RESEARCH ARTICLE

Genome-wide identification AINTEGUMENTA-like (AIL) genes in Brassica species and expression patterns during reproductive development in Brassica napus L.

Shulin Shen¹,²*, Fujun Sun¹,²*, Meichen Zhu¹,², Si Chen¹,², Mingwei Guan¹,², Rui Chen¹,², Fang Tang¹,², Nengwen Yin¹,², Xinfu Xu¹,², Zhanglin Tang¹,², Jiana Li¹,², Kun Lu¹,²*, Fang Tang¹,², Nengwen Yin¹,², Xinfu Xu¹,², Zhanglin Tang¹,², Jiana Li¹,², Kun Lu¹,²*, Cunmin Qu¹,²*

¹ Chongqing Rapeseed Engineering Research Center, College of Agronomy and Biotechnology, Southwest University, Beibei, Chongqing, China, ² Academy of Agricultural Sciences, Southwest University, Chongqing, China

* These authors contributed equally to this work.

* drlukun@swu.edu.cn (KL); drqucumin@swu.edu.cn (CQ)

Abstract

The AINTEGUMENTA-like (AIL) proteins, which belong to the AP2 family, play important roles in regulating the growth and development of plant organs. The AIL family has not yet been comprehensively studied in rapeseed (Brassica napus), an allotetraploid and model organism for the study of polyploid evolution. In the present study, 99 AIL family genes were identified and characterized from B. rapa, B. oleracea, B. napus, B. juncea, and B. nigra using a comprehensive genome-wide study, including analyses of phylogeny, gene structure, chromosomal localization, and expression pattern. Using a phylogenetic analysis, the AIL genes were divided into eight groups, which were closely related to the eight AtAIL genes, and which shared highly conserved structural features within the same subfamily. The non-synonymous/synonymous substitution ratios of the paralogs and orthologs were less than 1, suggesting that the AIL genes mainly experienced purifying selection during evolution. In addition, the RNA sequencing data and qRT-PCR analysis revealed that the B. napus AIL genes exhibited organ- and developmental stage-specific expression patterns. Certain genes were highly expressed in the developing seeds (BnaAIL1, BnaAIL2, BnaAIL5, and BnaAIL6), the roots (BnaANT, BnaAIL5, and BnaAIL6), and the stem (BnaAIL7). Our results provide valuable information for further functional analysis of the AIL family in B. napus and related Brassica species.

Introduction

Plant growth and developmental processes are influenced by the complex external environment and internal developmental factors [1, 2]. Deciphering the molecular networks contributing to plant growth and development is therefore an important research goal. Over the past...
few decades, many internal regulatory factors contributing to plant growth and development have been described, including AINTEGUMENTA (ANT) [1, 3–6], ANT-LIKE/PLETHORA (AIL/PLT) [4, 7, 8], AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS) [9, 10], and various growth-regulating factors (GRFs) [11–14]. Among them, the AIL genes comprise an APETALA 2 (AP2) subfamily known to be involved in growth-related processes in a variety of plants [3, 4, 15–18]. These genes play an important role in regulating the growth and development of organs (e.g., embryos and flowers) [1, 3, 19–22]; for example, loss-of-function mutations in ANT resulted in A. thaliana with smaller organs [18, 21, 23], while overexpression of the ANT genes caused increases in organ size [1, 24, 25]. In addition, some AIL genes were shown to be involved in the differentiation of embryogenic stem cells from somatic cells in A. thaliana [26–28], oil palm (Elaeis guineensis) [29], and coconut (Cocos nucifera) [30, 31]. The ectopic expression of EgAP2-1 alters leaf morphology and enhances the regeneration capacity of the oil palm [29], while coconut nucellar development is also regulated by AIL genes [30]. Transgenic Arabidopsis plants expressing AtAIL6 exhibited changes in floral organ size and morphology associated with alterations in the pattern and duration of cell divisions within the developing organs, while the ant ail6 double mutant displayed a premature differentiation of their floral meristem cells [28]. Numerous studies have also revealed that the AIL family genes are involved in root development and abiotic stress responses, including AIL6/PLETHORA3 (PLT3), PLT1, PLT2, and BABY BOOM (BBM) [17, 32]. As the excellent evolutionary model to investigate the expansion of gene families [33], the AIL family members have not been well studied in the Brassica genus.

In the present study, we investigated the AIL family members in various Brassica species (B. rapa, B. oleracea, B. napus, B. juncea, and B. nigra) using a genome-wide bioinformatics analysis, exploring their exon-intron organization, motif compositions, gene duplications, chromosome distribution, phylogeny, and synteny. We also examined the expression patterns of selected B. napus AIL genes in different tissues. Based on these data, the functions of the AIL genes in B. napus were predicted, providing a reference for the functional verification and utilization of the AIL family in B. napus and other Brassica species.

**Materials & methods**

**Identification and characterization of AIL family proteins**

The amino acid sequences of the AIL proteins were obtained from The Arabidopsis Information Resource (TAIR10) database (ftp://ftp.arabidopsis.org) and used as queries for a BLASTp search of the whole Brassica genome sequences stored in the Brassica database [34], including the B. rapa genome V3.0, B. oleracea genome V1.1, B. napus genome V5, B. juncea genome V1.5, and B. nigra genome V1.1 (http://brassicadb.org/brad/index.php), respectively. Candidate sequences with E-values \( \leq 1 \times 10^{-10} \) were identified and confirmed using the NCBI CD Searches-Tool (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi). A BLAST search of the Brassica protein database was performed to search for the AIL genomic sequences using the NCBI blast+ software (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/). The coding sequences (CDS) of the AILs were identified using BLASTn searches of the Brassica genome database. The candidate proteins were named using the species abbreviation of the source organism (italicized), the name of the clade determined in the subsequent phylogenetic analysis (below), and the position in the clade (e.g., BolaIL1A and BnaAIL1A). The physicochemical properties of each deduced protein, including the molecular weight (MW), isoelectric point (pI), and the grand average of hydropathy (GRAVY) value, were determined using the online ExPASy-ProtParam tool (http://web.expasy.org/protparam/).
Multiple sequence alignment and phylogenetic analysis of the AIL family

The deduced amino acid sequences of the AIL proteins from *A. thaliana* and various *Brassica* species, including *B. rapa*, *B. oleracea*, *B. napus*, *B. juncea*, and *B. nigra*, were subjected to multiple protein sequence alignments using the ClustalW software (Version 2.0) with default settings [35]. To illustrate the evolutionary relationships of the AILs in the *Brassica* genus, a neighbor-joining (NJ) phylogenetic tree was generated using the MEGA v6.0 program (Tokyo Metropolitan University, Tokyo, Japan) with the best-fit model, a Jones-Taylor-Thornton (JTT) matrix and incorporates a proportion of invariant sites (+I) and the gamma distribution for modeling rate heterogeneity (+G). We performed NJ analysis in MEGA v6.0 with bootstrap test with 1,000 replicates [36]. The phylogenetic trees were visualized using evolview v3 (https://www.evolgenius.info/evolview/) [37].

Conserved motif recognition and gene structure analysis

The exon-intron structures of the AIL family genes were analyzed using the TBtools software (https://github.com/CJ-Chen/TBtools). Conserved motifs were identified using Multiple Expectation Maximization for Motif Elucidation (MEME v4.12.0, http://meme-suite.org/tools/meme) with the following parameters: number of repetitions, any; maximum number of motifs, 10; and optimum width of each motif, between 6 and 300 residues [38].

Chromosomal distribution and gene duplication

All the identified AIL family genes were mapped to their respective chromosomes based on the physical location information from the *Brassica* species genome database using MapChart v2.0 (https://www.wur.nl/en/show/Mapchart.htm) [39]. A Multiple Collinearity Scan toolkit (MCScanX) was adopted to analyze the gene syntenic events, using the default parameters [40]. To examine the syntenic relationships of the orthologous AIL family genes obtained from *B. napus* and other selected species, synteny maps were constructed using the Circos software [41]. Non-synonymous (ka) and synonymous (ks) substitutions for each of the duplicated AIL family genes were calculated using KaKs_Calculator v2.0 [42].

Total RNA extraction and qRT-PCR analysis

Total RNA was isolated from the samples using a DNA away RNA Mini-Prep Kit (Sangon Biotech, Shanghai, China). For the tissue-specific expression analysis, RNA was extracted from the roots, stems, leaves, buds, flowers, and seeds, and pretreated with gDNA Eraser (Takara Bio, Kusatsu, Japan). Subsequently, 1 μg total RNA was used to synthesize the first-strand cDNA using an RNA PCR Kit (AMV) v3.0 (Takara Bio). The cDNA was subjected to a RT-qPCR analysis using SYBR Premix Ex Taq II (Takara Bio) on a Bio-Rad CFX96 Real-Time System (Bio-Rad Laboratories, Hercules, CA, USA), as previously described [43]. *BnACTIN7* (EV116054) was used as a reference gene to normalize the AIL gene expression levels via the $2^{-ΔΔCt}$ method [44]. Three technical replicates were performed for all experiments. The specific primer sequences used in this study were obtained from the qPCR Primer Database [45] and are listed in S1 Table.

Results

Identification and characterization of the AIL family genes

Eight *A. thaliana* AIL family protein sequences were acquired from the TAIR10 database and used as queries to identify the AIL family in various *Brassica* species (*B. rapa*, *B. oleracea*, *B. napus*, *B. juncea*, and *B. nigra*) using the BLASTp program. In total, 99 AIL family proteins...
were identified in these species, including 26 in *B. juncea*, 29 in *B. napus*, 15 in *B. oleracea*, 14 in *B. nigra*, and 15 in *B. rapa* (Table 1 and S2 Table). More AIL family proteins were identified in the *Brassica* species than in *A. thaliana*. Based on their homology with the corresponding *A. thaliana* AIL family genes, the identified *Brassica* AIL family genes were named ANT or AIL1–7 (Table 1 and S2 Table). A species-specific prefix was included, while a capital letter suffix was used to represent the gene number within each clade.

The lengths of the AIL protein sequences ranged from 290 (BnaAIL1A) to 652 (BjuANTD) amino acids were almost distributed across the whole chromosomes; the highest content was on chromosome BniB02, including 5 AIL genes. The MW varied from 32.52 (BnaAIL1A) to 71.78 kDa (BjuANTD), and the pIs ranged from 5.47 (BjuAIL3A) to 9.56 (BnaANTB and BniANTA), with 23 pIs > 7 and the remaining pIs ≤ 7 (S2 Table).

**Phylogenetic and classification analysis of the AIL proteins**

To investigate the evolutionary relationships among the AIL family, the protein sequences of AILs from the *A. thaliana* and various *Brassica* species were used to generate the phylogenetic tree in this study, we constructed a NJ phylogenetic tree using the *A. thaliana* AIL proteins as a reference. We showed that the 107 *A. thaliana* and *Brassica* AIL protein sequences were classified into eight clades: the ANT clade and clades AIL1 to AIL7 (Fig 1). However, no BniAIL1 homologs were found in *B. nigra*, and the AIL3 and AIL4 subgroups were located in the same phylogenetic branch, indicating that these genes are more closely related to each other than to the other clades. Their similarity may be related to their shared involvement in the development of the lateral root primordia [21]. In general, the AIL proteins in the allotetraploids (*B. napus* and *B. juncea*) and their diploid progenitors (*B. rapa*, *B. oleracea*, and *B. nigra*) were related to their corresponding *A. thaliana* homologs in each clade (Fig 1), suggesting that the AIL proteins among these species have close evolutionary relationships.

**Gene structure and conserved motif analysis of the AIL family genes**

To further investigate the AIL proteins in each clade, their corresponding gene structures and conserved motifs were analyzed (Fig 2). Accordingly, the numbers of exons/introns within each AIL family clade were similar to each other, whether they originated from the allotetraploids (*B. napus* and *B. juncea*) or their diploid progenitors (*B. rapa*, *B. oleracea*, and *B. nigra*). A statistical analysis revealed that their numbers of exons typically ranged from six to nine (in 91% (97/107) of AIL family genes), indicating that the structures of the AIL family genes were conserved during polyploidization. BnaAIL2B contained the fewest exons (five), while the highest exon numbers were found in BnaAIL1A (10), BnaAIL3B (10), BnaAIL6D (10), and

**Table 1. Statistics of AIL family genes between *A. thaliana* and five *Brassica* species.**

| Gen Family | *A. thaliana* | *B. rapa* | *B. oleracea* | *B. nigra* | *B. juncea* | *B. napus* |
|------------|--------------|-----------|---------------|-----------|-------------|-----------|
| ANT        | 1            | 3         | 3             | 3         | 5           | 5         |
| AIL1       | 1            | 1         | 1             | 0         | 1           | 3         |
| AIL2       | 1            | 2         | 2             | 2         | 4           | 2         |
| AIL3       | 1            | 2         | 2             | 2         | 4           | 5         |
| AIL4       | 1            | 2         | 2             | 2         | 4           | 3         |
| AIL5       | 1            | 1         | 1             | 1         | 2           | 2         |
| AIL6       | 1            | 3         | 3             | 3         | 5           | 6         |
| AIL7       | 1            | 1         | 1             | 1         | 2           | 2         |
| Total      | 8            | 15        | 15            | 14        | 26          | 29        |

https://doi.org/10.1371/journal.pone.0234411.t001
Fig 1. Neighbor-Joining (NJ) phylogenetic tree of the AIL family proteins in *A. thaliana* and various *Brassica* species. The AIL family was divided into eight clades (ANT and AIL1–7), which are indicated by different colors. The red, white, and black stars indicate *B. napus*, *B. rapa*, and *B. oleracea*, respectively. The red, white, and black circles indicate *A. thaliana*, *B. juncea*, and *B. nigra*, respectively.

https://doi.org/10.1371/journal.pone.0234411.g001
BnaAIL6E (11), respectively (Fig 2A and S2 Table). In general, the exon-intron patterns within the same phylogenetic classification group shared the high similarity between the A. thaliana and Brassica species (Fig 2A and S2 Table), indicating that they might be resulted by the replication of these sequences and supporting that the classification result is reliable.

In addition, the conserved motifs in the AIL proteins were predicted using MEME v4.12.0 (http://meme-suite.org/tools/meme). A total of ten conserved motifs were identified in the 107 AIL family members from A. thaliana and the selected Brassica species. The number of conserved motifs are different in each subgroup; for example, proteins in the AIL1 and AIL5 subfamilies contained eight motifs; those in the ANT, AIL2, AIL3, and AIL4 subfamilies had nine; and the AIL6 and AIL7 subfamilies contained ten (Fig 2B). Among these, the same conserved motifs were also widely observed in the paralogous/orthologous AIL family members; for instance, motifs 1 and 2 were found in every AIL family (Fig 2B), suggesting that they have a conserved position and functional similarity between A. thaliana and Brassica species. In addition, motif 9 was distributed in both the AIL6 and AIL7 subfamily, but motif 8 was not detected in the AIL5 subfamily (Fig 2B), indicating that these motifs were selectively distributed in certain AIL proteins. This specific distribution suggests that these motifs may have specific functions in the A. thaliana and Brassica AILs.

Conserved amino acid sequences within the AP2 domain

To investigate the sequences of the conserved AP2 domains in A. thaliana and the Brassica species, a multiple sequence alignment was performed using the 107 AIL proteins identified from A. thaliana, B. rapa, B. oleracea, B. nigra, B. napus, and B. juncea (Fig 3, S1 Fig and S3 Table). Two AP2 domains (AP2-R1 and AP2-R2) were located near the N- and C-terminal regions of the AIL proteins (Fig 3 and S1 Fig), which was consistent with previously published results [4, 6, 22]. These two AP2 domain regions were highly conserved in the AIL proteins. The lengths of the two AP2 domains were nearly constant between the AIL proteins, but varied in some of cases, such as for BnaAIL1C, BnaAIL2B, BnaAIL2D, BnaAIL3D, BjuAIL3A, BjuAIL3D, BjuAIL4B, BnaAIL4B, BnaAIL4D, and BjuAIL6F.

In addition, the AP2 domains all contained YRG and RAYD elements [46], the latter of which comprise a highly conserved 18-amino-acid core region predicted to form an amphipathic α-helix in the AP2 domains. The length of the RAYD α-helix was also highly conserved in most AP2 domains, except for BnaAIL1C, BnaAIL3D, BjuAIL4B, and BnaAIL4D (Fig 3 and S1 Fig). We found that the glycine residues within the RAYD element, which are involved in AP2 function [47], were identical in all the AP2-domain-containing proteins; therefore, the structure or function of the AP2 domains is likely to be associated with the invariant amino acid residues within the RAYD and RAYD elements [46, 47]. Additionally, the sequences (25 aa) between the AP2-R1 and AP2-R2 domain were named as the linker regions with highly conserved, except in BjuAIL2B, BolAIL2B, BjuAIL3C, BolAIL3B, BnaAIL4A, BnaAIL4C, and BraAIL4A (Fig 3 and S1 Fig). A detailed description of the two AP2 domains is provided in S3 Table.

Chromosomal localization analysis of the AIL genes among the Brassica species

The genome sequences of the allotetraploid species (B. napus and B. juncea) and their diploid progenitors (B. rapa, B. oleracea, and B. nigra) were acquired from the Brassica Database (http://brassicadb.org/brad/index.php), and the locations of the identified AIL family genes were drafted onto the corresponding chromosomes using Mapchart v2.0 software. As a result, 90 of the AIL family genes in the various Brassica plants could be mapped onto the A (39), B (26), and C (25) subgenomes, while nine were distributed onto different random chromosome
Fig 2. Characteristics of the identified AIL family genes and proteins in A. thaliana and selected Brassica species. A) Exons and introns are represented by green boxes and gray lines, respectively. B) The conserved motifs of the AIL proteins. A total of 10 motifs (number 1–10) were identified using MEME v4.12.0 (http://meme-suite.org/tools/meme), and are indicated as differently colored boxes. Yellow boxes represent upstream or downstream untranslated regions. Bra, B. rapa; Bol, B. oleracea; Bni, B. nigra; Bna, B. napus; Bju, B. juncea.

https://doi.org/10.1371/journal.pone.0234411.g002
and scaffold sequences that had not been assembled into the corresponding chromosomes (Fig 4, S2 Table). The *AIL* genes were unevenly distributed on the chromosomes, with between one and five genes on each. A comparison of the gene distributions of the allotetraploid species (*B. napus* and *B. juncea*) and their diploid progenitors (*B. rapa*, *B. oleracea*, and *B. nigra*) revealed the
important result that many AIL family genes retained their relative positions in A_{Bra}, A_{Bni}, and A_{Bju}; B_{Bni} and B_{Bju}; and C_{Bol} and C_{Bna}. For example, the ANT genes were located on chromosomes A01 and A08; the AIL6 genes were present on chromosomes A02, A03, and A10; and the AIL3 and AIL4 were located on chromosome A05; the same patterns was also repeated on the B and C subgenomes (Fig 4B and 4C). This similarity suggests that these genes might have undergone whole-genome duplication events during the evolutionary process, and might have similar functions. In addition, some genes (e.g., AIL1, AIL2, and AIL3 in A02; ANT in A03; and AIL4 in A06) might have been lost during the evolution of B. juncea and B. napus due to the incomplete assembly of their chromosomes during their hybridization and polyploidization. Together, these results shed light on the evolutionary patterns in these subfamilies among related species.

Synteny and duplicated gene analysis of the AIL family genes in B. rapa, B. oleracea, and B. napus

To investigate the patterns of retention or loss in the orthologous AIL family genes, we compared the relationships of the AIL genes between A. thaliana and B. rapa, B. oleracea, and B.
napus (S4 Table and S2 Fig). Genes of the same clade were identified on many chromosomes (S2 Fig), suggesting that they were evolutionarily related and that most AIL genes were preserved during polyploidization.

In addition, we compared the syntetic relationship of the AIL genes in A. thaliana, the allotetraploid B. napus (A<sub>Bna</sub> and C<sub>Bna</sub>) and its diploid progenitors B. rapa (A<sub>Bra</sub>) and B. oleracea (C<sub>Bol</sub>), according to their corresponding syntetic information obtained from the BRAD database. A total of 15 BraAIL genes and 13 BolAIL genes showed a syntetic relationship with the eight AtAIL genes and 22 BnaAIL genes (Fig 5, S4 Table). Furthermore, the numbers of orthologous pairs identified in the comparisons of AtAIL and BraAIL, AtAIL and BolAIL, AtAIL and BnaAIL, BraAIL and BnaAIL, and BolAIL and BnaAIL were 23, 19, 26, 39, and 48, respectively. These results showed that the syntetic AIL gene pairs were widely distributed on the genomes of the allotetraploid (B. napus) and its diploid progenitors (B. rapa and B. oleracea).

We also calculated the nonsynonymous substitutions (Ka), synonymous substitutions (Ks), and Ka/Ks ratios of the AIL gene pairs to identify the evolutionary constraints acting on the AIL gene pairs, revealing that the Ka/Ks values of all orthologous AIL gene pairs were less than 1 (S4 Table). This suggests that the AIL family genes in B. napus and its diploid progenitors might have experienced strong purifying selective pressure after the duplication events.

Expression profiles of the BnaAIL family genes in various B. napus organs

To investigate the putative functions of the BnaAIL family genes in regulating the growth and development of B. napus, we characterized the expression profiles of the BnaAIL genes in different tissues. This was achieved using the transcriptome sequencing datasets of B. napus ZS11 stored in National Genomics Data Center (BioProject ID PRJNA358784), which covered all stages of B. napus development and a variety of organs, including the roots, hypocotyl, cotyledon, stems, leaves, anthocaulus, buds, calyx, petals, pistil, stemans, anthers, capillament, initial apex, seeds, embryo, seed coat, and silique pericarp (Fig 6 and S5 Table). The expression profiles of these AIL family genes showed clear differences among these tissues, except for BnaAIL1C and BnaAIL3B that were not highly expressed in any of the tissues, suggesting that the genes of this family might perform a variety of biological functions in different tissues.

Furthermore, the members of specific AIL family clades displayed similar characteristics; for example, BnaAIL1 to BnaAIL4 were generally only expressed in the roots, hypocotyl, cotyledon, developing seed, embryo, and seed coat. The BnaAIL5, BnaAIL6, BnaAIL7, and BnaANT clades were widely expressed in all tissues, especially in the younger tissues; for example, the BnaAIL5 and BnaAIL6 family members showed higher expression levels in the developing seed, embryo, and seed coat. BnaAIL5, BnaAIL7, and BnaANT members were also expressed in the roots, hypocotyl, cotyledon, and stem, with BnaAIL5 and BnaANT also being expressed in the anthocaulus. BnaAIL5, BnaAIL6, and BnaANT were particularly highly expressed in the pistil and, in addition to BnaAIL7, in the initial apex. Our results suggest that the AIL family genes play important roles in the processes of growth and development in B. napus.

Expression patterns of the BnaAIL genes revealed using qRT-PCR analysis

To decipher the physiological functions of the B. napus AIL family genes, we analyzed the expressions of 20 randomly selected AIL genes in eight different B. napus tissues under normal growth conditions using qRT-PCR (Fig 7). Of these, 17 were more highly expressed in the developing seeds, which is consistent with the fact that the AIL proteins are master regulators of developmental processes, especially during embryogenesis [17, 26, 29]. In addition, AIL proteins are also required for the development of the floral and roots organs [1, 4, 7, 10, 48].
Some AIL genes, including *BnaAIL3B*, *BnaAIL4B*, and *BnaANTA*, also showed higher expression levels in the roots and flowers. Furthermore, *BnaAIL7B* was notably highly expressed in...
the stems. These results further highlight that the AIL family genes are involved in the vegetative and reproductive growth in *B. napus*, and especially in seed development.

**Discussion**

The AIL family genes belong to the AP2/ERF superfamily, the members of which are master regulators of plant growth and development, especially of embryogenesis [16, 17, 49]. Furthermore, the cruciferous plants arose from a common ancestor, and have undergone genome duplications and merging during the evolutionary process [50]. Using the eight *A. thaliana* AIL protein sequences as a reference, therefore, we identified 99 putatively AIL proteins from

---

Fig 6. Expression profiles of the *BnaAIL* family genes in different tissues and organs. The abbreviations above the heatmap represent the different tissues and organs/developmental stages of *B. napus* ZS11, and are listed in S5 Table. The bar represents the log2 expression levels (FPKM). Black boxes indicate that no expression was detected in an RNA-seq analysis.

https://doi.org/10.1371/journal.pone.0234411.g006
various *Brassica* species in this study. Among them, 15, 15, and 14 AIL family members were identified in the diploid species *B. rapa*, *B. oleracea*, and *B. nigra*, while 26 and 29 AIL family members were found in the allotetraploids *B. juncea* and *B. napus* (Table 1 and S2 Table). More AIL family members were identified in the *Brassica* species than in *A. thaliana*, suggesting that the AIL family genes had undergone a whole-genome triplication among them since

---

**Fig 7.** Expression patterns of the *B. napus* AIL family genes in different tissues, revealed using qRT-PCR. Ro: root; St: stem; Le: leaf; Bu: bud; Fl: flowers; Se_10d, Se_30d, and Se_50d: seeds 10, 30, and 50 days after flowering. The mean expression values were calculated from three independent replicates. Error bars indicate the standard deviation.

https://doi.org/10.1371/journal.pone.0234411.g007
their divergence from *A. thaliana*, resulting in a significant increase in the numbers of the duplicated genes [50, 51, 52]. Although the allotetraploid *B. napus* was formed by the natural hybridization and polyploidization of *B. rapa* and *B. oleracea* [50], the numbers of *AIL* family genes in these three species were almost equal, indicating that the expansion of the *AIL* family was largely a result of earlier whole-genome and segmental duplications [53]. However, different numbers of *AILs* were identified in the allotetraploid species (*B. napus* and *B. juncea*) and their parental species (*B. rapa*, *B. oleracea*, and *B. nigra*), suggesting that gene loss or duplication events might have occurred in the *AIL* family genes during the polyploidization of *B. napus* and *B. juncea*. Additionally, phylogenetic analysis revealed that all the *AIL* family genes could be divided into eight subgroups (Fig 2), which were closely associated with the *AtAIL* groups [17], suggesting that they might share similar functions in the same subgroup.

Previous research revealed that the *AIL* proteins are members of the AP2 subfamily, part of the AP2/ERF superfamily [16]. Our analysis revealed that two AP2 domains were conserved among all the *AIL* family proteins, and our sequence comparisons revealed two conserved motifs, referred to as the YRG and RAYD elements, within the AP2 domains (Fig 3 and S1 Fig). These results strongly suggested that the AP2 domain is an important and evolutionarily conserved region necessary for the correct structure or function of the *AIL* family proteins. The amphipathic α-helices in the RAYD elements were also highly conserved, except in BnaAIL1C, BnaAIL3D, BjuAIL4B, and BnaAIL4D (Fig 3 and S1 Fig), suggesting that these domains might be involved in DNA binding through the interaction of their hydrophobic face with the major groove of DNA [46, 54]. Additionally, the lengths of the AP2-R1 domains were different among the *AIL* family proteins (Fig 3 and S1 Fig), consistent with previous findings [15, 22], suggesting that they may contribute to differences in the functional specificities of these proteins.

Numerous studies have shown that the *AIL* proteins were widely involved in the plant growth and developmental processes in young, dividing tissues, including the roots, shoots, floral organs, leaves, and seeds [1, 4, 8, 22, 25, 55]. In the present study, most of the *AIL* genes were expressed at high levels in these tissues, especially in the seeds, embryos, roots, hypocotyls, and cotyledons (Figs 6 and 7), suggesting that these genes may play a role during their development. The expression patterns of some duplicated genes also displayed differences, suggesting that they might have undergone functional divergence after their duplication; for example, *BnaAIL1A* and *BnaAIL1B* were highly expressed in the developing seeds, *BnaAIL1C* was hardly expressed in any of the tissues, and the expression profile of *BnaAIL7A* was completely different to that of *BnaAIL7B* (Fig 6). Additionally, we noticed that *BnaANT* and *BnaAIL6* were expressed in the reproductive tissues (e.g., root, pistil, initial apex, developing seed, embryo, and seed coat), which was consistent with previous results [4, 7, 20, 28], indicating they may play similar roles in *B. napus* and *A. thaliana*. The expression patterns of these *AIL* family genes in the allotetraploid *B. napus* were similar to those observed in its diploid progenitor *B. rapa*, which had the higher expression levels in young tissues [22], suggesting they may play similar roles in both species. In addition, we found that *AIL* genes showed the similar expression patterns within the same subgroups (Figs 1, 6 and 7), implying that the invariant amino acid residues within the YRG and RAYD elements were controlled by the structure or function of the AP2 domains [46, 47]. Taken together, our results provide the new clues for investigating the roles of *AILs* in *B. napus*.

**Conclusions**

In this study, 99 *AIL* family genes were identified from five *Brassica* species, which could be divided into eight subgroups and had closely relationship with the *AtAILs*. Furthermore, the
AIL family genes shared a high similarity among the gene structure, conserved motifs within the same subgroups. The Ka/Ks ratios of orthologous AIL gene pairs among *A. thaliana* and *Brassica* indicates that the AIL genes had undergone strong purifying selection for retention. Additionally, RNA-Seq and qRT-PCR results indicated that the AIL family genes might be involved in regulating *B. napus* development, especially in the developing seeds. These results enhance the understanding of the evolution and function of AIL family genes in *B. napus*, providing valuable clues for further research.

**Supporting information**

S1 Fig. Sequence alignment of all identified AILs from *Arabidopsis* and various *Brassica*. The regions of AP2-R1 to AP2-R2 are shown with blue line. Blue shading represents identical conserved amino acid residues. Color shading represents an α-helix. A detailed description of the two AP2 domains is provided in S3 Table. Bra, *B. rapa*; Bol, *B. oleracea*; Bni, *B. nigra*; Bna, *B. napus*; Bju, *B. juncea.

(TIF)

S2 Fig. Genome-wide syntenic analysis of all identified AIL family genes among *A. thaliana*, *B. rapa*, *B. oleracea*, and *B. napus*. The syntenic genes are linked with the red (A subgenome) and light green lines (C subgenome), respectively.

(TIF)

S1 Table. Specific primers used to amplify the AIL and reference genes using a qRT-PCR analysis.

(XLSX)

S2 Table. List of AIL family genes identified from *A. thaliana* and *Brassica* species.

(XLSX)

S3 Table. Details of the AP2 domain in the AIL proteins in *A. thaliana* and *Brassica* species.

(XLSX)

S4 Table. The orthologous AIL gene pairs among *A. thaliana* and *Brassica* species.

(XLSX)

S5 Table. *B. napus* ZS11 tissues and organs used in this study.

(XLSX)

**Acknowledgments**

We would like to thank Kathy Farquharson for critical reading of this manuscript.

**Author Contributions**

**Data curation:** Shulin Shen.

**Formal analysis:** Fujun Sun, Meichen Zhu, Si Chen.

**Investigation:** Meichen Zhu.

**Methodology:** Si Chen, Fang Tang.

**Resources:** Xinfu Xu, Zhanglin Tang.

**Software:** Mingwei Guan, Rui Chen, Nengwen Yin.
Supervision: Jiana Li, Kun Lu, Cunmin Qu.

Writing – original draft: Shulin Shen, Fujun Sun, Kun Lu, Cunmin Qu.

Writing – review & editing: Jiana Li, Kun Lu, Cunmin Qu.

References

1. Mizukami Y, Fischer RL. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. Proc Natl Acad Sci USA. 2000; 97(2):942–7. https://doi.org/10.1073/pnas.97.2.942 PMID: 10639184.

2. Scheres B, Krizek BA. Coordination of growth in root and shoot apices by AIL/PLT transcription factors. Curr Opin Plant Biol. 2018; 41:95–101. https://doi.org/10.1016/j.pbi.2017.10.002 PMID: 29172412.

3. Elliott RC, Betzner AS, Huttner E, Oakes MP, Tucker WQ, Gerentes D, et al. AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. Plant Cell. 1996; 8(2):155–68. https://doi.org/10.1105/tpc.8.2.155 PMID: 8742707.

4. Nole-Wilson S, Tränby TL, Krizek BA. AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. Plant Mol Biol. 2005; 57(5):613–28. https://doi.org/10.1007/s11103-005-0955-6 PMID: 15988559.

5. Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. AP2/ERF family transcription factors in plant abiotic stress responses. Biochim Biophys Acta. 2012; 1819(2):86–96. https://doi.org/10.1016/j.bbagrm.2011.08.004 PMID: 21867785.

6. Chialva C, Eichler E, Grissi C, Muñoz C, Gomez-Talquinca S, Martínez-Zapater JM, et al. Expression of grapevine AINTEGUMENTA-like genes is associated with variation in ovary and berry size. Plant Mol Biol. 2016; 91(1):67–80. https://doi.org/10.1007/s11103-016-0443-1 PMID: 26843119.

7. Krizek B. AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. Plant Physiol. 2009; 150(4):1916–29. https://doi.org/10.1104/pp.109.141119 PMID: 19542297.

8. Rigal A, Yordanov YS, Perrone I, Karlberg A, Tisserant E, Bellini C, et al. The AINTEGUMENTA LIKE1 Homeotic Transcription Factor PtAIL1 Controls the Formation of Adventitious Root Primordia in Poplar. Plant Physiol. 2012; 160(4):1996–2006. https://doi.org/10.1104/pp.112.204453 PMID: 23077242.

9. Hu Y, Xie Q, Chua NH. The Arabidopsis Auxin-Inducible Gene ARGOS Controls Lateral Organ Size. Plant Cell. 2003; 15(9):1951–61. https://doi.org/10.1105/tpc.011119 PMID: 12953103.

10. Hu Y, Poh HM, Chua NH. The Arabidopsis ARGOS-LIKE gene regulates cell expansion during organ growth. Plant J. 2006; 47(1):1–9. https://doi.org/10.1111/j.1365-313X.2006.02750.x PMID: 16824178.

11. Wang F, Qiu N, Ding Q, Li J, Zhang Y, Li H, et al. Genomewide identification and analysis of the growth-regulating factor family in Chinese cabbage (Brassica rapa L. ssp. pekinensis). BMC Genomics. 2014; 15(1):1–12. https://doi.org/10.1186/1471-2164-15-807 PMID: 25262125.

12. Lee BH, Jeon JO, Lee MM, Kim JH. Genetic interaction between GROWTH-REGULATING FACTOR and CUP-SHAPED COTYLEDON in organ separation. Plant Signal Behav. 2015; 10(2): e988071. https://doi.org/10.4161/15592324.2014.988071 PMID: 25761011.

13. Cao Y, Han Y, Jin Q, Lin Y, Cai Y. Comparative Genomic Analysis of the GRF Genes in Chinese Pear (Pyrus bretschneideri Rehd), Poplar (Populus), Grape (Vitis vinifera), Arabidopsis and Rice (Oryza sativa). Front Plant Sci. 2016; 7(577):1750. https://doi.org/10.3389/fpls.2016.02750 PMID: 27933074.

14. Ma JQ, Jian HJ, Yang B, Lu K, Zhang AX, Liu P, et al. Genome-wide identification and expression profiling of the GRF gene family in oilseed rape (Brassica napus L.). Gene. 2017; 620:36–45. https://doi.org/10.1016/j.gene.2017.03.030 PMID: 28363779.

15. Kim S, Solitis PS, Kerr W, Solitis DE. Phylogeny and domain evolution in the APETALA2-like gene family. Mol Biol Evol. 2006; 23(1):107–20. https://doi.org/10.1093/molbev/msj014 PMID: 16151182.

16. Shigyo M, Hasebe M, Ito M. Molecular evolution of the AP2 subfamily. Gene. 2006; 366(2):256–65. https://doi.org/10.1016/j.gene.2005.08.009 PMID: 16388920.

17. Horstman A, Viola W, Kim B, Renze H. AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. Trends Plant Sci. 2014; 19(3):146–57. https://doi.org/10.1016/j.tplants.2013.10.010 PMID: 24260109.

18. Manchado-Rojo M, Weiss J, Egger-Cortines M. Validation of Ainetegumenta as a gene to modify floral size in ornamental plants. Plant Biotechnol J. 2015; 12(8):1053–65. https://doi.org/10.1111/pbi.12212 PMID: 24985495.
19. Klucher KM, Chow H, Reiser L, Fischer RL. The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. Plant Cell. 1996; 8(2):137–53. https://doi.org/10.1105/tpc.8.2.137 PMID: 8742706.

20. Han H, Krizek BA. AINTEGUMENTA-LIKE6 can functionally replace AINTEGUMENTA but alters Arabidopsis flower development when misexpressed at high levels. Plant Mol Biol. 2016; 92(4–5):597–612. https://doi.org/10.1007/s11103-016-0535-y PMID: 27605095.

21. Du Y, Scheres B. PLETHORA transcription factors orchestrate de novo organ patterning during Arabidopsis lateral root outgrowth. Proc Natl Acad Sci USA. 2017; 114(44):11709–14. https://doi.org/10.1073/pnas.1714410114 PMID: 29078398.

22. Ding Q, Cui B, Li J, Li H, Zhang Y, Lv X, et al. Ectopic expression of a Brassica rapa AINTEGUMENTA gene (BrANT-1) increases organ size and staminal density in Arabidopsis. Sci Rep. 2018; 8(1):10528. https://doi.org/10.1038/s41598-018-28606-4 PMID: 30002453.

23. Aida M, Beis D, Heidstra R, Willemse V, Bilou I, Galinha C, et al. The PLETHORA Genes Mediate Pattern- ing of the Arabidopsis Root Stem Cell Niche. Cell. 2004; 119(1):109–20. https://doi.org/10.1016/j.cell.2004.09.018 PMID: 15454085.

24. Dash M, Malladi A. The AINTEGUMENTA genes, MdANT1 and MdANT2, are associated with the regulation of cell production during fruit growth in apple (Malus × domestica Borkh.). BMC Plant Biol. 2012; 12(1):98. https://doi.org/10.1186/1471-2229-12-98 PMID: 22731507.

25. Confalonieri M, Carelli M, Galmberi V, Macovei A, Panara F, Bigiogera M, et al. Seed-Specific Expression of AINTEGUMENTA in Medicago truncatula Led to the Production of Larger Seeds and Improved Seed Germination. Plant Mol Biol Rep. 2014; 32(5):957–70. https://x.scihub.ltd/https://doi.org/10.1007/s11105-014-0706-4

26. Tsuchamoto R, Yokoi S, Takahata Y. Arabidopsis EMBRYOMAKER encoding an AP2 domain transcription factor plays a key role in developmental change from vegetative to embryonic phase. Plant Mol Biol. 2010; 73(4–5):481–92. https://doi.org/10.1007/s11103-010-9634-3 PMID: 20405311.

27. Krizek BA. Antegeuntea and Aintegumenta-Like6 regulate auxin-mediated flower development in Arabidopsis. BMC Res Notes. 2011; 4:176. https://doi.org/10.1186/1756-0500-4-176 PMID: 21645048.

28. Krizek BA, Eaddy M. AINTEGUMENTA-LIKE6 regulates cellular differentiation in flowers. Plant Mol Biol. 2012; 78(3):199–209. https://doi.org/10.1007/s11103-011-9844-3 PMID: 22076630.

29. Morcillo F, Gallard A, Pillot M, Jouanic S, Aberlenc-Bertossi F, Collin M, et al. EgAP2-1, an AINTEGU MENTA-like (AIL) gene expressed in meristematic and proliferating tissues of embryos in oil palm. Planta. 2007; 226(6):1353–62. https://doi.org/10.1007/s00425-007-0574-3 PMID: 17628826.

30. Dona DH, Dunwell J, Gibbings J. Characterization of Aintegumenta-like Gene in Coconut (Cocos nucifera L) and its Expression During Embryogenesis. In Vitro Cell Dev Biol.-Animal. 2010; 46:S123–S124. https://doi.org/10.1007/s11626-010-9339-6

31. Bandupriya HDD, Gibbings JG, Dunwell JM. Isolation and characterization of an AINTEGUMENTA-like gene in different coconut (Cocos nucifera L.) varieties from Sri Lanka. Tree Genet Genomes. 2013; 9(3):813–27. https://x.scihub.ltd/https://doi.org/10.1007/s11295-013-0600-5

32. Galinha C, Hugo H, Marlin J, Viola W, Ikram B, Renze H, et al. PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. Nature. 2007; 449(7165):1053–7. https://doi.org/10.1038/nature06206 PMID: 17960244.

33. Yu J, Hu F, Dossa K, Wang Z, Ke T. Genome-wide analysis of UDP-glycosyltransferase super family in Brassica rapa and Brassica oleracea reveals its evolutionary history and functional characterization. BMC Genomics. 2017; 18(1):474. https://doi.org/10.1186/s12864-017-3844-x PMID: 28645261.

34. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25(17):3389–402. https://doi.org/10.1093/nar/25.17.3389 PMID: 9254694.

35. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21):2947–8. https://doi.org/10.1093/bioinformatics/btm404 PMID: 17846036.

36. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol. 2013; 30(12):2725–9. https://doi.org/10.1093/molbev/mst197 PMID: 24132122.

37. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. Nucleic Acids Res. 2019; 47(W1):W270–W5. https://doi.org/10.1093/nar/gkx357 PMID: 31114888.

38. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009; 37:W202–W8. https://doi.org/10.1093/nar/gkp335 PMID: 19458158.
39. Voorrips RE. MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered. 2002; 93(1):77–8. https://doi.org/10.1093/jhered/93.1.77 PMID: 12011185.

40. Wang YP, Tang HB, DeBary JD, Tan X, Li JP, Wang XY, et al. MCScanX: a toolkit for detection and evolutionary analysis of syntenic and collinearity. Nucleic Acids Res. 2012; 40(7): e49. https://doi.org/10.1093/nar/gkr1293 PMID: 22217600.

41. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: An information aesthetic for comparative genomics. Genome Res. 2009; 19(9):1639–45. http://www.genome.org/cgi/doi/10.1101/gr.092759.109. PMID: 19541911.

42. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics. 2010; (10):2885–98. https://doi.org/10.1016/S1672-0229(10)60008-3 PMID: 20451164.

43. Qu CM, Fu FY, Lu K, Zhang K, Wang R, Xu XF, et al. Differential accumulation of phenolic compounds and expression of related genes in black- and yellow-seeded *Brassica napus*. J Exp Bot. 2013; 64 (10):2885–98. https://doi.org/10.1093/jxb/ert148 PMID: 23698630.

44. Wu G, Zhang L, Wu YH, Cao YL, Lu CM. Comparison of Five Endogenous Reference Genes for Specific-PCR Detection and Quantification of *Brassica napus*. J Agr Food Chem. 2010; 58(5):2812–7. https://doi.org/10.1021/jf904255b PMID: 20143854.

45. Lu K, Li T, He J, Chang W, Zhang R, Liu M, et al. qPrimerDB: a thermodynamics-based gene-specific qPCR primer database for 147 organisms. Nucleic Acids Res. 2018; 46(D1):D1229–D36. https://doi.org/10.1093/nar/gkx725 PMID: 28977518.

46. Okamura JK, Caster B, Villarroel R, Van Montagu M, Jofuku KD. The AP2 domain of *APETALA 2* defines a large new family of DNA binding proteins in *Arabidopsis*. Proc Natl Acad Sci USA. 1997; 94 (13):7076–81. https://doi.org/10.1073/pnas.94.13.7076 PMID: 9192694.

47. Jofuku KD, Boer BG, Den, Montagu M, Van, Okamura JK. Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA 2*. Plant Cell. 1994; 6(9):1211–25. https://doi.org/10.1105/tpc.6.9.1211 PMID: 7919989.

48. Mudunkothge JS, Krizek BA. Three *Arabidopsis AIL/PLT* genes act in combination to regulate shoot apical meristem function. Plant J. 2012; 71(1):108–21. https://doi.org/10.1111/j.1365-313X.2012.04975.x PMID: 22380923.

49. Shigyo M, Ito M. Analysis of gymnosperm two-AP2-domain-containing genes. Dev Genes Evol. 2004; 214(3):105–14. https://doi.org/10.1007/s00427-004-0385-5 PMID: 14986134.

50. Chalhoub B, France D, Shengyi L, Parkin IAP, Haibo T, Xiyan W, et al. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science. 2014; 345(6199):950–3. https://doi.org/10.1126/science.1253435 PMID: 25146293.

51. Cheng F, Sun R, Hou X, Zheng H, Zhang F, Zhang Y, et al. Subgenome parallel selection is associated with morphotype diversification and convergent crop domestication in *Brassica rapa* and *Brassica oleracea*. Nat Genet. 2016; 48:1218. https://doi.org/10.1038/ng.3634 PMID: 27526322.

52. Arias T, Beilstein MA, Tang M, McKain MR, Pires JC. Diversification times among Brassica (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence. Am J Bot. 2014; 101(1):86–91. https://doi.org/10.3732/ajb.1300312 PMID: 24388963.

53. Pan Y, Zhu M, Wang S, Ma G, Huang X, Qiao C, et al. Genome-Wide Characterization and Analysis of Metallothionein Family Genes That Function in Metal Stress Tolerance in *Brassica napus*. L. Int J Mol Sci. 2018; 19(8):2181. https://doi.org/10.3390/ijms19082181 PMID: 30049941.

54. Zubay G, Doty P. The isolation and properties of deoxyribonucleoprotein particles containing single nucleic acid molecules. J Mol Biol. 1959; 1(1):1–20. https://doi.org/10.1016/S0022-2836(59)80002-4

55. Chen B, Wang TX, Wang HZ, Li YC, Yan XH, Wang LJ, et al. Cloning and Expression Level Analysis of Two *BnaANT* Candidate Genes in *Brassica napus*. Agricultural Sciences in China. 2010; 9(4):488–96. https://doi.org/10.1016/S1671-2927(09)60121-8