDNIC_TV, which are important i-traits for biomass and yield prediction, respectively. Generally, heritability peaked at the fifth and sixth time points (Figure 3b), which suggested that plant growth was stably inherited and insensitive to environmental factors before the bud stage of rapeseed. However, heritability decreased sharply at the twelfth time point for most traits in top view (Figure 3b), which might be due to an impairment of the i-traits extraction by emerging buds and flowers. On the contrary, such results promote us to develop a specific image analysis pipeline for obtaining buds/flowers-related i-traits in the future. Of all traits, W_SV (reflecting the plant width in the side-view image) showed the highest heritability at almost all time points. Among plant compactness-related traits (PC1-PC6) during the 12 time points, PC6 showed the better heritability and ability to reflect the plant compactness, which had also been suggested for rice in our previous work (Yang et al., 2014).

Dynamic QTLs for 43 i-traits

QTL mapping was performed using ICI mapping software, and the physical intervals of QTLs were determined according to the newly published ZS11 genome sequence. The dynamic QTLs are illustrated in Figure 4. In total, 337 and 599 QTLs associated with rapeseed growth were identified in the 2015-2016 and 2016-2017 growing seasons across the 12 time points, respectively. These QTLs were further integrated into 28 and 29 nonredundant QTLs, respectively, which suggested a high percentage of overlapping QTL intervals for both same and different i-traits across the 12 time points and may be due to the high correlations between paired traits (Figure S3). The number of QTLs for each trait ranged from one to eight and one to nine across all 12 time points, with a mean of 1.0-2.8 and 1.0-3.3 QTLs per time point in the two

growing seasons, respectively (Table S11). The genetic effects of all QTLs were further analyzed. The phenotypic variation explained (PVE) by each QTL ranged from 2.4% to 31.1% and from 0.8% to 55.7%, with a mean of 11.1%-20.3% and 11.5%-20.5%, respectively for the 43 traits in the two growing seasons (Table S11). In summary, 31 (72.1%) and 38 (88.4%) i-traits at a unique stage were controlled by a number of QTLs in 2015-2016 and 2016-2017, respectively. QTLs simultaneously affected the same trait at multiple growth stages, ranging from 2 to 9 and from 2 to 8, for 31 and 37 i-traits in 2015-2016 and 2016-2017, respectively.

The mapping resolution, distribution and colocalization of QTLs were further analyzed. The interval length of 82.8% and 62.9% QTLs were smaller than 5 Mb in the two growing seasons, with a mean of 7.0 Mb and 9.7 Mb in 2015-2016 and 2016-2017, respectively (Figure S4). The QTLs were distributed nonrandomly throughout the rapeseed genome, although they were absent on chromosomes C01 and C07 in 2015-2016 and on A01 and C01 in 2016-2017. Interestingly, four QTL hotspots were observed across the rapeseed genome in each of the two seasons, which were distributed on chromosomes A01, A02, A04 and C08 in 2015-2016 and on A02, A03, A06 and A07 in 2016-2017 (Figure 5a), partially due to the high correlations between paired traits. The hotspots on A02 overlapped between the two growing seasons (Figure 5a), which suggests that this region was stable and might display pleiotropy for growth traits in rapeseed development. Moreover, 33.5% of QTLs in 2015-2016 and 36.1% QTLs in 2016-2017 were consistently detected in the two growing seasons for each trait through all 12 time points, respectively.

Two QTLs for PC1_SV-6 and PC3_SV-9 in 2016-2017 overlapped with the QTLs on chromosome A02 for seed number per silique, which was scored from the field in the same population (Yang et al., 2018). In addition, QTLs for PC6_SV-2, AC_SV-1 and R_SV-1 in 2015-2016 and for PAR_TV-11 in 2016-2017 also overlapped with the QTLs on chromosome A09 for seed number per silique, which was also scored from the field (Yang et al., 2018). Notably, of the eight i-traits selected as indicators to predict the final grain yield (Figure 1a), the QTLs for FDNIC_TV-7 and GPA_TV-7 as well as H_SV-12 colocalized with QTLs on chromosomes A07 and A05, respectively, for final yield assessed in the same greenhouse (Figure 5b). These results suggested that the potential candidate genes in these regions may control the

final yield by affecting early growth of rapeseed, potentially providing a new mechanism of yield regulation in addition to the genes associated with the three traditional components of yield. A large segregating population will be developed using the candidate ISL and recurrent parent ZY821 to identify the genes using map-based cloning.

High-throughput phenotyping benefits rapeseed breeding

Traditional phenotyping of rapeseed tends to be destructive, time consuming and labor intensive, and it cannot process large populations and lags far behind rapeseed functional genetics and breeding efforts. Our approach, which involves a high-throughput phenotyping platform and corresponding image analysis software, has the following advantages and yields the following interesting findings:

- (1) We developed a robust image analysis pipeline for rapeseed and documented many dynamic traits that are undetectable using traditional phenotyping methods, such as growth-related traits. With the derived growth models and growth-related traits, rapeseed populations whose plants grew rapidly and produced high yields could be accurately and rapidly screened, which would benefit rapeseed breeding in the future.
- (2) It was found that a combination of a few image-based traits (i-traits) measured at the multiple stages could be used as good predictors of final yield. For example, rather than other i-traits extracted from the side view, the new i-trait FDNIC_TV_7 (fractal dimension without image cropping in top view, reflecting the top-view projected shoot area and leaf shape) had a good correlation with final yield. In addition, as our image analysis pipeline is largely independent of the phenotyping platform, both indoors and outdoors, the pipeline and extracted i-traits could potentially be adapted for field phenotyping by acquiring top-view images of rapeseed plants with unmanned aerial vehicles (UAV) and other field-based phenotyping platforms. In future, it will be interesting to see if the selected i-traits could be used to predict the yield under field conditions. However, scoring of the selected i-traits in the field might pose some challenges, e.g. due to the overlap of plants caused by high planting density, which could be solved by 3D imaging (Guo et al., 2018), and a robust identification model (deep learning,

Singh et al., 2018). Other technical challenges using UAV include the variable wind speed, the effect of solar radiation, reasonable flight height, and so on (Yang, et al., 2020), which should be further investigated.

(3) A large-scale QTL analysis showed that 33.5% of QTLs in 2015-2016 and 36.1% of QTLs in 2016-2017 were consistently detected in the two growing seasons. Interestingly, the QTLs responsible for yield indicators colocalized with those for final yield, potentially providing a new mechanism of yield regulation. These results provide us with new insights into understanding the dynamic genetic basis of rapeseed development.

Experimental procedures

Plant materials and experimental design

In our study, a rapeseed ISL population that included 89 lines and its recurrent parent Zhongyou 821 (ZY821) (9 repeats) (Yang et al., 2018) was used for phenotyping in the HRPF. All lines were first planted in a two-row plot, with ten plants in each row, in the experimental field of Huazhong Agricultural University (Wuhan, Hubei Province, China). These lines were grown for 40 days, after which 5 replicates of each line were transplanted into the greenhouse, with 5 kg of soil per pot. Fertilizer was applied at sowing, as described previously (Zhang et al., 2017).

All rapeseed ISLs were screened from the early seedling to the initial flowering stage, at 12 time points (every ~7 days starting from 53 to 138 days after sowing). The trials were performed using a randomized block design with five replications in each growing season: 2015-2016 and 2016-2017. The screening generated a total of 1.62 terabytes of RGB images (16,986,93 images; PNG format), which are available in a database (http://plantphenomics.hzau.edu.cn/search_rape.action, 2015-2016-QTL and 2016-2017-QTL). A movie of the growth of the recurrent parent and two select ISLs is shown in Movies S1-S3. The inspected lines and inspection dates are shown in Table S1, where T1-T12 represent the twelve time points. In our greenhouse experiment, the final yield per plant was determined manually following harvest. Additionally, to evaluate the measurement accuracy, the plant height, fresh weight and dry weight were both screened nondestructively and measured destructively at the

same 12 time points using a subset of random populations that included the recurrent parent ZY821 and 9 ISLs. The experimental layout and temperature conditions of the two growing seasons are shown in Figure S5 and Table S12, respectively.

Image analysis pipeline and trait extraction for rapeseed

We developed specific rapeseed image analysis pipelines, including a program for side-view images and one for top-view images, to obtain 43 dynamic i-traits (Figure S6), including those such as plant morphological traits, histogram texture traits, color traits and biomass-related traits.

The workflow of the side-view image program was mainly as follows: (1) With the original side-view image (Figure S6a), a predefined value was used to remove the conveyor and other impurities to obtain the cropped image (Figure S6b); (2) HSI colour space (hue, saturation, and intensity) segmentation was used to obtain the segmented binary image (Figure S6c) and to calculate morphological traits and biomass traits; (3) The convex hull image (Figure S6d) could also be obtained by calling the enrolled OpenCV dynamic link library and calculating the convex hull image and other morphological traits; and (4) By the use of the binary image (Figure S6c) as the mask of the original side-view image (Figure S6a), the processed RGB image (Figure S6e) and the corresponding gray image of the intensity channel (Figure S6f) were obtained to calculate the histogram texture traits.

Similarly, the workflow of the top-view image program mainly was as follows: (1) With the original top-view image (Figure S6g), EG segmentation was used to obtain the segmented binary image (Figure S6h); (2) The impurities were removed, the final binary image was obtained (Figure S6i), and the morphological traits were calculated; and (3) By the use of the binary image (Figure S6i) as a mask of the original top-view image (Figure S6g), the processed RGB image (Figure S6j), gray image of the intensity channel (Figure S6k), and gray image of the green channel (Figure S6l) were obtained to calculate the histogram texture traits and color trait.

The programs were developed using LabVIEW 2015 (National Instruments, US) and involved a dynamic link library. The detailed instructions of the image analysis pipeline are shown in Note S1, and the operation procedures of the two programs are shown in Movies S4 and S5. The

download link of the source code of programs is shown in the “Availability of the data” section. The trait explanation and phenotypic data at all 12 time points in the two growing seasons are also shown in Table S13.

Biomass modeling

Two i-traits, including the total projected area in top view (TPA_TV) and the total projected area in side view (TPA_SV), and the destructively measured fresh weight and dry weight were collected from 120 plants across all 12 developmental time points. Eleven models, including linear, quadratic, exponential and power models (shown in Tables S2 and S3), were evaluated to determine the best model for measuring fresh weight and dry weight by calculating the adjusted coefficient of determination (adjusted R^2), mean absolute percentage error (MAPE) and standard deviation of the absolute percentage error (SD_{APE}). A 10-fold cross-validation approach was then introduced to test the prediction accuracy of the best models across different growth stages. Briefly, the 120 plants at all 12 time points were shuffled randomly and split equally into 10 groups. For each group, one group and the remaining nine groups were used as a testing set and training set, respectively. Finally, the model was evaluated by the average adjusted R^2 , MAPE and SD_{APE} values. All statistical analyses were performed with IBM SPSS statistics 20 (IBM, US).

Digital biomass accumulation modeling with dynamic TPA_SV and TPA_TV

Based on the i-traits of TPA_TV and TPA_SV at the 12 time points, we generated 30 growth-related traits to reflect the growth speed or growth size via 6 models (linear, power, exponential, logarithmic, quadratic, and sinusoidal models), which were implemented using LabVIEW 2015 (National Instruments, US).

Yield prediction using i-traits at the multiple growth stages

With all the dynamic i-traits at the 12 time points in the multiple growth stages, linear stepwise regression was used to determine the percentage of variance explained for the rapeseed final yield at the mature stage. The linear stepwise regression analysis was implemented with IBM SPSS

statistics 20 (IBM, US). The results of the yield predictions were validated using 5-fold cross-validation, which was implemented with LabVIEW 2015 (National Instruments, US).

Heritability analysis

Broad-sense heritability (H^2) was calculated as $H^2 = \delta_G^2 / (\delta_G^2 + \delta_{GE}^2/n + \delta_e^2/nr)$, where δ_G^2 is the genetic variance, δ_{GE}^2 is the variance of the genotype by environment interaction, δ_e^2 is the residual error variance, n is the number of environments, and r is the number of replicates within the environment. The values of δ_G^2 , δ_{GE}^2 and δ_e^2 were estimated by analysis of variance (ANOVA) using the lmer function in the lme4 package in the R environment (Team, 2012).

QTL analysis

The detailed characteristics of substituted chromosome segments of the ISL population are described in a previous study (Yang et al., 2018). In brief, the 89 ISLs were genotyped using the GBS approach, and 4214 SNP markers were found across the 19 chromosomes. These ISLs carried a total of 350 substituted chromosome segments. Significantly associated markers along with their amount of phenotypic variation explained (PVE) for all 43 i-traits at each time point were determined using the software QTL IciMapping 4.1 (Meng et al., 2015), for which the RSTEP-LRT-ADD model with a threshold of LOD=2.5 was used. To further validate the QTLs, Dunnett's test (Dunnett, 1955) was used for multiple comparisons to determine significant differences for each i-trait between the recurrent parent ZY821 and each ISL, which carried substituted chromosome segments with significantly associated markers. Last, the QTL intervals were defined as the overlapping region shared by target ISLs, according to the newly published ZS11 genome sequence (Song et al., 2020). All QTLs with overlapping QTL intervals were categorized as nonredundant QTLs. A permutation test using a 1 Mb interval was used to assess the statistical significance of the deviation of the observed QTL distribution from a uniform distribution. After the 10000-permutation test, the thresholds of QTL number per megabase ($p < 0.01$) were 32 and 51 in 2015-2016 and 2016-2017, respectively.

Availability of the data

The genotypic and phenotypic data of the two growing seasons in this study are available at <http://plantphenomics.hzau.edu.cn/search Rape.action> under the sections 2015-2016-QTL and 2016-2017-QTL. The phenotypic data are also shown in Table S13. All the source code, including that of LabVIEW programs, the dynamic link library, cpp documents, and test images, can be downloaded from <https://github.com/fenghuihf2006?tab=repositories>.

Acknowledgments

This work was supported by grants from the National Key Research and Development Program (2016YFD0100101-18), the National Natural Science Foundation of China (31770397), and the Fundamental Research Funds for the Central Universities (2662017PY058, 2662017QD044).

Conflicts of interest statement

The authors declare that they have no conflicts of interest.

Author Contributions

H.L., H.F., and W.Y. performed the experiments, analyzed the data and wrote the manuscript. C.G., S.Y., W.H., X.X., J.L. G.C., and Q.L. assisted in the data analysis and database information construction. W.Y., L.X., and K.L. supervised the project and designed the research.

Short supporting material legends

Figure S1 High-throughput phenotyping of rapeseed. (a) Rapeseed cultivation and transportation in the greenhouse. (b) RGB imaging chamber. (c) The exemplification of image acquisition at 12 time points (15 side-view images and 1 top-view image for each inspection).

Figure S2 Comparison of automatic digital measurements versus manual measurements of plant height (a), fresh weight (b), and dry weight (c).

Figure S3 Heatmap showing the correlation coefficients for 43 i-traits across 12 time points.

Figure S4 Distribution of interval length of the QTLs detected in the two growing seasons.

Figure S5 Experiment setup of rapeseed cultivation. (a) Rapeseed cultivation in the greenhouse. (b) System design and layout of the HRPF (Yang et al., 2014). (c) Space size of tested plant within and between trough in the greenhouse.

Figure S6 Image analysis pipeline for side-view images (a-f) and top-view images (g-l).

Table S1 Inspected lines and inspection dates during the growing seasons of 2015-2016 and 2016-2017.

Table S2 Statistical summary of the 11 developed models for dry weight estimation in 2015-2016 and 2016-2017.

Table S3 Statistical summary of the 11 developed models for fresh weight estimation in 2015-2016 and 2016-2017.

Table S4 Summary of the 10-fold cross-validation of model 9 for dry weight.

Table S5 Summary of the 10-fold cross-validation of model 9 for fresh weight.

Table S6 Statistical details of coefficients of the selected model for yield (combining 10 i-traits among 12 time points).

Table S7 Statistical details of coefficients of the selected model for yield (combining 8 i-traits at 3 time points).

Table S8 Performance evaluation of yield prediction of testing sets using ten random instances of 5-fold cross-validation.

Table S9 Detailed results of digital biomass modeling for TPA_SV and TPA_TV using six models.

Table S10 Phenotypic variation and broad-sense heritability of 43 i-traits of the ISL population.

Table S11 Summary of QTL information for all 43 i-traits across 12 time points.

Table S12 List of temperature conditions on the date of inspection in the two growing seasons.

Table S13 Original data and trait explanations of 43 i-traits across 12 time points.

Movie S1 Dynamic growth video of recurrent parent Zhongyou 821 (Phenotype ID 11).

Movie S2 Dynamic growth video of the ISL (Phenotype ID 36).

Movie S3 Dynamic growth video of the ISL (Phenotype ID 68).

Movie S4 Operational procedure of the image analysis pipeline for side-view images.

Movie S5 Operational procedure of the image analysis pipeline for top-view images.

Note S1 Instructions for the image analysis pipeline.

References

Bayer, P.E., Hurgobin, B., Golicz, A.A., Chan, C.-K.K., Yuan, Y., Lee, H., Renton, M., Meng, J., Li, R., Long, Y., Zou, J., Bancroft, I., Chalhoub, B., King, G.J., Batley, J. and Edwards, D. (2017) Assembly and comparison of two closely related *Brassica napus* genomes. *Plant Biotechnol. J.* **15**, 1602-1610.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, h., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corr ea, M., Silva, C.D., Just, J., Falentin, C., Koh, C.S., Clainche, I.L., Bernard, M., Bento, P., Noel, B., Labadie, K., Alberti, A., Charles, M., Arnaud, D., Guo, H., Daviaud, C., Alamery, S., Jabbari, K., Zhao, M., Edger, P.P., Chelaifa, H., Tack, D., Lassalle, G., Mestiri, I., Schnel, N., Le, P.M.-C., Fan, G., Renault, V., E, B.P., Golicz, A.A., Manoli, S., Lee, T.-H., Thi, V.H.D., Chalabi, S., Hu, Q., Fan, C., Tollenaere, R., Lu, Y., Battail, C., Shen, J., Sidebottom, C.H.D., Wang, X., Canaguier, A., Chauveau, A., B rard, A., Deniot, G., Guan, M., Liu, Z., Sun, F., Lim, Y.P., Lyons, E., Town, C.D., Bancroft, I., Wang, X., Meng, J., Ma, J., Pires, J.C., King, G.J., Brunel, D., Delourme, R., Renard, M., Aury, J.-M., Adams, K.L., Batley, J., Snowdon, R.J., Tost, J., Edwards, D., Zhou, Y., Hua, W., Sharpe, A.G., Paterson, A.H., Guan, C. and Wincker, P. (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome *Science* **345**, 950-953.

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T. and Klukas, C. (2014) Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* **26**, 4636-4655.

Chen, X., Li, X., Zhang, B., Xu, J., Wu, Z., Wang, B., Li, H., Younas, M., Huang, L., Luo, Y., Wu, J., Hu, S. and Liu, K. (2013) Detection and genotyping of restriction fragment associated polymorphisms in polyploid crops with a pseudo-reference sequence: a case study in allotetraploid *Brassica napus*. *BMC Genomics* **14**, 346.

Clarke, W.E., Higgins, E.E., Plieske, J., Wieseke, R., Sidebottom, C., Khedikar, Y., Batley, J., Edwards, D., Meng, J., Li, R., Lawley, C.T., Pauquet, J., Laga, B., Cheung, W., Iniguez-Luy, F., Dyrzka, E., Rae, S., Stich, B., Snowdon, R.J., Sharpe, A.G., Ganal, M.W. and Parkin, I.A.P. (2016) A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. *Theor. Appl. Genet.* **129**, 1887-1899.

Delourme, R., Laperche, A., Bouchet, A.-S., Jubault, M., Paillard, S., Manzanares-Dauleux, M.-J. and Nesi, N. (2018) Genes and quantitative trait loci mapping for major agronomic traits in *Brassica napus* L. In: *The Brassica napus genome* (Liu, S., Snowdon, R. and Chalhoub, B. eds), pp. 41-85. Cham, Switzerland: Springer Nature.

Dunnett, C.W. (1955) A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.* **50**, 1096-1121.

Friedt, W., Tu, J. and Fu, T. (2018) Academic and economic importance of *Brassica napus* rapeseed. In: *The Brassica napus genome* (Liu, S., Snowdon, R. and Chalhoub, B. eds), pp. 1-20. Cham, Switzerland: Springer Nature.

Guo, Q., Wu, F., Pang, S., Zhao, X., Chen, L., Liu, J., Xue, B., Xu, G., Li, L., Jing, H., et al. (2018). Crop 3D—a LiDAR based platform for 3D high-throughput crop phenotyping. *Sci. China Life Sci.* **61**: 328-339.

Hu, J., Guo, C., Wang, B., Ye, J., Liu, M., Wu, Z., Xiao, Y., Zhang, Q., Li, H., King, G.J. and Liu, K. (2018) Genetic properties of a nested association mapping population constructed with semi-winter and spring oilseed rapes. *Front. Plant Sci.* **9**, 1740.

Huang, X. and Han, B. (2014) Natural variations and genome-wide association studies in crop plants. *Annu. Rev. Plant Biol.* **65**, 531-551.

Knoch, D., Abbadi, A., Grandke, F., Meyer, R.C., Samans, B., Werner, C.R., Snowdon, R.J. and Altmann, T. (2020) Strong temporal dynamics of QTL action on plant growth progression revealed through high-throughput phenotyping in canola. *Plant Biotechnol. J.* **18**, 68-82.

Luo, X., Ma, C., Yue, Y., Hu, K., Li, Y., Duan, Z., Wu, M., Tu, J., Shen, J., Yi, B. and Fu, T. (2015) Unravelling the complex trait of harvest index in rapeseed (*Brassica napus* L.) with

association mapping. *BMC Genomics* **16**, 379.

Meng, L., Li, H., Zhang, L. and Wang, J. (2015) QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* **3**, 269-283.

Muraya, M.M., Chu, J., Zhao, Y., Junker, A., Klukas, C., Reif, J.C. and Altmann, T. (2017) Genetic variation of growth dynamics in maize (*Zea mays* L.) revealed through automated non-invasive phenotyping. *Plant J.* **89**, 366-380.

Singh, A.K., Ganapathysubramanian, B., Sarkar, S., and Singh, A. (2018). Deep learning for plant stress phenotyping: trends and future perspectives. *Trends Plant Sci.* **23**:883-898.

Song, J., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., Liu, D., Wang, B., Lu, S., Zhou, R., Xie, W., Cheng, Y., Zhang, Y., Liu, K., Yang, Q., Chen, L. and Guo, L. (2020) Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* **6**, 34-45.

Sun, F., Fan, G., Hu, Q., Zhou, Y., Guan, M., Tong, C., Li, J., Du, D., Qi, C., Jiang, L., Liu, W., Huang, S., Chen, W., Yu, J., Mei, D., Meng, J., Zeng, P., Shi, J., Liu, K., Wang, X., Wang, X., Long, Y., Liang, X., Hu, Z., Huang, G., Dong, C., Zhang, H., Li, J., Zhang, Y., Li, L., Shi, C., Wang, J., Lee, S.M.-Y., Guan, C., Xu, X., Liu, S., Liu, X., Chalhoub, B., Hua, W. and Wang, H. (2017) The high-quality genome of *Brassica napus* cultivar 'ZS11' reveals the introgression history in semi-winter morphotype. *Plant J.* **92**, 452-468.

Team, R.C. (2012) *R: a language and environment for statistical computing*: Vienna, Austria: R Foundation for Statistical Computing [WWW document] URL <http://www.R-project.org/>. [accessed 17 December 2015].

Wang, B., Wu, Z., Li, Z., Zhang, Q., Hu, J., Xiao, Y., Cai, D., Wu, J., King, G.J., Li, H. and Liu, K. (2018) Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *Brassica napus*. *Plant Biotechnol. J.* **16**, 1336-1348.

Yang, S., Zhang, B., Liu, G., Hong, B., Xu, J., Chen, X., Wang, B., Wu, Z., Hou, F., Yue, X., Wang, J., Zhang, Q., King, G.J. and Liu, K. (2018) A comprehensive and precise set of intervarietal substitution lines to identify candidate genes and quantitative trait loci in

oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* **131**, 2117-2129.

Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., Fang, W., Feng, H., Xie, W., Lian, X., Wang, G., Luo, Q., Zhang, Q., Liu, Q. and Xiong, L. (2014) Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* **5**, 5087.

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J.H., Batchelor, W.D., Xiong, L., and Yan, J. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Mol. Plant.* **13**, 187–214.

Yang, Y., Shi, J., Wang, X., Liu, G. and Wang, H. (2016) Genetic architecture and mechanism of seed number per pod in rapeseed: elucidated through linkage and near-isogenic line analysis. *Sci. Rep.* **6**, 24124.

Zhang, X., Huang, C., Wu, D., Qiao, F., Li, W., Duan, L., Wang, K., Xiao, Y., Chen, G., Liu, Q., Xiong, L., Yang, W. and Yan, J. (2017) High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiol.* **173**, 1554-1564.

Zou, J., Mao, L., Qiu, J., Wang, M., Jia, L., Wu, D., He, Z., Chen, M., Shen, Y., Shen, E., Huang, Y., Li, R., Hu, D., Shi, L., Wang, K., Zhu, Q., Ye, C., Bancroft, I., King, G.J., Meng, J. and Fan, L. (2019) Genome-wide selection footprints and deleterious variations in young Asian allotetraploid rapeseed. *Plant Biotechnol. J.* **17**, 1998-2010.

Figure legends

Figure 1 Yield prediction using i-traits. (a) Prediction of yield with 8 i-traits at 3 time points. These 8 i-traits include the following: FDNIC_TV-7, FDIC_SV-1, GPA_TV-7, PAR_SV_1, H_SV_12, PC3_SV-7, PC5_SV-7 and AC_TV-12. (b) R^2 value between FDNIC_TV and yield across 12 time points. (c) Dynamic change in plant height across the 12 time points. The error bars represent the standard deviation (SD) across all lines in the population. (d) The side-view and top-view images of rapeseed plants (Population ID 11) across 12 time points.

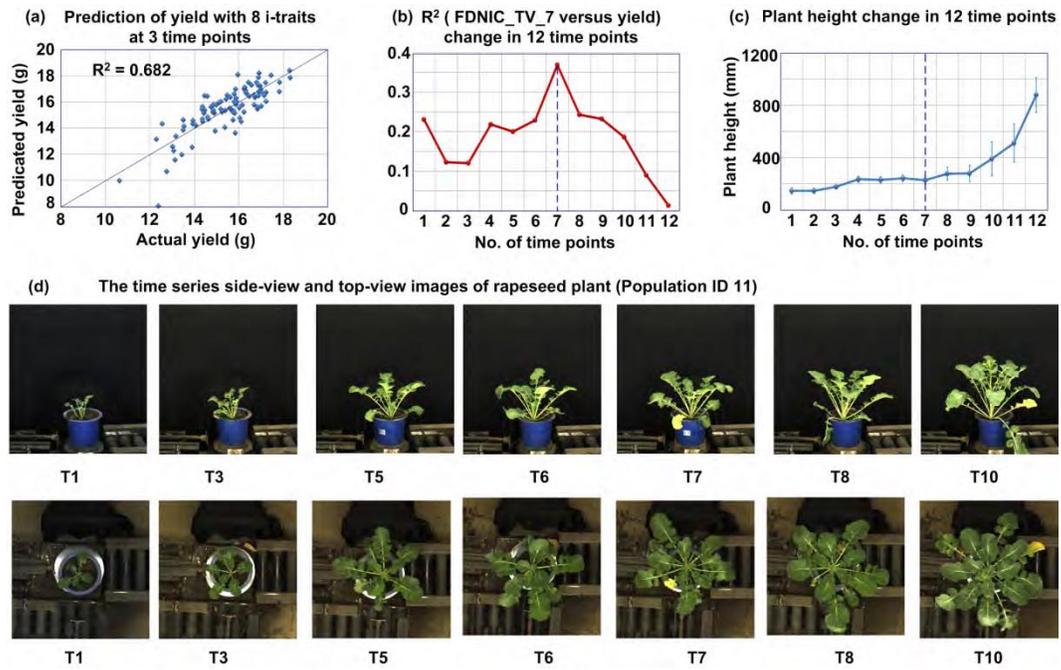
Figure 2 Correlations between growth-related traits and yield. (a) Comparison of actual digital

biomass (blue line) and predicted digital biomass (red line). (b) R^2 value between the actual yield and growth-related trait `a_Exp_SV`. (c) Heatmap of correlations between the yield and 30 growth-related traits.

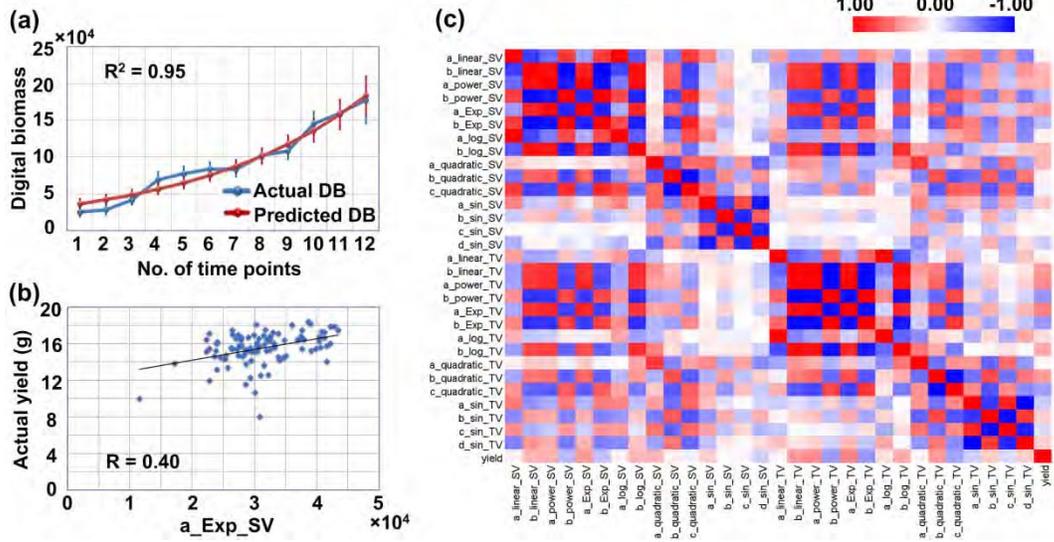
Figure 3 Phenotypic variation and heritability of 43 i-traits across 12 time points. Heatmap showing the phenotypic fold-change (a) and broad-sense heritability (H^2) of traits (b). T1-T12 time points are shown from the outer to inner circles.

Figure 4 Dynamic QTLs detected in the ISL population in the two growing seasons. (a, b) QTLs responsible for the biomass in 2015-2016 (a) and 2016-2017 (b). (c, d) QTLs responsible for the 23 i-traits extracted from side-view images in 2015-2016 (c) and 2016-2017 (d). `i1-i23` represent `AC_SV`, `E_TEX_SV`, `FDIC_SV`, `FDNIC_SV`, `GPA_SV`, `H_SV`, `HA_SV`, `HWR_SV`, `M_TEX_SV`, `MU3_TEX_SV`, `PAR_SV`, `PC1_SV`, `PC2_SV`, `PC3_SV`, `PC4_SV`, `PC5_SV`, `PC6_SV`, `R_SV`, `S_TEX_SV`, `SE_TEX_SV`, `TPA_SV`, `U_TEX_SV`, and `W_SV`, respectively. (e, f) QTLs responsible for the 18 i-traits extracted from top-view images in 2015-2016 (e) and 2016-2017 (f). `i1-i18` represent `AC_TV`, `E_TEX_TV`, `FDIC_TV`, `FDNIC_TV`, `GCV_TV`, `GPA_TV`, `H_TV`, `HA_TV`, `HWR_TV`, `M_TEX_TV`, `MU3_TEX_TV`, `PAR_TV`, `R_TV`, `S_TEX_TV`, `SE_TEX_TV`, `TPA_TV`, `U_TEX_TV` and `W_TV`, respectively. The T1-T12 time points are shown as circles with a color gradient from light to dark, as indicated in the legend.

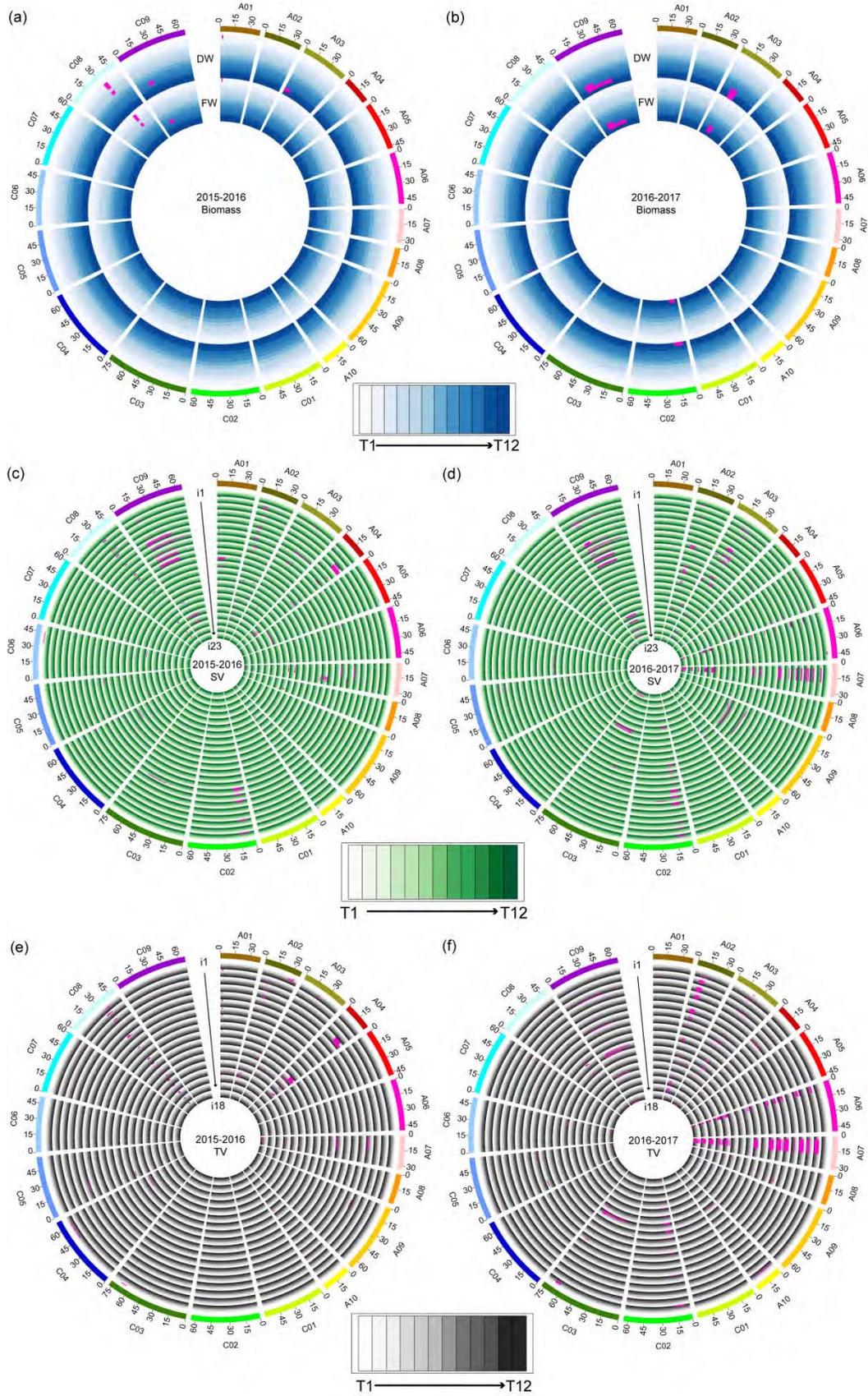
Figure 5 Chromosomal distribution of the QTLs in the two growing seasons. (a) Co-identified QTLs and hotspots in the two growing seasons. The QTL numbers are calculated in the 1 Mb window, and thresholds (orange lines) are 32 and 51 in 2015-2016 (G1) and 2016-2017 (G2), respectively. (b) Diagram showing the overlapping location of QTLs on chromosomes A07 and A05 for i-traits, including `FDNIC_TV-7`, `GPA_TV-7` and `H_SV-12`, and final yield, as assessed in the greenhouse.

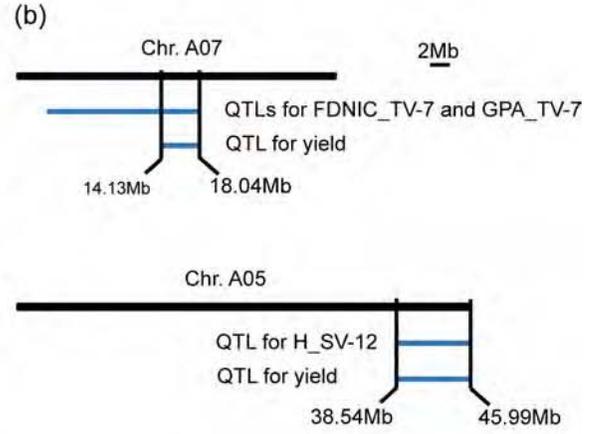
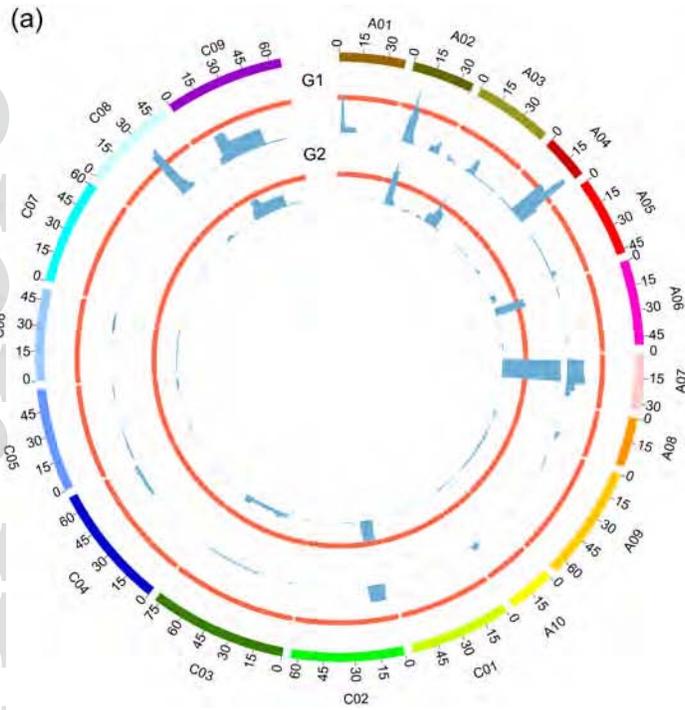


pbi_13396_f1.tif



pbi_13396_f2.tif





pbi_13396_f5.tif