RESEARCH NOTE

The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution

KANG ZHAO[1], HUALIN ZHAO[1], ZHU CHEN[1], LIN FENG[1], JIE REN[1], RONGHAO CAI[2] and YAN XIANG[1,2*]

[1]*Laboratory of Modern Biotechnology, School of Forestry and Landscape Architecture, and* [2]*Key Laboratory of Crop Biology of Anhui Province, School of Life Sciences, Anhui Agricultural University, Hefei 230036, People's Republic of China*

[Zhao K., Zhao H., Chen Z., Feng L., Ren J., Cai R. and Xiang Y. 2015 The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution. *J. Genet.* **94**, 317–321]

Introduction

RNA silencing is guided by numerous kinds of small silencing RNAs, such as micro RNAs (miRNAs), short-interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs) (Ding 2010). It is widely implicated as an important molecular mechanism to tune the expression of many genes in eukaryotic organisms. In this study, a total of five *PtDCL*, 15 *PtAGO* and nine *PtRDR* genes were identified in Poplar. These genes were categorized into four groups per family based on phylogenetic analysis and distributed across 12 chromosomes. Two *PtAGOs* and four *PtRDRs* were totally activated under specific stress treatment. This detailed investigation provides a foundation for future functional genomics studies of transcriptional silencing-associated gene families in Poplar.

Materials and methods

Phylogenetic analysis and identification of conserved motifs

In plants, small RNAs have been widely implicated in various developmental events and as guide RNAs in many gene silencing pathways (Hamilton and Baulcombe 1999; Mochizuki *et al.* 2002; Hall *et al.* 2003). These RNAs are usually generated by the activities of Dicers (DCLs), Argonautes (AGOs) and RNA-dependent RNA polymerases (RDRs), sometimes referred to as core proteins, which mediate RNA interference. Among these proteins, the plant DCL proteins, which are similar to RNaseIII endonuclease, mainly process long double-stranded RNAs into mature small RNAs (Bernstein *et al.* 2001; Hammond *et al.* 2001). The AGO proteins belong to the core components of RNAi effector complexes and play central roles in RNA silencing (Moazed 2009). All AGO proteins share domain structures include an N-terminal, PAZ, MID and C-terminal PIWI domains. In the signal amplification stage, RDR enzymes are responsible for synthesis of dsRNAs using the RNA as the template, and then the dsRNAs are cut into secondary small RNAs by DCL and give rise to a new round of RNA silencing (Djupedal and Ekwall 2009).

The Poplar genome database (http://www.Phytozome.net/) was searched to identify DCL, AGO and RDR proteins using BLASTP algorithms with the published *Arabidopsis* DCL, AGO and RDR protein sequences as query sequences. Candidate gene sequences were analysed separately by Pfam (http://pfam.sanger.ac.uk/search, $P = 0.001$) (Finn *et al.* 2006) and SMART (http://smart.embl-heidelberg.de/). Physicochemical parameters of each gene were calculated using ExPASy (http://www.expasy.org/tools/) (Gasteiger *et al.* 2003).

Phylogenetic trees were constructed using MEGA 6.0 with the neighbour–joining method (bootstrap = 1000). Structural motif annotation was performed using the MEME program with the following parameters: number of repetitions, any; maximum number of motifs, 20 and the optimum motif width, between 6 and 50 residues.

The physical locations of *DCL*, *AGO* and *RDR* genes were determined by initially confirming the starting position of the candidate genes from the phytozome database. Data from the recently identified duplicated blocks were obtained (Tuskan *et al.* 2006), and the segment duplication coordinates of the target genes were obtained from the Poplar JGI browser (http://genome.jgi-psf.org/). Genes separated by five or fewer gene loci in a 60-kb region were considered to be tandem duplicates.

*For correspondence. E-mail: xiangyanahau@sina.com.

Keywords. Dicer-like; Argonaute; RNA-dependent RNA polymerase; phylogenetic analysis; expression pattern; *Populus*.

Expression patterns of Poplar DCL, AGO and RDR genes

The gene expression omnibus database GSE12152 (GC-RMA algorithm) and VMatch tool at PLEXdb (http://www.plantgdb.org/) were used to prepare a compendium of transcriptome profiles for shoot organogenesis in Poplar. The tissue-specific transcript data (root, leaf, flower, xylem, http://bar.utoronto.ca/efppop/cgi-bin/efpWeb.cgi) of 29 Poplar genes were genewise normalized and the heat-map was drawn using Cluster 3.0 with the average linkage program (Cheng *et al.* 2013).

Plant stress treatment and semi-quantitative RT-PCR analysis

The plant material used was Poplar (Nanlin-95) tissue culture seedlings, whose rapid propagation system has been established previously. Three-week-old seedlings were subjected to three abiotic stress treatments. Drought stress was induced by 20% PEG-6000 (polyethylene glycerol) and seedling leaves were sampled at 24 h after the treatment. Seedling roots were submerged in 0.2 M NaCl solution for salt stress and seedling leaves were sampled at 6 h after the treatment. For cold stress, the seedlings were transferred to 4°C growth chamber for 30 min. Control seedlings were exposed to none of these treatments.

To examine the expression patterns of these predicted genes in Poplar and to further confirm their stress responsiveness to abiotic stresses, all 29 genes were subjected to semi-quantitative RT-PCR using specific primers designed using Primer 5.0 software. To adjust for RNA quality and differences in cDNA concentration, we amplified actin as an internal control with the following primers: PtActin-F (5′-TG AAGGAGAAACTTGCGTAT-3′) and PtActin-R (5′-GCA CAATGTTACCGTACAGAT-3′). These genes were amplified from first-strand cDNA using *Taq* polymerase (Promega, Madison, USA) on a thermal cycler (Tpersonal; Biometra, Germany), with the following profile: initial denaturation at 95°C for 10 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 51.6–57.8°C (table 1 in electronic supplementary material at http://www.ias.ac.in/jgenet/), for 45 s, polymerization at 72°C for 30 s and final elongation at 72°C for 10 min. Each PCR pattern was verified by triple replicate experiments; mixture without template was used as negative control and Poplar actin DNA fragment as positive control for each gene amplified. A 5-μL aliquot of the reaction was separated on 1% agarose gel.

Results and discussion

In total, five *DCL* genes, 15 *AGO* genes and nine *RDR* genes were identified in the Poplar genome. The candidate genes identified in this study are listed in table 2 in electronic supplementary material. The unrooted phylogenetic tree generated from the aligned full-length protein sequences of all candidate genes with those of *Arabidopsis*, rice and tomato, total protein sequences (figure 1). Based on phylogenetic analysis, these three gene families have been divided into four subfamilies. Most *PtDCL* and *AtDCL* genes were more closely related than *DCL*s in rice and tomato, most *PtAGO* genes shared more similarity with *SlAGO* genes. However, the DCL families in Poplar and *Arabidopsis* species are more likely to have descended from a common ancestor of the dicots. This implies that *DCL* and *AGO* genes originated after the divergence of monocots and dicots. Therefore, we infer that the two dicotyledons may have an analogous mode of selection and evolution.

Twenty-nine candidate genes were found to be distributed unevenly across all the 12 chromosomes in the Poplar genome (figure 1 in electronic supplementary material). One pair, *PtAGO2*/*3* were found in duplicated segments of the genome and shared 70.1% identity at the amino acid level between the partners. *PtRDR3*/*4*, *PtRDR6a*/*6b*, which shared 83.56% and 88.6% identity at amino acid level, respectively, appeared to represent segmental duplications. The identities of the other genes were not found located in tandem duplication.

MEME program clearly suggests that these motifs are well conserved and share major functional roles in the DCL, AGO and RDR proteins (figure 2 in electronic supplementary material). But there are some exceptions: in DCL2 subfamily, PtDCL2a lacked the dsRB domain in the C-terminal and PtDCL2b lacked the DExD domain compared to the DCL2 paralogues. All PtAGOs exhibited high similarity to AtAGO1; in the MEL1 and ZIPPY subfamilies, a subset of AGO members shared the absence or duplication of one or two motifs from the PIWI domain.

To analyse the expression profiles of *DCL*, *AGO* and *RDR* during shoot organogenesis in Poplar on a geneomewide scale, microarray datasets were utilized (table 3 in electronic supplementary material). Expression profile of almost all the genes in different families or subfamilies are shown in figure 2a. All *PtDCL* genes were highly expressed during the five phases of shoot organogenesis reaching their highest levels three days after callus induction. In addition, four *PtDCL* genes were consistently expressed at low levels. In PtAGO family, AGO1 subfamily showed high expression. It is noteworthy that PtAGO7 had a greater change in the amplitude during different stages.

Using the Poplar eFP Browser website data (table 4 in electronic supplementary material), we identified the expression of 29 genes in different tissues (figure 2b). No changes were detected in the expression of the PtDCL family genes in the xylem with the exception of *PtDCL2a* which was upregulated. *PtRDR1c* expression was upregulated in female catkins and downregulated in male catkins accounting for the differences in expressions between different organs.

Semiquantitative RT-PCR analyses on RNA isolated from Poplar leaves were performed. The results (figure 2c) revealed that these genes were differentially expressed in the leaves under either normal conditions (CK control) or stress conditions.

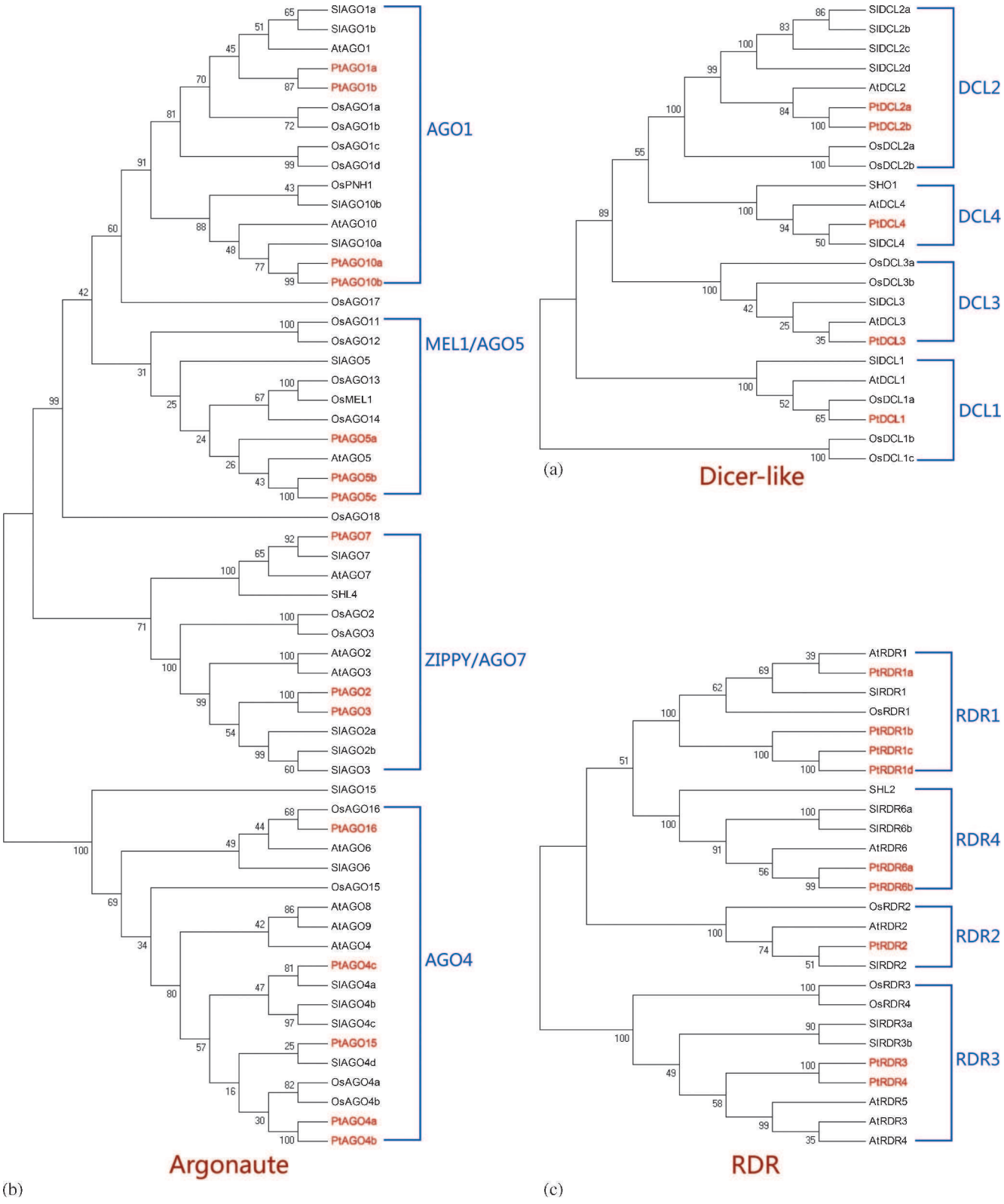

Figure 1. Unrooted NJ phylogenetic trees of Poplar, rice, tomato and *Arabidopsis* proteins; (a) *DCL*, (b) *AGO* and (c) RDR. Bootstrap values are shown for each clade. The Poplar DCLs, AGOs, or RDRs have been highlighted in red for each group. Each gene family is divided into four clades, which are shown in blue. Sequences of tomato, *Arabidopsis* and rice were downloaded from the Phytozome database.

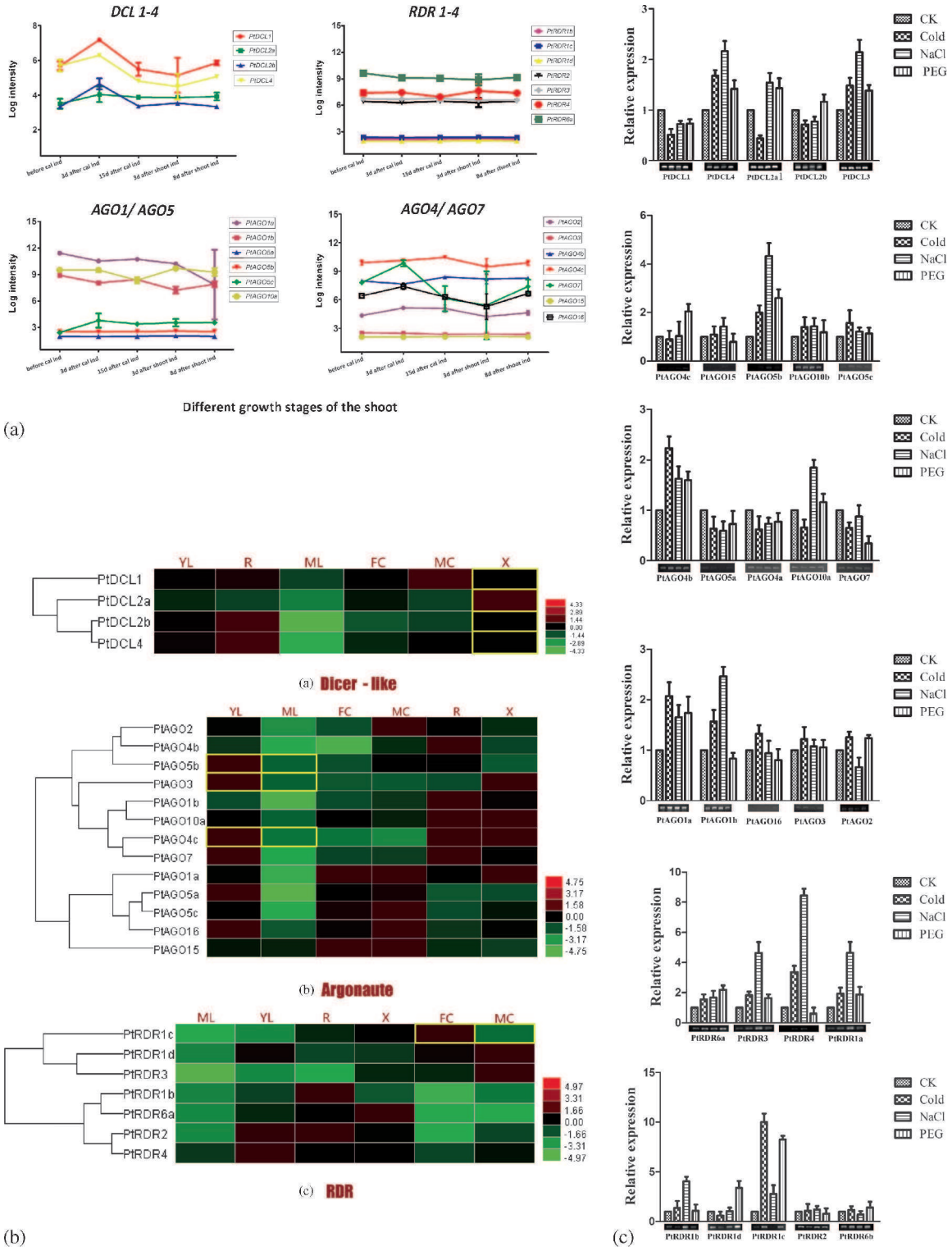

Figure 2. (a) Expression profiles of three family genes during five stages of shoot organogenesis (prior to callus induction, three days after callus induction, 15 days after callus induction, three days after the start of shoot induction and eight days after the start of shoot induction). The y-axis indicates the relative expression level and error bars represent standard deviation calculated based on two biological replicates. (b) The identified genes expression profiles have been analysed in vegetative tissues (young leaf, mature leaf, roots, female catkins, male catkins and xylem). Small squares of different colours represent different data. Red and green reflect gene upregulation and downregulation, respectively, of transcription under the same conditions. The absolute values are indicated by the colour intensity. (c) Semiquantitative RT-PCR analysis of expression levels of identified genes under abiotic stress treatments. Corresponding expression in each treatment obtained with semiquantitative RT-PCR is indicated below the x-axis. The expression levels are shown compared to the respective control (CK). Error bars represent the standard error.

Under cold treatment, 29 genes were significantly upregulated or downregulated, although no changes were detected for *PtAGO4c*/*5a*/*5b*/*15*/*16*. It should be noted that *PtAGO1a*/*1b* and *PtRDR1c*/*6a* were fully activated. Under high salinity conditions, there were no changes in the expression of *PtAGO4a*/*7*/*10a*/*10b*, but expression of *PtAGO1a*/*1b* and *PtRDR1a*/*1b*/*3* were fully activated.

The experimental results suggesting that these genes play important roles in plant RNAi regulation, especially those showing strong responses to the three abiotic stress conditions. In contrast, those genes (*PtAGO2*/*3*/*4c*/*5a*/*5b*/*15*/*16*, *PtRDR4*) that showed lower expression in response to stress treatments are likely to contribute to Poplar RNAi regulation by expressing only under specific conditions or in specific tissues other than seedling leaves. However, this conclusion remains to be further confirmed experimentally. Further, those genes (*PtAGO1a*/*1b*, *PtRDR1a*/*1b*/*1d*/*3*) exhibiting distinct expression patterns under different stress conditions might play vital roles in evolving specialized regulatory mechanisms in response to different abiotic stresses.

Acknowledgements

This work was supported by the National Natural Science Foundation (31370561) Specialized Research Fund for the Doctoral Programme of Higher Education (20133418110005), Anhui Provincial Natural Science Foundation (1308085MC36), Anhui Agricultural University Disciplinary Construction Foundation (XKTS2013001) and Opening Project of Anhui Taiping Experimental Station, ICBR.

References

Bernstein E., Caudy A. A., Hammond S. M. and Hannon G. J. 2001 Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366.

Cheng A.-X., Gou J.-Y., Yu X.-H., Yang H., Fang X., Chen X.-Y. and Liu C.-J. 2013 Characterization and ectopic expression of a populus hydroxyacid hydroxy cinnamoyl transferase. *Mol. Plant* **6**, 1889–1903.

Ding S.-W. 2010 RNA-based antiviral immunity. *Nat. Rev. Immunol.* **9**, 632–644.

Djupedal I. and Ekwall K. 2009 Epigenetics: heterochromatin meets RNAi. *Cell Res.* **19**, 282–295.

Finn R. D., Mistry J., Schuster-Böckler B., Griffiths-Jones S., Hollich V., Lassmann T. *et al.* 2006 Pfam: clans, web tools and services. *Nucleic Acids Res.* **34**, D247–D251.

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D. and Bairoch A. 2003 ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788.

Hall I. M., Noma K. and Grewal S. I. 2003 RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. *Proc. Natl. Acad. Sci. USA* **100**, 193–198.

Hamilton A. J. and Baulcombe D. C. 1999 A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952.

Hammond S. M., Boettcher S., Caudy A. A., Kobayashi R. and Hannon G. J. 2001 Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146–1150.

Moazed D. 2009 Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**, 413–420.

Mochizuki K., Fine N. A., Fujisawa T. and Gorovsky M. A. 2002 Analysis of a piwi related gene implicates small RNAs in genome rearrangement in Tetrahymena. *Cell* **110**, 689–699.

Tuskan G. A., Difazio S., Jansson S., Bohlmann J., Grigoriev I., Hellsten U. *et al.* 2006 The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604.

Received 25 March 2014, in final revised form 29 December 2014; accepted 5 January 2015
Unedited version published online: 7 January 2015
Final version published online: 8 June 2015