



# Genome-Wide Identification of Dicer-Like, Argonaute, and RNA-Dependent RNA Polymerase Gene Families in *Brassica* Species and Functional Analyses of Their Arabidopsis Homologs in Resistance to *Sclerotinia sclerotiorum*

### **OPEN ACCESS**

### Edited by:

Dirk Balmer, Syngenta Crop Protection, Switzerland

### Reviewed by:

Biswapriya Biswavas Misra, Texas Biomedical Research Institute, USA Airong Wang, Fujian Agriculture and Forestry University, China

> \***Correspondence:** Xin-Zhong Cai xzhcai@zju.edu.cn

### Specialty section:

This article was submitted to Plant Biotic Interactions, a section of the journal Frontiers in Plant Science

Received: 19 April 2016 Accepted: 12 October 2016 Published: 27 October 2016

### Citation:

Cao J-Y, Xu Y-P, Li W, Li S-S, Rahman H and Cai X-Z (2016) Genome-Wide Identification of Dicer-Like, Argonaute, and RNA-Dependent RNA Polymerase Gene Families in Brassica Species and Functional Analyses of Their Arabidopsis Homologs in Resistance to Sclerotinia sclerotiorum. Front. Plant Sci. 7:1614. doi: 10.3389/fpls.2016.01614 Jia-Yi Cao<sup>1</sup>, You-Ping Xu<sup>2</sup>, Wen Li<sup>1</sup>, Shuang-Sheng Li<sup>1</sup>, Hafizur Rahman<sup>1</sup> and Xin-Zhong Cai<sup>1\*</sup>

<sup>1</sup> Institute of Biotechnology, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China, <sup>2</sup> Center of Analysis and Measurement, Zhejiang University, Hangzhou, China

RNA silencing is an important mechanism to regulate gene expression and antiviral defense in plants. Nevertheless, RNA silencing machinery in the important oil crop Brassica napus and function in resistance to the devastating fungal pathogen Sclerotinia sclerotiorum are not well-understood. In this study, gene families of RNA silencing machinery in B. napus were identified and their role in resistance to S. sclerotiorum was revealed. Genome of the allopolyploid species B. napus possessed 8 Dicer-like (DCL), 27 Argonaute (AGO), and 16 RNA-dependent RNA polymerase (RDR) genes, which included almost all copies from its progenitor species B. rapa and B. oleracea and three extra copies of RDR5 genes, indicating that the RDR5 group in B. napus appears to have undergone further expansion through duplication during evolution. Moreover, compared with Arabidopsis, some AGO and RDR genes such as AGO1, AGO4, AGO9, and RDR5 had significantly expanded in these Brassica species. Twenty-one out of 51 DCL, AGO, and RDR genes were predicted to contain calmodulin-binding transcription activators (CAMTA)-binding site (CGCG box). S. sclerotiorum inoculation strongly induced the expression of BnCAMTA3 genes while significantly suppressed that of some CGCG-containing RNA silencing component genes, suggesting that RNA silencing machinery might be targeted by CAMTA3. Furthermore, Arabidopsis mutant analyses demonstrated that dcl4-2, ago9-1, rdr1-1, rdr6-11, and rdr6-15 mutants were more susceptible to S. sclerotiorum, while dcl1-9 was more resistant. Our results reveal the importance of RNA silencing in plant resistance to S. sclerotiorum and imply a new mechanism of CAMTA function as well as RNA silencing regulation.

Keywords: *Brassica napus*, RNA silencing, Dicer-like (DCL), Argonaute (AGO), RNA-dependent RNA polymerase (RDR), *Sclerotinia sclerotiorum*, CAMTA3

1

## INTRODUCTION

RNA silencing refers to a variety of mechanisms whereby a small RNA molecule interferes with a given nucleotide sequence. It was first discovered in plants and occurs widely in eukaryotic organisms (Tijsterman et al., 2002; Ullu et al., 2004). In plants, RNA silencing is triggered by double-stranded RNA (dsRNA) that gives rise to small RNAs known as microRNAs (miRNAs) or small-interfering RNAs (siRNAs). Generation and function of these small RNAs depend on three key protein families, Dicerlike proteins (DCLs), Argonautes (AGOs), and RNA-dependent RNA polymerases (RDRs; Baulcombe, 2004). A whole RNA silencing process comprises three stages: initiation, maintenance, and signal amplification. DCLs undergo RNase III-type activities to process complementary double-strand RNAs into small RNAs with 21-26 nucleotides in length (Carmell and Hannon, 2004). These small RNAs are then incorporated into AGO-containing RNA-induced silencing complexes (RISCs) to serve as the sequence specificity in RNA degradation, translational inhibition, or heterochromatin formation (Bologna and Voinnet, 2014). At the signal amplification stage, RDR enzymes are responsible for synthesis of dsRNAs from ssRNA templates to initiate a new round of RNA silencing (Sijen et al., 2001).

DCL, AGO, and RDR are key components of RNA silencing machinery. DCL proteins are key components in small RNA biogenesis. They are characterized by the presence of six domains: DEAD, helicase-C, DUF283, PAZ, RNase III, and dsRNA-binding motif (DSRM; Margis et al., 2006). DCL consists of a small gene family in higher plants, insects, protozoa, and some fungi, whereas only one Dicer protein exists in vertebrates, nematodes, Schizosaccharomyces pombe, and green alga Chlamydomonas reinhardtii (Liu et al., 2009). AGO proteins are core factors of the RISC that guide small RNAs to their targets by sequence complementarity, which results in target mRNA cleavage, translational repression, or chromatin modification (Hannon, 2002; Moazed, 2009). AGOs are large proteins (~90-100 kDa) consisting of several characteristic functional domains including DUF1785, PAZ, MID, and PIWI (Hutvagner and Simard, 2008). RDRs enhance the potency of RNAi by amplifying the aberrant RNA population. It is defined by the presence of a conserved RNA-dependent RNA polymerase catalytic domain and is required for initiation and amplification of the silencing signal (Schiebel et al., 1998). Multiple copies of AGO and RDR genes are known to exist in both plants and animals. Members of these gene families play different roles in RNA silencing. For example, the Arabidopsis thaliana genome contains four DCL proteins (DCL1-DCL4) that specifically produce different types and sizes of small RNAs (Bologna and Voinnet, 2014). The role of DCL1 is to extract a single small RNA duplex out of a RNA loop called the pri-miRNA. This gives rise to a miRNA, generally 21 nucleotides long, and is typically involved in regulating developmental genes (Parent et al., 2015). DCL2 can produce abundant 22-nt viral siRNAs and shares functional overlap with DCL4 in antivirus defense (Moissiard et al., 2007). DCL3 generates 24-nt repeat-associated siRNAs (ra-siRNAs) and is involved in antiviral defense against DNA viruses (Moissiard and Voinnet, 2006). DCL4 generates 21-nt trans-acting siRNAs

(ta-siRNA) and is the primary DCL component of antiviral defense against RNA viruses (Deleris et al., 2006). Likewise, AGO1, the most well-studied plant AGO gene, associates with miRNAs and some siRNAs such as ta-siRNAs to cleave target mRNA and/or inhibit translation (Yu and Wang, 2010). AGO2 protein is involved in antiviral defense by catalyzing viral RNA cleavage in Arabidopsis (Jaubert et al., 2011). AGO10, the closest paralog of AGO1, is functionally redundant with AGO1 in some aspects of development (Lynn et al., 1999) and also functions as a decoy for miR165/166 to prevent the formation of AGO1miR165/166 complexes and the subsequent repression of HDZIP III gene expression (Zhu et al., 2011). For RDRs, RDR2 converts ssRNAs generated from repetitive DNAs to precursor dsRNAs of ra-siRNAs (Xie et al., 2004), while RDR6 produces the ta-siRNA precursors (Yoshikawa et al., 2005). Gene families encoding these three key components of RNA silencing machinery have been identified only in several plant species such as A. thaliana, Oryza sativa (Kapoor et al., 2008), Zea mays (Qian et al., 2011), Solanum lycopersicum (Bai et al., 2012), Nicotiana benthamiana (Nakasugi et al., 2013), Setaria italica (Yadav et al., 2015), and Vitis vinifera (Zhao et al., 2015). Identification of these families in more plant species will enhance our understanding of RNA silencing.

RNA silencing plays multiple roles in regulating growth and development as well as abiotic and biotic stress responses. In higher plants, RNA silencing functions as an antiviral defense through the action of DCL, AGO, and RDR proteins (Ding and Voinnet, 2007). The importance of RNA silencing in plant viral defense is manifested by the fact that it has elicited counter defense measures from the viral pathogens to overcome it. Plant viruses have evolved various viral RNA silencing repressors (VSR) to counteract this defense mechanism by targeting different RNA silencing pathway components (Csorba et al., 2015). Apart from viral defense, evidence accumulates for RNA silencing to play a role in plant interactions with bacterial pathogens (Voinnet, 2008). The first example is a miRNA from Arabidopsis that contributes to basal defense against Pseudomonas syringae by regulating auxin signaling (Navarro et al., 2006). Similar to viruses, the bacteria has also developed mechanisms to suppress RNA silencing in order to cause disease (Navarro et al., 2008). Recently, through the use of mutants for key components of RNA silencing or functional analyses of miRNAs in plant defense, the potential role of RNA silencing in plant defense against fungal pathogens has been revealed. These fungi include Verticillium dahliae (Ellendorff et al., 2009), Verticillium longisporum (Shen et al., 2014), Magnaporthe oryzae (Li et al., 2014), and Botrytis cinerea (Jin and Wu, 2015).

*Brassica napus* is an allotetraploid and was formed about 7500 years ago by crossing between *B. oleracea* and *B. rapa*, followed by chromosome doubling (Chalhoub et al., 2014). It is one of the most important oil crops, yet few RNAi machinery components have been characterized to date. We have identified the miRNAs involved in the interactions between *B. napus* and *Sclerotinia sclerotiorum*, one of the most devastating fungal pathogens in oil and vegetable crops. Furthermore, we find that Arabidopsis *ago1* and *ago2* mutant plants exhibit enhanced susceptibility to *S. sclerotiorum* (Cao et al., 2016). Our results provide a clue to the important roles of RNA silencing in the interactions between

*B. napus* and *S. sclerotiorum*. In this study, taking advantage of the completion of the *B. napus* genome sequencing (Chalhoub et al., 2014), we performed comprehensive bioinformatics analyses to identify DCL, AGO, and RDR gene families that are the three key components of RNA silencing machinery in *B. napus*. Furthermore, we employed mutants to probe their functions in resistance to *S. sclerotiorum*. We revealed the significant difference in RNA silencing machinery composition between *B. napus* and Arabidopsis, demonstrated the important role of RNA silencing in resistance to *S. sclerotiorum* and indicated a possible regulating mechanism of RNA silencing.

## MATERIALS AND METHODS

# Identification of Putative *B. napus DCL*, *AGO*, and *RDR* Genes

Protein sequences of Arabidopsis DCLs, AGOs, and RDRs were downloaded from TAIR database (http://www.arabidopsis.org/) and scan for conserved domains were performed using National Center for Biotechnology Information Conserved Domain Database (NCBI-CDD; http://www.ncbi.nlm.nih.gov/Structure/ cdd/wrpsb.cgi). All these protein sequences were used as queries to search their orthologs in B. napus, B. rapa, and B. oleracea genomes using BLASTp program in NCBI database with default settings. All retrieved protein sequences were examined for the presence of conserved domains and redundant sequences were removed. All candidate sequences of B. napus were subsequently verified in the GENOSCOPE database (http://www.genoscope. cns.fr/blat-server/cgi-bin/colza/webBlat). The physico-chemical properties of BnDCL, BnAGO, and BnRDR proteins were then predicted using ExPASy Compute pI/Mw tool (http://au.expasy. org/tools/pi\_tool.html; Bjellqvist et al., 1993).

## **Phylogenetic Analysis and Nomenclature**

Multiple alignment of DCL, AGO, and RDR protein sequences from *A. thaliana, B. napus, B. rapa*, and *B. oleracea* was performed using MUSCLE program (Edgar, 2004). The phylogenetic trees were then constructed using MEGA 5.0 (Tamura et al., 2011) by Neighbor-Joining (NJ) method following the Jones-Taylor-Thornton (JTT) model. Bootstrap analysis was performed with 1000 replicates to assess statistical support for nodes. The candidate genes were renamed according to the phylogenetic relationship and sequence homologies with corresponding *A. thaliana* homologs. The detail information of the proteins used for phylogenetic tree construction was listed in **Table 1** and Table S1.

# Exon-Intron Structure Analysis and Promoter *Cis*-Acting Element Prediction

The exon-intron organization of *BnDCLs*, *BnAGOs*, and *BnRDRs* genes were determined using the online GSDS1.0 program (http://gsds.cbi.pku.edu.cn/) by comparing their full-length coding sequences (CDS) with their corresponding genomic sequences downloaded from the GENOSCOPE database. The upstream sequences (1.5 kb) of *BnDCL*, *BnAGO*, and *BnRDR* genes were searched for the presence of potential *cis*-acting

elements using PLACE database (http://www.dna.affrc.go.jp/ PLACE/signalup.html; Higo et al., 1999).

## Plant Materials and Inoculation Analyses

The Arabidopsis dcl, ago, and rdr mutants were provided by Prof. Shou-Wei Ding (Department of Plant Pathology and Microbiology, University of California, Riverside, USA) and Prof. Yi-Jun Qi (Tsinghua-Peking Center for Life Sciences, and School of Life Sciences, Tsinghua University, China). B. napus plants were grown in growth cabinets at 25°C under a 16/8 h light/dark photoperiod, while Arabidopsis plants of the wildtype and mutants of RNA-silencing related genes were grown at 23°C with a 12/12 h day/night photoperiod. Fresh sclerotia of S. sclerotiorum were cultured at 23°C on potato dextrose agar medium (PDA) to produce mycelia, which were transferred to new PDA plates and grown for 2 days. The PDA plugs containing the advancing edge of S. sclerotiorum mycelia were removed to inoculate the plant leaves. For gene expression analyses, leaves were collected at 0, 8, and 16 h post inoculation (hpi) and frozen immediately in liquid nitrogen. Diameter of disease lesions was measured at 24 hpi and statistically analyzed using SPSS (verson19.0) by Student's *t*-test (p < 0.05). For disease resistance evaluation, at least 10 plants for each genotype were examined and the experiments were conducted three times independently.

## **Real-Time Quantitative PCR**

Total RNA was extracted using Trizol reagent (Invitrogen, CA, USA) following the manufacturer's procedure. Real-time quantitative PCR (RT-qPCR) was carried out using SYBR Premix Ex Taq (TakaRa, China) on StepOne Real-Time PCR System (ABI, USA).The RT-qPCR analyses were conducted three times, with three replicates for each gene and the relative fold changes were calculated using the  $2^{-\Delta\Delta Ct}$  method as described (Zhao et al., 2013). A *B. napus* elongation factor gene was used as the reference gene and primers used for RT-qPCR are listed in Table S2. Significance of the differences between mean values was determined with Student's *t*-test (*p* < 0.05).

# RESULTS

# Genome-Wide Identification of *DCL*, *AGO*, and *RDR* Genes in *B. napus*

A BLASTp search was conducted against *B. napus* genome in NCBI database using well-characterized Arabidopsis AGO, DCL, and RDR protein sequences as query sequences. The retrieved sequences were further analyzed for domain composition. Finally, 8 *DCL*, 27 *AGO*, and 16 *RDR* genes were identified in *B. napus* genome (**Table 1**). Compared with *A. thaliana* which contains 4 *DCL*, 10 *AGO*, and 6 *RDR* genes (Table S1); the members for each gene family in *B. napus* expanded by two times or more (**Table 1**). To compare the composition of these RNA silencing machinery genes between *B. napus* and its progenitor species *B. rapa* and *B. oleracea*, similar BLASTp and domain identification analyses were performed for these two genomes. The results showed that *B. rapa* genome contained 4 *DCL*, 13 *AGO*, and 6 *RDR* genes (Table S1). Comparison analysis

### Gene no. Gene name Accession no. CDS length (bp) Predicted protein No. of introns **Genomic location** Length pl Μw (aa) (kDa) DCLs 1 BnDCL1A BnaA10g00800D 5439 1812 5.97 202.15 17 chrA10:390521..398004 2 BnDCL1C BnaC05q00860D 5430 1809 6.01 201.54 19 chrC05:435258..442575 З BnDCL2A BnaA05g32540D 4167 1388 6.77 156.70 21 chrA05:22308269..22314738 4 BnDCL2C BnaC05q47910D 4170 1389 7.65 156.94 21 chrC05:42659693..42666090 5 BnDCL3A XP\_013656716<sup>a</sup> 4572 1523 5.91 141.60 23 A8 NC\_027764.1 (14560641..14567368) 6 BnDCL3C BnaC03g54010D 4596 1531 6.04 172.25 24 chrC03:40673425..40680494 7 **RnDCI 4A** BnaA10g15080D 4929 1642 6 22 185.01 24 chrA10.11827194 11836581 8 BnDCL4C BnaC09g37430D 4926 184.48 24 chrC09:40709757..40719200 1641 6.22 AGOs 1 BnAGO1A1 BnaA08g03260D 3135 1044 9.40 115.85 22 chrA08:2681072..2686918 2 BnAGO1A2 BnaA05g17460D 3261 1086 9.38 120.33 22 chrA05:12290150..12296889 З BnaC08g46720D 20 BnAGO1C1 3159 1052 9.45 116 78 chrC08\_random:953598..958888 4 22 BnAGO1C2 BnaC05g25730D 2943 980 9.17 109.15 chrC05:21116236..21122145 5 BnAGO2A BnaA09q25290D 3072 1023 9.66 113.91 З chrA09:18324937..18328648 6 2664 2 BnAGO2C BnaCnng68320D 887 9.44 100.70 chrCnn\_random:67905947..67909115 7 BnAGO3A BnaA05g14760D 3033 1010 9.51 112.74 1 chrA05:9228612..9232112 8 BnAGO3C BnaC06g41790D 3108 1035 9.54 114.35 1 chrC06\_random:1066401..1069740 9 BnAGO4A1 BnaA04g15560D 2769 922 8.91 103.06 21 chrA04:12884285..12890171 BnaA07g13010D 21 10 BnAGO4A2 2772 923 8.82 103.36 chrA07:11653377..11659092 11 BnAGO4C1 BnaC04g38560D 2769 922 8.87 103.11 21 chrC04:39739228..39745117 12 BnAGO4C2 BnaC04g54830D 2772 923 8.87 103.36 21 chrC04\_random:2230442..2236093 13 BnAGO5A BnaA07g13430D 2874 957 9.52 106.71 19 chrA07:11907998..11912994 14 BnAGO5C BnaC04g16450D 2859 952 9.62 106.03 20 chrC04:14487678..14492897 15 BnaA03q15180D 9.03 97.20 21 chrA03:7005357..7010304 RnAGO6A 2604 867 16 BnAGO6C BnaC03g18310D 2604 867 8.99 97.37 21 chrC03:9391663..9396398 17 BnAGO7A BnaA07g24280D 2955 984 9.37 112.47 2 chrA07:18160385..18163845 18 BnAGO7C1 BnaC02g19190D 2931 976 9.32 111.58 2 chrC02:15451981..15455314 BnaC06g43420D 5 19 BnAGO7C2 2700 899 9.38 102.67 chrC06\_random:2865213..2868659 20 BnAGO8A BnaA02g05290D 2721 906 9.31 101.34 20 chrA02:2403187..2408757 21 BnaA10g14450D 2715 904 9.31 102.04 21 chrA10:11492941..11497567 BnAGO9A1 22 BnAGO9A2 BnaA10g14440D 2748 915 9.38 102.52 20 chrA10:11481378..11486279 23 BnAGO9C1 BnaCnng35060D 2721 906 9 42 102 57 21 chrCnn\_random:33265084..33269730 24 BnAGO9C2 BnaC09g36780D 2763 920 9.20 103.80 23 chrC09:40119310..40124830 25 BnaC09q36860D 19 chrC09.40226286\_40231475 BnAGO9C3 2649 882 9.23 99.37 26 BnAGO10A BnaA06g36540D 2928 975 9.38 109.27 16 chrA06:23915363..23920283 27 BnAGO10C BnaC07g17330D 2946 981 9.38 109.75 16 chrC07:23533982..23539648 RDRs 2 BnaA06g09600D 3330 7.50 126.63 1 BnRDR1A 1109 chrA06:5134289..5137990 2 XP\_013669643a 3 BnRDR1C1 3207 1068 5.93 84.40 C1 NC 027767.1 (42513981..42517438) З BnRDR1C2 BnaC05g10980D 3282 1093 6.62 124.83 3 chrC05:6358607..6362657 4 З BnRDR2A BnaA09g22040D 3390 1129 6.15 128.25 chrA09:14659131..14663050 5 BnRDR2C BnaCnng57100D 3378 1125 6.06 127.52 3 chrCnn\_random:56892147..56896339 6 BnRDR3A BnaA09g43930D 3003 1000 8.34 113.41 16 chrA09:30278290..30284646 7 BnRDR3C BnaC08q36490D 3000 999 8.05 113.19 16 chrC08:33743646..33750434

### TABLE 1 | List of *B. napus DCL*, AGO, and RDR genes.

(Continued)

### TABLE 1 | Continued

Gene no.	Gene name	Accession no.	CDS length (bp)	Predic	ted prot	ein	No. of introns	Genomic location
				Length (aa)	pl	Mw (kDa)		
8	BnRDR4A	BnaA10g12910D	2982	993	8.58	112.64	17	chrA10:1051602010520784
9	BnRDR5A1	BnaA07g00800D	2928	975	7.13	110.64	19	chrA07:581670586372
10	BnRDR5A2	XP_013652716 <sup>a</sup>	2853	951	6.32	84.22	16	A7 NC_027763.1 (14935631498604, complement)
11	BnRDR5A3	BnaA07g00770D	2757	918	6.08	103.58	16	chrA07:561541565534
12	BnRDR5C1	BnaC07g01150D	2934	977	6.10	110.60	17	chrC07:17649721769964
13	BnRDR5C2	BnaC07g01170D	2877	958	6.61	108.36	16	chrC07:18057111810421
14	BnRDR5C3	BnaC07g01120D	2586	861	5.65	97.84	16	chrC07:17386921746427
15	BnRDR6A	XP_013655642 <sup>a</sup>	3597	1198	6.93	137.07	1	A8 NC_027764.1 (31301353141222, complement)
16	BnRDR6C	XP_013720258 <sup>a</sup>	3597	1198	6.78	137.03	1	Unplaced scaffold NW_013650343.1 (584631589152)

<sup>a</sup>The noted sequences were from the NCBI BioProject: PRJNA293435, while all the others were from the NCBI BioProject: PRJEB5043.

indicated that *B. napus* genome possessed all copies of *DCL*, *AGO*, and *RDR* genes from the two progenitor species and contained three extra copies of *RDR* genes (**Table 1** and Table S1).

# Classification of *B. napus DCL*, *AGO*, and *RDR* Genes Based on Phylogenetic Analysis

In order to examine the phylogenetic relationship of DCL, AGO, and RDR families, we constructed unrooted phylogenetic trees of all BnDCL, BnAGO, and BnRDR protein sequences along with their A. thaliana, B. rapa, and B. oleracea homologs (Figure 1). The 8 BnDCLs were obviously divided into four groups as reported for AtDCLs. Each group comprised two members with one contributed by each subgenome (A and C; Figure 1A). Coincidently, B. rapa and B. oleracea contained 4 DCLs with one member for each group (Figure 1A). According to the phylogenetic relationship and sequence homology with AtDCLs, the 8 BnDCLs were named as BnDCL1A, BnDCL1C, BnDCL2A, BnDCL2C, BnDCL3A, BnDCL3C, BnDCL4A, and BnDCL4C in accordance with their genomic localization (in A or C). Besides, BnDCLs showed high sequence similarities (from 78.7 to 87.6% for each group) to their Arabidopsis counterparts (Table S3).

Based on the NJ phylogenetic tree and the protein sequence homologies with AtAGOs, the BnAGOs family consisted of 4 AGO1s (BnAGO1A1, BnAGO1A2, BnAGO1C1, and BnAGO1C2), 2 AGO2s (BnAGO2A and BnAGO2C), 2 AGO3s (BnAGO3A and BnAGO3C), 4 AGO4s (BnAGO4A1, BnAGO4A2, BnAGO4C1, and BnAGO4C2), 2 AGO5s (BnAGO5A and BnAGO5C), 2 AGO6s (BnAGO6A and BnAGO6C), 3 AGO7s (BnAGO7A, BnAGO7C1, and BnAGO7C2), 1 AGO8 (BnAGO8A), 5 AGO9s (BnAGO9A1, BnAGO9A2, BnAGO9C1, BnAGO9C2, and BnAGO9C3) and 2 AGO10s (BnAGO10A and BnAGO10C). The total *B. napus* AGO copies in each subgroup were corresponding to that in

the two progenitors B. rapa and B. oleracea except one extra AGO9 (BnAGO9C3) and one less AGO8 (Figure 1B). Notably, an uneven number of AGO gene copies from these three Brassica species was observed. Genomes of B. rapa and B. oleracea comprised two copies of AGO1, AGO4, and AGO9 genes, which was identical to the subgenomes A and C of B. napus except that the subgenome C of B. napus contained an extra copy of AGO9 (BnAGO9C3). In addition, B. oleracea genome and B. napus subgenome C possessed two copies of AGO7 genes. Thus, copy numbers of these AGO genes were higher in these Brassica species than in A. thaliana. Instead, B. napus subgenome C did not contained any AGO8 gene. Besides the exceptions herein described, the number (only one) of the remaining AGOs in genomes of B. rapa and B. oleracea and subgenomes A and C of B. napus was identical to A. thaliana. On the other hand, the distribution of gene members of AGO groups was generally even in the A and C subgenomes of B. napus except AGO7, AGO8, and AGO9 (Figure 1B and Table 1). Additionally, sequence similarity between BnAGOs and Arabidopsis homologs was generally high, ranging from 60.8 to 92.7%, while the similarity among sequences of the same group of BnAGOs was 87.9-91.1% for AGO1, 86.7% for AGO2, 90.0% for AGO3, 96.8-99.6% for AGO4, 94.5% for AGO5, 98.6% for AGO6, 76.5-91.0% for AGO7, 72.4-97.3% for AGO9 and 96.1% for AGO10 (Table S3).

Like DCLs and AGOs, RDRs in *B. napus* were named after the Arabidopsis homologs which were the phylogenetically closest in the NJ tree and showed the highest protein sequence homologies. Consequently, *B. napus* geneome comprised 3 RDR1s (BnRDR1A, BnRDR1C1, and BnRDR1C2), 2 RDR2s (BnRDR2A and BnRDR2C), 2 RDR3s (BnRDR3A and BnRDR3C), 1 RDR4 (RDR4A), 6 RDR5s (BnRDR5A1, BnRDR5A2, BnRDR5A3, BnRDR5C1, BnRDR5C2, and BnRDR5C3) and 2 BnRDR6s (BnRDR6A and BnRDR6C; **Figure 1C**). It is noteworthy that *B. napus* genome possessed 6 *RDR5* genes, 3 in each subgenomes (A and C), which is





more than that in genomes of the progenitors *B. rapa* (1) and *B. oleracea* (2), indicating the further multiplication of *RDR5* genes in *B. napus* genome during evolution. Besides, *B. oleracea* genome and *B. napus* subgenome C carried 2 *RDR1* genes but no any *RDR4* gene. Thus, composition of these RDRs in the three *Brassica* genomes was distinct from Arabidopsis. Furthermore, as observed for BnAGO family, members of RDR1 and RDR4 were unequally distributed in the A and C subgenomes of *B. napus*. In addition, protein sequence similarity between BnRDRs and AtRDRs was generally high, ranging from 63.1 to 90.4%, while the similarity within the same

group of BnRDRs was 61.5–95.2% for RDR1, 98% for RDR2, 97.6% for RDR3, 70.2–99.3% for RDR5 and 99.4% for RDR6 (Table S3).

Collectively, genome of the allopolyploid species *B. napus* possesses almost all copies of *DCL*, *AGO*, and *RDR* genes from its progenitor species *B. rapa* and the *RDR5* group in *B. napus* appears to have undergone further expansion through duplication during evolution. Furthermore, compared with Arabidopsis, some *AGO* and *RDR* genes such as *AGO1*, *AGO4*, *AGO9*, and *RDR5* have significantly expanded in these *Brassica* species.

## Physico-Chemical Properties and Domain Composition of *B. napus* DCL, AGO, and RDR Proteins

Physico-chemical properties and domain composition of *B. napus* DCL, AGO, and RDR proteins were generally similar to their Arabidopsis counterparts, though some differences were noticed. For DCLs, seven out of eight BnDCL proteins contained DEAD, Helicase-C, Duf283, PAZ, and RNaseIII domains as reported for Arabidopsis DCLs. The remaining one (BnDCL3C) lacked the DEAD domains, which is distinguishable from AtDCL3 in this regard (**Figure 2A**). Further, comparison of the DEAD domain region of all Arabidopsis and three *Brassica* species revealed two deletions of 31 and 9 amino acids (aa), respectively, in BnDCL3C (Figure S1A). The correct sequence of

BnDCL3C awaits further experimental confirmation. The gene length was highly similar within groups but considerably varied between groups of BnDCLs. BnDCL1s were the largest (5439 and 5430 bp) followed by BnDCL4s (4929 and 4926 bp) and BnDCL3s (4572 and 4596 bp), while BnDCL2s were the smallest (4167 and 4170 bp; **Table 1**). This group-wise variation in gene length is also observed in AtDCLs (Table S1).

The domain composition of BnAGOs was identical to that of AtAGOs. All BnAGOs possessed DUF1785, PAZ, and PIWI domains. Besides, all BnAGO1 proteins contained an additional Gly-rich Ago1 domain (**Figure 2B**). Furthermore, we examined BnAGOs for presence of the four key active residues (DDH/H) in the PIWI domain that are responsible for the endonuclease property of AGO proteins involved in RNAi. The PIWI domain



FIGURE 2 | Domain composition of *B. napus* and *A. thaliana* DCL (A), AGO (B), and RDR (C) protein sequences. Domains are indicated as boxes in various colors. The diagrams were drawn to scale.

sequence alignment revealed that all 11 members of groups BnAGO1s (4), BnAGO5s (2), BnAGO7s (3), and BnAGO10s (2) possessed the four key active residues, suggesting that they might have endonuclease activity (**Figure 3**). The remaining BnAGOs belonging to AGO2, AGO3, AGO4, AGO6, AGO8, and AGO9 groups contained substitutions of the two H residues, which is similar to their Arabidopsis AGO counterparts. Furthermore, the length of the *BnAGOs* CDS varied from 2604 bp for *BnAGO6* to 3261 bp for *BnAGO1A2*, potentially encoding 867 and 1086 amino acids, respectively (**Table 1**). All BnAGOs encode for ~100 kDa basic proteins with a pI ranging from 8.82 to 9.66. The physico-chemical properties of BnAGOs were generally similar to AtAGOs and conserved within all AGO groups.

All the 16 BnRDR proteins carried an RdRP domain (**Figure 2C**). Additionally, members of BnRDR1, BnRDR2, and BnRDR6 groups bore a RBD domain as observed in the same groups of AtRDRs (**Figure 2C**). Furthermore, we examined BnRDRs for presence of the catalytic motif in the RdRP domain. The sequence alignment demonstrated that all members of BnRDR1, BnRDR2, and BnRDR6 groups shared a common DLDGD motif in the catalytic domain, while all members of BnRDR3 and five out of six BnRDR5 proteins contained DFDGD motif. The exception was BnRDR5A3, which did not contain

	11/0	
Atacol ·BRIDLUSDRDTIEG	1100 1200 1240 1240 1240	878
BnAGO1A1 :RRIPLVSDRPTIIFG.	VINPHGEDS-SPSIAAVVASOWPEVTKVASUVCAQHAQLIQUITEKEWKDECKGVVIGGMIKELLIAFRRSTGHK-PIKIIFVGVSEGGFVVULVELDIMIRKACASLEAGVOPP	886
BnAGO1A2 :RRIPLVSDRPTIIFG.	NTHPHPGEDS-SPSIAAVVASÇDWPEITKYAGLVCAÇA <mark>k</mark> rçeliçdlfkewkdfçkgvvtggmikelliafrrstghk-plriifyr <mark>s</mark> gvsegçfyçvllyeldairkacasleagyçpp	924
BnAGO1C1 :RRIPLVSDRPTIIFG.	VTHPHPGEDS-SPSIAAVVASQDWPEVTKYAGLVCAQA <mark>H</mark> RQELIQDLFKEWKDFQKGVVTGGMIKELLIAFRRSTGHK-PLRIIFYR <mark>I</mark> GVSEGQFYQVLLYELDAIRKACASLEAGYQPP	882
BnAGO1C2 :RRIPLVSDRPTIIFG	VTHPHPGEDS-SPSIAAVVASQDWPEITKYAGLVCAQA <mark>H</mark> RQELIQDLFKEWKDFQKGVMTGGMIKELLIAFRRSTGHK-PLRIIFYR <mark>I</mark> GVSEGQFYQVLLYELDAIRKACASLEAGYQPP	808
AtAGO2 :-FFKKEDEVMFIG.	TVNPFAARDKM-SPSIVAVVGTLNWPEANRYAARVIAQPERREEIQGFGDACLELVKAHVQATGKR-PNKIVIFNGVSDAQFDMVLNVELLDVKLTFEKNGYNPK	849
BnAGO2A :SFFRKDDQVMFIG.	NUNPASRDKM-SPSIVAVVGTLNNPEANRYAARVIACPIRCEEIGGFGETCLELVKAHVQSTGKR-PNKIVIFRGVSBGQFDMVLNVELLDVKLTFEKNGVNPK	873
BHAGO2C :-FFRKDDQVMFIG.	VNHPASRDALSPSIVAVGTLNVPEARKIAKVIAQPEKLELQGFGETCLELVAAH-VQSTGAR-PNAIVIFNGVSDGQFDFUVINVELDVNTFFEX-NGSIAPA	1028
BnAGO3A SFEKREDEVMEIG	VNDPARDUND SPOLVAVVGIDNVPLANALARAKINAKSENLELUGE IVNDPARDUTT-GOSTVAVVGIDNVPLANALARVIADDRKFFIGEGDTCFFUKAHDNKTVFFGOSDGGEDNVINVELUNVLTVFFG-DNVVDK	848
BnAGO3C :SFFKREDEVMFIG	VNHPAARDOT-SPSIVAVVGTLNWPEANRYAARVIAOPRKEEIEGFGDTCLELVKAHFOATKKR-PNKIVIFR GVSDGOFDMVLNSELLDVKLTFGRNNYFPK	873
AtAGO4 :PAFTVISKVPTIILG	DVSHGSPGQSD-VPSIAAVVSSREWPLISKYRASVRTQPSKAEMIESLVKKNGTEDDGIIKELLVDFYTSSNKRKPEHIIIFREGVSESQFNQVLNIELDQIIEACKLLDANWNPK	774
BnAGO4A1 : PAFTVISKVPTIILG	DVSHGSPGQSD-VPSIAAVVSSRQWPLVSKYRASVRTQPSKAEMIESLVKKNGTEDDGIIKELLVDFYTSSGKRKPEHIIIFR <mark>H</mark> GVSESQFNQVLNIELDQIIEACKLLDENWNPK	772
BnAGO4A2 :PAFTVISKVPTIILG	NSHGSPGQSD-VPSIAAVVSSRQWPLISKYRASVRTQPSKAEMIESLVKKNGTEDDGIIKELLVDFYASSGKRKPEHIIIFR <mark>G</mark> GVSESQFNQVLNIELDQIIEACKLLDENWNPK	773
BnAGO4C1 : PAFTVISKVPTIILG	TVSHGSPGQSD-VPSIAAVVSSRQWPLVSKYRASVRTQPSKAEMIESLVKKNGTEDDGIIKELLVDFYTSSGKRKPEHIIIFR GVSESQFNQVLNIELDQIIEACKLLDENWNPK	772
BnAGO4C2 :PAFTVVSKVPTIILG	VSHGSPGQSD-VPSIAAVVSSRQMPLISKYRASVRTQPSKAEMIESIVKKNGTEDGIIKELLVDFYASSGKRRPENIIIFH GVSESQFNQVLNIELDQIIEACKLLDENWNPK	773
BnAGO5A :RRIPLISDRPTIIG	VINEYTEETSED STOLAA VASHWEELIKAINGU VASHAALEIIYUDILAIYUEYKSUUFKSGULFEATLAI — AKAIQYI-YXIIIIING VOEGYISYUUDHEMIAAKAKSUUSETIYDI VYHEGOGENS-GOSTAAVVASHWEETKYSGUVSGATTERFTIEDIIYUVYKONEGOGOUHEGATHEFTIIAIREATGGAL-DIRTIFYKGOSGAULHEMIAAKAASUSE	799
BnAGO5C :RRIPLISDRPTIIFG	VIHEOFGEDS-SPIAAVVASHDWPEITKYRGIVSACTEREIIEDINYLVODFGROUNGENIEHIIAF TKAGAK-PORIIFYN GVSEGOFSOULHENIAIRKACASLEERVLPF	794
AtAGO6 :YNIPLINKIPTLILG	NYSHGPPGRAD-VPSVAAVVGSKCWPLISRYRAAVRTOSPRLEMIDSLFOPIENTEKGDNGIMNELFVEFYRTSRARKPKOIIIFREGVSESOFEQVLKIEVDOIIKAYORLGESDVPK	740
BnAGO6A :SNIPLINKIPTLILG	VSHGSPGRAD-VPSIAAVVGSKNWPLISRYRAAVRTQSPKMEMIDSLFQPVEDFVNGDNGIMNELFVEFFKTSNARKPKQIIIFR GVSESQFNQVLNIEVDQIIKAYQRLGETDVPK	730
BnAGO6C :SNIPLINKIPTLILG	SYSHGSPGRAD-VPSVAAVVGSKNWPLISRYRAAVRTQSPKMEMIDSLFQPVEDFVNGDNGIMNELFVEFFKTSNARKPKQIIIFR <mark>I</mark> GVSESQFNQVLNIEVDQIIKAYQRLGETDVPK	730
AtAGO7 :SHIPRLLRPDEPVIFMG	VTHPHPFDDC-SPSVAAVVGSINWPEANRYVSRMRSQTERQEIIQDLDLMVKELLDDFYKAVKKL-PNRIIFFROVSETQFKKVLQEELQSIKTACSKF-QDYNPS	838
BnAGC7A :SHIPRLFRLDEPVIFMG.	NYTHPHPFDDC-SPSVAAVVGSINNPEANRYVSRMRSGTERGEIIGDLDVMVKELLEDFHKAKKL-PNRIIFFRGVSETGFKKVLGEELGAIKAACSNF-DUNNPT	829
BNAGO7C1 :SHIPRLIRVDEPVIFMG.	VIHPHPFDDC-SPSVAAVVGSINNPEANKHVSKMISGITHGEIIGDLDUMVKELLDDF-IKALNKL-PNKIIFFNGVSETGRKVLGEELGSIKAACSKF-GGIFT VMEUNDEPDC-SPSVAAVVGSINNPEANKYVGNBGGT BOFTTOTI	823
Atacos : GTMDLVMBVDTIIG	VINERFEDUCESFSVARVVGSTMRELAMIVSSUMRGLANGSVARVETIGUD TVSHCSGCGSRHTGETALVVGSTMREFEMULTSKVBLCVBRGSKVENTGSTEKDVSRKDDCTMBETILDEF-MSSCKKEDNITTTERGGSESGCRGVUNTFLDGMA	706
BnAGOSA :CAMPSVTOVPTIIVG	VSHGSPGSD-VPSVAAVVSSROWPLISKYRACVRTCSRVEMIDNLFKLVPCKEEMVDEGIFRSLLVDF-VSSSCKRKPDHIIIFNGVSESCFNOVINIELDOMMOACKFLDEKWNPK	754
AtAGO9 :PAMPKVTQVPTIIVG	NSHGSPGQSD-IPSIAAVVSSRQWPLISKYKACVRTQSRKMEMIDNLFKPVNGKDEGMFRELLLDFYYSSENRKPEHIIIFREGVSESQFNQVLNIELDQMMQACKFLDDTWHPK	746
BnAGO9A1 :AMMPLVTQVPTFIVG	NSHGSPNQAD-IPSIAAVVGSREWPLISKYRACVRTQSRKVEMIDNLFKLVPNEKGKLVDEGIFWELLFDFYTSSGKRRPEHIIIFR <mark>H</mark> GVSDSQFNQVLNIELDQIMQACKFVEENWEPK	752
BnAGC9A2 :AMMPLVTQVPTFIVG	NYSHGSPGQSD-IPSVAAVVGSIEWPLISKYRACVRTQSRKVEMIDNLFKPATNEKGEPVDDGIFRLTSSTFKLHSHSKKRKPEHIIIFR <mark>I</mark> GVSESQFNQVLNIELDQMMQACKFVEENWEPK	763
BnAG09C1 :AMMPLVTQVPTFIVG	TVSRGSPNQSD-IPSIAAVVGSREWPLISKYRACVRTQSRKTEMIDNLFKLVPNEKGKLVDEGIFWELLFDFYISSGKRRPEHIIIFNGVSDSQFNQVLNIELDQMMQACKFVEENWEPK	754
BnAGO9C2 :AMMPLVTQVYTFIVG	VSHGSPGGSD-IPSVAAVVGSREWPLISKYRACVF7GSRKVEMIDNLFKPV/DENGKPVDEGIFWELLEDFYTSSGNRRPEHNIIFFHGVSESKFNGVINIELDCMMGACKFVEENMEPK	811
Atagoio :CRIDLVSDIDTIIG	Vongesruged-folgen vosakevelisatiak vitugsatvelui datse velui denkatvegiskelde (111) angese istatus sonderververververververververververververve	827
BnAGO10A :CRIP-LVSDIPTIIFG	VINPENGEES-SPITATVASCOMPEVIKYAGIUCAGA RELIQUITIVETKOFVAGANEDILISF - REATGOR - DIRIFFNG OSEGOFVULIVELDAIREACHDETNYF	809
BnAGC10C :CRIPLVSDIPTIIFG.	VTHPENGEES-SPSIAAVVASQDWPEVTKYAGLVCAQA RQELIQDLYKTWQDFVRGTVSGGMIRDLLISFRKATGQK-PLRIIFYRGVSEGQFYQVLLYELDAIRKACASLEPNYQPP	816
1280	1320 1360 <b>1</b> 400	1014
1280 Atago1 :VTFVVVQKRHHTRLFAQ Bracolal :VTFVVVQKRHHTRLFAQ	1320 1360 1400 IHNDRHSVDRSGNILÞGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVUMDENNFTADGLQSLTNULCYTYARCTRSVSIVPPAYYA <mark>B</mark> LAAFRAFYMEPETSDSGMAGGMAGG HUNDRJSVDRSGNILÞGTVUDSKICHPTEFVLCSHAGIQGTSRPAHYHVUMDENNFTADGLQSLTNULCYTYARCTRSVSIVPPAYYA	1014
1280 Atago1 :VTFVVVQKRHHTRLFAQ Bnago1A1 :VTFVVVQKRHHTRLFAQ Bnago1A2 :VTFVVVQKRHHTRLFAH	1320 1360 1400 HNDERHSVDRSGNILPGTVVDSKICHPTEFDFVLCSHAGIQGTSRPAHYHVLWDENNFTADGLQSLTNULCYTYARCTRSVSIVPPAYYAELAAFRAFYMEPETSDSGSMASGSMARG HNDRNSVDRSGNILPGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLWDENNFSADGLQSLTNULCYTYARCTRSVSIVPPAYYAELAAFRAFYMEPETSDSGGMASGSMARG HNDRNSVDRSGNILPGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLWDENNFSADGLQSLTNULCYTYARCTRSVSIVPPAYYAELAAFRAFFYMEPETSDSGGMASGSMARG	1014 1021 1060
1280 AtAGO1 :VTFVVVQKRHHTRLFAQ BnAGO1A1 :VTFVVVQKRHHTRLFAQ BnAGO1A2 :VTFVVVQKRHHTRLFAD BnAGO1C1 :VTFVVVQKRHHTRLFAO	1320 HADDRHSVDRSGNILÞGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLMDENNFTADGLQSLTNNLCYTYARCTRSVSIVPPAYYA <mark>B</mark> LAAFRAFYMEPETSDSGNASGSNARG HHNDRNSVDRSGNILÞGTVVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLMDENNFSADGLQSLTNNLCYTYARCTRSVSIVPPAYYABLAAFRAFYMEPETSDSGNASGSNARG HHNDRNSVDRSGNILÞGTVDSKICHPTEFDFYLCSHAGIQGTSRPAHYHVLMDENNFSADGLQSLTNNLCYTYARCTRSVSIVPPAYYABLAAFRAFYMEPETSDSGNASGSNARG HNDRNSVDRSGNILÞGTVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCYTYARCTRSVSIVPPAYYABLAAFRAFFYMEPETSDSGNASGSNARG	1014 1021 1060 1018
1280 AtAGO1 :VTFVVQKRHHTRLFAQ BnAGO1A1 :VTFVVQKRHHTRLFAQ BnAGO1A2 :VTFVVQKRHHTRLFAH BnAGO1C1 :VTFVVQKRHHTRLFAH	1320 1360 1400 HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGNAGGRNMAGP HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGNAGGRNMAGP HHDDRNSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGSMASGSMARG HHDDRNSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGSMASGSMARG	1014 1021 1060 1018 944
1280 AtAGO1 :VTFVVVQKRHTRLFAQ BnAGO1A1 :VTFVVVQKRHTRLFAQ BnAGO1A2 :VTFVVVQKRHTRLFAQ BnAGO1C1 :VTFVVVQKRHTRLFAQ AtAGO2 :ITVVVQKRHTRLFAH	1320 IHNDRHSVDRSGNILDGTVVDSKICHPEEDFVLCSHAGIGGTSRPAHYHVLMDENNFTADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMAGGMARG IHNDRNSVDRSGNILDGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMAGGMARG IHNDRNSVDRSGNILDGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMASGMARG IHNDRNSVDRSGNILDGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMASGMARG IHNDRNSVDRSGNILDGTVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMASGMARG IHNDRNSVDRSGNILDGTVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCTYTARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMASGMARG IHNDRNSVDRSGNILDGTVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCTYTARCTRSVSIVPPAYYALAAFRARFYMEPETSDSGMASGMARG INNDGSDKGNVPSGTVDFVTVDFVFTSPCKGAFGASAFAA	1014 1021 1060 1018 944 987
1280 Atago1 : VTFVVVQRRHTRLFAQ Bhago1A1 : VTFVVVQRRHTRLFAQ Bhago1A2 : VTFVVVQRRHTRLFAH Bhago1C2 : VTFVVVQRHHTRLFAH Atago2 : ITVIVAQRRHTRFFA Bhago2A : ITVIVAQRRHTRFFA	1320 1400 HHDDRHSVDRSGNILÞGTVVDSKICHPTETDFYLCSHAGIGGTSRPAHYHVLMDENNFTADGLQSLTNNLCTYTARCTRSVSIVPPAYYA <mark>B</mark> LAAFRAFYMEPETSDSGNASGSMARG HHDDRHSVDRSGNILÞGTVVDSKICHPTETDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGNASGSMARG HHDDRNSVDRSGNILÞGTVDSKICHPTETPFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGNASGSMARG HHDDRNSVDRSGNILÞGTVDSKICHPTETPFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGNASGSMARG HHDDRNSVDRSGNILÞGTVDSKICHPTETPFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYMELAAFRAFYMEPETSDSGSMASGSMARG HHDDRNSVDRSGNILÞGTVDSKICHPTETPFYLCSHAGIGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYMELAAFRAFYMEPETSDSGSMASGSMARG HNDDGSDKGNVÞSGTVDTKVIHPEYDFYLCSHHGGIGTSKPTHYYTLMDEIGFTSDQVQKLIFEMCFTFTRCTKPVSLVPPYYADMVAFKGRMYHEASSREKNIKQQPFG	1014 1021 1060 1018 944 987 1003
1280 AtAGO1 :VTFVVVQKRHTRLFAQ BnAGO1A: VTFVVVQKRHTRLFAQ BnAGO1A: VTFVVVQKRHTRLFAA BnAGO1C: VTFVVVQKRHTRLFAA BnAGO2C :VTFVVQKRHTRFFA BnAGO2C :TTVVVQKRHTRFFA BnAGO2C :TTVVVQKRHTRFFA BnAGO2C :TTVVVQKRHTRFFA	1320 1360 1400 HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDDRNSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGSMASGSMARG HHDDRNSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGSMASGSMARG IHNDRNSVDRSGNILÞGTVVDSKICHPTEFDFVLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCTTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGSMASGSMARG INNDGSDKGNVÞSGTVUDTKVIHPFEYDFVLCSHAGIGGTSRPHTYTLMDELGFTSDQVQKLIFEMCFFFTRCTRFVSIVPPYYADMVAFRGMYHFASSRENFEND	1014 1021 1060 1018 944 987 1003 868
1280 AtaGO1 : VTFVVVQKRHTRLFAQ BhAGO1A1 : VTFVVVQKRHTRLFAQ BhAGO1A2 : VTFVVVQKRHTRLFAA BhAGO1C2 : VTFVVVQKRHTRLFAA AtaGO2 : ITVVVQKRHTRFFAA BhAGO2A : ITVVVQKRHTRFFAA BhAGO3 : ITVVVQKRHTRFFAA AtAGO3 : ITVVVQKRHTRFFAA AtAGO3 : ITVVVQKRHTRFFAA	1320 14000 1400 14000 1400 1400 1400 1400 1400 1400 1400 1400	1014 1021 1060 1018 944 987 1003 868 1167 983
1280 AtaGO1 : VTFVVVQKRHHTRLFAQ BhAGO1AI : VTFVVVQKRHHTRLFAQ BhAGO1A2 : VTFVVVQKRHHTRLFAQ BhAGO1C2 : VTFVVVQKRHHTRLFAQ BhAGO2C : ITVIVAQKRHTRFFAA BhAGO2C : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA	1320 14000 14000 14000 14000 14000 14000 14000 14000 14000 14000 140	1014 1021 1060 1018 944 987 1003 868 1167 983 1008
1280 AtaGO1 : VTFVVVQKRHITRLFAQ BnAGO1A1 : VTFVVVQKRHITRLFAQ BnAGO1A2 : VTFVVQKRHITRLFAQ BnAGO1C2 : VTFVVQKRHITRLFAH AtaGO2 : ITVVIQKRHITRFFAP BnAGO2A : ITVVIQKRHGTRFFAP BnAGO3C : ITVVIQKRHGTRFFAP BnAGO3C : ITVVIQKRHGTRFFAP BnAGO3C : ITVVIQKRHGTRFFAP AtaGO4 : FLLIVQKRHTFFAP AtaGO3 : FLLIVQKRHITRFFAP	1320 14000 14000 14000 14000 14000 14000 14000 14000 14000 140	1014 1021 1060 1018 944 987 1003 868 1167 983 1008 897
1280 AtagO1 : VTFVVVQKRHTRLFAQ BhAGO1A1 : VTFVVVQKRHTRLFAQ BhAGO1A2 : VTFVVVQKRHTRLFAA BhAGO1C2 : VTFVVVQKRHTRLFAA AtagO2 : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO4A : FLLLVAQNHHTKFFCP BhAGO4A : FLLLVAQNHHTKFFCP	1320 1400 1400 HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFTADGLQSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDDRHSVDRSGNILÞGTVVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDDRNSVDRSGNILÞGTVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLQSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSD	1014 1021 1060 1018 944 987 1003 868 1167 983 1008 897 895
1280 Atago1 : VTFVVVQKRHTRLFAQ BhaGo1A1 : VTFVVVQKRHTRLFAQ BhaGo1A2 : VTFVVVQKRHTRLFAA BhaGo1C2 : VTFVVVQKRHTRLFAA Atago2 : ITVIVAQKRHTRLFAA BhaGo2A : ITVIVAQKRHTRFFAA BhaGo2A : ITVIVAQKRHTRFFAA BhaGo3A : ITVIVAQKRHTRFFAA BhaGo3A : ITVIVAQKRHTRFFAA BhaGo3A : ITVIVAQKRHTRFFAA BhaGo3A : ITVIVAQKRHTRFFAA BhaGo3A : ITVIVAQKRHTRFFFAA BhaGo3A : ITVIVAQKRHTRFFFAA BhaGo4A1 : FLLLVAQKNHTKFFGFP BhaGo4A1 : FLLLVAQKNHTKFFGFP	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 987 1003 868 1167 983 1008 897 895 896
1280 AtaGO1 : TTYTVVQKRHHTRLFAG BnAGO1A: VTFVVVQKRHHTRLFAG BnAGO1A: VTFVVVQKRHTRLFAG BnAGO1C: VTFVVQKRHTRLFAG AtaGO2 : TTYTVQKRHTRLFAG BnAGO2A: ITVTVAQKHCTRFFAG BnAGO2A: ITVTVAQKHCTRFFAG BnAGO3C: ITVTVAQKHCTRFFAG BnAGO3C: ITVTVAQKHCTRFFAG BnAGO3C: ITVTVAQKHCTRFFAG BnAGO4A: FLLLVAQKNHTKFFCG BnAGO4A: FLLLVAQKNHTKFFCG BnAGO4A: FLLLVAQKNHTKFFCG BnAGO4A: FLLLVAQKNHTKFFCG	1320 14000 14000 14000 14000 14000 14000 14000 14000 14000 140	1014 1021 1060 944 987 1003 868 1167 983 1008 897 895 896 895
1280 AtagO1 : VTFVVVQKRHTRLFAQ BhAGO1A1 : VTFVVVQKRHTRLFAQ BhAGO1A2 : VTFVVVQKRHTRLFAQ BhAGO1C2 : VTFVVVQKRHTRLFAA AtagO2 : ITVIVAQKRHTRLFAA BhAGO2A : ITVIVAQKRHTRFFAA BhAGO2A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO3A : ITVIVAQKRHTRFFAA BhAGO4A : FLLLVAQKNHTKFFFCP BhAGO4A2 : FLLLVAQKNHTKFFFCP BhAGO4C2 : FLLLVAQKNHTKFFFCP BhAGO4C2 : FLLLVAQKNHTKFFFCP BhAGO4C2 : FLLLVAQKNHTKFFFCP BhAGO4C2 : FLLLVAQKNHTKFFFCAA BhAGO4C2 : FLLVAQKNHTKFFFCAA BhAGO4C2 : FLLVAQKNHTKFFFCAA BhAGO4C3 : FLLVAQKNHTKFFFCAA BhAGO4C3 : FLLVAQKNHTKFFFCAA BhAGO4C4 : FLLVAQKNHTKFFFCAA BhAGO4C4 : FLLVAQKNHTKFFFCAA BhAGO4C4 : FLLVAQKNHTKFFFCAA BhAGO4C4 : FLLVAQKNHTKFFFCAA BhAGO4C4 : FLLVAQKNHTKFFFCAAA BhAGO4C4 : FLLVAQKNHTKFFFCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 987 1003 868 1167 983 1008 895 895 896 895 896 895
1280 Atago1 : VTFVVVQRNHTRLFAQ BhaGo1A1 : VTFVVVQRNHTRLFAQ BhaGo1A2 : VTFVVVQRNHTRLFAA BhaGo1C2 : VTFVVVQRNHTRLFAA BhaGo2C : ITVIVAQRNHTRLFAA BhaGo2A : ITVIVAQRNHTRFFAA BhaGo2A : ITVIVAQRNHTRFFAA BhaGo3A : ITVIVAQRNHTRFFAA BhaGo3A : ITVIVAQRNHTRFFFAA BhaGo3A : ITVIVAQRNHTRFFFAA BhaGo4A1 : FLLLVAQRNHTRFFFCP BhaGo4A2 : FLLLVAQRNHTRFFFCP BhaGo4A3 : FLLVAQRNHTRFFFCP BhaGo4A3 : FLLVAQRNHTRFFFCA3 : FLLVAQRNHTRFFFCP BhaGO4A3 : FLLVAQRNHTRFFFCA3 : FLLVAQRNHTRFFFCA3 : FLUAA3 : F	1320 14000 1400 1400 1400 1400 1400 1400 1400 1400 1400 1400	1014 1021 1060 1018 944 987 868 1167 983 1008 897 895 895 895 895 895 895 895 895
1280 AtaGO1 : VTFVVVQKRHTRLFAQ BhaGO1A1 : VTFVVVQKRHTRLFAQ BhaGO1A2 : VTFVVVQKRHTRLFAQ BhaGO1C2 : VTFVVVQKRHTRLFAA AtaGO2 : VTFVVVQKRHTRLFAA BhaGO2A : ITVIVAQKRHCTRFFAA BhaGO2A : ITVIVAQKRHCTRFFAA BhaGO3C : ITVIVAQKRHCTRFFAA BhaGO3C : ITVIVAQKRHCTRFFAA BhaGO3C : ITVIVAQKRHCTRFFAA BhaGO3C : ITVIVAQKRHCTRFFAA BhaGO3C : ITVIVAQKRHCTRFFAA BhaGO4A2 : FLLIVAQKNHTKFFFFAA BhaGO4C2 : FLLIVAQKNHTKFFFFF BhaGO4C2 : VTFVVVQKRHTRLFAA AtaGO5 : VTFVVVQKRHTRLFAA	1320 1400 1400 HINDERHSVDRSGNILÞGYVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFTADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HINDERNSVDRSGNILÞGYVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HINDENNSVDRSGNILÞGYVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HINDENNSVDRSGNILÞGYVDSKICHPTEFDFYLCSHAGIGGTSRPAHYHVLMDENNFSADGLGSLTNNLCYTYARCTRSVSIVPPAYYALAAFRAFYMEPETSD	1014 1021 1060 1018 944 987 868 1167 983 1008 897 895 896 895 896 975 930
1280 AtagO1 : VTFVVVQRRHTRLFAQ BhAGO1A1 : VTFVVVQRRHTRLFAQ BhAGO1A2 : VTFVVVQRRHTRLFAG BhAGO1A2 : VTFVVVQRHTRLFAG AtagO2 : ITVIVAQRHTRLFAG BhAGO2A : ITVIVAQRHGTRFFAG BhAGO2A : ITVIVAQRHGTRFFAG BhAGO3A : ITVIVAQRHGTRFFAG BhAGO3A : ITVIVAQRHGTRFFAG BhAGO3A : ITVIVAQRHGTRFFAG BhAGO4A : FLLIVAQRHHTRFFCG BhAGO4A2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C2 : FLLIVAQRHHTRFFCG BhAGO4C3 : VTFVVUQRHHTRLFFEG BhAGO5C : VTFVVQRHHTRLFFEG BhAGO5C : VTFVVQRHHTRLFFEG BhAGO5C : FTVIVAQRHHTRFFGT	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 1018 944 987 1003 868 897 895 896 895 895 896 895 895 896 895 895 895 895 895 895 895 895 895 895
1280 AtaGO1 : TTYTVVQKRHHTRLFAG BnAGO1A1 : VTFVVVQKRHHTRLFAG BnAGO1A2 : VTFVVVQKRHHTRLFAG BnAGO1C2 : VTFVVQKRHHTRLFAG AtaGO2 : ITVIVAQKHGTRFFAG BnAGO2A : ITVIVAQKHGTRFFAG BnAGO2A : ITVIVAQKHGTRFFAG BnAGO3C : ITVIVAQKHGTRFFAG BnAGO3C : ITVIVAQKHGTRFFAG BnAGO4C1 : FLLLVAQKNHTKFFCG BnAGO4C1 : FLLLVAQKNHTKFFCG BnAGO4C2 : FLLLVAQKNHTKFFCG BnAGO4C2 : FLLLVAQKNHTKFFCG BnAGO4C2 : VTFVVQKRHTKLFFAG BNAGO5A : VTFVVQKRHTKLFFAG BNAGO5A : TTVIVAQKNHTKFFCG BNAGO5A : TTVIVAGKNHTKFFCG BNAGO5A : TTVIFFCG BNAGO5A : TTVIFFCG BNAGO5A : TTVIFFCG BNAGO5A : TTVIFFCG BNAGO5A : TTVIFFCG	1320 14000 1400 1400 1400 1400 14000 14000 14000 14000 14000	1014 1021 1060 1018 944 987 1003 868 1167 983 1008 897 895 896 895 896 895 895 895 895 895 895 895 895 895 895
1280 AtagO1 : VTFVVVQKRHTRLFAQ BhaGO1A1 : VTFVVVQKRHTRLFAQ BhaGO1A2 : VTFVVVQKRHTRLFAD BhaGO1C2 : VTFVVQKRHTRLFAD AtagO2 : VTFVVQKRHTRLFAD BhaGO2A : ITVIVAQKRHCTRFFAD BhaGO3A : ITVIVAQKRHCTRFFAD BhaGO3A : ITVIVAQKRHCTRFFAD BhaGO3A : ITVIVAQKRHCTRFFAD BhaGO3A : ITVIVAQKRHCTRFFAD BhaGO4A2 : FLLIVAQKNHTKFFCP BhaGO4A2 : FLLIVAQKNHTKFFCP BhaGO4A2 : FLLIVAQKNHTKFFCP BhaGO4A2 : FLLIVAQKNHTKFFCP BhaGO4A2 : FLLIVAQKNHTKFFCP BhaGO5A : VTFVVQKRHTRLFAD AtagO5 : VTFVVQKRHTRLFAD BhaGO5A : FTVIVAQKNHTKFFCP BhaGO5A : FTVIVAQKNHTKFFCP	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 1018 944 987 1003 868 1167 983 1008 897 895 896 975 930 896 935 930 849 849
1280 Atago1 : VTFVVVQKRHTRLFAQ BhAG01A1 : VTFVVVQKRHTRLFAQ BhAG01A2 : VTFVVVQKRHTRLFAQ BhAG01A2 : VTFVVVQKRHTRLFAA Atago2 : ITVIVQKRHTRLFAA Atago3 : ITVIVQKRHGTRFFAA BhAG02A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG03A : ITVIVQKRHGTRFFAA BhAG04A : FLLIVQKNHTKFFCP BhAG04A2 : FLLIVQKNHTKFFCP BhAG04C1 : FLLIVQKNHTKFFCP BhAG04C1 : FLLIVQKNHTKFFCP BhAG05C : VTFVVQKRHTRLFFAA BhAG05C : VTFVVQKRHTRLFFAA BhAG05C : VTFVVQKRHTRLFFAA BhAG05C : VTFVVQKRHTRLFFAA BhAG06A : FTVIVAQKRHTKLFQA BhAG06A : FTVIVQKRHTKLFQA BhAG06C : TTVIVQKRHTKLFQA	1320 1400 14	1014 1021 1060 944 987 1003 868 1167 983 1008 895 895 895 895 895 895 895 895 895 89
1280 AtaGG1 : VTFVVVQKRHHTRLFAG BhagG1A1 : VTFVVVQKRHHTRLFAG BhagG1A2 : VTFVVVQKRHHTRLFAG BhagG1C2 : VTFVVVQKRHTRLFAG AtaGG2 : VTFVVVQKRHTRLFAG BhagG2A : ITVIVAQKHGTRFFAG BhagG2A : ITVIVAQKHGTRFFAG BhagG2A : ITVIVAQKHGTRFFAG BhagG3A : ITVIVAQKHGTRFFAG BhagG3A : ITVIVAQKHGTRFFAG BhagG3A : ITVIVAQKHGTRFFAG BhagG4A : FLLIVAQKNHTKFFGT BhagG4A : FLLIVAQKNHTKFFFG BhagG4A : FLLIVAQKNHTKFFFG BhagG4A : FLLIVAQKNHTKFFFF BhagG4A : FLLIVAQKNHTKFFFF BhagG4A : FLLIVAQKNHTKFFFF BhagG4A : FLLIVAQKNHTKFFFF BhagG4A : FLLIVAQKNHTKFFFF BhagG4A : FLLIVAQKNHTKFFFF BhagG5A : VTFVVQKRHTRLFAG BhagG6C : TTVIVAQKHTRLFAG BhagG6C : FTVIVAQKHTRLFAG BhagG6C : FTVIVAQKHTRLFAG BhagG6C : TTVIVAQKHTRLFAG BhagG6C : TTVIVAQKHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG BhagG6A : ITFVVVQKHHTRLFAG	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 987 1003 868 1167 983 1008 897 895 896 895 896 895 895 896 895 896 849 849 849 849 849
1280 AtagO1 : VTFVVVQKRHHTKLFAQ BhAGO1A1 : VTFVVVQKRHHTKLFAQ BhAGO1A2 : VTFVVVQKRHHTKLFAQ BhAGO1C2 : VTFVVVQKRHHTKLFAA AtagO2 : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO4A : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A3 : VTFVVVQKNHTKLFFAA BhAGO5C : VTFVVVQKNHTKLFFAA BhAGO5C : VTFVVVQKNHTKLFFAA BhAGO5C : TTVIVAQKHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFAA BhAGO5C : TTVIVAQKNHTKLFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFAA BhAGO5C	1320 1400 1400 HHDCRHSVDRSGNILPGTVDSKICHPTEFDFYLCSHAGTQGTSRPAHYHVLMDENNFADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDCRHSVDRSGNILPGTVDSKICHPTEFDFYLCSHAGTQGTSRPAHYHVLMDENNFADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSDSGMASGSMARG HHDCRHSVDRSGNILPGTVDSKICHPTEFDFYLCSHAGTQGTSRPAHYHVLMDENNFSADGLQSLTNNLCTYTARCTRSVSIVPPAYYALAAFRAFYMEPETSD	1014 1021 1060 944 987 1003 868 1167 983 1008 895 896 895 896 895 896 930 869 849 849 849 962 957
1280 AtaGO1 : TTYTVVQKRHHTRLFAG BnAGO1A1 : VTFVVVQKRHHTRLFAG BnAGO1A2 : VTFVVVQKRHTRLFAG BnAGO1C2 : VTFVVVQKRHTRLFAG AtaGO2 : TTVIVQKRHTRLFAG BnAGO2A : TTVIVQKRHGTRFFAG BnAGO2A : TTVIVQKRHGTRFFAG BnAGO2A : TTVIVQKRHGTRFFAG BnAGO3C : TTVIVQKRHGTRFFAG BnAGO3C : TTVIVQKRHGTRFFAG BnAGO4C1 : FLLLVQKNHTKFFCG BnAGO4C2 : FLLLVQKNHTKFFCG BnAGO4C2 : FLLLVQKNHTKFFCG BnAGO4C2 : FLLLVQKNHTKFFCG BnAGO4C2 : FLLLVQKNHTKFFCG BnAGO4C2 : FLLLVQKNHTKFFCG BnAGO4C2 : TTVIVQKRHTRLFAG BnAGO5A : TTVIVQKRHTRLFAG BnAGO5A : TTVIVQKRHTRLFAG BNAGO5C : TTVIVQKRHTKLFGA BNAGO5C : TTVIVQKRHTKLFGA BNAGFAC : TTVIFFFFFFFFF	1320 14000 1400 14000 1400 1400 14000 14000 14000 14000 14000	1014 1021 1060 944 987 1003 868 993 1003 895 895 895 895 895 895 895 895 895 895
1280 AtaGO1 : VTFVVVQKRHHTRLFAC BhAGO1A1 : VTFVVVQKRHHTRLFAC BhAGO1A2 : VTFVVVQKRHHTRLFAC BhAGO1C2 : VTFVVVQKRHTRLFAF AtaGO2 : VTFVVVQKRHTRLFAF BhAGO2A : ITVIVAQKRHGTRFFAF BhAGO3A : ITVIVAQKRHGTRFFAF BhAGO3A : ITVIVAQKRHGTRFFAF BhAGO3C : ITVIVAQKRHGTRFFAF BhAGO3C : ITVIVAQKRHGTRFFAF BhAGO3C : ITVIVAQKRHGTRFFAF BhAGO3C : ITVIVAQKRHGTRFFAF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FTVIVAQKRHTRLFAF BhAGO5A : VTFVVQKRHTRLFAF BhAGO5A : VTFVVQKRHTRLFAF BhAGO5A : FTVIVAQKRHTKLFGA BhAGO5A : FTVIVAQKRHTKLFGA BhAGO5A : TTVIVAQKRHTKLFGA BhAGO5A : TTVIVAQKRHTKLFGA BhAGO7C1 : ITFSVQKRHTKLFFFF BhAGO7C1 : ITFSVQKRHTKLFFFF BhAGO7C1 : ITFSVQKRHTKLFFFFF BhAGO7C1 : ITFSVQKRHTKLFFFFF BhAGO7C1 : ITFSVQKRHTKLFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1320 14000 14000 14000 14000 140000 140000 140000 1400000 140000	1014 1021 1068 944 987 868 1167 983 1008 895 895 895 895 895 895 895 895 895 89
1280 AtagO1 : VTFVVVQKRHHTRLFAQ BhAGO1A1 : VTFVVVQKRHHTRLFAQ BhAGO1A2 : VTFVVVQKRHHTRLFAQ BhAGO1C2 : VTFVVVQKRHHTRLFAA AtagO2 : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO3A : ITVIVAQKRHGTRFFAA BhAGO4A : FLLIVAQKNHTKFFFCP BhAGO4A2 : FLLIVAQKNHTKFFFCP BhAGO4A2 : FLLIVAQKNHTKFFFCP BhAGO4A2 : FLLIVAQKNHTKFFFCP BhAGO4A2 : FLLIVAQKNHTKFFFCP BhAGO4A3 : VTFVVVQKNHTKLFFAA BhAGO5C : VTFVVVQKNHTKLFFAA BhAGO5C : VTFVVVQKNHTKLFFAA BhAGO5C : TTVIVAQKHTKLFFAA BhAGO5C : TTVIVAQKHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO5C : TTVIVAQKNHTKLFFAA BhAGO7C : ITFAVVQKRHTKLFFAA BhAGO7C : ITFAVVQKRHTKLFFFAA BHAGO7C : ITFAVVQKRHTKLFFAA BHAGO7C : ITFAVVQKRHTKLFFAA BHAGO	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 983 1003 868 895 896 895 896 895 895 896 895 895 896 849 975 930 849 962 954 877 823 869
1280 AtaGO1 : VTFVVVQKRHHTRLFAG BAGGO1A1 : VTFVVVQKRHHTRLFAG BAGGO1A2 : VTFVVVQKRHHTRLFAG BAGGO1C2 : VTFVVVQKRHTRLFAG AtaGO2 : VTFVVVQKRHTRLFAG BAGGO2 : ITVIVAQKRHCTRFFAG BAGGO2 : ITVIVAQKRHCTRFFAG BAGGO2 : ITVIVAQKRHCTRFFAG BAGGO3 : ITVIVAQKRHCTRFFAG BAGGO4 : FLLIVAQKNHTKFFGG BAGGO4 : FLLIVAQKNHTKFFGG BAGGO4 : FLLIVAQKNHTKFFGG BAGGO4 : FLLIVAQKNHTKFFGG BAGGO5 : VTFVVQKRHTRLFFAG BAGGO5 : VTFVVQKRHTRLFFAG BAGGO6 : FTVIVAQKHTRLFFAG BAGGO6 : FTVIVAQKHTRLFFAG BAGGO6 : FTVIVAQKHHTKFFGG BAGGO6 : FTVIVAQKHHTKLFFAG BAGGO6 : FTVIVAQKHHTKLFFAG BAGGO6 : FTVIVAQKHHTKLFFAG BAGGO7 : ITFAVVQKRHTRLFFAG BAGGO7 : ITFAVVQKRHTRLFFFAG BAGGO7 : ITFAVVQKRHTFFFFAG BAGGO7 : ITFAVVQKRHTFFFFFAG BAGGO7 : ITFAVVQKRHTFFFFFAG BAGGO7 :	1320 14000 14000 14000 14000 14000 14000 14000 14000 14000 1400	1014 1021 1068 944 1003 897 897 897 896 896 895 896 896 895 849 849 849 849 849 849 849 849 849 849
1280 AtagO1 : VTFVVVQKRHTRLFAQ BhaGO1A1 : VTFVVVQKRHTRLFAQ BhaGO1A2 : VTFVVVQKRHTRLFAQ BhaGO1A2 : VTFVVVQKRHTRLFAA AtaGO2 : VTFVVVQKRHTRLFAA AtaGO2 : ITVIVAQKRHGTRFFAA BhaGO2A : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA AtaGO3 : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA BhaGO3A : ITVIVAQKRHGTRFFAA AtaGO4 : FLLIVAQKNHTKFFGF BhaGO4A2 : FLLIVAQKNHTKFFGF BhaGO4A2 : FLLIVAQKNHTKFFGF BhaGO4A2 : VTFVVQKRHTRLFFAA BhaGO5A : VTFVVQKRHTRLFFAA BhaGO5A : TTVIVAQKNHTKFFGF BhaGO5A : TTVIVAQKNHTKFFGF BhaGO7CI : ITFSVVQKRHTRLFFAA BhaGO7CI : TTFSVQKRHTRLFFAA BhaGO7CI : TTFSVQKRHTRLFFFGF BhaGO3A : TTVIVAQKNHTKFFGFGF BhaGO3A : TTVIVAQKNHTKFFGFGF BhaGO3A : TTVIVAQKNHTKFFGFGF BhaGO3A : TTVIVAQKNHTKFFGFGF BhaGO3A : TTVIVAQKNHTKFFFGFA BhaGO7CI : TTFSVQKRHTRLFFFGFA BhaGO7CI : TTFSVQKRHTFLFFFGFA BhaGO7CI : TTFSVQKRHTFLFFFFA BhaGO7CI : TTFSVQKRHTFLFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 987 1003 868 997 895 935 935 935 935 935 936 930 850 856 896 895 849 954 849 954 849 954 849 849 849 854 849 856 849 856 856 856 856 856 856 856 856 856 856
1280 ALAGOI : UTTYUVUKKHHTRLEAG BNAGOIA: VTFVVVKKHHTRLEAG BNAGOIA: VTFVVVKKHHTRLEAG BNAGOIA: VTFVVVKKHHTRLEAG BNAGOIC: VTFVVVKKHHTRLEAG ALAGO: ITVIVAKHTRLEAG BNAGO2A: ITVIVAKHTRHEFF BNAGO2A: ITVIVAKHTRHEFF BNAGO3C: ITVIVAKHTRHEFF BNAGO3C: ITVIVAKHTRHEFF BNAGO3C: ITVIVAKHTRHEFF BNAGO4C: FLLLVAKHHTKFFC BNAGO4C: FLLLVAKHHTKFFC BNAGO4C: FLLLVAKHHTKFFC BNAGO4C: FLLLVAKHHTKFFC BNAGO4C: VTFVVVKHHTKLFFA BNAGO5: VTFVVVKHHTKLFFA BNAGO5: VTFVVVKHHTKLFFA BNAGO5: ITVIVAKHTKLFFA BNAGO5: ITVIVAKHTKLFFA BNAGO5: ITVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO5: ITTVIVAKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKLFFA BNAGO7: ITFAVVKHHTKFFFA BNAGO7: ITFAVKHHTKFFFA BNAGO7: ITFAVKHHTKFFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKHTKFFA BNAGO7: ITFAVKFA	1320 14000 14000 14000 14000 14000 14000 14000 14000 14000 1400	1014 1021 1060 944 867 1003 867 1008 897 868 895 896 895 895 895 895 895 895 895 895 895 895
1280 AtaGO1 : VTFVVVQKRHTRLFAG BTAGO1A1 : VTFVVVQKRHTRLFAG BTAGO1A2 : VTFVVVQKRHTRLFAG BTAGO1A2 : VTFVVVQKRHTRLFAG AtaGO2 : VTFVVVQKRHTRLFAG AtaGO2 : TTVIVAQKRHTRFFAG BTAGO2A : TTVIVAQKRHTRFFAG BTAGO2A : TTVIVAQKRHTRFFFAG BTAGO3A : TTVIVAQKRHTRFFFG BTAGO3A : TTVIVAQKRHTRFFFG BTAGO4A : FLLIVAQKNHTKFFFG BTAGO4A : FLLIVAQKNHTKFFFG BTAGO4A : FLLIVAQKNHTKFFFF BTAGO4A : FLLIVAQKNHTKFFFF BTAGO4A : FLLIVAQKNHTKFFFF BTAGO4A : FLLIVAQKNHTKFFFF BTAGO4A : FLLIVAQKNHTKFFFF BTAGO4A : FLLIVAQKNHTKFFFFF BTAGO4A : FLLIVAQKNHTKFFFFF BTAGO4A : FLLIVAQKNHTKFFFFF BTAGO4A : FLLIVAQKNHTKFFFFF BTAGO4A : FTVIVQKRHTRLFFFF BTAGO5A : VTFVVQKRHTRLFFFF BTAGO5A : TTVIVAQKHTKLFFFF BTAGO5A : TTVIVAQKNHTKFFFFF BTAGO5A : TTVIVAQKNHTKFFFFFF BTAGO5A : TTVIVAQKNHTKFFFFFF BTAGO5A : TTVIVAQKNHTKFFFFFF BTAGO5A : TTVIVAQKNHTKFFFFFF BTAGO5A : TTVIVAQKNHTKFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1320 1400 1400 1400 1400 1400 1400 1400 14	1014 1021 1060 944 987 1003 987 1003 889 1008 895 895 895 895 895 895 895 895 895 89
1280 AtaGO1 : VTFVVVQKRHTRLFAQ BhAGO1A1 : VTFVVVQKRHTRLFAQ BhAGO1A2 : VTFVVVQKRHTRLFAQ BhAGO1A2 : VTFVVVQKRHTRLFAA AtaGO2 : VTFVVVQKRHTRLFAA AtaGO2 : VTFVVQKRHTRLFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO2A : ITVIVAQKRHGTRFFAA BhAGO3 : ITVIVAQKRHGTRFFAA BhAGO3 : ITVIVAQKRHGTRFFAA BhAGO3 : ITVIVAQKRHGTRFFAA BhAGO4 : ILLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO4A2 : FLLIVAQKNHTKFFFFF BhAGO5A : VTFVVVQKRHTTLFAA BhAGO5A : VTFVVQKRHTTLFAA BhAGO5A : TTVIVAQKRHTTLFAA BhAGO5A : TTVIVAQKRHTKLFFFF BhAGO5A : TTVIVAQKNHTKFFFFF BhAGO5A : TTVIVAQKNHTKFFFFF BhAGO5A : TTVIVAQKRHTKLFFFF BhAGO7C1 : ITFSVVQKRHTLFAFFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFF BhAGO3A : TTVIVAQKNHTKFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1320 1407 1407 1	1014 1021 1060 944 967 1003 887 1003 887 983 1008 895 895 895 895 895 895 895 935 935 935 935 895 895 935 935 849 968 855 855 855 856 855 856 855 855 855 8
1280 AtaGO1 : VTFVVVQKRHHTRLFAG BnAGO1A1 : VTFVVVQKRHHTRLFAG BnAGO1A2 : VTFVVVQKRHTRLFAG BnAGO1C2 : VTFVVVQKRHTRLFAG AtaGO2 : ITVVIQKRHTRLFAG AtaGO2 : ITVVIQKRHTRLFAG AtaGO3 : ITVVIQKRHTRFFAG BnAGO2A : ITVVIQKRHTRFFAG BnAGO3C : ITVVIQKRHTRFFFAG BnAGO3C : ITVVIQKRHTRFFFG BnAGO4A1 : FLLIVAQKNHTKFFFG BnAGO4A2 : FLLIVAQKNHTKFFFG BnAGO5A : VTFVVQKRHTRLFAFA BnAGO5C : TVTVIQKRHTKFFFG BNAGO5C : TTVIVAQKNHTKFFFG BNAGO6C : FTVIVAQKNHTKFFFG BNAGO6C : FTVIVAQKNHTKFFFG BNAGO6C : FTVIVAQKNHTKFFFG BNAGO6C : TTFAVQKRHTKFFFG BNAGO7A : TTFVIVAQKNHTKFFFG BNAGO7A : TTVIIAQKNHTKFFFG BNAGO7A : TTV	1320 1400 HNDRHSVDRSGNILPGTVDSKICHPTEPTYLCSHAGIGGTSRPAHHVLWDENNFADGLGSLTNNLCYTYARCTRSVSIVPAYYA LAAFRAFYMEPTEDSGNAGGGMARG HNDRHSVDRSGNILPGTVDSKICHPTEPTYLCSHAGIGGTSRPAHHVLWDENNFADGLGSLTNNLCYTYARCTRSVSIVPAYYA LAAFRAFYMEPTEDSGNAGGMARG HNDRNSVDRSGNILPGTVDSKICHPTEPTYLCSHAGIGGTSRPAHHVLWDENNFADGLGSLTNNLCYTYARCTRSVSIVPAYYA LAAFRAFYMEPTEDSGNAGGMARG HNDRNSVDRSGNILPGTVDSKICHPTEPTYLCSHAGIGGTSRPAHHVLWDENNFADGLGSLTNNLCYTYARCTRSVSIVPAYYA LAAFRAFYMEPTED	1014 1021 1060 944 987 1003 987 1003 888 1167 888 895 895 895 895 895 895 895 895 895
1280 AtaGG1 : VTFVVVQKRHTRLFAG BhAGG1A: VTFVVVQKRHTRLFAG BhAGG1A: VTFVVVQKRHTRLFAH BhAGG1C: VTFVVVQKRHTRLFAH AtaGG2 : VTFVVQKRHTRLFAH AtaGG2 : VTFVVQKRHTRLFAH AtaGG2 : ITVIVAQKRHGTRFFAD BhAGG2A : ITVIVAQKRHGTRFFAD BhAGG3 : ITVIVAQKRHGTRFFAD BhAGG3 : ITVIVAQKRHGTRFFAD BhAGG3 : ITVIVAQKRHGTRFFAD BhAGG4A : FLLIVAQKNHTKFFGD BhAGG4A : FLLIVAQKNHTKFFGT BhAGG4A : FLLIVAQKNHTKFFFG BhAGG4A : FLLIVAQKNHTKFFFG BhAGG4A : FLLIVAQKNHTKFFFG BhAGG4A : FLLIVAQKNHTKFFFG BhAGG4A : FLLIVAQKNHTKFFFG BhAGG5A : VTFVVQKRHTRLFAD BhAGG5A : TTVIVAQKRHTKLFAD BhAGG6 : FTVIVAQKNHTKFFFG BhAGG6 : FTVIVAQKNHTKFFFG BhAGG6C : FTVIVAQKNHTKFFFG BhAGG6C : FTVIVAQKNHTKFFFG BhAGG6C : FTVIVAQKNHTKFFFG BhAGG7C : ITFAVQKRHTKLFAD BhAGG6C : FTVIVAQKNHTKFFFG BhAGG6C : FTVIVAQKNHTKFFFG BhAGG7C : ITFAVQKRHTKLFFG BhAGG7C : ITFAVQKRHTKLFFG BhAGG7C : ITFAVQKRHTKLFFG BhAGG7C : TTVIVAQKNHTKFFFG BhAGG7C : TTVIVAQKNHTKFFFG BhAGG7C : TTVIVAQKNHTKFFFG BhAGG7C : TTFAVQKRHTKLFFG BhAGG7C : TTFAVQKRHTKFFFG BhAGG7C : TTVIVAQKNHTKFFFG BhAGG7C : TTVIVAQKNHTKFFFFG BhAGG7C : TTVIVAQKNHTKFFFFG BhAGG7C : TTVIVAQKNHTKFFFFG BhAGG7C : TTFAVQKRHTKFFFFG BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTFVVQKRHTKFFFFF BhAGG7C : TTVVQKRHTFFFFFF BhAGG7C : TTVVQKRHTFFFFFFF BhAGG7C : TTVVQKRHTFFFFFFFF BhAGG7C : TTVVQKRHTFFFFFFF BhAGG7C : TTVVQKRHTFFFFFFF BhAGG7C : TTVVQKRHTFFFFFFFFFFFF BhAGG7C : TTVVQKRHTFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	1320 1400 1400 1400 1400 1400 1400 1500 150	1014 1021 1060 944 987 1003 887 1003 895 895 895 895 895 895 895 895 895 895
1280 ALAGO1 : 'UTFVVVQKRHITRLFAG BNAGO1A2 : VTFVVVQKRHITRLFAG BNAGO1A2 : VTFVVVQKRHITRLFAG BNAGO1C2 : VTFVVVQKRHITRLFAH ALAGO2 : ITVIVAQKRHITRLFAH BNAGO2A : ITVIVAQKRHITRFFA BNAGO2A : ITVIVAQKRHITRFFA BNAGO3C : ITVIVAQKRHITRFFA ALAGO3 : ITVIVAQKRHITRFFA BNAGO3C : ITVIVAQKRHITRFFA BNAGO3C : ITVIVAQKRHITRFFA BNAGO4A1 : FLLLVAQKNHITKFFOP BNAGO4A2 : FLLVAQKNHITKFFOP BNAGO4A2 : FLLVAQKNHITKFFOP BNAGO5A : VTFVVQKRHITRLFFA BNAGO5C : TITVIVAQKNHITKFFOP BNAGO5A : TITVIVAQKNHITKFFOP BNAGO5A : TITVIVAQKNHITKFFOP BNAGO5A : TITVIVAQKNHITKFFOP BNAGO5A : TITVIVAQKNHITKFFOP BNAGO7A : ITFAVVQKRHITKLFFA BNAGO72 : ITFAVVQKRHITKLFFA BNAGO72 : ITFAVVQKRHITKLFFA BNAGO72 : ITTFAVQKRHITKLFFA BNAGO72 : TITVIAQKNHITKFFOP BNAGO73 : FITVIVAQKNHITKFFOP BNAGO74 : FITVIVAQKNHITKFFOP BNAGO74 : FITVIVAQKNHITKFFOP BNAGO74 : FITVIVAQKNHITKFFFOP BNAGO74 : FITVIAQKNHITKFFFOP BNAGO74 : F	1320 1400 1400 1500 1500 1400 1500 1500 150	1014 1001 1018 944 987 1003 987 1003 868 1167 895 983 1008 889 889 889 889 889 869 955 955 860 8649 968 9968 9968 9968 9968 9968 9968

FIGURE 3 | Alignment of PIWI domains of *B. napus* and *A. thaliana* AGO proteins. The protein sequences were aligned using MEGA 5.0. The conserved Asp, Asp and His (DDH) triad residues as well as His (H) corresponding to H798 of Arabidopsis AGO1 are shaded black. Amino acid positions corresponding to each protein are indicated at the end of each line.

DFD in the characteristic catalytic motif, which requires further experimental confirmation (Figure S1B). These data indicated that BnRDR1, BnRDR2, and BnRDR6 proteins belong to RDR $\alpha$  class, while the BnRDR3 and BnRDR5 proteins are members of RDR $\gamma$  class. This is similar to what has been observed for AtRDRs. Additionally, the length of BnRDR proteins varied from 861 to 1198 aa (**Table 1**). Different groups of RDRs exhibited diverse pI-values; the pI-value of BnRDR3s and BnRDR4 was higher than 8.0, while that of the majority of the other BnRDRs was lower than 7.0 (**Table 1**).

Comparison of the protein pI-value indicated that AGOs are obviously basic proteins with a pI higher than 8.8, DCLs are generally acidic proteins with a pI lower than 6.8 except BnDCL2C, while RDRs may be acidic or basic with a pI ranging from 5.7 to 8.6 (**Table 1**).

# Exon-Intron Organization of *B. napus DCL*, *AGO*, and *RDR* Genes

The exon-intron structure of DCL, AGO, and RDR genes was examined to gain more insights into their possible structural evolution. Our results for all three gene families showed that intron number was generally conserved in members of the same groups while varied significantly in different groups of the same family. For BnDCL genes, the intron number varied from 17 to 24 and BnDCL3 and BnDCL4 groups contained 3~7 more introns than the remaining two groups (Table 1 and Figure S2A). In the case of BnAGOs, considerable variations in intron number were observed. BnAGO groups 1, 4, 5, 6, 8, 9, and 10 comprised similar number of introns ranging from 16 to 23, while BnAGO groups 2, 3, and 7 possessed only 1~5 introns, which were dramatically less than other *BnAGO* groups (Table 1 and Figure S2B). Similar to BnAGOs, remarkable intron number variation also occurred in BnRDRs. BnRDR1, BnRDR2, and BnRDR6 groups contained as few as  $1 \sim 3$  introns, while BnRDR3, BnRDR4, and BnRDR5 groups carried as many as 16~19 introns (Table 1 and Figure S2C). These group-dependent exon-intron structures were conserved for genes from both B. napus and Arabidopsis. The observation that exon-intron structure of DCL, AGO, and RDR genes was highly similar within members of the same groups but significantly divergent across different groups of the same family suggests that these gene families especially AGO and RDR families have undergone frequent gene duplication and recombination throughout evolution.

### Prediction of *Cis*-Acting Elements in Promoter of *B. napus DCL*, *AGO*, and *RDR* Genes

The 1.5 kb sequences upstream of the translation initiation codon of *BnDCL*, *BnAGO*, and *BnRDR* genes were retrieved and searched for *cis*-acting elements using PLACE database. This revealed the presence of various *cis*-acting elements related to phytohormone response, abiotic stress and defense response in *BnDCL*, *BnAGO*, and *BnRDR* genes (Table S4). All of these genes contained at least three *cis*-acting elements directly related to defense response, including ASF1MOTIFCAMV (S000024), GT1CONSENSUS (S000198), SEBFCONSSTPR10A

(S000391), WBOXATNPR1 (S000390), and WRKY71OS (S000447). Interestingly, 21 out of these 51 genes possessed a 6-bp CGCG element (A/C/G) CGCG (C/G/T; **Figure 4** and Table S4), with *BnAGO2A*, *BnAGO5C*, *BnAGO6A*, and *BnAGO8A* containing two copies of these *cis*-elements. It has been revealed that calmodulin-binding transcription activators (CAMTAs) contribute to plant defense responses by binding to CGCG *cis*-elements of the target gene promoters and thereby regulating their expression (Yang and Poovaiah, 2002; Du et al., 2009; Nie et al., 2012). Thus, intriguingly, our result predicts that RNA silencing might be regulated by CAMTAs.

### Expression Analyses Implied That CAMTA3 Might Mediate Regulation of *B. napus DCL*, *AGO*, and *RDR* Gene Expression in Response to Pathogen Inoculation

Owing to the high cDNA sequence identity in the same gene family, design of gene-specific primers for many of *B. napus DCL*, AGO, and RDR genes is unfeasible. Therefore, we chose all gene members of BnAGO4, BnRDR1, and BnDCL1 groups, for which ideal gene-specific primers could be designed, as representative to examine the expression pattern of these predicted genes in B. napus and their response to S. sclerotiorum inoculation. The analysis demonstrated that these genes were differentially expressed in leaves under normal growth conditions (Figure 5A and Figure S3). Six of them including four BnAGO4s, BnRDR1A, and BnDCL1A, exhibited moderate to high level expression in normal leaves, while the expression of BnRDR1C1, BnRDR1C2, and BnDCL1C was only detected when a second round PCR was conducted using the products of the first round PCR as templates (Figure 5A and Figure S3). These data demonstrated the difference in constitutive expression of these nine BnAGO4, BnRDR1, and BnDCL1 genes in B. napus leaves. Real-time quantitative PCR (RT-qPCR) was further used to gain insight into expression patterns of these genes in response to S. sclerotiorum inoculation at 0, 8, and 16 hpi. Interestingly, all of these genes, except BnRDR1A, BnRDR1C1, and BnDCL1A, were downregulated to various degrees by S. sclerotiorum inoculation. Notably, 4 AGO4s were all markedly reduced by 2.2~10.2 folds at 16 hpi (Figure 5B). This result suggested the possible involvement of these genes in the interactions between B. napus and S. sclerotiorum.

All these genes with the exception of 2 *BnRDR1* (*BnRDR1A* and *BnRDR1C1*), carried CGCG elements in the DNA sequences upstream of their coding regions (**Figure 4**) and we conjectured that the lowered expression of these genes might have resulted from up-regulation of *CAMTA* genes in response to *S. sclerotiorum* inoculation. To test this hypothesis, we examined the expression of *BnCAMTA3* genes after inoculation with this pathogen. *BnCAMTA3* genes were selected since *AtCAMTA3* has been reported to function in regulating plant defense (Du et al., 2009; Nie et al., 2012; Rahman et al., 2016). Four *BnCAMTA3* genes were identified through BLASTp searches against *B. napus* genome database using AtCAMTA3 as query. RT-qPCR assays demonstrated that expression of two out of four *BnCAMTA3* genes (*BnCAMTA3a* and *BnCAMTA3b*) increased





strongly by 14.7 and 10.4 folds, respectively, at 8 h after *S. sclerotiorum* inoculation (**Figure 5C**). Given that CAMTA3 negatively regulates plant disease resistance through direct binding and subsequent repression of target defense-related genes (Du et al., 2009; Nie et al., 2012), our results indicate that RNA silencing might be regulated by CAMTAs during *B. napus–S. sclerotiorum* interactions.

## Arabidopsis *dcl*, *ago*, and *rdr* Mutants Commonly Exhibited Altered Susceptibility to *S. sclerotiorum*

The observation in this study that expression of *BnDCL1*, *BnAGO4*, and *BnRDR1* genes altered significantly in response to *S. sclerotiorum* inoculation and our previous findings that

Arabidopsis *ago1* and *ago2* mutant plants exhibit enhanced susceptibility to *S. sclerotiorum* (Cao et al., 2016) prompted us to assess the possible role of DCL, AGO, and RDR-mediated RNA silencing in plant resistance to this pathogen. A total of 13 *A. thaliana dcl, ago,* and *rdr* mutants were tested. They included three *dcl,* seven *ago,* and three *rdr* mutants, all in Col-0 background except *dcl1-9* and *ago4-1* which are derived from Ler ecotype. Based on the phenotypes after *S. sclerotiorum* inoculation, the mutants could be divided into three classes. Class I comprised those exhibiting enhanced susceptibility, Class II contained mutants displaying enhanced resistance, while Class III included mutants showing similar disease phenotypes to wild-type plants (**Figure 6**). The mutants *dcl4-2, ago9-1, rdr1-1, rdr6-11,* and *rdr6-15* were more susceptible to *S. sclerotiorum* challenge by showing more severe necrosis and larger size of disease lesions



than wild-type plants (**Figure 6**). By contrast, the mutant *dcl1-9* was found to be more resistant since it displayed less necrosis and smaller size of disease lesions when compared with Ler plants upon *S. sclerotiorum* inoculation (**Figure 6**). However, the *dcl2-1*, *ago3-1*, *ago4-1*, *ago5-1*, *ago6-1*, *ago7-1*, and *ago10-3* mutant plants exhibited similar disease susceptibility as the wild-type plants with respect to severity of necrosis and size of disease lesions (**Figure 6**). These data demonstrate the involvement of DCL, AGO, and RDR-mediated RNA silencing in plant resistance to the necrotrophic fungal pathogen *S. sclerotiorum*.

### DISCUSSION

RNA silencing is a versatile molecular mechanism to regulate diverse biological processes through altering transcript accumulation of genes essential to these processes. In this study, we identified three gene families encoding key components of RNA silencing machinery in *B. napus* and its progenitors *B. rapa* and *B. oleracea*. We demonstrated the important role of RNA silencing in plant resistance to the fungal pathogen *S. sclerotiorum* and indicated a possible new molecular mechanism to regulate RNA silencing in response to this pathogen. Our results provide insights into composition, function, and mechanism of RNA silencing in plants.

### **RNA Silencing Machinery in** *B. napus*

DCL, AGO, and RDR are key components of plant RNA silencing machinery. Our results demonstrate that B. napus possesses 8 BnDCL, 27 BnAGO, and 16 BnRDR genes. This gene copy numbers are much larger than those in A. thaliana although the two species belong to the same family (Brassicaceae). However, considering that B. napus is tetraploid while A. thaliana is diploid, it is clear that each haploidy in the two species carries the same number (4) of DCL genes but different number of AGO and RDR genes. The haploid B. napus contains more AGO and RDR genes than A. thaliana. B. napus is an allotetraploid from crossing between B. oleracea and B. rapa, followed by chromosome doubling (Chalhoub et al., 2014). Our results demonstrate that B. napus genome includes almost all copies of these RNA silencing machinery genes from its progenitor species B. rapa and B. oleracea. In addition, B. napus genome contains three extra copies of RDR5 genes, indicating that the RDR5 group in *B. napus* appears to have undergone further expansion through duplication during evolution. Moreover, compared with A. thaliana, some AGO and RDR genes such as AGO1, AGO4, AGO9, and RDR5 have significantly expanded in these Brassica species. The B. napus subgenomes A and C bear more RDR5 genes but similar AGO genes compared with genomes of its progenitor species B. rapa and B. oleracea, suggesting that the expansion of RDR5 genes is likely to occur after B. napus formation while that of AGO1, AGO4, and AGO9 may have occurred before B. napus formation. Additionally, members of AGO7, AGO8, and AGO9 as well as RDR4 are unevenly distributed in the A and C subgenomes of B. napus. This asymmetric distribution may have resulted from homeologous exchanges, which is also most likely the source of further expansion of these two families and it is consistent with the previous report (Chalhoub et al., 2014). Considering that *B. napus* genome contains at least two copies of each *DCL*, *AGO*, and *RDR* genes except *AGO8*, and that all four members of BnAGO4 are expressed constitutively and/or in response to pathogen inoculation (**Figure 5**) and are thus most probably functional, it is interesting to understand whether these genes functions redundantly or differentially.

Variation in composition of RNA machinery, especially composition of different gene groups (DCL1, DCL2, DCL3, and DCL4 for DCL; AGO2/3/7, AGO1/5/10, and AGO4/6/8/9 for AGO as well as RDR1/2/6 and RDR3/4/5 for RDR) seems to widely exist in plants. For example, the tomato genome contains four DCL2 but only a single ortholog of the other DCLs (Bai et al., 2012). Similarly, tomato AGO1/5/10 group consists of two AGO1 and AGO10 each, AGO2/3/7 group comprises two AGO2, AGO4/6/8/9 group is constituted of four AGO4s, and a new AGO member, but lacks AGO8 and AGO9 (Bai et al., 2012). Another species of the Solanaceae family, N. benthamiana, has identical number of AGO1/5/10 group genes, a similar number of AGO4/6/8/9 group genes but different number of AGO2/3/7 group genes when compared with tomato. The N. benthamiana AGO4/6/8/9 group bears two AGO4s, while the AGO2/3/7 group lacks AGO3 (Nakasugi et al., 2013). In addition, grape AGO1/5/10 group consists of two AGO10s, while AGO2/3/7 group includes two AGO2s (Zhao et al., 2015). Regarding RDR, tomato RDR1/2/6 group consists of two RDR6s, while this group in grape contains two RDR1s. The N. benthamiana genome apparently does not carry any RDR3/4/5 group genes, while tomato and grape RDR3/4/5 group both lack RDR4 and RDR5 but contain two and one RDR3, respectively (Bai et al., 2012; Nakasugi et al., 2013; Zhao et al., 2015).

## Function of RNA Silencing in Plant Resistance to Fungal Pathogens

It is well-known that RNA silencing can protect plants from viral infection (Ding and Voinnet, 2007; Llave, 2010; Incarbone and Dunoyer, 2013; Wu et al., 2015). This antiviral immunity involves production of virus-derived small interfering RNAs (viRNAs) and results in specific silencing of viruses by viRNAguided effector complexes. Apart from defense against viruses, RNA silencing also plays a role in plant defense against bacterial pathogens (Katiyar-Agarwal et al., 2006; Navarro et al., 2006; Voinnet, 2008; Robert-Seilaniantz et al., 2011). Similar to viruses, bacteria have also developed mechanisms to suppress RNA silencing in order to infect successfully (Navarro et al., 2008). More recently, functions of RNA silencing in plant resistance to fungal pathogens are being revealed. Plant miRNAs are differentially expressed in response to inoculation with fungal pathogens, such as Erysiphe graminis (Xin et al., 2010), Fusarium virguliforme (Radwan et al., 2011), V. dahliae (Yin et al., 2012; Yang et al., 2013), V. longisporum (Shen et al., 2014), M. oryzae (Li et al., 2014), and B. cinerea (Jin and Wu, 2015). More importantly, mutants of key components of the RNA silencing machinery exhibit altered susceptibility to fungal pathogens including two species of Verticillium (Ellendorff et al., 2009; Shen et al., 2014). Moreover, it has been reported that B. cinerea

produces small RNAs to suppress plant defense by hijacking host RNA interference pathways (Weiberg et al., 2013). It is interesting to elucidate the function and mechanism of RNA silencing in the interactions between the important oil crop B. napus and the devastating fungal pathogen S. sclerotiorum. Previously, we have identified the miRNAs involved in this plant-pathogen interaction, many of which targeting genes involved in plant defense. Besides, three miRNAs (ath-miR168a\_1ss21AC, alymiR403a-3p\_L+1, and bna-miR403) target AGO1 and AGO2 which are two key components of RNA silencing. We further found that Arabidopsis ago1-27, ago1-33, and ago2-1 mutant plants exhibit enhanced susceptibility to S. sclerotiorum (Cao et al., 2016), thus providing a clue to the important roles of RNA silencing in the interactions between B. napus and S. sclerotiorum. In this study, we identified gene families encoding DCL, AGO, and RDR, three key components of RNA silencing in B. napus. We found that these genes, represented by all members of three groups of genes (4 BnAGO4, 3 BnRDR1, and 2 BnDCL1 genes), are differentially expressed in response to S. sclerotiorum inoculation (Figure 5B). Our results showed that the expression divergence was present among members belonging to the same gene group after S. sclerotiorum inoculation (Figure 5B), which suggested that gene members within the same groups may have diverse functions in this plant-pathogen interaction. AtAGO4 is one of the critical components in the transcriptional genesilencing pathway associated with siRNA that directs DNA methylation (Zilberman et al., 2004; Qi et al., 2006) and is required for resistance to P. syringae (Agorio and Vera, 2007). AtRDR1 is elicited by salicylic acid (SA) treatment and viral infection (Yu et al., 2003; Diaz-Pendon et al., 2007; Qi et al., 2009). Together, these results suggest that AGO4 and RDR1 may also contribute to B. napus defense against S. sclerotiorum. Furthermore, we assessed susceptibility toward S. sclerotiorum in 13 dcl, ago, and rdr A. thaliana mutants. As many as six dcl, ago, and rdr mutants exhibit altered susceptibility (Figure 6). These mutants include dcl4-2, ago9-1, rdr1-1, rdr6-11, rdr6-15, and dcl1-9. The mutated genes encode different RNA-silencing components. This is similar to what was found in the study involving another fungal pathogen V. dahliae (Ellendorff et al., 2009). Collectively, these results indicate that RNA silencing may be involved in interactions between plants and fungal pathogens. However, how these DCLs, AGOs, and RDRs regulate resistance to S. sclerotiorum is unclear. Recently, it has been reported that rice AGO18 functions through regulating AGO1 accumulation by direct binding with the AGO1-targeted miRNA (miR168; Wu et al., 2015). Whether a similar mechanism exists for AGOs or even RDRs and DCLs in Arabidopsis and B. napus is worthy of further study.

### **Regulation of RNA Silencing Machinery**

A remaining challenge for RNA silencing is to dissect the mechanism through which the RNA silencing pathway itself can be regulated. Up to now, little of this mechanism has been uncovered except the evidence that the expression of components of the RNA silencing pathway is subject to negative feedback regulation by their own miRNA products. For example, miR162 targets DCL1, miR168 targets AGO1, and miR403

targets AGO2. However, there is no more information regarding regulatory mechanisms for the other RNA silencing components. Considering the importance of RNA silencing pathway, we are curious to know whether other mechanisms exist that confer additional layers of regulation on the RNA silencing machinery. Intriguingly, in present study, we found that substantial number of B. napus DCL, AGO, and RDR genes (21 out of a total of 51) possessed CAMTA/SR binding sites [(A/C/G) CGCG (C/G/T)] in their promoter sequences (Figure 4 and Table S4). CAMTAs, especially CAMTA3, contribute to plant defense responses by direct binding and thereby regulating the expression of the target genes (Yang and Poovaiah, 2002; Du et al., 2009; Nie et al., 2012). Therefore, we suspect that the expression of these CGCG-element-containing RNA silencing components may be regulated by CAMTAs. To confirm this possibility, we examined the expression of CAMTA genes and these CGCGelement-containing RNA silencing genes in B. napus in response to S. sclerotiorum inoculation. The results of this expression analysis indicate that S. sclerotiorum inoculation strongly induced the expression of BnCAMTA3 genes while it significantly suppressed that of many CGCG-element-containing BnAGO, BnDCL and BnRDR genes (Figure 5). Moreover, another work in our laboratory has revealed that Atcamta3 mutant plants exhibit enhanced resistance to S. sclerotiorum (Rahman et al., unpublished data). Taken together, our results suggest that RNA silencing might be regulated by CAMTA3. Nevertheless, further confirmation of CAMTA binding activity with CGCGelement-containing RNA silencing genes by other assays such as EMSA and its effect on expression of these genes will provide more straightforward evidence to support this intriguing likely mechanism for regulation of RNA silencing machinery.

It is well-known that the functions of CAMTAs are dependent on their interaction with Ca<sup>2+</sup>/CaM (Choi et al., 2005; Du et al., 2009) and these genes, especially CAMTA3, are widely involved in plant defense (Yang and Poovaiah, 2002; Du et al., 2009; Nie et al., 2012; Rahman et al., 2016). For instance, knockout of AtCAMTA3 leads to increased accumulation of salicylic acid and enhanced host disease resistance to both bacterial (Du et al., 2009) and fungal pathogens (Nie et al., 2012) and nonhost resistance to bacterial pathogen (Rahman et al., 2016) but reduced resistance against insect herbivores (Qiu et al., 2012). Similarly, one rice CAMTA mutant (oscbt-1) exhibits enhanced resistance to blast fungal pathogen and leaf blight bacterial pathogen (Koo et al., 2009). Given that expression of some RNA silencing machinery may be regulated by CAMATs and both of CAMTA as well as RNA silencing contribute to plant resistance, it will be intriguing to explore whether the role of CAMTAs in regulating plant defense response can be attributed to its regulation of RNA silencing and thus opening a possibility to link Ca<sup>2+</sup> signaling and RNA silencing together to provide a novel facet of Ca<sup>2+</sup> signaling in regulation of plant disease resistance.

## CONCLUSIONS

The B. napus genome possessed 8 Dicer-like (DCL), 27 Argonaute and 16 RNA-dependent RNA polymerase (RDR) genes. The B.

*napus* genome expanded *RDR5* genes compared with that of its progenitors *B. oleracea* and *B. rapa* and all genomes of three *Brassica* species expanded *AGO1*, *AGO4*, and *AGO9* genes compared with the Arabidopsis genome. *B. napus DCL*, *AGO*, and *RDR* genes widely (21 out of 51) harbored a CAMTA-binding site (CGCG box) in their promoter regions. Inoculation with *S. sclerotiorum* strongly induced the expression of *BnCAMTA3* genes while significantly reduced that of many CGCG-containing RNA silencing component genes. Our results suggested that RNA silencing machinery might be targeted by CAMTA3. Furthermore, mutant analyses demonstrated that six out of 13 *dcl, ago*, and *rdr* mutants exhibited altered susceptibility to *S. sclerotiorum*. These results indicate the important role of RNA silencing in plant resistance to this devastating fungal pathogen.

## **AUTHOR NOTE**

During the review process of this manuscript, a paper on bioinformatics identification of *DCL*, *AGO*, and *RDR* families in *Brassica napus* was accepted for publication (Zhao et al., 2016).

## **AUTHOR CONTRIBUTIONS**

The project was coordinated by X-ZC. J-YC, and Y-PX conducted the bioinformatics and phylogenetic analyses. J-YC, Y-PX, and

### REFERENCES

- Agorio, A., and Vera, P. (2007). ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in Arabidopsis. *Plant Cell* 19, 3778–3790. doi: 10.1105/ tpc.107.054494
- Bai, M., Yang, G. S., Chen, W. T., Mao, Z. C., Kang, H. X., Chen, G. H., et al. (2012). Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene* 501, 52–62. doi: 10.1016/j.gene.2012.02.009
- Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363. doi: 10.1038/ nature02874
- Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., et al. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* 14, 1023–1031. doi: 10.1002/elps.11501401163
- Bologna, N. G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu. Rev. Plant Biol. 65, 473–503. doi: 10.1146/annurev-arplant-050213-035728
- Cao, J. Y., Xu, Y. P., Zhao, L., Li, S. S., and Cai, X. Z. (2016). Tight regulation of the interaction between *Brassica napus* and *Sclerotinia sclerotiorum* at the microRNA level. *Plant Mol. Biol.* 92, 39–55. doi: 10.1007/s11103-016-0494-3
- Carmell, M. A., and Hannon, G. J. (2004). RNase III enzymes and the initiation of gene silencing. Nat. Struct. Mol. Biol. 11, 214–218. doi: 10.1038/nsmb729
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435
- Choi, M. S., Kim, M. C., Yoo, J. H., Moon, B. C., Koo, S. C., Park, B. O., et al. (2005). Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). J. Biol. Chem. 280, 40820–40831. doi: 10.1074/jbc.M504616200
- Csorba, T., Kontra, L., and Burgyán, J. (2015). Viral silencing suppressors: tools forged to fine-tune host-pathogen coexistence. *Virology* 479, 85–103. doi: 10.1016/j.virol.2015.02.028

HR carried out the gene expression assays. J-YC, WL, and S-SL performed disease resistance evaluation analyses. J-YC designed and performed the statistical analysis. X-ZC conceived of the study, and participated in its design and coordination. X-ZC and J-YC prepared the manuscript. All authors read and approved the final manuscript.

### ACKNOWLEDGMENTS

We are grateful to Prof. Shou-Wei Ding (Department of Plant Pathology and Microbiology, University of California, Riverside, USA) and Prof. Yi-Jun Qi (Tsinghua-Peking Center for Life Sciences, and School of Life Sciences, Tsinghua University, China) for providing the *A. thaliana* mutants. This work was financially supported by grants from the National Natural Science Foundation of China (nos. 31371892, 31672014), the Special Fund for Agro-scientific Research in the Public Interest (no. 201103016) and the SRFDP (no. 201101110092).

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016. 01614/full#supplementary-material

- Deleris, A., Gallego-Bartolome, J., Bao, J., Kasschau, K. D., Carrington, J. C., and Voinnet, O. (2006). Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313, 68–71. doi: 10.1126/science.1128214
- Diaz-Pendon, J. A., Li, F., Li, W. X., and Ding, S. W. (2007). Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* 19, 2053–2063. doi: 10.1105/tpc.106.047449
- Ding, S. W., and Voinnet, O. (2007). Antiviral immunity directed by small RNAs. Cell 130, 413–426. doi: 10.1016/j.cell.2007.07.039
- Du, L., Ali, G. S., Simons, K. A., Hou, J., Yang, T., Reddy, A. S. N., et al. (2009). Ca<sup>2+</sup>/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature* 457, 1154–1158. doi: 10.1038/nature07612
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/ gkh340
- Ellendorff, U., Fradin, E. F., de Jonge, R., and Thomma, B. P. H. J. (2009). RNA silencing is required for Arabidopsis defence against Verticillium wilt disease. *J. Exp. Bot.* 60, 591–602. doi: 10.1093/jxb/ern306
- Hannon, G. J. (2002). RNA interference. *Nature* 418, 244–251. doi: 10.1038/418 244a
- Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* 27, 297–300. doi: 10.1093/nar/27.1.297
- Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol. 9, 22–32. doi: 10.1038/nrm2321
- Incarbone, M., and Dunoyer, P. (2013). RNA silencing and its suppression: novel insights from in planta analyses. *Trends Plant Sci.* 18, 382–392. doi: 10.1016/j.tplants.2013.04.001
- Jaubert, M., Bhattacharjee, S., Mello, A. F. S., Perry, K. L., and Moffett, P. (2011). ARGONAUTE2 mediates RNA-silencing antiviral defenses against *Potato virus* X in Arabidopsis. *Plant Physiol*. 156, 1556–1564. doi: 10.1104/pp.111.178012
- Jin, W., and Wu, F. (2015). Characterization of miRNAs associated with *Botrytis* cinerea infection of tomato leaves. *BMC Plant Biol*.15:1. doi: 10.1186/s12870-014-0410-4

- Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., et al. (2008). Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics* 9:451. doi: 10.1186/1471-2164-9-451
- Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A. Jr., Zhu, J. K., et al. (2006). A pathogen-inducible endogenous siRNA in plant immunity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18002–18007. doi: 10.1073/pnas.0608258103
- Koo, S. C., Choi, M. S., Chun, H. J., Shin, D. B., Park, B. S., Kim, Y. H., et al. (2009). The calmodulin-binding transcription factor OsCBT suppresses defense responses to pathogens in rice. *Mol. Cells* 27, 563–570. doi: 10.1007/s10059-009-0081-4
- Li, Y., Lu, Y. G., Shi, Y., Wu, L., Xu, Y. J., Huang, F., et al. (2014). Multiple rice microRNAs are involved in immunity against the blast fungus *Magnaporthe* oryzae. Plant Physiol. 164, 1077–1092. doi: 10.1104/pp.113.230052
- Liu, Q., Feng, Y., and Zhu, Z. (2009). Dicer-like (DCL) proteins in plants. *Funct. Integr. Genomics* 9, 277–286. doi: 10.1007/s10142-009-0111-5
- Llave, C. (2010). Virus-derived small interfering RNAs at the core of plant-virus interactions. *Trends Plant Sci.* 15, 701–707. doi: 10.1016/j.tplants.2010.09.001
- Lynn, K., Fernandez, A., Aida, M., Sedbrook, J., Tasaka, M., Masson, P., et al. (1999). The PINHEAD/ZWILLE gene acts pleiotropically in Arabidopsis development and has overlapping functions with the ARGONAUTE1 gene. *Development* 126, 469–481.
- Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J., et al. (2006). The evolution and diversification of Dicers in plants. *FEBS Lett.* 580, 2442–2450. doi: 10.1016/j.febslet.2006.03.072
- Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. *Nature* 457, 413–420. doi: 10.1038/nature07756
- Moissiard, G., Parizotto, E. A., Himber, C., and Voinnet, O. (2007). Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. RNA 13, 1268–1278. doi: 10.1261/rna.541307
- Moissiard, G., and Voinnet, O. (2006). RNA silencing of host transcripts by cauliflower mosaic virus requires coordinated action of the four Arabidopsis Dicer-like proteins. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19593–19598. doi: 10.1073/pnas.0604627103
- Nakasugi, K., Crowhurst, R. N., Bally, J., Wood, C. C., Hellens, R. P., and Waterhouse, P. M. (2013). *De novo* transcriptome sequence assembly and analysis of RNA silencing genes of *Nicotiana benthamiana*. *PLoS ONE* 8:e59534. doi: 10.1371/journal.pone.0059534
- Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., et al. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312, 436–439. doi: 10.1126/science.1126088
- Navarro, L., Jay, F., Nomura, K., He, S. Y., and Voinnet, O. (2008). Suppression of the microRNA pathway by bacterial effector proteins. *Science* 321, 964–967. doi: 10.1126/science.1159505
- Nie, H., Zhao, C., Wu, G., Wu, Y., Chen, Y., and Tang, D. (2012). SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating NDR1 and EIN3. *Plant Physiol.* 158, 1847–1859. doi: 10.1104/pp.111.192310
- Parent, J. S., Bouteiller, N., Elmayan, T., and Vaucheret, H. (2015). Respective contributions of Arabidopsis DCL2 and DCL4 to RNA silencing. *Plant J.* 81, 223–232. doi: 10.1111/tpj.12720
- Qi, X., Bao, F. S., and Xie, Z. (2009). Small RNA deep sequencing reveals role for *Arabidopsis thaliana* RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS ONE* 4:e4971. doi: 10.1371/journal.pone.0004971
- Qi, Y., He, X., Wang, X. J., Kohany, O., Jurka, J., and Hannon, G. J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* 443, 1008–1012. doi: 10.1038/nature05198
- Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., and Cheng, B. (2011). Identification and characterization of Dicer-like, Argonaute and RNAdependent RNA polymerase gene families in maize. *Plant Cell Rep.* 30, 1347–1363. doi: 10.1007/s00299-011-1046-6
- Qiu, Y., Xi, J., Du, L., Suttle, J. C., and Poovaiah, B. W. (2012). Coupling calcium/calmodulin-mediated signaling and herbivore-induced plant response through calmodulin-binding transcription factor AtSR1/CAMTA3. *Plant Mol. Biol.* 79, 89–99. doi: 10.1007/s11103-012-9896-z

- Radwan, O., Liu, Y., and Clough, S. J. (2011). Transcriptional analysis of soybean root response to *Fusarium virguliforme*, the causal agent of sudden death syndrome. *Mol. Plant Microb. Interact.* 24, 958–972. doi: 10.1094/MPMI-11-10-0271
- Rahman, H., Yang, J., Xu, Y. P., Munyampundu, J. P., and Cai, X. Z. (2016). Phylogeny of plant CAMTAs and role of AtCAMTAs in nonhost resistance to *Xanthomonas oryzae* pv. oryzae. Front. Plant Sci. 7:117. doi: 10.3389/fpls. 2016.00177
- Robert-Seilaniantz, A., MacLean, D., Jikumaru, Y., Hill, L., Yamaguchi, S., Kamiya, Y., et al. (2011). The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates. *Plant J.* 67, 218–231. doi: 10.1111/j.1365-313X.2011.04591.x
- Schiebel, W., Pélissier, T., Riedel, L., Thalmeir, S., Schiebel, R., Kempe, D., et al. (1998). Isolation of an RNA-Directed RNA polymerase-specific cDNA clone from tomato. *Plant Cell* 10, 2087–2101.
- Shen, D., Suhrkamp, I., Wang, Y., Liu, S., Menkhaus, J., Verreet, J. A., et al. (2014). Identification and characterization of microRNAs in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using Brassica AA (*Brassica rapa*) and CC (*Brassica oleracea*) as reference genomes. *New Phytol.* 204, 577–594. doi: 10.1111/nph.12934
- Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., et al. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* 107, 465–476. doi: 10.1016/S0092-8674(01)00576-1
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Tijsterman, M., Ketting, R. F., and Plasterk, R. H. A. (2002). The genetics of RNA silencing. Annu. Rev. Genet. 36, 489–519. doi: 10.1146/annurev. genet.36.043002.091619
- Ullu, E., Tschudi, C., and Chakraborty, T. (2004). RNA interference in protozoan parasites. *Cell. Microbiol.* 6, 509–519. doi: 10.1111/j.1462-5822.2004.00399.x
- Voinnet, O. (2008). Post-transcriptional RNA silencing in plant-microbe interactions: a touch of robustness and versatility. *Curr. Opin. Plant Biol.* 11, 464–470. doi: 10.1016/j.pbi.2008.04.006
- Weiberg, A., Wang, M., Lin, F. M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342, 118–123. doi: 10.1126/science.1239705
- Wu, J., Yang, Z., Wang, Y., Zheng, L., Ye, R., Ji, Y., et al. (2015). Viral-inducible Argonaute18 confers broad-spectrum virus resistance in rice by sequestering a host microRNA. *Elife* 4:e05733. doi: 10.7554/eLife.05733
- Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., et al. (2004). Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* 2:E104. doi: 10.1371/journal.pbio.0020104
- Xin, M., Wang, Y., Yao, Y., Xie, C., Peng, H., Ni, Z., et al. (2010). Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 10:123. doi: 10.1186/1471-2229-10-123
- Yadav, C. B., Muthamilarasan, M., Pandey, G., and Prasad, M. (2015). Identification, characterization and expression profiling of Dicer-Like, Argonaute and RNA-dependent RNA polymerase gene families in foxtail millet. *Plant Mol. Biol. Rep.* 33, 43–55. doi: 10.1007/s11105-014-0736-y
- Yang, L., Jue, D., Li, W., Zhang, R., Chen, M., and Yang, Q. (2013). Identification of miRNA from eggplant (*Solanum melongena* L.) by small RNA deep sequencing and their response to *Verticillium dahliae* infection. *PLoS ONE* 8:e72840. doi: 10.1371/journal.pone.0072840
- Yang, T. B., and Poovaiah, B. W. (2002). A calmodulin-binding/CGCG box DNAbinding protein family involved in multiple signaling pathways in plants. J. Biol. Chem. 277, 45049–45058. doi: 10.1074/jbc.M207941200
- Yin, Z., Li, Y., Han, X., and Shen, F. (2012). Genome-wide profiling of miRNAs and other small non-coding RNAs in the *Verticillium dahliae*-inoculated cotton roots. *PLoS ONE* 7:e35765. doi: 10.1371/journal.pone.0035765
- Yoshikawa, M., Peragine, A., Park, M. Y., and Poethig, R. S. (2005). A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev.* 19, 2164–2175. doi: 10.1101/gad.1352605
- Yu, B., and Wang, H. (2010). Translational inhibition by microRNAs in plants. Prog. Mol. Subcel. Biol. 50, 41–57. doi: 10.1007/978-3-642-03103-8\_3

- Yu, D., Fan, B., MacFarlane, S. A., and Chen, Z. (2003). Analysis of the involvement of an inducible Arabidopsis RNA-dependent RNA polymerase in antiviral defense. *Mol. Plant Microb. Interact.* 16, 206–216. doi: 10.1094/MPMI.2003.16.3.206
- Zhao, H., Zhao, K., Wang, J., Chen, X., Chen, Z., Cai, R., et al. (2015). Comprehensive analysis of Dicer-Like, Argonaute, and RNA-dependent RNA polymerase gene families in grapevine (*Vitis Vinifera*). J. Plant Growth Regul. 34, 108–121. doi: 10.1007/s00344-014-9448-7
- Zhao, X., Zheng, W., Zhong, Z., Chen, X., Wang, A., and Wang, Z. (2016). Genome-wide analysis of RNA-interference pathway in *Brassica napus*, and the expression profile of BnAGOs in response to *Sclerotinia sclerotiorum* infection. *Eur. J. Plant Pathol.* 146, 565. doi: 10.1007/s10658-016-0942-6
- Zhao, Y., Liu, W., Xu, Y. P., Cao, J. Y., Braam, J., and Cai, X. Z. (2013). Genomewide identification and functional analyses of calmodulin genes in Solanaceous species. *BMC Plant Biol*. 13:70. doi: 10.1186/1471-2229-13-70
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S. H., Liou, L. W., et al. (2011). Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot

apical meristem development. Cell 145, 242–256. doi: 10.1016/j.cell.2011. 03.024

Zilberman, D., Cao, X., Johansen, L. K., Xie, Z., Carrington, J. C., and Jacobsen, S. E. (2004). Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* 14, 1214–1220. doi: 10.1016/j.cub.2004.06.055

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Cao, Xu, Li, Li, Rahman and Cai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.