



# Further Elucidation of the Argonaute and Dicer Protein Families in the Model Grass Species *Brachypodium distachyon*

Ena Šečić, Silvia Zanini and Karl-Heinz Kogel\*

Institute of Phytopathology, Centre for BioSystems, Land Use and Nutrition, Justus Liebig University, Giessen, Germany

## OPEN ACCESS

### Edited by:

Hailing Jin,  
University of California,  
United States

### Reviewed by:

Zhaoqing Chu,  
Shanghai Chenshan Plant Science  
Research Center (CAS), China

Ling Li,  
Mississippi State University,  
United States

### \*Correspondence:

Karl-Heinz Kogel  
Karl-Heinz.Kogel@agr.uni-giessen.de

### Specialty section:

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

**Received:** 13 April 2019

**Accepted:** 25 September 2019

**Published:** 22 October 2019

### Citation:

Šečić E, Zanini S and Kogel K-H  
(2019) Further Elucidation of the  
Argonaute and Dicer Protein  
Families in the Model Grass Species  
*Brachypodium distachyon*.  
Front. Plant Sci. 10:1332.  
doi: 10.3389/fpls.2019.01332

RNA interference (RNAi) is a biological process in which small RNAs regulate gene silencing at the transcriptional or posttranscriptional level. The trigger for gene silencing is double-stranded RNA generated from an endogenous genomic locus or a foreign source, such as a transgene or virus. In addition to regulating endogenous gene expression, RNAi provides the mechanistic basis for small RNA-mediated communication between plant hosts and interacting pathogenic microbes, known as cross-kingdom RNAi. Two core protein components, Argonaute (AGO) and Dicer (DCL), are central to the RNAi machinery of eukaryotes. Plants encode for several copies of AGO and DCL genes; in *Arabidopsis thaliana*, the AGO protein family contains 10 members, and the DCL family contains four. Little is known about the conservation and specific roles of these proteins in monocotyledonous plants, which account for the most important food staples. Here, we utilized *in silico* tools to investigate the structure and related functions of AGO and DCL proteins from the model grass *Brachypodium distachyon*. Based on the presence of characteristic domains, 16 BdAGO- and 6 BdDCL-predicted proteins were identified. Phylogenetic analysis showed that both protein families were expanded in *Brachypodium* as compared with *Arabidopsis*. For BdDCL proteins, both plant species contain a single copy of DCL1 and DCL4; however, *Brachypodium* contains two copies each of DCL2 and DCL3. Members of the BdAGO family were placed in all three functional clades of AGO proteins previously described in *Arabidopsis*. The greatest expansion occurred in the AtAGO1/5/10 clade, which contains nine BdAGOs (BdAGO5/6/7/9/10/11/12/15/16). The catalytic tetrad of the AGO P-element-induced wimpy testis domain (PIWI), which is required for endonuclease activity, is conserved in most BdAGOs, with the exception of BdAGO1, which lacks the last D/H residue. Three-dimensional modeling of BdAGO proteins using tertiary structure prediction software supported the phylogenetic classification. We also predicted a provisional interactome network for BdAGOs, their localization within the cell, and organ/tissue-specific expression. Exploring the specifics of RNAi machinery proteins in a model grass species can serve as a proxy for agronomically important cereals such as barley and wheat, where the development of RNAi-based plant protection strategies is of great interest.

**Keywords:** protein, structure, prediction, RNAi, Argonaute, Dicer, *Brachypodium*

## INTRODUCTION

RNA interference (RNAi) is a regulatory mechanism utilized by most eukaryotes for endogenous gene silencing and protection against mobile repetitive sequences, transposons, and viruses (Fire et al., 1998; Wilson and Doudna, 2013). In contrast to transcriptional gene silencing (TGS), which results in the methylation of DNA and/or histones, posttranscriptional gene silencing (PTGS) operates by transcript degradation or translation inhibition. Selection of the target for silencing is governed by sequence complementarity between a single-stranded small RNA (sRNA) and the target RNA. Beyond its native role, the RNAi machinery has been exploited for developing novel plant protection strategies based on double-stranded (ds)RNA applications. Delivery of artificial dsRNA through transgene expression [host-induced gene silencing (HIGS)] or exogenous application [spray-induced gene silencing (SIGS)] was proven effective against fungal pathogens (Nowara et al., 2010; Koch et al., 2013; Koch et al., 2016; Wang et al., 2016), nematodes (Dutta et al., 2015), insects (Coleman et al., 2014; Abdellatef et al., 2015; Head et al., 2017), and parasitic plants (Tomilov et al., 2008; for review, see Andrade and Hunter, 2016; Cai et al., 2018). A recent discovery revealed that RNAi also is involved in natural cross-kingdom RNA communication (ckRNAi), where sRNA molecules function as mediators that are exchanged bidirectionally between a host plant and a microbial pathogen to silence their target transcripts and impact the outcome of the plant–pathogen interaction (Weiberg et al., 2013; Zhang et al., 2016; Wang et al., 2017a; Wang et al., 2017b).

Regardless of which RNAi-based process or application is involved, evolutionarily conserved protein components, including Dicer [termed Dicer-like (DCL) in plants] and Argonaute (AGO), play key roles in dsRNA processing. Dicers and DCLs are RNase III endonucleases that process exogenously supplied or endogenously generated ds- or hairpin (hp)-containing RNA precursors into various species of dsRNAs, commonly 21–24 nucleotides (nt) in length. These sRNAs are then loaded onto specific AGO proteins, which are components of the RNA-induced silencing complex (RISC). The loaded sRNA is processed into a single-stranded sRNA molecule, which then guides the RISC to complementary targets in the cytoplasm or nucleus. Depending on the biological context, target recognition leads to PTGS *via* RNA degradation, which may be mediated by the AGO protein's slicer activity, or inhibition of translation, or to TGS *via* genomic DNA and/or histone methylation (Carthew and Sontheimer, 2009; Poulsen et al., 2013; Borges and Martienssen, 2015; Fang and Qi, 2016).

Phylogenetic analysis of genes belonging to the Dicer family suggests that they arose early in the evolution of eukaryotes and that their duplication and diversification correlated with the development of multicellularity and the need for complex gene regulation (Mukherjee et al., 2013). In plants, the structure and function of DCL proteins have been investigated most intensively in *Arabidopsis*, which expresses four DCLs (Schauer et al., 2002). The domain architecture of these proteins, like that of other eukaryotic Dicers, generally consists of an amino-terminal DEXDc and helicase-C (HELICc) domain, which are thought

to mediate processive movement along a dsRNA, a dicer-dimer (heterodimerization) domain that facilitates binding with protein partners (Qin et al., 2010), a P-element-induced wimpy testis (PIWI)–Argonaute–Zwille (PAZ) domain, which binds the 3' end of the dsRNA, two RIBOc (ribonuclease III family) domains and at least one dsRNA-binding motif (DSRM) domain at the C terminus (Schauer et al., 2002; Mukherjee et al., 2013; Bologna and Voinnet, 2014; Song and Rossi, 2017).

Analyses of *Arabidopsis* mutants revealed that the four DCLs generate different types of sRNAs, although some functional redundancy was observed (Gascioli et al., 2005; Bologna and Voinnet, 2014; Borges and Martienssen, 2015). AtDCL1 produces microRNAs (miRNAs), a class of sRNAs that regulates endogenous gene expression via PTGS (Bartel, 2004). The remaining AtDCLs generate various subclasses of small interfering RNAs (siRNAs), including i) natural-antisense-transcript (nat)-siRNAs generated by AtDCL2 (Borsani et al., 2005), ii) trans-activating (ta)-siRNAs produced by AtDCL4 (Dunoyer et al., 2005), and iii) TGS-related 24-nt siRNAs generated by AtDCL3, which are responsible for silencing transposons and other repeated DNA sequences (Xie et al., 2004).

Despite the diversity of sRNAs, their association with AGO proteins and the RISC complex is a common feature. AGO proteins were named after the tube-shaped leaves of *Arabidopsis ago1* mutants, which resemble the tentacles of the pelagic octopus, *Argonauta argo* (Bohmert et al., 1998). AGO proteins are highly conserved in nature, although the size of this family varies substantially between species (Höck and Meister, 2008; Zhang et al., 2015; Fang and Qi, 2016; You et al., 2017).

AGOs have a high level of structure and domain conservation between the prokaryotic and eukaryotic variants, even when the biological function clearly differs (Willkomm et al., 2015). Several prokaryotic (Wang et al., 2009; Liu et al., 2018) and eukaryotic (Lingel et al., 2003; Lingel et al., 2004; Boland et al., 2011) complete AGO structures or individual domains have been crystallographically resolved. Several human AGO proteins have been crystallized, namely, AGO2 in complex with a miRNA (Elkayam et al., 2012), AGO1 (Faehnle et al., 2013), and AGO3 (Park et al., 2017), showing that the AGO activity is dependent on conservation of active site residues and their interaction with other protein regions. The structures of AtAGO proteins have been partly resolved, especially the middle (MID) domain of AtAGO1, AtAGO2, and AtAGO5 (Frank et al., 2012; Zha et al., 2012). Functional domains characteristic of all AGO proteins, including the *Arabidopsis* AGOs, are the PAZ, MID, and PIWI domains governing the binding of sRNA ends and the slicer activity (Höck and Meister, 2008; Frank et al., 2012).

In *Arabidopsis*, 10 different AGOs have been identified. Phylogenetic analyses have divided them into three clades, comprising AGO1/5/10, AGO2/3/7, and AGO4/6/8/9 (Vaucheret, 2008). The different clades contain AGOs that mediate PTGS or TGS after they load specific types of sRNAs, which are selected based on length and identity of the 5' nt (Bologna and Voinnet, 2014; Zhang et al., 2015; Fang and Qi, 2016). For example, AtAGO1 is involved in endogenous developmental regulation by miRNAs (Vaucheret et al., 2004), antiviral defense (alongside AtAGO2, Harvey et al., 2011), as well as bidirectional ckRNAi

(Weiberg et al., 2013). AtAGO4 and AtAGO6 are involved in DNA and histone methylation (Zilberman et al., 2003; Zheng et al., 2007). AtAGO9 is known to be involved in female gametogenesis (Olmedo-Monfil et al., 2010), while AtAGO10 competes with AtAGO1 for sRNA loading in regulation of shoot apical meristem development (Zhu et al., 2011). AtAGO7 plays a role in defense against viruses (Qu et al., 2008).

In comparison to *Arabidopsis*, little is known about the RNAi machinery components in monocots. The number of AGO proteins is expanded in cereals, as there are 17 AGOs in maize (*Zea mays*) and 19 in rice (*Oryza sativa*; Fang and Qi, 2016; Mirzaei et al., 2014; Patel et al., 2018). The copy number of DCL2 and/or DCL3 genes also differs between monocot species and *Arabidopsis*. Six predicted DCLs were identified in rice, while five were identified in maize, wheat, and barley (Margis et al., 2006). Furthermore, the DCL3b gene has diverged significantly from its DCL3a paralog (Margis et al., 2006) and, thus, is considered a distinct, monocot-specific class of Dicer, termed DCL5 (Fukudome and Fukuhara, 2017; Borges and Martienssen, 2015).

Cereals are major staple crops worldwide; however, a plethora of pathogens and pests threaten their production (Savary et al., 2012). Recent efforts to develop environmentally friendly plant protection strategies have demonstrated that HIGS and SIGS can be used in major cereal crops, such as barley and wheat, to control necrotrophic fungi (Koch et al., 2013; Koch et al., 2016; Koch et al., 2018) and aphid pests (Abdellatef et al., 2015). Developing a better understanding of the cereal AGO and DCL protein family members and their specific functions is a prerequisite for clarifying the mechanisms undergirding RNAi-mediated plant protection. Here, we use the model species for temperate grass plants, *Brachypodium distachyon* (*Brachypodium*), to investigate cereal AGO and DCL proteins. *Brachypodium* is self-fertile, has a small genome (~272 Mb), a short life cycle, and established transformation protocols (Vogel et al., 2006). The commonly used diploid inbred line Bd21 is fully sequenced (The International Brachypodium Initiative, 2010). In addition, literature data reveal strong responsiveness of *Brachypodium* sRNA pools to abiotic stress, suggesting that the RNAi machinery is sensitive to environmental changes (Wang et al., 2015).

Based on genomic database searches and *in silico* analysis, we identified six BdDCL proteins, as well as 16 previously reported AGO protein sequences in *Brachypodium* (Mirzaei et al., 2014). Since the structure of proteins closely relates to function and thus can serve as an indication of interaction patterns and redundancy in large protein families, we especially looked into domain structure conservation in *Brachypodium* relative to *Arabidopsis*. Similar to the protein structure and interactome analysis applied in Secic et al. (2015), we subjected Bd AGO and DCL proteins to a series of *in silico* analysis steps. The focus of this study is the structures and related functions of the AGO-like and DCL proteins of *B. distachyon*, with special regard to analysis of the phylogeny and three-dimensional (3D) structure modeling of the AGO family, as compared with the more familiar *Arabidopsis thaliana* AGO protein family. Given that the At AGO1/5/10

clade contains proteins involved in ckRNAi, we were especially interested to define the BdAGO proteins that are structurally most related to this clade and thus potentially have a key function in plant immunity and RNAi-based plant protection.

## MATERIALS AND METHODS

### Acquisition of Sequences and Database Search

AGO and DCL protein sequences corresponding to the primary transcripts of specific genes were acquired by searching the Plant Comparative Genomics portal Phytozome 12 (Goodstein et al., 2012) *B. distachyon* v3.1 database (The International Brachypodium Initiative, 2010). Proteins whose domain architecture resembled those of *Arabidopsis* AGO and DCL proteins were considered. The *Arabidopsis* AGO and DCL protein sequences were taken from The *Arabidopsis* Information Resource database (Rhee et al., 2003; Berardini et al., 2015). Information on resolved protein structures was acquired from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (Berman et al., 2000; Burley et al., 2018). The Brachypodium eFP Browser (Sibout et al., 2017; Winter et al., 2007) was used to assess the expression of transcripts corresponding to proteins involved in this study, based on the expression atlas detailing different organs and developmental stages.

### Phylogenetic Analysis, Interactome Analysis, and Localization

The phylogenetic analysis and tree rendering were done by the Phylogeny.fr web server (Dereeper et al., 2008; Dereeper et al., 2010). The operational sequence is composed of MUSCLE 3.8.31 (Edgar, 2004) for alignment with default settings, Gblocks 0.91b for removal of ambiguous regions (Castresana, 2000), PhyML 3.1/3.0 aLRT for phylogeny (Guindon and Gascuel, 2003; Anisimova and Gascuel, 2006), based on maximum likelihood, and TreeDyn 198.3 for graphical representation (Chevenet et al., 2006). Multiple sequence alignment (MSA) was done using Clustal Omega at European Molecular Biology Laboratory—European Bioinformatics Institute (Sievers et al., 2011; Goujon et al., 2010) and the conserved residues and domains visualized by the Mview multiple alignment viewer (Brown et al., 1998). Pairwise sequence alignments were done using EMBOSS Needle (Rice et al., 2000), utilizing the Needleman–Wunsch algorithm for global alignment. Domain search was conducted using Simple Modular Architecture Research Tool (SMART) in normal SMART mode (Schultz et al., 1998; Letunic and Bork, 2017) and visualized with the Illustrator for Biological Sequences (IBS) online illustrator (Liu et al., 2015). Prediction of protein location was done using the plant subcellular localization integrative predictor (PSI), which shows an integrative result based on the output of an 11-member predictor community (Liu et al., 2013). Prediction of the interactome was done using the STRING database of protein–protein associations, while searching by protein

sequence (Szklarczyk et al., 2019). Resulting associations/possible interactions that originate from text mining have been excluded, and the results show only associations supported by co-expression and/or experimental data.

### Three-Dimensional Structure Modeling and Validation

SWISS-MODEL, a homology-based modeling software available at the ExPASy web server (Waterhouse et al., 2018), and CPHmodels 3.2 protein homology modeling server (Nielsen et al., 2010) were both used for 3D structure prediction from the sequence data. BLAST (Camacho et al., 2009) and HHBlits (Remmert et al., 2011) template search through the SWISS-MODEL template library was done and the models built using ProMod3 (Waterhouse et al., 2018) and the target-template alignment.

QMEAN, used for validation of the predicted 3D structures, is a scoring function that considers single residues and the global model, delivering an estimation of absolute quality of the prediction (Benkert et al., 2011). In order to check the stereochemical quality of predicted structures, we used the PROCHECK program (Morris et al., 1992; Laskowski et al., 1993). One of the stereochemical parameters considered is the fitness of the model in a Ramachandran plot, which maps the allowed backbone dihedral angles of amino acids (aa) in a protein structure (Ramachandran et al., 1963). Further on, we used the WHATCHECK software (Hooft et al., 1996) to calculate the Ramachandran Z-score, which compares the quality of the query structure to structures with high confidence (Hooft et al., 1997). Lastly, we used the dDFIRE/DFIRE2 energy calculation (Yang and Zhou, 2008) to calculate free energy scores for our structure predictions.

PyMOL (The Py-MOL Molecular Graphics System) was used for visualization of the predicted structures (Schrödinger, 2010, Open-Source PyMOL 1.3).

## RESULTS

### Argonaute and Dicer Protein Families Are Expanded in *Brachypodium* Relative to *Arabidopsis*

To identify AGO and DCL proteins, the *B. distachyon* v3.1 database (The International Brachypodium Initiative, 2010) was searched for transcripts whose encoded proteins contain the characteristic domain architecture of each protein family. The accession numbers of the acquired sequences, the names assigned to the corresponding BdAGO proteins, the location of the encoding genes, and a description of the primary transcripts are shown in **Table 1**. The naming convention is similar to that used by Mirzaei et al. (2014) for 16 AGO proteins identified by primary transcripts in the *B. distachyon* Bd21 v3.1 annotation (The International Brachypodium Initiative, 2010). Our search for BdDCL candidates within the Bd21 v3.1 database revealed nine sequences. Clear lack of functional domains or insufficient length of the deduced aa sequence reduced the number of putative *DCL* genes to six (Bradi1g15440, Bradi1g77087, Bradi2g23187, Bradi5g15337, Bradi1g21030, and Bradi3g29287). Accession numbers and assigned names for the encoded BdDCLs are shown in **Table 2**. The putative AtAGO and AtDCL protein sequences were downloaded from The *Arabidopsis* Information Resource database (**Table S1**) and included in the MSA and phylogenetic analysis.

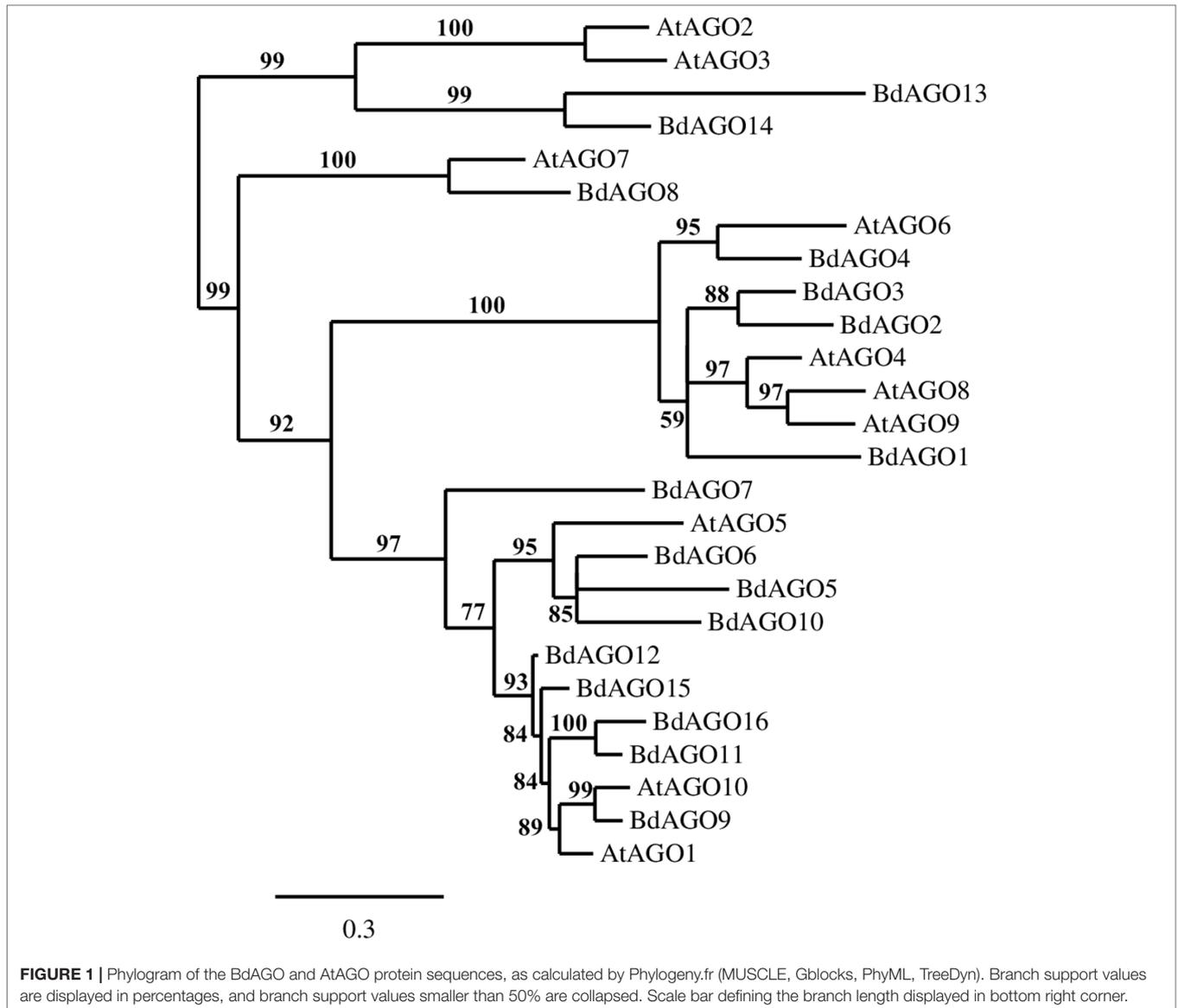
A phylogenetic analysis of the inferred BdAGO protein sequences relative to those of *Arabidopsis* AGOs is shown in **Figure 1**. BdAGO proteins were placed in all three AtAGO clades. Some were grouped with a specific AtAGO member within a clade (e.g., BdAGO8 was grouped with AtAGO7, and BdAGO5/6/7/10 were grouped with AtAGO5), whereas other BdAGOs were distributed throughout an entire clade (e.g., BdAGO1/2/3/4 within the AtAGO4/6/8/9 clade). In the AtAGO1/5/10 clade, BdAGO9/11/12/15/16 were interspersed

**TABLE 1** | Assigned names and accession numbers of BdAGO proteins as well as genomic location and description of the primary transcript (as acquired from Phytozome Bd21 v3.1 database).

Assigned name of protein	Primary transcript ID (Phytozome)	Location	Description (Phytozome)
BdAGO1	Bradi2g10360.2	Bd2:8611187.8615652 reverse	PTHR22891//PTHR22891:SF44 – Eukaryotic translation initiation factor 2C
BdAGO2	Bradi2g14147.1	Bd2:12806099.12812784 reverse	PTHR22891:SF20 – Protein AGO 4-related
BdAGO3	Bradi2g10370.1	Bd2:8620394.8628745 reverse	AGO family, subfamily AGO4
BdAGO4	Bradi4g08587.1	Bd4:7715921.7724879 reverse	PTHR22891:SF35 – Protein AGO 6
BdAGO5	Bradi1g12431.2	Bd1:9307067.9313002 forward	PTHR22891//PTHR22891:SF49 – Eukaryotic translation initiation factor 2C
BdAGO6	Bradi1g05162.2	Bd1:3447373.3455769 forward	PTHR22891//PTHR22891:SF24 – Eukaryotic translation initiation factor 2C
BdAGO7	Bradi1g28260.3	Bd1:23482384.23489131 reverse	AGO family, subfamily monocot-AGO1
BdAGO8	Bradi1g16060.3	Bd1:12986117.12991032 reverse	AGO family, subfamily AGO7
BdAGO9	Bradi1g36907.2	Bd1:32760045.32772130 reverse	PTHR22891//PTHR22891:SF25 – Eukaryotic translation initiation factor 2C
BdAGO10	Bradi1g54977.1	Bd1:53536162.53543236 forward	PTHR22891//PTHR22891:SF36 – Eukaryotic translation initiation factor 2C
BdAGO11	Bradi1g29577.1	Bd1:25162908.25171156 reverse	PTHR22891//PTHR22891:SF57 – Eukaryotic translation initiation factor 2C
BdAGO12	Bradi5g18540.1	Bd5:21720455.21728815 reverse	AGO family, subfamily AGO1
BdAGO13	Bradi5g21810.1	Bd5:24487261.24492250 forward	AGO family, subfamily AGO2/3
BdAGO14	Bradi5g21800.1	Bd5:24479944.24484383 forward	AGO family, subfamily AGO2/3
BdAGO15	Bradi3g51077.3	Bd3:51944662.51956527 forward	PTHR22891//PTHR22891:SF34 – Eukaryotic translation initiation factor 2C
BdAGO16	Bradi3g60697.5	Bd3:59325332.59333596 reverse	PTHR22891//PTHR22891:SF34 – Eukaryotic translation initiation factor 2C

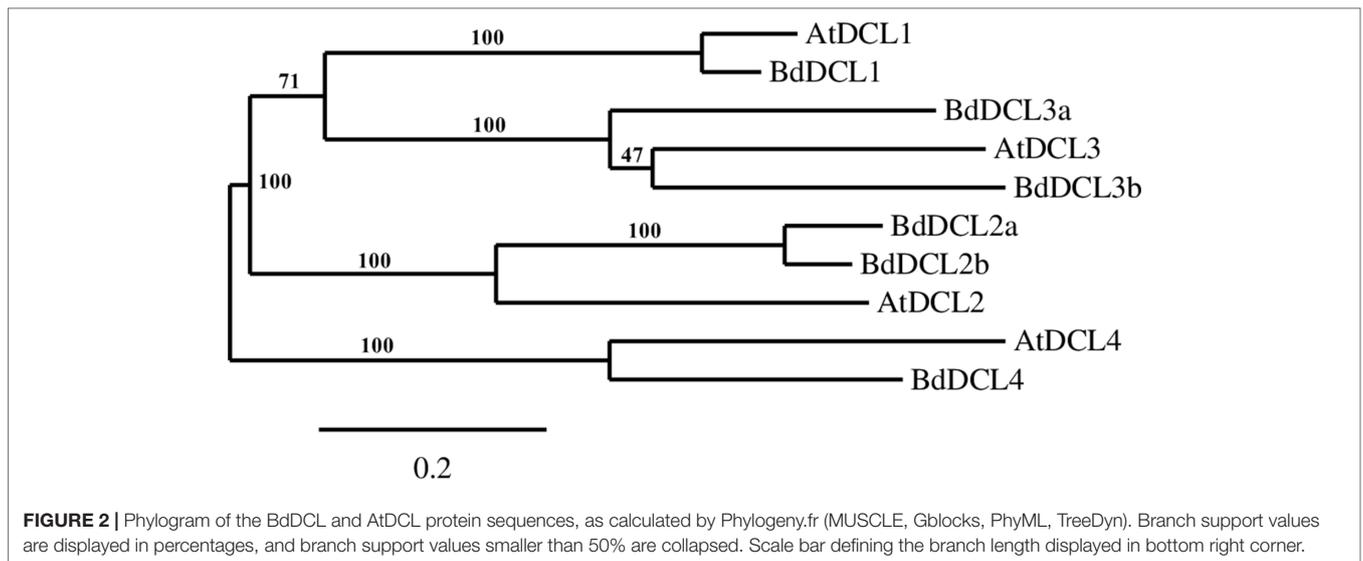
**TABLE 2** | Assigned names and accession numbers of BdDCL proteins as well as genomic location and description of the primary transcript (as acquired from Phytozome Bd21 v3.1 database).

Assigned name of protein	Primary transcript ID (Phytozome)	Location	Description (Phytozome)
BdDCL1	Bradi1g77087.1	Bd1:73701094.73713218 forward	PTHR14950:SF3 – ENDORIBONUCLEASE DICER HOMOLOG 1
BdDCL2a	Bradi1g15440.1	Bd1:12353426.12376799 forward	DCL family, subfamily DCL2
BdDCL2b	Bradi1g21030.3	Bd1:16934990.16948923 reverse	PTHR14950:SF19 - ENDORIBONUCLEASE DICER HOMOLOG 2
BdDCL3a	Bradi3g29287.1	Bd3:31008845.31020951 forward	PTHR14950//PTHR14950:SF31 - HELICASE-RELATED//SUBFAMILY NOT NAMED
BdDCL3b	Bradi2g23187.3	Bd2:20726122.20733365 reverse	PF00636//PF02170//PF03368 – Ribonuclease III domain (Ribonuclease_3)//PAZ domain (PAZ)//Dicer dimerization domain (Dicer_dimer)
BdDCL4	Bradi5g15337.3	Bd5:18845215.18867304 reverse	PTHR14950:SF15 – DCL 4



with AtAGO1 and AtAGO10. These findings suggest that the structural and functional differences of AtAGO proteins are translated to the expanded *Brachypodium* family. Phylogenetic analysis of the inferred BdDCL proteins showed that they

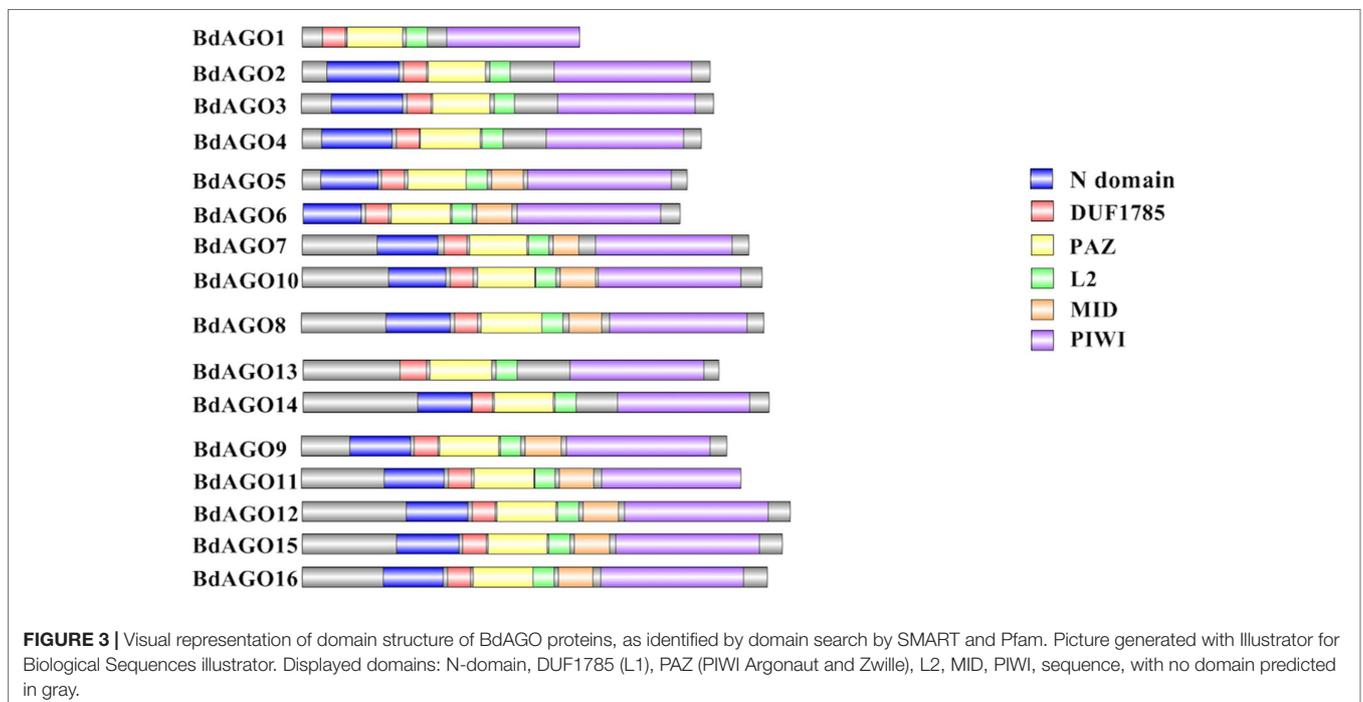
strongly aligned with individual members of the *Arabidopsis* DCL family (Figure 2). Like *Arabidopsis*, *Brachypodium* contains a single ortholog of DCL1 and DCL4; however, expansion of the BdDCL family has led to the presence of two copies of both

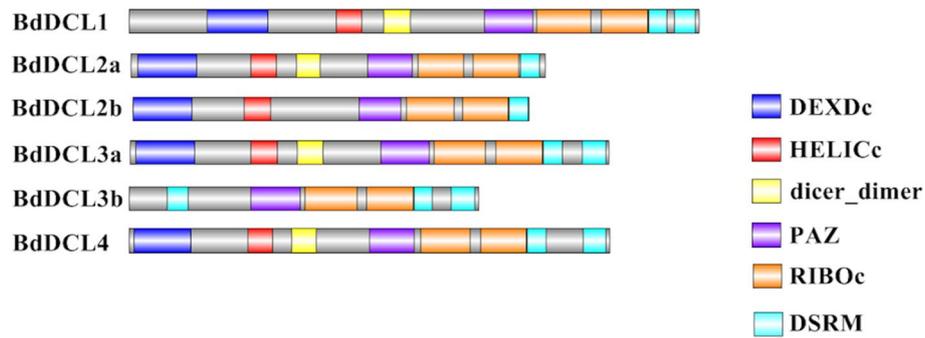


DCL2 and DCL3, as compared with the single ortholog present in *Arabidopsis*. Sequence comparisons revealed that DCL2a and DCL2b share 82.5% similarity at the aa level, while BdDCL3a and BdDCL3b share 44.2% similarity (aa, global alignment). Together, the phylogenetic trees show distinct branches interspersing *Arabidopsis* and *Brachypodium* homologues in functional clades; to our knowledge, this is the first indication of how the expansion of AGO and DCL protein families in *Brachypodium* relates to the specific clades and/or functional diversity of the corresponding *Arabidopsis* proteins.

### Predicted Domains of BdAGO and BdDCL Proteins Indicate Structure Conservation

Next, we executed a domain search using SMART to elucidate the structures and functions of the 16 BdAGO proteins and six BdDCLs. The domain structure visualization of BdAGO (Figure 3) and BdDCL (Figure 4) proteins highlights the differences between members of each protein family with respect to the positions and presence/absence of the typically conserved domains. Detailed domain prediction data, as acquired by SMART/Pfam search, and the corresponding confidence values

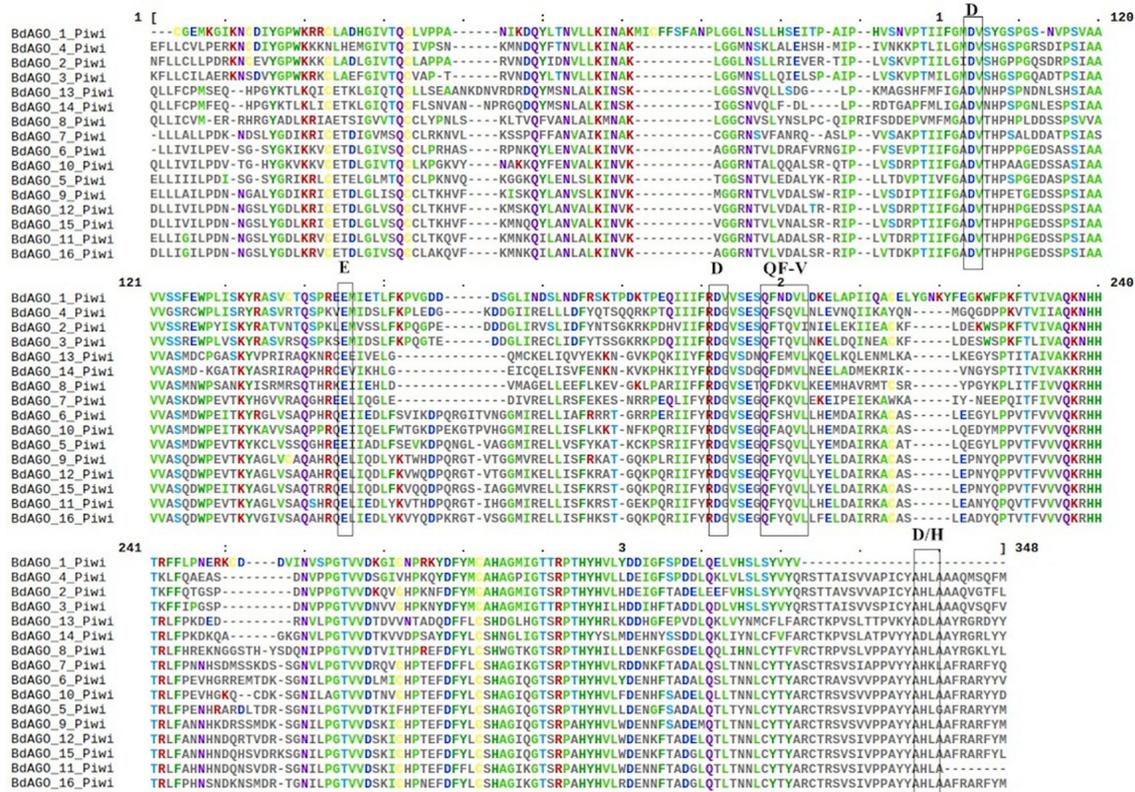




**FIGURE 4 |** Visual representation of domain structure of Bd DCL proteins, as identified by domain search by SMART and Pfam. Picture generated with Illustrator for Biological Sequences illustrator. Displayed domains: DEAD-like helicase superfamily (DEXDc), helicase superfamily c-terminal domain (HELICc), dicer\_dimer, PIWI Argonaut and Zwiille (PAZ), ribonuclease III family (RIBOc), double-stranded RNA-binding motif (DSRM), sequence with no domain predicted in gray.

are shown for BdAGO (Table S2) and BdDCL proteins (Table S3). Consistent with other eukaryotic AGO proteins, many members of the BdAGO family are predicted to have four characteristic functional domains, including the N-terminal domain, PAZ, MID, and PIWI domain (Zhang et al., 2014). However, while the domain prediction results identified a variable N-t domain in most BdAGOs that consisted of both an N-domain and a DUF1785 domain (Poulsen et al., 2013), BdAGO1 and BdAGO13

contained only the DUF1785 domain. In addition, BdAGO1, BdAGO2, BdAGO3, BdAGO4, BdAGO13, and BdAGO14 were not predicted to contain a MID domain, in comparison to a previous report (Mirzaei et al., 2014). MSA performed by Clustal Omega on the PIWI domain of BdAGO proteins (Figure 5) showed a typical pattern of conservation for the DEDD/H catalytic tetrad required for slicer activity and a conserved QF-V motif in all aligned sequences except BdAGO1, which has the



**FIGURE 5 |** Multiple sequence alignment (MSA) of the PIWI domain of BdAGO proteins, as acquired by Clustal Omega and Mview visualization. The catalytic tetrad DEDD/H (Asp-Glu-Asp-Asp or Asp-Glu-Asp-His) and the QF-V (Gln-Phe-Val) motifs are boxed.

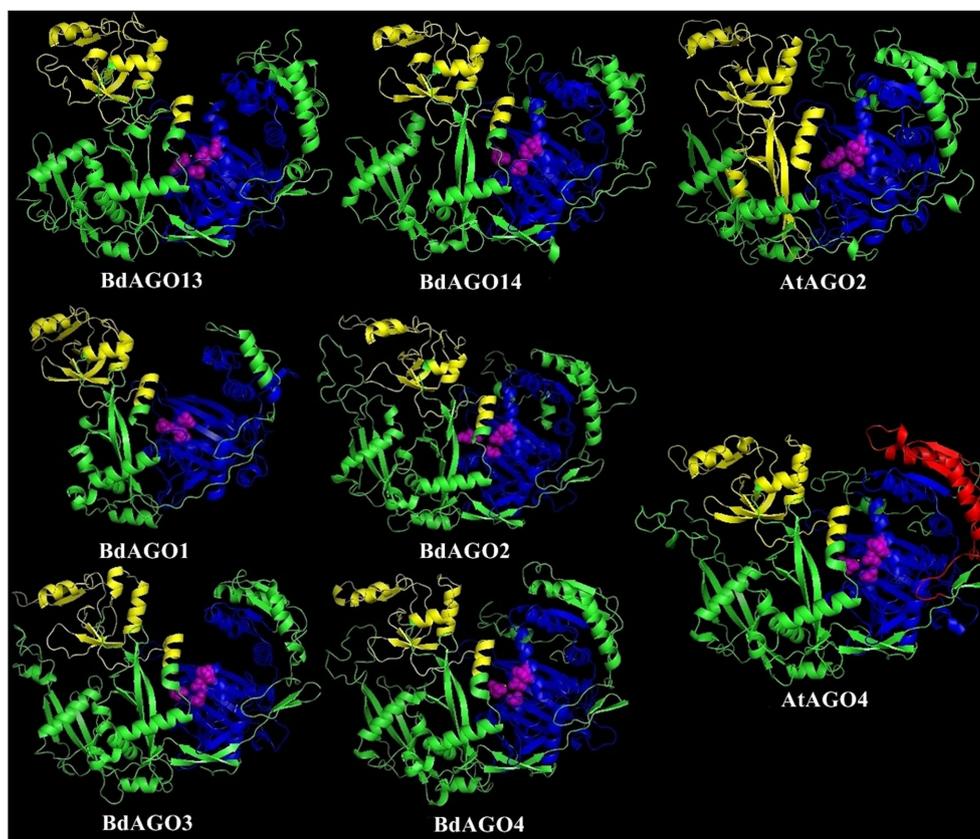
shortest protein sequence of all the BdAGO proteins and lacks the D/H residue of the catalytic tetrad.

Analysis of the predicted domains in BdDCL proteins revealed that the characteristic DEXDc, HELICc, Dicer-dimer (DUF283, Qin et al., 2010), PAZ, RIBOc, and DSRM domains are present in most family members. However, BdDCL2b and BdDCL3b lack the dimerization domain; BdDCL3b additionally lacks both the DEXDc and HELICc domains and instead contains an additional DSRM domain at the N terminus (position: 131-218, E-value:  $8.6e-17$ ; **Figure 4** and **Table S3**). By contrast, BdDCL2a and BdDCL2b contain only one DSRM domain (**Figure 4**, **Table S3**).

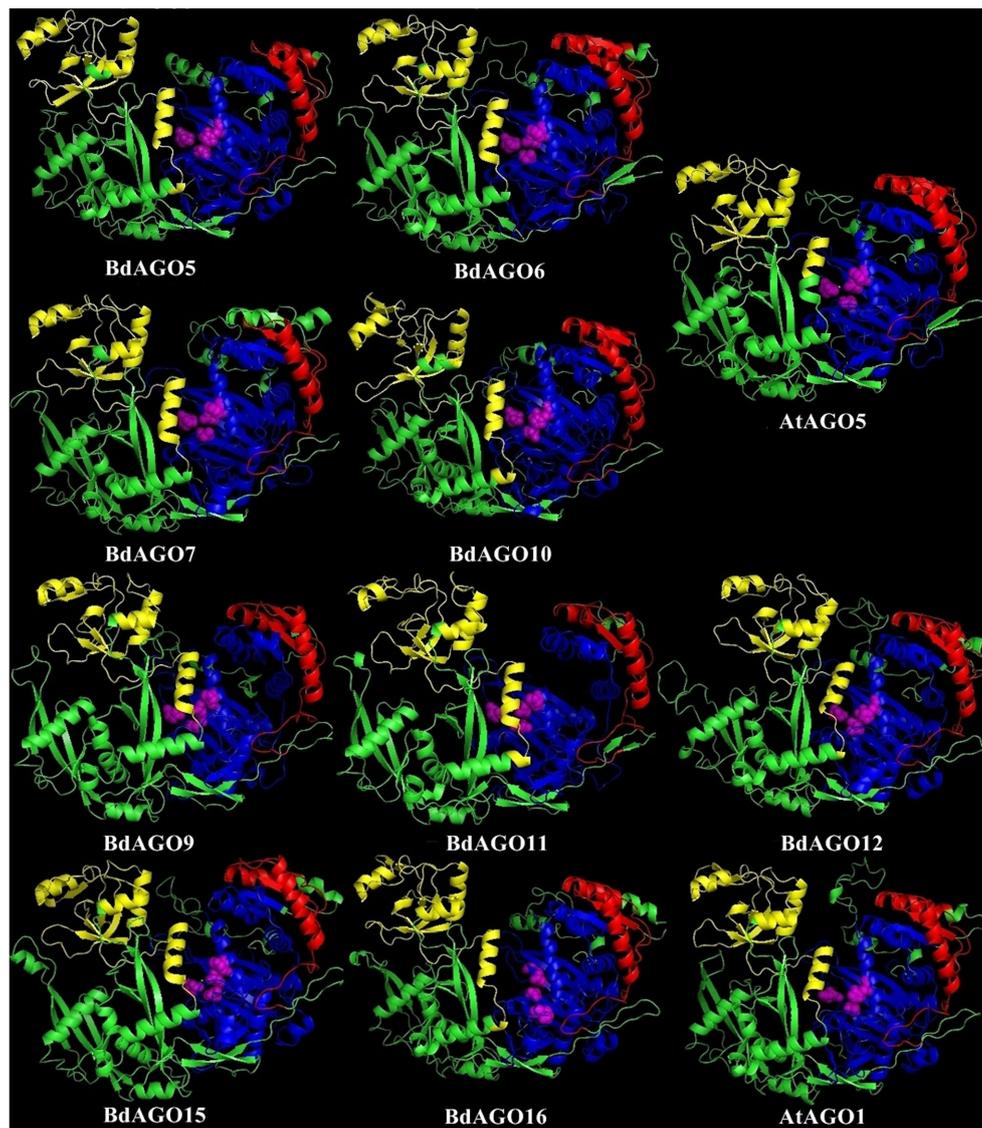
### Three-Dimensional Modeling Supports Phylogenetic Data Showing a Strong Expansion in the BdAGOs in the AGO1/5/10-Related Clade

In order to obtain an optimal homology-based 3D model of the studied proteins, we used SWISS-MODEL and CPHmodels 3.2. When choosing between models generated by alternative software programs or based on different templates, validation of the predicted structures is crucial for generating a consensus on the optimal model and further comparison. In case of

BdAGOs, validation of the predicted structures was done using four different measurements, the results of which are shown in **Table S4**. While a 0-1 QMEAN value gives an absolute scoring of the predicted model, the Z-score shown in **Table S4** serves as a comparison of the quality of the prediction of the query model relative to expected from a high-resolution X-ray crystallography structure. Typically, the more negative the Z-score is, the lower the quality of the predicted structure. Using PROCHECK, we report on the percentage of residues that fall into the most favored regions of the Ramachandran plot. The free energy score of the conformation of the predicted protein calculated by dDFIRE usually indicates lower values for a better model. Based on validation of the 3D models by the software, SWISS-MODEL was chosen as the preferred modeling tool for BdAGO proteins (**Table S4**). The corresponding AtAGO 3D structures, predicted and validated in the same fashion (**Table S5**), were subsequently used alongside the visualization of the BdAGO proteins by PyMOL. **Figures 6**, **7**, and **S1** display the models, in which the PAZ, MID, and PIWI domains (where predicted) and residues comprising the DEDD/H catalytic tetrad are indicated for all BdAGOs and a corresponding AtAGO representing the appropriate branch of the phylogenetic tree depicted in **Figure 1**. Overall, the predicted structures of the BdAGOs mirror the



**FIGURE 6 |** Three-dimensional structure predictions for BdAGO13 and BdAGO14 (with AtAGO2 as the closest homolog in *Arabidopsis*) and BdAGO1, BdAGO2, BdAGO3, BdAGO4 (with AtAGO4 as the closest homolog in *Arabidopsis*), as modeled by SWISS-MODEL. PAZ (yellow), PIWI (blue), and MID (red) domains as predicted by SMART and Pfam displayed. The catalytic tetrad within the PIWI domain (DEDD) is marked by magenta spheres. Visualization by PyMOL.



**FIGURE 7 |** Three-dimensional structure predictions for BdAGO5, BdAGO6, BdAGO7, and BdAGO10 (with AtAGO5 as the closest homolog in *Arabidopsis*) and BdAGO9, BdAGO11, BdAGO12, BdAGO15, and BdAGO16 (with AtAGO1 as the closest homolog in *Arabidopsis*), as modeled by SWISS-MODEL. PAZ (yellow), PIWI (blue), and MID (red) domains as predicted by SMART and Pfam displayed. The catalytic tetrad within the PIWI domain (DEDD) is marked by magenta spheres. Visualization by PyMOL.

corresponding AtAGO structures, suggesting a functional conservation. The PIWI domain and the catalytic tetrad especially show similarity between the clade members shown together in **Figures 6** and **7**. The BdDCL proteins did not have successfully modeled structures predicted by either software and thus are not shown.

### Expression Analysis and Putative Interactors of BdAGO Proteins

We addressed the question of tissue-specific expression of *BdAGO* and *BdDCL* genes by utilizing the *B. distachyon* eFP Browser (Sibout et al., 2017; Winter et al., 2007) in **Table 3**.

Stronger expression of *BdAGO* and *BdDCL* genes was observed in seed and stem tissue compared with roots or leaves. The plant subcellular localization integrative predictor used for protein localization predicted that all *BdAGO*s reside in the cytosol, except for *BdAGO3*, *BdAGO14* (predicted to localize in the nucleus), and *BdAGO7* (predicted to localize in plastids), with varying scores of confidence (**Table S7**).

Finally, prediction of proteins that interact with *BdAGO*s was carried out using STRING (**Table S6**). All predicted *BdAGO*s were found to be either co-expressed or experimentally shown to interact with three proteins: Bradi1g36340.1, Bradi2g30160.1, and Bradi4g45065.1. BLASTP search of these protein sequences identified them as a 110-kDa U5 small nuclear ribonucleoprotein

**TABLE 3 |** Gene expression data as displayed on the *B. distachyon* eFP Browser (Winter et al., 2007; Sibout et al., 2017).

Gene ID	Assigned name	<sup>1</sup> Highest expression signal	<sup>2</sup> Peduncle, spikelet and stem nodes	<sup>2</sup> Root	<sup>2</sup> Leaf	<sup>2</sup> Seed
Bradi2g10360	BdAGO1	Whole_grain_11_DAF				
Bradi2g14147	BdAGO2	First_node_27_DAG				
Bradi2g10370	BdAGO3	First_node_10_DAG				
Bradi4g08587	BdAGO4	Whole_grain_2_years				
Bradi1g12431	BdAGO5	Not found in browser				
Bradi1g05162	BdAGO6	Endosperm_31_DAF				
Bradi1g28260	BdAGO7	Endosperm_11_DAF				
Bradi1g16060	BdAGO8	First_node_10_DAG				
Bradi1g36907	BdAGO9	First_node_10_DAG				
Bradi1g54977	BdAGO10	Whole_grain_11_DAF				
Bradi1g29577	BdAGO11	Upper_part_of_inclined_node_42_DAG				
Bradi5g18540	BdAGO12	Last_internode_35_DAG				
Bradi5g21810	BdAGO13	Last_node_35_DAG				
Bradi5g21800	BdAGO14	Last_internode_35_DAG				
Bradi3g51077	BdAGO15	Roots_10_DAG				
Bradi3g60697	BdAGO16	Roots_10_DAG				
Bradi1g77087	BdDCL1	Whole_grain_2_years				
Bradi1g15440	BdDCL2a	First_internode_27_DAG				
Bradi1g21030	BdDCL2b	Whole_grain_2_years				
Bradi3g29287	BdDCL3a	Lower_part_of_inclined_node_42_DAG				
Bradi2g23187	BdDCL3b	First_node_10_DAG				
Bradi5g15337	BdDCL4	Whole_grain_2_years				

<sup>1</sup>The tissue with the highest absolute expression level per gene ID is indicated (DAF, day after fertilization; DAG, day after germination).

<sup>2</sup>Summary of relative expression level per gene ID displayed for tissues; color indicates relative expression levels (log2, the control is calculated from all the samples displayed on the particular eFP browser view); dark blue (high for the transcript relative to control) to light blue (low for the transcript relative to control).

component CLO (Bradi1g36340.1), a putative GTP-binding/transcription factor (Bradi2g30160.1), and DNA-directed RNA polymerase V subunit 1 or DNA-directed RNA polymerase V subunit 1 (Bradi4g45065.1). In addition to the aforementioned proteins, BdAGO9 (classified in the AtAGO1/5/10 clade) was predicted to interact with seven other proteins, identified as three homeobox proteins knotted-1-like (Bradi1g12677.1, Bradi1g12690.1, Bradi1g57607.1), two GATA transcription factors (Bradi2g14890.1, Bradi2g45750.1), and two putative uncharacterized proteins (Figure S2).

## DISCUSSION

In the present work, we investigated the phylogenetic relationships, domain, structure conservation, and predicted redundancy of AGO and DCL proteins in the model grass plant *B. distachyon*. Our findings imply that BdAGOs and BdDCLs have more copies and possibly greater diversification relative to *Arabidopsis*. One known example of such diversification in monocotyledonous plants is the rice AGO18, which confers antiviral immunity by sequestration of an miRNA (Wu et al., 2015). Since the presence of domains typical for AGO and DCL protein families serves as a selection criterion for proteins within this uninvestigated grass model species, we discuss phylogenetic relationships and predicted domain occurrence in detail.

Our analyses show that *Brachypodium*, like other grasses, contains one protein (BdDCL1) whose sequence groups with AtDCL1, one with AtDCL4 (BdDCL4), and two proteins each that group with AtDCL2 and AtDCL3 (Margis et al., 2006). Analysis

of their predicted domain structures showed that BdDCL2b and BdDCL3b lack the dicer-dimer (DUF283) domain, known to mediate heterodimerization of AtDCL4 with its protein partners (Qin et al., 2010), but it is partially missing in two other DCLs (AtDCL3 and OsDCL2b, Margis et al., 2006). The second DSRM domain also was not predicted in either of the BdDCL2s. This finding is consistent with the previous discovery that AtDCL2 in *Arabidopsis* and OsDCL2a and OsDCL2b in rice also contain only one DSRM (Margis et al., 2006). This second DSRM domain has only a weak affinity for dsRNA, but it specifically binds to proteins of the HYPONASTIC LEAVES 1/dsRNA-binding protein family (Hiraguri et al., 2005; Margis et al., 2006). Since the DSRM domains mediate the transfer of the newly generated sRNA to the appropriate AGO protein (Parker et al., 2008), variations in the C-terminal architecture may influence which downstream partners and RNAi pathways are utilized by specific DCLs. The high level of divergence between DCL3a and DCL3b in several monocot species has led to the classification of DCL3b as a distinct type of DCL, termed DCL5. This monocot-specific class of DCLs has been retained for over 60 million years (Margis et al., 2006). It is responsible for generating 24-nt-phased sRNAs in the male reproductive organs (Song et al., 2012). Interestingly, the predicted domain structure of BdDCL3b differs substantially from that of BdDCL3a, as it lacks both the DEXDc and HELICc domains (alongside the dicer-dimer domain) but contains an additional N-terminal DSRM (Figure 4, Table S3). Since the helicase domains are thought to mediate unwinding of the dsRNA (Zhang et al., 2004), the functionality of BdDCL3b is unclear. Mutations in AtDCL1 that impair helicase activity were previously shown to suppress miRNA accumulation (Liu

et al., 2012). However, comparable levels of transcripts for two splice variants of AtDCL2, one of which contains an altered helicase region, were detected throughout the *Arabidopsis* life cycle (Margis et al., 2006). Additional structural and biochemical analyses are therefore required to assess the role of BdDCL3b.

Phylogenetic analysis of the BdAGO protein family placed members in all three clades defined by *Arabidopsis* AGOs. Thus, the structural and functional differences of AtAGO proteins appear to be translated to the expanded *Brachypodium* family. For two of the three clades, the number of BdAGO and AtAGO family members was equivalent. By contrast, the AtAGO1/5/10 clade was highly expanded in *Brachypodium*, with four members (BdAGO9/11/12/15/16) grouped with AtAGO1. This member of the *Arabidopsis* family is associated with a range of functions, including processing of dsRNA from transgenes and exogenous sources, and RNAs involved in ckRNAi (Vaucheret et al., 2004; Weiberg et al., 2013). If the corresponding BdAGO members of this clade are found to have similar functions in PTGS-mediated transgene silencing, this information would be highly useful for developing RNAi-based protection strategies for cereal crops. AtAGO10 groups with the same BdAGOs, as expected considering the clade association with AtAGO1. AtAGO5, which is the third member of the AtAGO1/5/10 clade, groups with BdAGO5/6/7/10. By contrast, BdAGO1/2/3/4 were interspersed within the AtAGO4/6/8/9 clade, raising the possibility that these *Brachypodium* proteins are involved in TGS.

As displayed in our domain visualization (Figure 3), all 16 BdAGOs have a predicted PAZ domain. In AGOs, this domain recognizes the 3' end of the guide sRNA molecule, made accessible to the hydrophobic pocket of this nucleotide-binding domain by the typical 2'-O-methyl modification of the final sugar (Lingel et al., 2003; Cenik and Zamore, 2011). The MID domain recognizes the 5' nucleotide of the sRNA, thus giving preference of an AGO protein into which the sRNA will be loaded (Frank et al., 2012). In *Arabidopsis*, sRNA with a 5' U are sorted into AtAGO1, while AtAGO2 and AtAGO4 load sRNAs with a 5' A and AtAGO5 loads sRNAs with a 5' C (Mi et al., 2008). Our SMART/Pfam domain architecture search failed to identify a MID domain in any of the BdAGOs grouped in the AtAGO4/6/8/9 clade (BdAGO1/2/3/4) and with the AtAGO2/3 (BdAGO13/14) (Table S2), although this domain was reported in these proteins in a different study (Mirzaei et al., 2014). The specificity of sRNA sorting into particular AGOs can be further determined by the recognition of the sRNA secondary structure/base pairing by a QF-V motif present in the PIWI domain (Zhang et al., 2014). All *Arabidopsis* AGOs have the conserved QF-V motif, as do all 16 BdAGOs (Figure 5). The DEDD/H catalytic tetrad in the PIWI domain is also present in all but one of the BdAGOs. These active-site residues are critical for the RNase H-like endonuclease (slicer) activity exhibited by certain AGOs, which mediates sequence-specific cleavage of the target transcript. AtAGO1, AtAGO2, AtAGO4, AtAGO7, and AtAGO10 have been shown to have endonucleolytic activity toward target RNAs (Fang and Qi, 2016). Originally identified as a catalytic triad consisting of the residues DDH in most AtAGOs, but DDD in AtAGO2 and AtAGO3 (Höck and Meister, 2008), studies of yeast AGO revealed the importance of an invariant glutamate (E)

residue, creating a catalytic tetrad (Nakanishi et al., 2012). This E residue is conserved in all *Arabidopsis* AGOs (Zhang et al., 2014). Consistent with these findings, MSA visualization of the BdAGO PIWI domains indicated that the majority of BdAGO proteins have the DEDH tetrad, except BdAGO13 and BdAGO14, which like their closest homologues AtAGO2 and AtAGO3, contain the DEDD tetrad (Figure 5). The only exception is BdAGO1, which is a short protein that terminates after 624 residues and lacks the last catalytic residue of the tetrad. Without the conserved catalytic residues, a specific AGO protein might induce gene silencing through means other than cutting, but Höck and Meister (2008) also discuss that the presence of a conserved catalytic triad does not mean the protein indeed has endonuclease activity. If an AGO does not display endonuclease activity, it may mediate PTGS via translation inhibition of the target RNA (Carthew and Sontheimer, 2009). Interestingly, the L1 and L2 linkers are predicted in all 16 BdAGOs as well (Table S2).

3D structure visualizations of all BdAGOs (Figures 6, 7 and S1) reinforce the conservation of the PAZ, PIWI, and MID domains (when predicted by SMART) and the catalytic tetrad residues in proximity within the PIWI domain (magenta spheres). The differences in the folding and looping linker regions within a certain group, relative to *Arabidopsis* AGOs, are shown in the model visualizations. Furthermore, the similarity between the 3D structures of BdAGOs that were predicted either to contain or lack a MID domain by the SMART/Pfam domain architecture search (e.g., Figure 6) reinforces the importance of comparing entire 3D models in order to gain insight into structure/function conservation. These structures are based on templates with better-known functional specificity and thus hint at the functions of the orthologs in Bd. As shown in Table S4, the templates used for modeling are based on either Argonaute 1 or Argonaute 2, with varying coverage and confidence values.

To assess the expression levels and locations of BdAGO and BdDCL family members, we analyzed the microarray-based expression data in the *B. distachyon* eFP browser (Table 3). Expression of *BdAGO* genes was observed in all four tissues assayed, although to varying extents. The expression patterns across the gene families indicate potential for functional redundancy. Notably, all members of the AtAGO1 clade (*BdAGO9/11/12/15/16*) show high and intermediate levels of expression in stem nodes and root tissue, respectively, while the *BdAGO1/2/3/4* proteins generally display high expression in stem nodes and seeds. Analysis of *BdDCL* gene expression revealed that most members of this family are highly expressed in stem nodes and/or seeds. *In vivo* experimental approaches are necessary to decipher whether the apparent co-expression of these genes indicates specific compartmentalization or complete/partial redundancy in the various RNAi processes, including environmental RNAi and ckRNAi pathways.

Finally, we used STRING to predict the interactome for members of the BdAGO family. This analysis indicates that all BdAGOs interact with three proteins (Table S6), as was expected because of the domain conservation within the family. In addition, several potential interactors were identified for BdAGO9, based on co-expression or experimental data (Figure S2, Table S6). Of these, DNA-directed RNA polymerase V subunit 1 was previously

shown to co-localize with, and possibly directly bind to, AtAGO4 via a so-called “Ago hook” (GW-rich domain), in order to facilitate the recruitment of AGO4 to chromatin to mediate TGS (El-Shami et al., 2007; Fang and Qi, 2016). Poulsen et al. (2013) have discussed that the binding of GW interactors to AGO make the loop with the E residue of the catalytic tetrad unavailable to the otherwise rigid DDD/H triad within the PIWI domain, thus offering an explanation of how the slicing activity is prevented in cases of silencing by translational inhibition. Moreover, GW containing proteins Needed for RDR2-independent DNA methylation and Silencing Defective 3 have been indicated in pathways bringing DNA/chromatin silencing together with RNAi proteins (Garcia et al., 2012; Pontier et al., 2012). Protein co-expression and interaction studies *in vivo* are necessary to confirm the identity and locations of these putative BdAGO interacting proteins. Due to the stringency of the prediction (excluding text mining data) and the lack of knowledge about the *Brachypodium* RNAi machinery, we were unable to predict additional interactions or to detect RNAi-related proteins that are known to interact with members of the AGO family in *Arabidopsis*. These include DCLs, HEN1 (involved in the methylation of sRNA 3' ends to prevent degradation), RDRs (RNA-dependent RNA polymerases that synthesize dsRNAs from single-stranded RNAs), and HSP90, the heat shock protein that binds to AtAGO1 and AtAGO4 to aid the loading of the sRNAs and RISC assembly (Fang and Qi, 2016; Nakanishi, 2016). Moreover, the predicted localization of the BdAGOs places the majority of them in the cytosol, except for BdAGO3 and BdAGO14, which are predicted to localize in the nucleus (Table S7). From what is known about *Arabidopsis* AGOs, AtAGO1 is proposed to have a localization in the nucleus and the cytoplasm (Vaucheret, 2008), while AtAGO4 localizes to the nuclear Cajal bodies (Höck and Meister, 2008).

In sum, based on *in silico* prediction, our data provide the first detailed functional insight into the AGO and DCL protein families in *Brachypodium*. In the context of plant–microbe interactions and ckRNAi, the *Brachypodium* orthologs of AtAGO1 are of special interest because microbial sRNAs are shown to be loaded onto AtAGO1 (Weiberg et al., 2013). Our predictions indicate a clade of BdAGOs structurally similar to AtAGO1, consisting of BdAGO9/11/12/15/16 (Figure 7). Elaborating on such similarities with the well-established clades of *Arabidopsis* AGOs and DCLs is a valuable basis for testing the hypothesis that BdAGO9/11/12/15/16 proteins are required for exogenous and endogenous dsRNA processing in HIGS, SIGS, and bidirectional ckRNAi in the grass model. Beyond what is

predictable by *in silico* analysis, more data on expression patterns and interacting proteins are needed to further understand the role of these pillar proteins of RNAi pathways in cereals.

## DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/Supplementary Files.

## AUTHOR CONTRIBUTIONS

EŠ, SZ, and KHK wrote the text. EŠ and SZ performed analysis and designed the figures.

## FUNDING

This work was funded by the Deutsche Forschungsgemeinschaft in the program GRK2355 to KHK and by of the European Union in the Marie Skłodowska-Curie Actions CEREALPATH to KHK and SZ.

## ACKNOWLEDGMENTS

We thank Sebastien Santini—CNRS/AMU IGS UMR7256, for administration of Phylogeny.fr, used for phylogenetic analysis within this study. We are very thankful to D'Maris Dempsey for her assistance in text editing. This manuscript summarizes EŠ's and KHK's contribution during the Organisation for Economic Cooperation and Development Conference on RNAi-Based Pesticides, which was sponsored by the Organisation for Economic Cooperation and Development Co-operative Research Programme: Biological Resource Management for Sustainable Agricultural Systems whose financial support made it possible for the author KHK to participate in the conference.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2019.01332/full#supplementary-material>

## REFERENCES

- Abdellatef, E., Will, T., Koch, A., Imani, J., Vilcinskas, A., and Kogel, K.-H. (2015). Silencing the expression of the salivary sheath protein causes transgenerational feeding suppression in the aphid *Sitobion avenae*. *Plant Biotechnol. J.* 13 (6), 849–857. doi: 10.1111/pbi.12322
- Andrade, E. C., and Hunter, W. B. (2016). “RNA interference—natural gene-based technology for highly specific pest control (HiSPeC)”. *RNA interference (Ch. 19)*. Ed. I. Y. Abdurakhmonov (IntechOpen), 391–409. doi: 10.5772/61612
- Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55 (4), 539–552. doi: 10.1080/10635150600755453
- Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116 (2), 281–297. doi: 10.1016/S0092-8674(04)00045-5
- Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27 (3), 343–350. doi: 10.1093/bioinformatics/btq662
- Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., et al. (2015). The *aArabidopsis* information resource: Making and mining the “gold standard” annotated reference plant genome. *Genesis* 53 (8), 474–485. doi: 10.1002/dvg.22877
- Berman, H., Gilliland, G. M., Weissig, H., Shindyalov, I. N., Westbrook, J., Bourne, P. E., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28 (1), 235–242. doi: 10.1093/nar/28.1.235

- Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *EMBO J.* 17 (1), 170–180. doi: 10.1093/emboj/17.1.170
- Boland, A., Huntzinger, E., Schmidt, S., Izaurre, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc. Natl. Acad. Sci.* 108 (26), 10466 LP–10471. doi: 10.1073/pnas.1103946108
- 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
- Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16 (12), 727–741. doi: 10.1038/nrm4085
- Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R., and Zhu, J.-K. (2005). Endogenous siRNAs derived from a pair of natural cis-Antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* 123 (7), 1279–1291. doi: 10.1016/j.cell.2005.11.035
- Brown, N. P., Leroy, C., and Sander, C. (1998). MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* 14 (4), 380–381. doi: 10.1093/bioinformatics/14.4.380
- Burley, S. K., Yang, H., Tan, L., Sala, R., Hudson, B. P., Bhikadiya, C., et al. (2018). RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* 47 (D1), D464–D474. doi: 10.1093/nar/gky1004
- Cai, Q., He, B., Kogel, K.-H., and Jin, H. (2018). Cross-kingdom RNA trafficking and environmental RNAi — nature's blueprint for modern crop protection strategies. *Curr. Opin. Microbiol.* 46, 58–64. doi: 10.1016/j.mib.2018.02.003
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10 (1), 421. doi: 10.1186/1471-2105-10-421
- Carthew, R. W., and Sontheimer, E. J. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell* 136 (4), 642–655. doi: 10.1016/j.cell.2009.01.035
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17 (4), 540–552. doi: 10.1093/oxfordjournals.molbev.a026334
- Cenik, E. S., and Zamore, P. D. (2011). Argonaute proteins. *Curr. Biol.* 21 (12), R446–R449. doi: 10.1016/j.cub.2011.05.020
- Chevenet, F., Brun, C., Bañués, A.-L., Jacq, B., and Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7, 439. doi: 10.1186/1471-2105-7-439
- Coleman, A. D., Wouters, R. H. M., Mugford, S. T., and Hogenhout, S. A. (2014). Persistence and transgenerational effect of plant-mediated RNAi in aphids. *J. Exp. Bot.* 66 (2), 541–548. doi: 10.1093/jxb/eru450
- Dereeper, A., Audic, S., Claverie, J.-M., and Blanc, G. (2010). BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol. Biol.* 10, 8. doi: 10.1186/1471-2148-10-8
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36 (Web Server issue), W465–W469. doi: 10.1093/nar/gkn180
- Dunoyer, P., Himber, C., and Voinnet, O. (2005). DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nat. Genet.* 37, 1356. doi: 10.1038/ng1675
- Dutta, T. K., Papolu, P. K., Banakar, P., Choudhary, D., Sirohi, A., and Rao, U. (2015). Tomato transgenic plants expressing hairpin construct of a nematode protease gene conferred enhanced resistance to root-knot nematodes. *Front. Microbiol.* 6, 260. doi: 10.3389/fmicb.2015.00260
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32 (5), 1792–1797. doi: 10.1093/nar/gkh340
- Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A. D., Greene, E. M., Hannon, G. J., et al. (2012). The structure of human argonaute-2 in complex with miR-20a. *Cell* 150 (1), 100–110. doi: 10.1016/j.cell.2012.05.017
- El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., et al. (2007). Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev.* 21 (20), 2539–2544. doi: 10.1101/gad.451207
- Faehle, C. R., Elkayam, E., Haase, A. D., Hannon, G. J., and Joshua-Tor, L. (2013). The making of a slicer: activation of human Argonaute-1. *Cell Rep.* 3 (6), 1901–1909. doi: 10.1016/j.celrep.2013.05.033
- Fang, X., and Qi, Y. (2016). RNAi in Plants: an argonaute-centered view. *Plant Cell.* 28 (2), 272 LP–272285. doi: 10.1105/tpc.15.00920
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391 (6669), 806–811. doi: 10.1038/35888
- Frank, F., Hauver, J., Sonenberg, N., and Nagar, B. (2012). Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. *EMBO J.* 31 (17), 3588 LP–3595. doi: 10.1038/emboj.2012.204
- Fukudome, A., and Fukuhara, T. (2017). Plant dicer-like proteins: double-stranded RNA-cleaving enzymes for small RNA biogenesis. *J. Plant Res.* 130 (1), 33–44. doi: 10.1007/s10265-016-0877-1
- Garcia, D., Garcia, S., Pontier, D., Marchais, A., Renou, J. P., Lagrange, T., et al. (2012). Ago hook and RNA helicase motifs underpin dual roles for SDE3 in antiviral defense and silencing of nonconserved intergenic regions. *Mol. Cell* 48 (1), 109–120. doi: 10.1016/j.molcel.2012.07.028
- Gascioli, V., Mallory, A. C., Bartel, D. P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol.* 15 (16), 1494–1500. doi: 10.1016/j.cub.2005.07.024
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40 (Database issue), D1178–D1186. doi: 10.1093/nar/gkr944
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., et al. (2010). A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Res.* 38 (suppl\_2), W695–W699. doi: 10.1093/nar/gkq313
- Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520
- Harvey, J. J. W., Lewsey, M. G., Patel, K., Westwood, J., Heimstädt, S., Carr, J. P., et al. (2011). An antiviral defense role of AGO2 in plants. *PLoS One* 6 (1), e14639. doi: 10.1371/journal.pone.0014639
- Head, G. P., Carroll, M. W., Evans, S. P., Rule, D. M., Willse, A. R., Clark, T. L., et al. (2017). Evaluation of SmartStax and SmartStax PRO maize against western corn rootworm and northern corn rootworm: efficacy and resistance management. *Pest Manag. Sci.* 73 (9), 1883–1899. doi: 10.1002/ps.4554
- Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., et al. (2005). Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. *Plant Mol. Biol.* 57 (2), 173–188. doi: 10.1007/s11103-004-6853-5
- Höck, J., and Meister, G. (2008). The Argonaute protein family. *Genome Biol.* 9 (2), 210. doi: 10.1186/gb-2008-9-2-210
- Hoof, R. W. W., Sander, C., and Vriend, G. (1997). Objectively judging the quality of a protein structure from a Ramachandran plot. *Bioinformatics* 13 (4), 425–430. doi: 10.1093/bioinformatics/13.4.425
- Hoof, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996). Errors in protein structures. *Nature* 381 (6580), 272. doi: 10.1038/381272a0
- Koch, A., Biedenkopf, D., Furch, A., Weber, L., Rossbach, O., Abdellatif, E., et al. (2016). An RNAi-based control of fusarium graminearum infections through spraying of long dsRNAs involves a plant passage and is controlled by the fungal silencing machinery. *PLOS Pathog.* 12 (10), e1005901. doi: 10.1371/journal.ppat.1005901
- Koch, A., Kumar, N., Weber, L., Keller, H., Imani, J., and Kogel, K.-H. (2013). Host-induced gene silencing of cytochrome P450 lanosterol C14 $\alpha$ -demethylase-encoding genes confers strong resistance to Fusarium species. *Proc. Natl. Acad. Sci.* 110 (4), 19324 LP–19329. doi: 10.1073/pnas.1306373110
- Koch, A., Stein, E., and Kogel, K.-H. (2018). RNA-based disease control as a complementary measure to fight Fusarium fungi through silencing of the azole target Cytochrome P450 Lanosterol C-14  $\alpha$ -Demethylase. *Eur. J. Plant Pathol.* 152 (4), 1003–1010. doi: 10.1007/s10658-018-1518-4
- Laskowski, R., MacArthur, M. W., Moss, D. S., and Thornton, J. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. doi: 10.1107/S0021889892009944
- Letunic, I., and Bork, P. (2017). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46 (D1), D493–D496. doi: 10.1093/nar/gkx922
- Lingel, A., Simon, B., Izaurre, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* 426 (6965), 465–469. doi: 10.1038/nature02123

- Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol.* 11, 576. doi: 10.1038/nsmb777
- Liu, C., Axtell, M. J., and Fedoroff, N. V. (2012). The helicase and RNaseIIIa domains of Arabidopsis Dicer-Like1 modulate catalytic parameters during microRNA biogenesis. *Plant Physiol.* 159 (2), 748–758. doi: 10.1104/pp.112.193508
- Liu, L., Zhang, Z., Mei, Q., and Chen, M. (2013). PSI: a comprehensive and integrative approach for accurate plant subcellular localization prediction. *PLOS ONE* 8 (10), e75826. doi: 10.1371/journal.pone.0075826
- Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., et al. (2015). IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics.* 31 (20), 3359–3361. doi: 10.1093/bioinformatics/btv362
- Liu, Y., Eshyunina, D., Olovnikov, I., Teplova, M., Kulbachinskiy, A., Aravin, A. A., et al. (2018). Accommodation of helical imperfections in rhodobacter sphaeroides argonaute ternary complexes with guide RNA and target DNA. *Cell Rep.* 24 (2), 453–462. doi: 10.1016/j.celrep.2018.06.021
- 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 (10), 2442–2450. doi: 10.1016/j.febslet.2006.03.072
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., et al. (2008). Sorting of Small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* 133 (1), 116–127. doi: 10.1016/j.cell.2008.02.034
- Mirzaei, K., Bahramnejad, B., Shamsifard, M. H., and Zamani, W. (2014). In silico identification, phylogenetic and bioinformatic analysis of argonaute genes in plants. *Int. J. Genomics.* 2014, 967461. doi: 10.1155/2014/967461
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins Struct. Funct. Bioinf.* 12 (4), 345–364. doi: 10.1002/prot.340120407
- Mukherjee, K., Campos, H., and Kolaczowski, B. (2013). Evolution of animal and plant dicers: early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. *Mol. Biol. Evol.* 30 (3), 627–641. doi: 10.1093/molbev/mss263
- Nakanishi, K. (2016). Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip. Rev. RNA.* 7 (5), 637–660. doi: 10.1002/wrna.1356
- Nakanishi, K., Weinberg, D. E., Bartel, D. P., and Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. *Nature* 486, 368. doi: 10.1038/nature11211
- Nielsen, M., Lundegaard, C., Lund, O., and Petersen, T. N. (2010). CPHmodels-3.0—remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res.* 38 (Web Server issue), W576–W581. doi: 10.1093/nar/gkq535
- Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., et al. (2010). HIGS: Host-Induced Gene Silencing in the obligate biotrophic fungal pathogen *blumeria graminis*. *Plant Cell.* 22 (9), 3130 LP–3141. doi: 10.1105/tpc.110.077040
- Olmedo-Monfil, V., Durán-Figueroa, N., Arteaga-Vázquez, M., Demesa-Arévalo, E., Autran, D., Grimanelli, D., et al. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. *Nature* 464, 628. doi: 10.1038/nature08828
- Park, M. S., Phan, H.-D., Busch, F., Hinckley, S. H., Brackbill, J. A., Wysocki, V. H., et al. (2017). Human Argonaute3 has slicer activity. *Nucleic Acids Res.* 45 (20), 11867–11877. doi: 10.1093/nar/gkx916
- Parker, G. S., Maity, T. S., and Bass, B. L. (2008). dsRNA binding properties of RDE-4 and TRBP reflect their distinct roles in RNAi. *J. Mol. Biol.* 384 (4), 967–979. doi: 10.1016/j.jmb.2008.10.002
- Patel, P., Mathioni, S., Kakrana, A., Shatky, H., and Meyers, B. (2018). Reproductive phasiRNAs in grasses are compositionally distinct from other classes of small RNAs. *New Phytol.* 220, 851–864. doi: 10.1111/nph.15349
- Pontier, D., Picart, C., Roudier, F., Garcia, D., Lahmy, S., Azevedo, J., et al. (2012). NERD, a Plant-Specific GW Protein, Defines an Additional RNAi-Dependent Chromatin-Based Pathway in Arabidopsis. *Mol. Cell* 48 (1), 121–132. doi: 10.1016/j.molcel.2012.07.027
- Poulsen, C., Vaucheret, H., and Brodersen, P. (2013). Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell.* 25 (1), 22–37. doi: 10.1105/tpc.112.105643
- Qin, H., Chen, F., Huan, X., Machida, S., Song, J., and Yuan, Y. A. (2010). Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. *RNA* 16 (3), 474–481. doi: 10.1261/rna.1965310
- Qu, F., Ye, X., and Morris, T. J. (2008). Arabidopsis DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc. Natl. Acad. Sci. U. S. A.* 105 (38), 14732–14737. doi: 10.1073/pnas.0805760105
- Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* 7 (1), 95–99. doi: 10.1016/S0022-2836(63)80023-6
- Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2011). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* 9, 173. doi: 10.1038/nmeth.1818
- Rhee, S. Y., Beavis, W., Berardini, T. Z., Chen, G., Dixon, D., Doyle, A., et al. (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Res.* 31 (1), 224–228. doi: 10.1093/nar/gkg076
- Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16 (6), 276–277. doi: 10.1016/S0168-9525(00)02024-2
- Savary, S., Ficke, A., Aubertot, J.-N., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Secur.* 4, 519–537. doi: 10.1007/s12571-012-0200-5
- Schauer, S. E., Jacobsen, S. E., Meinke, D. W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in Arabidopsis development. *Trends Plant Sci.* 7 (11), 487–491. doi: 10.1016/S1360-1385(02)02355-5
- Schrödinger, L. L. C. (2010). The PyMOL Molecular Graphics System, Version~1.3.
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci.* 95 (11), 5857 LP–5864. doi: 10.1073/pnas.95.11.5857
- Secic, E., Sutkovic, J., and Abdel Gawwad, M. (2015). Interactome analysis and docking sites prediction of radiation sensitive 23 (RAD 23) proteins in Arabidopsis thaliana. *Curr. Proteomics* 12 (1), 28–44. doi: 10.2174/1570164612666150225234240
- Sibout, R., Proost, S., Hansen, B. O., Vaid, N., Giorgi, F. M., Ho-Yue-Kuang, S., et al. (2017). Expression atlas and comparative coexpression network analyses reveal important genes involved in the formation of lignified cell wall in *Brachypodium distachyon*. *New Phytol.* 215 (3), 1009–1025. doi: 10.1111/nph.14635
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7 (1), 539. doi: 10.1038/msb.2011.75
- Song, M. S., and Rossi, J. J. (2017). Molecular mechanisms of Dicer: endonuclease and enzymatic activity. *Biochem. J.* 474, 10) 1603–1618. doi: 10.1042/BCJ20160759
- Song, X., Li, P., Zhai, J., Zhou, M., Ma, L., Liu, B., et al. (2012). Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J.* 69 (3), 462–474. doi: 10.1111/j.1365-313X.2011.04805.x
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi: 10.1093/nar/gky1131
- The International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763–768. doi: 10.1038/nature08747
- Tomilov, A. A., Tomilova, N. B., Wroblewski, T., Michelmore, R., and Yoder, J. I. (2008). Trans-specific gene silencing between host and parasitic plants. *Plant J.* 56 (3), 389–397. doi: 10.1111/j.1365-313X.2008.03613.x
- Vaucheret, H. (2008). Plant ARGONAUTES. *Trends Plant Sci.* 13 (7), 350–358. doi: 10.1016/j.tplants.2008.04.007
- Vaucheret, H., Vazquez, F., Crété, P., and Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* 18 (10), 1187–1197. doi: 10.1101/gad.1201404
- Vogel, J. P., Garvin, D. F., Leong, M. O., and Hayden, D. M. (2006). Agrobacterium-mediated transformation and inbred line development in the model grass *Brachypodium distachyon*. *Plant Cell Tissue Organ Culture* 84, 199–211. doi: 10.1007/s11240-005-9023-9

- Wang, B., Sun, Y., Song, N., Zhao, M., Liu, R., Feng, H., et al. (2017b). Puccinia striiformis f. sp. tritici microRNA-like RNA 1 (Pst-milR1), an important pathogenicity factor of Pst, impairs wheat resistance to Pst by suppressing the wheat pathogenesis-related 2 gene. *New Phytol.* 215 (1), 338–350. doi: 10.1111/nph.14577
- Wang, H.-L. V., Dinwiddie, B. L., Lee, H., and Chekanova, J. A. (2015). Stress-induced endogenous siRNAs targeting regulatory intron sequences in *Brachypodium*. *RNA (New York, N.Y.)* 21 (2), 145–163. doi: 10.1261/rna.047662.114
- Wang, M., Weiberg, A., Dellota, E., Yamane, D., and Jin, H. (2017a). Botrytis small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi. *RNA Biol.* 14 (4), 421–428. doi: 10.1080/15476286.2017.1291112
- Wang, M., Weiberg, A., Lin, F.-M., Thomma, B. P. H. J., Huang, H.-D., and Jin, H. (2016). Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nat. Plants* 2, 16151. doi: 10.1038/nplants.2016.151
- Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G. S., Tuschl, T., et al. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 461, 754. doi: 10.1038/nature08434
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46 (W1), W296–W303. doi: 10.1093/nar/gky427
- 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 (6154), 118 LP–118123. doi: 10.1126/science.1239705
- Willkomm, S., Zander, A., Gust, A., and Grohmann, D. (2015). A prokaryotic twist on argonaute function. *Life (Basel, Switzerland)* 5 (1), 538–553. doi: 10.3390/life5010538
- Wilson, R. C., and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* 42, 217–239. doi: 10.1146/annurev-biophys-083012-130404
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2 (8), e718. doi: 10.1371/journal.pone.0000718
- 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 (5), e104. doi: 10.1371/journal.pbio.0020104
- Yang, Y., and Zhou, Y. (2008). Specific interactions for ab initio folding of protein terminal regions with secondary structures. *Proteins Struct. Funct. Bioinf.* 72 (2), 793–803. doi: 10.1002/prot.21968
- You, C., Cui, J., Wang, H., Qi, X., Kuo, L.-Y., Ma, H., et al. (2017). Conservation and divergence of small RNA pathways and microRNAs in land plants. *Genome Biol.* 18, 158. doi: 10.1186/s13059-017-1291-2
- Zha, X., Xia, Q., and Adam Yuan, Y. (2012). Structural insights into small RNA sorting and mRNA target binding by Arabidopsis argonaute mid domains. *FEBS Lett.* 586 (19), 3200–3207. doi: 10.1016/j.febslet.2012.06.038
- Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human dicer and bacterial RNase III. *Cell* 118 (1), 57–68. doi: 10.1016/j.cell.2004.06.017
- Zhang, H., Xia, R., Meyers, B. C., and Walbot, V. (2015). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Curr. Opin. Plant Biol.* 27, 84–90. doi: 10.1016/j.pbi.2015.06.011
- Zhang, T., Zhao, Y.-L., Zhao, J.-H., Wang, S., Jin, Y., Chen, Z.-Q., et al. (2016). Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat. Plants* 2, 16153. doi: 10.1038/nplants.2016.153
- Zhang, X., Niu, D., Carbonell, A., Wang, A., Lee, A., Tun, V., et al. (2014). ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nat. Commun.* 5, 5468. doi: 10.1038/ncomms6468
- Zheng, X., Zhu, J., Kapoor, A., and Zhu, J. (2007). Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO J.* 26 (6), 1691 LP–1701. doi: 10.1038/sj.emboj.7601603
- 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 (2), 242–256. doi: 10.1016/j.cell.2011.03.024
- Zilberman, D., Cao, X., and Jacobsen, S. E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 2995607, 716 LP–716719. doi: 10.1126/science.1079695

**Conflict of Interest:** 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 © 2019 Šečić, Zanini and Kogel. 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) and the copyright owner(s) 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.