

# Genome-wide analysis of the auxin/indoleacetic acid (Aux/IAA) gene family in allotetraploid rapeseed (*Brassica napus* L.)

Li Haitao, Wang Jing

(College of Plant Science and Technology, Huazhong Agricultural University)

**Abstract:** Auxin/Indoleacetic acid (Aux/IAA) genes participate in the auxin signaling pathway and play key roles in plant growth and development. Although the Aux/IAA gene family has been identified in many plants, within allotetraploid *Brassica napus* little is known. In this study, a total of 119 non-redundant Aux/IAA genes were found in the genome of *B. napus*. They were distributed non-randomly across all 19 chromosomes and other non-anchored random scaffolds, with a symmetric distribution in the A and C subgenomes. Evolutionary and comparative analysis revealed that 111 (94.1%) *B. napus* Aux/IAA genes were multiplied due to ancestral *Brassica* genome triplication and recent allotetraploidy from *B. rapa* and *B. oleracea*. Phylogenetic analysis indicated seven subgroups containing 29 orthologous gene sets and two *Brassica*-specific gene sets. Structures of genes and proteins varied across different genes but were conserved among homologous genes. Furthermore, analysis of transcriptional profiles revealed that the expression patterns of Aux/IAA genes in *B. napus* were tissue dependent. Auxin-responsive elements tend to be distributed in the proximal region of promoters, and are significantly associated with early exogenous auxin up-regulation. This comprehensive analysis provides a deeper understanding of diversification of Aux/IAA gene family and will facilitate further dissection of Aux/IAA gene function in *B. napus*.

**Key words:** *Brassica napus*; Aux/IAA gene; Evolutionary analysis; Expression pattern; Auxin response

## 0 Introduction

Auxins were the first class of phytohormone discovered, and are prevalent signal chemicals produced within all vascular plants, and comprise a group of molecules with an indole ring. Auxins are able to regulate many aspects of plant growth and development, including cell division and elongation, and organ development at both cellular and whole plant level. They also play a role in regulating plant responses to environment such as phototropism, gravitropism, thigmotropism, and shade avoidance<sup>[1,2]</sup>. This physiological regulation is achieved by changes in expression of many responsive genes resulting from auxin perception and signal transduction, including the well described transport inhibitor response1/auxin signaling F-Box (TIR1/AFB) auxin signalling pathway<sup>[3]</sup>. The Aux/IAA protein is a central repressor in this pathway and can interact with both TIR1/AFB and auxin response factors (ARFs). Aux/IAA proteins interact strongly with ARF activators, and may also have weak or no interaction with ARF repressors<sup>[4]</sup>. ARF activators induce the transcription of auxin responsive genes by an amino-terminal DNA binding domain (DBD) that binds to auxin response elements (AuxREs) on promoters of target genes<sup>[5]</sup>. Auxin signaling is regulated by repression/de-repression mechanism with TIR1/AFB, Aux/IAA and ARF proteins in the TIR1/AFB pathway. At basal concentrations in plants, the heterodimer Aux/IAA-ARF represses expression of auxin-responsive genes<sup>[4,6,7]</sup>. Increased auxin can induce the ubiquitin–ligase complex SCF<sup>TIR1/AFB</sup>, which targets Aux/IAA proteins for ubiquitination and degradation, thus allowing ARF homodimers to activate the transcription of auxin-responsive genes<sup>[4,6,7]</sup>.

Foundations: The work was supported by the National Natural Science Foundation of China (31601332), National Basic Research and Development Program of China (2015CB150200) ; China Postdoctoral Science Foundation (2015M582240) and Research Fund for the Doctoral Program of Higher Education of China (20130146120034)

Brief author introduction:Li Haitao (1986), Male, Post-doctor,Crop genetics and breeding

Correspondance author: Wang Jing (1979), Female, Associate Professor,Crop genetics and breeding. E-mail:

wangjing@mail.hzau.edu.cn

Canonical Aux/IAA proteins contain four highly conserved domains designated I, II, III, and IV [8, 9]. Recent molecular and crystallographic studies indicate that these domains contribute to distinct roles of repression, degradation or interaction. Domain I contains a conserved LxLxLx motif that can recruit TOPLESS (TPL)/TPL-related (TPR) corepressors [10, 11] and is responsible for the repression property of the proteins [12]. Domain II contains a conserved degron GWPPV motif that interacts with the SCF<sup>TIR1/AFB</sup> complex [13, 14] and confers instability to the proteins [15]. Domains III and IV contain a carboxy-terminal PB1 (Phox and Bem1) domain that is also located in the carboxy-terminal of ARFs, and thus are responsible for homo- and hetero-dimerization between Aux/IAA and ARF proteins [4, 16].

It has been found that gain-of-function mutations within domain II which confer instability to the Aux/IAA proteins can result in auxin-related developmental defects, suggesting that *Aux/IAA* genes are important in the auxin signaling pathway [17]. Eleven *Aux/IAA* gain-of-function mutants have been obtained since the 1990s, by forward genetic screening, mostly by observing root elongation response under exogenous auxin treatment, in the model plant *A. thaliana*. Map-based cloning from these mutants has identified the sequence of the underlying genes, including *iaa1/axr5* [18], *iaa3/shy2* [19, 20], *iaa6/shy1* [21], *iaa7/axr2* [22, 23], *iaa12/bdl* [24, 25], *iaa14 /slr* [26], *iaa16* [27]; *iaa17/axr3* [28, 29], *iaa18* [30], *iaa19/msg2* [31], and *iaa28* [32]. Recently, several *Aux/IAA* gain-of-function mutants with auxin-related developmental defects have also been identified by map-based cloning in the model crop rice, including *Osiaa11* [33], *Osiaa13* [34] and *Osiaa23* [35]. Analysis of these mutants revealed that the *Aux/IAA* genes, both in monocots and dicots, have pleiotropic effects on diverse aspects of auxin-related plant growth and development, including embryogenesis, growth and development of many tissues including root, hypocotyl, leaf, stem and flower, tropism and apical dominance. The phenotypes of these mutants are similar but not identical, indicating that the *Aux/IAA* members overlap but do not replace each other in plants. For example, *iaa12/bdl* mutants uniquely have an observable defect in early embryogenesis and *iaa16* uniquely abolishes fertility, although both *iaa12/bdl* and *iaa16* mutants display reduced stature, reduced shoot apical dominance and curled leaves [24, 25, 27]. In addition, these analyses also revealed that the *Aux/IAA* genes mediate the interaction of auxin and signaling of other hormones such as abscisic acid [27], cytokinin and ethylene [32].

The *Aux/IAA* genes were first isolated from cDNAs induced by auxin in soybean [36], and orthologues subsequently identified in the *A. thaliana* genome using PCR-based and yeast two-hybridization approaches [8, 37]. In the past ten years, the release of reference genomes for different species has provides an opportunity for genome-wide identification of *Aux/IAA* gene families using bioinformatic approaches. *Aux/IAA* gene families have been isolated and analyzed from many diverse plant species, including not only *A. thaliana* (Brassicaceae, 29 members) [38], but also the forest tree *Populus* (35 members) [39]; cereal crops such as rice (31 members) [40], sorghum (26 members) [41], maize (31 members) [42] and wheat (84 members) [43]; legume crops such as *Medicago truncatula* (17 members) [44], chickpea (22 members) and soybean (63 members) [45]; and some vegetable/fruit crops such as tomato (26 members) and potato (27 members) [46], cucumber (27 members) [47] and the diploid *Brassica rapa* (55 members) [48].

The *B. rapa* A genome comprises 10 of the 19 chromosomes of *B. napus* (canola, oilseed rape, rapeseed), which is one of the most important oil crops in the world. The diploid *Brassica* genomes arose following whole genome triplication (WGT) from an ancestor in common with *A. thaliana* ~15.9 MYA. This was followed by divergence of *B. rapa* (AA, 2n=20) from *B. oleracea* (CC, 2n=18) ~4.6 MYA [49]. The allopolyploid *B. napus* (AACC, 2n=38) formed ~7500 years ago, most likely in domestication, by hybridization between *B. rapa* and *B. oleracea* [50]. To date, little

is known about the *Aux/IAA* gene family in *Brassica* species, and none of the *Aux/IAA* genes have been isolated and functionally analyzed from mutants of allotetraploid *B. napus*. In 2014, the draft genome of *B. napus* of cultivar Darmor-‘*bzh*’ was released, generated by both Sanger and next generation sequencing<sup>[50]</sup>, which provided a good opportunity for genome-wide identification and analysis of the *Aux/IAA* genes in *B. napus*.

Given the significant importance of *Aux/IAA* genes in plant development, the objectives of this study were to: (1) comprehensively identify and map the *Aux/IAA* genes in the *B. napus* genome by sequence similarity; (2) analyze the nature of duplication events and phylogenetic relationships of the *Aux/IAA* gene family using the homologous *Aux/IAA* proteins in *B. napus* compared with the related species *A. thaliana*, *B. rapa* and *B. oleracea*; (3) investigate and compare the gene structure and protein composition of *Aux/IAAs*; (4) profile the *Aux/IAA* gene expression pattern in four major tissues and determine differential expression under auxin treatment using RNA-seq data. These results will provide useful information for further studies in the crop *B. napus* on the function of *Aux/IAA* family in diverse growth and development processes.

## 1 Materials and Methods

### 1.1 Plant materials and auxin treatment

Seeds of *B. napus* were sterilized, rinsed three times in sterile water and sown on Murashige and Skoog (MS) culture medium with 1% agar. The seedlings were grown at 22°C with 16h light/8h dark. In addition, seeds of rapeseed were sown in the field to collect tissues of root, leaves, stem and flowering bud. For the IAA treatment, two sets of 30-day-old seedlings were incubated in liquid MS medium with and without 1 μM IAA for three hours, respectively. Each of ten seedlings for the both treatment were then pooled together. All materials were frozen in liquid nitrogen and subsequently stored at -80°C until RNA isolation.

### 1.2 Identification of *Aux/IAA* genes

All protein sequences of *B. napus* cultivar ‘Darmor-*bzh*’ were downloaded from the Genoscope Genome Database (<http://www.genoscope.cns.fr/brassicapapus/>). Initially, they were searched against the Pfam library of Hidden Markov Model (HMM) profiles (<http://pfam.xfam.org/>) using hmmer3 software locally (<http://hmmer.org/>). The genes corresponding to proteins with the *Aux/IAA* domain (PF02309) were extracted from the *B. napus* genome and genes with both ARF and *Aux/IAA* domains were removed. Meanwhile, all proteins of *B. napus* were searched again via BLASTP algorithms with E value lower than 1e-10 using 29 *A. thaliana* *Aux/IAA* protein sequences as queries<sup>[38]</sup>. Then all non-redundant protein sequences obtained were then checked for the presence of the *Aux/IAA* domain by Conserved Domain (CD) search service on the NCBI website (<https://www.ncbi.nlm.nih.gov/>). Combining all of the results from above analysis, we identified all members of *Aux/IAA* genes in the currently available *B. napus* genome. To investigate the duplication and evolution profile of *Aux/IAA* genes in the *B. napus* genome, all members of *Aux/IAA* genes in the *B. oleracea* genome, one of the two diploid progenitor genomes of *B. napus*, were obtained using the same workflow as indicated for *B. napus*. For these, the protein sequences of *B. oleracea* (TO1000) were downloaded from the Ensembl Plants database (<http://plants.ensembl.org/index.html>). The *Aux/IAA* genes in *B. rapa*, the other diploid progenitor genome of *B. napus*, were obtained from a previous study<sup>[48]</sup>.

### 1.3 Chromosome mapping and homologous identification of *Aux/IAA*

All *Aux/IAA* genes were mapped to specific *Brassica* chromosomes according to the location

135 information retrieved from relevant GFF files. These *Aux/IAA* genes were allocated a unique name  
according to the functional gene nomenclature for the *Brassica* genus<sup>[51]</sup>, based on their position  
on each chromosome. To investigate the whole genome duplication of *Aux/IAA* genes, most  
140 homologous sets of *Aux/IAA* genes among *B. rapa*, *B. oleracea*, *B. napus* and their putative  
orthologs in *A. thaliana* were extracted from a previous study of the *B. napus* genome<sup>[50]</sup>, and  
others were established by reciprocal BLASTP analysis. All *Aux/IAA* genes and their whole  
genome duplication distribution were displayed by Circos software<sup>[52]</sup>.

#### 1.4 Gene structure, motif scanning and phylogenetic analysis of *Aux/IAA* genes

To illustrate the exon/intron composition of *Aux/IAA* genes, the structures of *Aux/IAA* genes  
were displayed using Gene Structure Display Server (GSDS 2.0) software  
145 (<http://gsds.cbi.pku.edu.cn/index.php>) using the gene annotation described in the GFF3 format.  
The physical and chemical parameters of each protein were calculated using ProtParam  
(<http://www.expasy.ch/tools/protparam.html>). Motifs of the *Aux/IAA* proteins were investigated  
with the MEME tool (<http://meme-suite.org/index.html>). Default parameters were used, apart from  
the number of found motifs was set as four, with motif width ranging from 6 to 60. All conserved  
150 domains and signal peptide were investigated based on multiple sequence alignment of *Aux/IAA*  
proteins performed by Cluster Omega program (<http://www.ebi.ac.uk/Tools/msa/clustalo/>), and  
displayed by Jalview 2.0<sup>[53]</sup>. Phylogenetic relationships were established using MEGA 5.2<sup>[54]</sup> by  
the Neighbor-Joining (NJ) method based on p-distance model of amino acid substitutions type. A  
non-parametric bootstrap method was performed with bootstrap replication of 1000.

#### 1.5 Motif analysis in the promoter regions of *Aux/IAA* genes

To identify the *cis*-elements in the promoter region of each *Aux/IAA* gene, 2,000 bp of  
genomic sequence upstream of the translation start site was retrieved from the available *B. napus*  
genome sequence and analyzed by PlantPAN 2.0 (<http://plantpan2.itps.ncku.edu.tw>). However, for  
genes where the length of promoter was less than 2,000 bp, the inter-genic sequence up to the  
160 neighbouring upstream gene was extracted for analysis.

#### 1.6 Transcriptome analysis based on RNA-seq data

To reveal the expression profiling of *Aux/IAA* genes in *B. napus*, RNA-seq data from major  
tissues, including roots, stem, leaves and flower buds, were obtained from a previous study<sup>[50]</sup>. All  
reads were mapped to the *B. napus* Darmor-‘*bzh*’ genome v4.2 using hisat2 (v2.0.4). Fragments  
165 per kilobase per million mapped reads (FPKM) were calculated using cufflinks (v2.2.1) to  
estimate gene expression levels. Heat maps were constructed by R package pheatmap based on  
normalized expression values of *Aux/IAA* gene. For the auxin response of *Aux/IAA* genes, the  
RNA integrity of indicated samples were determined with Agilent 2100 Bioanalyzer (Agilent  
Technologies) and RNA concentration was measured by Qubit<sup>®</sup> 2.0 (Life Technologies). The  
170 libraries were constructed according to TruSeq<sup>®</sup> RNA Sample Preparation v2 Guide (Illumina)  
and sequenced on the HiSeq3000 platform. Read mapping and FPKM calculation were performed  
as described above. Differential expression genes (DEGs) were identified using the R package  
DEGseq based on the read count for each gene with thresholds that adjusted *p* value (*q* value) <  
0.05 and absolute value of  $\log_2(\text{fold change}) \geq 1$ .

#### 1.7 qRT-PCR analysis

Total RNA was extracted using the ultrapure total RNA isolation kit (BioTeke, Beijing). The  
quality was checked on 2% agarose gel and concentration was quantified in a spectrophotometer.

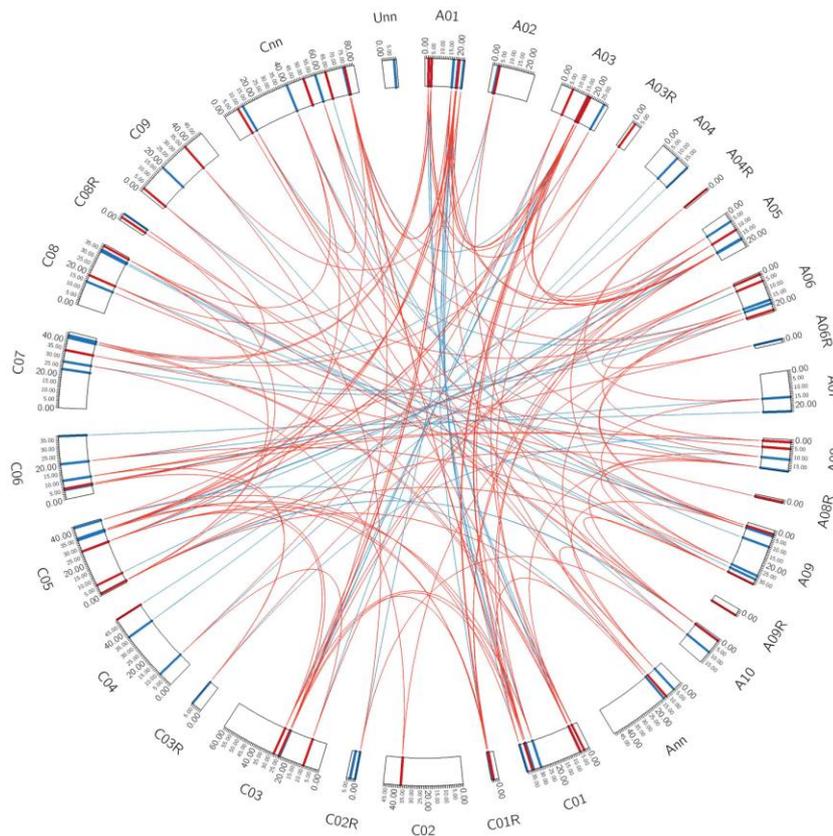
Genomic DNA was removed by digestion with DNase I and the first strand cDNA was synthesized from total RNA according to the manufacturer's instructions (Thermo Scientific). cDNA was diluted 10-fold as the templates of RT-PCR. The qRT-PCR reactions were performed on a CFX96 Touch Real-Time PCR detection system (Bio-Rad) using SYBR Green Supermix (Bio-Rad). Each sample was performed by three biological and two technical repeats. The PCR reaction and cycling protocol were carried out according to instructions. Melting curves were generated from 65 to 95 °C with 0.5 °C increments at 5 sec/step to estimate the specificity of product. The gene expression level of target genes was analyzed following deltaCt method with *B. napus* *ENTH* gene as reference for normalization<sup>[55]</sup>. The primers for qRT-PCR were listed in Supplemental Table S1.

## 2 Results

### 2.1 Genome-wide identification and chromosomal distribution of *Aux/IAA* genes in *B. napus*

In order to identify all the *Aux/IAA* gene members in *B. napus*, 101,040 available protein sequences were obtained from the *B. napus* database. 182 non-redundant genes were extracted as candidate *Aux/IAA* members by a Hidden Markov Model (HMM)-based search with *Aux/IAA* domain (PF02309) and BLASTP with 29 *A. thaliana* *Aux/IAA* genes as queries. Of these, 58 genes were removed that contained additional B3 and ARF domains characteristic of the ARF gene family. After checking for the presence of *Aux/IAA* domains using the Conserved Domain (CD) search platform, a total of 119 *Aux/IAA* members were identified in the *B. napus* genome (details in Supplementary Table S2).

The chromosomal location and direction of transcription for each *Aux/IAA* gene were established, with 94 on the 19 chromosomes and 25 located on random scaffolds of the 'Darmor-*bzh*' reference sequence<sup>[50]</sup> (Fig. 1). We analyzed the distributions of the 118 *Aux/IAA* genes across the genome of *B. napus*, excluding *BnaX.IAA.1* on an unassigned scaffold which lacked location information. The *Aux/IAA* genes have a non-random distribution across the 19 chromosomes, but have a symmetric distribution in the A and C subgenomes (57 and 61 genes, respectively) (Fig. 1; Supplementary Table S2). The number of *Aux/IAA* genes in the A and C subgenomes appears almost identical to that in the ancestor genomes *B. rapa* (Ar genome, 55)<sup>[48]</sup> and *B. oleracea* (Co genome, 60, Supplementary Table S3). The number of *Aux/IAA* gene varied dramatically among chromosomes, with a minimum of two on chromosome A07 and a maximum of ten on chromosomes C01 and C05 (Fig. 1; Supplementary Table S2). In addition, three and eight *Aux/IAA* genes were present on unassigned scaffolds in the A subgenome and the C subgenome, respectively. Five tandem *Aux/IAA* gene pairs located on chromosomes A03, A10, C03, C05 and C08 (Fig. 1). Similar tandem pairs have been observed in the *Populus*, chickpea and soybean genomes<sup>[39, 45]</sup>. Notably, for the tandemly duplicated genes *BnaC03.IAA.4* and *BnaC03.IAA.5* on chromosome C03, *BnaC03.IAA.5* may represent a recently truncated gene generated by gene conversion between the extensively homoeologous A3 and C3 chromosomes (Supplementary Fig. S2). This is supported by the phylogenetic distance of the non-homologous pair of *BnaA03.IAA.5/BnaC03.IAA.5* which was closer than that of homologous pair *BnaA03.IAA.5/BnaC03.IAA.4* that resulted from the recent allotetraploidy of *B. napus* (Supplementary Fig. S2).



**Fig.1** Distribution and duplicated nature of *Aux/IAA* genes in *B. napus* genome.

Outer boxes represent chromosomes in the *B. napus* genome. Colored lines in boxes indicate location of *Aux/IAA* genes in each chromosome, where blue and red lines show forward and reverse transcription direction, respectively. The homeologous gene pairs resulting from *Brassica* WGT and recent allotetraploidy are linked by red and blue lines, respectively.

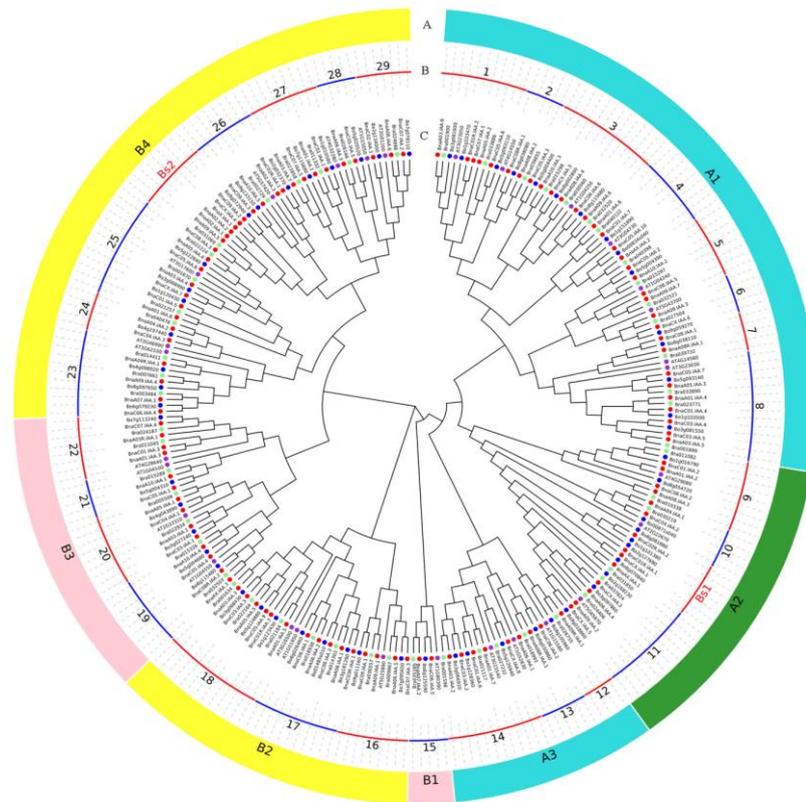
## 2.2 Comparative analysis of *Aux/IAA* genes in *Brassica* species and their duplication in *B. napus*

The evolution and duplication of *Aux/IAA* genes in *Brassica* was analyzed using gene models from genomes of *A. thaliana*, *B. napus* (An and Cn genome) and its diploid progenitors *B. rapa* (Ar genome) and *B. oleracea* (Co genome). Previous analyses identified 53 of the 55 *Aux/IAA* genes in the *B. rapa* genome to be orthologs corresponding to 29 *A. thaliana* genes [48], compared with 55 of the 60 *Aux/IAA* genes in *B. oleracea* (Supplementary Table S3). Most of the *Arabidopsis Aux/IAA* genes had one or two orthologs in the *B. rapa* and *B. oleracea* genomes. However, seven *A. thaliana Aux/IAA* genes each had three orthologs in both *B. rapa* and *B. oleracea* genomes, respectively, and another two *A. thaliana Aux/IAA* genes had three orthologs in only *B. oleracea* genomes (Supplementary Table S3), which indicated that most *Aux/IAA* genes experienced gene loss after WGT. Compared to *B. rapa* and *B. oleracea*, 50 (87.7%) and 53 (86.8%) orthologous *Aux/IAA* gene pairs were observed between the An and Cn subgenomes and their respective progenitor genomes. All orthologous pairs where chromosome location was known, including 47 An-Ar pairs and 43 Cn-Co pairs, were located on the same chromosomes of *Brassica* A and C genome (Supplementary Table S3). Most of the orthologous *Aux/IAA* gene pairs (43, 84.3%) between *B. rapa* and *B. oleracea* remain as homeologous pairs in *B. napus* (Fig. 1). These observations suggest that most of the *Aux/IAA* genes were retained intact during the recent formation of the allotetraploid *B. napus* from *B. rapa* and *B. oleracea*. In summary, homologous analysis of *Aux/IAA* genes revealed that as a result of both *Brassica* WGT and recent

allotetraploidy, 111 different *Aux/IAA* gene family members were represented by two to six copies each on different chromosomes of the *B. napus* genome, with the remaining eight *Aux/IAA* genes being unique within the *B. napus* genome (Fig. 1).

### 2.3 Phylogenetic relationship of *Aux/IAA* genes in *Arabidopsis* and *Brassica*

250 To examine the evolutionary relationships among the *Aux/IAA* genes from *B. napus*, *B. rapa*,  
*B. oleracea*, and *A. thaliana*, a rooted phylogenetic tree was generated based on the alignment of  
amino acid sequences for 261 *Aux/IAA* genes, including 29 *A. thaliana*, 54 *B. rapa*, 60 *B. oleracea*  
and 118 *B. napus* members. *BnaC08.IAA.4* could not be clustered and was removed from further  
analysis. The phylogenetic tree could be divided into Group A and Group B, which could be  
255 further divided into three (A1-A3) and four (B1-B4) subgroups respectively (Fig. 2). This pattern  
of two major groups for *Aux/IAA* gene family members in the phylogenetic tree was similar to that  
reported for other plants including the monocots wheat [43], maize [42] and rice [40], and the dicots  
tomato [46] as well as *B. rapa* [48], which suggests that the *Aux/IAA* genes have been widely  
conserved in different taxa.



260 **Fig.2** Phylogenetic relationship of *Aux/IAA* genes among *A. thaliana* and *B. rapa*, *B. oleracea*, *B. napus*  
The unrooted tree was generated using MEGA5.2 based on alignment of full-length amino acid sequences of 29 *A.*  
*thaliana*, 54 *B. rapa*, 60 *B. oleracea* and 118 *B. napus* *Aux/IAA* proteins. (A) Seven subgroups were displayed by  
colored arcs; (B) The 29 homologous sets and two *Brassica* specific sets are indicated with colored lines; (C) The  
265 *Aux/IAA* genes from *A. thaliana* and *B. rapa*, *B. oleracea*, *B. napus* were indicated by yellow, lightgreen, blue and  
red dots, respectively.

Group A and B consisted of 125 and 136 *Aux/IAA* genes, respectively. Twenty nine  
orthologous gene sets were identified with *A. thaliana* *Aux/IAA* genes as a reference and  
distributed in all seven subgroups (Fig. 2). Sixteen *Brassica*-specific *Aux/IAA* genes, which did not  
270 have any homologs in the *A. thaliana* genome, (Supplementary Table S3) were clustered into two  
sets in subgroup A2 and B4 and designated as A2Bs1 and B4Bs2, respectively (Fig. 2). This

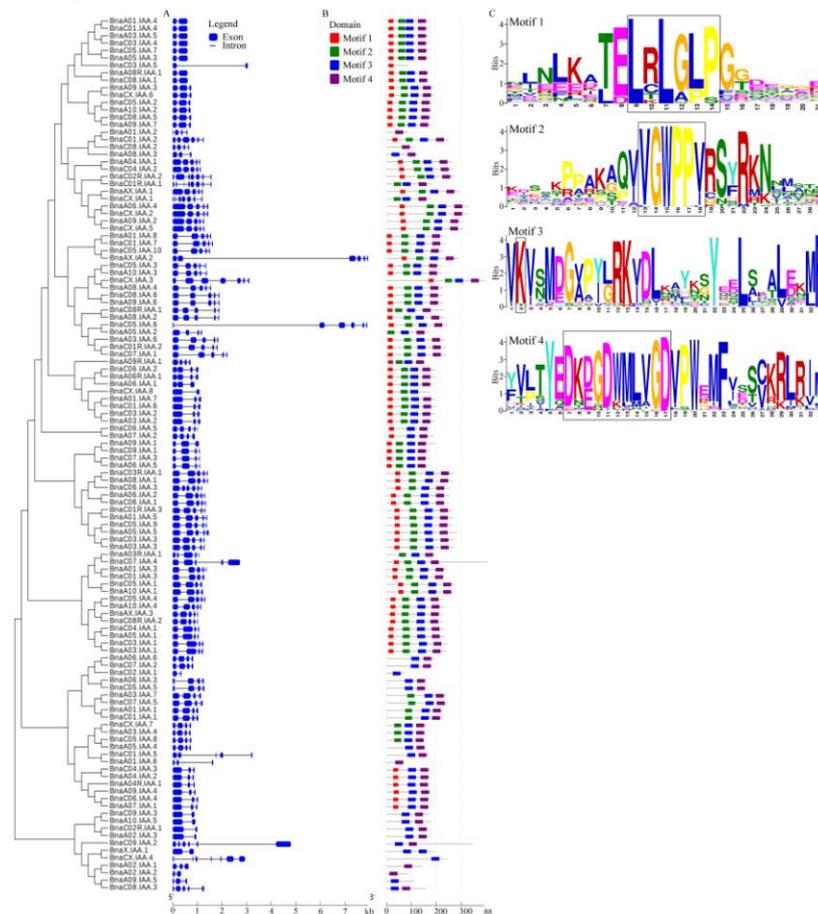
suggests that these two sets of *Brassica*-specific *Aux/IAA* genes may have similar function. All subgroups except subgroup B1 included multiple sets of orthologs. It was apparent that all 29 orthologous sets grouped together in neighboring branches. A sister pair indicates the closest relatives within a phylogenetic tree. Within this tree, a total of 97 sister pairs were found, consisting of 45 and 52 pairs in group A and B (Fig. 2). Most of the sister pairs were orthologous *Aux/IAA* gene pairs between the An and Cn subgenomes and their respective progenitor genomes, with 43 An-Ar pairs and 47 Cn-Co pairs. These two observations add further support to the results of the *Aux/IAA* gene duplication analysis. For two-thirds of the orthologous sets where two or three copies of the *Aux/IAA* genes were present in the A/C genome, the evolutionary relationship between *A. thaliana* genes with one of the *Brassica* orthologs derived from the genome triplication was closer than between the homologs themselves (Fig. 2). This phenomenon has also been observed for other genes<sup>[56]</sup> and is consistent with the two steps of WGT leading to the *Brassica* subgenomes.

## 2.4 Gene and protein structure of IAAs in *B. napus*

The open reading frame (ORF) length of the 118 *Aux/IAA* genes ranged from 340 to 7,942 bp, with an average of 1,336 bp. This corresponded to the coding domain sequence (CDS) length of the 118 *Aux/IAA* genes which ranged from 216 to 1,213 bp, with an average of 632 bp (Supplementary Table S2). This suggests that variation of intron length was more extensive than that of flanking exons as seen in the schematic diagram of the genes (Fig. 3A). The number of exons varied from two to eight, with most genes (108; 91.5%) having two to five exons. The distribution of exons and introns is complex, with a different structural pattern of exon/intron composition even within the same phylogenetic subgroup. However, the homologous genes have a similar pattern of gene structure (Fig. 3A; Supplementary Fig. S1). We therefore take the 17 homologous gene sets in *B. napus* that have an identical exon number as in *B. rapa*, *B. oleracea* and *A. thaliana* to correspond to ancestral genes, with the remainder having only one or two differences (Fig. 3A; Supplementary Fig. S1). Moreover, 71.7% (124/173) of the homologous *Aux/IAA* gene pairs in the *B. napus* genome had an almost similar structural pattern with respect to exon number and CDS length, which was a much higher proportion than for all paralogous gene pairs reported in the ancestral *B. rapa* and *B. oleracea* genome<sup>[49]</sup>. This observation indicates that the *Aux/IAA* gene families may be more conserved in *Brassica*, possibly due to their importance in facilitating plasticity of plant development.

The various physical and chemical properties of *Aux/IAA* proteins, including polypeptide length, molecular weight, instability index and Grand average of hydropathicity (GRAVY), were calculated (Supplementary Table S2). The polypeptide length ranged from 72 to 405 aa, with molecular weight ranging from 8.39 to 44.47 kD. The GRAVY index of all proteins was negative, indicating that all *Aux/IAA* proteins in *B. napus* are hydrophilic. Seventy seven (65.3%) *Aux/IAA* proteins had an instability index of more than 40, and may be unstable *in vitro*. Multiple amino acid alignment showed that there were four conserved regions designated as domain I, II, III and IV in *B. napus* *Aux/IAA* proteins (Supplementary Fig. S2). Domain I contained an LxLxLx motif, an EAR-like repression motif recruiting TOPLESS (TPL) co-repressors. Domain II contained VGWPPV(I) motifs, the conserved degron of *Aux/IAA* protein in auxin signaling. Domain III and IV contained putative canonical PB1 domains, including both an invariant lysine typical of type II PB1 domains and a type I PB1 series of acidic residues (D-x-D/E-x-D-x<sub>n</sub>-D/E)<sup>[16]</sup>, mediating interaction of ARF and *Aux/IAA*. Four conserved *Aux/IAA* protein domains were represented with four motifs generated from MEME analysis (Fig. 3B and C). Seventy five (63.6%) of the

proteins contained four canonical domains, compared to 20 (16.9%) that contained three domains (eleven for II, III and IV; eight for I, III and IV; one for I, II and III). Thirteen (11.0%) proteins contained two domains (III and IV), with the remaining ten (8.5%) only having a single domain (five for IV; four for III; one for I). It is notable that the proteins belonging to the same subgroup have a similar domain distribution. Almost all Aux/IAA proteins in subgroup A1, A2, A3, B2 and B3 contain the canonical four domains, whereas all Aux/IAA proteins in subgroup B1 and B4 contain non-canonical domains (Supplementary Fig. S3). However, some Aux/IAA proteins that belong to the same orthologous sets in subgroup A1 and A2 have truncated domains, such as BnaA05.IAA.2 in set 1, BnaC03.IAA.5 in set 8 and BnaA08.IAA.3, BnaC08.IAA.2 and BnaA01.IAA.2 in set 9. The truncated Aux/IAA proteins in subgroup A1 and A2 and B1 appear to have been formed after the divergence between *A. thaliana* and *Brassica*, since the orthologs in *A. thaliana* contained the canonical four domains (Supplementary Fig. S3). In addition, the majority of Aux/IAA proteins in *B. napus* had two nuclear localization signals (NLSs), one being a bipartite NLS and the other a SV40-like NLS (Supplementary Fig. S2). Interestingly, a subset of 20 Aux/IAA proteins appeared to contain a second repression domain (LxLxLx motif) between domain I and II, which also has repressive capacity in the auxin signaling pathway (Supplementary Fig. S2) [57].



335 **Fig.3 Gene and conserved motif structure of the *B. napus* Aux/IAA gene family.**

(A) Exon-intron organization of *B. napus* Aux/IAA genes. The blue solid boxes represent exons and black lines represent introns. (B) Conserved motif analysis of *B. napus* Aux/IAA proteins. The motifs representing four conserved domains are marked in red, green, blue and purple solid boxes, respectively. (C) The amino acid sequence of four motifs. The bits indicate amino acid conservation in each position. LxLxLx motifs in domain I, GWPPV/i motifs in domain II and conserved lysine and the OPCA-like motif phosphorylation sites in domain III/IV are highlighted with black boxes.

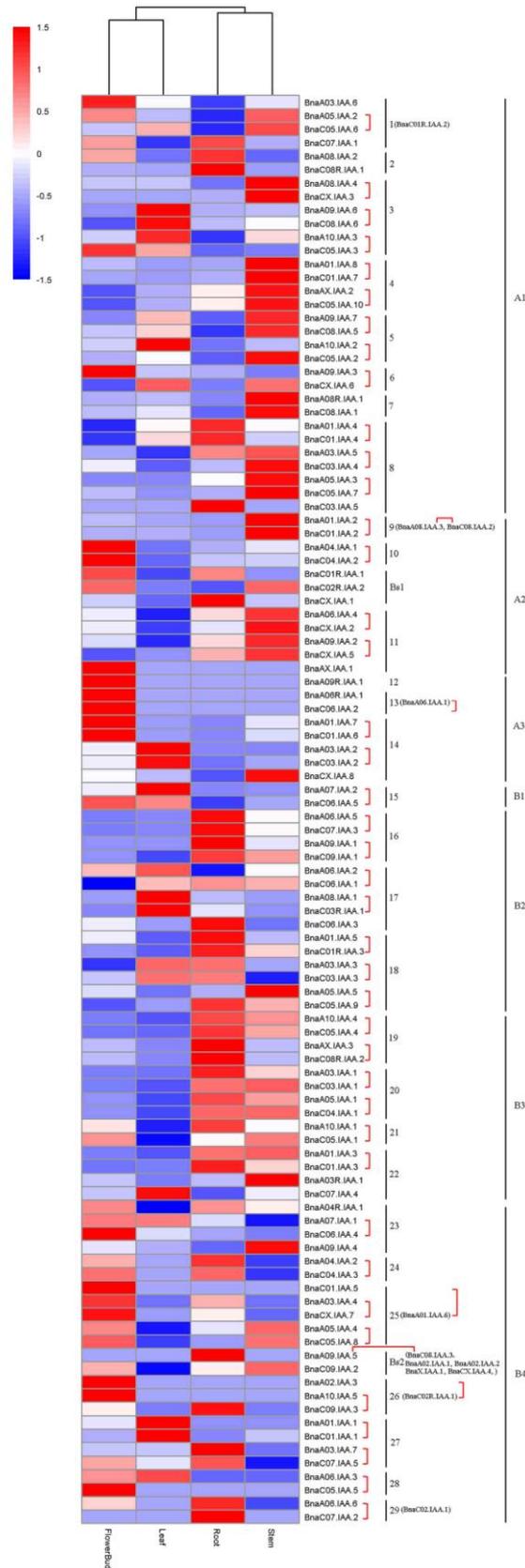
340

We assigned GO annotation to the 118 *Aux/IAA* genes to investigate the biological processes they potentially regulate, (Supplementary Fig. S4). Based on the cellular components, most of the genes were localized in cellular (organelle), consistent with the NLS identification based on amino acid alignment. Based on biological process, most of genes participated in response to stimulus, cellular process, pigmentation, biological regulation and metabolic process. In addition, some genes were also specifically involved in other biological process, such as reproduction process and immune system process. These results suggest that *Aux/IAA* genes have acquired multiple biological roles in *B. napus*. Based on molecular function, most of the *Aux/IAA* genes had binding and transcription regulator activity, which was consistent with their role as repressors in auxin signaling pathway.

## 2.5 Expression patterns of *Aux/IAA* genes in *B. napus*

To gain insights into the putative functions of *Aux/IAA* genes in development of *B. napus*, we analyzed their expression patterns in four major tissues including roots, stem, leaves and flower buds based on RNA-seq data. A subset of 12 genes lacked expression in any of these four tissues (Fig. 4; Supplementary Table S4), of which almost half were clustered in the B4Bs2 group. This suggests that most of the *Brassica*-specific genes might be non-functional or have inducible expression. The RNA-seq data indicated that the expression of *Aux/IAA* members was highly variable among tissues, supporting the diversification of functions for the *Aux/IAA* genes during *B. napus* development. Cluster analysis showed that a similar expression pattern was observed between flowering bud and leaf, root and stem, (Fig. 4). Twenty four *Aux/IAA* genes had tissue(s)-specific expression patterns (Fig. 4, Supplementary Table S4), although most of these had a very low expression level (FPKM <1), suggesting that these genes may be less important for growth and development in *B. napus*. We validated six *Aux/IAA* genes (*BnaA04.IAA.1*, *BnaC04.IAA.2*, *BnaA06.IAA.6*, *BnaA09R.IAA.1*, *BnaAX.IAA.1* and *BnaC01.IAA.5*) by quantitative RT-PCR, and found that two of these (*BnaA04.IAA.1* and *BnaC04.IAA.2*) express in all four tissues, which is consistent with the RNA-seq data (Supplementary Fig. S5). The expression level and specificity of the other four tissue(s)-specific genes was also consistent with the RNA-seq data (Supplementary Fig. S5).

The phylogenetic analysis had generated seven subgroups with 29 homologous sets. In general, *Aux/IAA* genes in the same homologous sets had distinct expression patterns across different tissues, suggesting acquisition of subfunctionalization. However, most of the homologous pairs (An and Cn) from the recent allotetraploidy shared a similar expression pattern across tissues, although some had a diverged pattern (Fig. 4). For example, within homologous set 3, the *BnaA08.IAA.4* and *BnaCX.IAA.3* homologous pair from the allotetraploid exhibited a similar expression pattern, with higher transcription in stem. Likewise, another homologous pair, *BnaA09.IAA.6* and *BnaC08.IAA.6*, were both highly expressed in leaf. In contrast, *BnaA10.IAA.3* and *BnaC05.IAA.3* were found to be differentially expressed across tissues, with a higher level in leaf and flower bud respectively, although they represented homologs within the allotetraploid (Fig. 4). Moreover, distinct expression patterns were observed among the subgroups. The majority of genes in subgroup A1, A2, A3, B1, B2 and B3 exhibited higher expression level across four tissues. In contrast, most of the genes in subgroup B4 exhibited lower or even no expression (Supplementary Table S4).



385

Fig.4 Heatmap of expression profiles of *B. napus* Aux/IAA genes.

The information subgroup and homologous sets is as shown in Fig. 2. The genes homologous between An and Cn genomes resulting from recent allotetraploidy are highlighted by red square brackets

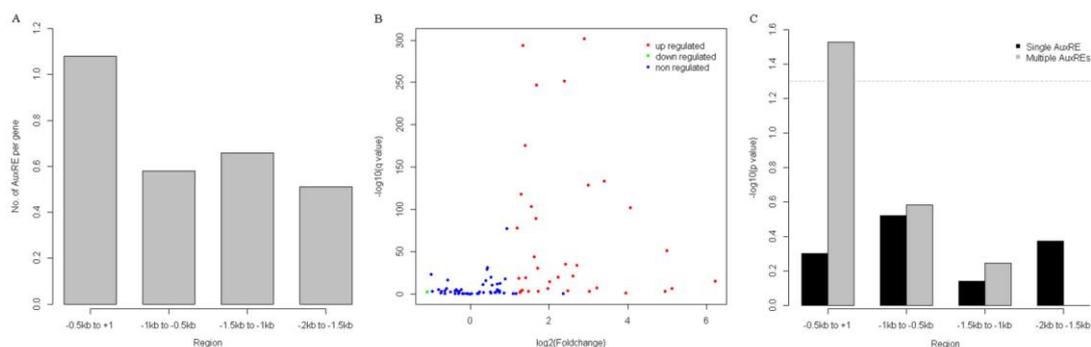
## 2.6 Expression analysis of *Aux/IAA* genes during seedling stage under auxin treatment

390

*Aux/IAA* genes are members of one of the three major early auxin-inducible gene families, and many auxin-responsive *cis*-elements (AuxREs) have been found in promoters of these genes [9]. To find potential AuxREs, we performed a motif search in the 2 kb promoter region of all *B. napus Aux/IAA* genes. AuxREs were found in the promoter of 101 genes (85.59%), with AuxRE counts ranging from one to nine. Of these, 21 (17.80%) *Aux/IAA* genes contained single and 80 (67.79%) multiple AuxREs in the promoters. The distribution of AuxREs along the promoters is non-uniform, with a significantly higher frequency in the proximal promoter (-0.5 kb to +1) of the *Aux/IAA* genes (Fig. 6A). At the same time, we analyzed the expression of *Aux/IAA* genes under auxin treatment. A total of 34 genes was significantly up-regulated by exogenous auxin treatment (Fig. 6B), all of which had AuxREs in the promoters. A detailed analysis was performed to analyze the association between the presence of AuxREs and auxin responsiveness of *Aux/IAA* genes. Significantly up-regulated genes were enriched in those with multiple AuxREs in the proximal promoter region (Fig. 6C).

395

400



405

**Fig.5 Analysis of AuxREs in promoters and auxin-induced expression of *B. napus Aux/IAA* genes.**

(A) Distribution of AuxREs in the 2.0 kb regions upstream of translation initiation site. The location of translation initiation site was designed as +1 bp and the numbers of AuxREs were calculated at intervals of 500 bp. (B) Scatterplot showing the significantly differentially expressed *Aux/IAA* genes. The fold change of *Aux/IAA* genes expression in auxin treated seedling relative to control was expressed as  $\log_2(\text{foldchange})$  indicated in x-axis. The adjusted p value calculated by DEGseq program was expressed as  $-\log_{10}(\text{q-value})$  indicated in y-axis. (C) Enrichment analysis of auxin up-regulation of *Aux/IAA* genes. The associations between presence of AuxREs and auxin up-regulation were analyzed by Pearson's chi-square test where a *p*-value < 0.05 indicates a significant association. The horizontal dotted gray line indicates a threshold of  $-\log_{10}(\text{p value}) = 1.30$ .

410

## 3 Discussion

415

As a large distinct plant gene family, *Aux/IAA* genes regulate auxin-induced gene expression and diverse aspects of plant development via the Aux/IAA-ARF and Aux/IAA-TIR complex [21]. Recent study in *A. thaliana* has shown that members of the *Aux/IAA* gene family have different properties. Firstly, *Aux/IAA* proteins differ in their capacity to interact with ARF activators *in vitro*, and most *Aux/IAA*-ARF interacting pairs have a diverse co-expression pattern that might lead to co-functions in particular processes or tissues [58]. Secondly, different combinations of TIR1 and *Aux/IAA* proteins display a wide range of auxin-binding affinities that are largely determined by the *Aux/IAA* identity [59]. Thirdly, the *Aux/IAA* protein family has diversified in degradation and auxin responsiveness, features that are dependent on sequences within and outside of Domain II [60]. Hence, it is necessary to identify the complete set of *Aux/IAA* gene family members within a species, as far as possible, to help develop a comprehensive understanding of *Aux/IAA* biological functions. Although in the past several experimental approaches have been

420

425

used to characterize gene families in the complex crop amphidiploid *B. napus*<sup>[61, 62]</sup>, genome-wide search based on the available draft genome sequence has provided a more convenient and effective approach for isolation of *B. napus* gene families such as the *SBP-box* and *LEA* families<sup>[63, 64]</sup>. In this study, we identified all 119 *Aux/IAA* genes present within the genome of *B. napus* using a similar genome-wide search (Fig. 1). This number in *B. napus* represents the largest *Aux/IAA* gene family identified in plants to date, and may have contributed to conferring higher phenotypic plasticity on this crop species.

Tandem duplication and segmental duplication both contribute to the dramatic variation in gene family number and distribution<sup>[65]</sup>. We mapped all 119 *Aux/IAA* genes on the chromosomes *in silico* and found five tandemly duplicated gene clusters (Fig. 1). Each of these clusters also represented ancient tandem genes in the *A. thaliana* genome. However, the number of *Aux/IAA* gene family in *B. napus* is much higher than that in *A. thaliana* (29 members), *B. rapa* (55 members) and *B. oleracea* (60 members). These observations suggest that the expansion of the *B. napus* *Aux/IAA* gene family may be independent of tandem duplication, and only affected by segmental duplication resulting from *Brassica* WGT and allopolyploidy. The segmental duplication has also led to a much higher number of *Aux/IAA* members in hexaploid bread wheat (84 members)<sup>[43]</sup> and the palaeopolyploid soybean (63 members)<sup>[45]</sup>. With 29 *Aux/IAA* members in the *A. thaliana* genome, one would expect ~90 and ~180 *Aux/IAA* genes to be present in the *B. rapa/B. oleracea* and *B. napus* genomes. However, only 55, 60 and 119 genes remain in these three genomes, respectively and the number of genes in the An and Cn subgenomes of *B. napus* is almost the same as that in the diploid progenitors *B. rapa* (Ar genome) and *B. oleracea* (Co genome) (Fig. 1 and Supplementary Table S3). These findings indicate that the loss of *B. napus* *Aux/IAA* members mainly occurred during the *Brassica* WGT process which resulted in widespread reshuffling of conserved genomic blocks<sup>[49, 66]</sup>, rather than the more recent allopolyploidy from *B. rapa* and *B. oleracea*.

The higher load of *Aux/IAA* genes in *B. napus* increases the probability of divergence within this family. It has been suggested that differentiation of the *Aux/IAA* genes in *A. thaliana* may depend both on molecular properties of proteins as well as expression patterns<sup>[67]</sup>. In the present study, 43 non-canonical *Aux/IAA* proteins were found to lack single or multiple domains that may contribute to their divergence (Fig. 3). 35 *Aux/IAA* proteins did not contain Domain I or Domain III/IV and might have lost capacity in recruiting TPL co-repressors or interaction with ARFs, implying that these proteins could not act as a repressor in auxin signaling, and may function differently in other processes. Notably, most (13/16) of the *Brassica*-specific *Aux/IAA* genes did not contain Domain I, suggesting that these genes may not contribute to classical auxin signal transduction. In addition, eight *Aux/IAA* proteins did not contain Domain II. A similar non-canonical *Aux/IAA* protein lacking domain II has been identified in all other reported plant species. This is consistent with evidence from the *A. thaliana* *IAA20* which lacks domain II and cannot be rapidly degraded in the presence of basal or increased levels of auxin<sup>[60]</sup>. Theoretically, such *Aux/IAA* proteins lacking domain II can repress auxin response gene expression due to the presence of conserved domain I and domain III/IV, and lead to developmental defects similar to *Aux/IAA* dominant mutants in wild-type plants. However, this is not the case in wild-type plants, and some alternative explanations should be sought. One explanation is that these *Aux/IAA* proteins could interact with additional unknown components and be degraded by a novel process in the auxin signal transduction cascade. Another explanation is that these *Aux/IAA* proteins have a very low expression level across tissues, and thus little effect on plant growth and development. Within our RNA-seq analysis we found that the expression level of all eight *Aux/IAA* genes

lacking Domain II was very low in all four major tissues compared with that of the canonical Aux/IAA protein (Supplementary Table S3). Sub-functionalization often depends on changes in the *cis*-regulatory elements of duplicated genes, which primarily leads to a divergence of gene expression<sup>[68]</sup>. Transcription profiles of these genes showed distinct patterns of expression between different tissues for the *B. napus* Aux/IAA genes possessing the canonical four domains (Fig. 4), which also supports the existence of their sub-functionalization. For example, *BnaC01.IAA.2* and *BnaC01.IAA.6* had much higher expression in stem and flower bud respectively, which suggests that these genes may play a key role in stem and reproductive organ development. This divergence of expression pattern is also observed among homologous Aux/IAA genes in *B. napus*. It is noted that a much higher expression divergence was found among homologs resulting from the earlier WGT, although most of the homoeologous pairs (An and Cn) from recent allotetraploidy shared similar expression patterns. This difference in functional divergence may be a consequence of the longer time that the older duplicated genes have had to accumulate more changes in promoter regions during their evolutionary history.

In general, orthologs have similar biological function in plant growth and development. In our study, phylogenetic analysis generated 29 orthologous gene sets containing Aux/IAA genes from *B. napus*, *B. rapa*, *B. oleracea* and *A. thaliana* (Fig. 2). The biological function of numerous Aux/IAA genes has been well studied in the model plant *A. thaliana*, which can also provide a valuable framework for functional prediction of Aux/IAA genes in *Brassica* species. For example, *AtIAA16* (*At3g04730*) was involved in root gravitropism and hair development, stem elongation and apical dominance and fertility<sup>[27]</sup>, suggesting that the all *Brassica* Aux/IAA genes in orthologous set 4 may have similar functions in each species.

In order to unravel whether Aux/IAA members in *B. napus* are auxin early-response genes, we analyzed the AuxRE distribution by bioinformatics tools as well as gene response under experimental auxin treatment. Most of Aux/IAA genes (85.59%) have AuxRE within a 2 kb region upstream of the translation initiation site, similar to observations in other plants<sup>[44, 46]</sup>. Within our experiment, 34 Aux/IAA genes were significantly up-regulated by auxin treatment (Fig. 6B). However, there were a further 28 Aux/IAA genes where there was some evidence of partial up-regulation by auxin treatment. This suggests that the expression of Aux/IAA genes could be induced by an exogenous auxin signal. In this study, we found that significantly up-regulated genes were enriched in the Aux/IAA gene family having multiple AuxREs in the proximal promoter region. This suggests that the number and location of AuxREs may partially account for the differential expression patterns of Aux/IAAs under IAA treatment, and that multiple AuxREs near the translation initiation site are more likely to mediate transcriptional activation of *B. napus* Aux/IAA genes in response to auxin.

## 4 Conclusion

In present study, a genome-wide analysis of Aux/IAA gene family was performed in *B. napus*, which included chromosomal distribution, duplication, phylogeny, gene and protein structure, expression pattern and response to exogenous auxin. A total of 119 Aux/IAA genes were identified in the *B. napus* genome by bioinformatics. These genes were non-randomly distributed across the 19 chromosomes and other unassigned scaffolds and symmetrically distributed in the A and C subgenomes. Comparative analysis between *A. thaliana*, *B. napus* and its diploid progenitors *B. rapa* and *B. oleracea* revealed the Aux/IAA genes undergone serious gene loss during *Brassica* WGT but almost unchanged during the recent allotetraploidy in *B. napus*, which resulted that most of Aux/IAA genes have multiple copies. Phylogenetic analysis generated two major group

and seven subgroups. Of which, we identified 29 orthologous gene sets using *A. thaliana* Aux/IAA genes as a reference and two *Brassica*-specific sets. The gene structure was different among all  
520 Aux/IAA gene members but similar between homologous genes. Four conserved regions designated as domain I, II, III and IV were identified in *B. napus* Aux/IAA proteins. 63.6% of the proteins contained four canonical domains and other ones lacked one or more domains. The variable expression of was observed between Aux/IAA genes among different tissues. And the expression pattern was distinct among all genes but similar between homologous pairs from the  
525 recent allotetraploidy. These complex motif distribution and expression profile suggest that there were sub-functionalization and redundancy for Aux/IAA gene family in *B. napus*. Furthermore, 34 genes, which enriched in ones with multiple AuxREs in proximal region of promoter, could be significantly up-regulated by exogenous auxin treatment in seedling. The results about Aux/IAA genes presented in this study will be useful for future functional dissection of them in *B. napus*.

### 530 References

- [1] Davies PJ: Plant Hormones: Kluwer, Dordrecht, The Netherlands; 1995.
- [2] Teale WD, Paponov IA, Klaus P: Auxin in action: signalling, transport and the control of plant growth and development[J]. Nat Rev Mol Cell Biol 2006, 7(11):847-859.
- [3] Enders TA, Strader LC: Auxin activity: past, present, and future[J]. Am J Bot 2015, 102(2):180-196.
- 535 [4] Chandler JW: Auxin response factors[J]. Plant Cell Environ 2016, 39:1014-1028.
- [5] Guilfoyle TJ, Hagen G: Auxin response factors[J]. Curr Opin Plant Biol 2007, 10(5):453-460.
- [6] Quint M, Gray WM: Auxin signaling[J]. Curr Opin Plant Biol 2006, 9(5):448-453.
- [7] Woodward AW, Bartel B: Auxin: regulation, action, and interaction[J]. Ann Bot 2005, 5(5):707-735.
- [8] Abel S, Nguyen MD, Theologis A: The PS-IAA4/5-like family of early auxin-inducible mRNAs in  
540 Arabidopsis thaliana[J]. J Mol Biol 1995, 251(4):533-549.
- [9] Hagen G, Guilfoyle T: Auxin-responsive gene expression: genes, promoters and regulatory factors[J]. Plant Mol Biol 2002, 49(3-4):373-385.
- [10] Causier B, Ashworth M, Guo W, Davies B: The TOPLESS interactome: a framework for gene repression in Arabidopsis[J]. Plant physiol 2012, 158(1):423-438.
- 545 [11] Szemenyei H, Hannon M, Long JA: TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis[J]. Science 2008, 319(319):1384-1386.
- [12] Tiwari SB, Hagen G, Guilfoyle TJ: Aux/IAA proteins contain a potent transcriptional repression domain[J]. Plant Cell 2004, 16(2):533-543.
- [13] Gray WM, Kepinski S, Rouse D, Leyser O, Estelle M: Auxin regulates SCFTIR1-dependent degradation of AUX/IAA proteins[J]. Nature 2001, 414(6861):271-276.
- 550 [14] Tan X, Calderonvillalobos LIA, Sharon M, Zheng C, Robinson CV, Estelle M, Zheng N: Mechanism of auxin perception by the TIR1 ubiquitin ligase[J]. Nature 2007, 446:640-645.
- [15] Ramos JA, Zenser N, Leyser O, Callis J: Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent[J]. Plant Cell 2001, 13(10):2349-2360.
- 555 [16] Korasick DA, Westfall CS, Soon Goo L, Nanao MH, Renaud D, Gretchen H, Guilfoyle TJ, Jez JM, Strader LC: Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression[J]. Proc Natl Acad Sci USA 2014, 111(14):5427-5432.
- [17] Kepinski S, Leyser O: Ubiquitination and auxin signaling: a degrading story[J]. Plant cell 2002, 14(1):S81-S95.
- 560 [18] Yang X, Lee S, So Jh, Dharmasiri S, Dharmasiri N, Ge L, Jensen C, Hangarter R, Hobbie L, Estelle M: The IAA1 protein is encoded by AXR5 and is a substrate of SCFTIR1[J]. Plant J 2004, 40(5):772-782.
- [19] Kim BC, Soh MS, Kang BJ, Furuya M, Nam HG: Two dominant photomorphogenic mutations of Arabidopsis thaliana identified as suppressor mutations of hy2[J]. Plant J 1996, 9(4):441-456.
- [20] Tian Q, Reed JW: Control of auxin-regulated root development by the Arabidopsis thaliana SHY2/IAA3 gene[J]. Development 1999, 126(4):711-721.
- 565 [21] Reed JW: Roles and activities of Aux/IAA proteins in Arabidopsis[J]. Trends Plant Sci 2001, 6(9):420-425.
- [22] Timpte C, Wilson A, Estelle M: The axr2-1 mutation of Arabidopsis thaliana is a gain-of-function mutation that disrupts an early step in auxin response[J]. Genetics 1994, 138(4):1239-1249.
- [23] Nagpal P, Walker LM, Young JC, Sonawala A, Timpte C, Estelle M, Reed JW: AXR2 encodes a member of the Aux/IAA protein family[J]. Plant physiol 2000, 123(2):563-574.
- 570 [24] Hamann T, Mayer U, Jurgens G: The auxin-insensitive bodenlos mutation affects primary root formation and apical-basal patterning in the Arabidopsis embryo[J]. Development 1999, 126(7):1387-1395.
- [25] Hamann T, Benkova E, Baurle I, Kientz M, Jurgens G: The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning[J]. Genes Dev 2002, 16(13):1610-1615.
- 575 [26] Fukaki H, Tameda S, Masuda H, Tasaka M: Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of Arabidopsis[J]. Plant J 2002, 29(2):153-168.
- [27] Rinaldi MA, Liu J, Enders TA, Bartel B, Strader LC: A gain-of-function mutation in IAA16 confers reduced

- responses to auxin and abscisic acid and impedes plant growth and fertility[J]. *Plant Mol Biol* 2012, 79(4-5):359-373.
- 580 [28] Leyser O, Pickett FB, Dharmasiri S, Estelle M: Mutations in the AXR3 gene of Arabidopsis result in altered auxin response including ectopic expression from the SAUR-AC1 promoter[J]. *Plant J* 1996, 10(3):403-413.
- [29] Rouse D, Mackay P, Stirnberg P, Estelle M, Leyser O: Changes in Auxin Response from Mutations in an AUX/IAA Gene[J]. *Science* 1998, 279(5355):1371-1373.
- 585 [30] Uehara TO, Yoko, Mimura T, Tasaka M, Fukaki H: Domain II mutations in CRANE/IAA18 suppress lateral root formation and affect shoot development in Arabidopsis thaliana[J]. *Plant Cell Physiol* 2008, 49(7):1025-1038.
- [31] Tatematsu K, Kumagai S, Muto H, Sato A, Watahiki MK, Harper RM, Liscum E, Yamamoto KT: MASSUGU2 encodes Aux/IAA19, an auxin-regulated protein that functions together with the transcriptional activator NPH4/ARF7 to regulate differential growth responses of hypocotyl and formation of lateral roots in Arabidopsis thaliana[J]. *Plant cell* 2004, 16(2):379-393.
- 590 [32] Rogg L, Lasswell J, Bartel B: A gain-of-function mutation in IAA28 suppresses lateral root development[J]. *Plant cell* 2001, 13(3):465-480.
- [33] Zhu Z, Liu Y, Liu S, Mao C, Wu Y, Wu P: A gain-of-function mutation in OsIAA11 affects lateral root development in rice[J]. *Mol Plant* 2012, 5(1):154-161.
- [34] Kitomi Y, Inahashi H, Takehisa H, Sato Y, Inukai Y: OsIAA13-mediated auxin signaling is involved in lateral root initiation in rice[J]. *Plant Sci* 2012, 190:116-122.
- 595 [35] Ni J, Wang G, Zhu Z, Zhang H, Wu Y, Wu P: OsIAA23 - mediated auxin signaling defines postembryonic maintenance of QC in rice[J]. *Plant J* 2011, 68(3):433-442.
- [36] Walker JC, Key JL: Isolation of cloned cDNAs to auxin-responsive poly(A)+RNAs of elongating soybean hypocotyl[J]. *Proc Natl Acad Sci USA* 1982, 79(23):7185-7189.
- 600 [37] Kim J, Harter K, Theologis A: Protein-protein interactions among the Aux/IAA proteins[J]. *Proc Natl Acad Sci USA* 1997, 94(22):11786-11791.
- [38] Overvoorde PJ, Okushima Y, Alonso JM, Chan A, Chang C, Ecker JR, Hughes BJ, Liu A, Onodera C, Quach H: Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana[J]. *Plant cell* 2005, 17(12):3282-3300.
- 605 [39] Kalluri UC, DiFazio SP, Brunner AM, Tuskan GA: Genome-wide analysis of Aux/IAA and ARF gene families in Populus trichocarpa[J]. *BMC Mol Biol* 2007, 7(1):59; 10.1186/1471-2229-1187-1159.
- [40] Jain M, Kaur N, Garg RK, Thakur JK, Tyagi AK, Khurana JP: Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (*Oryza sativa*)[J]. *Funct Integr Genomics* 2006, 6(1):47.
- [41] Wang S, Bai Y, Shen C, Wu Y, Zhang S, Jiang D, Guilfoyle T, Chen M, Qi Y: Auxin-related gene families in abiotic stress response in Sorghum bicolor[J]. *Funct Integr Genomics* 2010, 10(4):533.
- 610 [42] Wang Y, Deng D, Bian Y, Lv Y, Xie Q: Genome-wide analysis of primary auxin-responsive Aux/IAA gene family in maize (*Zea mays* L.)[J]. *Mol Biol Rep* 2010, 37(8):3991-4001.
- [43] Qiao L, Zhang X, Han X, Zhang L, Li X, Zhan H, Ma J, Luo P, Zhang W, Cui L: A genome-wide analysis of the auxin/indole-3-acetic acid gene family in hexaploid bread wheat (*Triticum aestivum* L.)[J]. *Front Plant Sci* 2015, 6:770; 710.3389/fpls.2015.00770.
- 615 [44] Shen C, Yue R, Yang Y, Zhang L, Sun T, Xu L, Tie S, Wang H: Genome-wide identification and expression profiling analysis of the Aux/IAA gene family in *Medicago truncatula* during the early phase of *Sinorhizobium meliloti* infection[J]. *PLoS One* 2014, 9(9):e107495; 107410.101371/journal.pone.0107495.
- [45] Singh VK, Jain M: Genome-wide survey and comprehensive expression profiling of Aux/IAA gene family in chickpea and soybean[J]. *Front Plant Sci* 2015, 6:918; 910.3389/fpls.2015.00918.
- 620 [46] Wu J, Peng Z, Liu S, He Y, Cheng L, Kong F, Wang J, Lu G: Genome-wide analysis of Aux/IAA gene family in Solanaceae species using tomato as a model[J]. *Mol Genet Genomics* 2012, 287(4):295-311.
- [47] Wu J, Liu S, Guan X, Chen L, He Y, Wang J, Lu G: Genome-wide identification and transcriptional profiling analysis of auxin response-related gene families in cucumber[J]. *BMC Res Notes* 2014, 7(1):218; 210.1186/1756-0500-1187-1218
- 625 [48] Paul P, Dhandapani V, Rameneni JJ, Li X, Sivanandhan G, Choi SR, Pang W, Im S, Lim YP: Genome-wide analysis and characterization of Aux/IAA family genes in Brassica rapa[J]. *PLoS One* 2016, 11(4):e0151522; 0151510.0151371/journal.pone.0151522.
- [49] Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IAP, Zhao M, Ma J, Yu J, Huang S et al: The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes[J]. *Nature Communications* 2014, 5:doi:10.1038/ng.1919.
- 630 [50] Chalhou B, Denoed F, Liu S, Parkin IAP, 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 [J]. *Science* 2014, 345:950-953.
- [51] Østergaard L, King GJ: Standardized gene nomenclature for the Brassica genus[J]. *Plant Methods* 2008, 4:10.
- 635 [52] Krzywinski M, Schein JE, Birol I, Connors JM, Gascoyne RD, Horsman DE, Jones SJ, Marra MA: Circos: An information aesthetic for comparative genomics[J]. *Genome Res* 2009, 19(9):1639-1645.
- [53] Waterhouse A, Procter JB, Martin DMA, Clamp M, Barton GJ: Jalview Version 2-a multiple sequence alignment editor and analysis workbench[J]. *Bioinformatics* 2009, 25(9):1189-1191.
- [54] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. *Mol Biol Evol* 2011, 28(10):2731-2739.
- 640 [55] Liu J, Hua W, Hu Z, Yang H, Zhang L, Li R, Deng L, Sun X, Wang X, Wang H: Natural variation in ARF18 gene simultaneously affects seed weight and silique length in polyploid rapeseed[J]. *Proc Natl Acad Sci USA* 2015, 112:E5123-E5132.

- 645 [56] Lu Y, Arnaud D, Belcram H, Falentin C, Rouault P, Piel N, Lucas M-O, Just J, Renard M, Delourme R: A Dominant Point Mutation in a RINGv E3 Ubiquitin Ligase Homoeologous Gene Leads to Cleistogamy in *Brassica napus*[J]. *Plant Cell* 2012, 24(12):4875-4891.
- [57] Li H, Hagen G, Guilfoyle TJ: Do some IAA proteins have two repression domains?[J]. *Plant Signal Behav* 2011, 6(6):858-860.
- 650 [58] Piya S, Shrestha SK, Binder B, Jr SC, Hewezi T: Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in *Arabidopsis*[J]. *Front Plant Sci* 2014, 5:744.
- [59] Villalobos LIAC, Lee S, Oliveira DC, Ivetac A, Brandt W, Armitage L, Sheard LB, Tan X, Parry G, Mao H et al: A combinatorial TIR1/AFB-Aux/IAA co-receptor system for differential sensing of auxin[J]. *Nat Chem Biol* 2012, 8(5):477-485.
- 655 [60] Dreher KA, Brown J, Saw RE, Callis J: The *Arabidopsis* Aux/IAA protein family has diversified in degradation and auxin responsiveness[J]. *Plant Cell* 2006, 18(3):699-714.
- [61] Wei Y, Li J, Lu J, Tang Z, Pu D, Chai Y: Molecular cloning of *Brassica napus* TRANSPARENT TESTA 2 gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis[J]. *Mol Biol Rep* 2006, 34(2):105-120.
- 660 [62] Hegedus DD, Yu M, Baldwin D, Gruber MY, Sharpe AG, Parkin IAP, Whitwill S, Lydiat DJ: Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress[J]. *Plant Mol Biol* 2003, 53(3):383-397.
- [63] Cheng H, Hao M, Wang W, Mei D, Tong C, Wang H, Liu J, Fu L, Hu Q: Genomic identification, characterization and differential expression analysis of SBP-box gene family in *Brassica napus*[J]. *BMC Plant Biol* 2016, 16(1):196.
- 665 [64] Liang Y, Xiong Z, Zheng J, Xu D, Zhu Z, Xiang J, Gan J, Raboanatahiry NH, Yin Y, Li M: Genome-wide identification, structural analysis and new insights into late embryogenesis abundant (LEA) gene family formation pattern in *Brassica napus*[J]. *Sci Rep* 2016, 6:24265.
- [65] Cannon SB, Mitra A, Baumgarten A, Young ND, May G: The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*[J]. *BMC Plant Biol* 2004, 4(1):10.
- 670 [66] Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun J-H, Bancroft I, Cheng F et al: The genome of the mesopolyploid crop species *Brassica rapa*[J]. *Nature genetics* 2011, 43(10):1035-1039.
- [67] Muto H, Watahiki MK, Nakamoto D, Kinjo M, Yamamoto KT: Specificity and similarity of functions of the Aux/IAA genes in auxin signaling of *Arabidopsis* revealed by promoter-exchange experiments among
- 675 MSG2/IAA19, AXR2/IAA7, and SLR/IAA14[J]. *Plant Physiol* 2007, 144(1):187-196.
- [68] Zhang J: Evolution by gene duplication: an update[J]. *Trends Ecol Evol* 2003, 18(6):292-298.

## 甘蓝型油菜 Aux/IAA 基因家族的全基因组分析

李海涛, 王晶

(华中农业大学, 植物科学技术学院, 武汉 430070)

- 摘要: 本研究利用生物信息学的方法在甘蓝型油菜基因组中鉴定到 119 个 Aux/IAA 基因。这些基因非随机地分布于全基因组的 19 个染色体和其他未定位的 scaffolds, 并且在 A 和 C 亚基因组之间匀称分布。进化与比较分析发现, 甘蓝型油菜在历史上经历的芸薹属基因组三倍化事件以及异缘四倍化事件导致了 111 (94.1%) Aux/IAA 基因以多拷贝的形式存在。系统进化分析将这些基因分成了 7 个组, 包含 29 个同源基因群和 2 个芸薹属特异的基因群。基因结构和蛋白质结构域分布在不同的 Aux/IAA 基因间存在很大差异。转录组测序分析表明 Aux/IAA 基因在不同的组织间表达模式存在很大差异。生长素响应元件倾向分布于启动子区域的近端, 且与 Aux/IAA 基因对外源生长素的响应有很大关系。
- 685
- 690
- 关键词: 甘蓝型油菜; Aux/IAA 基因; 进化分析; 表达模式, 生长素反应
- 中图分类号: S