Research Article

Genome-Wide Comparative Analysis of Flowering-Related Genes in Arabidopsis, Wheat, and Barley

Fred Y. Peng, Zhiqiu Hu, and Rong-Cai Yang

1 Feed Crops Branch, Alberta Agriculture and Forestry, 7000-113 Street, Edmonton, AB, Canada T6H 5T6
2 Department of Agricultural, Food and Nutritional Science, University of Alberta, 410 Agriculture/Forestry Centre, Edmonton, AB, Canada T6G 2P5

Correspondence should be addressed to Rong-Cai Yang; rong-cai.yang@ualberta.ca

Received 4 June 2015; Revised 24 July 2015; Accepted 10 August 2015

1. Introduction

Allohexaploid wheat (Triticum aestivum L., 2n = 6x = 42) and diploid barley (Hordeum vulgare L., 2n = 2x = 14) are two major temperate cereal crop species. The polyploid wheat originated from a two-step natural hybridization of three diploid species, each with seven basic chromosomes (x = 7). The first step was the natural hybridization between Triticum urartu Tumanian ex Gandilyan (2n = 2x = 14 AA, the A genome) and Aegilops speltoides Tausch (2n = 2x = 14 BB, the B genome) to form a tetraploid wheat species, Triticum turgidum L. [1, 2]. In the second step, the natural hybridization between Triticum turgidum L. (2n = 4x = 28 AABB) and Aegilops tauschii Coss. (2n = 2x = 14 DD, the D genome) occurred to form the hexaploid wheat (AABBDD), which, like many other allopolyploid plant species, has a diploid-like meiotic behavior to prevent the formation of multivalent associations of more than two homologous or homoeologous chromosomes at meiosis [3]. The hexaploid wheat has a very large genome, with an estimated size of about 17 Gb [4] and with more than 80% of the genome consisting of repetitive DNA sequences [5, 6]. Similarly, the diploid barley also has a large genome with an estimated size of about 5.3 Gb and with approximately 84% of the genome being comprised of mobile elements or other repeated structures [7]. Thus, despite recent constructions of physical maps for wheat and barley [6–8], genome-wide characterization of gene functions in these species remains challenging.

Both wheat and barley are widely cultivated mainly for human food, beverages, and animal feed and they are among the top five cereal crops in the world, with a global production of 713 and 145 million tons in 2014 (International Grains Council, http://www.igc.int/en/grainsupdate/sd.aspx). The timing of flowering is one of the most critical agronomic traits influencing grain yield and quality. Early flowering and maturing wheat and barley cultivars are desired in...
high-latitude regions with short growing seasons and long summer days [9–12]. Additionally, synchronous flowering and maturity can help timely crop harvest to prevent lowered yield and quality due to frost and preharvesting sprouting [13]. Therefore, control of flowering time and the adaptation of flowering to diverse growing environments are vitally important for sustainable production of wheat and barley under changing climate conditions or in different geographical regions.

Most of our understanding of the genetic components and environmental factors triggering floral initiation is gained in the diploid model organism Arabidopsis (Arabidopsis thaliana (L.) Heynh., 2n = 2x = 10), which, like wheat and barley, is a long-day plant, is widely distributed in northern temperate regions, and requires both vernalization (extended exposure to low temperatures) and long photoperiod to stimulate flowering [10, 14–16]. To date, more than 180 genes involved in flowering time control have been identified in Arabidopsis [17–26]. In contrast, only a small number of flowering genes have been studied in temperate cereal crops wheat and barley. These include the pseudoresponse regulator gene Ppd-h1 (on 2D) [12, 27–30], TaGII (GIGANTEA homolog) [31] and the vernalization genes VERNALIZATION 1 (VRN1) and VRN2 in wheat [15, 32–34], and Ppd-H1 (on 2H) [35], HvGI [36], HvVRN1 and HvVRN2 [37], HvCO1 (an ortholog of Arabidopsis CONSTANS) [38], EARLY MATURITY 8 (an ortholog of ELF3 in Arabidopsis) [39], and EARLY FLOWERING 3 [40] in barley. Recently, Alqudah et al. [41] compiled a list of 60 genes for their genome-wide association study (GWAS) of photoperiod response in barley. In addition, several reviews about the genetic control of flowering, including those in temperate cereals, have also been published in recent years [10, 16, 42–50], highlighting not only functional conservation but also divergence in molecular mechanisms underlying the floral transition between Arabidopsis and cereal crops. For example, the common ancestor of Arabidopsis and barley is estimated to possess two-thirds of the key circadian clock genes identified in Arabidopsis [51]. The functional orthologs of Arabidopsis CONSTANS (CO) and FLOWERING LOCUS T (FT) have been identified in wheat and barley [21, 48, 52, 53]. However, it is important to note the difference of flowering pathways (most notably the vernalization response) in the dicots and monocots [21]. It should also be recognized that genes with the same name in Arabidopsis and cereals may not be functionally related and vice versa. For example, the VRN1 gene in wheat and barley is not related to VRN1 in Arabidopsis but homologous to AP1/CAL/FUL [48], and the VRN3 gene in wheat and barley is an ortholog of FT [54].

In addition to experimental identification and characterization of flowering-related genes, computational genomic analysis has become a popular strategy to identify flowering-related genes in economically important crop species, usually using Arabidopsis as the reference. For example, such comparative genomic analyses have been carried out in dicot species including long-day garden pea (Pisum sativum) [55], short-day soybean (Glycine max) [22, 56], day-neutral mung bean (Vigna radiata) [57], and cotton (Gossypium hirsutum, cultivated cotton’s day-neutral flowering is due to domestication and selective breeding but its wild progenitors require short days) [58], as well as in monocot species including short-day rice [59, 60] and long-day Brachypodium (Brachypodium distachyon) [61], which is a small temperate grass (purple false brome) with a potential to serve as a new model species for temperate cereal crops [62] and diverged from wheat around 32–39 million years ago (MYA) [63]. These comparative genomic analyses have provided researchers with candidate genes for further molecular characterization to advance our understanding on the genetic control of flowering time in crops. To our knowledge, however, no similar genomic-scale analysis has been reported in wheat; the CCT domain gene family, including CONSTANS-like (COL) and PREDUORRESPONSE REGULATOR (PRR) gene families, core circadian clock genes, and a MYB transcription factor (HvLUXI) involved in transcriptional regulation within the circadian clock have been analyzed in barley [51, 64–67].

The genome sequences of bread wheat and barley were released in 2012 [6–8], laying a foundation for identification and comparative analyses of flowering-related genes between Arabidopsis, wheat, and barley on a genome-wide scale. This study has two objectives. The primary objective was to predict putative orthologs of Arabidopsis flowering genes in wheat and barley using a bioinformatic approach that combines reciprocal BLAST searches [68] and OrthoMCL clustering [69, 70]. InterPro domains in all these flowering relevant proteins were compared in Arabidopsis versus wheat or barley [71, 72] and phylogeny analysis was used to validate our approach to ortholog prediction. The secondary objective was to determine whether or not orthologous genes exhibit expression similarities using microarray data analysis. This was achieved by examining gene expression profiles of the flowering genes in different organs and developmental stages using three similar, public transcriptome datasets obtained from the Plant Gene Expression Database PLEXdb [73–76]. Our work was initiated to create a comprehensive collection of flowering-related genes in wheat and barley and their expression profiles in different tissues and developmental stages. This collection will help researchers to select additional genes for further study on genetic control of flowering time in these two important temperate cereal crops.

2. Materials and Methods

2.1. Identification of Flowering-Related Genes in Wheat and Barley. The 204 flowering genes in Arabidopsis were compiled manually through searches on TAIR [77] and previous studies (Supplemental file 1 in Supplementary Material available online at http://dx.doi.org/10.1155/2015/874361), which include genes with GO (gene ontology) biological process containing one or more terms of circadian rhythm, flowering, flower (floral) development, regulation of flower development, photoperiodism, or vernalization response. The flowering-related genes in wheat and barley were identified using reciprocal BLAST searches followed by OrthoMCL clustering [68–70]. To enable batch BLAST searches, a standalone version of the BLAST tool (version
2.2. InterPro Domain Analysis. All protein sequences of the flowering-related genes in Arabidopsis, wheat, and barley were analyzed with a standalone version InterProScan 5 [71, 72]. The default parameters were used, and its InterPro lookup option (iprlookup) was switched on to generate InterPro annotation. For gene encoding multiple proteins (i.e., from alternatively spliced transcripts), its longest sequence was chosen for this analysis.

2.3. Multiple Sequence Alignment and Phylogenetic Analysis. Multiple sequence alignment (MSA) was performed using predicted protein sequences with Clustal X (version 2.0) [79] and manually examined with Jalview (version 2.0) [80]. For phylogenetic analysis using a Bayesian approach with BEAST (v1.8.2) [81], input files were first generated using the alignment files from Clustal X (saved as NEXUS format) with BEAUTi (Bayesian Evolutionary Analysis Utility), and the phylogeny was analyzed with BEAST under default settings (1,000,000 generations, four Markov chains, and two runs). The first 25% of the tree from the runs was discarded as burn-in. Then, the tree topology was annotated with TreeAnnotator (both BEAUTi and TreeAnnotator are within the BEAST package). Finally, the phylogenetic tree was viewed with FigTree v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). The moss homologs PpIs34_16V6 (for PEBP) and Phpat.004G002000.1 (for MADS orthogroup OG5:178217) were used as outgroup to root phylogenetic trees.

2.4. Expression Analysis of Flowering-Related Genes. The raw data files (.CEL files) for the transcriptome datasets of Arabidopsis, wheat, and barley were retrieved from the plant expression database PLEXdb [76], with experiment AT40 for a gene expression atlas during Arabidopsis development [73], BB3 for transcriptional changes throughout during barley development [74], and TA3 for comparative transcriptomics in the Triticeae [75]. The three raw datasets were analyzed using the same procedure with Bioconductor packages [82] in the open-source statistical R environment [83]. Briefly, the raw data files were imported into Bioconductor using the Simpleaffy package [84] and normalized and transformed to the log₂ values with the GCRMA package [85]. To get the expression values of flowering genes, each Affymetrix probe set was mapped to an Ensembl gene identifier through BLAST using the flowering gene sequences identified in this study against the target sequences Affymetrix used for the design of these three GeneChips, downloaded from NetAffx Analysis Center (https://www.affymetrix.com/analysis/index.affx). The heatmaps were generated using the heatmap.2 function in the gplots package [83].

3. Results

3.1. Flowering-Related Genes in Arabidopsis, Wheat, and Barley. Table 1 presents a list of 204 Arabidopsis flowering genes compiled through searches on The Arabidopsis Information Resource (TAIR) and the literature [17, 22, 26, 58, 77]. Of these 204 Arabidopsis flowering genes, 190 genes are known to encode proteins and they were broadly (and somewhat arbitrarily in some cases) classified into seven functional groups as in [22, 56, 58]: autonomous (including ambient temperature pathway), flower development, gibberellin, photoperiod, pathway integration, regulation, and vernalization (see Supplemental file I for details). The autonomous pathway consists of genes promoting flowering independently of day-length. The category of flower development includes genes with roles in floral meristem identity and tissue development. The gibberellin (GA) pathway contains genes in GA biosynthesis and metabolism, important for floral transition, and likely inhibits flower formation [86, 87]. The genes in the photoperiod pathway are involved in circadian clock and light signaling. The pathway integration is composed of genes that integrate signals from various flowering pathways. The regulation category contains genes that regulate
other flowering genes at transcriptional, posttranscriptional, epigenetic, and posttranslational levels. The vernalization pathway comprises genes for the prolonged exposure of cold temperature required for flowering. The remaining 14 genes are microRNA genes, which are known to regulate the flowering time [88], but these noncoding genes were excluded for subsequent identification of orthologous protein-coding genes in wheat and barley.

A total of 144 distinct ortholog groups (OGs) for all the flowering proteins in these three species were identified (Supplemental file 2). On average, ~1.5 barley and nearly 5.0 wheat copies were identified for each Arabidopsis flowering gene. The identification of the barley gene set may be incomplete [7, 51] and, as a result, the number of orthologous flowering genes we predicted in barley may be underestimated, which is also due to our conservative approach including both reciprocal BLAST searches and OG clustering. In comparison, 491 flowering genes were identified in soybean, a partially diploidized tetraploid, but with a smaller genome size of 1.1–1.15 Gb [22].

As might be expected, complex orthology relationships exist between the flowering genes identified in these three species, including one-to-one, one-to-many, many-to-one, and many-to-many. The vast majority of OGS contain less than 10 genes in Arabidopsis, wheat, and barley, with all OGS containing <10 Arabidopsis flowering genes (Supplemental file 2). A noteworthy exception is OG5_127136 with only one Arabidopsis gene (AT4G39400), but with 75 wheat orthologs and 45 barley orthologs, each of which represents the largest number of flowering genes identified in these two cereal species. The AT4G39400 gene encodes BRASSINOSTEROID INSENSITIVE 1, which is involved in the autonomous pathway that regulates the transition to flowering, mainly through its effects on FLC gene expression levels [89]. On the other hand, OG5_139532, an ortholog group (OG) known to contain the soybean FLC [22], includes six Arabidopsis genes: FLOWERING LOCUS C (FLC; AT5G1040), AT1G77080 (MAF1/AGL27), AT5G65050 (MAF2/AGL31), AT5G65060 (MAF3), AT5G65070 (MAF4), and AT5G65080 (MAF5/AGL68). But we only detected one ortholog in barley (MLOC_57890), which matches HvOS1 (ODDSOCl; GenBank accession: HM130526) and HvOS2 (ODDSOC2; HM130525) [90, 91], and wheat (Traes_4AS_EliE60C5E5), which matches TaAGL33 (DQ512366), TaAGL41 (DQ512357), and TaMADS2 (DQ534490) [90]. As in Jung et al. [22], we also tested the whole proteome of wheat and barley (instead of proteins containing at least one Arabidopsis flowering gene remains the same (144). This suggests that we have identified the majority of flowering gene orthologs in wheat and barley. This difference between our approach and that of Jung et al. [22] may lead to different false positive and false negative rates in orthology prediction, because using the whole proteome for clustering will likely produce more false orthologs.

Several known flowering genes in wheat and barley have been identified and they offer an opportunity for validation of our approach to ortholog identification. For example, Traes_3B_2A454DB62 and MLOC_68576 represent the FLOWERING LOCUS T (FT) in wheat (TaFT) and barley (HvFT), with the latter already annotated as HvFT in Ensembl. Another example is LFY (AT5G61850), with TaLFY represented by Traes_2AL_83D0D0C3F and Traes_2BL_8DEC0E9BF in wheat and HvLFY represented by MLOC_14305 in barley; all of these three genes have been annotated as LFAY in the Ensembl database. In addition, Traes_2DS_2A961F39D and MLOC_81154 are putative PPD in wheat (Ppd-D1) and barley (Ppd-Hl), respectively. For AtLHY (LATE ELONGATED HYPOCOTYL, AT1G01060), we identified three orthologs in wheat and one in barley (MLOC_14118) (Supplemental file 2). And for AtCCA1 (CIRCADIAN CLOCK ASSOCIATED 1, AT2G46830), we only predicted one ortholog in barley (MLOC_10707) but not in wheat. Previous studies have shown that one homolog of CCA1/LHY exists in grass species including Brachypodium, rice, barley, and wheat [15, 59, 61]. However, discrepancy may exist in our analysis compared with other similar studies, which is generally caused by differences in sequence analysis methods, genomic databases, and parameter settings. For example, Calixto et al. [51] did not find any ortholog of the Arabidopsis ELF4 gene (AtELF4; AT2G40080) in barley and suggested that it might be specific to dicots. However, we identified one putative ortholog each in wheat (Traes_5BL_ECI3571BI on chromosome 5B) and barley (MLOC_58590 on 5H; ELF4-like protein annotated by Ensembl), both of which are single-exon genes like AtELF4.

During our reciprocal BLAST process (using flowering candidate genes identified in wheat and barley to BLAST the Arabidopsis genome), we identified 101 additional Arabidopsis genes that are related to flowering inferred from sequence similarity (Supplemental file 1). Some of these genes may represent those missed in our manual assembly of Arabidopsis flowering genes based on TAIR and literature searches, while the roles of others in flowering will need to be investigated. Because more than 90% of wheat and barley flowering genes are annotated as "uncharacterized protein" or "predicted protein" in Ensembl (Table 2), we annotated these putative flowering genes identified in wheat and barley (Supplemental file 3), using the annotation of their top BLAST hits in Arabidopsis.

### Table 2: Flowering-related genes in barley and wheat that are annotated on the basis of top BLAST hits in Arabidopsis thaliana expressed as the percentage of characterized and uncharacterized proteins/enzymes in current Ensembl annotation.

| Ensembl annotation status | New annotation | Barley | Wheat |
|---------------------------|----------------|--------|--------|
| Characterized             | 6.9% (19)      | 3.8% (34) |
| Uncharacterized           | 93.1% (256)    | 96.2% (866) |
| Total                     | 275            | 900    |
gene clusters are noticeable (Supplemental Figure S1). In Arabidopsis, 50 and 58 flowering genes, respectively, are located on the two longest chromosomes (AT1 and AT5). It is known that four MADS Affecting Flowering (MAF) genes (MAF2, MAF3, MAF4, and MAF5) are clustered in a ~1.4 Mb (mega base pairs) region on AT5 [92]. In barley, chromosome 2H harbors the most (45) flowering genes, which are mainly located at or near the telomere regions. In wheat, the longest chromosome, 3B, contains the largest number (82) of predicted flowering genes. Nevertheless, since physical positions of all 82 flowering genes on 3B and 293 flowering genes on all other chromosomes were unknown, they were randomly assigned on the respective chromosomes as represented by dashed lines in Figure S1. This lack of information on gene position is caused by the incomplete assembly status of the wheat genome: many assemblies have only been performed to the scaffold level (instead of chromosome level). As a result, only 58% (525/900) of the wheat flowering genes have chromosome positions in the latest GFF3 (General Feature Format for genomic features) file released by Ensembl [78], compared with 97% (265/275) of barley flowering genes and 100% of Arabidopsis flowering genes with chromosome positions. Additionally, the orthologs of flowering genes in wheat are often located on the same group of chromosomes. For instance, the ELF3 (AT2G25930) has three wheat orthologs Traes_1AL_52C553IA4, Traes_1BL_B95F8C666, and Traes_1DL_96D83DE2D, which are located on A1, B1, and D1, respectively. The chromosomal locations of the 101 Arabidopsis genes and their corresponding barley and wheat genes are shown in Supplemental Table S1 and Figure S2.

3.3. Exon Intron Organization of Flowering Genes in Arabidopsis, Wheat, and Barley. Motivated by a previous study showing the relationship between gene structure and gene expression in wheat [93], the structural features of the flowering genes in these three species were examined, using the GFF3 files downloaded from Ensembl [78]. As shown in Table 3 (see Supplemental file 4 for details), each Arabidopsis flowering gene has an average of 1.4 transcripts (maximum five transcripts for LHY AT1G01060) with a length of 3161 bp. On average, a barley flowering gene has 2.8 transcripts (MLOC_56110 has 27 transcripts, the biggest number of transcripts in barley flowering genes) with an average length of 4328 bp, and a wheat flowering gene only has one transcript (i.e., no alternative splicing) with an average length of 3815 bp. Arabidopsis flowering genes have 6.5 exons on average, with an average length of 466 bp, while barley and wheat have an average number of 4.5 and 5.7 exons with average length of 878 and 565 bp, respectively. The introns are the longest in wheat flowering genes (924 bp), compared with 468 bp in Arabidopsis and 856 bp in barley. On average, the Arabidopsis flowering proteins are the longest (529 amino acids), compared to 444 and 500 in wheat and barley.

Moreover, the intron length variation in VRN-H1 has been shown to affect vernalization sensitivity in barley [94]. We performed a more detailed intron length analysis in the ortholog groups of these flowering genes. Our results show that, overall, genes in wheat and barley have larger intron sizes than their Arabidopsis homologs in the same ortholog group. For example, the OG OG5 170388 includes API (AT1G69120) and CAL/AGLI0 (AT1G26310, which is known to be homologous to API) with an average of intron length of ~599 bp, nine wheat homologs, with an average intron length of 1761 bp, and one barley gene (MLOC_61901) with average intron length of 2251 bp, which matches VRN-H1 in GenBank (BM5a; AAW82994). However, there are exceptions; in OGS 147177, for example, two Arabidopsis genes (AT1G5550/GA3OX1 and AT1G80340/GA3OX2) have an average intron length of 1598 bp, compared with 486 bp in wheat (six genes: Traes_2AL_B8AB4810, Traes_2BL_9E115B19F, Traes_2BL_FF2BB4810, Traes_2DL_66F9CEA3C, Traes_2DL_F2C4569D7, and Traes_3B_791A6E8DF) and 814 bp in barley (MLOC_12855). Additionally, three OGs (OG5_153242, OG5_160203, and OG5_160330) contain only single-exon genes in all these three species (Supplemental file 4; intron length 0 indicates intronless genes).

Table 3: Structural characteristics of flowering-related genes in Arabidopsis thaliana (AT), Triticum aestivum (TA), and Hordeum vulgare (HV).

| Gene Feature | AT (n = 190) | TA (n = 525) | HV (n = 265) |
|-------------|-------------|-------------|-------------|
| Transcripts per gene | 1.4 (1–5) | 1.0 (1–1) | 2.8 (1–27) |
| Gene length (bp) | 316 (182–16871) | 3815 (240–20952) | 4328 (404–15512) |
| Exons per gene | 6.5 (1–48) | 5.7 (1–42) | 4.5 (1–20) |
| Exon size (bp) | 466 (79–4165) | 565 (42–5550) | 878 (87–5211) |
| Intron size (bp) | 468 (78–2316) | 924 (58–7291) | 856 (44–5912) |
| Protein length (aa) | 529 (77–3259) | 444 (32–3250) | 500 (50–2056) |

The numbers of flowering genes used for the summary statistics are shown in parentheses. Single-exon genes (no introns) were excluded for intron size calculation. bp, base pair; aa, amino acid.
OG5_136555 (an OG known to be involved in light signaling) in these three species share 13 known domains: IPR016132 (Phytochrome chromophore attachment domain), IPR013515 (Phytochrome, central region), IPR03018 (GAF domain), IPR03661 (Signal transduction histidine kinase EnvZ-like, dimerisation/phosphoacceptor domain), IPR001294 Phytochrome, IPR029016 (GAF domain-like), IPR000014 (PAS domain), IPR012129 (Phytochrome A/B/C/D/E), IPR013516 (Phytochrome chromophore binding site), IPR03767 (PAS fold), IPR03654 (PAS fold-2), IPR005467 (Signal transduction histidine kinase, core), and IPR003594 (Histidine kinase-like ATPase, C-terminal domain). This domain analysis further provides confidence in our approach for orthology detection.

The multiple sequence alignments for the MADS-box and PEBP (for phosphatidyethanolamine-binding protein) family proteins show that these domains are more conserved than noncritical regions (Figure 1). The MADS-box near their N-termini is conserved among the genes in the orthogroup OG5_178217, consistent with the fact that the MADS-box is a highly conserved DNA-binding domain; the K-box regions in them are less conserved (Figure 1(a)). In comparison, the highly conserved DNA-binding domain; the K-box regions (Figure 1). The MADS-box near their C-termini contains six genes: FLOWERING LOCUS T (FT), TERMINAL FLOWER 1 (TFL1), and MFT (MOTHER OF FT). While FT induces flowering, TFL1 suppresses flowering, and MFT mainly regulates seed germination [95, 96]. In Arabidopsis, the PEBP family contains six genes: FLOWERING LOCUS T (FT), TERMINAL FLOWER1 (TFL1), TWIN SISTER OF FT (TSF), BROTHER OF FT AND TFL1 (BFT), CENTRORADIALIS (ATC), and MOTHER OF FT AND TFL1 (MFT). We identified nine PEBP genes in barley and 58 PEBP genes in wheat (Supplemental file 2). Five of the nine barley PEBP genes match those reported in [52, 95]: MLOC_68576 ( HvFT) matches HvFT1 (DQ100327), HvFT2 (DQ297407), HvFT3 (DQ411319), and HvFT5 (EF012202); MLOC_13102 and MLOC_71606 are related to HvMFT1 (AB447466); MLOC_74854 is similar to HvFT4 (DQ411320); and MLOC_35818 corresponds to HvTFL1(AB447465). No HvCEN or HvBFT was reported in [52, 95]. In Arabidopsis, there are 19 flowering genes containing a MADS-box domain, including FLOWERING LOCUS C (FLC; AT5G10140) and MAF2 to MAF5 (AT5G65050, AT5G65060, AT5G65070, and AT5G65080). We identified eight and 44 MADS proteins in barley and wheat, respectively. Most of these flowering MADS proteins usually also contain a K-box region (IPR002487) near their C-termini (Figure 1(a)).

In the phylogenetic tree, the three subfamilies are clearly divided into three clades (Figure 2(a)), a topology similar to the phylogenetic relationship of FT proteins in Arabidopsis, Brachypodium, rice, and barley previously reported [61]. Interestingly, PEBP genes with known antagonistic roles in flowering responses are in different clades: FT and TFL, two floral inducers, are in one clade, whereas ATC and TFL, two floral inhibitors, are in another. Also, the wheat gene Traes_3B_2A454DB62 is phylogenetically close to AtFT (AT1G65480) and AtTSF (AT4G20370). It is annotated as “uncharacterized protein” in Ensembl, and from our BLAST analysis it is a good hit of Arabidopsis FT (E-value = 5.00e-48; Supplemental file 3). For the MADS-box proteins, we carried out phylogenetic analysis of OG5_178217, which includes Arabidopsis AGL12 (ATIG71692). As shown in Figure 2(b), two cladest were formed, one for AtAGL12 and the other for the five MADS genes in wheat and barley. The latter is further divided into two branches: one for the two wheat genes on the group of 2 chromosomes and the other for four genes on the group of 7 chromosomes (barley MLOC_53973 on 7H).

3.6. Expression Profiles of PEBP, MADS, and B-Box Family Genes in Different Organs of Arabidopsis, Wheat, and Barley. Three similar, independent microarray gene expression datasets for Arabidopsis [73], wheat [75], and barley [74] are available in PLEXdb [76], thereby enabling us to analyze the expression profiles of these flowering genes in a wide range of tissues and developmental stages. These three transcriptome datasets were all obtained using the Affymetrix GeneChip platforms and tissues and developmental stages sampled throughout a plant life cycle (Table 4). Additionally, the experimental design of wheat TA3 mirrored that of barley BB3, with 13 of 15 nearly identical tissues [74, 75]. According to our analysis of the 273 raw data files (three replicates for each sample), 189 of 190 Arabidopsis flowering genes were expressed in at least one of the 65 tissues and developmental stages. In barley, 248 (~91%) of the 275 flowering genes are expressed in at least one of the 15 tissue types. Likewise, 676 (~75%) of the 900 putative flowering genes in wheat were expressed in at least one of the 13 tissue types. These percentages for wheat and barley were lower because not all flowering genes we identified were on these two microarrays that were designed using EST (expressed sequence tag) collections (rather than whole genome sequences) in both species (Table 4) [74, 75, 97]. The normalized expression values of flowering genes in Arabidopsis, wheat, and barley are shown in Supplemental file 5. An overview of flowering gene expression in different tissues and development stages of Arabidopsis, wheat, and barley are shown in Figure 3. As evident from the tissue dendrograms, the pollen in Arabidopsis and anthers (before anthesis) in both wheat and barley showed drastically different expression profiles from other tissues. We identified three, 21, and 23 highly expressed flowering genes in Arabidopsis, wheat, and barley, respectively, as represented by the green bands in the heat maps with average log2 expression values >10 across all analyzed samples. All three Arabidopsis genes belong to the photoperiod pathway, the green-coded (highly expressed) wheat genes include eight regulatory genes and eight genes in flower development, and the barley green-coded genes include 10 regulatory genes, four photoperiod genes, and five genes related to flower development.

The expression patterns of genes in the different OGs were further compared among the PEBP, MADS-box, and B-box families. These three important gene families contain
key genes in the control of flowering time, such as CO (CONSTANS), FLC, FT, FUL (FRUITFULL), and SOC1 (SUPPRESSOR OF OVEREXPRESSION OF CONSTANS) [23, 95, 96, 98–104]. As no similar Arabidopsis tissues corresponding to those used for expression profiling experiments in wheat and barley were used in the AtGenExpress experiments, we only compared the tissue-specific expression patterns of flowering genes between wheat and barley. Two additional samples (10 DAP caryopsis and 16 DAP caryopsis) exist in barley BB3, which were removed in the barley gene expression dataset in order to compare expression of flowering genes in equivalent tissues of these two cereal species.

Figure 1: The multiple sequence alignment of OG5_J78217 including conserved MADS-box domain and K-box region (a) and PEBP family proteins (b). (a) The MADS-box domain (IPR002100) is marked blue, and the K-box region (IPR002487) marked red. (b) The PEBP domain (IPR008914) in Arabidopsis FT protein encoded by AT1G65480 spans from 27 to 161 amino acids.
Figure 2: Phylogeny of PEBP (a) and MADS (b) family proteins in Arabidopsis, wheat, and barley. The PEBP proteins include 11 sequences in OG5,146543 (see Supplemental file 2), and a PEBP protein Pp1s34,16V6 in moss was used as an outgroup to root the phylogenetic tree. The MADS AGT12 proteins include six sequences in OG5,178217, and a MADS protein Phpat.004G002000.1 in moss was used as an outgroup to root the phylogenetic tree. The support value on each node is the Bayesian posterior probability. The scale bar denotes the number of nucleotide replacements per site.

Figure 4 shows the expression profiles of the major OGs in these three families. The PEBP family proteins were clustered into four OGs: OG5,127642, OG5,146543, OG5,158796, and OG5,163093. Among the three genes in OG5,158796 (Figure 4(a)), Traes.5BL_E6535628C and MLOC.44160 (HvCEN) show higher expression in seedling roots, and the other barley gene MLOC.35818 (HvTFL1) was relatively weakly and stably expressed in all these tissues. OG5,146543 includes three wheat genes and one barley gene. Traes.3B.2A454DB62 (putative TaFT) shows higher expression in immature inflorescence, floral bracts (before anthesis), 3–5 DAP caryopsis, and 22 DAP endosperm, and the barley FT gene MLOC.68576 (HvFT) has relatively high expression in all the tissues, especially in the 22 DAP endosperm (Figure 4(b)). The comparative expression profiles of additional PEBP genes in ortholog groups OG5,127642 and OG5,163093 are shown in Supplemental Figures S3(A) and S3(B), which also include genes with similar expression patterns, such as Traes.3DS.E0EF3E9AB and MLOC.74854 (HvFT4) in OG5,127642.

The MADS-box family of flowering genes was clustered into 11OGs; yet only six OGs (OG5,135817, OG5,177438, OG5,190130, OG5,144912, OG5,170388, and OG5,178217) have expression data for both wheat and barley flowering genes; OG5,212214, OG5,212591, OG5,139532, OG5,164556, and OG5,211687 have no barley and/or wheat gene expression data for this comparison. Clearly, many MADS flowering genes in each OG show similar expression patterns in the tissues examined in both wheat and barley (Figures 4(c) and 4(d)). For example, the wheat gene Traes.2DL.71120931 and its barley orthologous gene MLOC.35973 in the ortholog group OG5,178217 exhibit...
similar expression patterns (Figure 4(c)). Three wheat genes 
Traes_5AL_13E2DEC48, Traes_5DS_B05596869 (TaVRN1), and Traes_2DL_903A29C6A and their barley orthologous gene MLOC_61901 (VRN-H1) in OG5_170388 show strikingly similar expression profiles, with elevated expression levels in reproductive tissues including immature inflorescence, floral bracts, pistil, anthers, and 3–5 DAP caryopsis (Figure 4(d)). We also analyzed two microarray datasets after cold and/or light treatments in wheat (NCBI GEO accession: GSE11774) and barley (Plexdb accession: BB94) and found that both Traes_5DS_B05596869 and MLOC_61901 exhibited an expression profile consistent with that of TaVRN1 and HvVRN1, respectively, as in [86, 105]. Most additional MADS genes in OG5_144912, OG5_177438, OG5_135817, and OG5_190310 also show similar expression profiles (Supplemental Figures S3(C)–S3(F)).

The B-box (BBX) family of transcription factors contains a zinc-finger and B-box domain (IPR000315) with one or two B-box motifs and sometimes also includes a CCT (CONSTANS, CO-like, and TOC1) domain (IPR010402) [104]. The BBX family proteins were clustered into five OGs: OG5_139246, OG5_156319, OG5_178368, OG5_170758, and OG5_170476 (no barley and wheat flowering genes in this OG). The expression profiles of the orthologous BBX genes in OG5_178368 and OG5_170758 are shown in Figures 4(e) and 4(f). Again, similar expression profiles exist in the BBX family genes. The two wheat genes Traes_5DL_8CE2482E6, Traes_5AL_852A1474C and their barley ortholog MLOC_57021 (HvPrr95) in OG5_178368 exhibit comparable expression profiles across the 13 tissues (Figure 4(e)). As shown in Figure 4(f), two wheat genes Traes_6AL_A0A31AA9F and Traes_6DL_C215BACFD, as well as their barley orthologous gene MLOC_52387 (HvTOC1) in OG5_170758, were all relatively highly expressed in these 13 tissues. Moreover, two wheat genes Traes_2AS_2FCD9730 and Traes_4DL_EE41726EA and the two barley genes MLOC_81154 and MLOC_12732 in OG5_139246, as well as Traes_6DL_036293C55 and MLOC_6921 (putative HvCO) in OG5_156319, share similar expression profiles (Supplemental Figures S3(G) and S3(H)). The orthologous genes with similar expression patterns (together with sequence-based homology) in a variety of tissues and development stages are more likely to maintain similar functions related to flowering in wheat and barley.

In addition, when multiple wheat paralogs exist in an OG, some of them are virtually unexpressed (inactive) in the examined tissues. For example, in the PEBP family, both Traes_3B_C8DBBCDC0E and Traes_7AS_EBD5F1F54 in OG5_146543 were nearly unexpressed in all these tissues (Figure 4(b)). Both Traes_2BL_E99788B8C in the MADS family and Traes_6BL_ED40C8806 in BBX family also appear to be unexpressed (Figures 4(c) and 4(f); more examples in Supplemental Figure 3S). Taken together, our expression analysis is consistent with previous studies indicating that homoeologous genes in hexaploid bread wheat can be expressed from one, two, or three homeologues [75, 106].

4. Discussion

The release of genomic sequences of wheat and barley [6–8] provides a new opportunity for inferring genes and their functions that are agronomically and economically important but yet poorly characterized in these crops through comparative assessment of sequence similarity with the same genes that are well characterized in the model plants. In this study, we used a bioinformatic approach (i.e., reciprocal BLAST searches coupled with OrthoMCL clustering) for identification of putative flowering-related genes in wheat and barley from the known flowering genes in Arabidopsis. Further comparative genomics analyses of these flowering genes in Arabidopsis, wheat, and barley enabled the formation of ortholog groups. Orthologous flowering genes in wheat and barley are often clustered on the same chromosomes, and their exon–intron architectures and key domains are generally conserved.

The intron length of flowering genes in barley and particularly in wheat is generally larger than that of their
Figure 3: Overview of expression of flowering genes in different tissues and development stages in Arabidopsis (a), wheat (b), and barley (c). The expression data Arabidopsis, wheat, and barley was from 63, 13, and 15 tissue types, respectively (Table 4, Supplemental file 5). The heat maps were created by hierarchical clustering using complete linkage method with the heatmap.2 function in R. The same color key shown is used for all the three heat maps.
Figure 4: Continued.
Arabidopsis homologs (Supplemental file 4), consistent with the comparison of the size of introns in 21 clock genes in Arabidopsis and barley [51]. Szucs et al. [94] showed that the intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity in barley, and thus the consequence of large introns in many cereal genes will need further investigation. Our domain analysis showed that most of the orthologous flowering proteins share one or more known InterPro domains (Supplemental file 2). As the complete sequence of many cereal genes exhibits low sequence similarity to Arabidopsis genes but shares a higher degree of sequence conservation within protein functional domains [48], domain analysis may play a more important role in prediction of flowering orthologous proteins in monocot crop species.

As our analysis was based on a list of known Arabidopsis flowering genes, we could only find genes with sequence similarity above the threshold in wheat and barley. However, it is known that genes in the vernalization response are not conserved in dicot and monocot between them [48], and our sequence analysis indicates that, compared to those in other pathways, vernalization genes show lowest sequence similarity between Arabidopsis and wheat or barley (Table 5). The parameter setting in reciprocal BLAST [68] and OrthoMCL clustering [69, 70] can also affect the results. In addition, the gene prediction in these two cereal genomes is still incomplete, particularly for barley [7]. As a consequence, our approach shows different performance in different groups of flowering genes. Future studies may focus on a gene family or genes in a flowering pathway, taking into account other sets of genome neighborhood information such as synteny (colinearity), which is particularly important for genes that are less conserved at the sequence level. For example, Ruelens

**Table 5:** The average percentage of protein sequence similarity of flowering genes in the seven functional groups in Arabidopsis, wheat, and barley.

| Functional group | AT versus TA | AT versus HV | HV versus TA |
|-----------------|--------------|--------------|--------------|
| Autonomous | 62.72 (16.91) | 53.36 (14.48) | 97.32 (2.28) |
| Flower development | 68.77 (11.83) | 63.23 (12.39) | 95.19 (6.21) |
| Gibberellin | 61.27 (15.07) | 54.71 (10.95) | 96.85 (2.01) |
| Pathway integration | 74.75 (13.39) | 58.95 (16.97) | 95.44 (3.81) |
| Photoperiod | 61.44 (17.99) | 57.37 (16.25) | 94.48 (5.99) |
| Regulation | 69.78 (13.04) | 61.97 (13.49) | 96.49 (3.75) |
| Vernalization | 58.88 (8.31) | 44.63 (13.04) | 81.64 (15.04) |

AT versus TA, comparison of flowering protein sequences between Arabidopsis and wheat; AT versus HV, comparison of flowering protein sequences between Arabidopsis and barley; HV versus TA, comparison of flowering protein sequences between barley and wheat. The values in the parentheses are standard deviations.
et al. [91] identified two and three FLC-like genes, respectively, in barley and wheat, using an approach that combines phylogenetic reconstruction and genome synteny.

It is evident from our in silico expression analyses that many orthologous genes showed similar expression profiles in different tissues of wheat and barley, and sometimes one or more wheat paralogs in an OG were virtually unexpressed in all the thirteen tissues (Figure 4 and Figure S3). These results suggest potential functional conservation and divergence of flowering genes in these two Triticaceae species. The unexpressed paralogs in more than 10 developmental stages likely represent pseudogenes.

Several factors affected our in silico expression analysis of flowering genes in these species. First, the orthology between barley and wheat genes can be a one-to-many or many-to-many (i.e., not a simple one-to-one) relationship, which can complicate the comparison of their expression profiles. Second, the reliability of gene expression data obtained with the wheat GeneChip can be affected by the fact that wheat is hexaploid with approximately 80% repeats in the genome [6, 75]. Third, as Druka et al. [74] pointed out, the spatial resolution over which they have measured gene expression is low and only a single barley cultivar (Morex) was used. (Similar expression data for the barley cv. Golden Promise can be found but only six tissue types were surveyed.) Lastly and importantly, it would be more interesting to compare the expression profiles of these flowering genes in genotypes with various photoperiod sensitivity and/or vernalization requirements or after different daylength and/or cold temperature treatments, as the expression of many flowering genes is induced by external conditions suitable for flowering [107–109]. For example, two public microarray data sets exist for transcriptomic changes in wheat and barley under the inductive conditions required for flowering [86, 105]; yet the differences in treatments (both cold and light treatments in [86] versus cold treatment in [105]) and tissues sampled (leaf/crown in [86] versus whole plant in [105]) for these two profiling experiments make it difficult to compare the expression patterns of orthologous genes in wheat and barley.

This study has important implication for genetic improvement of early flowering and related traits in wheat, barley, or other cereals. We annotated functions of many flowering-related genes in wheat and barley from known flowering genes identified in Arabidopsis. Of all the annotated genes, those responsible for vernalization and photoperiod are the two most important functional gene groups, accounting for about 70–75% and 20–25% of the genetic variability in the flowering time of wheat, for example, [110, 111]. In western Canada where the growing season for cereal crops is short (95–125 days) with long daylength (>14 h), breeding for early flowering would be most effective with its focus on the use of vernalization genes. For the 20 vernalization genes in Arabidopsis, only eight genes were found in barley and 31 in wheat (cf. Table I and Supplemental file 2).

However, while such annotation of vernalization genes is an important first step towards genetic improvement of early flowering in cereal crops, these annotated genes need to be verified before incorporating them into practical breeding programs. Our sequence analysis (Table 5) and several other studies [10, 43, 45, 50] indicated divergence of genes responsible for vernalization response between monocots (e.g., wheat and barley) and dicots (e.g., Arabidopsis). For example, as described earlier, one of the major vernalization genes, Vrn2, in wheat and other cereals does not have a clear ortholog in Arabidopsis whereas another main vernalization gene in cereals, Vrn1, is homologous to genes that encode proteins APETTALA1 and FRUITFUL with no role in vernalization response in Arabidopsis. For this reason, recent attempts (e.g., [63]) have been made to use phylogenetically more similar cereal species (e.g., rice or Brachypodium) as a more immediate model organism for characterization of flowering genes in wheat and barley. However, genetic resources for gene annotation and characterization in rice or Brachypodium remain limited in comparison to those in Arabidopsis. Thus, molecular breeding for early flowering and other agronomically important traits in wheat and barley will continue to benefit from comparative genomic analysis with Arabidopsis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is funded by the Growing Forward 2 Research Opportunities and Innovation Internal Initiatives of Alberta Agriculture and Forestry and by the Natural Sciences and Engineering Research Council of Canada Grant OGP0183983 to Rong-Cai Yang.

References

[1] J. Dvořák, P. di Terlizzi, H.-B. Zhang, and P. Resta, “The evolution of polyploid wheats: identification of the A genome donor species,” Genome, vol. 36, no. 1, pp. 21–31, 1993.
[2] J. Dvořák and H.-B. Zhang, “Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and G genomes,” Proceedings of the National Academy of Sciences of the United States of America, vol. 87, no. 24, pp. 9640–9644, 1990.
[3] T. Naranjo and E. Corredor, “Clustering of centromeres precedes bivalent chromosome pairing of polyploid wheats,” Trends in Plant Science, vol. 9, no. 5, pp. 214–217, 2004.
[4] K. M. Devos, J. Doležel, and C. Feuillet, “Genome organization and comparative genomics,” Wheat Science and Trade, pp. 327–367, 2009.
[5] G. Moore, “Cereal genome evolution: Pastoral pursuits with “Lego” genomes,” Current Opinion in Genetics and Development, vol. 5, no. 6, pp. 717–724, 1995.
[6] R. Brenchley, M. Spannagl, M. Pfeifer et al., “Analysis of the bread wheat genome using whole-genome shotgun sequencing,” Nature, vol. 491, no. 7426, pp. 705–710, 2012.
[7] The International Barley Genome Sequencing Consortium, “A physical, genetic and functional sequence assembly of the barley genome,” Nature, vol. 491, no. 7426, pp. 711–716, 2012.
[8] The International Wheat Genome Sequencing Consortium, “A chromosome-based draft sequence of the hexaploid bread
wheat (Triticum aestivum) genome,” *Science*, vol. 345, no. 6194, Article ID 1251788, 2014.

[9] A. N. Dodd, N. Salathia, A. Hall et al., “Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage,” *Science*, vol. 309, no. 5734, pp. 630–633, 2005.

[10] J. Cockram, H. Jones, F. J. Leigh et al., “Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity,” *Journal of Experimental Botany*, vol. 58, no. 6, pp. 1231–1244, 2007.

[11] R. Nitcher, S. Pearce, G. Tranquilli, X. Zhang, and J. Dubcovsky, “Effect of the hope FT-B1 allele on wheat heading time and yield components,” *Journal of Heredity*, vol. 105, no. 5, pp. 666–675, 2014.

[12] S. Thepot, G. Restoux, I. Goldringer et al., “Efficiently tracking selection in a multiparental population: the case of earliness in wheat,” *Genetics*, vol. 199, no. 2, pp. 609–623, 2015.

[13] R.-C. Yang and B. J. Ham, “Stability of genome-wide QTL effects on malt α-amylase activity in a barley doubled-haploid population,” *Euphytica*, vol. 188, no. 1, pp. 131–139, 2012.

[14] I. R. Henderson and C. Dean, “Control of Arabidopsis flowering: the chill before the bloom,” *Development*, vol. 131, no. 16, pp. 3829–3838, 2004.

[15] A. S. Turner, S. Faure, Y. Zhang, and D. A. Laurie, “The effect of day-neutral mutations in barley and wheat on the interaction between photoperiod and vernalization,” *Theoretical and Applied Genetics*, vol. 126, no. 9, pp. 2267–2277, 2013.

[16] S. Fjellheim, S. Boden, and B. Trevaskis, “The role of seasonal flowering responses in adaptation of grasses to temperate climates,” *Frontiers in Plant Science*, vol. 5, 2014.

[17] F. Fornara, A. de Montaigu, and G. Coupland, “SnapShot: control of flowering in Arabidopsis,” *Cell*, vol. 141, no. 3, pp. 550–550.e2, 2010.

[18] N. Nakamichi, T. Kiba, R. Henriques, T. Mizuno, N.-H. Chua, X. Zhao, M. S. Liu, J. R. Li, and et al., “The wheat TüGII, involved in photoperiodic flowering, encodes an Arabidopsis GI ortholog,” *Plant Molecular Biology*, vol. 58, no. 1, pp. 53–64, 2005.

[19] J. Dubcovsky, A. Loukoianov, D. Fu, M. Valarik, A. Sanchez, and L. Yan, “Effect of photoperiod on the regulation of wheat vernalization genes VRN1 and VRN2,” *Plant Molecular Biology*, vol. 60, no. 4, pp. 469–480, 2006.

[20] A. Chen and J. Dubcovsky, “Wheat TILLING mutants show that the vernalization gene VRN1 down-regulates the flowering repressor VRN2 in leaves but is not essential for flowering,” *PLoS Genetics*, vol. 8, no. 12, Article ID e1003134, 2012.

[21] D. Gomez, L. Vanzetti, M. Helguera, L. Lombardo, J. Fraschini, and D. J. Miralles, “Effect of Vrn-1, Ppd-1 genes and earliness per se on heading time in Argentinean bread wheat cultivars,” *Field Crops Research*, vol. 158, pp. 73–81, 2014.

[22] A. Turner, J. Beales, S. Faure, R. P. Dunford, and D. A. Laurie, “The pseudo-response regulator Ppd-HI provides adaptation to photoperiod in barley,” *Science*, vol. 310, no. 5750, pp. 1031–1034, 2005.

[23] R. P. Dunford, S. Griffiths, V. Christodoulou, and D. A. Laurie, “Characterisation of a barley (Hordeum vulgare L.) homologue of the Arabidopsis flowering time regulator GIGANTEA,” *Theoretical and Applied Genetics*, vol. 110, no. 5, pp. 925–933, 2005.

[24] B. Trevaskis, M. N. Hemming, W. J. Peacock, and E. S. Dennis, “HvVRN2 responds to daylength, whereas HvVRN1 is regulated by vernalization and developmental status,” *Plant Physiology*, vol. 140, no. 4, pp. 1397–1405, 2006.

[25] C. Campoli, B. Drosos, I. Searle, G. Coupland, and M. Von Korff, “Functional characterisation of HvCO1, the barley (Hordeum vulgare) flowering time ortholog of CONSTANS,” *Plant Journal*, vol. 69, no. 5, pp. 868–880, 2012.

[26] S. Faure, A. S. Turner, D. Gruszka et al., “Mutation at the flowering time regulator EARLY FLOWERING3 regulates flowering in spring barley by mediating Gibberellin production and FLOWERING LOCUS T expression,” *The Plant Cell*, vol. 26, no. 4, pp. 1557–1569, 2014.

[27] A. M. Alqudah, R. Sharma, R. K. Pasam, A. Graner, B. Kilian, and T. Schnurbusch, “Genetic dissection of photoperiodic response based on gwas of pre-anthesis phase duration in spring barley,” *PLoS ONE*, vol. 9, no. 11, Article ID e113120, 2014.
[42] A. Distelfeld, C. Li, and J. Dubcovsky, "Regulation of flowering in temperate cereals," Current Opinion in Plant Biology, vol. 12, no. 2, pp. 178–184, 2009.

[43] A. Greenup, W. J. Peacock, E. S. Dennis, and B. Trevaskis, "The molecular biology of seasonal flowering-responses in Arabidopsis and the cereals," Annals of Botany, vol. 103, no. 8, pp. 1165–1172, 2009.

[44] B. Trevaskis, "The central role of the VERNALIZATION1 gene in the vernalization response of cereals," Functional Plant Biology, vol. 37, no. 6, pp. 479–487, 2010.

[45] B. Trevaskis, M. N. Hemming, E. S. Dennis, and W. J. Peacock, "The molecular basis of vernalization-induced flowering in cereals," Trends in Plant Science, vol. 12, no. 8, pp. 352–357, 2007.

[46] C. R. McClung, "A modern circadian clock in the common angiosperm ancestor of monocots and eudicots," BMC Biology, vol. 8, article 55, 2010.

[47] C. R. McClung and R. A. Gutierrez, "Network news: prime time for systems biology of the plant circadian clock," Current Opinion in Genetics and Development, vol. 20, no. 6, pp. 588–598, 2010.

[48] M. Blümel, N. Dally, and C. Jung, "Flowering time regulation in crops—what did we learn from Arabidopsis?" Current Opinion in Biotechnology, vol. 32, pp. 121–129, 2015.

[49] Y. H. Song, J. S. Shim, H. A. Kinmonth-Schultz, and T. Imaizumi, "Photoperiodic flowering: time measurement mechanisms in leaves," Annual Review of Plant Biology, vol. 66, pp. 441–464, 2015.

[50] J. Colasanti and V. Coneva, "Mechanisms of floral induction in grasses: something borrowed, something new," Plant Physiology, vol. 149, no. 1, pp. 56–62, 2009.

[51] C. P. G. Calixto, R. Waugh, and J. W. S. Brown, "Evolutionary relationships among barley and Arabidopsis core circadian clock and clock-associated genes," Journal of Molecular Evolution, vol. 80, no. 2, pp. 108–119, 2015.

[52] S. Faure, J. Higgins, A. Turner, and D. A. Laurie, "The FLOWERING LOCUST-like gene family in barley (Hordeum vulgare)," Genetics, vol. 176, no. 1, pp. 599–609, 2007.

[53] R. Nitcher, A. Distelfeld, C. Tan, L. Yan, and J. Dubcovsky, "Increased copy number at the HvFT1 locus is associated with accelerated flowering time in barley," Molecular Genetics and Genomics, vol. 288, no. 5–6, pp. 261–275, 2013.

[54] L. Yan, D. Fu, C. Li et al., "The wheat and barley vernalization gene VRN3 is an orthologue of FT," Proceedings of the National Academy of Sciences of the United States of America, vol. 103, no. 51, pp. 19581–19586, 2006.

[55] V. Hecht, F. Foucher, C. Ferrándiz et al., "Conservation of Arabidopsis flowering genes in model legumes," Plant Physiology, vol. 137, no. 4, pp. 1420–1434, 2005.

[56] M. Y. Kim, Y. J. Kang, T. Lee, and S.-H. Lee, "Divergence of flowering-related genes in three legume species," The Plant Genome, vol. 6, no. 3, 2013.

[57] S. K. Kim, T. Lee, Y. J. Kang et al., "Genome-wide comparative analysis of flowering genes between Arabidopsis and mungbean," Genes & Genomics, vol. 36, no. 6, pp. 799–808, 2014.

[58] C. E. Grover, J. P. Gallagher, and J. F. Wendel, "Candidate gene identification of flowering time genes in Cotton," The Plant Genome, 2015.

[59] M. Murakami, Y. Tago, T. Yamashino, and T. Mizuno, "Comparative overviews of clock-associated genes of Arabidopsis thaliana and Oryza sativa," Plant and Cell Physiology, vol. 48, no. 1, pp. 110–121, 2007.

[60] H. Tsuji, K.-I. Taoka, and K. Shimamoto, "Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation," Current Opinion in Plant Biology, vol. 14, no. 1, pp. 45–52, 2011.

[61] J. A. Higgins, P. C. Bailey, and D. A. Laurie, "Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate Grasses," PLoS ONE, vol. 5, no. 4, Article ID e00605, 2010.

[62] J. Brljacić, E. Grotewold, R. Scholl et al., "Brachypodium as a model for the grasses: today and the future," Plant Physiology, vol. 157, no. 1, pp. 3–13, 2011.

[63] L. A. J. Mur, J. Allaingguillaume, P. Catalán et al., "Exploiting the brachypodium tool box in cereal and grass research," New Phytologist, vol. 191, no. 2, pp. 334–347, 2011.

[64] S. Griffiths, R. P. Dunford, G. Coupland, and D. A. Laurie, "The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis," Plant Physiology, vol. 131, no. 4, pp. 1855–1867, 2003.

[65] C. Campoli, M. Shtaya, S. J. Davis, and M. von Korff, "Expression conservation within the circadian clock of a monocot: natural variation at barley Ppd-H1 affects circadian expression of flowering time genes, but not clock orthologs," BMC Plant Biology, vol. 12, article 97, 2012.

[66] J. Cockram, T. Thiel, B. Steuernagel et al., "Genome dynamics explain the evolution of flowering time CCT domain gene families in the Poaceae," PLoS ONE, vol. 7, no. 9, Article ID e45307, 2012.

[67] C. Campoli, A. Pankin, B. Drosse, C. M. Casao, S. J. Davis, and M. Von Korff, "HvLUX1 is a candidate gene underlying the early maturity 10 locus in barley: phylogeny, diversity, and interactions with the circadian clock and photoperiodic pathways," New Phytologist, vol. 199, no. 4, pp. 1045–1059, 2013.

[68] S. F. Altschul, T. L. Madden, A. A. Schaffer et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," Nucleic Acids Research, vol. 25, no. 17, pp. 3389–3402, 1997.

[69] L. Li, C. J. Stoeckert Jr., and D. S. Roos, "OrthoMCL: identification of ortholog groups for eukaryotic genomes," Genome Research, vol. 13, no. 9, pp. 2178–2189, 2003.

[70] F. Chen, A. J. Mackey, C. J. Stoeckert Jr., and D. S. Roos, "OrthoMCL-DB: query a comprehensive multi-species collection of ortholog groups," Nucleic Acids Research, vol. 34, pp. D363–D368, 2006.

[71] S. Hunter, P. Jones, A. Mitchell et al., "InterPro in 2011: new developments in the family and domain prediction database (vol 40, pg D306, 2011)," Nucleic Acids Research, vol. 40, no. 10, p. 4725, 2012.

[72] P. Jones, D. Binns, H.-Y. Chang et al., "InterProScan 5: genome-scale protein function classification," Bioinformatics, vol. 30, no. 9, pp. 1236–1240, 2014.

[73] M. Schmid, T. S. Davison, S. R. Henz et al., "A gene expression map of Arabidopsis thaliana development," Nature Genetics, vol. 37, no. 5, pp. 501–506, 2005.

[74] A. Druka, G. Muehlbauer, I. Druka et al., "An atlas of gene expression from seed to seed through barley development," Functional & Integrative Genomics, vol. 6, no. 3, pp. 202–211, 2006.

[75] A. W. Schreiber, T. Sutton, R. A. Caldo et al., "Comparative transcriptomics in the Triticaceae," BMC Genomics, vol. 10, article 285, 2009.
[76] S. Dash, J. Van Hemert, L. Hong, R. P. Wise, and J. A. Dickerson, "PLEXdb: gene expression resources for plants and plant pathogens," *Nucleic Acids Research*, vol. 40, no. 1, pp. D1194–D1201, 2012.

[77] P. Lamesch, T. Z. Berardini, D. Li et al., "A gene expression map of Arabidopsis thaliana development," *Nucleic Acids Research*, vol. 40, no. 1, pp. D1202–D1210, 2012.

[78] F. Cunningham, M. R. Amode, D. F. Cunningham et al., "Ensembl 2015," *Nucleic Acids Research*, vol. 43, no. 1, pp. D662–D669, 2015.

[79] M. A. Larkin, G. Blackshields, N. P. Brown et al., "Clustal W and Clustal X version 2.0," *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, 2007.

[80] A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, and G. J. Barton, "Jalview Version 2-A multiple sequence alignment editor and analysis workbench," *Bioinformatics*, vol. 25, no. 9, pp. 1189–1191, 2009.

[81] A. J. Drummond, M. A. Suchard, D. Xie, and A. Rambaut, "Bayesian phylogenetics with BEAUti and the BEAST 1.7," *Molecular Biology and Evolution*, vol. 29, no. 8, pp. 1969–1973, 2012.

[82] R. C. Gentleman, V. J. Carey, D. M. Bates et al., "Biocductor: open software development for computational biology and bioinformatics," *Genome Biology*, vol. 5, R80, 2004.

[83] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014, http://www.R-project.org/.

[84] C. L. Wilson and C. J. Miller, "Simpleaxe: a BioConductor package for affymetrix quality control and data analysis," *Bioinformatics*, vol. 21, no. 18, pp. 3683–3685, 2005.

[85] Z. Wu, R. A. Irizarry, R. Gentleman, F. Martinez-Murillo, and F. Spencer, "A model-based background adjustment for oligonucleotide expression arrays," *Journal of the American Statistical Association*, vol. 99, no. 468, pp. 909–917, 2004.

[86] M. O. Winfield, C. Lu, I. D. Wilson, J. A. Coghill, and K. J. Edwards, "Cold- and light-induced changes in the transcriptome of wheat leading to phase transition from vegetative to reproductive growth," *BMC Plant Biology*, vol. 9, article 55, 2009.

[87] N. Yamaguchi, C. M. Winter, M.-F. Wu et al., "Gibberellin acts positively then negatively to control onset of flower formation in Arabidopsis," *Science*, vol. 344, no. 6184, pp. 638–641, 2014.

[88] E. Spanudakis and S. Jackson, "The role of microRNAs in the control of flowering time," *Journal of Experimental Botany*, vol. 65, no. 2, pp. 365–380, 2014.

[89] L. Hategan, B. Godza, L. Kozma-Bognar, G. J. Bishop, and E. Spanudakis and S. Jackson, "The role of microRNAs in the tandem origin of angiosperm-specific MADS-box genes," *Molecular Phylogenetics and Evolution*, vol. 75, no. 2, pp. 3081–3098, 2012.

[90] R. Kikuchi, H. Kawahigashi, T. Ando, T. Tonooka, and H. Handa, "Molecular and functional characterization of PEBP genes in barley reveal the diversification of their roles in flowering," *Plant Physiology*, vol. 149, no. 3, pp. 1341–1353, 2009.

[91] A. Karlgren, N. Gyllenstrand, T. Källman et al., "Evolution of the PEBP gene family in plants: functional diversification in seed plant evolution," *Plant Physiology*, vol. 156, no. 4, pp. 1967–1977, 2011.

[92] T. J. Close, S. I. Wánamaker, R. A. Caldo et al., "A new resource for cereal genomics: 22K barley GeneChip comes of age," *Plant Physiology*, vol. 134, no. 3, pp. 960–968, 2004.

[93] A. Becker and G. Theissen, "The major clades of MADS-box genes and their role in the development and evolution of flowering plants," *Molecular Phylogenetics and Evolution*, vol. 29, no. 3, pp. 464–489, 2003.

[94] Y. Nemoto, M. Kisaka, T. Fuse, M. Yano, and Y. Ogihara, "Characterization and functional analysis of three wheat genes with homology to the CONSTANS flowering time gene in transgenic rice," *The Plant Journal*, vol. 36, no. 1, pp. 82–93, 2003.

[95] B. Trevaskis, D. J. Bagnall, M. H. Ellis, W. J. Peacock, and E. S. Dennis, "MADS box genes control vernalization-induced flowering in cereals," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 22, pp. 13099–13104, 2003.

[96] C. Smaczniak, R. G. H. Immink, G. C. Angenent, and K. Kaufmann, "Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies," *Development*, vol. 139, no. 17, pp. 3081–3098, 2012.

[97] C. Smaczniak, R. G. H. Immink, J. M. Muñó et al., "Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 5, pp. 1560–1565, 2012.

[98] S. Torti, F. Fornera, C. Vincent et al., "Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering," *The Plant Cell*, vol. 24, no. 2, pp. 444–462, 2012.

[99] S. N. Gangappa and J. F. Botto, "The BBX family of plant transcription factors," *Trends in Plant Science*, vol. 19, no. 7, pp. 460–470, 2014.

[100] A. G. Greenup, S. Sasani, S. N. Oliver, S. A. Walford, A. A. Millar, and B. Trevaskis, "Transcriptome analysis of the vernalization response in barley (Hordeum vulgare) seedlings," *PLoS ONE*, vol. 6, no. 3, Article ID e7900, 2011.

[101] L. J. Leach, E. J. Belfield, C. Jiang, C. Brown, A. Mithani, and N. P. Harberd, "Patterns of homoeologous gene expression shown by RNA sequencing in hexaploid bread wheat," *BMC Genomics*, vol. 15, no. 1, article 276, 2014.

[102] D. B. Fowler, G. Breton, A. E. Limin, S. Mahfoozi, and F. Sarhan, "Photoperiod and temperature interactions regulate low-temperature-induced gene expression in barley," *Plant Physiology*, vol. 127, no. 4, pp. 1676–1681, 2001.
[108] M. F. Covington, J. N. Maloof, M. Straume, S. A. Kay, and S. L. Harmer, “Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development,” *Genome Biology*, vol. 9, no. 8, article R130, 2008.

[109] S. Sasani, M. N. Hemming, S. N. Oliver et al., “The influence of vernalization and daylength on expression of flowering-time genes in the shoot apex and leaves of barley (*Hordeum vulgare)*,” *Journal of Experimental Botany*, vol. 60, no. 7, pp. 2169–2178, 2009.

[110] A. F. Stelmakh, “Genetic systems regulating flowering response in wheat (reprinted from wheat: prospects for global improvement, 1998),” *Euphytica*, vol. 100, no. 1–3, pp. 359–369, 1998.

[111] M. Iqbal, A. Navabi, R.-C. Yang, D. F. Salmon, and D. Spaner, “Molecular characterization of vernalization response genes in Canadian spring wheat,” *Genome*, vol. 50, no. 5, pp. 511–516, 2007.