Theor Appl Genet (2013) 126:1273–1283
DOI 10.1007/s00122-013-2052-6

ORIGINAL PAPER

Structure, transcription and post-transcriptional regulation of the bread wheat orthologs of the barley cleistogamy gene *Cly1*

Shunzong Ning · Ning Wang · Shun Sakuma · Mohammad Pourkheirandish · Jianzhong Wu · Takashi Matsumoto · Takato Koba · Takao Komatsuda

Received: 2 October 2012 / Accepted: 19 January 2013 / Published online: 5 February 2013

Abstract The majority of genes present in the hexaploid bread wheat genome are present as three homoeologs. Here, we describe the three homoeologous orthologs of the barley cleistogamy gene *Cly1*, a member of the *AP2* gene family. As in barley, the wheat genes (designated *TaAP2-A*, *-B* and *-D*) map to the sub-telomeric region of the long arms of the group 2 chromosomes. The structure and pattern of transcription of the *TaAP2* homoeologs were similar to those of *Cly1*. Transcript abundance was high in the florets, and particularly in the lodicule. The *TaAP2* message was cleaved at its *miR172* target sites. The set of homoeolog-specific PCR assays developed will be informative for identifying either naturally occurring or induced cleistogamous alleles at each of the three wheat homoeologs. By combining such alleles via conventional crossing, it should be possible to generate a cleistogamous form of bread wheat, which would be advantageous both with respect to improving the level of the crop's resistance against the causative pathogen of fusarium head blight, and for controlling pollen-mediated gene flow to and from genetically modified cultivars.

Communicated by I. Godwin.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2052-6) contains supplementary material, which is available to authorized users.

S. Ning · N. Wang · S. Sakuma · M. Pourkheirandish · J. Wu · T. Matsumoto · T. Komatsuda (✉)
Plant Genome Research Unit, National Institute of Agrobiological Sciences (NIAS), 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan
e-mail: takao@affrc.go.jp

S. Ning · S. Sakuma · T. Koba
Graduate school of Horticulture, Chiba University, 648 Matsudo, Matsudo, Chiba 271-8510, Japan

Introduction

Along with rice and maize, wheat is the one of the world's major cereals. The diploid progenitor of the bread wheat (*Triticum aestivum* L.) A genome was *Triticum urartu* (Chapman et al. 1976) and that of the D genome was *Aegilops tauschii* (Kihara 1944); the progenitor of the B genome is yet to be established, but must have been an extant or extinct species belonging to the *Sitopsis* section of *Aegilops* (Riley et al. 1958; Petersen et al. 2006; Kilian et al. 2007). The hexaploidy of bread wheat has ensured that the majority of its single copy genes are represented as three similar (homoeologous) copies. A major consequence of the rather recent origin of the (AB) × D hybridization/polyploidization event (~10,000 years ago), followed by subsequent intensive farmer and breeder selection, is that the crop's genetic base is particularly narrow, implying a substantial level of vulnerability to fluctuations in environmental stress and pathogen variability (Nevo 2009, 2011; Fu and Somers 2009).

The development of the angiosperm flower passes through three phases, namely meristem induction, determination of meristem identity and determination of floral organ identity (Coen and Meyerowitz 1991). The leading genetic model underlying this development postulates five classes of homoeotic genes, referred to as A through E (Theissen and Saedler 2001). Class A, B and E genes specify petals in the second whorl, and in the dicotyledonous species *Arabidopsis thaliana*, they all, except for the class A gene *APETALA2* (*AP2*), encode a MADS box transcription factor. The monocotyledonous equivalent of the petal is the lodicule (Glover 2007), a pair of which forms at the base of the floret. Their expansion around the time of anthesis forces the lemma and palea apart, allowing first the emergence of the anthers and later that of the stigma (Heslop-Harrison and Heslop-Harrison 1996). If the lodicules fail to swell, gaping of the floret does not generally occur until well after fertilization. In this situation, the exertion of the anther filament is restricted, and the pollen is shed within the closed floret.

The barley (*Hordeum vulgare* L.) *cleistogamy 1* gene encodes an AP2 protein (Nair et al. 2010). The cleistogamous *cly1* allele features a synonymous single nucleotide change within a specific microRNA (miR172) site. In cultivars carrying this allele, the lodicules fail to develop normally, resulting in a cleistogamous phenotype. MicroRNAs are small (~22 nt) sequences which induce the degradation of a specific target mRNA and thereby inhibit translation. They have been implicated in a number of regulatory processes in both plant and animal cells (Bartel 2009). Their specificity relies on their sharing sequence complementarity with their target mRNA. A number of *AP2* genes are known to be regulated by miR172, largely via translational repression (Chen 2004; Chuck et al. 2007) induced following mRNA cleavage (Aukerman and Sakai 2003; Chen 2004). An analysis of allelic variation at *Cly1* among a substantial number of barley accessions has shown that the two cleistogamous alleles detected both include a sequence variant within their miR172 target site; while one of these alleles originated in Northern Europe, the other appears to have arisen in the Western Mediterranean region (Nair et al. 2010). One of the major significant advantages of the cleistogamous phenotype has been highlighted by its pleoitropic effect on resistance to the pathogen causing the destructive disease fusarium head blight (FHB) in barley (Hori et al. 2005; Sato et al. 2008) and wheat (Kubo et al. 2010).

Cleistogamy has been observed in wheat as well as in barley. A screen of *T. durum* accessions carried out by Sethi and Chhabra (1990) identified two sources, and further analysis showed that the trait was monogenic and recessive, was largely independent of the environment, and was caused by the under-development of the lodicules and the formation of a stiff perianth (Chhabra and Sethi 1991). While the mechanism underlying lodicule development in barley is now well understood, this is not the case for wheat. The aim of the present study was therefore to investigate whether the wheat orthologs of *cly1* are key genes driving lodicule development and flower opening. Here, we describe the identification of the three wheat homoeologs of *Cly1*, along with their structure and transcription profile. An analysis of miR172-guided mRNA cleavage is also provided.

Materials and methods

Plant materials and the measurement of lodicule size

The bread wheat lines used are detailed in Table 1. The cultivar ‘Shinchunaga’ shows a level of resistance to FHB (Ban and Suenaga. 2000) and has high cross-compatibility to rye (Ma et al. 1996). Grain of cv. Chinese Spring (‘CS’) and its derived sets of nullisomic-tetrasomic (NT) and ditelosomic (Dt) lines (Sears 1954, 1966; Sears and Sears 1978), as well as 15 deletion lines stocks involving one of the homoeologous group 2 chromosomes (Fig. 2) (Endo and Gill 1996) were used to determine intrachromosomal locations. All the material was autumn-sown in the field at Tsukuba, Japan. Immediately prior to anthesis, three spikes per line (still attached to the peduncle and including the flag leaf) were detached and maintained in 100 mg/l 2,4-D for 24 h at room temperature to preserve the swollen state

Table 1 Wheat materials used in this study

Lines	Origin or source	GenBank accession no.		
		TaAP2-A	*TaAP2-B*	*TaAP2-D*
Shinchunaga	Obtained from S. Taketa, Okayama Univ., Japan	AB749305	AB749306	AB749307
KU-515	Tibet, China	AB761163	AB761180	AB761194
Fukuho	Collection of Lab. of Genetics and Plant Breeding, Facul. of Horticulture, Chiba Univ., Japan	AB761164	AB761188	AB761189
Norin 61	Collection of Lab. of Genetics and Plant Breeding, Facul. of Horticulture, Chiba Univ., Japan	AB761165	AB761181	AB761190
KU-163	Collection of Col Agr. Hokkaido Univ., Japan	AB761172	AB761176	AB761191
KU-165	Correns, Germany	AB761159	AB761177	AB761192
KU-265	Collection of Lab. of Plant Breeding, Facul. of Agr., Kyoto Univ., Japan	AB761160	AB761179	AB761193
Chinese spring	Strain ID: LPGKU2269 from National Bioresource Project (NBRP) of Japan	AB749311	AB749312	AB749313

The lines starting with the code KU were kindly provided by National Bioresource Project (NBRP) of Japan of the lodicules. This measure facilitated the assessment of lodicule width and depth. The lemma from the first floret of a spikelet in the middle portion of each spike was removed to permit the imaging of the lodicules. The resulting images provided a means of estimating lodicule width and depth with the aid of Makijaku v1.1 software (cse.naro.affrc.go.jp/iwatah).

PCR primer design, amplification and amplicon sequencing

Genomic DNA was extracted from young leaves according to Komatsuda et al. (1998). Relevant PCR primers were designed based on the barley *Cly1* sequence using either Oligo 6 (W. Rychlick, National Bioscience, Plymouth, MN, USA) or DNAMAN v6.0 (Lynnon Biosoft, Quebec, Canada) software. Amplification of the DNA was carried out in 10 µl reactions under conditions detailed in Supplementary Table 1. Each reaction was exposed to an initial denaturation (94 °C/5 min), followed by 30 cycles of 94 °C/30–60 s, 57–68 °C (primer-dependent)/30–60 s, 72 °C/30–120 s, and a final extension of 72 °C/7–10 min. The resulting amplicons were electrophoresed through 1.0–3.0 % (amplicon-dependent) agarose (Iwai Kagaku, Tokyo) in 0.5× TBE, and visualized by EtBr staining. Amplicons were then purified using a QIAquick PCR purification kit (QIAGEN, Germantown, MD, USA) and cycle sequenced using Big Dye Terminator technology (Applied Biosystems, Foster, CA, USA). The sequencing reactions were purified by Agencourt CleanSEQ (Beckman, Beverly, MA, USA) and analysed with an ABI prism 3130 genetic analyzer (Applied Biosystems). Sequence data were aligned using DNAMAN v6.0 software.

Bacterial artificial chromosome (BAC) library analysis and annotation

A BAC library made from cv. ‘CS’ (obtained from the John Innes Center Genome Laboratory) was screened using PCR, and positive BAC clones were sequenced according to Ishikawa et al. (2009) and Wu et al. (2002). Repetitive element (retrotransposons and DNA transposons) sequence was excluded by Repeat Masker (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) analysis, and the remaining sequence subjected to in silico gene prediction, based on GeneMark.hmm v2.2a (http://www.opal.biology.gatech.edu/GeneMark/eukhmm.cgi) software and the NCBI plant EST database (http://www.blast.ncbi.nlm.nih.gov/Blast.cgi). Softberry Bacterial Genome Explorer (http://www.linux1.softberry.com/berry.phtml) software was then used to allow the simultaneous comparison of distinct annotated genomes.

Phylogenetic analysis

Protein sequence data were aligned using ClustalW2 software (http://www.ebi.ac.uk/Tools/clustalw2/) and a phylogeny was inferred by applying the neighbour-joining method implemented in the software package MEGA v5 (Tamura et al. 2011).

RNA extraction and *cDNA* synthesis

Total RNA was extracted from the developing spikes of cv. ‘Shinchunaga’ sampled at the eight developmental stages defined by Kirby and Appleyard (1981), using the TRIzol reagent (Invitrogen, Carlsbad, CA). At the green anther stage, each spikelet was partitioned into the lemma (including awn), palea, lodicule, anther and pistil, and RNA was extracted separately from each of these, as well as from the glumes. First-strand cDNA was synthesized from 5 µg DNase treated total RNA by priming with oligo (dT), according to the Invitrogen RT-PCR first-strand synthesis protocol.

Quantitative real-time PCR (qRT-PCR)

The transcript abundance of each target was estimated by an analysis based on the StepOne Real-Time PCR system (Applied Biosystems) and THUNDERBIRD SYBR qPCR mix kit (Toyobo, Osaka) according to the manufacturers’ protocols. Each gene fragment (primers for their amplification given in Supplementary Table 1 and their genome specificity illustrated in Supplementary Fig. 6) was inserted into pCR4-TOPO (Invitrogen), which was then used to generate a standard curve based on a dilution series (4.0×10^{-2}–5.1×10^{-7} ng plasmid for each of the three *cly1* homoeolog sequences, and from 2.5 to 3×10^{-5} ng for *Actin*) in order to estimate absolute quantification and the amplification efficiency of each primer pair (Supplementary Fig. 7). At least three independent biological replicates were performed, and at least two technical replicates per biological replicate. A portion of the wheat *Actin* sequence (NCBI accession number CJ932475) was used as the reference sequence.

Mapping the *miR172*-guided cleavage site

Total RNA extracted from spikes at the terminal spikelet stage was subjected to an RNA-ligase mediated 5′ RACE (Kasschau et al. 2003) reaction, employing a GeneRacer kit (Invitrogen). This developmental stage was chosen because it is analogous to the stamen primordium stage in barley, which is when miR172-guided cleavage of *Cly1* was detectable (Nair et al. 2010). The dephosphorylation and decapping steps were both omitted, so that only the 5′ ends of the truncated transcripts were ligated to the GeneRacer RNA oligomer. A nested PCR was based on a primer targeting the GeneRacer RNA oligomer, initially in combination with a gene-specific reverse primer, and subsequently with an internal gene-specific primer (Supplementary Table 1). The resulting amplicons were electrophoresed through 1 % agarose, inserted into the TA vector (TOPO TA Cloning Kit, Invitrogen), and thence into *E. coli* (DH5α) competent cells. Randomly selected clones (without any prior size selection) were chosen for DNA sequencing.

Results

Floret gaping and lodicule swelling in cv. 'Shinchunaga'

The appearance of the pre-anthesis lodicules in cv. 'Shinchunaga' was normal, with the palea and lemma closely aligned with one another (Fig. 1c, e). As anthesis approached, the lodicules became swollen (Fig. 1b, d, f), thereby forcing open the floret. At the same time, the filaments elongated sufficiently to push the anthers out of the floret (Fig. 1a), whereupon they dehisced and the pollen was shed. Between the day prior to anthesis and anthesis itself, lodicule width expanded from 0.7 to 1.1 mm (Fig. 1c, d) and its depth from 0.4 to 1.2 mm (Fig. 1e, f).

Isolation of the bread wheat *cly1* homoeologs

Based on the high level of sequence homology between the cv. 'CS' cDNA sequence AK331198 and *Cly1*, a pair of primers (F695 and R1428, see Supplementary Table 1) was designed to amplify the sequence lying between exons 2 and 6 (Nair et al. 2010), which produced a 760-bp amplicon in cv. 'CS'. When the cv. 'CS' BAC library was screened using the same primer pair, a 760-bp amplicon was amplified from nine independent clones. When these clones were digested with *Hin*dIII, they formed three distinct contigs: the first of these involved clones WCS0770D05, WCS0842L23, WCS1336M06 and WCS1899I02; the second WCS0173D15, WCS1225B11 and WCS1471K05; and the third WCS0049K23 and WCS0230G03 (Supplementary Fig. 1). When the sequences of the nine 760 bp amplicons were aligned with one another, three haplotypes were evident. Both these results implied the presence in the bread wheat genome of three *Cly1* homoeologs. One member of each BAC group (WCS0842L23, WCS1471K05 and WCS0049K23) was taken forward for full sequencing. The length of the WCS0842L23 sequence was 111.5 kb (DDBJ accession number AB749308), that of WCS1471K05 was 97.5 kb (AB749309), and that of WCS0049K23 was 152.3 kb (AB749310). A highly conserved *AP2*-like sequence was present in each (Supplementary Tables 2–4, accession numbers AB749308–AB749310). The other regions in the three BACs which shared appreciable homology with one another comprised retrotransposon and DNA transposon sequence (Supplementary Fig. 2; Supplementary Tables 2–4).

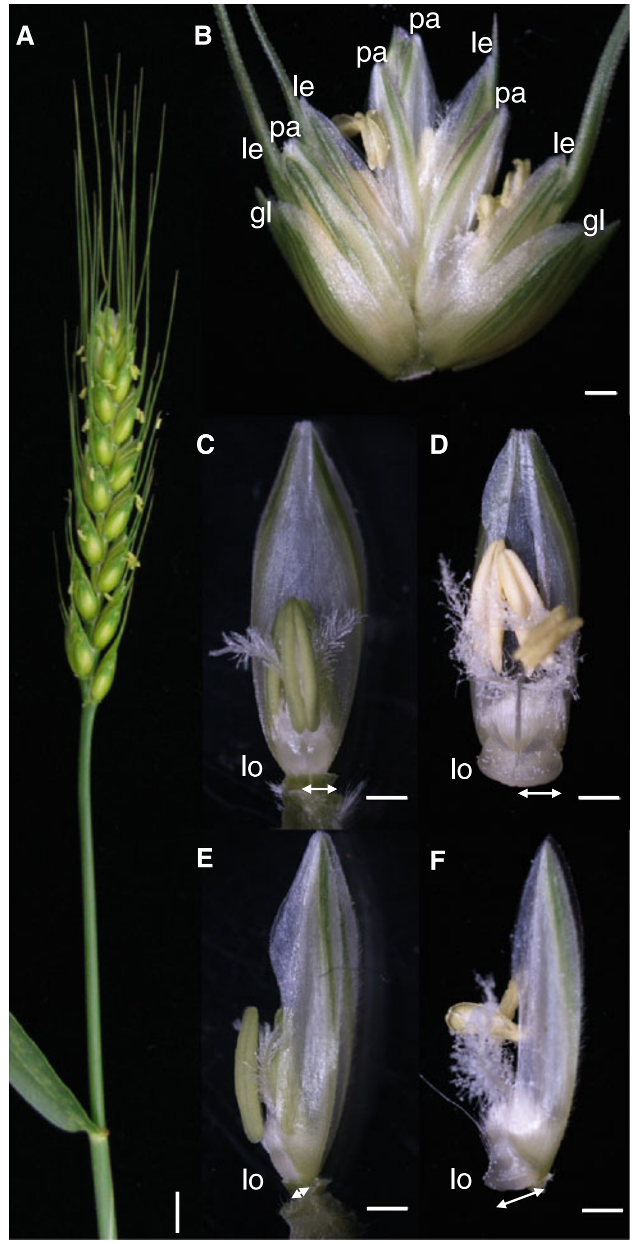

Fig. 1 Lodicule development in bread wheat cv. 'Shinchunaga'. **a**, **b** Gaping of the floret and anther exertion at anthesis. The *double-headed arrows* indicate **c**, **d** the width, and **e**, **f** the depth of the lodicule. *gl* glume, *le* lemma, *pa* palea, *lo* lodicule. (*Bar* represents 1 cm in **a**, 1 mm in **b**–**f**)

The chromosomal origin of the three BACs was explored using a standard aneuploid analysis based on the cv. 'CS' NT, Dt and deletion lines. For this purpose, a set of BAC-specific primer pairs targeting the region flanking the miR172 site was designed (Supplementary Table 1; Supplementary Fig. 3). Primer pair F-est1320/L3721A19 (targeting WCS0842L23) generated an identical size amplicon whether the template was WCS0842L23 or genomic DNA from cv. 'CS'. Among the aneuploid lines, the only ones which did not amplify this fragment were N2AT2D, Dt2AS and the 2A deletion lines lacking the distal end of the long arm (Supplementary Fig. 4). This allowed the origin of the WCS0842L23 sequence to be assigned to the distal (sub-telomeric) region of the long arm of chromosome 2A, so that the *AP2*-like gene present on this BAC was designated *TaAP2-A* (Fig. 2). In the same way, the *Cly1* ortholog present on WCS1471K05 was located to the sub-telomeric region of chromosome arm 2BL (Supplementary Fig. 4) and designated *TaAP2-B* (Fig. 2), and the one on WCS0049K23 to the sub-telomeric region of 2DL (Supplementary Fig. 4) and designated *TaAP2-D* (Fig. 2).

Structure of the *TaAP2* homoeologs

Genome-specific primers were designed from an alignment of the three *TaAP2* gene sequences obtained from the BAC clones (Supplementary Table 1) in order to amplify the corresponding gDNA and cDNA copies in both cv. 'Shinchunaga' (accession numbers AB749305–AB749307) and cv. 'CS' (accession numbers AB749311–AB749313). All three homoeologs comprised ten exons (Supplementary Fig. 5). A sequence comparison between each cv. 'CS' and cv. 'Shinchunaga' homoeolog with that of the *Cly1* (non-cleistogamous) barley allele showed that nucleotide identity ranged from 79.4 to 81.0 % in the gDNA and 90.2 to 91.2 % in the coding sequence. At the peptide level, the homology range was 88.3–89.1 % (Table 2).

The cv. 'Shinchunaga' *TaAP2* products shared key conserved sequence features with both barley Cly1 (ACY29532) and *A. thaliana* AP2 (NP195410) (Jofuku et al. 1994; Tang et al. 2007) (Fig. 3), namely a highly basic ten-residue nuclear localization signal, and two AP2 domains each comprising two copies of a 68-residue direct repeat, the AASSGF box (corresponding to the miR172 target site) and three other motifs (motifs 1–3). All these motifs were highly conserved among the grasses. There was no variation for any of these features among the three *TaAP2* homoeologs present in cvs. 'CS' and 'Shinchunaga'. Comparison between the deduced cv. 'CS' and cv. 'Shinchunaga' TaAP2-A sequences showed that the former

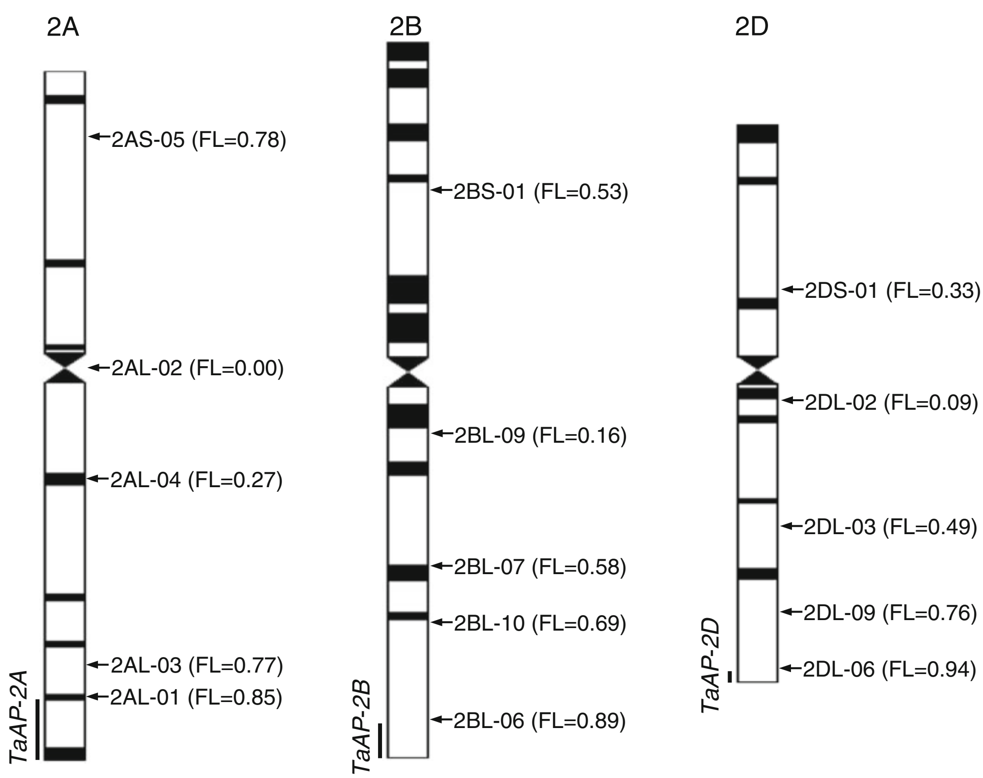

Fig. 2 The intrachromosomal location of the TaAP2 homoeologs. The C-banding karyotype of cv. 'CS' was taken from Endo and Gill (1996). Deletion breakpoints (*arrowed*) defined by the associated fraction length (*FL*). *TaAP2* genotyping profiles shown in Supplementary Fig. 4

Table 2 Homology between *Cly1* and its bread wheat orthologs in both cv. 'Shinchunaga' and cv. 'CS'

	Genomic DNA					ORF			Protein				Accession no.
	Length (bp)	GC (%)	No. introns	No. exons	Identity (%)	Length (bp)	GC (%)	Identity (%)	Length (aa)	MW (kDa)	*PI*	Identity (%)	
H. vulgare cv. Azumamugi													
Cly1	2,691	60.8	9	10	100	1,464	70.5	100	487	51.752	7.18	100	GQ403050
T. aestivum cv. Shinchunaga													
TaAP2-A	2,652	62.1	9	10	79.6	1,431	71.3	90.7	476	50.28	7.47	88.6	AB749305
TaAP2-B	2,706	61.9	9	10	80.0	1,455	71.1	90.2	484	51.066	7.14	89.1	AB749306
TaAP2-D	2,755	60.7	9	10	81.0	1,452	71.4	90.4	483	50.81	7.11	89.0	AB749307
T. aestivum cv. Chinese spring													
TaAP2-A	2,643	62.0	9	10	79.4	1,422	71.1	90.3	473	49.989	7.47	88.3	AB749311
TaAP2-B	2,706	61.9	9	10	80.0	1,455	71.1	90.2	484	51.066	7.14	89.1	AB749312
TaAP2-D	2,786	60.8	9	10	80.1	1,461	71.7	91.2	486	51.061	7.11	88.9	AB749313

The percentage of identity was estimated by alignment of each sequences with the *H. vulgare* cv. AZ (GenBank: GQ403050)

Genomic DNA: from start to stop codons; ORF: open reading frame; PI: isoelectric points

differed from the latter by a run of five, rather than eight proline residues around position 435 (Fig. 3); the same comparison involving the two TaAP2-B sequences identified no polymorphisms whatsoever, while the cv. 'CS' TaAP2-D sequence differed from that of its cv. 'Shinchunaga' homolog at positions 108 (G/V), 329 (E/D) and 436 (L/P), featured an extra proline residue between positions 440 and 441, and had a string of seven (rather than five) alanine residues around position 353.

Resequencing of the three TaAP2 homoeologs from six other wheat cultivars showed that at *TaAP2-A*, cvs. KU-515, Fukuho Komugi and Norin 61 were identical with the cv. 'Shinchunaga' type, carrying a GCCGCCGCC insertion in exon 10 (encoding three prolines), while cvs. KU-163, KU-165 and KU-265, like cv. 'CS', lacked this insert. The only polymorphism among the *TaAP2-B* sequences was a G/A variant in intron 4 in cv. Fukuho Komugi. All six *TaAP2-D* sequences were identical to that present in cv. 'Shinchunaga.

Phylogeny of the *TaAP2* homoeologs

A protein-based phylogeny (Fig. 4) showed that the three cv. 'Shinchunaga' products were highly similar both to one another, to that of the non-cleistogamous barley Cly1 protein (Nair et al. 2010) and to the rice SHAT1 (OsAP2) protein (Zhou et al. 2012). This cluster of cereal sequences was less strongly related to those of the two *A. thaliana* proteins AP2 (NP195410) and TOE3 (NP201519) (an ethylene-responsive transcription factor), *Brassica napus* APETALA2 (ADU04499), *Betula platyphylla* APETALA2 (AEL29576) and *Ricinus communis* putative APETALA2 (XP002534399) (data not shown). The product of the wheat gene *Q* and that of its two homoeologs, together with barley HvAP2-like, were phylogenetically closely related to one another, but clearly belong to a distinct lineage, along with the maize proteins IDS1 and SID1, and the rice protein SNB.

Transcription profiling of the *TaAP2* homoeologs

The transcription profiles of each of the three *TaAP2* genes carried by cv. 'Shinchunaga' were explored using qRT-PCR, based on *TaAP2* homoeolog-specific primer pairs (Supplementary Fig. 6). Performing qRT-PCR allowed for a direct comparison of *TaAP2* mRNA copy number within each given biological sample. The transcription of *TaActin* was stable throughout spike development, and was used as a reference to compare *TaAP2* mRNA transcript abundance between different stages and different organs. The analysis revealed that each homoeolog was transcribed throughout spike development (from the glume primordium stage to anthesis) (Fig. 5a). The abundance of the *TaAP2-A* and *TaAP2-B* transcript between the green anther stage and anthesis was about double that observed earlier during development, but that of *TaAP2-D* remained throughout at a rather constant level below that of either *TaAP2-A* or *TaAP2-B*). All three homoeologs were also transcribed at the green anther stage in each of the floral organs sampled (lemma, palea, lodicule, anther, pistil and glume). The transcript abundance of *TaAP2-A* and *TaAP2-B* in the lodicule was more than double that in any other organ.

miR172-guided cleavage of *TaAP2* transcript

All six *TaAP2* sequences (three each in each of the two cultivars) shared identical *miR172* targeting site sequence (Supplementary Fig. 5). The modified 5′ RACE experiment

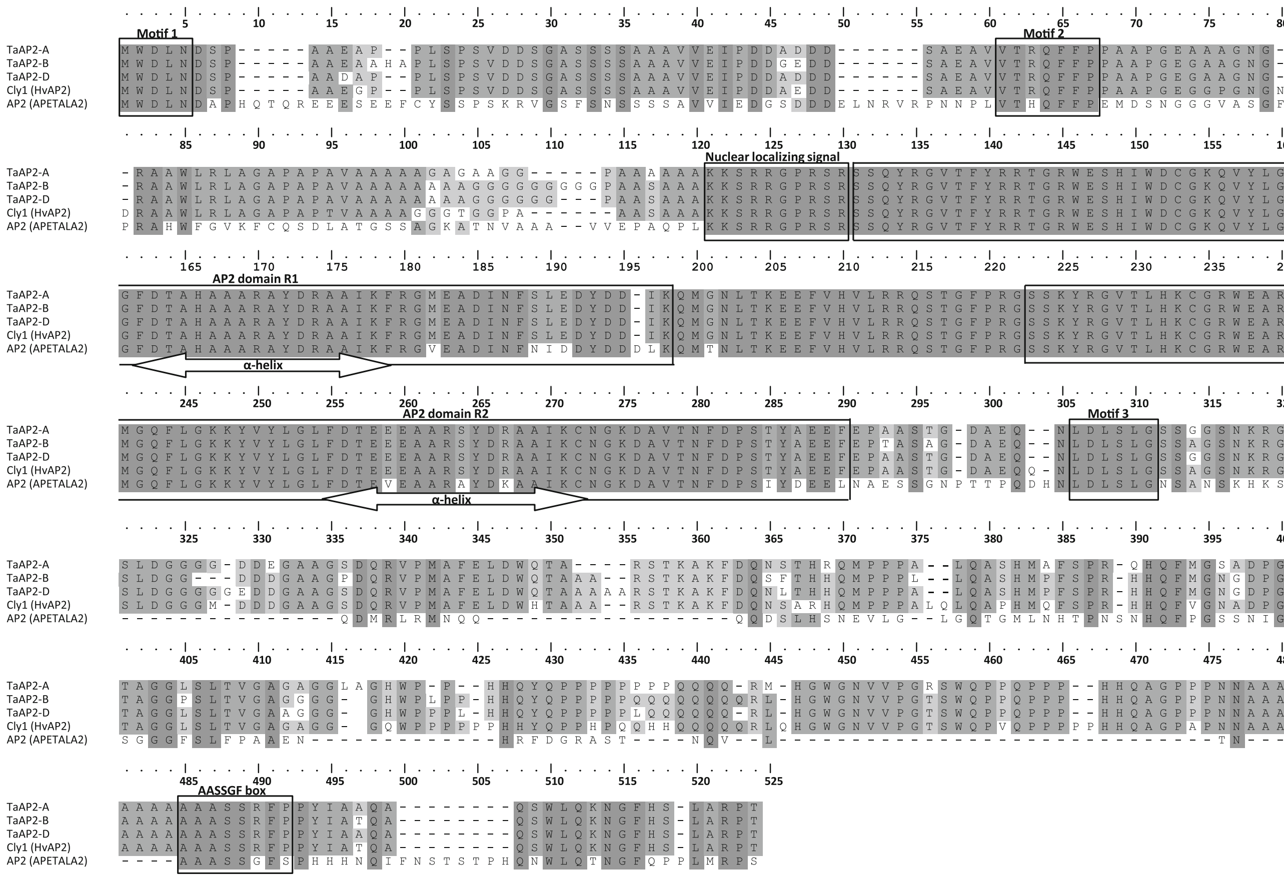

Fig. 3 Peptide alignment of the TaAP2 proteins from bread wheat cv. 'Shinchuanga', Cly1 from barley cv. 'Azumamugi' and AP2 from *A. thaliana*. Key features of the sequence (motifs 1–3, nuclear localization signal, R1 and R2 domains and the AASSGF box) are shown *boxed*. The α-helical structures formed by the core region of each AP2 domain are delimited by *arrows*

carried out to detect *miR172*-guided cleavage was expected to generate fragments of sizes 210, 370 and 255 bp from, respectively, *TaAP2-A, TaAP2-B and TaAP2-D*. The derived sequences of the majority of 5′ RACE clones analysed were consistent with cleavage within the miRNA172 targeting site. This applied to 39/64 *TaAP2-A* clones, 32/52 *TaAP2-B* clones and 17/43 *TaAP2-D* clones, with cleavage most frequently occurring between the A and U nucleotides (Fig. 6), as also occurs in *Cly1* (Nair et al. 2010). The other clones comprised 3′ UTR sequence, consistent with random mRNA breakage, again as has been observed in barley (Nair et al. 2010).

Discussion

The three bread wheat *TaAP2* genes all mapped to the distal region of the long arm of the group 2 chromosomes (Fig. 2), a region syntenous with that harbouring *Cly1* in barley (Turuspekov et al. 2004). The three homoeologs share a highly similar structure to that of *Cly1* (Table 2; Fig. 3), and their sequences are phylogenetically strongly related both to one another's and to that of *Cly1* (Fig. 4). Each was transcribed during spike development, and was particularly abundantly expressed in the lodicule (Fig. 5b). The mRNA extracted at the terminal spikelet stage was effectively cleaved at the miR172 target site (Fig. 6). Cleavage of *HvAP2* mRNA was detectable only in non-cleistogamous cultivars (Nair et al. 2010), while here the proportion of cleaved mRNA in the non-cleistogamous cv. 'Shinchunaga' was significantly greater than 0, the level expected for a cleistogamous type (Nair et al. 2010). Thus, it is clear that wheat *TaAP2* and barley *Cly1* represent a set of orthologous loci. As in barley, floret gaping in the non-cleistogamous wheat cv. 'Shinchunaga' was triggered by the expansion of the lodicules (Fig. 1), so the implication is that *TaAP2* and *Cly1* share both structure and function.

In barley, the only biologically significant difference between the *Cly1* (non-cleistogamous) and the *cly1* (cleistogamous) allele is the synonymous single nucleotide change at the miR172 targeting site; in the presence of the latter allele, miR172 is unable to cleave the relevant mRNA (Nair et al. 2010). The miR172 targeting sites in the *TaAP2* genes (in both cv. 'Shinchunaga' and cv. 'CS') as in

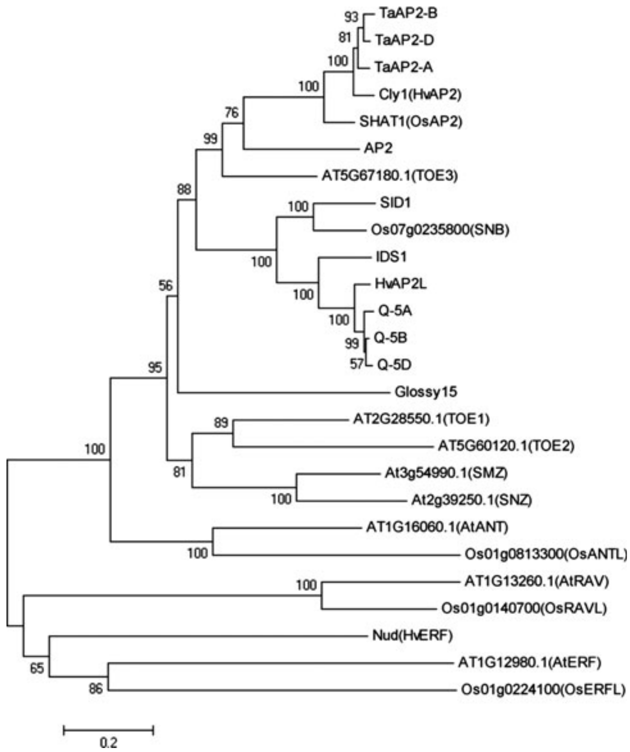

Fig. 4 Phylogeny of *A. thaliana*, rice, maize, wheat cv. ‘Shinchunaga’ and barley cv. ‘Azumamugi’ *AP2* homologs obtained using the neighbour-joining method. Only bootstrap values >50%, as calculated from 1,000 replicates, are shown

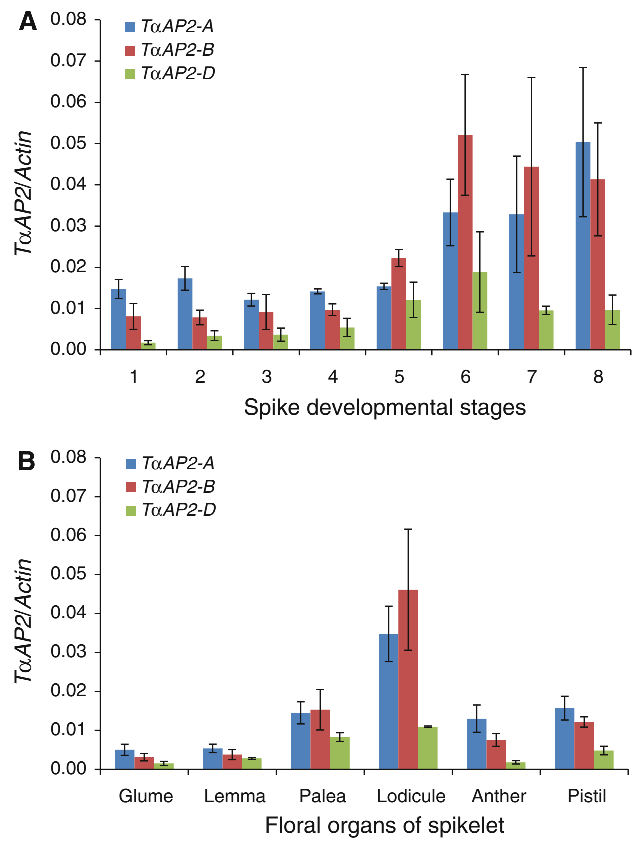

Fig. 5 Transcription profiling of *TaAP2* homoeologs in wheat cv. ‘Shinchunaga’. **a** In the developing *spike* (*1* glume primordium stage, *2* lemma primordium stage, *3* floret primordium stage, *4* terminal spikelet stage, *5* white anther stage, *6* green anther stage, *7* yellow anther stage, *8* anthesis). **b** In various floral organs at the green anther stage. Mean ± SE of three biological replicates are shown

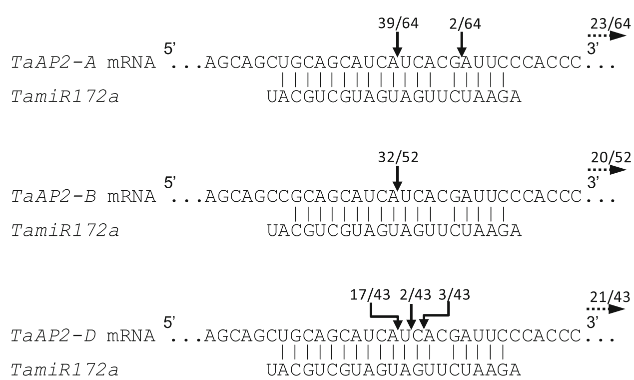

Fig. 6 *miR172*-guided mRNA cleavage of *TaAP2* mRNA. The 5′ termini of the cleaved products were identified using a modified 5′ RACE approach. *Vertical arrows* indicate the inferred 5′ termini of *miR172*-guided cleavage, and the number above each *arrow* the proportion of clones containing that site. The *horizontal arrows* refer to cleavage downstream of the *miR172* site

Cly1 (e.g. in the non-cleistogamous barley cv. ‘Azumamugi’) are highly similar to one another; the exceptions relate to the second nucleotide, namely C in *TaAP2-B* but U in *TaAP2-A*, *TaAP2-D* and *cly1* (Supplementary Fig. 5). This nucleotide is also variable within barley, but variants are not associated with any suppression of miR172-guided cleavage and/or cleistogamy (Nair et al. 2010). As expected therefore, the three *TaAP2* mRNAs were all readily cleaved by miR172 (Fig. 6), consistent with the non-cleistogamous phenotype of cv. ‘Shinchunaga’.

A rational strategy for the induction of cleistogamy in bread wheat would be to identify naturally occurring or induced mutants at the miR172 targeting site for each *TaAP2* homoeolog, and then to combine these within a single plant by conventional crossing. Since the primer sets developed here specifically amplify the miR172 targeting site from each homoeolog, their deployment should be effective for the detection of such variants in both hexaploid and tetraploid materials. No such variation was apparent among the *TaAP2* alleles resequenced from eight cultivars, as might be expected from such a highly conserved sequence (Fig. 7; Nair et al. 2010). The wheat cv. ‘U24′ is cleistogamous (Kubo et al. 2010), the frequency of floret gaping is low, its anther filaments do not elongate, and its anthers are not extruded (unpublished data). However, because its lodicules are fully swollen at anthesis, its *TaAP2* alleles were not considered for resequencing in the

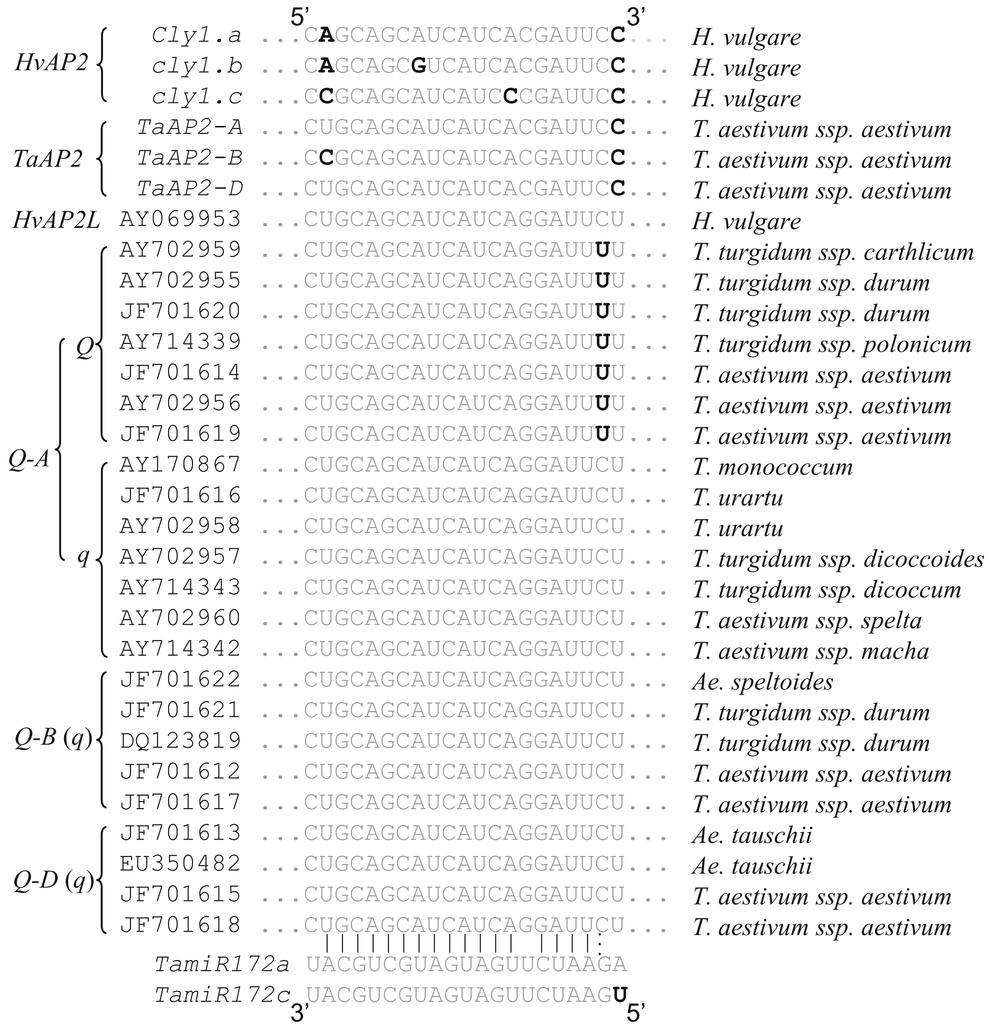

Fig. 7 *miR172* target sites in *AP2* and *AP2-like* mRNA in barley and wheat. Wheat miRNAs *TamiR172a* and *c* have 3′–5′ sequence complementarity to the binding site of *AP2* and *Q* in barley and wheat. Sequence variants identified in *bold*

present study. We were unable to access grain of the two cleistogamous *T. durum* accessions reported by Chhabra and Sethi (1991). In barley, two naturally occurring distinct cleistogamous alleles (*cly1.b* and *cly1.c*) have been identified, both of which are thought to be of relatively recent origin (Nair et al. 2010). The recessive nature of these mutants implies that their wheat equivalents may be represented in material which is phenotypically non-cleistogamous. The TILLING approach (Henikoff et al. 2004) could provide an attractive platform for detecting allelic variants. The abundance of *TaAP2-D* transcript was lower than that of either *TaAP2-A* or *TaAP2-B* throughout the development of the spike. It has been well established (Comai et al. 2000; Bottley et al. 2006; Shitsukawa et al. 2007; Zhang et al. 2011) that polyploidization results in an appreciable level of homoeolog silencing or re-functionalization.

Site-specific DNA binding domains are favoured targets for certain genetic engineering strategies aiming at crop improvement. Site-specific nucleases have been designed by fusing the DNA cleavage domain of *Fok*I and a custom-designed DNA binding domain, such as the C2H2 zinc-finger motif for zinc-finger nucleases (ZFNs) (Urnov et al. 2010) and the truncated transcription activator-like effector (TALE) domain for TALE nucleases (Miller et al. 2011). Both ZFNs and TALE nucleases induce double-strand breaks at a target locus, which are subsequently repaired by error-prone non-homologous end-joining; the intention is to induce small indels at the breakage site, thereby enabling targeted mutagenesis to be applied in non-model organisms (Wood et al. 2011). The approach could lend itself readily to engineering the *miR172* targeting site in wheat. An alternative approach could be to attempt the modification of the *miR172* gene. One possibility, already pioneered in both maize (Chuck et al. 2007) and barley (Brown and Bregitzer 2011), could be to induce the insertion of a copy of a *Ds* transposon into *miR172*, with the intention of down-regulating its transcription. Based on its homology to maize *IDS1* and the similarity of the maize *ts4* mutant to the *miR172* mutant (Brown and Bregitzer 2011), the barley

Cly1 and *AP2-like* gene are both likely targets of miR172. By inference, its targets in wheat include both *TaAP2* and the major domestication gene *Q*. *Q* and its homoeoloci form a sister clade along with barley *HvAP2-like* (Fig. 4). *Q* is responsible for the free-threshing trait (as well as acting pleiotropically on a number of other characters). One of its two homoeologs has evolved into a pseudogene and the other has been sub-functionalized (Simons et al. 2006; Zhang et al. 2011). The *Q* sequence includes a 21 nt *miR172* targeting site in exon 10 (Zhang et al. 2011), suggesting its regulation by this miRNA. The *miR172* target site within the A, B and D genome *Q* homoeologs is perfectly complementary to miR172 except for a single nucleotide mismatch at position 15 of the miRNA binding site. The mismatch is completely conserved among the cereals, suggesting a likely role in mRNA cleavage (Nair et al. 2010). All *Q* alleles resequenced to date include a single nucleotide variant of C to U at position 20 of the miRNA binding site (Simons et al. 2006, Zhang et al. 2011). *TamiR172a* and *c* enjoy near perfect complementarity to their respective mRNA targets, and so are expected to be effective as cleavage agents (Fig. 7). Mutation in *miR172* genes may disrupt the functionality of *Q* as does *Ds* transposition into *HvmiR172* on *HvAP2-like* in barley (Brown and Bregitzer 2011). If it also disrupted genes within the ABCDE network, then modifying the *TamiR172* genes would not be an attractive approach for engineering cleistogamy in wheat. Instead, altering the miR172 targeting site in *TaAP2* might represent a less risky way of achieving this goal.

FHB can be a devastating disease of wheat (Snijders 1990; Parry et al. 1995; McMullen et al. 1997). As for most plant diseases, a rational control strategy should take advantage of genetically based resistance, and because the pathogen commonly enters the host through the floret around the time of anthesis, a plant with a flowering habit involving minimal floret gaping is more likely to escape infection than one in which the floret gapes (Gilsinger et al. 2005; Kubo et al. 2010). Furthermore, as genetic modification (GM) technology becomes more widely exploited in crop breeding, controlling gene flow between GM and non-GM cultivars needs to be managed. Although wheat is a predominantly self-pollinating species (De Vries 1971), its pollen can move up to 2.75 km from its source (Matus-Cádiz et al. 2007). Cleistogamy delivered by allelic variation at the *TaAP2* homoeologs would simultaneously improve the resistance of the crop to FHB and minimize the risk of pollen-mediated gene flow between GM and non-GM wheat cultivars.

Acknowledgments We thank Y. Nagamura, J. Song, G. Chen, C. Liu and C. Li, H. Sassa, S. Kikuchi for their help and advice and K. Kakeda for comments on the manuscript. This research was funded by the Japanese Ministry of Agriculture, Forestry and Fisheries (Genomics for Agricultural Innovation grant no. TRG1004) to T.K. S.N. appreciates the award of a Japanese Government (Monbukagakusho: MEXT) scholarship.

References

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. Plant Cell 15:2730–2741

Ban T, Suenaga K (2000) Genetic analysis of resistance to fusarium head blight caused by *Fusarium graminearum* in Chinese wheat cultivar Sumai 3 and the Japanese cultivar Saikai 165. Euphytica 113:87–99

Bartel DP (2009) MicroRNAs: Target recognition and regulatory functions. Cell 136:215–233

Bottley A, Xia GM, Koebner RMD (2006) Homoeologous gene silencing in hexaploid wheat. Plant J 47:897–906

Brown RH, Bregitzer P (2011) A Ds insertional mutant of a barley *mir172* gene results in indeterminate spikelet development. Crop Sci 51:1664–1672

Chapman V, Miller TE, Riley R (1976) Equivalence of the a genome of bread wheat and that of Triticum urartu. Genet Res 27:69–76

Chen X (2004) A microRNA as a translational repressor of *APETALA2* in *Arabidopsis* flower development. Science 303: 2022–2025

Chhabra AK, Sethi SK (1991) Inheritance of cleistogamic flowering in durum wheat (*Triticum durum*). Euphytica 55:147–150

Chuck G, Meeley R, Irish E, Sakai H, Hake S (2007) The maize tasselseed4 microRNA controls sex determination and meristem cell fate by targeting *Tasselseed6/indeterminate spikelet1*. Nat Genet 39:1517–1521

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. Nature 353:31–37

Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B (2000) Phenotypic instability and rapid gene silencing in newly formed Arabidopsis allotetraploids. Plant Cell 12:1551–1567

De Vries AP (1971) Flowering biology of wheat, particularly in view of hybrid seed production—a review. Euphytica 20:152–170

Endo TR, Gill BS (1996) The deletion stocks of common wheat. J Hered 87:295–307

Fu YB, Somers DJ (2009) Genome-wide reduction of genetic diversity in wheat breeding. Crop Sci 49:161–168

Gilsinger J, Kong L, Shen X, Ohm H (2005) DNA markers associated with low fusarium head blight incidence and narrow flower opening in wheat. Theor Appl Genet 110:1218–1225

Glover B (2007) Understanding flowers and flowering. New York, Oxford, p 227

Henikoff S, Till BJ, Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol 135: 630–636

Heslop-Harrison Y, Heslop-Harrison JS (1996) Lodicule function and filament extension in the grasses: potassium ion movement and tissue specialization. Ann Bot 77:573–582

Hori K, Kobayashi T, Sato K, Takeda T (2005) QTL analysis of fusarium head blight resistance using a high-density linkage map in barley. Theor Appl Genet 111:1661–1672

Ishikawa G, Nakamura T, Ashida T, Saito M, Nasuda S, Endo TR, Wu J, Matsumoto T (2009) Localization of anchor loci representing five hundred annotated rice genes to wheat chromosomes using PLUG markers. Theor Appl Genet 118:499–514

Jofuku KD, den Boer BG, Montagu MV, Okamuro JK (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene *APETALA2*. Plant Cell 6:1211–1225

Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function. Dev Cell 4:205–217

Kihara H (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. Agric Hortic 19:889–890

Kilian B, Özkan H, Deusch O, Effgen S, Brandolini A, Kohl J, Martin W, Salamini F (2007) Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. Mol Biol Evol 24:217–227

Kirby EJM, Appleyard M (1981) Cereal development guide. National Agricultural center Cereal Unit, Stoneleigh, Warwickshire, UK

Komatsuda T, Nakamura I, Takaiwa F, Oka S (1998) Development of STS markers closely linked to the *vrs1* locus in barley, *Hordeum vulgare*. Genome 41:680–685

Kubo K, Kawada N, Fujita M, Hatta K, Oda S, Nakajima T (2010) Effect of cleistogamy on fusarium head blight resistance in wheat. Breeding Sci 60:405–411

Ma R, Zheng DS, Fan L (1996) The crossability percentages of 96 bread wheat landraces and cultivars from Japan with rye. Euphytica 92:301–330

Matus- Cádiz MA, Hucl P, Dupuis B (2007) Pollen-mediated gene flow in wheat at the commercial scale. Crop Sci 47:573–579

McMullen M, Jones R, Gallenberg D (1997) Scab of wheat and barley: a re-emerging disease of devastating impact. Plant Dis 81:1340–1348

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148

Nair SK, Wang N, Turuspekov Y, Pourkheirandish M, Sinsuwongwat S, Chen G, Sameri M, Tagiri A, Honda I, Watanabe Y, Kanamori H, Wicker T, Stein N, Nagamura Y, Matsumoto T, Komatsuda T (2010) Cleistogamous flowering in barley arises from the suppression of microRNA-guided *HvAP2* mRNA cleavage. Proc Natl Acad Sci USA 107:490–495

Nevo E (2009) Ecological genomics of natural plant populations: the Israeli perspective. In: Somers DJ, Langridge P, Gustafson JP (eds) Methods in molecular biology, plant genomics, vol 513. Human Press, A part of Springer Science + Business Media, pp 321–344

Nevo E (2011) *Triticum*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources, cereals. Springer, Berlin, pp 407–456

Parry DW, Jenkinson P, McLeod L (1995) Fusarium ear blight (scab) in small grain cereals. Plant Pathol 44:207–238

Petersen G, Seberg O, Yde M, Berthelsen K (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). Mol Phylogenet Evol 39:70–82

Riley R, Unrau J, Chapman V (1958) Evidence on the origin of the B genome of wheat. J Hered 49:91–98

Sato K, Hori K, Takeda K (2008) Detection of fusarium head blight resistance QTLs using five populations of top-cross progeny derived from two-row × two-row crosses in barley. Mol Breeding 22:517–526

Sears ER (1954) The aneuploids of common wheat. Univ Mo Agric Exp Stn Bull 572:1–58

Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: Rilly R, Lewis KR (eds) Chromosome manipulations and plant genetics. Oliver and Boyd, Edinburgh, pp 29–45

Sears ER, Sears MS (1978) The telocentric chromosomes of commonwheat. In: Ramanujam S (ed) Proceedings of the 5th international wheat genetics symposium. Indian Society of Genetics and Plant Breeding, New Delhi, pp 389–407

Sethi K, Chhabra AK (1990) Cleistogamy in wheat. Rachis 9:34–36

Shitsukawa N, Tahira C, Kassai KI, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, Ogihara Y, Murai Y (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS box gene in hexaploid wheat. Plant Cell 19:1723–1737

Simons KJ, Fellers JP, Trick HN, Zhang ZC, Tai YS, Gill BS (2006) Molecular characterization of the major wheat domestication gene *Q*. Genetics 172:547–555

Snijders CHA (1990) Fusarium head blight and mycotoxin contamination of wheat, a review. Eur J Plant Pathol 96:187–198

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tang M, Li G, Chen M (2007) The phylogeny and expression pattern of *APETALA2*-like genes in rice. J Genet Genomics 34:930–938

Theissen G, Saedler H (2001) Floral quartets. Nature 409:469–471

Turuspekov Y, Mano Y, Honda I, Kawada N, Watanabe Y, Komatsuda T (2004) Identification and mapping of cleistogamy genes in barley. Theor Appl Genet 109:480–487

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11:636–646

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ (2011) Targeted genome editing across species using ZFNs and TALENs. Science 333:307

Wu J, Maehara T, Shimokawa T, Yamamoto S, Harada C, Takazaki Y, Ono N, Mukai Y, Koike K, Yazaki J, Fujii F, Shomura A, Ando T, Kono I, Waki K, Yamamoto K, Yano M, Matsumoto T, Sasaki T (2002) A comprehensive rice transcript map containing 6591 expressed sequence tag sites. Plant Cell 14:525–535

Zhang Z, Belcram H, Gornicki P, Charles M, Just J, Huneau C, Magdelenat G, Couloux A, Samain S, Gill BS, Rasmussen JB, Barbe V, Faris JD, Chalhoub B (2011) Duplication and partitioning in evolution and function of homoeologous *Q* loci governing domestication characters in polyploid wheat. Proc Natl Acad Sci USA 108:18737–18742

Zhou Y, Lu D, Li C, Luo J, Zhu BF, Zhu J, Shangguan Y, Wang Z, Sang T, Zhou B, Han B (2012) Genetic Control of Seed Shattering in Rice by the APETALA2 Transcription Factor *SHATTERING ABORTION1*. Plant Cell 24:1034–1048