Plant Biotechnology Journal
Open Access

Plant Biotechnology Journal (2021) **19**, pp. 1065–1079 doi: 10.1111/pbi.13531

A novel motif in the 5′-UTR of an orphan gene '*Big Root Biomass*' modulates root biomass in sesame

Komivi Dossa[1,2,*], Rong Zhou[1], Donghua Li[1], Aili Liu[1], Lu Qin[1], Marie A. Mmadi[1], Ruqi Su[1], Yujuan Zhang[1,3], Jianqiang Wang[1], Yuan Gao[1], Xiurong Zhang[1,*] and Jun You[1,*]

[1]*Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture and Rural Affairs, Wuhan, China*
[2]*Laboratory of Genetics, Horticulture and Seed Sciences, Faculty of Agronomic Sciences, University of Abomey-Calavi, Cotonou, Benin*
[3]*Cotton Research Center, Shandong Academy of Agricultural Sciences, Jinan, China*

Received 31 July 2020;
revised 30 November 2020;
accepted 8 December 2020.
**Correspondence* (Tel +86 21 64322008; fax +86 21 86811836; email dossakomivi@gmail.com (K. D.), Tel +86 21 64322008; fax +86 21 86811836; email zhangxr@oilcrops.cn (X. Z.), Tel +86 21 64322008; fax +86 21 86811836; email junyou@caas.sn (J. Y.))

Keywords: genetic architecture, non model crop, plant adaptation, root biomass, *Sesamum indicum*.

Summary

Developing crops with improved root system is crucial in current global warming scenario. Underexploited crops are valuable reservoirs of unique genes that can be harnessed for the improvement of major crops. In this study, we performed genome-wide association studies on seven root traits in sesame (*Sesamum indicum* L.) and uncovered 409 significant signals, 19 quantitative trait loci containing 32 candidate genes. A peak SNP significantly associated with root number and root dry weight traits was located in the promoter of the gene named '*Big Root Biomass*' (*BRB*), which was subsequently validated in a bi-parental population. *BRB* has no functional annotation and is restricted to the Lamiales order. We detected the presence of a novel motif 'AACACACAC' located in the 5′-UTR of *BRB* in single and duplicated copy in accessions with high and small root biomass, respectively. A strong expression level of *BRB* was negatively correlated with high root biomass, and this was attributed to the gene *SiMYB181* which represses the activity of *BRB* by binding specifically to the single motif but not to the duplicated one. Curiously, the allele that enhanced *BRB* expression has been intensively selected by modern breeding. Overexpression of *BRB* in *Arabidopsis* modulates auxin pathway leading to reduced root biomass, improved yield parameters under normal growth conditions and increased drought stress sensitivity. Overall, *BRB* represents a solid gene model for improving the performance of sesame and other crops.

Introduction

Climate change-related risks coupled with the rapid population growth and the rarity of natural resources are posing serious threats on food and nutritional security. This has catalysed global efforts to increase crop productivity and resilience through various crop improvement strategies. Among these, molecular plant breeding plays a central role in the identification of the genetic basis of important agronomic traits and their deployment for the development of improved cultivars (Moose and Mumm, 2008; Varshney *et al*., 2006). However, modern crop breeding has mainly focused on the above-ground plant components, largely neglecting the hidden half of the plant, that is the root system (Voss-Fels *et al*., 2018). As the root system plays a crucial role in anchoring the plant, exploring and exploiting soil resources such as nutrients and water to sustain growth and productivity under various environmental conditions (Lynch, 2013; Wang *et al*., 2015; Svacina *et al*., 2014; Xie *et al*., 2017; Liu *et al*., 2018; Su *et al*., 2019), more studies and efforts are needed to deliver crops with improved below-ground traits.

The complex nature of root traits and the opaque nature of soil are the main factors that confound large-scale and in-depth studies on plant root system. Nonetheless, by adapting artificial growth conditions, phenotyping platforms and novel data processing tools, the genetic architecture of root traits is progressively uncovered in many plant species (Hochholdinger and Tuberosa, 2009; Mai *et al*., 2014; Downie *et al*., 2014; Atkinson *et al*., 2019). For example, using hydroponics system and root imaging, Kitomi *et al*. (2018) identified two quantitative trait loci (QTL) associated with root length in rice. The gel imaging platform was employed to detect several QTLs controlling root traits in maize (Zurek *et al*., 2015) and rice (Topp *et al*., 2013). Beyer *et al*. (2019) recently employed a paper roll-supported hydroponic system to perform genome-wide association studies (GWAS) in wheat that led to the identification of 68 marker–trait associations associated with various root traits. Transparent Plexiglas nailboard sandwiches filled with glass bead, and a nutrient solution was adapted to grow a rice diversity panel and perform GWAS for root traits (Courtois *et al*., 2013). Several studies were also conducted in *Arabidopsis thaliana* to dissect the genetic network controlling root traits by using petri dish or hydroponics (Gifford *et al*., 2013; Meijón *et al*., 2014; Kobayashi *et al*., 2015; Jia *et al*., 2019; Ogura *et al*., 2019). In wheat, the filter paper/polycarbonate screening plates and the 'clear pot' methods were used to discover major QTLs for root system architecture traits (Alahmad *et al*., 2019; Maccaferri *et al*., 2016).

Discoveries made so far on the genetic architecture of root traits have been essentially focused on major crops and model plants while minor crops have been overlooked. Besides their role in the diversification of food production, minor crops represent untapped gene pool, which could be harnessed for improvement of major crops (Mayes *et al*., 2012). Sesame (*Sesamum indicum* L.) is a typical example of an underexploited crop with special attributes such as high-quality vegetable oil (Anilakumar *et al*., 2010; Miyahara *et al*., 2001; Noguchi *et al*., 2001; Sankar *et al*., 2005), good adaptation to marginal lands (Langham, 2007) and resistance to abiotic stresses such as drought and salt stresses (Dossa *et al*., 2019a, b; Islam *et al*., 2016; Lakhanpaul *et al*., 2012). In particular, its extensive root system is thought to play a foremost role in its adaptation to different environments (Dossa *et al*., 2017a,b; Gloaguen *et al*., 2019; Lakhanpaul *et al*., 2012; Langham, 2007). In addition, Su *et al*. (2019) recently reported a positive contribution of high root biomass to increased seed yield in sesame. However, the genetic basis of sesame root system is still unknown.

In a previous study, we designed and successfully adapted a hydroponic growth platform to demonstrate the high variability of root traits in sesame (Su *et al*., 2019). The availability of the genome sequence (Wang *et al*., 2014) and high-quality single-nucleotide polymorphism genotyping data of 705 diverse accessions (Wei *et al*., 2015) offer the opportunity to explore the genetic variants controlling the natural variation of sesame root traits. Herein, we performed GWAS based on a large-scale phenotyping of root traits in a sesame diversity panel. We identified key genomic regions and several candidate genes associated with root traits in sesame. In particular, we identified and functionally characterized a novel gene named '*Big Root Biomass*' (*BRB*), which modulates root biomass in sesame.

Result

Natural variation of root traits in sesame

An association panel composed of 705 diverse sesame (*Sesamum indicum* L.) accessions was previously assembled in our group and used for the detection of genetic variants governing several agronomic traits (Wei *et al*., 2015; Li *et al*., 2018; Zhou *et al*., 2018; Dossa *et al*., 2019a). To investigate variation in root traits in this study, we extracted 327 landraces and modern cultivars with large morphological diversity (Wei *et al*., 2015) and originating from 28 different countries in Africa, Asia, Europe and America (Table S1). A total of seven root traits, including root number (RN), root volume (RV), root length (RL), root surface area (SA), root dry weight (RDW), shoot dry weight (SDW) and root–shoot ratio (RSR), were evaluated on 1-month-old seedlings grown hydroponically under controlled conditions.

Large variation and normal distribution of all the assayed traits were observed in the association panel ($P < 0.001$) (Table S2; Figure S1). The broad-sense heritability (H^2) values of the traits were moderate to high, ranging from 0.58 to 0.83 (Table S2). Moreover, strong correlations were observed between the traits (Figure S2), suggesting the existence of pleiotropic loci.

GWAS for sesame root traits

To uncover the genetic variants underlying the root traits in sesame, we performed a genome-wide association study (GWAS) using one million high-quality single-nucleotide polymorphisms (SNPs) with a minor allele frequency ≥ 0.05 (Wei *et al*., 2015). The mixed model taking into account the population structure and parental relatedness was employed (Dossa *et al*., 2019a). We identified 409 SNPs significantly associated with the different traits at $P < 1 \times 10^{-5}$ (Table S3; Figure S3). Most of these associated loci were distributed in clusters across ten linkage groups (LG). The significant SNPs were resolved into 19 quantitative trait loci (QTL) based on the estimated linkage disequilibrium (LD) decay (88 kb) in the sesame genome (Wei *et al*., 2015; Dossa *et al*., 2019a) (Figure 1a; Table S3). We identified seven pleiotropic QTLs associated with up to five different traits (Table S3). The top five peak SNPs for each trait contributed from ~ 16.7 (RSR) to 67% (RN) of the phenotypic variation in the entire population (Table S3; Figure S4). Collectively, the QTLs harboured 318 genes (Table S4), 30 of which contained in total 90-genic and 59-promoter significant SNPs (Table S5). Candidate genes in the different QTLs were inferred by prioritizing the genes containing significant associated SNPs or closer to the peak SNPs (Dossa *et al*., 2019a). In total, 32 candidate genes controlling root traits and involved in various biological functions related to hormone signalling, transcription regulation, growth and development, cell organization, were uncovered in this study (Figure 1b; Table S6).

BRB gene is significantly associated with root number and root dry weight

The most significant SNPs associated with the root traits were located on the LG15 within the co-localized QTLs *qRN15.2* and *qRDW15.1* (Figure 1c; Table S3). The peak locus SNP5024573 ($P < 6.6 \times 10^{-10}$) fell within the promoter region of the gene *SIN_1025576*, dubbed *BRB* ('*Big Root Biomass*'). *BRB* contained one synonymous SNP SNP5022360 (A/T) located at 393 bp in the coding region, one SNP SNP5022842 (C/G) located at −90 bp in the 5′-UTR and 11 SNPs topped by the SNP SNP5024573 (G/A) located at −1821 bp in the promoter (Figure 2a; Table S7). The two loci, SNP5022842 and SNP5024573, were found in complete LD, resulting into two haplotypes (Hap1 and Hap2). Accessions with Hap1 representing the smallest group ($n = 63$) had significantly higher RN and RDW ($P < 0.001$) than those with Hap2 ($n = 253$) (Figure 2b,c).

To further confirm that variants in *BRB* modulate root biomass in sesame, we developed 197 $F_{2:3}$ population from the accessions G340 (RDW = 0.121 ± 0.04 g; RN = 1137±109) and G441 (RDW = 0.040 ± 0.019 g; RN = 566±72) having the Hap1 and Hap2, respectively. We tried unsuccessfully to sequence 2 kb promoter region and the coding region (417 bp) of *BRB* in the $F_{2:3}$ population. However, after decreasing the promoter size, we sequenced and analysed 292 bp upstream of the translation start codon (ATG), the full coding sequence and 180 bp downstream of the stop codon (TGA). A total of 146 SNPs and 53 insertions and deletions (InDel) were identified. Four InDels (InDel5022798, InDel5022796, InDel5022797 and InDel5022799) topped by IndDel5022798 located at −46 bp in the 5′-UTR of *BRB* contributed significantly to the variation of RDW in the $F_{2:3}$ population grown in pots containing soil as substrate ($P < 1.67 \times 10^{-4}$) (Figure 2d; Table S8). All these InDels were found in strong LD ($r^2 > 0.9$). Similarly, a bunch of six InDels (InDel5022804, InDel5022801, InDel5022800, InDel5022798, InDel5022799 and InDel5022796) and the GWAS-based identified significant locus SNP5022842 were found significantly associated with RN, peaked by InDel5022804 located at −52 bp upstream of the ATG ($P < 2.3 \times 10^{-4}$) (Figure 2e; Table S8). All these loci were also found in strong LD ($r^2 > 0.9$). Sesame lines with a nucleotide deletion at the peak InDels

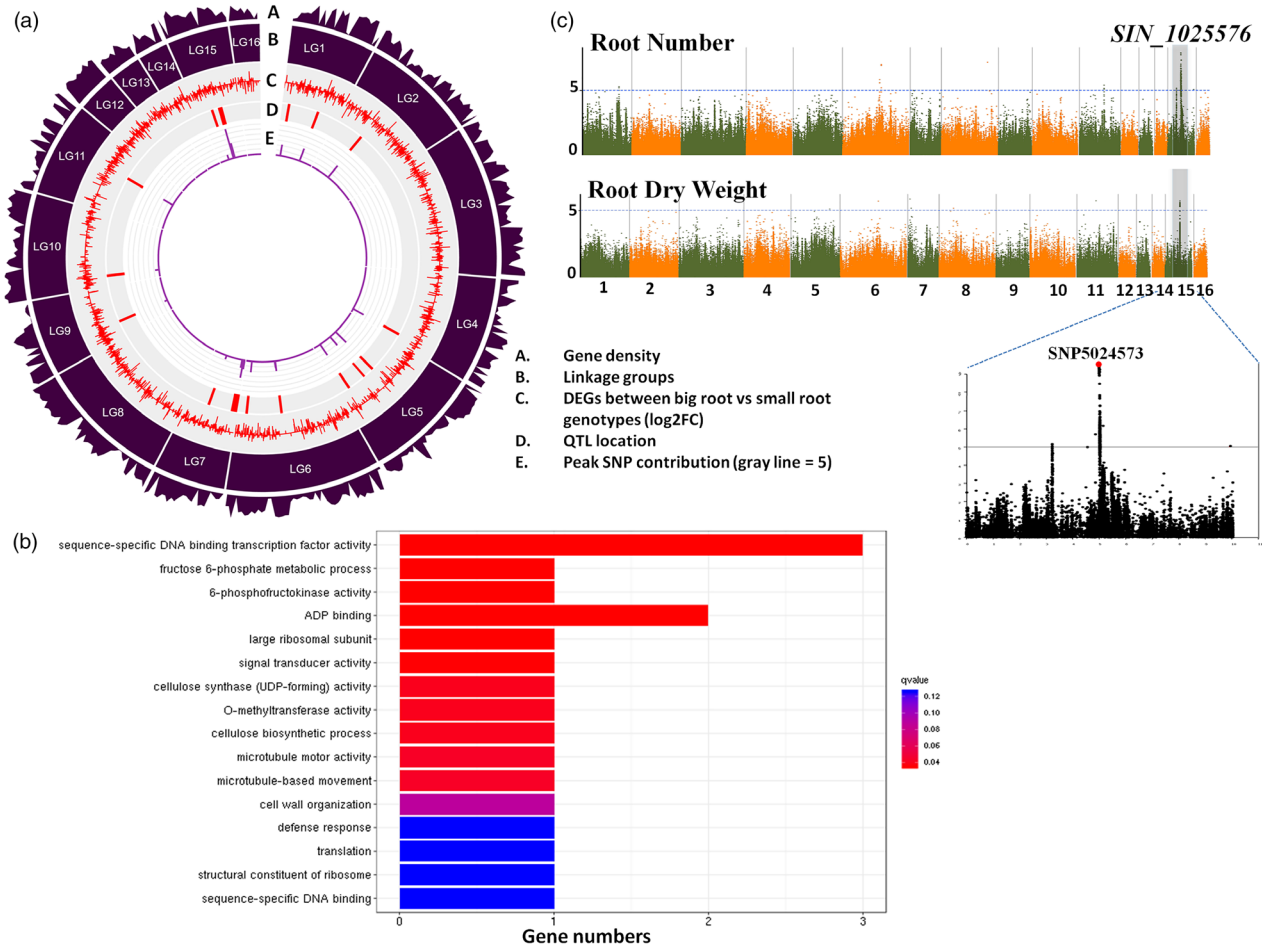

Figure 1 Genome-wide association study (GWAS) for root traits in sesame. (a) Circos plot summarizing the GWAS results; (b) Gene Ontology enrichment analysis of the candidate genes identified for root traits in sesame; (c) Manhattan plot for root number and root dry weight traits. A strong association was co-detected for the two traits on the linkage group 15, and the regional plot displays the top SNP (SNP5024573) located in the gene *SIN_1025576*.

(InDel5022798 and InDel5022804) displayed significantly high RDW and RN ($P < 0.001$) than those with a nucleotide insertion (Figure 2f,g). Taken together, our results demonstrated that variants in the 5′-UTR of *BRB*, particularly the InDels, contribute to the natural variation of RN and RDW in sesame, independently of the growing media.

The significant InDels correspond to a novel motif in the 5′-UTR altering the expression of *BRB*

Since the detected InDels contributing to the variation of RDW and RN in sesame were closely located in the 5′-UTR of *BRB* based on the $F_{2:3}$ population analysis, we sequenced and analysed the 5′-UTR in 50 accessions from the association panel, including 25 accessions with Hap1 and 25 accessions with Hap2. We found that all the clustered significant InDels for RDW and RN correspond to the presence/absence of a 9 bp sequence duplicated motif 'AACACACAC' (Figure 3a). Blast search in various databases (http://www.dna.affrc.go.jp/PLACE/; http://arabidopsis.med.ohio-state.edu) for the identification of the motif did not return any hit, so we presumed that it is an unreported or an uncharacterized motif. All sesame accessions with Hap1 displayed the single AACACACAC motif while accessions with Hap2 had a duplicated motif (18 bp), indicating that the duplication pattern of the motif is associated with variations of RDW and RN in sesame (Figure S5).

BRB is preferentially expressed in sesame root as compared to other organs such seed, leaf, flower and capsule (Figure 3b). Similarly, β-glucuronidase (gus) gene driven by the promoter of *BRB* was expressed in *Arabidopsis* main root and root hair but was not expressed in the root tip (Figure 3c). Because of the specific position of the repeated motif in the 5′-UTR of *BRB*, we hypothesized that it may affect the gene expression level among the accessions with Hap1 and Hap2. To test our hypothesis, we analysed the expression level of *BRB* using our previously released root transcriptome data from G546 (Hap1) and G259 (Hap2) (Su *et al.*, 2019). We found that the expression level of *BRB* was ~ 45-fold lower in G546 than in G259 (Figure 3d), suggesting that *BRB* is a negative regulator of root biomass in sesame. To further confirm this finding, we performed a quantitative reverse transcription PCR (qRT-PCR) of root RNA samples from 25 Hap1 and 25 Hap2 accessions. The qRT-PCR results validated that accessions with the single motif (Hap1) exhibited significantly

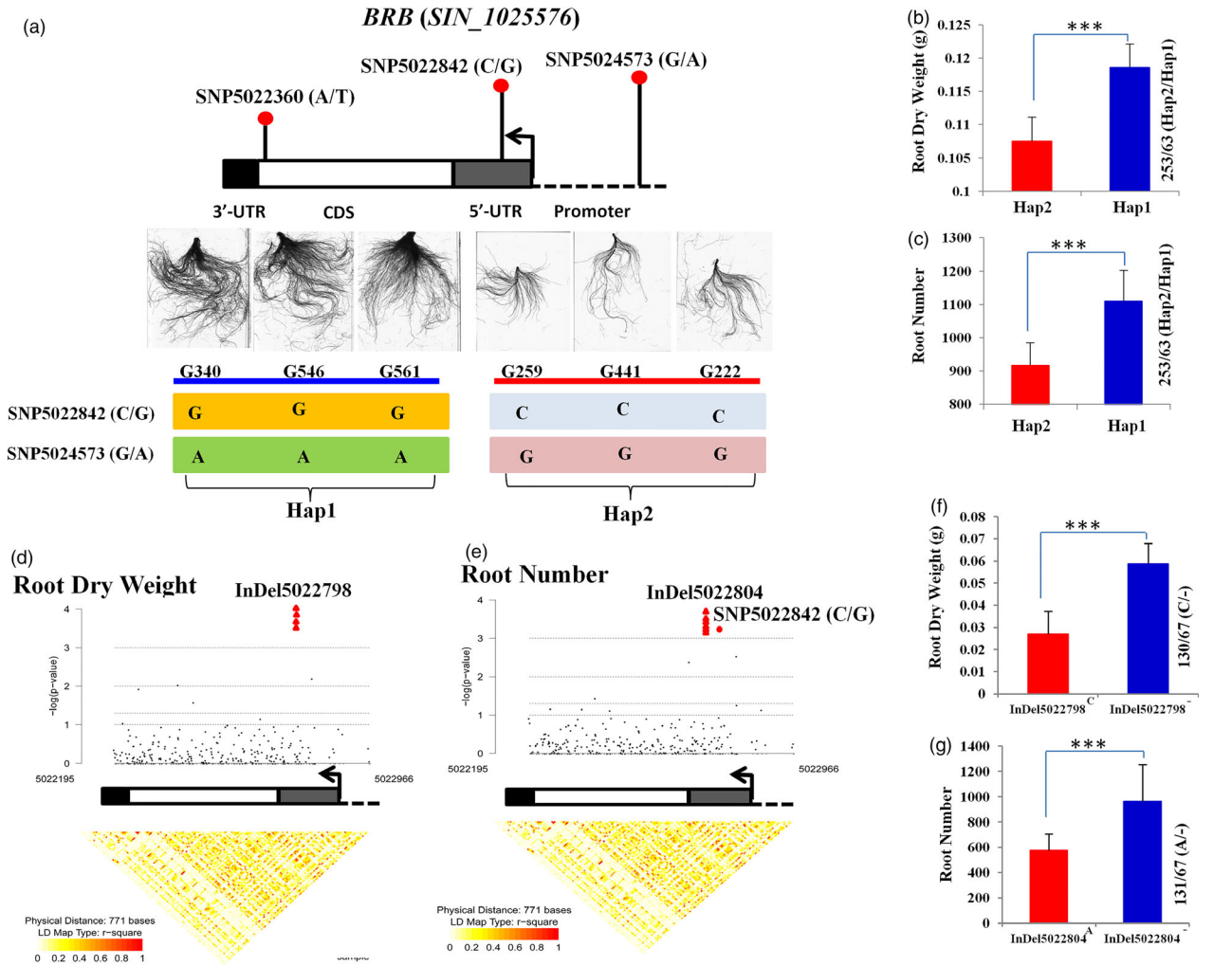

Figure 2 Natural variations in the gene *BRB* (*SIN_1025576*) are associated with root traits. (a) Schematic representation of the significant SNPs located in coding and promoter regions of *BRB*. The SNP5022360 (A/T) was synonymous SNP so was excluded from further analyses. The two other SNPs, SNP5022842 (C/G) and SNP5024573 (G/A), were in perfect linkage disequilibrium leading to two haplotypes (Hap1 and Hap2). Sesame genotypes with Hap1 display dense root biomass while genotypes with Hap2 exhibited small root biomass; (b) root dry weight was significantly higher in Hap1 genotypes than those with Hap2; (c) root number was significantly higher in Hap1 genotypes than those with Hap2; (d–e) association analysis of root dry weight and root number traits in 197 $F_{2:3}$ sesame population. The significant associated InDels and SNPs were marked in red triangle and dot, respectively. (f–g) sesame lines with a nucleotide deletion at the peak InDels (InDel5022798 and InDel5022804) displayed significantly high root dry weight and root number than those with a nucleotide insertion. Error bars indicate the SD (*** significant difference at $P < 0.001$).

lower level of *BRB* transcript as compared to those with the duplicated motif (Hap2) ($P < 0.001$) (Figure 3e) with similar observations in leaf samples (Data not shown).

The single motif AACACACAC facilitates the binding of an MYB transcription factor in the 5′-UTR of *BRB*

Gene expression levels are known to be modulated by transcription factors (TF), binding to specific motifs in the promoter. We speculated that the duplication pattern of the motif AACACACAC in the 5′-UTR of *BRB* may affect the binding of its regulator TF, leading to the differential expression level between Hap1 and Hap2. To examine this possibility, we first identified the TFs co-expressed with *BRB* by performing a weighted gene co-expression network analysis using 60 RNA-seq datasets previously released by our group from root samples of two sesame genotypes at 10 time points of root growth (National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) accession number: SRP181800). *BRB* gene was co-expressed with two WRKY and two MYB TFs, which represent the potential regulators (Table S9). Next, we performed yeast one-hybrid assay using the cloned 5′-UTR sequences from the Hap1 and Hap2. In addition, we also designed two different synthetic 5′-UTR sequences by switching the alleles at the locus SNP5022842 (C/G) to assess the effect of this polymorphism on TF binding ability. The MYB gene *SIN_1023179* (*SiMYB181*; Mmadi *et al.*, 2017) was able to bind to the 5′-UTR of *BRB* but specifically to the Hap1 sequences (Figure 3f). Also, the base change at the locus SNP5022842 did not affect the *SiMYB181* binding. Altogether, our results suggest that the deletion of the motif AACACACAC in Hap1 accessions facilitates the binding of *SiMYB181* to the 5′-

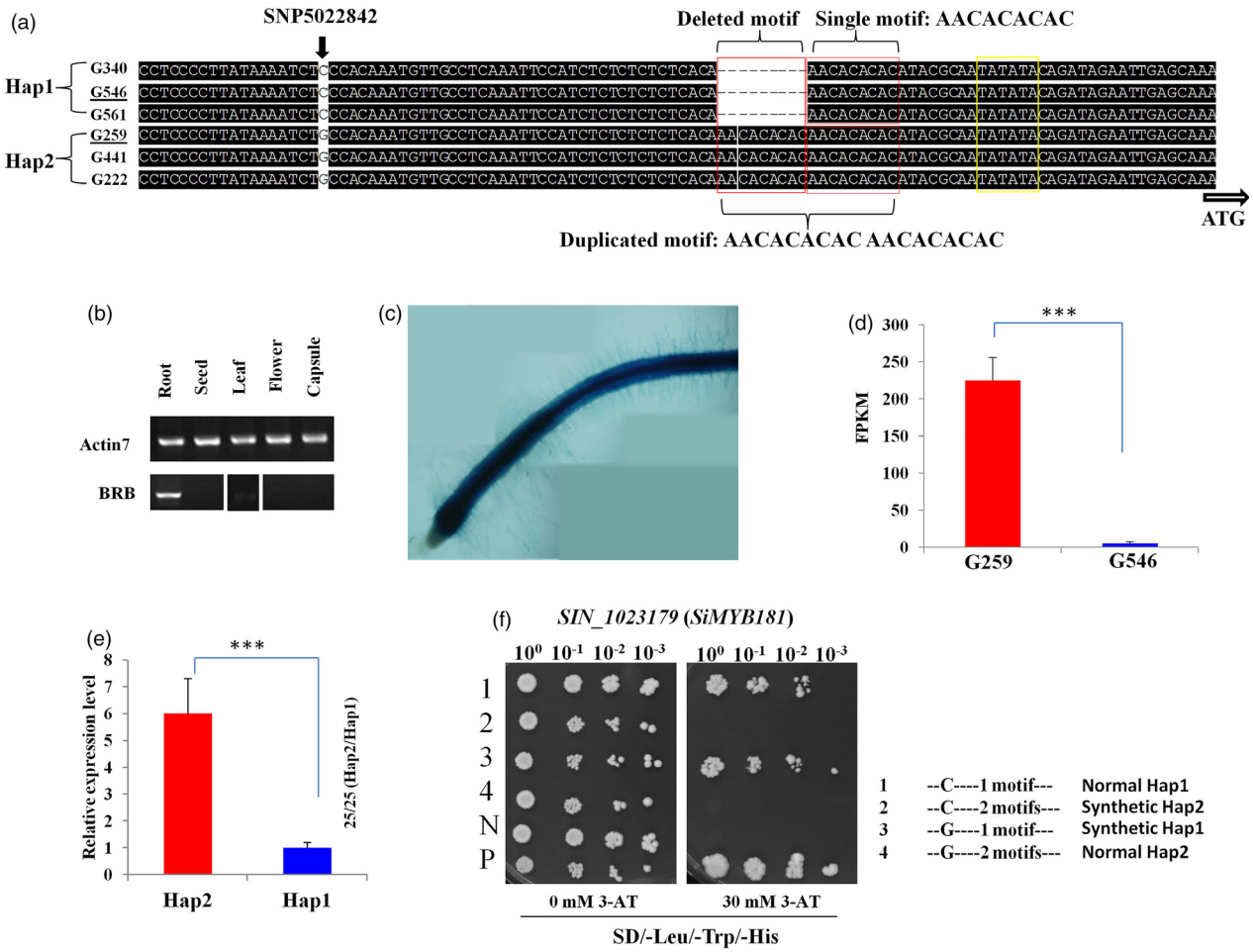

Figure 3 A motif in the 5′-UTR alters the expression level of *BRB* gene. (a) Alignment of the reverse complemented sequences of genotypes from Hap1 and Hap2. The red box shows the location of the conserved motif (AACACACAC) in Hap2 and Hap1. The yellow box shows the TATA box. The genotype names underlined are those used later for RNA-sequencing; (b) RT-PCR analysis of transcript levels of the *BRB* gene in various sesame tissues. The band for Leaf in BRB row was rearranged from the initial gel images and this has been highlighted by a white space around the band; (c) β-glucuronidase (*gus*) gene driven by the promoter of *BRB* was expressed in *Arabidopsis* main root and root hair but not in the root tip. The composite picture was generated by combining multiple pictures from different root sections; (d) expression levels of *BRB* gene in two contrasting genotypes for root trait based on RNA-seq. G546 (Hap1) has dense root biomass while G259 (Hap2) has small root biomass; (e) qRT-PCR analysis of the transcript levels of *BRB* in 25 genotypes (Hap1) and 25 genotypes (Hap2). Error bars indicate the SD (*** significant difference at $P < 0.001$); (f) Yeast one-hybrid assay for the interactions between the transcription factor *SIN_1023179* (*SiMYB181*) and the 5′-UTR sequences of *BRB* (Hap1, Hap2 and synthetic sequences). The co-transformed yeasts strain Y187 containing the bait and prey were cultivated on the synthetic dextrose minimal medium (SD) lacking leucine, tryptophan and histidine (SD/-Leu/-Trp/-His) supplemented with 0 or 30 mM 3-amino-1,2,4-triazole (3-AT). N = negative control and P = positive control.

UTR of *BRB*, impeding the gene normal transcription. Therefore, *SiMYB181* is a negative regulator of *BRB*. Nonetheless, it is still unclear how the duplicated motif affects the *SiMYB181* binding to the 5′-UTR of *BRB* in Hap2 genotypes.

BRB is a Lamiales-specific gene, and the duplicated motif in the 5′-UTR is under positive selection in sesame

The full-length genomic DNA sequence of *SIN_1025576* (*BRB*) gene consists of 417 bp long with a single exon. The ORF encodes a protein of 138 amino acids with a predicted molecular mass of 15.96 kDa which has no functional annotation. Searching for recognizable protein domains or signatures of conserved motifs in the protein sequence in publicly available databases did not result in any significant hits. We searched for *BRB* homologs within the plant kingdom through blastn of the NCBI non-redundant nucleotide sequence collection database (E-value < 10^{-10} and Identity > 70%). Two homologs (*LOC105953437* and *LOC105975223*) were only detected in *Erythranthe guttatus* (Phrymaceae) and one (*LOC111397704*) in *Olea europeae* (Oleaceae), none being functionally annotated. *Erythranthe guttatus* and *Olea europea* are close relatives of sesame, and all belong to the Lamiales order. This suggests that *BRB* is an orphan gene restricted to the Lamiales order. We further confirmed this finding by successfully cloning the homologous of *BRB* using genomic DNA of three wild sesame species (Pedaliaceae): *Ceratotheca sesamoides* (Genbank number: MN336257), *Sesamum radiatum* (Genbank number: MN336258) and *Sesamum alatum* (Genbank number:

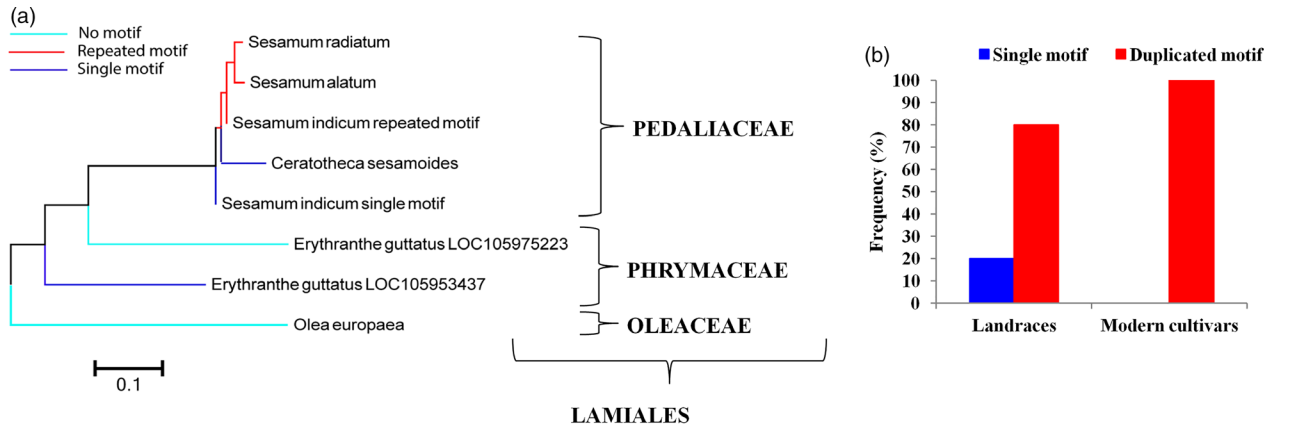

Figure 4 Evolutionary analysis of *BRB*. (a) *BRB* was only detected in few plant species of the order Lamiales. Analysis of the 5′-UTR of the *BRB* genes from different species showed the presence of no motif, single motif (Hap1) and duplicated motif (Hap2); (b) frequency of the presence of single motif (Hap1) and duplicated motif (Hap2) in sesame landraces and modern cultivars.

MN336259) (Figure 4a). The motif AACACACAC was present as single or duplicated forms only in the 5′-UTR of *LOC105953437* from *Erythranthe guttatus* and the homologous *BRB* from all wild sesame species. The homologous *BRB* genes have the same structure, similar size and conserved nucleotide sequences (Figure S6; Figure S7).

We further compared the frequency of the single/duplicated motif between landraces and modern cultivars to assess the impact of modern breeding at this locus. Intriguingly, we noticed that none of the modern cultivars harboured the single motif (Hap1) associated with the high root biomass trait while 20% of the landraces harboured the single motif (Figure 4b). This finding indicates that the duplicated motif (Hap2) was intensively selected by modern breeding; therefore, we hypothesized that Hap2 favours another desirable trait in the above-ground part of sesame.

Overexpression of *BRB* in *Arabidopsis* modulates root and shoot traits and drought stress response

In order to confirm the function of *BRB* in another plant system, we created *BRB*-overexpressing *Arabidopsis* mutants. There was not obvious difference in the main root length of wild-type (WT) and *BRB*-overexpressing lines (Figure 5a). However, the lateral root number, total lateral root length and root fresh weight were significantly reduced in the *BRB*-overexpressing lines as compared to WT plants under control growth conditions (Figure 5a,b,c,d). This result demonstrates that *BRB* modulates root biomass and has potential applications in other plant systems. Since root traits are crucial for plant response to drought stress, we further tested whether overexpression of *BRB* will affect the drought response in *Arabidopsis* mutants. We observed that all mutant lines died just after 7 days of drought stress while no death was recorded in WT plants until 15 days (Figure 5e), indicating that *BRB* overexpression induces drought stress sensitivity.

Because we suspected that high expression of *BRB* favours another desirable trait in the above-ground part of sesame, we evaluated the phenotypes of *Arabidopsis* mutants at the reproductive stage. Interestingly, we observed that under normal growth conditions, mutant lines displayed clustered silique and short internodes (Figure 6a). In addition, they had significantly higher plant height, silique number, branch number and seed yield per plant as compared to WT plants (Figure 6b,c,d,e,f). These observations confirm that *BRB* expression level not only modulates root traits but also impacts on the plant above-ground part.

BRB is modulates the auxin pathway

BRB gene has no functional annotation. In order to understand the functional pathway involving *BRB*, we compared transcriptome of WT and *BRB*-overexpressing lines based on RNA-sequencing. A total of 4685 and 3579 differentially expressed genes (DEG) were identified in WT_vs_*BRB*-OE1 and WT_vs_*BRB*-OE2, respectively (Tables S10 and S11). Gene Ontology enrichment analysis of the DEGs revealed various enriched biological pathways such as response to wounding, response to cold, response to abscisic acid (ABA), response to water deprivation, response to chitin and response to sucrose (Figure 7a,b). These enriched pathways hint that *BRB* modulates plant hormone signal transduction. Therefore, we examined the DEGs related to plant hormone signal transduction. We detected 68 DEGs related to various hormones such as abscisic acid, auxin/Indole-3-acetic acid (AUX/IAA), brassinosteroid, cytokinin, ethylene, gibberellic acid, jasmonic acid and salicylic acid (Figure 7c; Table S12). Notably, AUX/IAA-related genes were the most altered by *BRB* overexpression in *Arabidopsis*, which were further validated by qRT-PCR analysis (Figure S8). Hence, we deduce that *BRB* modulates the auxin pathway with consequences on root and shoot traits.

Discussion

Deciphering the genetic architecture of root-related traits in crops is crucial for improving productivity and abiotic stress tolerance (Liu *et al*., 2018; Svacina *et al*., 2014; Wang *et al*., 2015). Genome-wide association study (GWAS) has been applied to unlock the genetic loci associated with root traits in crops such as wheat, maize, barley, rice (Abdel-Ghani *et al*., 2019; Li *et al*., 2017; Li *et al*., 2019; Wang *et al*., 2019) and extensively in the model plant species *Arabidopsis thaliana* (Niu *et al*., 2018; Ogura *et al*., 2019; Ristova *et al*., 2018). However, most genes discovered for root traits in crops so far have only minor contributions (Uga and Yano, 2013). Also, key findings on root architecture in *Arabidopsis* are not directly transferable to crops. Studies on minor crops and non-model plants have been largely neglected although they represent great sources of untapped

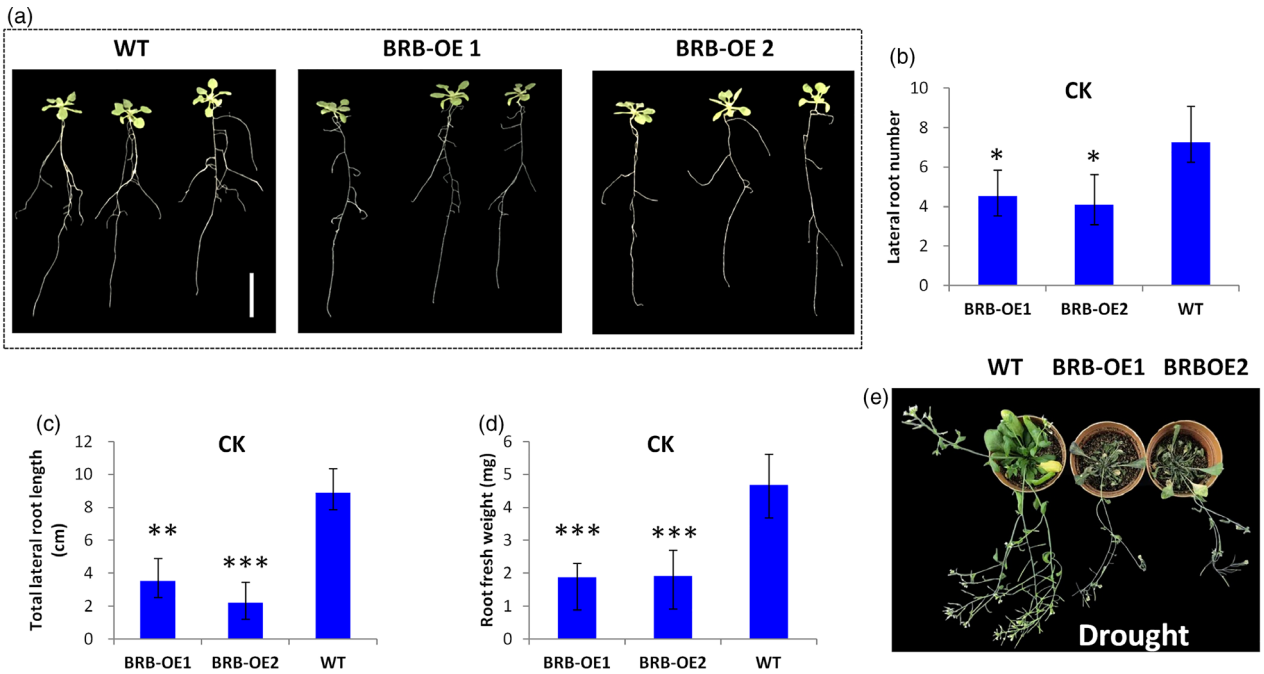

Figure 5 Functional analysis of *BRB* gene in *Arabidopsis thaliana*. (a) Phenotypes of 14 days old wild-type (WT) plants and *BRB*-overexpressing lines (*BRB*-OE1 and *BRB*-OE2). *BRB*-OE1 and *BRB*-OE2 display reduced root lateral root number. Bar = 1 cm; (b–d) lateral root number, total lateral root length and root fresh weight were significantly lower in *BRB*-OE1 and *BRB*-OE2 compared to WT; (e) phenotypes of 30 days old WT, *BRB*-OE1 and *BRB*-OE2 plants after 7 days drought stress. Error bars indicate the SD (*,**,*** significant difference at $P < 0.05$, 0.01, 0.001).

alleles for improving root traits in major crops (Mayes *et al*., 2012). Herein, we studied the non-model plant sesame and observed high variability and heritability for root traits, which is advantageous for GWAS implementation. Using a previously developed root traits phenotyping platform (Su *et al*., 2019) combined with high-quality genotyping data (Wei *et al*. 2015), we successfully dissected the genetic architectures of seven root traits using GWAS. With the strong and positive correlations between root traits, we identified several QTLs with concurrent effects which will be beneficial for multi-traits breeding in sesame as reported in jatropha (Sun *et al*., 2012). For instance, the QTL cluster *QTL11* located on LG7 encompassed five different QTLs and could be applied for simultaneously improving root dry weight, root length, root volume, root surface area and shoot dry weight traits in sesame. Similarly, we identified several major loci ($R^2 > 15\%$), indicating that breeding efforts for high root biomass can be simplified by pyramiding favourable alleles of major genes (Ye *et al*., 2019). Moreover, we reported for the first time 32 candidate genes for root traits in sesame. Homologs of some of these genes have been shown to control root traits in *Arabidopsis*. In *qRL5.1* detected for root length, the candidate gene identified was *SIN_1016062* and *Arabidopsis* loss-of-function mutant of its homologous gene (*AT2G25010*, MAIL1) developed short primary roots (Ühlken *et al*., 2014). Also, we identified the gene *SIN_1024379* in *qRSR9.1* for root–shoot ratio and its homolog in *Arabidopsis* (*AT2G30520*, RPT2) is involved in root phototropism (Sakai *et al*., 2000). The gene *AT1G02730* (*ATCSLD5*) was shown to be involved in normal root growth (Bernal *et al*., 2007), and Yoo *et al*. (2012) later demonstrated its function in root hair development. In this study, we identified in *qRV2.1* for root volume the candidate gene *SIN_1005391*, which is the homolog of *ATCSLD5*. Some candidate genes, however, have no clear functions related to root growth and development and require further functional investigations to uncover their exact contribution to root traits in sesame. One of these genes is *SIN_1025576* named as '*Big Root Biomass*' (*BRB*) in this study. *BRB* is an orphan gene restricted to the order Lamiales, and two independent association studies using different statistical models, populations and growth media demonstrated that a natural variation in the 5′-UTR of *BRB* is strongly associated (with a major contribution $R^2 > 30\%$) to root number and root dry weight traits in sesame. Furthermore, *BRB* overexpression in *Arabidopsis* confirmed not only its function in root biomass modulation but also highlighted the potential of this gene in non-Lamiales species. Orphan genes do not possess a known protein domains (Khalturin *et al*., 2009); therefore, their functions remain largely unknown (Arendsee *et al*., 2014). However, it has been proven that they are important players in key agronomic traits. For example, Yadeta *et al*. (2014) identified the Brassicaceae-specific gene *EWR1* conferring resistance to vascular wilt pathogens which was also functional in *Nicotiana benthamiana*, a member of the Solanaceae family. Also, the grass-specific gene *Ms2* conferring male sterility (Ni *et al*., 2017) and the Pooideae orphan gene *TaFROG* which modulates resistance to diseases (Perochon *et al*., 2015) have been disclosed.

More interestingly, we have discovered a novel motif 'AACACACAC' in the 5′-UTR of *BRB* and a polymorphism of the motif was found strongly associated with root biomass in sesame. Insertion–deletion of DNA sequences in the regulatory regions of key genes has been linked to different phenotypes. For

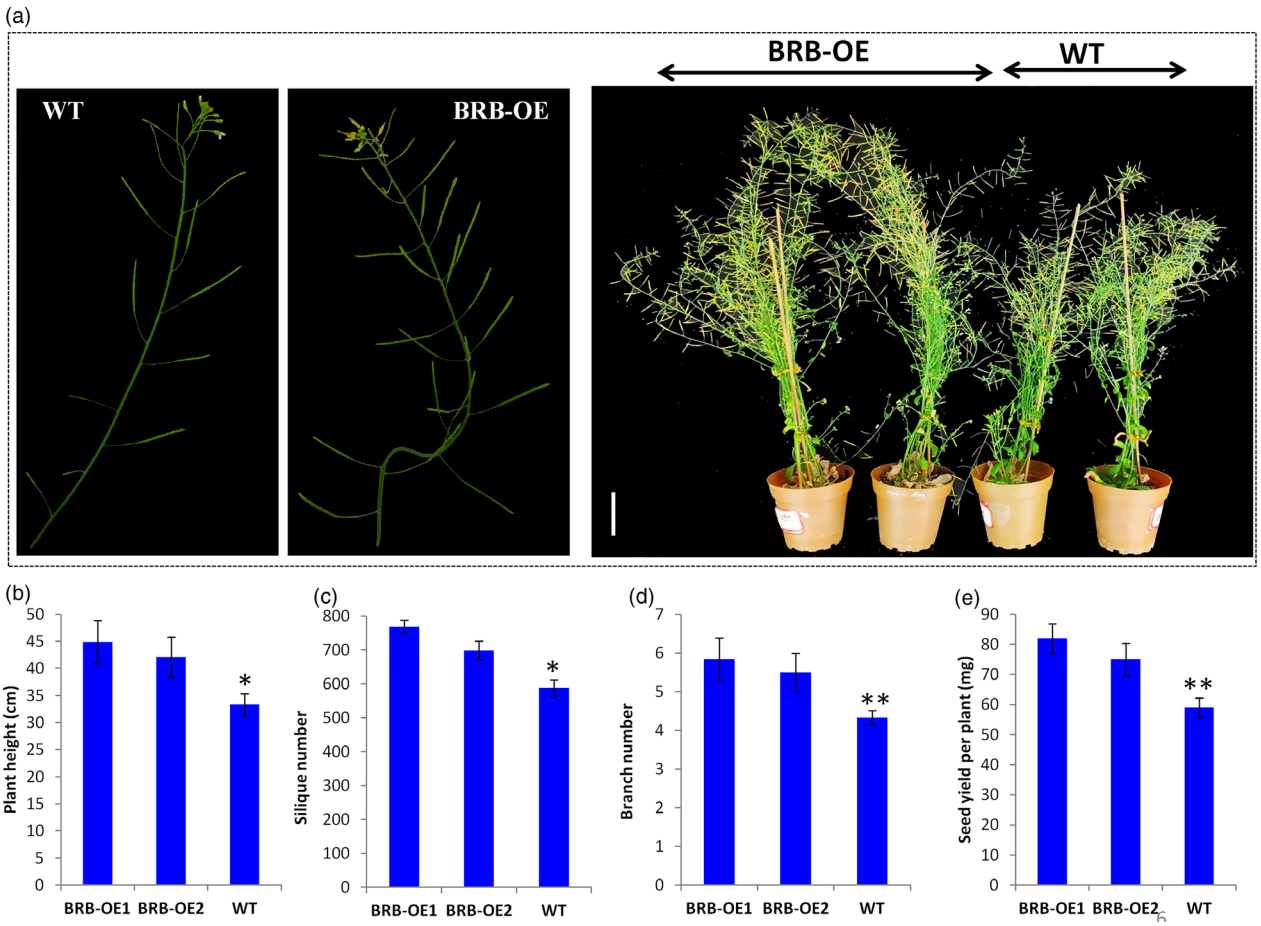

Figure 6 Overexpression of *BRB* affects shoot and yield-related traits. (a) Phenotypes of the main stem and shoot of the *BRB*-overexpressing (*BRB*-OE) lines and wild-type (WT) plants. Bar = 1 cm; (b–e) comparison of plant height, silique number, branch number and seed yield per plant between WT, *BRB*-OE1 and *BRB*-OE2 at maturation stage. Error bars indicate the SD (*,**significant difference at $P < 0.05$, 0.01).

example, the insertion of a terminal-repeat retrotransposon in miniature (TRIM) element in the promoter of *Ms2* is responsible for the anther-specific *Ms2* activation that confers male sterility in wheat (Xia *et al*., 2017). Ye *et al*. (2019) found a 8-bp InDel in the 5′-UTR of *SlbHLH59* which regulates ascorbate biosynthesis in tomato. Likewise, the presence of tandem repeats of a 23-bp forming a minisatellite-like structure in the upstream regulatory region of the gene *MYB10* directly controls anthocyanin biosynthesis in apple (Espley *et al*., 2009). In this study, we found that variations in the 5′-UTR of *BRB* regulate its expression level through interaction with the negative regulator *SiMYB181*. In particular, the presence of a single motif facilitates the binding of *SiMYB181*, repressing *BRB* expression while the duplicated motif prevents the binding of *SiMYB181*, leading to a normal transcription of *BRB*. It is also possible that the duplicated motif promotes the binding of an inhibitor of *SiMYB181*. Similar observations were reported by Espley *et al*. (2009) and Ye *et al*. (2019) who showed that variations in the repeat numbers or insertion/deletion of DNA sequences affect the binding of a transcription factor to the causative gene and consequently alter its expression level.

The phytohormone auxin regulates many aspects of root traits, including root elongation, gravitropism and lateral root development (Davies, 1995). It is well known that when inhibitors of auxin transport are applied to root, it impedes lateral root initiation (Casimiro *et al*., 2001). In this study, we observed that *BRB* mainly controls the root lateral number then impacting the total root weight. We deduce that *BRB* contributes to a normal polar auxin transport in sesame root. Several genes of the auxin/Indole-3-acetic acid (Aux/IAA) family have been characterized as controlling root traits. For example, *IAA1*, *IAA19*, *IAA17*, *IAA3*, *IAA7*, *IAA28*, *IAA14* control lateral root formation, root hair development and root gravitropic response (Fukaki *et al*., 2002; Leyser *et al*., 1996; Nagpal *et al*., 2000; Rogg *et al*., 2001; Tatematsu *et al*., 2004; Tian and Reed, 1999). In addition, functional characterization of these genes revealed overlapping, different or opposite effects on lateral root formation, indicating a complex regulation of auxin pathway to obtain a specific root phenotype. *BRB* has no signature of IAA proteins but significant changes were noticed in the auxin pathway in *BRB*-overexpressing *Arabidopsis* lines as compared to wild-type plants. Based on this observation, we conclude that *BRB* modulates the auxin pathway similar to reports of Ogura *et al*. (2019) who discovered an exocytosis factor, *EXO70A3*, as a modulator of the auxin pathway, affecting root system architecture and depth.

BRB overexpression in *Arabidopsis* showed altered shoot phenotype and improved seed yield under normal growth

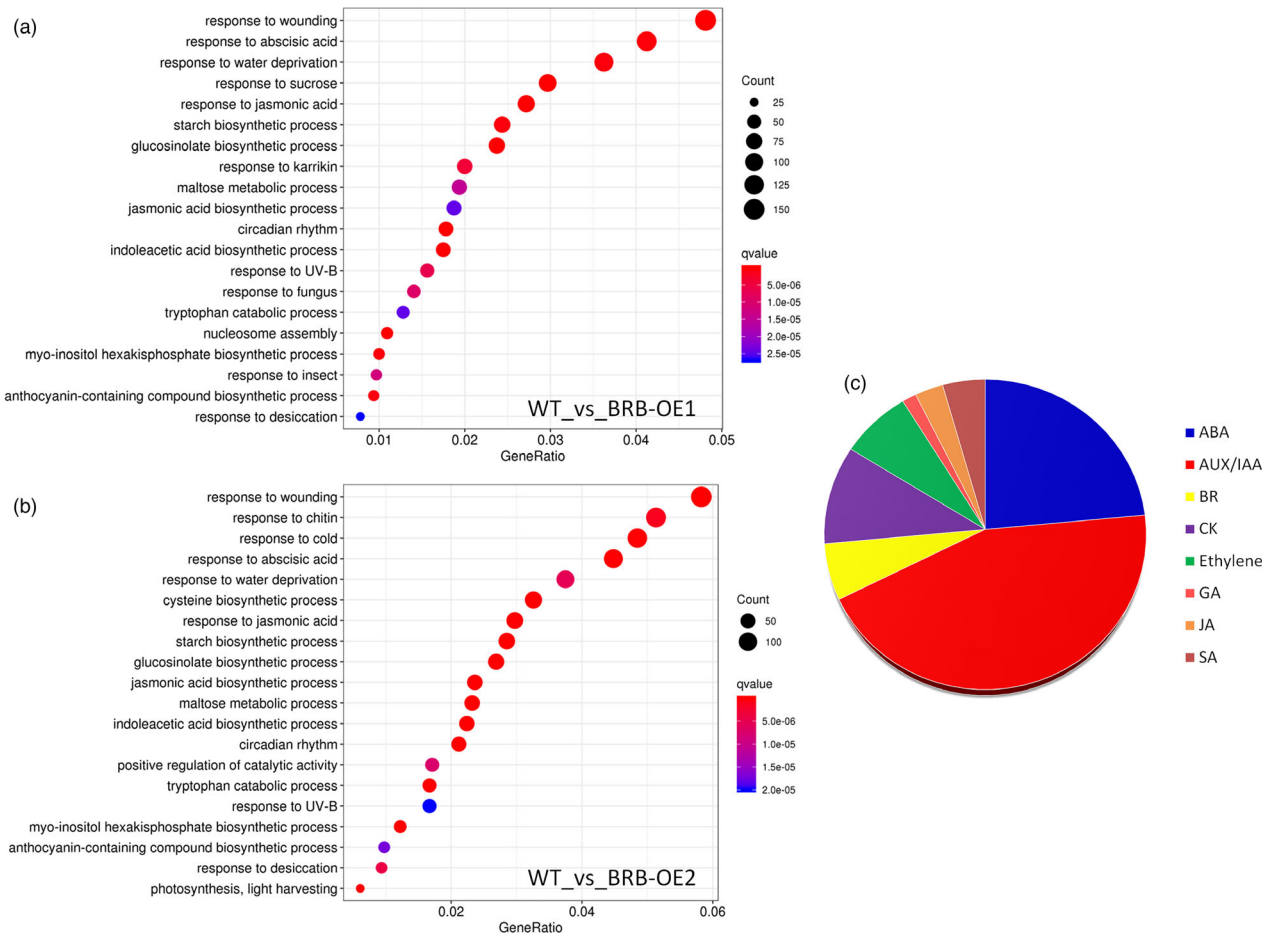

Figure 7 Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEG) between wild-type (WT) plants and *BRB*-overexpressing (*BRB*-OE) lines. (a) GO analysis of DEG between WT and *BRB*-OE1; (b) GO analysis of DEG between WT and *BRB*-OE2; (c) proportion of phytohormone-related genes detected within the DEGs. ABA: abscisic acid, AUX/IAA: auxin/Indole-3-acetic acid, BR: brassinosteroid, CK: cytokinin, ET: ethylene, GA: gibberellic acid, JA: jasmonic acid, and SA: salicylic acid.

condition. Our experimental set-up unfortunately does not allow us to record seed yield traits in our GWAS population in order to compare the two haplotypes. Nonetheless, based on our results in *Arabidopsis* mutants, we infer that *BRB* not only affects root traits but also impacts shoot and yield traits. It is well documented that auxins solely or in combination with other phytohormones shape plant architecture and regulate seed development, size and yield (Cao *et al.*, 2020; Gallavotti, 2013; Shirley *et al.*, 2019). Therefore, it is probable that high expression levels of *BRB* in above-ground tissues are beneficial to the plant productivity and may explain why the Hap1 (allele conferring high expression level of *BRB*) has been intensively selected by modern breeding in sesame. In addition, we observed that *Arabidopsis* lines overexpressing *BRB* were highly sensitive to drought stress. Lateral root development is important for plant responses to abiotic stresses such as drought (Comas *et al.*, 2013).

Functional characterization of orphan genes is challenging (Arendsee *et al.*, 2014), especially in non-model plant species such as sesame with limited tools for functional genomics and biotechnology. Hence, various unanswered questions remain on *BRB*. For example, after several attempts we were unable to clarify the subcellular localization of the product of this gene. Also, why *SiMYB181* cannot bind to the repeated motif located in the 5′-UTR of *BRB* (Hap2) and how *BRB* modulates the auxin pathway will need additional in-depth investigations.

In summary, we report the genetic loci associated with various root traits in sesame and propose several candidate genes. *BRB*, a gene containing significant SNPs associated with root number and root dry weight, has been functionally investigated and our results confirm that it controls root biomass (Figure 8). Besides, *BRB* appears to affect differentially shoot traits, seed yield and drought stress response, depending on its expression levels. Further understanding of the way to control the expression levels of *BRB* either through its regulator *SiMYB181* or the novel motif discovered in the 5′-UTR region will provide crucial functional tool for developing new sesame lines with combined enhanced attributes, including dense root system, high seed yield, improved shoot traits and tolerance to abiotic stress. Beyond sesame, it is envisioned that *BRB* might have a significant contribution to the improvement of major crops.

Methods

Plant materials

We extracted 327 sesame (*Sesamum indicum* L.) accessions comprising landraces and modern cultivars (Table S1) from our

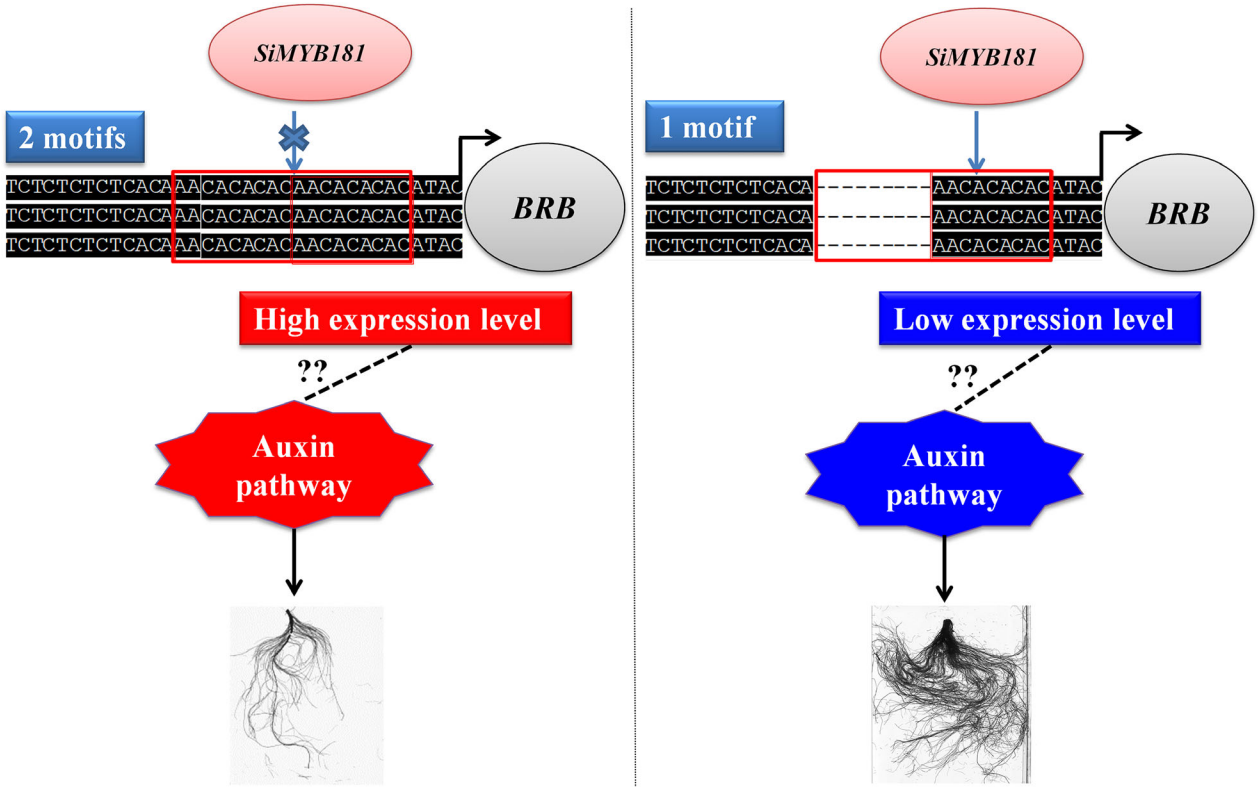

Figure 8 Proposed hypothetical model. The presence of the duplicated motif AACACACAC in the 5′-UTR of *BRB* gene impedes the binding of the negative regulator *SiMYB181*, leading to a normal transcription of *BRB*. In contrast, the presence of the single motif AACACACAC in the 5′-UTR of *BRB* gene facilitates the binding of *SiMYB181*, leading to a weak transcription of *BRB*. High and low expression levels of *BRB* differentially affect the auxin pathway to reduce and increase root biomass, respectively, through a yet undetermined mechanism.

association panel established at the Oil Crops Research Institute Chinese Academy of Agricultural Sciences (OCRI-CAAS) (Wei *et al*., 2015). Based on the root phenotyping data, two contrasting genotypes for root biomass (G430 with a big root biomass and G441 with a small root biomass) were selected to develop 197 $F_{2:3}$ population. In addition, three wild sesame including *Sesamum alatum*, *Sesamum radiatum* and *Ceratotheca sesamoides* were obtained from the Sesame Germplasm Resource Preserving Center of OCRI-CAAS.

Growth conditions of the sesame association panel

The association panel was grown hydroponically in blue plastic basins (34 × 26 × 12 cm, length × width × height) containing modified half-strength Hoagland's solution (Hoagland and Arnon, 1950) following technical descriptions of Su *et al*. (2019). Twenty-four seedlings of six accessions (four replicated seedlings for each accession) were grown in each basin. The experiment was led under controlled environmental conditions with the temperature and relative humidity kept at 35 °C and 60%, respectively, and under long-day condition (16/8 h day/night). A completely randomized design was employed with three replications, resulting in a total of ~ 3600 seedlings for each experiment. Plants were grown for a week in ¼ strength Hoagland nutrient solution and subsequently transferred into ½ strength nutrient solution for three additional weeks. The nutrient solution was renewed once a week, and the whole experiment was replicated over three different time periods during 2016–2017. At the end of each experiment, the whole root was delicately separated from the shoot, and both samples were placed into paper bags for later use.

Growth conditions of the $F_{2:3}$ population

Sterile seeds of the 197 $F_{2:3}$ population and the two parents were grown in pots (25 cm diameter and 45 cm depth) filled with 6 kg of experimental soil composed of 1.5 kg vermiculite, 2 kg soil, 2 kg sand and 0.5 kg nutritive soil as detailed by Su *et al*. (2019). The experiment was led under controlled environmental conditions with the temperature and relative humidity kept at 35 °C and 60%, respectively, and under long-day condition (16/8 h day/night) at a PPFD of 200 µmol/m^2/s. Plants were watered regularly to keep normal growth. The experiment was laid out in a completely randomized design with four replications and single plant per pot. The root of 4-week-old plants was delicately separated from the shoot, thoroughly washed with tap water, and both samples were put into paper bags for later use.

Root image scanning and data acquisition

After harvest, root samples were quickly scanned using a desktop scanner (EPSON Perfection V800 Photo, EPSON, Carson, CA, USA). Intact root systems were spread to minimize overlaps in a tray (A4 size) with a transparent glass bottom which was placed on the scanner. Next, the root systems were scanned to produce

digital images in uncompressed tagged image file (TIF) format (400 dpi). Root images were analysed using the WinRHIZO Pro software (Regent Instruments Inc., Quebec, QC, Canada). Four root data including the total root length (RL, cm), root surface area (SA, cm^2), root volume (RV, cm^3) and root number (RN) were extracted. Shoots and roots were oven dried at 80 °C for 72 h and weighed to determine shoot dry weight (SDW, g) and root dry weight (RDW, g). Root–shoot ratio (RSR) data were then estimated. Broad-sense heritability (H^2) was estimated following descriptions of Dossa *et al*. (2019a).

GWAS implementation

To perform the genome-wide association analysis (GWAS), the best linear unbiased prediction (BLUP) value of each phenotypic trait was computed using mixed linear model (MLM) in R 3.1 (R Core Team, 2015). Next, we extracted from the sesame HapMap project (www.ncgr.ac.cn/SesameHapMap/), one million high-quality SNPs with a minor allele frequency ≥ 0.05 and a good coverage of the whole genome (Wei *et al*., 2015). GWAS was conducted based on the mixed model with the SNP data, the kinship matrix, population structure (first three PCA components) using the EMMAX package (Kang *et al*., 2010).

Suggestive *P* value threshold was set at $P = 1 \times 10^{-5}$ to identify significant SNPs because root traits are complex polygenic and only small effects of the individual underlying loci were expected. A similar suggestive *P* value was previously used for dissecting complex traits in sesame (Li *et al*., 2018; Dossa *et al*., 2019a). For clustered significant SNPs, the peak SNP was defined as the one having the lowest *P* value in the linkage disequilibrium (LD) region. The LD decay in sesame has been estimated at ~ 88 kb (Wei *et al*., 2015; Dossa *et al*., 2019a), and LD windows upstream and downstream of the peak SNPs were defined as quantitative trait loci (QTL). To estimate the phenotypic variance explained (PVE) of each peak SNP, the value R^2 derived from linear regressions were calculated in the R package (R Core Team, 2015).

BRB gene sequencing and candidate gene association analysis

The full coding sequence (417 bp), 292 bp upstream of the translation start codon and 180 bp downstream of the stop codon of the gene *BRB* (*SIN_1025576*) were sequenced in the 197 $F_{2:3}$ population and the two parents using the primers A1 (Table S13). With the DNAMAN software (Lynnon Biosoft, Quebec, Canada), we assembled the sequences and the polymorphisms including SNPs and insertion–deletion (MAF ≥ 0.05) were detected using the DNASP software v6.11.01 (Rozas *et al*., 2017). Next, the association between the polymorphisms and two root traits (RDW and RN) was tested using TASSEL5.0 (Bradbury *et al*., 2007) based on the general linear model. Only significant variants with $P \leq 0.001$ were retained.

Gene co-expression analysis

A total of 60 root RNA-seq datasets generated from two sesame genotypes at 10 time points (0, 4, 8, 12, 16, 20, 24, 40, 44, 48 h beginning 15 days after the initiation of flowering) of root growth under control condition released by our group (NCBI SRA accession number: SRP181800) (Dossa *et al*., 2019a; Wang *et al*., 2020) were employed for gene co-expression analysis according to descriptions of Lv *et al*. (2019). The analysis was conducted with the WGCNA package version: 1.61 (Langfelder and Horvath, 2008) based on the normalized $\log_2$-transformed FPKM matrix. The co-expressed module containing the *BRB* gene (*SIN_1025576*) was extracted, and GO/KEGG enrichment analyses of the module were performed using clusterProfiler version 3.8.

Quantitative real-time PCR

The qRT-PCR analysis of target genes was performed according to descriptions by Dossa *et al*. (2019a) using the ChamQ SYBR qPCR Master Mix (Vazyme Biotec, Nanjing China) on a Light Cycler 480 II (Roche, Basel, Switzerland). The relative expression levels of target genes were normalized to the expression levels of two endogenous genes: *SiActin 7* and *SiHistone 3* for sesame and one endogenous gene *AtActin 2* for *Arabidopsis* using the primers A2-A32, (Table S13). Data are presented as relative transcript level based on the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) with three technical and three biological replicates. For RT-PCR, high-quality RNA was extracted from various tissues using the EASYspin Plus kit (Aidlab, Beijing, China) according to the manufacturer's instructions and reverse transcribed using the Superscript III reverse transcription kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The primer A2 (Table S13) was used to amplify *BRB* gene, and PCR products were visualized on 2% agarose gel.

Yeast one-hybrid assay

The 5′-UTR sequences from Hap1 and Hap2 containing a single and repeated AACACACAC motif, respectively, were isolated by PCR using the primer A33 (Table S13). Further, we synthesized two mutated synthetic 5′-UTR sequences by switching the alleles at the locus SNP5022842 (C/G). The four sequences were cloned into the bait vector pHIS2 (Clontech, Palo Alto, California, USA) between EcoRI/SmaI and EcoRI/SacI sites. Next, the complete CDS of the transcription factors (TF) co-expressed with *BRB* were amplified by the primers A34-A37 (Table S13) and cloned into the prey vector pGADT7-Rec2 (Clontech, Palo Alto, California, USA) using EcoRI/XhoI sites. Then, the co-transformed yeasts strain Y187 containing the bait and prey were cultivated on the SD/-Leu/-Trp/-His selective media supplemented with 0 or 3 mM 3-amino-1,2,4-triazole (3-AT) for 3 days according to the instructions for the Matchmaker™ Gold Yeast One Hybrid System (Clontech, Palo Alto, California, USA). Yeasts co-transformed with pGADT7-Rec2 −53 (the pGADT7-Rec2-53 plasmid containing a murine p53 and GAL4 AD domain fusion gene) and p53HIS2 (the pHIS2 harbouring three tandem copies of p53 recognized DNA consensus motifs) was used as positive control. The negative control was pGADT7-Rec2-TF and p53HIS2 co-transformation. The interaction between prey and bait was observed according to the growth of the yeast transformants in a series of 10-fold dilution.

Arabidopsis transgenics experiment

To functionally characterize the *BRB* gene in *Arabidopsis thaliana*, we isolated the coding sequence by PCR from G546 (Hap1) using the primer A38 (Table S13). Gene cloning and *Arabidopsis* transformation were performed as described by Dossa *et al*. (2019a) using the pCAMBIA 1301s vector. Positive T1 plants were screened on Murashige and Skoog (MS) medium containing 1% agar, and 1% sucrose and 50 μg/mL hygromicin and further confirmed by PCR, RT-PCR and β-glucuronidase (GUS) staining. Three independent T3 transgenic homozygous lines were used for

the evaluation of root and yield traits and RNA-seq analysis. We also generated *BRB*-promoter:GUS expression cassette, which contains the GUS reporter gene under the control of the promoter-containing DNA fragment from *BRB* (292 bp) in pCAMBIA-1381z binary vector. The resulting recombinant vector was then introduced into *Agrobacterium tumefaciens* strain LBA4404 and transferred into *Arabidopsis*. Histochemical localization of GUS activity in the generated transgenic *Arabidopsis* lines was performed by incubating whole seedlings with X-Gluc (5-bromo-4-chloro-3-indolyl-β-d-glucuronide) as described by Jefferson *et al*. (1987).

Evaluation of the transgenic lines

Seeds of wild-type (WT) and three T3 lines were germinated on solid MS agar medium. The seeds were stratified for 2 days in the dark at 4 °C and then transferred to growth chamber at 22 °C under long-day conditions (16 h light/8 h dark) with light intensity of 120–150 μmol/m^2 sec and relative humidity of 70%. Thirty seedlings (10 days old) of WT and *BRB*-overexpressing lines were transferred into new MS plates. Plates were placed vertically, and after seven days, lateral root length, lateral root number and estimated total lateral root length were obtained by summing up lengths of all lateral roots. In addition, root fresh weight of bulked seedling was recorded. At the reproductive stage, drought stress was applied by withholding water supply for 15 days and the number of survived plants out of 50 investigated plants was recorded. In parallel, 50 plants of each line were grown under well-watered conditions, and plant height, branch number, silique number per plant and seed yield per plant data were recorded at the maturity stage. The entire experiment was repeated two more times with four replicates in each experiment for statistical analysis.

Transcriptome sequencing and analysis

Total RNAs were extracted from whole plants of two independent *Arabidopsis* mutant lines and WT plants in two biological replicates using the EASYspin Plus kit (Aidlab) according to the manufacturer's instructions and reverse transcribed using the Superscript III reverse transcription kit (Invitrogen) according to the manufacturer's instructions. The cDNA libraries were pair-end sequenced on an Illumina Hiseq 4000 platform according to the methods described by Dossa *et al*. (2017b). The raw data were processed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to filter out adapters and low-quality sequences. Thereafter, the clean reads were mapped to the reference genome of *Arabidopsis* (TAIR10) using HISAT (Kim *et al*., 2015). The gene expression level for each sample expressed as fragments per kilobase of transcript per million fragments mapped (FPKM) was estimated employing the RSEM package v1.3.0 (Li and Dewey, 2011). Comparison of the gene expression levels between samples helped to identify the differentially expressed genes with the following parameters: Fold change ≥ 2, Probability ≥ 0.8, false discovery rate-adjusted P value < 0.05 (Tarazona *et al*., 2011). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses for the DEGs were performed using clusterProfiler version 3.8.

Statistical analysis

All the data were analysed with the R software (www.r-project.org). One-way analysis of variance was performed by comparing each transgenic line to the wild-type plants. This was followed by Tukey HSD test for mean comparison.

Acknowledgement

We thank all students and technical staff of the sesame genomics laboratory of OCRI-CAAS for their valuable help during experiment set-up and data collection. This work was supported by China Agriculture Research System (CARS-14), Central Public-interest Scientific Institution Basal Research Fund (1610172018007), and the Agricultural Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (CAAS-ASTIP-2016-OCRI).

Conflict of interest

They declare no conflict of interest.

Author contribution

Komivi Dossa, Xiurong Zhang and Jun You conceived and designed this project. Komivi Dossa, Rong Zhou, Donghua Li, Aili Liu, Lu Qin, Marie A. Mmadi, Ruqi Su, Yujuan Zhang, Jianqiang Wang and Yuan Gao conducted the experiments and collected data. Komivi Dossa performed data analysis and drafted the manuscript. Xiurong Zhang and Jun You supervised the study, provided funding and technical support, and revised the drafts of the manuscript. All authors have read and approved the final version of this manuscript.

Data availability statement

Sequence data supporting this study are available at NCBI GenBank under the accession numbers MN336257, MN336258 and MN336259. Transcriptome data are available at NCBI SRA under the accession numbers SRP181800, PRJNA638763 and PRJNA552167. The authors also declare that all other data that support the findings of this study are available within the manuscript and its Supplementary Files. Plant materials request could be sent to the corresponding authors.

References

Abdel-Ghani, A.H., Sharma, R., Wabila, C., Dhanagond, S., Owais, S.J., Duwayri, M.A., Al-Dalain, S.A. *et al*. (2019) Genome-wide association mapping in a diverse spring barley collection reveals the presence of QTL hotspots and candidate genes for root and shoot architecture traits at seedling stage. *BMC Plant Biol.* **19**, 216.

Alahmad, S., El Hassouni, K., Bassi, F.M., Dinglasan, E., Youssef, C., Quarry, G., Aksoy, A. *et al*. (2019) A major root architecture QTL responding to water limitation in Durum Wheat. *Frontiers in Plant Science* **10**, 436.

Anilakumar, K.R., Pal, A., Khanum, F. and Bawas, A.S. (2010) Nutritional, medicinal and industrial uses of sesame (*Sesamum indicum* L.) seeds. *Agriculturae Conspectus Science* **75**, 159–168.

Arendsee, Z.W., Li, L. and Wurtele, E.S. (2014) Coming of age: orphan genes in plants. *Trends Plant Sci.* **19**, 698–708.

Atkinson, J.A., Pound, M.P., Bennett, M.J. and Wells, D.M. (2019) Uncovering the hidden half of plants using new advances in root phenotyping. *Current opinion in biotechnology* **55**, 1–8. https://doi.org/10.1016/j.copbio.2018.06.002

Bernal, A.J., Jensen, J.K., Harholt, J., Sørensen, S., Moller, I., Blaukopf, C., Johansen, B. *et al*. (2007) Disruption of *ATCSLD5* results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in Arabidopsis. *Plant Journal* **52**, 791–802.

Beyer, S., Daba, S., Tyagi, P., Bockelman, H., Brown-Guedira, G., IWGSC, and Mohammadi, M. (2019) Loci and candidate genes controlling root traits in wheat seedlings-a wheat root GWAS. *Functional & integrative genomics* **19** (1), 91–107. https://doi.org/10.1007/s10142-018-0630-z

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635.

Cao, J., Li, G., Qu, D., Li, X. and Wang, Y. (2020) Into the seed: Auxin controls seed development and grain yield. *Int. J. Mol. Sci.* **21**, 1662.

Casimiro, I., Marchant, A., Bhalerao, R.P. *et al*. (2001) Auxin transport promotes Arabidopsis lateral root development. *Plant Cell* **13**, 843–852.

Comas, L.H., Becker, S.R., Cruz, V.M., Byrne, P.F. and Dierig, D.A. (2013) Root traits contributing to plant productivity under drought. *Frontiers in Plant Science* **4**, 442.

Courtois, B., Audebert, A., Dardou, A., Roques, S., Ghneim-Herrera, T., Droc, G., Frouin, J., Rouan, L., Gozé, E., Kilian, A., Ahmadi, N. and Dingkuhn, M. (2013) Genome-wide association mapping of root traits in a japonica rice panel. *PloS one* **8**(11), e78037. https://doi.org/10.1371/journal.pone.0078037

Davies, P.J. (1995) *Plant Hormones: Physiology, Biochemistry and Molecular Biology*, 2nd ed. Dordrecht, The Netherlands: Kluwer.

Dossa, K., Li, D., Wang, L., Zheng, X., Liu, A., Yu, J., Wei, X. *et al*. (2017b) Transcriptomic, biochemical and physio-anatomical investigations shed more light on responses to drought stress in two contrasting sesame genotypes. *Sci. Rep.* **7**, 8755.

Dossa, K., Li, D., Yu, J., Wang, L., Zhang, Y., You, J., Zhou, R. *et al*. (2019a) The genetic basis of drought tolerance in the high oil crop *Sesamum indicum*. *Plant Biotechnol. J.* **17**, 1788–1803.

Dossa, K., Yehouessi, L.W., Likeng-Li-Ngue, B.C., Diouf, D., Liao, B., Zhang, X., Cissé, N. *et al*. (2017a) Comprehensive screening of some west and central African sesame genotypes for drought resistance probing by agromorphological, physiological, biochemical and seed quality traits. *Agronomy* **7**, 83.

Dossa, K., You, J., Wang, L., Zhang, Y., Li, D., Zhou, R., Yu, J. *et al*. (2019b) Transcriptomic profiling of sesame during waterlogging and recovery. *Scientific Data* **6**, 204.

Downie, H.F., Adu, M.O., Schmidt, S., Otten, W., Dupuy, L.X., White, P.J. and Valentine, T.A. (2015) Quantifying the rhizosphere using image analysis. *Plant Cell Environ* **38**, 1213–1232. https://doi.org/10.1111/pce.12448

Espley, R.V., Brendolise, C., Chagné, D., Kutty-Amma, S., Green, S., Volz, R., Putterill, J. *et al*. (2009) Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. *Plant Cell* **21**, 168–183.

Fukaki, H., Tameda, S., Masuda, H. and Tasaka, M. (2002) Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT/IAA14* gene of Arabidopsis. *Plant Journal* **29**, 153–168.

Gallavotti, A. (2013) The role of auxin in shaping shoot architecture. *J. Exp. Bot.* **64**, 2593–2608.

Gifford, M.L., Banta, J.A., Katari, M.S., Hulsmans, J., Chen, L., Ristova, D., Tranchina, D., Purugganan, M.D., Coruzzi, G.M. and Birnbaum, K.D. (2013) Plasticity regulators modulate specific root traits in discrete nitrogen environments. *PLoS genetics* **9**(9), e1003760. https://doi.org/10.1371/journal.pgen.1003760

Gloaguen, R.M., Couch, A., Rowland, D.L., Bennett, J., Hochmuth, G., Langham, D.R. and Brym, Z.T. (2019) Root life history of non-dehiscent sesame (*Sesamum indicum* L.) cultivars and the relationship with canopy development. *Field. Crop. Res.* **241**, 107560.

Hoagland, D.R. and Arnon, D.I. (1950) The water-culture method of growing plants without soil. *Circular, California Agricultural Experiment Station* **347**, 1–47.

Hochholdinger, F. and Tuberosa, R. (2009) Genetic and genomic dissection of maize root development and architecture. *Current opinion in plant biology* **12** (2), 172–177. https://doi.org/10.1016/j.pbi.2008.12.002

Islam, F., Gill, R.A., Ali, B., Farooq, M.A., Xu, L., Najeeb, U. and Zhou, W. (2016) Sesame. In *Breeding oilseed crop for sustainable production: opportunities and constraints*. (Gupta, S.K., ed), pp. 135–147.Cambridge, MA: Academic Press.

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: betaglucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal* **6**, 3901–3907.

Jia, Z., Liu, Y., Gruber, B.D., Neumann, K., Kilian, B., Graner, A. and von Wirén, N. (2019) Genetic Dissection of Root System Architectural Traits in Spring Barley. *Frontiers in plant science* **10**, 400. https://doi.org/10.3389/fpls.2019.00400

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Sabatti, C. *et al*. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348–354.

Khalturin, K., Hemmrich, G., Fraune, S., Augustin, R. and Bosch, T.C.G. (2009) More than just orphans: are taxonomically-restricted genes important in evolution? *Trends Genet.* **25**, 404–413.

Kim, D., Langmead, B. and Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360.

Kitomi, Y., Nakao, E., Kawai, S., Kanno, N., Ando, T., Fukuoka, S., Irie, K., & Uga, Y. (2018). Fine Mapping of QUICK ROOTING 1 and 2, Quantitative Trait Loci Increasing Root Length in Rice. *G3: Genes, Genomes, Genetics* (Bethesda, MD), **8**(2), 727–735. https://doi.org/10.1534/g3.117.300147

Kobayashi, Y., Sadhukhan, A., Tazib, T., Nakano, Y., Kusunoki, K., Kamara, M., Chaffai, R. *et al*. (2016) Joint genetic and network analyses identify loci associated with root growth under NaCl stress in Arabidopsis thaliana. *Plant, Cell & Environment* **39**(4), 918–934. https://doi.org/10.1111/pce.12691

Lakhanpaul, S., Singh, V., Kumar, S., Bhardwaj, D. and Bhat, K.V. (2012). Sesame: overcoming the abiotic stresses in the queen of oilseed crops. In *improving crop resistance to abiotic stress* (eds Tuteja, N., Gill, S.S., Tiburcio, A.F. and Tuteja, R., (pp. 1251-1283). Weinheim, Germany: Wiley Blackwell. https://doi.org/10.1002/9783527632930.ch48.

Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.

Langham, D.R. (2007) Phenology of sesame. In *Issues in new crops and new uses*. (Janick, J. and Whipkey, A., eds), pp. 144–182.Alexandria, VA: ASHS Press.

Leyser, H.M.O., Pickett, F.B., Dharmasiri, S. and Estelle, M. (1996) Mutations in the *AXR3* gene of Arabidopsis result in altered auxin response including ectopic expression from the *SAURAC1* promoter. *Plant J.* **10**, 403–413.

Li, B. and Dewey, C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323.

Li, D., Dossa, K., Zhang, Y., Wei, X., Wang, L., Zhang, Y., Liu, A., Zhou, R. and Zhang, X. (2018) GWAS Uncovers Differential Genetic Bases for Drought and Salt Tolerances in Sesame at the Germination Stage. *Genes* **9**(2), 87. https://doi.org/10.3390/genes9020087

Li, X., Guo, Z., Lv, Y., Cen, X., Ding, X., Wu, H., Li, X. *et al*. (2017) Genetic control of the root system in rice under normal and drought stress conditions by genome-wide association study. *PLoS Genet.* **13**, e1006889.

Li, L., Peng, Z., Mao, X., Wang, J., Chang, X., Reynolds, M. and Jing, R. (2019) Genome-wide association study reveals genomic regions controlling root and shoot traits at late growth stages in wheat. *Ann. Bot.* **124**, 993–1006.

Liu, K., He, A., Ye, C. *et al*. (2018) Root morphological traits and spatial distribution under different nitrogen treatments and their relationship with grain yield in super hybrid rice. *Sci. Rep.* **8**, 1–9.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{\Delta\Delta Ct}$ method. *Methods* **25**, 402–408.

Lv, Y., Xu, L., Dossa, K., Zhou, K., Zhu, M., Xie, H., Tang, S. *et al*. (2019) Identification of putative drought-responsive genes in rice using gene co-expression analysis. *Bioinformation* **15**, 480–489.

Lynch, J.P. (2013) Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. *Annals of botany* **112**(2), 347–357. https://doi.org/10.1093/aob/mcs293

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M.A., Colalongo, M.C., Stefanelli, S. *et al*. (2016) Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* **67**, 1161–1178.

Mai, C.D., Phung, N.T., To, H.T., Gonin, M., Hoang, G.T., Nguyen, K.L., Do, V.N., Courtois, B. and Gantet, P. (2014) Genes controlling root development in rice. *Rice* **7**, 30. https://doi.org/10.1186/s12284-014-0030-5

Meijón, M., Satbhai, S.B., Tsuchimatsu, T. and Busch, W. (2014) Genome-wide association study using cellular traits identifies a new regulator of root development in Arabidopsis. *Nature genetics* **46**, 77–81.

Mayes, S., Massawe, F.J., Alderson, P.G., Roberts, J.A., Azam-Ali, S.N. and Hermann, M. (2012) The potential for underutilized crops to improve security of food production. *J. Exp. Bot.* **63**, 1075–1079.

Miyahara, Y., Hibasami, H., Katsuzaki, H., Imai, K. and Komiya, T. (2001) Sesamolin from sesame seed inhibits proliferation by inducing apoptosis in human lymphoid leukemia Molt 4B cells. *Int. J. Mol. Med.* **7**, 369–371.

Mmadi, M.A., Dossa, K., Wang, L., Zhou, R., Wang, Y., Cisse, N., Sy, M.O. *et al.* (2017) Functional characterization of the versatile MYB gene family uncovered their important roles in plant development and responses to drought and waterlogging in sesame. *Genes* **8**, 362.

Moose, S.P. and Mumm, R.H. (2008) Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* **147**, 969–977.

Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M. and Reed, J.W. (2000) *AXR2* encodes a member of the Aux/IAA protein family. *Plant Physiol.* **123**, 563–573.

Ni, F., Qi, J., Hao, Q., Lyu, B., Luo, M.-C., Wang, Y., Chen, F. *et al.* (2017) Wheat *Ms2* encodes for an orphan protein that confers male sterility in grass species. *Nat. Commun.* **8**, 15121.

Niu, Y., Chen, P., Zhang, Y., Wang, Z., Hu, S., Jin, G., Tang, C. *et al.* (2018) Natural variation among Arabidopsis thaliana accessions in tolerance to high magnesium supply. *Sci. Rep.* **8**, 13640.

Noguchi, T., Ikeda, K., Sasaki, Y., Yamamoto, J., Seki, J., Yamagata, K., Nara, Y. *et al.* (2001) Effects of vitamin E and sesamin on hypertension and cerebral thrombogenesis in stroke-prone spontaneously hypertensive rats. *Hypertens. Res.* **24**, 735–742.

Ogura, T., Goeschl, C., Filiault, D., Mirea, M., Slovak, R., Wolhrab, B., Satbhai, S.B. *et al.* (2019) Root system depth in Arabidopsis is shaped by *EXOCYST70A3* via the dynamic modulation of auxin transport. *Cell* **178**, 400–412.

Perochon, A., Jianguang, J., Kahla, A., Arunachalam, C., Scofield, S.R., Bowden, S., Wallington, E. *et al.* (2015) *TaFROG* encodes a Pooideae orphan protein that interacts with SnRK1 and enhances resistance to the mycotoxigenic fungus *Fusarium graminearum*. *Plant Physiol.* **169**, 2895–2906.

Ristova, D., Giovannetti, M., Metesch, K. and Busch, W. (2018) Natural genetic variation shapes root system responses to phytohormones in Arabidopsis. *Plant J.* **96**, 468–481.

Rogg, L.E., Lasswell, J. and Bartel, B. (2001) A gain-of-function mutation in *IAA28* suppresses lateral root development. *Plant Cell* **13**, 465–480.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. and Sánchez-Gracia, A. (2017) DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol.* **34**, 3299–3302.

Sakai, T., Wada, T., Ishiguro, S. and Okada, K. (2000) RPT2. A signal transducer of the phototropic response in Arabidopsis. *Plant Cell* **12**, 225–236.

Sankar, D., Sambandam, G., Ramakrishna, R.M. and Pugalendi, K.V. (2005) Modulation of blood pressure, lipid profiles and redox status in hypertensive patients taking different edible oils. *Clin. Chim. Acta* **355**, 97–104.

Shirley, N.J., Aubert, M.K., Wilkinson, L.G., Bird, D.C., Lora, J., Yang, X. and Tucker, M.R. (2019) Translating auxin responses into ovules, seeds and yield: Insight from Arabidopsis and the cereals. *J. Integr. Plant Biol.* **61**, 310–336.

Su, R., Zhou, R., Mmadi, M.A., Li, D., Qin, L., Liu, A., Wang, J. *et al.* (2019) Root diversity in sesame (*Sesamum indicum* L.): insights into the morphological, anatomical and gene expression profiles. *Planta* **250**, 1461–1474.

Sun, F., Liu, P., Ye, J., Lo, L.C., Cao, S., Li, L., Yue, G.H. *et al.* (2012) An approach for jatropha improvement using pleiotropic QTLs regulating plant growth and seed yield. *Biotechnol. Biofuels* **5**, 42.

Svacina, P., Streda, T. and Chloupek, O. (2014) Uncommon selection by root system size increases barley yield. *Agron. Sustain. Dev.* **34**, 545–551.

Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A. and Conesa, A. (2011) Differential expression in RNA-seq: a matter of depth. *Genome Res.* **21**, 2213–2223.

Tatematsu, K., Kumagai, S., Muto, H., Sato, A., Watahiki, M.K., Harper, R.M., Liscum, E. *et al.* (2004) *MASSUGU2* encodes *Aux/IAA19*, an auxin-regulated protein that functions together with the transcriptional activator *NPH4/ARF7* to regulate differential growth responses of hypocotyl and formation of lateral roots in *Arabidopsis thaliana*. *Plant Cell* **16**, 379–393.

Tian, Q. and Reed, J.W. (1999) Control of auxin-regulated root development by the *Arabidopsis thaliana SHY2/IAA3* gene. *Development* **126**, 711–721.

Topp, C.N., Iyer-Pascuzzi, A.S., Anderson, J.T., Lee, C.R., Zurek, P.R., Symonova, O., Zheng, Y., *et al.* (2013) 3D phenotyping and quantitative trait locus mapping identify core regions of the rice genome controlling root architecture. *Proceedings of the National Academy of Sciences of the United States of America* **110**(18), E1695–E1704. https://doi.org/10.1073/pnas.1304354110

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M., Hara, N., Kitomi, Y. *et al.* (2013) Control of root system architecture by *DEEPER ROOTING* 1 increases rice yield under drought conditions. *Nat. Genet.* **45**, 1097–1102.

Ühlken, C., Horvath, B., Stadler, R., Sauer, N. and Weingartner, M. (2014) *MAIN-LIKE1* is a crucial factor for correct cell division and differentiation in *Arabidopsis thaliana*. *Plant J.* **78**, 107–120.

Voss-Fels, K.P., Qian, L., Gabur, I., Obermeier, C., Hickey, L.T., Werner, C.R., Kontowski, S., *et al.* (2018) Genetic insights into underground responses to Fusarium graminearum infection in wheat. *Scientific reports* **8**(1), 13153. https://doi.org/10.1038/s41598-018-31544-w

Varshney, R.V., Hoisington, D.A. and Tyagi, A.K. (2006) Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* **24**, 490–499.

Wang, L., Dossa, K., You, J., Zhang, Y., Li, D., Zhou, R., Yu, J. *et al.* (2020) High-resolution temporal transcriptome sequencing unravels ERF and WRKY as the master players in the regulatory networks underlying sesame responses to waterlogging and recovery. *Genomics* **113**, 276–290.

Wang, Z.G., Ma, B.-L., Gao, J. and Sun, J. (2015) Effects of different management systems on root distribution of maize. *Can. J. Plant Sci.* **95**, 21–28.

Wang, H., Wei, J., Li, P., Wang, Y., Ge, Z., Qian, J., Fan, Y. *et al.* (2019). Integrating GWAS and gene expression analysis identifies candidate genes for root Morphology traits in maize at the seedling stage. *Genes* **10**, 773.

Wang, L., Yu, S., Tong, C., Zhao, Y., Liu, Y., Song, C., Zhang, Y., *et al.* (2014) Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis. *Genome biology* **15**(2), R39. https://doi.org/10.1186/gb-2014-15-2-r39

Wei, X., Liu, K., Zhang, Y., Feng, Q., Wang, L., Zhao, Y., Li, D., *et al.* (2015) Genetic discovery for oil production and quality in sesame. *Nature communications* **6**, 8609. https://doi.org/10.1038/ncomms9609

Xia, C., Zhang, L., Zou, C., Gu, Y., Duan, J., Zhao, G. *et al.* (2017) A TRIM insertion in the promoter of Ms2 causes male sterility in wheat. *Nat. Commun.* **8**, 15407.

Yadeta, K.A., Valkenburg, D.-J., Hanemian, M., Marco, Y. and Thomma, B.P.H.J. (2014) The Brassicaceae-specific *EWR1* gene provides resistance to vascular wilt pathogens. *PLoS One* **9**, e88230.

Ye, J., Li, W., Ai, G., Li, C., Liu, G., Chen, W., Wang, B. *et al.* (2019) Genome-wide association analysis identifies a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato. *PLoS Genet.* **15**, e1008149.

Yoo, C.M., Quan, L. and Blancaflor, E.B. (2012) Divergence and redundancy in *CSLD2* and *CSLD3* function during *Arabidopsis thaliana* root hair and female gametophyte development. *Front. Plant Sci.* **3**, 111.

Zhou, R., Dossa, K., Li, D., Yu, J., You, J., Wei, X. and Zhang, X. (2018) Genome-Wide Association Studies of 39 Seed Yield-Related Traits in Sesame (Sesamum indicum L.). *International journal of molecular sciences* **19**(9), 2794. https://doi.org/10.3390/ijms19092794.

Zurek, P.R., Topp, C.N. and Benfey, P.N. (2015) Quantitative trait locus mapping reveals regions of the maize genome controlling root system architecture. *Plant physiology* **167**(4), 1487–1496. https://doi.org/10.1104/pp.114.251751

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Frequency distribution of root traits in 327 *Sesamum indicum* accessions.

Figure S2 Pearson correlation analysis of the different root-related traits. Total root length (RL), root surface area (SA), root volume (RV), root number (RN), shoot dry weight (SDW), root dry weight (RDW) and root–shoot ratio (RSR).

Figure S3 Manhattan plots of genome-wide association studies using the mixed model for seven root-related traits. The top significant trait-associated loci were marked with a red star.
Figure S4 Cumulative contribution (Phenotypic variance explained = PVE) of the top 5 significant SNPs associated with the different root traits. Total root length (RL), root surface area (SA), root volume (RV), root number (RN), shoot dry weight (SDW), root dry weight (RDW) and root–shoot ratio (RSR).
Figure S5 Alignment of the 5′-UTR reverse complemented sequences from 25 sesame genotypes of Hap1 group and 25 sesame genotypes from Hap2 group. The blue box shows the conserved motif (AACACACAC) between Hap2 and Hap1 groups, while the red box shows the duplicated motif only observed in Hap2 group.
Figure S6 Gene structure of *BRB* from various species.
Figure S7 Alignment of the coding sequences of the *BRB* gene from various species.
Figure S8 qRT-PCR analysis of the auxin-related genes differentially expressed between wild-type plants and *BRB*-overexpressing *Arabidopsis* lines. Data are represented as mean ± SD of three biological replicates.

Table S1 Full list of the 327 accessions used in this study, their origin and their breeding status.
Table S2 Summary of descriptive statistics of the seven root traits investigated in this study. CV: coefficient of variation; H^2: broad-sense heritability; SD: standard deviation, Min: minimum value; Max: maximum value.
Table S3 SNPs significantly associated with root traits in *Sesamum indicum*. PVE: phenotypic variance explained, LG: linkage group. The QTL names in red are those significantly associated with various root traits.
Table S4 List of genes located within QTLs detected for root traits through GWAS in sesame.
Table S5 List of genes containing significant SNPs in coding or regulatory regions and their functional annotation.
Table S6 Candidate genes identified for root traits in sesame and their homologs in *Arabidopsis thaliana*.
Table S7 The 11 significant SNPs located in *BRB* (*SIN_1025576*).
Table S8 Association analysis of root number (RN) and root dry weight (RDW) traits in 197 $F_{2:3}$ sesame population. The markers highlighted in red are those significantly associated with RN an RDW.
Table S9 Genes co-expressed with *BRB* (*SIN_1025576*) and their annotation. The four transcription factors were highlighted in red.
Table S10 List of the differentially expressed genes identified between *BRB*-overexpressing line 1 (*BRB*-OE1) and wild-type (WT) plants.
Table S11 List of the differentially expressed genes identified between *BRB*-overexpressing line 2 (*BRB*-OE2) and wild-type (WT) plants
Table S12 Differentially expressed genes related to phytohormones between *BRB*-overexpressing lines (*BRB*-OE) and wild-type (WT) plants.
Table S13 Sequences of the different primers used in this study.

