

RESEARCH

Open Access



# Fine-mapping and transcriptome analysis of a candidate gene controlling plant height in *Brassica napus* L.

Xiaodong Wang<sup>1,2†</sup>, Ming Zheng<sup>1†</sup>, Hongfang Liu<sup>1</sup>, Liang Zhang<sup>1</sup>, Feng Chen<sup>2</sup>, Wei Zhang<sup>2</sup>, Shihang Fan<sup>1</sup>, Menlu Peng<sup>2</sup>, Maolong Hu<sup>2</sup>, Hanzhong Wang<sup>1</sup>, Jiefu Zhang<sup>2\*</sup> and Wei Hua<sup>1\*</sup>

## Abstract

**Background:** *Brassica napus* provides approximately 13–16% of global vegetable oil for human consumption and biodiesel production. Plant height (PH) is a key trait that affects plant architecture, seed yield and harvest index. However, the genetic mechanism of PH in *B. napus* is poorly understood.

**Results:** A dwarf mutant *df59* was isolated from a large-scale screening of an ethyl methanesulphonate-mutagenized rapeseed variety Ningyou 18. A genetic analysis showed that the dwarfism phenotype was controlled by one semi-dominant gene, which was mapped on C9 chromosome by quantitative trait loci sequencing analysis and designated as *BnaDwf.C9*. To fine-map *BnaDwf.C9*, two F<sub>2</sub> populations were constructed from crosses between conventional rapeseed cultivars (Zhongshuang 11 and Holly) and *df59*. *BnaDwf.C9* was fine-mapped to the region between single-nucleotide polymorphism (SNP) markers M14 and M4, corresponding to a 120.87-kb interval of the *B. napus* 'Darmor-bzh' genome. Within this interval, seven, eight and nine annotated or predicted genes were identified in "Darmor-bzh", "Ningyou 7" and "Zhongshuang 11" reference genomes, respectively. In addition, a comparative transcriptome analysis was performed using stem tips from Ningyou 18 and *df59* at the stem elongation stage. In total, 3995 differentially expressed genes (DEGs) were identified. Among them, 118 DEGs were clustered in plant hormone-related signal transduction pathways, including 81 DEGs were enriched in auxin signal transduction. Combining the results of fine-mapping and transcriptome analyses, *BnaC09g20450D* was considered a candidate gene for *BnaDwf.C9*, which contains a SNP that co-segregated in 4746 individuals. Finally, a PCR-based marker was developed based on the SNP in *BnaC09g20450D*.

**Conclusions:** The combination of quantitative trait loci sequencing, fine-mapping and genome-wide transcriptomic analysis revealed one candidate gene located within the confidence interval of 120.87-kb region. This study provides a new genetic resource for semi-dwarf breeding and new insights into understanding the genetic architecture of PH in *B. napus*.

**Keywords:** *Brassica napus* L., Plant height, Quantitative trait loci sequencing, Fine-mapping, Transcriptome analysis, Candidate gene, Molecular marker

\*Correspondence: jiefu\_z@163.com; huawei@oilcrops.cn

<sup>†</sup>Xiaodong Wang and Ming Zheng contributed equally to this work

<sup>1</sup>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

<sup>2</sup>Provincial Key Laboratory of Agrobiotechnology, Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture and Rural Affairs, Nanjing, China

## Background

Rapeseed (*Brassica napus*, AACC, 2n=38) is not only an important oilseed crop worldwide, but also an emerging biofuel crop. Rapeseed oil is an ideal vegetable oil for human consumption, because it contains ~65% oleic, ~20% linoleic, ~9% linolenic and a very low level



of stearic acid [1]. Vegetable oils are triglyceride sources for biodiesel production. In Europe, biodiesel has been produced mainly from rapeseed oil [2]. In addition, rape straw is an abundant lignocellulosic material for the production of liquid biofuel, particularly ethanol. In China, the cultivated area of rapeseed is ~67 million hectares, with an annual seed yield of ~4.5 million tons every year [3], and the collectable production of rape straw was 38.17 million tons in 2013 [4]. Plant height (PH) is a key trait that affects the plant architecture, seed yield, dry weight and harvest index [5]. Moreover, PH is the major contributor to lodging tolerance, a serious abiotic stress during rapeseed production. Lodging makes *B. napus* unsuitable for mechanical harvesting and causes dramatic decreases in yield and seed quality [6]. Therefore, it is important to understand the genetic bases of PH to breed new cultivars with an ideal plant architecture and to maximize *B. napus*' economic benefits as an oil and bioenergy crop.

Plant growth and development consist of very precise and complicated procedures. The molecular mechanisms regulating PH in the model plant *Arabidopsis thaliana* and rice (*Oryza sativa*) have been well recorded, and most of the genes are involved in phytohormone-related pathways [7–11]. Gibberellic acid (GA) is an essential endogenous regulator of PH. For example, the rice 'green revolution' gene *sd1* encodes GA20ox2, which is an oxidase enzyme involved in the biosynthesis of GA [12]. The wheat 'green revolution' gene *Rht-B1/Rht-D1* encodes a DELLA protein that acts as a negative regulator in the GA-signaling pathway [13]. Auxins play pivotal functions in developmental processes, because they are involved in controlling virtually every aspect of plant biology [14]. Indole-3-acetic acid (IAA) is the key auxin in most plants, and it is mainly biosynthesized from tryptophan (Trp) through a Trp-dependent pathway [14, 15]. Loss-of-function mutations in IAA-related Trp-dependent biosynthetic genes (e.g., *TAA1* or *YUCCA*) can seriously affect PH [16, 17]. In the auxin-signaling pathway, transport inhibitor resistant1/auxin-signaling F-box (TIR1/AFB), auxin/indole acetic acid proteins (Aux/IAAs) and auxin response factors (ARFs) are three key factors regulating auxin-induced gene expression [18, 19]. At low auxin levels, the transcriptional repressor Aux/IAA proteins and the corepressor TOPLESS interact with ARF proteins, resulting in auxin-induced gene repression. At high auxin levels, auxin interacts with TIR1, leading to the formation of TIR1/AFB-Aux/IAA-ARF complexes, then ubiquitin-ligase (E3) ubiquitinates AUX/IAAs, forming the ubiquitin 26S proteasome responsible for the degradation of AUX/IAAs [16, 18, 20]. Free ARF proteins can activate the transcription of auxin-response genes. Mutations that occur in the genes involved in auxin

signaling may result in typical dwarf phenotypes, such as *IAA7* in *B. napus* [21].

*Brassica napus* is a tetraploid species, which originated from a hybridization between *B. rapa* (AA,  $2n = 20$ ) and *B. oleracea* (CC,  $2n = 18$ ) around 7500 years ago [22]. In *B. napus*, many studies have focused on quantitative trait loci (QTLs) detection using linkage mapping or genome-wide association studies (GWASs), and hundreds of QTLs have been identified across all 19 chromosomes [23–31]. However, most QTLs explain a small percentage of the total phenotypic variance (PV) [23–31]. Only a few QTLs affecting PH have been fine-mapped and cloned in *B. napus*. Wang et al. obtained two dwarf mutants "*Bndwf1*" and "*Bndwf1/dcl1*", and fine-mapped the associated QTLs on the A9 chromosome to a 152-kb interval and on the C5 chromosome to a 175-kb interval, respectively [32, 33]. Liu et al. mapped the semi-dominant gene *ds-1* (*BnaA06.RGA*) on the A6 chromosome. It encodes a DELLA protein that functions as a repressor in GA signaling [34]. A single proline (P)-to-leucine (L) change was identified in the VHYNP motif of a DELLA protein in *ds-1* that leads to a gain-of-function and caused a dwarf phenotype [34]. Subsequently, the *DS-3* gene, which is syntenic to *ds-1* with a similar gene function but a weaker effect on PH, was mapped on the C7 chromosome [35]. In the auxin-signaling pathway, only two genes that encode Aux/IAA proteins were identified and functionally validated in *B. napus*, *BnaA3.IAA7* on the A3 chromosome and its homolog *BnaC05.IAA7* on the C5 chromosome [21, 36]. The molecular mechanisms regulating the PH of *B. napus* remain elusive, and elucidating the mechanism of a new dwarf gene has important scientific significance and applicable value.

The rapeseed ideotype is a semi-dwarf stature with a plant height of ~120–140 cm, narrow branch angles (<30°), and a middle-long silique length [21, 37]. A dwarf mutant of *df59* with a PH of ~65 cm was obtained from ethyl methanesulphonate (EMS)-mutagenized Ningyou 18 (NY18). The *df59* is an excellent germplasm resource for semi-dwarf breeding, with the average PH of F<sub>1</sub> hybrid lines from crosses between NY18 and *df59* being  $126.75 \pm 4.3$  cm, which is in accordance with the ideotype criteria of *B. napus* [21, 37]. The aims of the present study were to: (1) fine-map the gene responsible for dwarf architecture in *df59* using QTL sequencing (QTL-seq) and map-based cloning strategies; (2) elucidate the patterns of gene expression between NY18 and *df59* using comparative transcriptomic analyses; and (3) develop a stable single-nucleotide polymorphism (SNP) marker tightly linked to the dwarf gene that could be used for marker-assisted selection. The present study provides a new gene source for the semi-dwarf breeding of new varieties, and the findings contribute to a better

understanding of the molecular mechanisms underlying dwarfism.

## Results

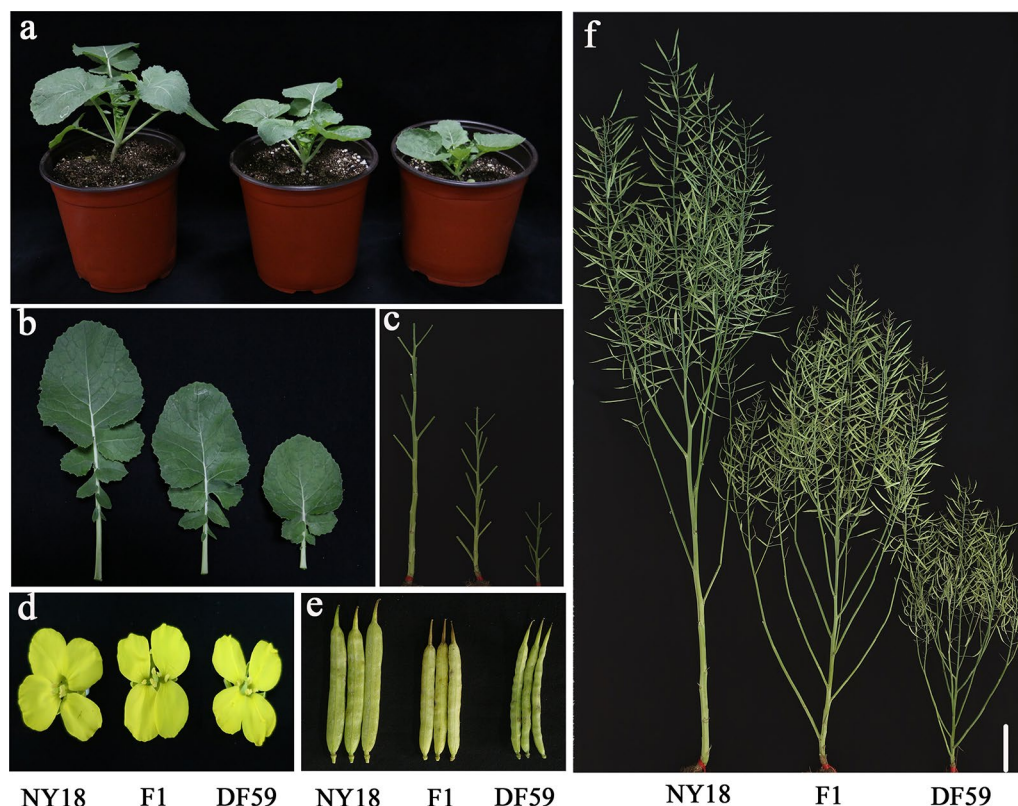
### Phenotypic variation and genetic analysis of plant height

At the seedling stage, *df59* was already significantly shorter than NY18, with smaller leaves and shorter petioles (Fig. 1a, b). The internode length and PH were different between *df59* and NY18 at the mature stage (Fig. 1c), while the petals and siliques of *df59* were smaller than those of NY18 (Fig. 1d, e). At maturity, agronomic traits and seed yield-related traits were investigated in NY18, *df59* and the F<sub>1</sub> of their cross (Table 1). Among the 15 traits, 12 were significantly higher in NY18 than in *df59* (Table 1), with the exception of the first effective branch number, pod number of main inflorescence and silique number per plant. The PH of the F<sub>1</sub> was 126.75 ± 4.3 cm, which meets the ideotype requirement of ~120–140 cm and will be beneficial to high-density planting (Fig. 1f, Table 1). In addition, the seed yield per F<sub>1</sub> plant was 28.4 ± 2.20 g, which was 84.02% of the NY18 yield (33.8 ± 2.51 g). The phenotypic values of seed oil content

and seed fatty acid concentrations for the NY18, *df59* and the F<sub>1</sub> are provided in Additional file 1: Table S1. F<sub>1</sub> and NY18 had similar C16:0, C18:0, C18:2, C18:3, C20:1 and oil content, but had different C18:1 and C22:1 content. Theoretically, if the planting density of the F<sub>1</sub> increased to 1.5 times, the seed yield would increase ~26% per unit of area compared with the normal planting density of NY18.

Root-related traits, including total root length, root surface area, root volume and number of root tips, were measured for NY18, *df59* and their F<sub>1</sub> at 10 days after germination (Additional file 2: Figure S1). The values of the four traits in *df59* were significantly less than in NY18, and the F<sub>1</sub> had approximate mean values of the two parents, which were in accordance with the PH value (Table 1).

The 165 individuals of the F<sub>2</sub> population from a cross between NY18 and *df59* (named NY–DF) observed contained 43 tall plants, 88 medium plants and 34 dwarf plants. A Chi-squared test indicated that the segregation pattern agreed with the Mendelian segregation ratio of 1:2:1 ( $\chi^2 = 1.719$ ,  $P > 0.05$ ). The mixed major-gene plus polygenes inheritance model in the software package



**Fig. 1** Phenotypic characterization of NY18, *df59*, and their F<sub>1</sub> at different developmental stages. **a** Morphology of the NY18, *df59* and their F<sub>1</sub> at the seedling stage; **b** comparison of leaf phenotypes at the seedling stage; **c** comparison of internode lengths at the mature stage; **d** comparison of petal phenotypes at the flowering stage; **e** comparison of silique-related traits at the mature stage; **f** comparison of plant heights and plant architecture at the mature stage. Scale bars = 10 cm

**Table 1 Phenotypic values of agronomic traits, seed yield-related traits and root-related traits for NY18, *df59* and their F<sub>1</sub>**

	Traits	NY18	<i>df59</i>	NY18 × <i>df59</i>	F <sub>1</sub> /NY18 (%)
Agronomic traits	Plant height (cm)	180.4 ± 3.6	65.2 ± 4.6	126.75 ± 4.3	70.26
	First effective branch height (cm)	59.4 ± 5.4	6.4 ± 2.6	20.25 ± 5.3	34.09
	First effective branch number	6.8 ± 0.4	10.4 ± 4.4	10.25 ± 2.6	150.73
	Length of main inflorescence (cm)	76.2 ± 3.4	40.2 ± 4.5	67.4 ± 10.8	88.45
	Pod number of main inflorescence	75.0 ± 10.9	61.6 ± 5.7	76.6 ± 5.0	102.13
	Biomass yield (g)	99.0 ± 3.7	46.9 ± 3.1	68.7 ± 2.7	69.41
	Internode length (cm)	8.04 ± 1.9	2.69 ± 0.6	5.53 ± 1.5	68.78
	Harvest index	0.34 ± 0.03	0.31 ± 0.05	0.41 ± 0.04	120.58
Seed yield-related traits	Seed yield (g)	33.8 ± 2.51	14.7 ± 1.44	28.4 ± 2.20	84.02
	Thousand seed weight (g)	5.23 ± 0.07	3.23 ± 0.06	3.56 ± 0.05	68.07
	Silique number per plant	245.8 ± 30.3	277.6 ± 60.2	303 ± 44.9	123.27
	Silique length (cm)	6.7 ± 0.3	5.5 ± 0.4	5.8 ± 0.2	86.57
	Silique breadth (mm)	7.31 ± 0.25	4.3 ± 0.35	5.43 ± 0.51	74.28
	Silique volume (cm <sup>3</sup> )	1.34 ± 0.12	0.44 ± 0.12	0.95 ± 0.12	70.89
	Seed number per silique	25.7 ± 2.2	14.6 ± 3.2	18.9 ± 2.1	73.54
Root-related traits	Total root length (cm)	37.9 ± 7.23	24.67 ± 5.04	31.0 ± 6.07	81.79
	Root surface area (cm <sup>2</sup> )	3.12 ± 0.55	2.25 ± 0.46	2.61 ± 0.47	83.65
	Root volume (cm <sup>3</sup> )	0.020 ± 0.004	0.0165 ± 0.003	0.018 ± 0.003	90.0
	Number of root tips	42.0 ± 8.3	27.8 ± 8.8	33.4 ± 7.1	79.52

SEA-G4F<sub>2</sub> [38], was used to perform the genetic analysis of PH for the NY–DF F<sub>2</sub> population along with the two parents and the F<sub>1</sub>. The results showed that PH is controlled by one major gene with additive-dominant effects. The major gene heritability was 80.04%, with an additive effect of 28.73 and a dominant effect of −3.12 (Additional file 3: Table S2). The high heritability of the major gene indicated that PH in *df59* was relatively stable and not greatly influenced by environment; therefore, it can be selected in the early generations of a breeding process.

#### QTL-seq of the NY–DF population

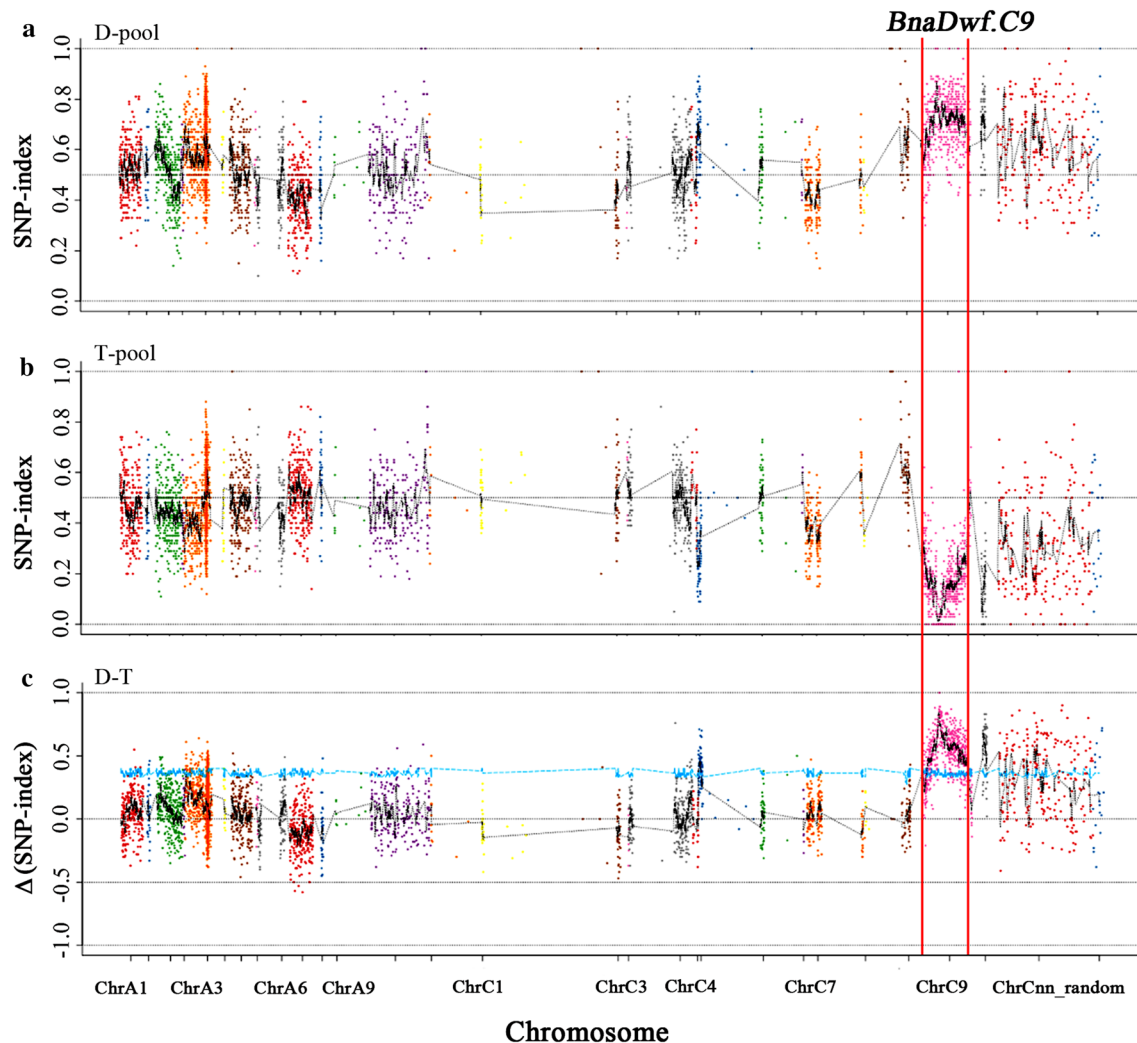
Two contrasting DNA pools were constructed from 16 extremely tall lines (T-pool) and 24 extremely dwarf lines (D-pool) in NY–DF F<sub>2</sub> population. Illumina high-throughput sequencing generated a total of 92.986 Gb clean data for the two pools and two parental lines, with average Q20 ≥ 95.66% and Q30 ≥ 93.14%. When the obtained reads were matched to the *B. napus* “Darmor-*bzh*” reference genome, the results showed that the depths of sequencing coverage for T-pool, D-pool, NY18 and *df59* were 35.85-fold, 31.27-fold, 16.79-fold and 17.84-fold, respectively.

Based on the genotyping, 4520 polymorphic SNPs with haplotype differences between the two parents were identified. The region (17.00–27.30 Mb) on chromosome C9 had an average SNP-index of 0.726 in the D-pool, with the highest being 1 (Fig. 2a), while the SNP-index in T-pool was 0.128, with the lowest being 0 (Fig. 2b). Then,

the genome sequence of the parent NY18 was used as a reference to calculate the Δ(SNP-index) of the 4520 SNPs by combining the SNP-indices of the T- and D-pools. At a 95% significance level, the genomic region from 17.00 to 27.30 Mb had an average Δ(SNP-index) of 0.60 (Fig. 2c), suggesting that this region harbored a major QTL for PH, which was designated *BnaDwf.C9*.

#### Fine-mapping the *BnaDwf.C9* locus

To fine-map the *BnaDwf.C9* locus, two flanking pentaprimer amplification refractory mutation system (PARMS) SNP markers, M1 (physical position of the *B. napus* ‘Darmor-*bzh*’ genome: 17,001,732) and M11 (27,299,112), were first designed to investigate recombinants in the F<sub>2</sub> population (HO–DF) from a cross between Holly and *df59* (Additional file 4: Table S3). The results showed that 81 recombinants were obtained among the 2356 F<sub>2</sub> individuals, including 56 tall recombinants and 25 dwarf recombinants (Additional file 4: Table S3). Based on the re-sequencing information of NY18 and *df59* at the target region, nine polymorphic SNP markers (M2–M10) between M1 and M11 were designed and used for genotyping the 81 recombinant individuals. Recombinant genotypes showed that M2 was located on the left side of *BnaDwf.C9*, while six SNP markers (M5–M10) were mapped on the right side (Fig. 3, Additional file 4: Table S3). In addition, M3 and M4 co-segregated with *BnaDwf.C9* and were consistent with the PH phenotype. Eventually, the location of



**Fig. 2** SNP-index and  $\Delta(\text{SNP-index})$  graphs from the QTL-seq analysis. SNP-indices of **a** D-pool and **b** T-pool; **c**  $\Delta(\text{SNP-index})$ . The x-axes represent the *B. napus* chromosomes and the y-axes represent the SNP-index (**a, b**) or  $\Delta(\text{SNP-index})$  (**c**). The SNP-index and  $\Delta(\text{SNP-index})$  calculations are based on the descriptions in the “Methods.” The genomic region (17.00–27.30 Mb) on chromosome C9 had an average  $\Delta(\text{SNP-index})$  of 0.60 and is regarded as the candidate QTL, with a 95% significance level ( $P < 0.05$ )

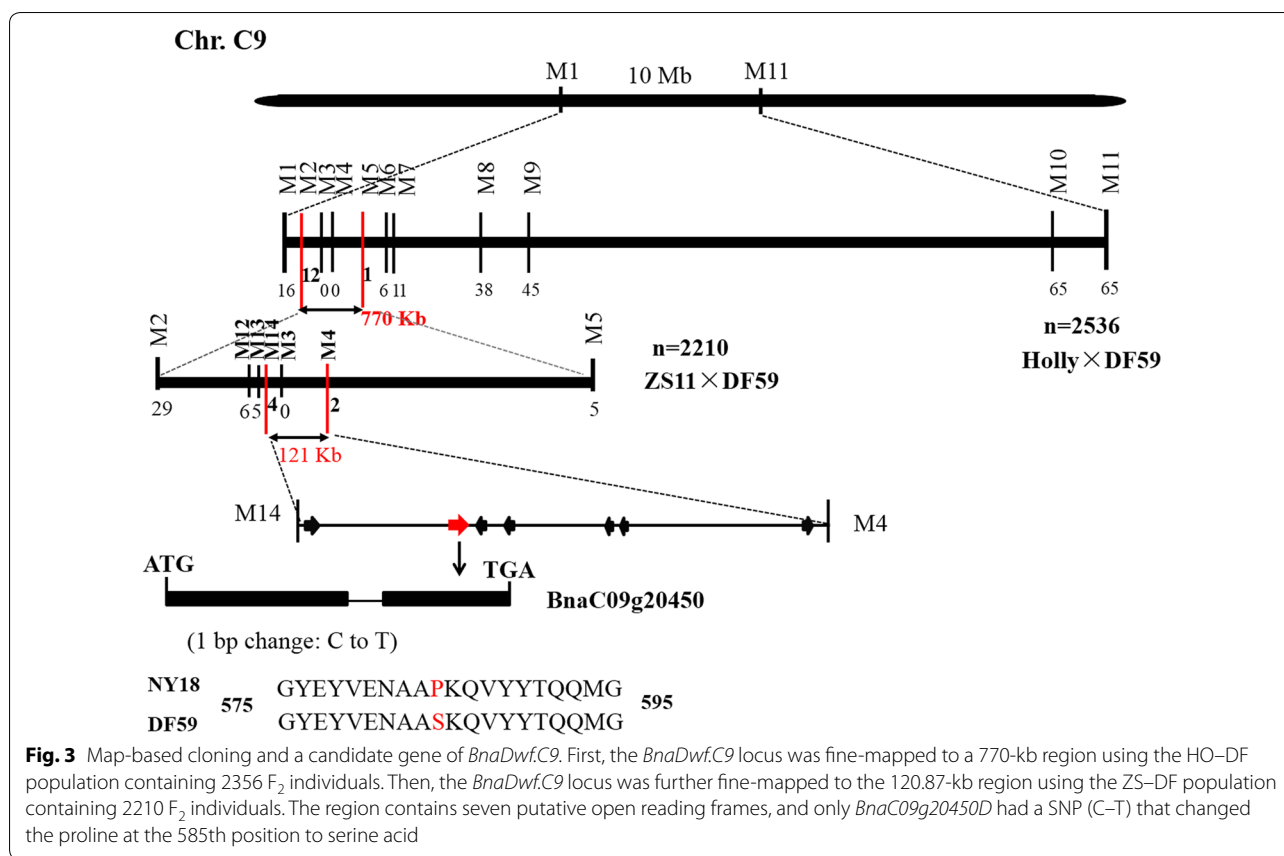
*BnaDwf.C9* was narrowed down to the interval between M2 and M5 in a region of 770.72 kb in the ‘Darmor-*bzh*’ genome. No polymorphic SNP markers were able to narrow the interval in the HO–DF population.

For the further fine-mapping of *BnaDwf.C9*, a population of 2210  $F_2$  individuals was constructed from a cross between Zhongshuang 11 (ZS11) and *df59* (ZS–DF population). Using the methods described above, 34 recombinant individuals were detected between M2 and M5 in the ZS–DF population (Table 2). Subsequently, three new polymorphic SNP markers in this region (M12–M14) were designed, together with co-segregating SNP markers (M3 and M4) in the HO–DF population and used to screen the 34 recombinants (Additional file 5: Table S4),

resulting in 6, 5, 4, 0 and 2 recombinants, respectively (Fig. 3, Table 2). These results suggested that the locus *BnaDwf.C9* was fine-mapped to the region between M14 and M4, corresponding a 120.87 kb interval in the *B. napus* ‘Darmor-*bzh*’ genome (Fig. 3).

#### Genome-wide transcriptomic analyses of NY18 and *df59*

High-throughput RNA sequencing (RNA-seq) generated 20.32–26.83 million raw reads for each sample. After trimming the low-quality sequences, 16.29–22.37 million clean reads were aligned to the *B. napus* ‘Darmor-*bzh*’ genome (Additional file 6: Table S5). A total of 3995 differentially expressed genes (DEGs) were identified between NY18 and *df59* stem tip transcriptomes



using the threshold false discovery rate <0.005 and at least a 2.0-fold expression changed. In total, 1266 genes were significantly upregulated and 2729 genes were significantly down-regulated in *df59* compared with NY18 (Additional file 7: Table S6).

To understand gene functions associated with the dwarf phenotype in *df59*, a Gene ontology (GO) enrichment analysis of the DEGs was performed. For both up- and down-regulated genes, the most three significantly enriched GO terms in the ‘biological process’, ‘cellular component’ and ‘molecular function’ groups were the same (Additional file 8: Table S7). For example, “metabolic process”, “cellular process” and “single-organism process” were the three most enriched GO terms of the “biological process” group.

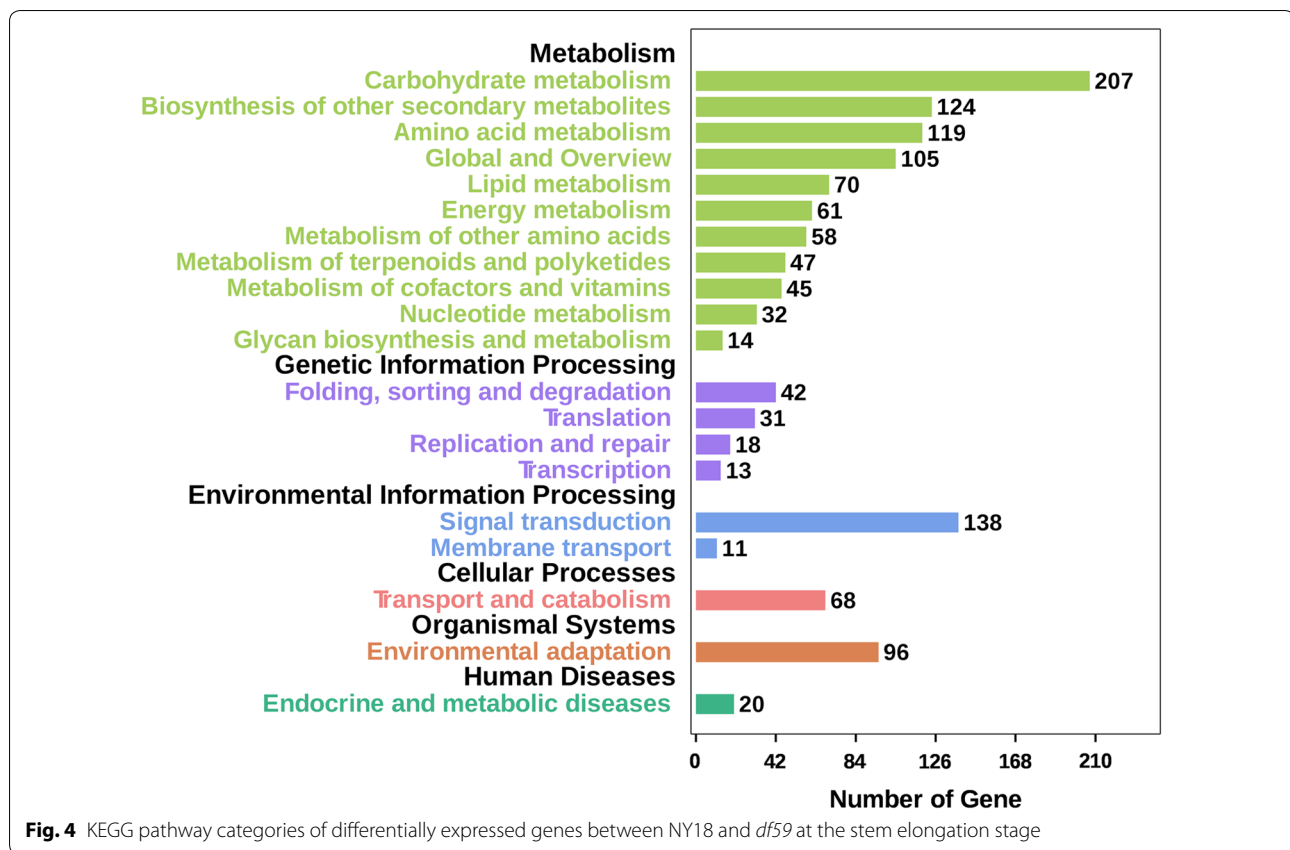
To discern DEG functions, we also conducted a Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. The 1266 up- and 2729 down-regulated genes were enriched in 92 and 114 pathways, respectively (Additional file 9: Table S8). The top 20 enriched pathways for all of the up- and down-regulated DEGs are shown in Fig. 4. The five most significantly enriched pathways were carbohydrate metabolism, signal transduction, biosynthesis of other secondary metabolites, amino acid

metabolism, and global and overview (Fig. 4). As the PHs in *A. thaliana* and rice were mainly regulated by plant hormones [7–11], we focused on the DEGs involved in signal transduction pathway (138 genes). Among them, 118 DEGs were clustered in plant hormone-related signal transduction pathways. In total, 81 DEGs were enriched in auxin signal transduction, accounting for 68.6% of the 118 DEGs (Additional file 10: Table S9). The remaining 37 DEGs were enriched in brassinosteroid (10 genes), abscisic acid (seven genes), GA (five genes), ethylene (five genes), jasmonic acid (four genes), salicylic acid (four genes) and cytokinin (two genes) signaling pathways (Additional file 10: Table S9). This result strongly indicated that the mutation leading to the dwarf phenotype of *df59* may occur in an important gene in the auxin signal transduction. The KEGG pathway enrichment analysis provided an important clue for identifying candidate genes.

#### Identification of a candidate gene

Candidate genes underlying the 17.42–17.54 Mb of *BnaDwf.C9* were analyzed based on the *B. napus* “Darmor-bzh” reference genome [22]. The segment harboring the *BnaDwf.C9* locus contained seven annotated or





**Fig. 4** KEGG pathway categories of differentially expressed genes between NY18 and *df59* at the stem elongation stage

**Table 3** Genes on the mapped 120.87 kb interval of “Darmor-*bzh*” reference genome and their annotation

Gene of <i>B. napus</i>	Chromosome position	Orthologous gene of <i>A. thaliana</i>	Gene function
<i>BnaC09g20440D</i>	17,421,172–17,423,263	AT2G01180	Phosphatidic acid phosphatase 1
<i>BnaC09g20450D</i>	17,461,819–17,464,062	AT2G01190	Octicosapeptide/Phox/Bem1p family protein
<i>BnaC09g20460D</i>	17,464,702–17,466,229	AT2G01250	Ribosomal protein L30/L7 family protein
<i>BnaC09g20470D</i>	17,469,391–17,470,034	AT2G01300	Na
<i>BnaC09g20480D</i>	17,481,324–17,484,464	AT2G01330	Nucleotide binding
<i>BnaC09g20490D</i>	17,493,553–17,493,792	AT1G18370	ATP binding microtubule motor family protein
<i>BnaC09g20500D</i>	17,534,536–17,535,744	AT2G01340	Na

the ORFs nor expression differences between NY18 and *df59*.

In addition, “Ningyou 7” [39] and “ZS11” [40] reference genomes were used for the candidate gene analysis within the 120.87-kb interval harboring *BnaDwf.C9*. A new gene (*chrC09g002459*) in the “Ningyou 7” genome that was not in “Darmor-*bzh*” was identified (Additional file 11: Table S10). However, the gene was not expressed in either NY18 or *df59*. In the “ZS11” genome, *BnaC09G0251200ZS* was a new gene with no expression, while *BnaC09G0251600ZS* was a new predicted gene with a sequence length of only 240 bp, which showed

neither a sequence nor expression difference between NY18 and *df59* (Additional file 11: Table S10). Therefore, these genes were excluded as candidate genes for *BnaDwf.C9*.

*BnaC09g20450D*, named *chrC09g002455* in “Ningyou 7” and divided into *BnaC09G0251300ZS* and *BnaC09G0251400ZS* in “ZS11”, is homologous to *AT2G01190* in *Arabidopsis*. *AT2G01190* belongs to the octicosapeptide/Phox/Bem1p family of proteins and encodes a protein of unknown function (Table 3). The PB1 domain (aa 69–167) is an important functional domain in *BnaC09g20450D* (Additional file 12: Figure



S2). In plants, PB1 mediates interactions of ARF and Aux/IAA to modulate auxin-regulated gene transcription [41, 42]. The KEGG pathway enrichment analysis showed that most of the DEGs related to signal transduction were enriched in the auxin signal transduction pathway. The RNA-seq analysis revealed that *BnaC09g20450D* had similar transcript levels between NY18 and *df59* (Additional file 11: Table S10), while the re-sequencing analysis revealed a SNP in NY18 and *df59* (Fig. 3). Subsequently, we amplified and sequenced the promoter region (2.0-kb upstream of the start codon) and the ORF of *BnaC09g20450D* from NY18 and *df59*. No polymorphism was identified in the promoter sequences. The ORF is 1887-bp in length, encoding a protein of 628 amino acids, and a single nucleotide substitution (C to T) was identified in the second exon, which changed the proline at the 585th position to serine acid (P585S) (Fig. 3). However, the SNP did not occur in PB1 domain or any other domain (Additional file 12: Figure S2). The SNP in *BnaC09g20450D* was also developed as PARMS SNP marker M3, which co-segregated with *BnaDwf.C9* in 2536 individuals of the HO–DF population and in 2210 individuals of the ZS–DF population (Fig. 3). Therefore, we speculated that *BnaC09g20450D* was the most likely candidate gene of *BnaDwf.C9*.

#### Development of a PCR-based SNP marker specific for *BnaDwf.C9*

The SNP in the candidate gene *BnaC09g20450D* was targeted to develop a molecular marker. A SNP marker named BnaPHC9-SNP containing four primers (BnaM3pcr-F/BnaM3pcr-R/BnaM3pcr-Fc/BnaM3pcr-Rt) was designed based on the 400-bp flanking sequence of the SNP. BnaPHC9-SNP was first used to amplify NY18, *df59* and their F<sub>1</sub>, which produced a 351-bp fragment in NY18, a 179-bp fragment in *df59*, and both 351-bp and 179-bp fragments in the F<sub>1</sub>. Subsequently, 21 (tall PH), 21 (dwarfism PH) and 22 (medium PH) individuals were randomly selected from the ZS–DF F<sub>2</sub> population to test the SNP marker amplification. The results of the agarose gel electrophoresis showed that only a 179-bp PCR fragment was present in individuals with a dwarf PH (Fig. 5a), only a 351-bp fragment was amplified in individuals with a tall PH (Fig. 5b), and both 179-bp and 351-bp fragments were produced in individuals with a medium PH (Fig. 5c). These results suggested that BnaPHC9-SNP could produce specific PCR amplicons from different alleles of the SNP and that it can be used for the rapid identification of the *BnaDwf.C9* locus, which confers the dwarfing trait, in breeding programs.

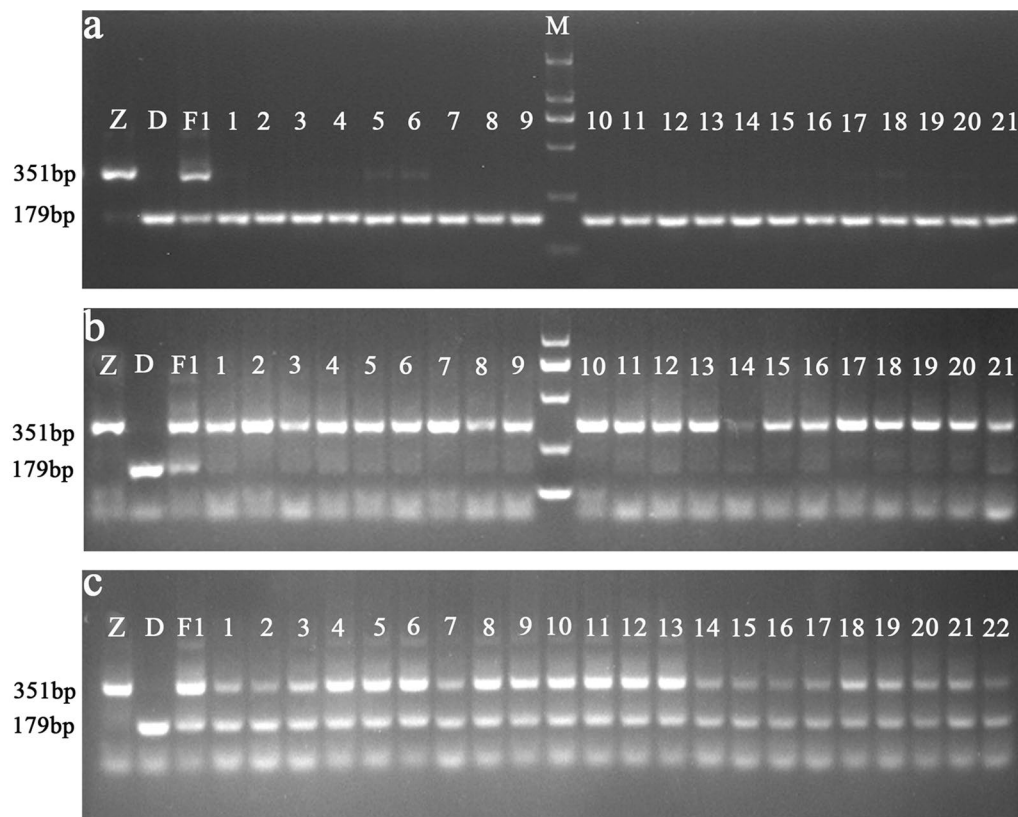
## Discussion

### *df59* is an elite genetic resource for semi-dwarf breeding

In the 1960s, the introduction of dwarfing traits into wheat and rice, combined with the application of improved cultivation methods, led to spectacular increases in grain yields, the so-called “green revolution” [12, 13]. *B. napus* is the third most important oilseed crop, providing 13–16% of vegetable oil globally [43], and rapeseed oil has a strong potential for use in biodiesel production [44]. In China, up to 70% of the total rapeseed cultivation areas were planted with hybrid rapeseed, because they normally produce a 25% greater seed yield and have greater yield stability [37]. However, the utilization of heterosis in rapeseed also led to the PH increasing significantly, resulting in an increased risk of lodging and a decrease in the harvest index. Owing to the lack of excellent dwarf *B. napus* germplasm resources, there is no high-yielding semi-dwarf variety widely planted, as there is for wheat and rice. *df59* is a dwarf mutant obtained from a large-scale screening of EMS-mutagenized NY18. The PH of F<sub>1</sub> individuals of a cross between *df59* and NY18 was  $126.75 \pm 4.3$  cm (Table 1), and the seed yield was 84% that of NY18, suggesting that the F<sub>1</sub> are suitable for high-density planting and mechanized harvest [37]. The average harvest index in *B. napus* of 0.28 [45, 46], was much lower than those of wheat and rice (0.4–0.6) [45]. The harvest index of F<sub>1</sub> individuals was 0.41, which was significantly greater than that of NY18 (0.34), indicating that the F<sub>1</sub> more efficiently utilized water and soil resources (Table 1). In addition, the PH of ZS11 was  $162.6 \pm 5.5$  cm, while the F<sub>1</sub> PH of a ZS11 × *df59* cross was  $114.4 \pm 3.9$  cm (data not shown), suggesting that allele in *df59* showed a dwarfing effect in multiple genetic backgrounds. In summary, *df59* is an outstanding male parent for breeding new semi-dwarf hybrid varieties that can be densely planted and machine harvested.

### *BnaDwf.C9* is a new locus associated with plant height in *B. napus*

In *B. napus*, hundreds of QTLs associated with PH have been identified and located on all 19 chromosomes [23–31]. Several QTLs on the A3, A6, A9, C5 and C7 chromosomes have been fine-mapped or cloned [21, 32–36]. However, only a few QTLs with minor PVs were obtained on the C9 chromosome in previous studies. For example, Shi et al. detected a minor QTL with a PV range of 3.2–4.3% [23]; Ding et al. identified an environment-specific QTL with a PV of 10.9% [25]; and Wang et al. detected four QTLs at the mature stage, with a PV range of 3.69–9.87% [26]. Using a population containing 520 diverse rapeseed accessions, Sun et al. identified 68 loci, which were distributed over the chromosomes, except for the



**Fig. 5** Electrophoretic profiles obtained for individuals with different plant heights from the developed SNP marker. Z: Zhongshuang 11; D: *df59*; F<sub>1</sub>: the F<sub>1</sub> of a cross between Zhongshuang 11 and *df59*; M: DNA marker. The SNP marker can amplify **a** specific PCR products in *df59* and dwarf individuals (179-bp fragment), **b** Zhongshuang 11 and tall individuals (351-bp fragment), and **c** F<sub>1</sub> and middle individuals (both 351-bp and 179-bp fragments). PCR products were analyzed by electrophoresis in 2.5% agarose gel

C8 and C9 chromosomes, significantly associated with PH using a GWAS [29]. Luo et al. and Li et al. also performed GWASs for PH, and no locus significantly associated with PH was identified on the C9 chromosome [27, 28]. Thus, it appeared that there was no major QTL controlling PH on the C9 chromosome in *B. napus*. In the present study, *BnaDwf.C9* was fine-mapped to a 120.87-kb region on the C9 chromosome (Fig. 3). Although it difficult to directly compare *BnaDwf.C9* with those reported QTLs owing to the lack of common markers, we still believe that *BnaDwf.C9* is a potentially new locus, because it showed an obvious effect on PH.

#### The candidate gene may regulate plant height of *df59* through the auxin-signaling pathway

The plant hormone auxin plays an essential role in most aspects of growth and developmental processes [18, 19, 47]. In most plants, auxin regulates the transcription of auxin-responsive genes through the well-established TIR1/AFB-Aux/IAA-ARF pathway [18, 19]. The Aux/IAAs contain four functional domains [18, 19]: Domain

I is a repressor domain that contains a leucine repeat EAR motif [48]; Domain II is an internal domain that contains the degron motif GWPPV; Domains III and IV share high homology and contain C-terminal regions that form a PB1 protein–protein interaction domain [49]. In *B. napus*, a G-to-A mutation in the GWPPV motif of domain II causes multiple phenotypic alterations, including a reduction in PH and branch angles [21]. Among the seven genes underlying the mapped interval of *BnaDwf.C9* (Table 3, Additional file 11: Table S10), only *BnaC09g20450D* has a SNP (C-to-T mutation) between NY18 and *df59*. *BnaC09g20450D* also contains a PB1 domain (aa 69–167) (Additional file 12: Figure S2); however, the P585S substitution did not occur in the PB1 domain or any other domain. The PB1 domain can facilitate the formation of ARF-ARE, ARF-Aux/IAA, and Aux/IAA-Aux/IAA homo- and hetero-oligomers, because the ARF and Aux/IAA proteins contain similar PB1 domains [49]. We speculate that the P585S substitution in *BnaC09g20450D* disrupts the normal function of the PB1 domain in an unknown way and then affects

auxin signaling resulting in the dwarf stature of *df59*. The RNA-seq analysis also showed 81 DEGs in the auxin-signaling pathway between NY18 and *df59* (Additional file 10: Table S9). Combining the results of fine-mapping and RNA-seq, we hypothesized that the candidate gene of *BnaDwf.C9* regulated PH through the auxin-signaling pathway, but the molecular mechanism remains unknown and requires further study.

#### The developed SNP marker can facilitate the application of *BnaDwf.C9* in breeding

Traditional breeding, which largely depends on the breeder's experience and subjective judgment, is a relatively resource- and time-consuming process. Marker-assisted breeding is an effective and accurate approach to perform target-trait selection in breeding, because traits can be examined at any developmental stage, providing results without incurring environmental impacts [50, 51]. In the present study, we determined that the dwarf *B. napus* mutant *df59* has important potential application value in breeding. However, it is important to determine how to efficiently and extensively use the dwarfing gene in semi-dwarf breeding. To address this question, a molecular marker BnaPHC9-SNP was developed for *BnaDwf.C9* based on a SNP that co-segregated in two populations, containing 4746 individuals in total (Table 2, Additional file 4: Table S3). BnaPHC9-SNP was a dominant, allele-specific functional marker, which could amplify specific PCR products in *df59* (179-bp fragment), NY18 (351-bp fragment) and their  $F_1$  (both 351-bp and 179-bp fragments). The marker was subsequently confirmed using individuals with different PHs in the ZS-DF  $F_2$  population (Fig. 5), suggesting that BnaPHC9-SNP can be used to select the target PH without morphological characterization, which would speed up the breeding process. To date, a variety of high-throughput SNP genotyping methods has been developed, including the TaqMan system [52] and Kompetitive allele-specific PCR. However, for most breeders, the equipment needed in these methods is prohibitory and more expensive than using normal primers. The SNP marker developed in this study requires no fluorescence-tagged probes or real-time PCR instruments, and PCR products can be correctly detected using agarose gel electrophoresis. Therefore, BnaPHC9-SNP can be widely used in marker-assisted selection of *BnaDwf.C9* and speed up the breeding process.

#### Conclusions

*Brassica napus* provides not only edible vegetable oil for human consumption, but also a triglyceride source for biofuel and lubricant production. In the present study, we isolated the dwarf mutant *df59* from an EMS-mutagenized NY18. The main agronomic traits in NY18, *df59*

and their  $F_1$  showed that *df59* is an elite genetic resource for semi-dwarf breeding. Subsequently, the combination of QTL-seq and fine-mapping revealed a candidate gene located within an interval of 120.87 kb on the C9 chromosome. The transcriptome analysis suggested that the PH of *df59* was most likely influenced by a gene involved in auxin signal transduction. In addition, a comprehensive analysis revealed that *BnaC09g20450D* was the most likely candidate gene. Then, a molecular marker was developed based on the SNP in *BnaC09g20450D*. These results enrich our knowledge of the genetic architecture underlying PH in *B. napus* and also provide valuable resources for semi-dwarf breeding.

#### Methods

##### Plant materials

The dwarf mutant *df59* was isolated from EMS-mutagenized lines of NY18. NY18 is a variety cultivated by Jiangsu Academy of Agricultural Sciences, China. Mature seeds of NY18 were mutagenized with 1.0% EMS solution (W/V, Sigma-Aldrich) at pH 7.0 phosphate buffer for 12 h at 25 °C, according to the descriptions of Li et al. with modifications [53]. The mutagenized seeds ( $M_1$  generation) were rinsed with water for 4 h and sown in the field. Individual plants were bagged at the flowering stage. Self-pollination seeds of approximately 10,000 individual plants were harvested. Each  $M_2$  seeds were then sown into independent line. Dwarf lines were bagged to harvest seeds for evaluation in the  $M_3$  generation.

The NY-DF  $F_2$  population containing 165 individual lines derived from a cross between NY18 and *df59* was used for genetic inheritance and QTL-seq. Two conventional rapeseed cultivars, Holly and ZS11, were used as parental lines to develop segregating populations for fine-mapping of the QTLs associated with dwarf architecture. Holly is a Canadian spring variety, while ZS11 is an elite Chinese semi-winter rapeseed cultivar. The HO-DF  $F_2$  population derived from a cross between Holly and *df59* contained 2536 lines, while the ZS-DF  $F_2$  population derived from a cross between ZS11 and *df59* contained 2210 lines.

##### Trait measurement

NY-DF, HO-DF and ZS-DF populations, as well as their parents and the  $F_1$  were all planted in the field of Jiangsu Academy of Agricultural Sciences, Nanjing, Jiangsu Province, China. No specific permissions were required for the field trials. The field experiments were conducted in accordance with Wang et al. with 20 plants per row and 40 cm between the rows [54]. The PH values for all the materials were measured at the mature stage.

At maturity, five open-pollinated plants each of NY18, *df59* and their  $F_1$  growing in the middle of the plot were

selected. Agronomic traits and seed yield-related traits were measured in accordance with the description of Zhao et al. [55] and Wang et al. [56], including PH, biomass yield, seed yield, thousand seed weight, first effective branch height, first effective branch number, internode length, harvest index, length of main inflorescence, pod number of main inflorescence and silique number per plant. The harvest index was calculated as the ratio of seed yield to biomass yield [45]. In total, 10 well-developed siliques, which were randomly selected from the first branch adjacent to the main inflorescence, were used to determine the silique-related traits of NY18, *df59* and their  $F_1$ , including silique length, silique breadth, seed number per silique and silique volume [56]. The seed oil content was measured by nuclear magnetic resonance using standard methods and fatty acid compositions were determined by near infrared reflectance spectroscopy in accordance with Chen et al. [57].

To investigate root traits at the seeding stage, four plump seeds each of NY18, *df59* and the  $F_1$  were planted in a seed germination pouch (CYG-19LB; Pheno Trait Technology, Co., Ltd, Beijing, China) at 25 °C, and three repetitions were conducted. After 10 days, total root length (cm), root surface area (cm<sup>2</sup>), root volume (cm<sup>3</sup>) and number of root tips for each seeding were measured using LA-S Root Analysis software (Wanshen Ltd, Hangzhou, China).

### Statistical and QTL-seq analyses

The mixed major-gene plus polygenes inheritance model of the software package SEA-G4 $F_2$  was used to identify the inheritance of PH in the NY-DF  $F_2$  population [38].

From the 165 lines of the NY-DF  $F_2$  population, 16 extremely tall lines and 24 extremely dwarf lines were selected. Genomic DNA was extracted from young leaves using a Plant Genomic DNA Kit (Tiangen, Beijing, China), and T-pool and D-pool were constructed by mixing equal ratios of appropriate individual DNAs. Sequencing libraries of the two bulks and two parents were constructed, and sequence data were generated using the Illumina HiSeq™ PE150 (Illumina, Inc; San Diego, CA, USA) platform. Both data sequencing and data analyses were performed by Novogene Bioinformatics Technology Co. Ltd. (Beijing, China). The raw data were filtered through a series of quality controls which resulted in the removal of reads with  $\geq 10\%$  unidentified nucleotides,  $>50\%$  bases having phred quality  $<5$  or  $>10$  nt aligned to the adapter. The clean reads of each sample were aligned against the *B. napus* “Darmor-*bzh*” reference genome [22] using BWA software [58], and the SAMtools command “rmdup” was used to remove multiple read pairs [59]. Variant calling was performed for all the samples using the Unified Genotyper function in

GATK software [60]. SNPs were determined using the Variant Filtration parameter in GATK. Using NY18 as the reference parent, the above two bulks SNP-indices were calculated as the proportion of reads containing SNPs not found in NY18. The  $\Delta(\text{SNP-index})$  was calculated using the formula:  $\Delta(\text{SNP-index}) = \text{SNP-index (D-pool)} - \text{SNP-index (T-pool)}$ . A QTL was considered a candidate associated with PH if the  $\Delta(\text{SNP-index})$  was significantly different ( $P < 0.05$ ).

### Fine-mapping of the QTL for plant height

The major QTL for PH was identified based on QTL-seq and named *BnaDwf.C9*, and SNPs underlying the confidence interval of *BnaDwf.C9* were obtained. Total DNAs of HO-DF and ZS-DF  $F_2$  individuals were extracted from fresh leaves using a modified cetyl-trimethylammonium bromide method [61]. PARMS was used to screen recombinant plants for fine-mapping among HO-DF and ZS-DF  $F_2$  populations. The principle of PARMS SNP genotyping is similar to that of Kompetitive allele-specific PCR assays [62]. The master mix for PARMS markers was purchased from Gentides Biotech Co., Ltd. (Wuhan, China). Detailed information for conducting the qPCR-based PARMS assay was previously published by Zhang et al. [62] and Liu et al. [63].

PARMS SNP markers flanking the confidence interval of *BnaDwf.C9* were screened from the two parents, the  $F_1$ , and 10 individuals each of the extremely tall and dwarf lines of the HO-DF population. Then two polymorphic markers (M1 and M11) were used to screen the 2536 individuals of the HO-DF  $F_2$  population (Fig. 3). Combined with phenotypes, plants containing recombinants between the two markers were selected. Furthermore, the recombinant plants were analyzed with newly developed polymorphic PARMS SNP markers (M2–M10), and *BnaDwf.C9* was finally mapped to an interval between M2 and M5 (Fig. 3).

For the further fine-mapping of *BnaDwf.C9*, M2 and M5 were used to screen the 2210 individuals of the ZS-DF  $F_2$  population. Using the same method as described above, plants containing recombinants between M2 and M5 were selected. Three new polymorphic PARMS SNP markers (M13–M15) together with co-segregating SNP markers (M3 and M4), were used to analyze the selected recombinant plants (Fig. 3, Additional file 5: Table S4). Finally, we narrowed down *BnaDwf.C9* to a genomic region between the markers M14 and M4.

### RNA library construction and sequencing

Equivalent amounts of stem tips from NY18 and *df59* at the stem elongation stage were collected for RNA extraction, and two biological replicates were performed. Total RNA was extracted using a MiniBEST Plant RNA

Extraction Kit (TaKaRa, Dalian, China). RNA-seq and data analysis were carried out by Novogene Bioinformatics Technology Co. Ltd. In brief, the RNA concentration was measured using a Qubit<sup>®</sup> RNA Assay Kit and a Qubit<sup>®</sup> 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). The RNA integrity was assessed using an RNA Nano 6000 Assay Kit and the Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). Each of the two sequencing libraries for NY18 and *df59* were constructed using an NEBNext<sup>®</sup> Ultra<sup>™</sup> RNA Library Prep Kit for Illumina<sup>®</sup> (NEB, USA) following the manufacturer's instructions, and the library quality was assessed on the Agilent Bioanalyzer 2100 system. The four libraries were sequenced on an Illumina HiSeq 2000 platform, and 100 bp paired-end reads were generated. These methods were described by Yu et al. with modifications [64].

### RNA-seq data analysis

Raw RNA-seq reads were processed to remove reads containing the adapter, reads containing ploy-Ns and low-quality reads. Paired-end clean reads were aligned to the *B. napus* "Darmor-*bzh*" reference genome [22] using TopHat version 2.0.6 [65]. The read numbers mapped to each gene were counted using HTSeq version 0.6.1 [66], and reads per kilobase of exon per million reads of each gene were calculated. Differential expression analyses of NY18 and *df59* (two biological replicates per sample) were performed using the DESeq R package (1.10.1). Genes with a false discovery rate < 0.005 and  $\log_2(\text{fold change}) \geq 1$  were declared DEGs.

The GO annotation of DEGs was performed using the GOseq R package [67], and GO terms with corrected  $P < 0.05$  were considered to be significantly enriched terms. In addition, DEGs were submitted to the KEGG (<https://www.genome.jp/kegg/>) website, and KEGG enrichment pathways of DEGs were determined using KOBAS online analysis database [68].

### Development of an allele-specific marker for traditional PCR amplification

Based on the results of the fine-mapped *BnaDwfC9*, the PARMS SNP marker M3 co-segregated with the *BnaDwfC9* gene associated with PH. To confirm this SNP, the coding region of the gene that contained the SNP was amplified and sequenced from NY18 and *df59*. A region, including 200 bp both upstream and downstream of the SNP, was considered the target region. A SNP marker containing four primers and named BnaPHC9-SNP was designed for the allele-specific amplification of the SNP. The primers were a forward locus primer (BnaM3pcr-F: 5'-GAGAAATACTCCGCAACCTACG-3'), reverse locus primer (BnaM3pcr-R: 5'-ATGTTCCGAAACCAA CCAGAG-3'), allele primer 1 (BnaM3pcr-Fc: 5'-TAT

GAATATGTGGAAAATGAGC-3') and allele primer 2 (BnaM3pcr-Rt: 5'-GCGTGTAGTATACCTGCTTGGA-3'). BnaM3pcr-F began amplifying from 157-bp upstream of the SNP, while BnaM3pcr-R began amplifying from 193-bp downstream of the SNP. BnaM3pcr-Fc and BnaM3pcr-Rt were upstream and downstream, respectively, with 3'-terminal bases of C and A (according to the allele of NY18 or *df59*), respectively. A mismatch base was introduced at the third-to-last base of the primer BnaM3pcr-Fc. All the primer oligonucleotides were synthesized by Tsingke Biological Technology Co., Ltd. (Wuhan, China).

The PCR mixture (including dNTPs, Taq buffer and Taq enzyme) was purchased from Gentides Biotech Co., Ltd. The PCR reagent mixture (20  $\mu\text{L}$  total volume) contained: 2  $\times$  PCR MIX: 10  $\mu\text{L}$ , BnaM3pcr-F: 0.8  $\mu\text{L}$ , BnaM3pcr-R: 0.8  $\mu\text{L}$ , BnaM3pcr-Fc: 0.8  $\mu\text{L}$ , BnaM3pcr-Rt: 0.8  $\mu\text{L}$ , DNA: 1  $\mu\text{L}$  (50 ng/ $\mu\text{L}$ ), ddH<sub>2</sub>O: 5.8  $\mu\text{L}$ . The PCR assay was conducted as described by Zhang et al. with modifications [62], as follows: denaturation at 94 °C for 15 min, followed by 10 cycles of 94 °C for 20 s and 65 °C (−0.8 °C per cycle) for 1 min, followed by 30 cycles of 94 °C for 20 s and 57 °C for 1 min, and a final extension at 72 °C for 5 min. PCR products were analyzed by electrophoresis in 2.5% agarose gel for 40 min at 100 V in TAE buffer (40 mM Tris–acetate, 1 mM EDTA, pH 8.0).

### Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s13068-020-01687-y>.

**Additional file 1: Table S1.** The phenotypic values of seed oil content and seed fatty acid concentrations for the NY18, *df59* and the F<sub>1</sub>.

**Additional file 2: Figure S1.** Root related traits of NY18, *df59* and their F<sub>1</sub> at 10 days after germination.

**Additional file 3: Table S2.** Genetic parameters estimated in one major gene with additive-dominant model in the NY–DF F<sub>2</sub> population.

**Additional file 4: Table S3.** Eighty-one recombinants and their genotypes detected in HO–DF F<sub>2</sub> population.

**Additional file 5: Table S4.** PARMS SNP markers used for fine-mapping of the *BnaDwfC9* locus in ZS–DF F<sub>2</sub> population.

**Additional file 6: Table S5.** Summary of transcriptome sequencing data.

**Additional file 7: Table S6.** The differentially expressed genes between NY18 and *df59*.

**Additional file 8: Table S7.** GO terms for the differentially expressed genes between NY18 and *df59*.

**Additional file 9: Table S8.** KEGG pathways for the differentially expressed genes between NY18 and *df59*.

**Additional file 10: Table S9.** The 118 differentially expressed genes clustered in plant hormone-related signal transduction pathways.

**Additional file 11:** Genes on the mapped 120.87 kb interval of "Darmor-*bzh*", "Ningyou 7" and "Zhongshuang 11" reference genomes and their expression pattern.

**Additional file 12: Figure S2.** The conserved domains for the candidate gene *BnaC09g20450D*.

## Abbreviations

ARF: Auxin response factor; Aux/IAA: Auxin/indole acetic acid protein; *B. napus*: *Brassica napus*; DEG: Differentially expressed gene; D-pool: Dwarf pool; EMS: Ethyl methanesulphonate; GA: Gibberellic acid; GO: Gene ontology; GWAS: Genome-wide association Study; KEGG: Kyoto Encyclopedia of Genes and Genomes; IAA: Indole-3-acetic acid; NY18: Ningyou 18; ORF: Open reading frames; PH: Plant height; PV: Phenotypic variance; QTL: Quantitative trait locus; SNP: Single-nucleotide polymorphism; TIR1/AFB: Transport inhibitor resistant/auxin signaling F-box; T-pool: Tall pool; ZS11: Zhongshuang 11.

## Acknowledgements

Not applicable.

## Authors' contributions

XW and MZ designed the research and wrote the manuscript. HL and LZ carried out the statistical analysis and QTL-seq analysis. FC and WZ performed the fine-mapping of candidate gene. MP participated in the field experiment. SF and MH developed the molecular markers. HW, JZ and WH led and coordinated the overall study. All authors read and approved the final manuscript.

## Funding

The work was supported by National Natural Science Foundation of China (31971973, 31801402), the Earmarked Fund for China Agriculture Research System (CARS-12), Jiangsu Agricultural Science and Technology Innovation Fund (CX(19)3053), China Postdoctoral Science Foundation (2018M630231) and Jiangsu Collaborative Innovation Center for Modern Crop Production.

## Availability of data and materials

The raw sequence data have been deposited in the NCBI (<https://www.ncbi.nlm.nih.gov/sra/>) Sequence Read Archive (SRA) under Accession numbers SRR10915207, SRR10915208, SRR10915209 and SRR10915210. All other relevant data during this study are included in the manuscript and additional files.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

All authors consent for publication.

## Competing interests

The authors declare that they have no competing interests.

Received: 19 January 2020 Accepted: 22 February 2020

Published online: 10 March 2020

## References

- Weiss EA. Oilseed crops. London: Blackwell Publishing Limited; 2000.
- Pullen J, Saeed K. Investigation of the factors affecting the progress of base-catalyzed transesterification of rapeseed oil to biodiesel FAME. *Fuel Process Technol.* 2015;130:127–35.
- Fan C, Tian J, Hu Z, Wang Y, Lv H, Ge Y, Wei X, Deng X, Zhang L, Yang W. Advances of oilseed rape breeding (in Chinese with an English abstract). *J Plant Genet Resour.* 2018;19(3):447–54.
- Zhang B, Ma Y, Geng W, Cui J, Mu K, Hu L. Assessment of rape straw resources for biomass energy production in China (in Chinese with an English abstract). *Renew Energy Resour.* 2017;35(1):126–34.
- Islam N, Evans EJ. Influence of lodging and nitrogen rate on the yield and yield attributes of oilseed rape (*Brassica napus* L.). *Theor Appl Genet.* 1994;88(5):530–4.
- Khan S, Anwar S, Kuai J, Noman A, Shahid M, Din M, Ali A, Zhou G. Alteration in yield and oil quality traits of winter rapeseed by lodging at different planting density and nitrogen rates. *Sci Rep.* 2018;8(1):634.
- Zhou F, Lin Q, Zhu L, Ren Y, Zhou K, Shabek N, Wu F, Mao H, Dong W, Gan L, et al. D14-SCF<sup>D3</sup>-dependent degradation of D53 regulates strigolactone signalling. *Nature.* 2013;504:406–10.
- Jiang L, Liu X, Xiong G, Liu H, Chen F, Wang L, Meng X, Liu G, Yu H, Yuan Y, et al. DWARF 53 acts as a repressor of strigolactone signalling in rice. *Nature.* 2013;504:401–5.
- Sun T. Gibberellin metabolism, perception and signaling pathways in *Arabidopsis*. *The Arabidopsis Book.* 2008;6:e0103.
- Sazuka T, Kamiya N, Nishimura T, Ohmae K, Sato Y, Imamura K, Nagato Y, Koshihara T, Nagamura Y, Ashikari M, et al. A rice *tryptophan deficient dwarf* mutant, *tdd1*, contains a reduced level of indole acetic acid and develops abnormal flowers and organless embryos. *Plant J.* 2009;60(2):227–41.
- Wang Y, Li J. Molecular basis of plant architecture. *Annu Rev Plant Biol.* 2008;59:253–79.
- Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, et al. Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature.* 2002;416(6882):701–2.
- Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, et al. 'Green revolution' genes encode mutant gibberellin response modulators. *Nature.* 1999;400(6741):256–61.
- Lehmann T, Hoffmann M, Henrich M, Pollmann S. Indole-3-acetamide-dependent auxin biosynthesis: a widely distributed way of indole-3-acetic acid production? *Eur J Cell Biol.* 2010;89(12):895–905.
- Petrasek J, Hoyerova K, Motyka V, Hejatk J, Dobrev P, Kaminek M, Vankova R. Auxins and cytokinins in plant development 2018. *Int J Mol Sci.* 2019;20(4):909.
- Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, Weigel D, Chory J. A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science.* 2001;291(5502):306–9.
- Tao Y, Ferrer J, Ljung K, Pojer F, Hong F, Long J, Li L, Moreno JE, Bowman ME, Ivans LJ, et al. Rapid synthesis of auxin via a new tryptophan-dependent pathway is required for shade avoidance in plants. *Cell.* 2008;133(1):164–76.
- Calderon-Villalobos LI, Tan X, Zheng N, Estelle M. Auxin perception-structural insights. *Csh Perspect Biol.* 2010;2(7):5546.
- Lavy M, Estelle M. Mechanisms of auxin signaling. *Development.* 2016;143(18):3226–9.
- Szemenyei H, Hannon M, Long J. TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science.* 2008;319(5868):1384–6.
- Li H, Li J, Song J, Zhao B, Guo C, Wang B, Zhang Q, Wang J, King G, Liu K. An auxin signaling gene *BnaA3.1AA7* contributes to improved plant architecture and yield heterosis in rapeseed. *New Phytol.* 2019;222(2):837–51.
- Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, et al. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science.* 2014;345(6199):950–3.
- Shi J, Li R, Qiu D, Jiang C, Long Y, Morgan C, Bancroft I, Zhao J, Meng J. Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. *Genetics.* 2009;182(3):851–61.
- Basunanda P, Radoev M, Ecke W, Friedt W, Becker HC, Snowden RJ. Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theor Appl Genet.* 2010;120(2):271–81.
- Ding G, Zhao Z, Liao Y, Hu Y, Shi L, Long Y, Xu F. Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in *Brassica napus*. *Ann Bot.* 2012;109(4):747–59.
- Wang X, Wang H, Long Y, Liu L, Zhao Y, Tian J, Zhao W, Li B, Chen L, Chao H, et al. Dynamic and comparative QTL analysis for plant height in different developmental stages of *Brassica napus* L. *Theor Appl Genet.* 2015;128(6):1175–92.
- Luo X, Ma C, Yue Y, Hu K, Li Y, Duan Z, Wu M, Tu J, Shen J, Yi B, et al. Unravelling the complex trait of harvest index in rapeseed (*Brassica napus* L.) with association mapping. *BMC Genomics.* 2015;16(1):379.
- Li F, Chen B, Xu K, Gao G, Yan G, Qiao J, Li J, Li H, Li L, Xiao X, et al. A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant Sci.* 2016;242:169–77.
- Sun C, Wang B, Yan L, Hu K, Liu S, Zhou Y, Guan C, Zhang Z, Li J, Zhang J, et al. Genome-wide association study provides insight into the genetic control of plant height in rapeseed (*Brassica napus* L.). *Front Plant Sci.* 2016;7:1102.
- Zheng M, Peng C, Liu H, Tang M, Yang H, Li X, Liu J, Sun X, Wang X, Xu J, et al. Genome-wide association study reveals candidate genes for control

- of plant height, branch initiation height and branch number in rapeseed (*Brassica napus* L.). *Front Plant Sci.* 2017;8:1246.
31. Shen Y, Xiang Y, Xu E, Ge X, Li Z. Major co-localized QTL for plant height, branch initiation height, stem diameter, and flowering time in an alien introgression derived *Brassica napus* DH population. *Front Plant Sci.* 2018;9:390.
  32. Wang Y, He J, Yang L, Wang Y, Chen W, Wan S, Chu P, Guan R. Fine mapping of a major locus controlling plant height using a high-density single-nucleotide polymorphism map in *Brassica napus*. *Theor Appl Genet.* 2016;129(8):1479–91.
  33. Wang Y, Chen W, Chu P, Wan S, Yang M, Wang M, Guan R. Mapping a major QTL responsible for dwarf architecture in *Brassica napus* using a single-nucleotide polymorphism marker approach. *BMC Plant Biol.* 2016;16(1):178.
  34. Liu C, Wang J, Huang T, Wang F, Yuan F, Cheng X, Zhang Y, Shi S, Wu J, Liu K. A missense mutation in the VHYNP motif of a DELLA protein causes a semi-dwarf mutant phenotype in *Brassica napus*. *Theor Appl Genet.* 2010;121(2):249–58.
  35. Zhao B, Li H, Li J, Wang B, Dai C, Wang J, Liu K. *Brassica napus* DS-3, encoding a DELLA protein, negatively regulates stem elongation through gibberellin signaling pathway. *Theor Appl Genet.* 2017;130(4):727–41.
  36. Zheng M, Hu M, Yang H, Tang M, Zhang L, Liu H, Li X, Liu J, Sun X, Fan S, et al. Three *BnaIAA7* homologs are involved in auxin/brassinosteroid-mediated plant morphogenesis in rapeseed (*Brassica napus* L.). *Plant Cell Rep.* 2019;38(8):883–97.
  37. Fu T, Zou Y. Progress and future development of hybrid rapeseed in China. *Eng Sci.* 2013;11(5):13–8.
  38. Cao X, Liu B, Zhang Y. SEA: a software package of segregation analysis of quantitative traits in plants (in Chinese with an English abstract). *J Nanjing Agric University.* 2013;36(6):1–6.
  39. Zou J, Mao L, Qiu J, Wang M, Jia L, Wu D, He Z, Chen M, Shen Y, Shen E, et al. Genome-wide selection footprints and deleterious variations in young Asian allotetraploid rapeseed. *Plant Biotechnol J.* 2019;17(10):1998–2010.
  40. Song J, Guan Z, Hu J, Guo C, Yang Z, Wang S, Liu D, Wang B, Lu S, Zhou R, et al. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat Plants.* 2020. <https://doi.org/10.1038/s41477-019-0577-7>.
  41. Korasick DA, Westfall CS, Lee SG, Nanao MH, Dumas R, Hagen G, Guilfoyle TJ, Jez JM, Strader LC. Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression. *P Natl Acad Sci USA.* 2014;111(14):5427–32.
  42. Nanao MH, Vinos-Poyo T, Brunoud G, Thévenon E, Mazzoleni M, Mast D, Lainé S, Wang S, Hagen G, Li H, et al. Structural basis for oligomerization of auxin transcriptional regulators. *Nat Commun.* 2014;5(1):3617.
  43. Wang B, Wu Z, Li Z, Zhang Q, Hu J, Xiao Y, Cai D, Wu J, King GJ, Li H, et al. Dissection of the genetic architecture of three seed-quality traits and consequences for breeding in *Brassica napus*. *Plant Biotechnol J.* 2018;16(7):1336–48.
  44. Uttamaprakrom W, Reubroycharoen P, Vitidsant T, Charusiri W. Catalytic degradation of rapeseed (*Brassica napus*) oil to a biofuel using MgO: an optimization and kinetic study. *Journal of the Japan Institute of Energy.* 2017;96(6):190–8.
  45. Chao H, Raboanatahiry N, Wang X, Zhao W, Chen L, Guo L, Li B, Hou D, Pu S, Zhang L, et al. Genetic dissection of harvest index and related traits through genome-wide quantitative trait locus mapping in *Brassica napus* L. *Breeding Sci.* 2019;69(1):104–16.
  46. Lu K, Xiao Z, Jian H, Peng L, Qu C, Fu M, He B, Tie L, Liang Y, Xu X, et al. A combination of genome-wide association and transcriptome analysis reveals candidate genes controlling harvest index-related traits in *Brassica napus*. *Sci Rep-UK.* 2016;6(1):36452.
  47. Salehin M, Bagchi R, Estelle M. SCFTIR1/AFB-based auxin perception: mechanism and role in plant growth and development. *Plant Cell.* 2015;27(1):9–19.
  48. Kagale S, Rozwadowski K. EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. *Epigenetics-US.* 2011;6(2):141–6.
  49. Guilfoyle TJ. The PB1 domain in auxin response factor and Aux/IAA proteins: a versatile protein interaction module in the auxin response. *Plant Cell.* 2015;27(1):33–43.
  50. Zhao X, Ma W, Gale KR, Lei Z, He Z, Sun Q, Xia X. Identification of SNPs and development of functional markers for LMW-GS genes at Glu-D3 and Glu-B3 loci in bread wheat (*Triticum aestivum* L.). *Mol Breeding.* 2007;20(3):223–31.
  51. Jiang G. Molecular marker-assisted breeding: a plant breeder's review. *Advances in plant breeding strategies: breeding, biotechnology and molecular tools.* Switzerland: Springer; 2015.
  52. De la Vega FM, Lazaruk KD, Rhodes MD, Wenz MH. Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP genotyping assays and the SNPlex genotyping system. *Mutat Res.* 2005;573:111–35.
  53. Li H, Li J, Zhao B, Wang J, Yi L, Liu C, Wu J, King G, Liu K. Generation and characterization of tribenuron-methyl herbicide-resistant rapeseed (*Brassica napus*) for hybrid seed production using chemically induced male sterility. *Theor Appl Genet.* 2015;128(1):107–18.
  54. Wang X, Yu K, Li H, Peng Q, Chen F, Zhang W, Chen S, Hu M, Zhang J. High-density SNP map construction and QTL identification for the apetalous character in *Brassica napus* L. *Front Plant Sci.* 2015;6:1164.
  55. Zhao W, Wang X, Wang H, Tian J, Li B, Chen L, Chao H, Long Y, Xiang J, Gan J, et al. Genome-wide identification of QTL for seed yield and yield-related traits and construction of a high-density consensus map for QTL comparison in *Brassica napus*. *Front Plant Sci.* 2016;7:17.
  56. Wang X, Chen L, Wang A, Wang H, Tian J, Zhao X, Chao H, Zhao Y, Zhao W, Xiang J, et al. Quantitative trait loci analysis and genome-wide comparison for silique related traits in *Brassica napus*. *BMC Plant Biol.* 2016;16(1):71.
  57. Chen F, Zhang W, Yu K, Sun L, Gao J, Zhou X, Peng Q, Fu S, Hu M, Long W, et al. Unconditional and conditional QTL analyses of seed fatty acid composition in *Brassica napus* L. *BMC Plant Biol.* 2018;18(1):49.
  58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
  59. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9.
  60. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–303.
  61. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep.* 1997;15(1):8–15.
  62. Zhang B, Zhao N, Liu Y, Jia L, Fu Y, He X, Liu K, Xu Z, Bao B. Novel molecular markers for high-throughput sex characterization of *Cynoglossus semilaevis*. *Aquaculture.* 2019;513:734331.
  63. Lu J, Hou J, Ouyang Y, Luo H, Zhao J, Mao C, Han M, Wang L, Xiao J, Yang Y, et al. A direct PCR-based SNP marker-assisted selection system (D-MAS) for different crops. *Mol Breeding.* 2020;40(9):1–10.
  64. Yu K, Wang X, Chen F, Chen S, Peng Q, Li H, Zhang W, Hu M, Chu P, Zhang J, et al. Genome-wide transcriptomic analysis uncovers the molecular basis underlying early flowering and apetalous characteristic in *Brassica napus* L. *Sci Rep-UK.* 2016;6(1):30576.
  65. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;25(9):1105–11.
  66. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–9.
  67. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010;11(2):14.
  68. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li C, Wei L. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 2011;39:316–22.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.