Plant Biotechnology Journal

Plant Biotechnology Journal (2012) **10**, pp. 67–82 doi: 10.1111/j.1467-7652.2011.00645.x

Molecular, phylogenetic and comparative genomic analysis of the *cytokinin oxidase/dehydrogenase* gene family in the Poaceae

Sabine Mameaux[1,†], James Cockram[1,*,†], Thomas Thiel[2], Burkhard Steuernagel[2], Nils Stein[2], Stefan Taudien[3], Peter Jack[4], Peter Werner[5], John C. Gray[6], Andy J. Greenland[1] and Wayne Powell[1,‡]

[1]*National Institute of Agricultural Botany (NIAB), Cambridge, UK*
[2]*Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany*
[3]*Leibniz Institute of Age Research—Fritz-Lipmann-Institute (FLI), Jena, Germany*
[4]*RAGT Seeds Ltd., Ickelton, UK*
[5]*KWS UK Ltd., Thriplow, Royston, UK*
[6]*Department of Plant Sciences, University of Cambridge, Cambridge, UK*

Received 7 April 2011;
accepted 20 June 2011.
**Correspondence*
(Tel +44 0 1223 342473;
fax +44 0 1223 277602;
email james.cockram@niab.com)
†These authors contributed equally to the article.
‡Present address: Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University, Aberystwyth, SY23 3DA, UK.

Keywords: wheat, barley, molecular phylogeny, comparative mapping, cytokinin oxidase, cytokinin oxidase/dehydrogenases 2.

Summary

The genomes of cereals such as wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*) are large and therefore problematic for the map-based cloning of agronomicaly important traits. However, comparative approaches within the Poaceae permit transfer of molecular knowledge between species, despite their divergence from a common ancestor sixty million years ago. The finding that null variants of the rice gene *cytokinin oxidase/dehydrogenase 2* (*OsCKX2*) result in large yield increases provides an opportunity to explore whether similar gains could be achieved in other Poaceae members. Here, phylogenetic, molecular and comparative analyses of *CKX* families in the sequenced grass species rice, brachypodium, sorghum, maize and foxtail millet, as well as members identified from the transcriptomes/genomes of wheat and barley, are presented. Phylogenetic analyses define four Poaceae *CKX* clades. Comparative analyses showed that *CKX* phylogenetic groupings can largely be explained by a combination of local gene duplication, and the whole-genome duplication event that predates their speciation. Full-length *OsCKX2* homologues in barley (*HvCKX2.1*, *HvCKX2.2*) and wheat (*TaCKX2.3*, *TaCKX2.4*, *TaCKX2.5*) are characterized, with comparative analysis at the DNA, protein and genetic/physical map levels suggesting that true *CKX2* orthologs have been identified. Furthermore, our analysis shows *CKX2* genes in barley and wheat have undergone a Triticeae-specific gene-duplication event. Finally, by identifying ten of the eleven *CKX* genes predicted to be present in barley by comparative analyses, we show that next-generation sequencing approaches can efficiently determine the gene space of large-genome crops. Together, this work provides the foundation for future functional investigation of *CKX* family members within the Poaceae.

Introduction

Cytokinins (CKs) are a group of plant hormones first identified as promoting cell proliferation and shoot formation. Subsequently, a long list of developmental and physiological functions under the control of CKs has been established, including cell differentiation, leaf expansion, apical dominance, reproductive development, seed germination, delay of senescence, de-etiolation, chloroplast differentiation, plant–pathogen interactions, as well as flower and fruit development (Sakakibara, 2006). A group of cytokinin-degrading enzymes, termed cytokinin oxidase/dehydrogenases (CKXs), are thought to control levels of CK, inactivating the phyto-hormone in a single enzymatic step by cleavage of the N^6-substituted isoprene-side chain to produce adenine and the corresponding aldehyde (Brownlee *et al.*, 1975). CKXs are characterized by the presence of two conserved regions within their proteins: the flavin adenine dinucleotide (FAD)-binding and cytokinin-binding domains, both of which are essential for fully functional enzymatic activity (Malito *et al.*, 2004). Investigations in various plant species suggest that *CKXs* respond differently to various stresses and that their expression may be tissue specific (Vaseva-Gemisheva *et al.*, 2004). Recently, positional cloning of the rice yield quantitative trait locus (QTL) *Gn1a* has shown it to be controlled by *OsCKX2*, encoding a CKX (Ashikari *et al.*, 2005). Natural *OsCKX2* genetic variants were shown to confer increased grain number per panicle, leading to a 21% increase in grain number between the parental line and a near isogenic line harbouring the contrasting *OsCKX2* allele. The positive effects of *OsCKX2* on grains per panicle are due either to an exonic deletion resulting in truncation of the OsCKX2 protein or to reduced *OsCKX2* expression.

Rice, a major provider of global human nutrition, is a member of the grass family (Poaceae). The Poaceae is divided into around twelve subfamilies (with species from the Ehrhartoideae, Panicoideae and Pooideae subfamilies investigated here) and includes other major global crop species such as bread wheat (*Triticum aestivum*), barley (*H. vulgare*) and maize (*Zea mays*). The Poaceae

also includes many minor or regionally important crops, such as sorghum (*Sorghum bicolour*) and foxtail millet (*Setaria italica*), as well as the model Pooideae species, brachypodium (*Brachypodium distachyon*). Since their divergence from a common ancestor ~60 million years ago (mya) (Salse *et al.*, 2008), members of the Poaceae have undergone major changes in the size and structural arrangement of their genomes. For example, while the genomes of brachypodium (355 Mbp), rice (430 Mbp), foxtail millet (515 Mbp) and sorghum (750 Mbp) have remained relatively small, those of barley (5500 Mbp) and bread wheat (16 000 Mbp) are an order of magnitude larger. One of the factors driving increased genome size is polyploidization: while barley ($2n = 2x = 14$), sorghum ($2n = 2x = 20$), brachypodium ($2n = 2x = 10$), foxtail millet ($2n = 2x = 18$) and rice ($2n = 2x = 24$) are diploid, bread wheat is hexaploid ($2n = 6x = 42$), with the three subgenomes originating from diploid A, B and D progenitor genomes. Although maize (2500 Mbp) behaves genetically as a diploid with ten pairs of chromosomes ($2n = 2x = 20$), it actually represents an allotetraploid, resulting from the hybridization of two slightly diverged progenitors ~5 mya (Wei *et al.*, 2007). Another major factor affecting genome size is the rate of accumulation of repetitive sequence, because of differences in transposable element (TE) activity and redaction/elimination (Bennetzen *et al.*, 2005). The high TE frequency in wheat, barley and maize has been a major factor in the inflation of their genome sizes. Despite these substantial structural changes that have taken place during the evolutionary radiation of the Poaceae, comparative analyses have shown high degrees of colinearity between grass genomes (Gale and Devos, 1998). The ability to 'cross-reference' Poaceae genomes provides a framework within which advances in genetic understanding in one species can potentially be transferred to related grass species. This is of particular relevance to barley and wheat, whose large genomes present a major obstacle for the creation of whole-genome sequence assemblies. The ability to align multiple cereal genomes has led to the comparative mapping (and in many cases, map-based cloning) of series of orthologous loci controlling related traits within the Poaceae. Examples include the colinear *VRN-1* and *VRN-2* flowering time loci in diploid wheat (*Triticum durum*), tetraploid wheat (*Triticum durum*), and hexaploid wheat (*T. aestivum*), barley (*H. vulgare*) and rye (*Secale cereale*) (reviewed by Cockram *et al.*, 2007), the photoperiod response *PPD-1* loci in barley (Turner *et al.*, 2005) and bread wheat (Beales *et al.*, 2007) and genes controlling anthocyanin in maize and barley (Goff *et al.*, 1992; Cockram *et al.*, 2010a). More recently, the availability of high-density genetic maps and genome sequence assemblies has led to the identification of whole-genome duplication (WGD) events in rice (Salse *et al.*, 2008), brachypodium (International Brachypodium Initiative, 2010), sorghum (Paterson *et al.*, 2009), maize (Wei *et al.*, 2007), bread wheat (Salse *et al.*, 2008) and barley (Thiel *et al.*, 2009), estimated to have occurred ~20 million years prior to their divergence from a common ancestor (Paterson *et al.*, 2004). After the ancestral WGD, structural rearrangements are thought to have led to diploidization and stabilization of the genome (Paterson *et al.*, 2004). Taken together, inter- and intra-species colinearity provides the evolutionary context with which to investigate the genetic control of related traits within the Poaceae, allowing efficient exploitation of natural and artificial variants within orthologous genes from related grasses.

Several studies have investigated Poaceae *CKX* family members to varying extents in one or more grass species: Ashikari *et al.*, 2005 (rice *CKX* family); Bilyeu *et al.*, 2001 (*ZmCKX1*); Galuszka *et al.*, 2004 [barley and wheat *CKX* expressed sequence tags (ESTs)]; Houba-Hérin *et al.*, 1999 (*ZmCKX1*); Schmülling *et al.*, 2003 (rice *CKX* family, *ZmCKX1*, *HvCKX9*); Vyroubalová *et al.*, 2009 (Maize *CKX* genes, 2008 assembly); Zhang *et al.*, 2010 (rice *CKX* family, *ZmCKX1-3*, *TaCKX2.1*, *TaCKX2.2*). However, although the cloning of *OsCKX2* has highlighted the potential importance of this gene family in plants, a systematic analysis of the *CKX* family in grasses has been lacking to date. As a basis from which informed investigation of cereal *CKX* members can be undertaken, we describe the identification, molecular characterization, comparative mapping and phylogenetic analysis of *CKX* gene families in the Poaceae, with particular emphasis on wheat and barley *OsCKX2* orthologs.

Results

Identification of *CKX* genes in sequenced cereal genomes

Previous studies identified 11 *CKX* members in rice (Ashikari *et al.*, 2005). To verify copy number in the current rice genome assembly, *CKX* coding regions (CDS) were used for BLASTn searches of Michigan State University (MSU) Rice Genome Annotation Release v6.1 (http://rice.plantbiology.msu.edu/). As expected, 11 gene models were identified (Table 1). Of these, *OsCKX6* and *OsCKX7* represent tandemly duplicated genes on chromosome Os2, while *OsCKX1* and *OsCKX2* are located within ~570 kb of each other on Os1. The primary predicted splice forms were subsequently used for BLASTn analysis of related grass species.

Searches of the brachypodium genome (evolutionarily the most closely related to rice of the sequenced grasses analysed here) identified 11 *CKX* genes, with predicted CDS of 1149–1695 bp (Table 1). In addition, two putative *CKX* pseudogenes were identified, as assumed by severe truncations in their predicted proteins: the first, *BdCKX8b* (Bradi5g16090.1, CDS of 438 bp), is a tandem duplication of *BdCKX8a* (Bradi5g16080.1, CDS of 1149 bp). The second shows sequence similarity to *BdCKX10* and is located on chromosome Bd2 (Bradi2g05210.1, CDS of 399 bp). BLASTn analysis of the sorghum genome also identified 11 *CKX* genes, plus one *CKX* pseudogene (Sb04g007745.1, CDS of 150 bp) (Table 1). The latter is one of a cluster of three tandemly duplicated genes, colinear with clusters of two *CKX* genes in rice and brachypodium (*CKX6*, *CKX7*). Searches of the foxtail millet genome identified 12 *CKX* family members (Table 1). Further analysis of two adjacent gene models with high sequence similarity to *OsCKX2* (SiPROV034880m and SiPROV019828m) showed they are homologous to the 5′ and 3′ ends of the rice gene. Extraction of the genomic region spanning these two genes, and re-analysis using the gene prediction software FGENESH, resulted in a single-gene model prediction (Figure S1), suggesting an error in the foxtail millet gene annotation. The updated gene model was used in all subsequent analysis, resulting in a total of 11 *SiCKX* family members. Searches of the current maize genome assembly (v. 5b.60) returned 12 *CKX* family members, as well as three pseudogenes (*CKX4*, *CKX5* and *CKX8*) (Table 1). Two maize homologues were identified for *CKX4* and *CKX6*. Single copies were found for *CKX1*, *CKX2*, *CKX3*, *CKX7*, *CKX8* and *CKX11*.

Although no genome assembly is currently available for barley, BLASTn searches of genomic sequence data from our shotgun next-generation sequencing (NGS) of barley cv. Morex (Nils Stein, unpublished data) identified 14 contigs ranging in

Table 1 Rice *CKX* gene family members, and their homologues identified in the sequenced genomes of brachypodium, sorghum, foxtail millet and maize. Full-length barley *CKX* genes identified in this study are also shown

Gene	Chr (Mbp)	Stra-nd	Gene model	Genomic (bp)	cDNA (bp)	Exo-ns	Protein (aa)	*e*-value (% identity)*
Oryza sativa								
OsCKX1	1 (4.70)	–	Os01g09260.1	2479	1599	3	532	N/A
OsCKX2	1 (5.30)	–	Os01g10110.1	5576	1698	4	565	N/A
OsCKX3	10 (18.20)	+	Os10g34230.1	4809	1584	5	527	N/A
OsCKX4	1 (41.30)	+	Os01g71310.1	2842	1590	5	529	N/A
OsCKX5	1 (32.79)	–	Os01g56810.1	4265	1605	5	534	N/A
OsCKX6	2 (6.68)	+	Os02g12770.1	2283	1506	3	501	N/A
OsCKX7	2 (6.69)	+	Os02g12780.1	1829	1575	3	524	N/A
OsCKX8	4 (26.01)	+	Os04g44230.1	4505	1599	5	532	N/A
OsCKX9	5 (17.97)	+	Os05g31040.1	2149	1566	5	521	N/A
OsCKX10	6 (22.20)	–	Os06g37500.1	2931	1587	4	528	N/A
OsCKX11	8 (22.62)	–	Os08g35860.1	2949	1863	4	620	N/A
Brachypodium distachyon								
BdCKX1	2 (4.06)	–	Bradi2g05580.1	2835	1566	3	521	0.0 (73)
BdCKX2	2 (4.53)	–	Bradi2g06030.1	3435	1695	3	564	0.0 (66)
BdCKX3	3 (31.06)	–	Bradi3g29130.1	3528	1566	5	521	0.0 (90)
BdCKX4	2 (57.80)	–	Bradi2g60456.1	1933	1572	5	523	0.0 (88)
BdCKX5	2 (51.12)	–	Bradi2g51530.1	3838	1650	5	549	0.0 (84)
BdCKX6	3 (6.40)	+	Bradi3g08220.1	1359	1359	1	452	4e–179 (71)
BdCKX7	3 (6.41)	+	Bradi3g08230.1	1599	1599	1	532	0.0 (75)
BdCKX8a	5 (19.49)	+	Bradi5g16080.1	1781	1149	4	382	7e–122 (66)
BdCKX8b†	5 (19.49)	+	Bradi5g16090.1	1758	438	2	145	2e–66 (74)
BdCKX9	2 (25.97)	–	Bradi2g27170.1	2055	1566	5	521	0.0 (87)
BdCKX10a	1 (33.31)	+	Bradi1g37470.1	1903	1617	3	538	0.0 (73)
BdCKX10b†	2 (3.78)	+	Bradi2g05210.1	399	399	1	132	6e–42 (73)
BdCKX11	3 (40.25)	–	Bradi3g37690.1	2580	1542	4	513	0.0 (80)
Sorghum bicolor								
SbCKX1	3 (3.35)	+	Sb03g003280.1	2465	1590	3	529	0.0 (81)
SbCKX2	3 (2.72)	+	Sb03g002810.1	4634	1734	2	577	0.0 (77)
SbCKX3	1 (19.99)	–	Sb01g019000.1	4462	1587	6	528	0.0 (86)
SbCKX4	3 (72.68)	+	Sb03g045410.1	2160	1572	5	523	0.0 (85)
SbCKX5	3 (64.20)	–	Sb03g036160.1	3538	1647	5	548	0.0 (85)
SbCKX6	4 (8.09)	–	Sb04g007730.1	1863	1614	3	537	0.0 (77)
SbCKX7a	4 (8.12)	+	Sb04g007740.1	1846	1605	5	534	0.0 (78)
SbCKX7b†	4 (8.13)	+	Sb04g007745.1	150	150	1	49	4e–126 (78)
SbCKX8	6 (52.15)	+	Sb06g022930.1	4869	1593	5	530	0.0 (81)
SbCKX9	9 (46.59)	+	Sb09g018640.1	1954	1569	5	522	0.0 (82)
SbCKX10	10 (50.61)	–	Sb10g022590.1	1817	1650	2	549	0.0 (77)
SbCKX11	7 (57.12)	–	Sb07g022530.1	2965	1563	4	520	0.0 (83)
Setaria italica								
SiCKX1	3 (2.46)	+	SiPROV006865m	2069	1569	3	522	0.0 (81)
SiCKX2	3 (1.73)	+	FGENESH model‡	4240	1665	3	554	0.0 (79)
SiCKX3	1 (15.84)	–	SiPROV006823m	3649	1572	5	523	0.0 (87)
SiCKX4	3 (44.77)	+	SiPROV006863m	2378	1569	5	522	0.0 (86)
SiCKX5	3 (37.14)	–	SiPROV006381m	4286	1617	5	538	0.0 (87)
SiCKX6	4 (3.24)	+	SiPROV033306m	1836	1614	3	537	0.0 (78)
SiCKX7	4 (3.25)	+	SiPROV027697mg	1772	1464	4	488	0.0 (82)
SiCKX8	6 (25.33)	+	SiPROV0027343m	1707	1245	4	415	0.0 (82)
SiCKX9	9 (14.87)	+	SiPROV006866m	2087	1569	5	522	0.0 (83)
SiCKX10	10 (34.36)	–	SiPROV028586m	1867	1620	3	539	0.0 (79)
SiCKX11	7 (29.83)	–	SiPROV006958m	3143	1557	3	518	0.0 (83)
Zea mays								
ZmCKX1	3 (7.59)	–	GRMZM2G146644	>4377	1632	7	543	4e–158 (81)

Table 1 Continued

Gene	Chr (Mbp)	Stra-nd	Gene model	Genomic (bp)	cDNA (bp)	Exo-ns	Protein (aa)	*e*-value (% identity)*
ZmCKX2	8 (22.25)	–	GRMZM2G325612	3620	1755	3	584	0.0 (75)
ZmCKX3	1 (230.01)	–	GRMZM2G404443	3223	1629	3	542	0.0 (79)
ZmCKX4a	3 (152.20)	–	GRMZM2G050997	1929	1560	5	519	0.0 (84)
ZmCKX4b	8 (159.19)	–	GRMZM2G167220	1982	1578	5	525	0.0 (77)
ZmCKX4c†	4 (226.39)	+	GRMZM2G170446	4054	1056	8	351	2e–154 (86)
ZmCKX5a	8 (172.36)	+	GRMZM2G024476	3519	1605	5	534	0.0 (85)
ZmCKX5b†	3 (191.84)	–	GRMZM5G817173	2076	942	5	313	0.0 (88)
ZmCKX6a	4 (230.02)	–	GRMZM2G114427	1825	1620	3	539	0.0 (76)
ZmCKX6b	4 (230.11)	–	GRMZM2G134634	1825	1620	3	539	0.0 (76)
ZmCKX7	4 (230.19)	–	GRMZM2G162048	1803	1605	3	534	0.0 (77)
ZmCKX8a†	10 (131.7)	+	GRMZM2G122340	7976	804	3	267	0.0 (81)
ZmCKX8b	2 (27.55)	–	GRMZM2G008792	4540	1587	5	528	0.0 (82)
ZmCKX10	9 (80.49)	–	GRMZM2G303707	2747	2066	3	568	0.0 (77)
ZmCKX11	1 (207.71)	+	GRMZM2G348452	2775	1578	4	525	0.0 (84)
Hordeum vulgare								
HvCKX2.1§	3H	N/A	*HvCKX2.1*	4087	1704	3	553	0.0 (77)
HvCKX2.2	3H	N/A	*HvCKX2.2*	3145	1695	3	564	0.0 (77)
HvCKX7	N/A	N/A	*HvCKX7*	1602	1602	1	533	0.0 (77)
HvCKX9	N/A	N/A	*HvCKX9*	1955	1581	5	526	0.0 (84)

*BLASTn comparison between the gene and the closest rice homologue.

†Truncated genes.

‡Re-analysis of gene models SiPROV034880m and SiPROV019828m.

§Oregon Wolfe Barley Recessive allele, FGENESH gene model prediction, utilizing alternative intron 1 splice site.

N/A, not applicable; CKX, cytokinin oxidase/dehydrogenases.

size from 1319 to 8641 bp, representing putative homologues of ten (*OsCKX1-5*, *7*-11) of the eleven rice *CKX* family members (Table S1A). For the majority of *CKX* genes, single contigs with partial gene coverage were identified. However, *OsCKX1* identified two barley contigs, homologous to the 5′ (contig_1005647) and 3′ (contig_1010998) ends of *OsCKX1* (3e-159 and 7e–159, respectively). Similarly, four barley contigs with high sequence similarity (≤7e–129) to *OsCKX2* were identified. A partial imperfect overlap between contig_47875 and contig_104715 suggests these represent the 5′ and 3′ ends of the same gene, while the remaining discrete contigs (contig_1005571 and contig_2164028) were hypothesized to represent 5′ and 3′ genomic fragments of a second *CKX2*-like gene in barley. Of the 14 contigs identified, two were predicted to contain a full-length *CKX* genes: contigs 1021400 and 1008654 (displaying high sequence similarity to *OsCKX7* and *OsCKX9*, respectively) (Table 1). Finally, although 5× sequence coverage of the hexaploid wheat genome (cv. 'Chinese Spring') has very recently been made publicly available (http://www.cerealsdb.uk.net/), our analyses find that sequence depth is currently insufficient to provide detailed bioinformatic analysis of the *CKX2* genes (data not shown). However, we utilize alternative approaches [EST databases and bacterial artificial chromosome (BAC) library screening] to identify *OsCKX2* orthologues in wheat.

Identification of *CKX* ESTs in the transcriptomes of wheat and barley

Extensive publicly available transcriptome sequence databases are available for wheat and barley, containing ESTs from many different tissues, developmental stages and growth regimes. The grouping of ESTs into transcript assemblies (TAs) provides an invaluable resource for the identification of genes and gene families in large-genome species. Here, the eleven rice *OsCKX* CDS were used to identify homologues in wheat and barley databases using BLASTn analyses. Homology was verified by back-BLASTn comparison to the rice genome. Putative wheat homologues of *OsCKX1*, *OsCKX2*, *OsCKX3*, *OsCKX4*, *OsCKX5* and *OsCKX11* were identified, while searches of barley TAs identified homologues for *OsCKX1*, *OsCKX4*, *OsCKX5*, *OsCKX9* and *OsCKX11* (Table S2). The ESTs were derived from cDNA libraries constructed from a range of tissues, including callus, roots, leaves and kernels, with 2–8 ESTs identified within each homologous *CKX* cluster. Although *CKX2* homologues were not identified in barley, the four wheat *CKX2* ESTs all originated from developing kernel tissue (GenBank accessions BQ235927, BQ238832, BQ903062 and CA705202). In addition, two full-length *TaCKX2*-like cDNA sequences were identified: *TaCKX2.1* (FJ648070) and *TaCKX2.2* (GU084177), both of which have recently been cloned by 5′ and 3′ rapid amplification of cDNA ends (RACE) using primers designed from EST CA705202 and coarsely mapped to a deletion bin on chromosome 3D (Zhang *et al.*, 2010).

Molecular characterization of *CKX2* homologues in wheat and barley

Among the four wheat ESTs homologous to *OsCKX2*, two ESTs (BQ238832 and BQ235928) originated from the same cDNA (TaE05040F02), representing forward and reverse sequences spanning the *OsCKX2* start and stop codons. To determine the relationship of this wheat cDNA with the two *TaCKX2*-like genes described by Zhang *et al.* (2010), the cDNA was sequenced by primer walking. We found the cDNA to be

1865 bp (GenBank accession JF293079), with CDS of 1638 bp, displaying high sequence similarity to *OsCKX2* (*e* = 0.0, identities = 78%). Alignment of the CDS with *TaCKX2.1* and *TaCKX2.2* (Zhang *et al.*, 2010) showed sequence identity of 84% and 85% identity, respectively, but clearly indicated its origin from a different *CKX2* gene (Figure S2). Accordingly, the *CKX2* gene represented by cDNA TaE05040F02 has been designated *TaCKX2.3*. To obtain the complete genomic *TaCKX2.3* sequence, as well as the surrounding genomic context for comparative analyses, primers were designed for PCR screening of a subset of a wheat BAC library, identifying four positive BAC clones. *Hin*dIII digests of BAC DNAs found two clones to overlap with each other (1991B18 and 2018M8), while the remaining two clones (470M29 abd 400N24) represented nonoverlapping genomic regions (data not shown). Accordingly, we fully sequenced BAC clones 1991B18, 470M29 and 400N24, finding them contain inserts of 157, 128 and 164 kb, with a mean GC-content of 44%, 47% and 69%, respectively. For clone 400N24 (GenBank accession JF292901), four predicted gene models [400.1 (*TaCKX2.3*), 400.2, 400.3, 400.4] were identified, all of which were found within a gene island of ~30 kb (Table 2). Alignment of cDNA and genomic sequences found *TaCKX2.3* CDS to be arranged over three exons of 646, 422 and 588 bp, with no sequence polymorphisms identified between the coding regions of *TaCKX2.3* CDS and gene 400.1. The positions of introns 1 and 2 are perfectly conserved between *OsCKX2* and *TaCKX2.3*, whereas the third intron found in *OsCKX2* is absent from *TaCKX2.3* (Figure 1). BAC clones 1991B18 and 470M29 each contained a single predicted gene model, annotated *TaCKX2.4* (3 exons, CDS = **1659** bp) and *TaCKX2.5* (3 exons, CDS = **1638** bp), respectively.

As no *OsCKX2* homologues were identified in barley EST databases, the partial *CKX2*-like NGS barley genomic contigs were used to identify full-length *HvCKX2* gene sequences from cvs. Oregon Wolfe Barley Recessive (OWB-R) and Oregon Wolfe Barley Dominant (OWB-D) by PCR amplification and primer walking. This resolved the four NGS contigs into two genes, designated *HvCKX2.1* and *HvCKX2.2*, with predicted CDS of 1704 and 1695 bp, respectively (Table S1B). Both barley *CKX2* genes were predicted to contain three exons located at conserved positions relative to *OsCKX2*, and like all other Poaceae *CKX2* genes investigated here, lack the third intron found in *OsCKX2* (Figure 1).

HvCKX2.1 (cv. Morex), *HvCKX2.2* (cv. OWB-D), *TaCKX2.3*, *TaCKX2.4* and *TaCKX2.5* (cv. Chinese Spring) encode predicted proteins of 567, 556, 551 552 and 545 amino acid residues, respectively (Figure S3). Protein domain analyses predicted all five proteins to possess FAD-binding (≤6.2e-20) and cytokinin-binding (≤1e-108) domains, which are common to all known CKX proteins. Alignment of wheat/barley CKX2 proteins to OsCKX2 showed that while the positions of the predicted stop codons are conserved across the species, the N-terminus shows variation in length (Figure S3). Of the seven known wheat and barley CKX2 proteins, TaCKX2.3 is the most similar to OsCKX2 (identities = 69%), followed by TaCKX2.1/TaCKX2.4/TaCKX2.5 (68%) and HvCKX2.1/HvCKX2.2 /TaCKX2.2 (67%).

Genetic mapping of *CKX2* homologues in barley and wheat

Rice chromosome Os1 is colinear with the *Triticeae* group 3 chromosomes, with the short arms colinear with each other between both clades (Devos, 2005; Salse *et al.*, 2008; Thiel *et al.*, 2009). To support the putative orthology of Triticeae *CKX2*-like genes with *OsCKX2*, we genetically mapped the identified *OsCKX2* orthologues in wheat (*TaCKX2.3*, *TaCKX2.4*, *TaCKX2.5*), and barley (*HvCKX2.1* and *HvCKX2.2*). *TaCKX2.4*, *TaCKX2.5* and *TaCKX2.3* were mapped to chromosomes 3A, 3B and 3D, respectively, using gene-specific primers to amplify products from the nullisomic/tetrasomic (NT) aneuploid series developed from cv. Chinese Spring (Sears, 1966) (Figure 2a). Sequencing *TaCKX2.3* PCR products amplified from the wheat varieties 'Avalon' and 'Cadenza' identified a synonymous single nucleotide polymorphism (SNP) within exon 1 (G + 1084/T). This allowed intra-chromosomal mapping of *TaCKX2.3* to chromosome 3D, 2 cM distal to SSR markers GWM456/GWM52, and 28 cM proximal to amplified fragment length polymorphism (AFLP) marker stm02TCAC (Figure 2b). A nonsynonymous SNP in *TaCKX2.4* (C + 893/T, exon 2, Ala → Val) was identified in the 'Opata 85' × 'synthetic' parental lines, placing it 4 cM distal to SSR marker gwm32. Finally, a synonymous SNP within *TaCKX2.5* (G + 234/A) was identified between the parental lines of the 'Helidur' × 'TD161' population, mapping this gene to chromosome 3B between markers gwm533 and gwm285. Alignment of all three mapped *TaCKX2* genes to the wheat consensus genetic map (Somers *et al.*, 2004) shows they are located on the short arms of the group 3 chromosomes, close to the nondistal low-recombining regions within which the centromeres are predicted to be located.

In Barley, sequencing *HvCKX2.1* PCR amplicons from the parental lines of the OWB mapping population (Costa *et al.*, 2001) identified six polymorphisms between OWB-D (JF293078) and OWB-R (JF293074). Using a 32-bp insertion/deletion (InDel)

Table 2 Predicted genes on sequenced wheat *TaCKX2* bacterial artificial chromosome clones 400N24 (JF292901), 470M29 (JN381556) and 1991B18 (JN381555)

Wheat gene	CDS (bp) [exons]	Corresponding TIGR TA/EST singleton	Orthologous rice gene model	Rice annotation	*Ta* versus *Os* predicted CDS: *e*-value (% identities)
400.1	1656 [3]	JF293079	LOC_Os01g10110.1	OsCKX2	0.0 (78)
400.2	861 [6]	TA111747_4565	LOC_Os01g10100.1	Expressed protein	5e−157 (88)
400.3	1935 [1]	None	LOC_Os01g10090.1	Pentatricopeptide	0.0 (82)
400.4	1335 [7]	TA81187_4565	LOC_Os01g10070.1	Expressed protein	2e−54 (72)
1991.1	1659 [3]	None	LOC_Os01g10110.1	OsCKX2	0.0 (78)
470.1	1638 [3]	None	LOC_Os01g10110.1	OsCKX2	0.0 (78)

TA, transcript assembly; *Ta*, wheat; *Os*, rice; CDS, coding regions.

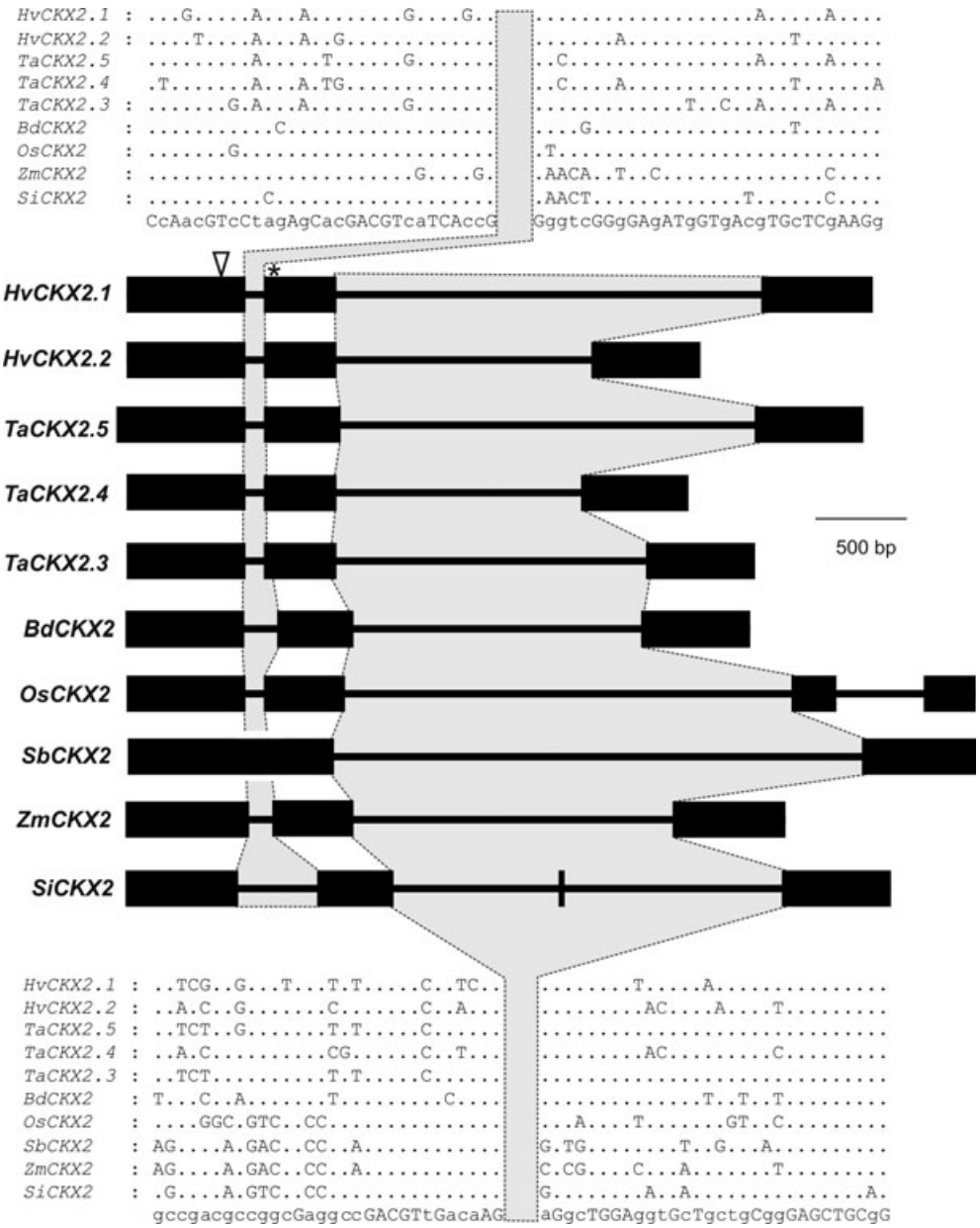

Figure 1 Diagram of predicted *cytokinin oxidase/dehydrogenases 2* intron/exon structure. Exonic sequence conservation surrounding intron/exon boundaries is shown. Species: barley (*Hv*), wheat (*Ta*), brachypodium (*Bd*), rice (*Os*), sorghum (*Sb*), maize (*Zm*) and foxtail millet (*Si*). The position of the 32-bp deletion within *HvCKX2.1* in cv. Oregon Wolfe Barley Recessive is indicated by the triangle, resulting in a premature stop-codon (indicated by a star) relative to cvs. Morex and Oregon Wolfe Barley Dominant.

within exon 1, *HvCKX2.1* was mapped to the short arm of chromosome 3H, cosegregating with SNP markers 1_1102, 3_0583 and 3_0721 at 56.4 cM. Full-length sequencing of *HvCKX2.2* in OWB-D (JF293076) and OWB-R (JF293077) identified ten polymorphisms (3 exonic, 1 nonsynonymous) across ~3.2 kb. Genetic mapping of a selected SNP (T + 1583/C, intron 2) showed *HvCKX2.2* to cosegregate with *HvCKX2.1* on chromosome 3H. Integration of *HvCKX2.1* and *HvCKX2.2* into the barley consensus map (Close *et al.*, 2009) showed these genes to be located at 50.2, ~2 cM proximal to the low-recombining region of chromosome 3H predicted to span the centromere (Figure 2c). Of the 455 gene-based genetic markers distributed across 173.2 cM on the 3H consensus map, 75% identify homologous genes on the colinear rice chromosome Os1 (based on highest-scoring matches from BLASTn analyses). A scatter plot of markers on the barley 3H genetic map (cM) versus the rice Os1 physical map (Mbp) shows that *HvCKX2* orthologs fit in well with the overall conservation of gene order between the two species (Figure 2c). Indeed, two of the genetic markers found to cosegregate with *HvCKX2.1* in the OWB population are orthologous to genes on rice chromosome Os1, 3_0721 (LOC_Os01g10450.1) and 1_1102 (LOC_Os01g11810.1), and are located just 240 and 1120 kb from *OsCKX2* (LOC_Os01g10110.1), respectively. In addition to the genetic mapping of the six *CXK2* orthologues in wheat and barley, we determined the genetic map positions of eight of the remaining

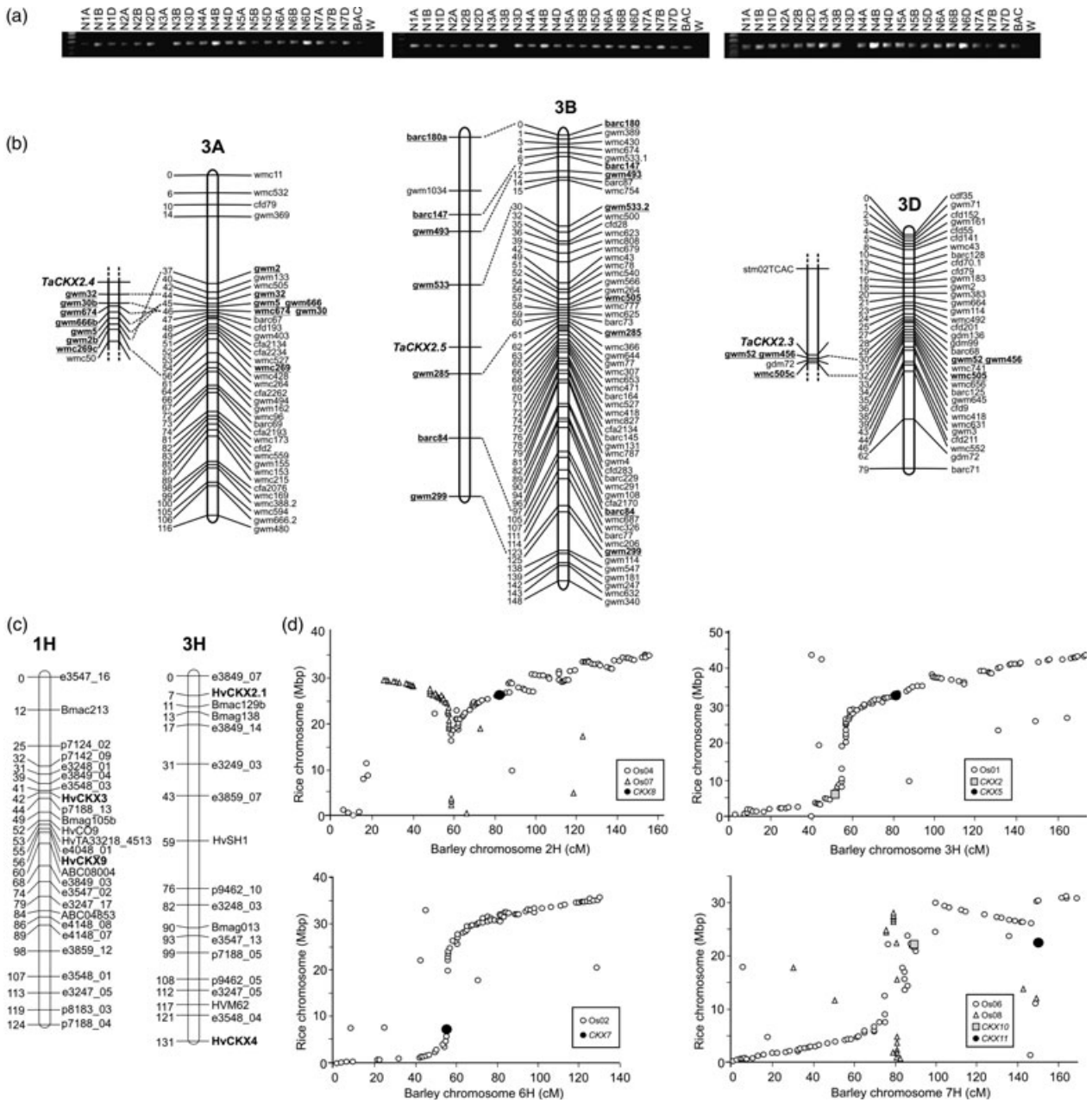

Figure 2 Genetic mapping of *cytokinin oxidase/dehydrogenases* genes in wheat and barley. (a) Inter-chromosomal mapping of wheat genes *TaCKX2.3*, *TaCKX2.4* and *TaCKX2.5* to the group 3 chromosomes (using gene-specific primers to amplify products from the wheat nullisomic/tetrasomic chromosome addition lines). Deleted chromosomes indicated. W = water control. (b) Intra-chromosomal mapping of *TaCKX2.3*, *TaCKX2.4* and *TaCKX2.5* (using the 'Avalon' × 'Cadenza', 'Opata 85' × 'synthetic' and 'Helidur' × 'TD161' populations, respectively), aligned to the consensus genetic map of Somers *et al.*, 2004;. (c) Genetic mapping of barley *CKX* genes in the 'Arta' × '*H. spontaneum*' population. (d) Genetic mapping of barley *CKX* genes in the 'Oregon Wolfe Barley Dominant' × 'Oregon Wolfe Barley Recessive' populations, allowing integration into the consensus map of Close *et al.*, 2009. Colinearity between rice physical maps and homologous barley EST-based markers positioned in the barley consensus genetic map is indicated. *HvCKX2.1* and *HvCKX2.2* cosegregate with each other on chromosome 3H. No polymorphism was identified for *HvCKX1*.

nine barley *CKX* family members identified in our analysis of the draft genome assembly (Figure 2, mapped polymorphisms listed in Table S3). Using the 'Arta' × '*H. spontaneum*' (A × S) population (Baum *et al.*, 2003), *HvCKX3* (42 cM) and *HvCKX9* (56 cM) mapped to the short and long arms of chromosome 1H, respectively, while *HvCX8* mapped to 95.2 cM, between marker 11_11388 and 11_10786 on the long arm of chromosome 2H. *HvCKX4* (131 cM in the A×S population) and *HvCKX5* (85.5 cM in the OWB population) both mapped to 3H, while *HvCKX7* mapped to the nondistal region of low genetic recombination predicted to span the 6H centromere (cosegregating with 11_10659 at 64.4 cM in the OWB population). Finally, *HvCKX10* and *HvCKX11* both mapped to the long arm of chromosome 7H in the OWB population, cosegregating with markers

11_20896 (169.3 cM) and 11_10896 (128.3 cM), respectively. No polymorphism was identified for *HvCKX1* in the amplicons and mapping populations investigated.

Comparative genomic analysis of cereal *CKX* gene family

Using the established inter-species grass comparative framework for rice, brachypodium, barley, sorghum, foxtail millet and maize (Devos, 2005; International Brachypodium Initiative, 2010; Paterson *et al.*, 2009; Salse *et al.*, 2008; Schnable *et al.*, 2009; Thiel *et al.*, 2009) (Figure 3), as well as the confirmation of micro-colinearity around each *CKX* gene (Figures 2d and 4, Table S4), cross-species comparison showed that the identified full-length grass *CKX* members map to colinear genomic locations. However, chromosome-specific rearrangements within the maize genome have resulted in deletion/disruption of almost all of the duplicated *CKX* genes arising because of allotetraploidization, with the exceptions of maize orthologues of *CKX4* and *CKX8* (Figure 3, Table 1).

Duplications within the Ehrhartoideae (rice; Paterson *et al.*, 2004; Salse *et al.*, 2008; Yu *et al.*, 2005; Thiel *et al.*, 2009), Panicioideae (sorghum, maize; Paterson *et al.*, 2009; Wei *et al.*, 2007) and Pooideae (barley, brachypodium, wheat; International Brachypodium Initiative, 2010; Salse *et al.*, 2008; Thiel *et al.*, 2009) originated from a WGD in the putative ancestral cereal genome (Bolot *et al.*, 2009; Paterson *et al.*, 2004; Salse *et al.*, 2008). The most recent rice genome assembly was used to update the WGD blocks within the rice genome (Figure 3). Combining inter- and intra-species colinearity duplications indicated the extent of shared WGD in the remaining grass species relative to the genomic locations of *CKX* genes (Figure 3). In addition, WGD-derived blocks of colinearity were verified at the micro-colinearity level by analysis of genes surrounding each rice gene (Table S4, Figure 4). Two *CKX* gene pairs (*CKX4/CKX9* and *CKX7/CKX10*) appear to have evolved as a result of the WGD. In addition, *OsCKX3* and *OsCKX8* are segmentally duplicated within a small region of 26 gene pairs on rice chromosomes Os10 and Os4 (http://rice.plantbiology.msu.edu/). Although *OsCKX1*, *OsCKX2*, *OsCKX5* and *OsCKX11* are located within regions of intra-specific duplication, no paralogous genes were found in the corresponding chromosomal locations in rice (or any of the other sequenced grass genomes), indicating these may have been lost in the ancestral genome prior to the divergence of the modern-day grass species.

With the aim of clarifying the orthology of *TaCKX2.3* with other *CKX2*-like genes, a 154-kb region of rice chromosome Os1 containing ten genes flanking *OsCKX2* on either side was used to investigate the extent of micro-colinearity with the sequenced physical regions of brachypodium, sorghum, foxtail millet, maize and wheat (BAC clone 400N24). Conservation of gene/gene order varied across the species investigated: rice/brachypodium and brachypodium/sorghum shared the highest number of orthologous gene pairs (12), while the lowest conservation was observed between maize chromosomes Zm8 and Zm3, which shared just one gene (Figure 5, Table S5). Gene orientation is conserved in all but one (GRMZM2G014284, maize chromosome Zm3) of the 63 genes conserved between ≥2 species. The four genes on the wheat *TaCKX2.3* BAC are conserved in all species, with the exception of maize Zm3,

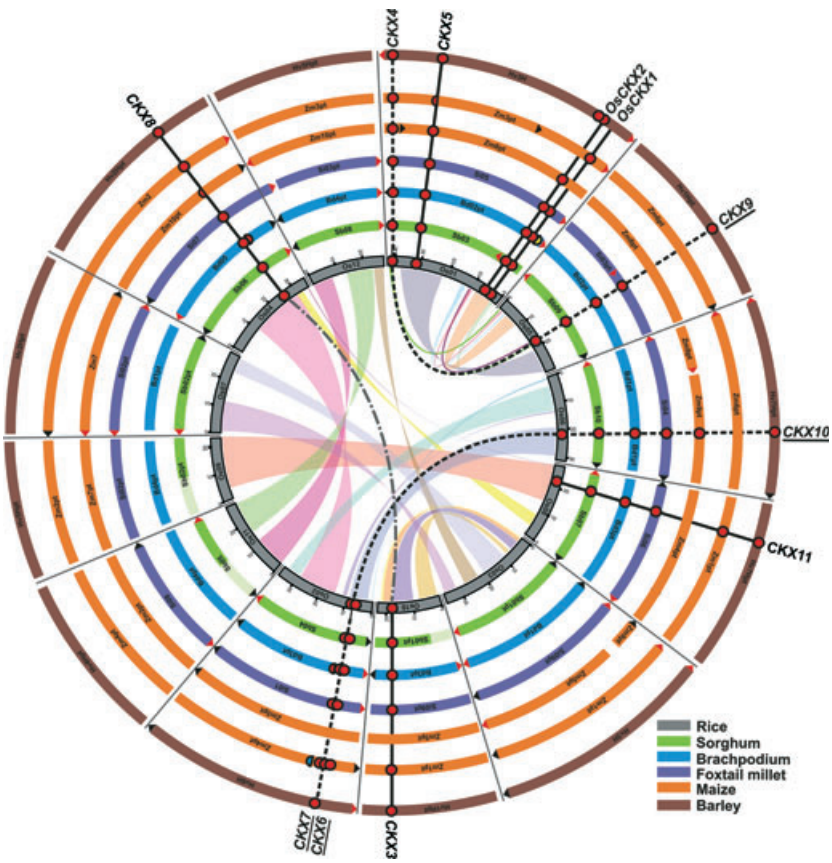

Figure 3 Diagrammatic representation of the position of *cytokinin oxidase/dehydrogenases* (*CKX*) genes within the framework of Poaceae inter- and intra-species colinearity. Duplicated segments of the rice genome originating from the ancestral whole-genome duplication (WGD) event are shown in the centre, surrounded by inter-specific genome colinearity between other grass species (adapted from Devos, 2005; International Brachypodium Initiative, 2010; Schnable *et al.*, 2009). The positions of *CKX* genes are shown (truncated genes indicated by semi-circles), with solid lines linking colinear genes, and dashed lines linking colinear genes involved in the ancestral WGD event. *OsCKX8* and *OsCKX3* lie within a segmentally duplicated region not involved in the WGD, as determined by our methodology (grey dashed line). Telomeres on the short and long arms are indicated by red and black triangles, respectively. Pseudo-genes *ZmCKX4b* and *BdCKX10b* are shown in blue and yellow, respectively.

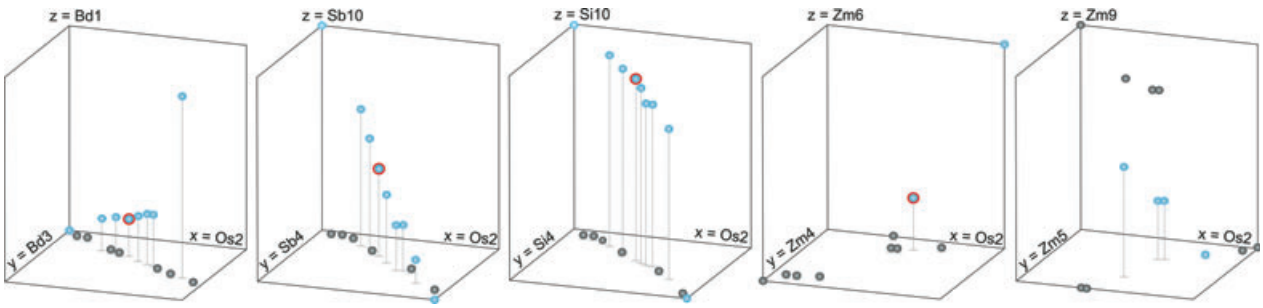

Figure 4 3-D scatter plot illustrating micro-colinearity between physical regions spanning *CKX6/7* and *CKX10* in sequenced grass species, highlighting regions of segmental duplication thought to have arisen because of the whole-genome duplication (WGD) event in the ancestral grass genome. For each scatter plot, the physical positions of rice (Os) genes in the *CKX6/7* region are plotted against the *x*-axis (the 'reference genome'). For each scatter plot, orthologous genes at colinear physical positions at the *CKX6/7* regions in sequenced cereals (the 'target genomes') are plotted against the *y*-axis, resulting in plots along the horizontal *x*-*y*-axis: from left to right: brachypodium (Bd), sorghum (Sb), foxtail millet (Si) and maize (Zm, for which two graphs are shown, because of its allotetraploid genomic structure). The positions of paralogous genes in 'target genomes' around the *CKX10* region (thought to have arisen because of the WGD event in the ancestral cereal genome) are plotted against the *z*-axis. Thus, where a homologous gene is present at all three genomic locations, the plotted point is boosted from the horizontal axis, according to its position on the z-axis. The positions of such genes, shared between all three chromosomal regions, are shown in blue. *CKX* genes are circled in red. Where tandem duplications occur (as is the case for *CKX6/CKX7*), only the gene with highest homology between species is illustrated. Colinearity in maize is generally less well conserved, likely due to extensive genomic modifications post tetraploidization. CKX, cytokinin oxidase/dehydrogenases.

which possessed just one orthologous gene. As expected, the physical sizes of colinear regions varied depending on species: those in rice, brachypodium and sorghum were the smallest, and broadly equal in size (~120 kb), followed by foxtail millet (170 kb) and maize (450 kb, Zm8). Although the increase in physical size in colinear maize chromosomal regions is predominantly because of the increased abundance of TEs, these regions also contain a greater proportion on noncolinear genes (Figure 5). All of the diploid sequenced cereals investigated possessed a single *CKX2* gene at colinear positions.

Phylogenetic analysis of cereal CKX proteins

To help clarify the evolutionary relationships of the *CKX* gene family in grasses, a phylogenetic tree was constructed based on protein sequences from the conserved FAD and cytokinin-binding protein domains (Figure 6). After exclusion of truncated *CKX* family members (Bradi2g05210, Bradi5g16090, Sb04g007745, GRMZM2G817173, GRMZM2G170446, GRMZM2G122340), a total of 66 full-length cereal CKX proteins were analysed, originating from rice (11 members), brachypodium (11), sorghum (11), foxtail millet (11), maize (12), wheat (six) and barley (four). CKX proteins from sequenced grasses grouped into four major phylogenetic clusters, with clades I and II each divided into two subclades. The seven full-length wheat/barley CKX2 proteins are found within clade Ia, with TaCKX2.3 clustering with HvCKX2.1 and HvCKX2.2, while TaCKX2.1 and TaCKX2.2 group separately. The occurrence of two or more CKX2 paralogs is specific to the Triticeae. The overall tree topography of CKX2 proteins agrees with established evolutionary relationships between the grass species investigated, with those from the Poideae (wheat, barley, brachypodium) clustering together, followed by members from the Ehrhartoideae (rice) and Panicoideae (foxtail millet, sorghum and maize). CKX1 proteins from all five sequenced grasses cluster within the second major phylogenetic grouping within clade Ia, supporting their close evolutionary relationship with the CKX2 proteins (as implied by their physical proximity to each other in the five sequenced grass species investigated). Clade Ib includes orthologs of OsCKX6 and OsCKX7, as well as OsCKX10 orthologs. CKX4 and CKX9 proteins group within clade IIa. Clades III and IV contain two (CKX3, CKX8) and one (CKX11) CKX family member, respectively. Phylogenetic analysis of barley and wheat CKX members for which only partial protein sequence data were available showed all members to group within the expected clades (Figure S4).

Discussion

Gene duplication, because of genomic processes such as WGD and polyploidization, provides the raw materials upon which mutation, selection and genetic drift result in the evolution of gene families. Following duplication, functional redundancy means that one or more paralog may become silent/nonfunctional because of the accumulation of deleterious mutations or be lost because of other localized genomic rearrangements. This is especially true of *CKX* family members in the allotetraploid species maize, where in all but two cases, genetic mutation has resulted in the elimination or severe truncation of paralogous *CKX* genes since the allotetraploidization event ~5 mya. This agrees with previous reports documenting frequent loss of one duplicated gene (e.g. Bruggmann *et al.*, 2006). Evidence for the loss of paralogous grass *CKX* genes assumed to have originated during the ancestral grass WGD event is also apparent, with comparative mapping indicating gene loss prior to speciation. Alternatively, paralogous genes may be retained, after which they often diverge in terms of their temporal/spatial expression (subfunctionalization) or function (neo-functionalization) (Wei *et al.*, 2007). Although *OsCKX2* is the only member within the grasses for which natural intra-specific variation has been proven to affect phenotype, simultaneous analysis of genomic, phylogenetic, molecular and transgenic data from the Poaceae (and arabidopsis) can provide insights into the evolution and possible function of grass *CKX2* orthologs and other Poaceae CKX family members.

Clade Ia

Phylogenetic and sequence analyses show *CKX1* and *CKX2* to be closely related, grouping within clade Ia. Although they are

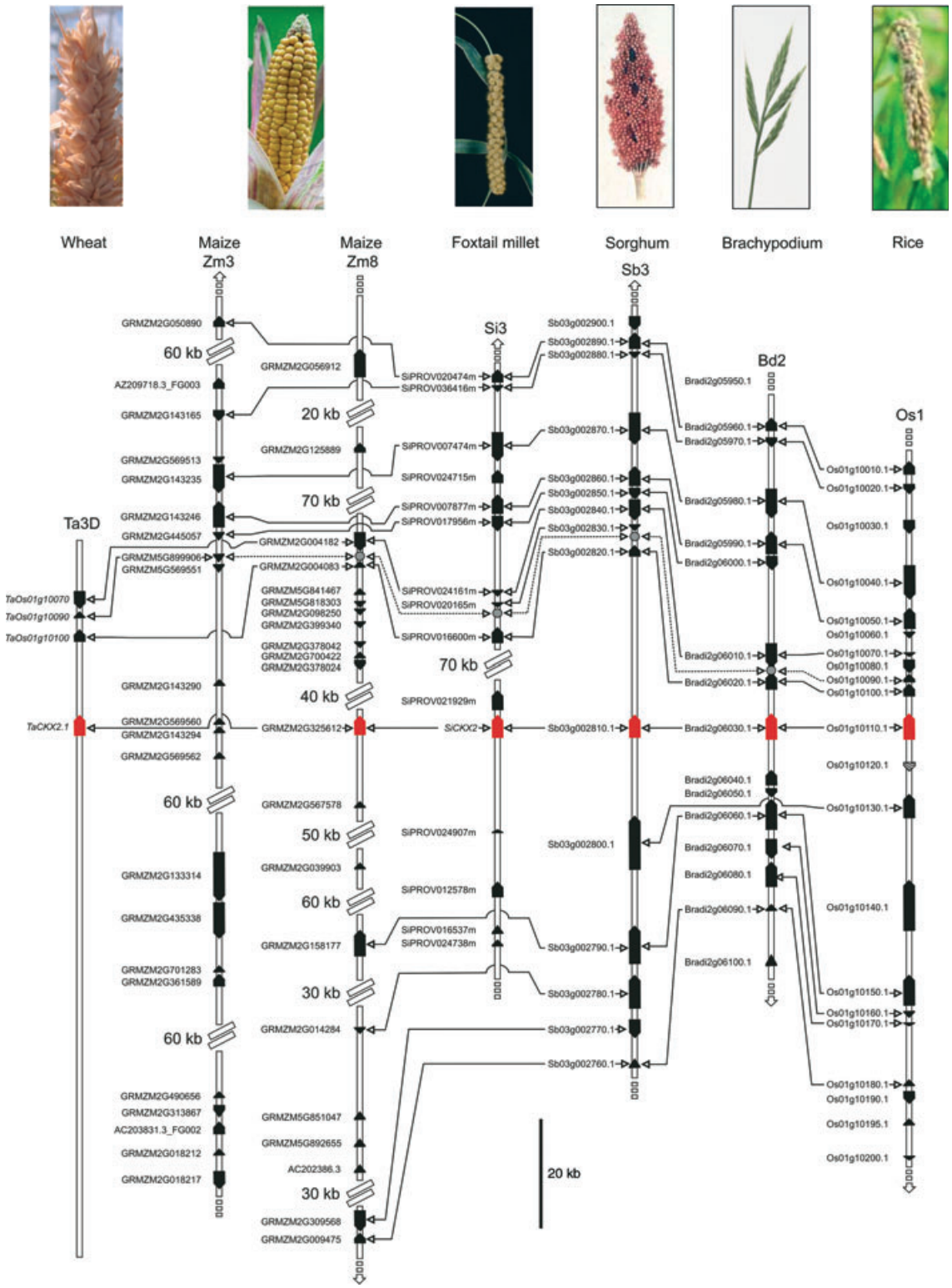

Figure 5 Comparative analysis of physical maps around *cytokinin oxidase/dehydrogenases 2* (*CKX2*) loci in cereal genomes. Sequenced wheat bacterial artificial chromosome clones for *TaCKX2.3* (chromosome 3D, JF292901), *TaCKX2.4* (chromosome 3A, JN381556) and *TaCKX2.5* (chromosome 3B, JN381555) are included. Dashed lines link gene models with significant sequence similarity to genomic regions which lack predicted genes (grey circles). *CKX2* genes are highlighted in red. Chromosome orientations (direction to the long arm telomere) are indicated.

physically linked in all sequenced grass species investigated, their depth of phylogenetic divergence is the greatest amongst all subclades. Our findings indicate that this local gene duplication arose within the ancestral grass genome, prior to the WGD event. Their older origins, relative to family members duplicated during the ancestral WGD, provide greater evolutionary time

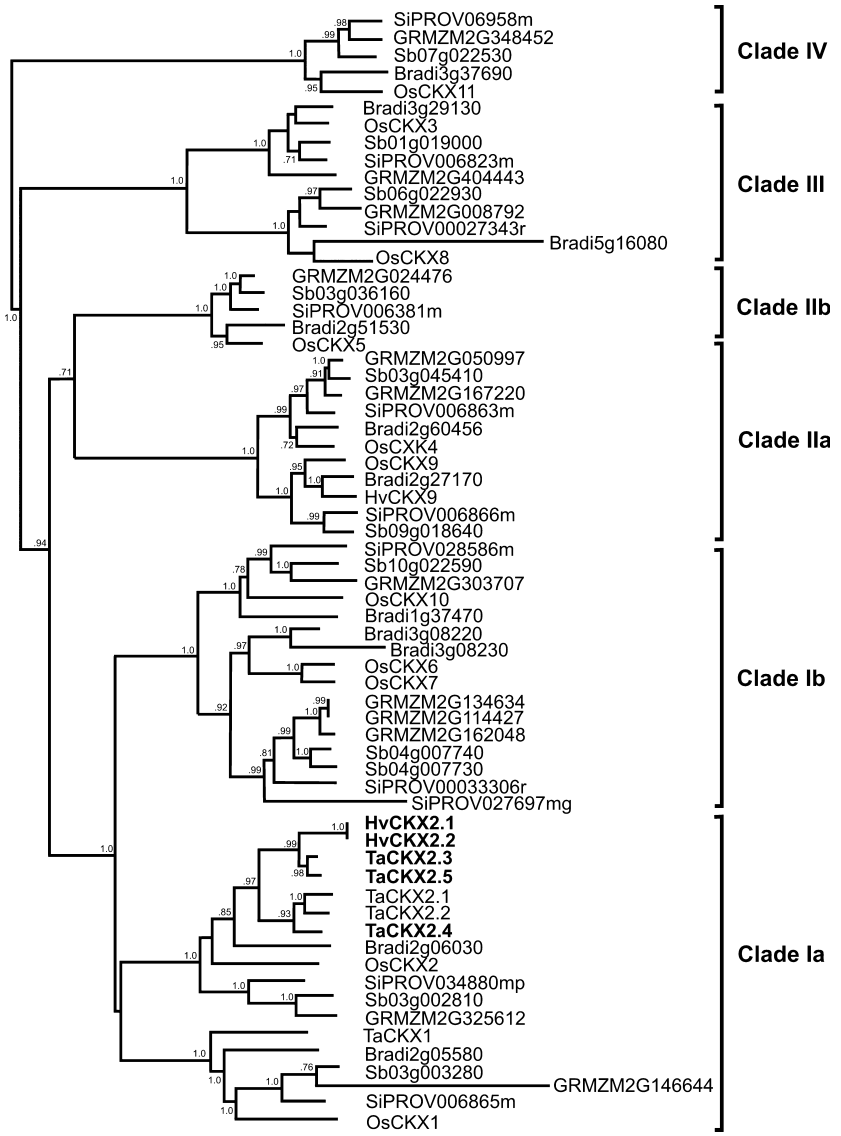

Figure 6 Phylogenetic tree of all 66 full-length cytokinin oxidase/dehydrogenases 2 (CKX2) family members from rice, brachypodium, sorghum, maize and foxtail millet, as well as full-length members from wheat and barley. CKX2 sequences identified in this study are highlighted in bold. TaCKX1 = Feng *et al.* (2008). Bootstrap proportions (1000 replicates) ≥0.70 are indicated.

for sub- or neo-functionalization to potentially occur. In the grasses, *CKX1* genes are expressed in a wide variety of tissues, with *TaCKX1* expressed in leaves, sheaths and roots (Feng *et al.*, 2008), *HvCKX1* in the roots, leaves and grain (Galuszka *et al.*, 2004; Zalewski *et al.*, 2010) and *ZmCKX1* in kernels and the vascular system of seedling shoots, roots and coleoptiles (Brugière *et al.*, 2003; Galuszka *et al.*, 2005; Smehilova *et al.*, 2009). Genetic mapping resolved the rice grain number locus, *Gn1*, into two closely linked QTL of roughly equal effect: *Gn1a* (*OsCKX2*) and *Gn1b* (Ashikari *et al.*, 2005). While *Gn1a* is controlled by *OsCKX2*, *Gn1b* remains to be identified, mapping distal to marker R3192. *OsCKX1*, which phylogenetic analysis shows to be closely related to *OsCKX2*, is located ~410 kb distal to R3192 and therefore represents a possible candidate for *Gn1b*. This hypothesis is supported by the recent finding that RNAi silencing of *HvCKX1* in barley results in higher grain yield, as well as root weight and height (Zalewski *et al.*, 2010).

Clade 1a: *CKX2* orthologs

The following findings indicate true *OsCKX2* orthologs have been identified in the six grass species investigated: (i) high levels of sequence conservation at the DNA/protein level, (ii) conservation of intron/exon structure, (iii) close phylogenetic clustering and (iv) macro-/micro-colinear physical/genetic map locations. As *OsCKX2* belongs to a gene family of eleven members, it has been suggested that spatial and temporal differences in expression may have contributed to functional differentiation (Ashikari *et al.*, 2005; Zalewski *et al.*, 2010). Indeed, tissue- and organ-specific expression has been demonstrated for arabidopsis *CKX* genes using promoter:*GUS* fusions (Werner *et al.*, 2003), with AtCKX enzymes differing in their biochemical properties and subcellular localization (Schmülling *et al.*, 2003). In rice, *OsCKX2* expression is low. It is predominantly expressed in the culms, inflorescence meristems and spikelets and was not represented in rice EST libraries prior to map-based cloning (Ashikari *et al.*, 2005). Despite large EST libraries in the species investigated here (ranging from 20 323 accessions in brachypodium to 1 032 224 in maize), evidence of low *CKX2* transcript abundance is mirrored in wheat (4 *CKX2* ESTs), barley (0), brachypodium (1, DV475865), sorghum (0) and maize (1, CO452786), with all of these transcripts originating from ear/grain tissue. Furthermore, the recently characterized

genes *TaCKX2.1* and *TaCKX2.2* have been found to be expressed predominantly in the wheat culm and spike, (Zhang *et al.*, 2010).

The identification of a Triticeae-specific *CKX2* duplication that is likely to have occurred prior to the divergence of wheat and barley ~11 mya (Huang *et al.*, 2002) suggests the potential phenotypic effects of Triticeae *CKX2* variants may be masked, because of functional redundancy. Together with the occurrence of natural rice *OsCKX2* gene deletions that result in increased grain number, these findings indicate that intra-specific *CKX2* copy number variation may be of importance to agronomic traits in cereals. Indeed, copy number variation of paralogous *CONSTANS*-like genes that arose during the ancestral WGD event has previously been shown to play a critical role in the agronomic attributes of related grass crops (Cockram *et al.*, 2010b). The establishment of high Triticeae *CKX2* copy number will inform future functional validation using reverse genetic approaches, favouring RNA interference (RNAi) approaches over TILLING (McCallum *et al.*, 2000), as combining closely linked TILLed genes in a single genetic background would be problematic. The finding that *HvCKX2.2* possesses an exonic deletion that is predicted to result in the truncation of the cytokinin-binding domain in cv. OWB-R indicates EcoTILLING (Comai *et al.*, 2003) may prove fruitful, at least in diploid grasses. Despite the identification of the truncated *HvCKX2.2* allele, no grain number QTL overlapping its genetic map position has been reported in the OWB population (Chutimanitsakun *et al.*, 2011), suggesting that *HvCKX2.2* is functionally redundant with *HvCKX2.1*, or that it does not control grain number in barley. Interestingly, analysis of twelve wheat varieties, Zhang *et al.* (2010) reported an association between increased *TaCKX2.1/TaCKX2.2* expression and increased grain number per spike. However, this is opposite to the effects on floral meristem development observed in rice and arabidopsis, where overexpression of *CKX2* genes results in reduced floret/flower primordia establishment, while null variants and transgenic plants expressing *OsCKX2* antisense constructs display increased grains per panicle (Ashikari *et al.*, 2005). Antagonsitic effects of arabidopsis *CKX* genes between different tissues have also been observed, in which silencing of *CKX* genes stimulated shoot growth and repressed root development (Werner *et al.*, 2003). We note that as well as local function in the tissues within which they are synthesized, CKs act as long distance signalling messengers (Kudo *et al.*, 2010) and are able to establish or modulate sink strength (Argueso *et al.*, 2009; Kuiper, 1993; Guivarc'h *et al.*, 2002; Werner *et al.*, 2008). As plant development combines the effects of cell proliferation/expansion with whole plant sink/source balance plant (Vercruyssen *et al.*, 2011), it has been suggested that the observed species-specific differences in the physiological effects of *CKX* silencing could be due to differences in sink strength (Zalewski *et al.*, 2010). To clarify the function of *CKX2* in wheat, an RNAi silencing approach has been initiated to investigate its role *in planta*.

Clade 1b

Clade Ib contains two to three tandemly duplicated *CKX* genes (*OsCKX6* and *OsCKX7* in rice), as well as *CKX10* members in all species investigated, with the exception of barley, where no *OsCKX6* orthologue was identified. The presence of variable numbers of genes orthologous to the *OsCKX6/OsCKX7* cluster in multiple grass species indicates this genomic region is prone to localized genomic rearrangements such as unequal crossing-over (Achaz *et al.*, 2000), which act to increase or decrease copy number of tandemly duplicated genes. In rice, brachypodium, sorghum and maize, *CKX10* orthologs are located within regions of intra-specific colinearity with the *CKX6/CKX7* cluster, indicating they are related by evolutionary descent after arising as part of the ancestral WGD event. While the lack of ESTs for any of the clade Ib genes in rice, wheat and barley indicates they display low or tissue/treatment-dependent expression, RT-PCR expression analysis in rice found very low *OsCKX6* and *OsCKX10* transcription in the shoot apical meristem (SAM) (Ashikari *et al.*, 2005).

Clades II–IV

Poaceae *CKX4* and *CKX9* members all cluster within clade IIa and are likely to have originated because of the ancestral grass WGD event. Little is known about these members, although *HvCKX9* has been found to be expressed in the grain and 7-day-old leaves (Galuszka *et al.*, 2004). The deep phylogenetic roots of clades III and IV suggest that they have distinct evolutionary histories relative to clades I and II. Within clade III, grass *CKX3* and *CKX8* members group together, despite the genomic regions in rice not having been identified as participating in the recent WGD event (based on nonoverlapping regions of duplication and similar Ks distribution). However, these regions are nevertheless involved in a small region of segmental duplication (http://rice.plantbiology.msu.edu/), indicating that they share a common evolutionary origin. This suggests a local duplication, or remnants of an older WGD event, that occurred before the evolutionary divergence of the Poaceae. However, analysis of micro-synteny around *CKX3* and *CKX10* genes in other sequenced grass species found no further evidence of such duplications, indicating further investigation is needed to establish the evolutionary timing of this event. Finally, clade IV contains just one member (*CKX11*), which has previously been phylogenetically grouped with *AtCKX7* (Zhang *et al.*, 2010), which encodes a nonsecretory cytosolic enzyme in arabidopsis (Kowalska *et al.*, 2010). Although little is known about clade IV members, genomic/cDNA Triticeae *CKX11* homologues have been identified here, indicating that *CKX11* is expressed in diverse tissues in wheat/barley (root, leaf, grain), while in rice, it is relatively highly expressed in the SAM (Ashikari *et al.*, 2005).

Future prospects

Studies of *CKX3* and *CKX5* double mutants in arabidopsis show that CKs delay cellular differentiation in the apical meristem and terminal floral cellular differentiation, as well as a promotion of placental growth, leading to an increased seed set per silique (Bartrina *et al.*, 2011). The resulting increase in cytokinin content was associated with an increase in seed yield of over 50%, illustrating the role of sink strength as a yield factor. Recent findings in arabidopsis are also beginning to shed light on the interaction of CKs with other developmental pathways that regulate growth and biomass. Crosstalk between CKs and the brassinosteroid, gibberellin and cell proliferation pathways have been found to affect biomass accumulation in leaves and roots (Vercruyssen *et al.*, 2011). Similarly, the role of CKs in the control of flowering time has been shown to act via transcriptional regulation of the *FLOWERING LOCUS T* (*FT*) paralog, *TWIN SISTER OF FT* (*TSF*), as well as *FD* and the floral pathway integrator, *SUPPRESSOR OF EVEREXPRESSION OF CONSTANS1* (*SOC1*) (D'Aloia *et al.*,

2011), allowing integration of CKs into the otherwise relatively well-defined floral molecular pathway. These studies highlight the potential impact that detailed genetic and physiological investigation of the effects of *CKX* genetic variation may have in crop species. The work presented here provides the molecular and phylogenetic basis from which systematic evaluation of grass *CKX* gene family function can be undertaken, allowing further understanding of the roles CKs play in the modulation of sink strength, biomass and yield in cereal crops, and the basis from which yield gains comparable to those grained in rice could potentially be achieved in other Poaceae crops. Finally, by identifying what is most likely the complete barley *CKX* gene family, we demonstrate that NGS shotgun sequencing of the nuclear genome provides an efficient method of determining the gene space of large-genome crop species.

Experimental procedures

Bioinformatic identification and analysis of *CKX* genes in cereals

CKX copy number in the rice (*Oryza sativa* ssp. *japonica* cv. Nipponbare) genome (MSU Osa1 assembly v6.1, http://rice.plantbiology.msu.edu/) was verified by BLASTn (Altschul *et al.*, 1997) searches (match/mismatch scores = 2,−3; gap costs: existence = 5, extension = 2) using *OsCKX1* to *OsCKX11* CDS (Ashikari *et al.*, 2005). An expectation-value threshold (the number of times that a match, or a better match, occurs by chance within the database) of 1e−40 was used, and the results compared to those reported by Ashikari *et al.* (2005). While *e*-values of zero are reported here, these actually represent positive *e*-values too small for the precision of the software used. Identification of *CKX* genes within the genome assemblies of sequenced cereals was conducted using *OSCKX1* to *OsCKX11* CDS for BLASTn searches against the gene models of *B. distachyon* accession BD21 (assembly v1.0, using sequence data produced by the US Department of Energy Joint Genome Institute, http://modelcop.org/), *S. bicolour* cv. Moench (v1.0, http://phytozome.net/), *Z. mays* (cv. B73, v5b.60, http://www.maizesequence.org/) and *S. italica* cv. Yugu1 (assembly v1.0, using sequence data produced by the US Department of Energy Joint Genome Institute, http://phytozome.net/). Where necessary, gene models were reassessed using FGENESH (http://www.softberry.com/). Barley (*H. vulgare*) *CKX* genomic sequence fragments were identified by BLASTn searches of rice *CKX* CDS versus a 28× NGS genomic survey of cv. Morex (Stein *et al.*, unpublished). In addition, identification of *CKX* genes within the transcriptomes of barley and wheat (*T. aestivum*) was conducted using *OsCKX1* to *OsCKX11* as queries for BLASTn searches of TIGR TA databases (release 2.0, http://www.tigr.org/). To clarify homology groupings, hits for all species were used for back-BLASTn analysis of the rice genome (i.e. used as queries for BLASTn searches of the rice genome determine whether they showed highest homology to the original *CKX* gene query). Predicted protein sequences were obtained using the VectorNTI Advance package v10.1.1 (Invitrogen, Paisley, UK), and protein domains identified using pfam (http://pfam.sanger.ac.uk/). Inclusion criteria for CKX proteins identified in species with sequenced genomes were the presence of the FAD and cytokinin-binding domains. Protein alignments were performed using CLUSTALW (Thompson *et al.*, 1994), with manual adjustment/editing using GENEDOC v2.6 (http://www.nrbsc.org/gfx/genedoc/) and Jalview v2.6.1 (Waterhouse *et al.*, 2009). *CKX* genes are prefixed with the corresponding genus and species initials. *CKX* numbering follows orthology with rice *CKX* genes (Ashikari *et al.*, 2005). Accordingly, the nomenclature of barley/wheat *CKX* genes differs to that listed in Galuszka *et al.* (2004).

Molecular characterization of *CKX2* homologues in wheat and barley

Plasmid DNA was extracted using the Qiaprep Kit (Qiagen, Crawley, UK) and the insert sequenced by primer walking (details listed in Table S6), initiated with generic T7 and SP6 primers. Sequencing was performed using BigDye kit v3.1 (Applied Biosystems) and the resulting extension products separated and visualized using a 3730 DNA Analyzer (Applied Biosystems). Sequence editing and assembly was performed using Vector NTI. Inter-chromosomal mapping of *TaCKX2* genes was performed using the wheat NT chromosome deletion lines (Sears, 1966). Genomic DNA was extracted using the DNeasy 96 Plant Kit (Qiagen) and used as a template for PCR amplification of *TaCKX2* genes using *Pfu*-turbo (Agilent, Edinburgh, UK) and primers listed in Table S6. PCR products were separated by agarose gel electrophoresis. Amplicon specificity was verified by direct sequencing of PCR products, as described by Chiapparino *et al.* (2006). Mapped SNPs were identified by direct sequencing of PCR products amplified from parental DNAs of the mapping populations listed below, using primers TaCKX2.3-F4/TaCKX2.3-R4, TaCKX2.4-F1/TaCKX2.4-R1 and TaCKX2.5-F1/TaCKX2.5-R1 (Table S6). Intra-chromosomal mapping of *TaCKX2* genes was conducted by fluorescent single-base extension using the ABI Prism SNaPshot Multiplex Kit (Applied Biosystems) as previously described (Cockram *et al.*, 2010b). *TaCKX2.3* was mapped in the Avalon × Cadenza doubled-haploid mapping population (http://www.wgin.org.uk/), genotyped using amplification primers TaCKX2.3-F4/-R4 and single-base extension primer TaCKX2.3-SNAPR. *TaCKX2.4* was mapped in the 'Opata 85' × 'synthetic W7984' population (http://wheat.pw.usda.gov/ggpages/maps.html/) using amplification primers TaCKX2.40F1/TaCKX2.4-R1 and extinction primer TaCKX2.4-SNAPF. *TaCKX2.5* was mapped in the 'Helidur' × TD161' (RAGT, unpublished data) using amplification primers TaCKX2.5-F1/TaCKX2.5-R1 and extension primer TaCKX2,5-SNAPF. Details of SNPs used for mapping are shown in Table S7. All genetic mapping was performed using JoinMap v3.0 (Van Ooijen and Voorrips, 2004). *TaCKX2* primers designed from fully sequenced EST clone TaEo5040F02 were used to PCR screen a subset of the 4× cv. Chinese Spring *T. aestivum* BAC library (http://cnrgv.toulouse.inra.fr/), using methods described by Alfares *et al.* (2009). BAC clones were sized and fingerprinted as described by Foote *et al.* (2004) and selected clones fully sequenced (completed as a service provided by Washington State University). BAC sequence annotation was completed following the protocols of Sabot *et al.* (2005).

HvCKX2.1 and *HvCKX2.2* genomic sequences were obtained using barley NGS contigs to design primers for PCR amplification/primer walking (Table S6). All additional barley genomic contigs are part of a CLC assembly (http://www.clcbio.com/) using 2×730 683 996 sequences (2 × 101 bp; total 147.6 Gb; ~28× genomic coverage) from three barley (cv. Morex) Illumina-Paired-End libraries (500 bp distance). Genetic mapping was conducted using the Oregon-Wolfe barley (OWB) mapping

population (Costa *et al.*, 2001). A 32-bp InDel identified between the parental lines OWB-D and OWB-R was used to genotype the mapping population, by separation of PCR products (amplified using primer pair HvCKX2.1-F4/-R4) on ethidium bromide stained 1% agarose gels. SNP T + 1583/C was used to genetically map *HvCKX2.2* by fluorescent single-base extension using the SNaPshot Multiplex Kit in combination with amplification primers HvCKX2.2-F3/-R3 and single base-pair extension primer HvCKX2.2-SNAPF. Genetic mapping was performed using JoinMap v3.0, and map positions integrated into the barley consensus map (Close *et al.*, 2009). Additional barley *CKX* genes were genetically mapped using the OWB or 'Arta' × '*H. spontaneum*' (Baum *et al.*, 2003) populations, using the primers and polymorphisms listed in Table S3.

Phylogenetic and comparative analysis

Protein alignments were constructed using CLUSTALW, and phylogenetic analysis undertaken using the PHYLIP package, v3.5 (Felsenstein, 1989). Unrooted phylogenies were estimated by the distance matrix method, with tree topographies supported by bootstrapping (1000 replicates). Analysis was based on alignments of FAD- and cytokinin-binding domains separately (for partial predicted proteins based on ESTs or partial NGS genomic sequence), or on both domains simultaneously (for full-length predicted proteins). Macro-colinearity between cereal genomes is as previously described (Devos, 2005; International Brachypodium Initiative, 2010). Verification of micro-colinearity around each *CKX* member was performed using ±10 genes either side of each *OsCKX* gene as queries for BLASTn searches of the genomes of each sequenced grass species, using the BLAST parameters described above. Location of rice WGD genome duplications was determined as described by Thiel *et al.* (2009), updated using the most recent rice assembly and plotted using Circos (Krzywinski *et al.*, 2009). Colinearity between the genetic map of barley chromosome 3H and the physical map of rice chromosome Os1 was performed as previously described (Cockram *et al.*, 2010a).

DNA sequences have been deposited in GenBank: accessions JF292901, JF293074 to JF293079, JF495479 to JF495489, JN381555 and JN381556.

Acknowledgements

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) Industrial CASE studentship (BBS/S/M/2006/13029) awarded to RAGT, KWS and NIAB. Additional financial support was provided by the NIAB Trust. Barley NGS data were generated by a project funded by German Ministry of Education and Research, BMBF (GABI-BARLEX 0314000). We thank the Cereal Research Centre at the Agriculture and Agri-food Canada for provision of the wheat EST clone, INRA-CNRGV for supply of BAC clones, J Bennetzen for permission to analyse the *CKX* gene family in the unpublished foxtail millet genome, and Mike Ambrose and Phillipe Vain (John Innes Centre) for images of brachypodium and foxtail millet. The population of doubled-haploid individuals derived from F_1 progeny of a cross between cvs. Avalon and Cadenza was developed by Clare Ellerbrook, Liz Sayers and the late Tony Worland (John Innes Centre, UK), as part of a Defra funded project led by ADAS.

References

Achaz, G., Coissac, E., Viari, A. and Netter, P. (2000) Analysis of intrachromosomal repeats in yeast *Saccharomyces cerevisiae*: a possible model for their origin. *Mol. Biol. Evol.* **17**, 1268–1275.

Alfares, W., Bouguennec, A., Balfourier, F., Gay, G., Bergès, H., Vautrin, S., Sourdille, P., Bernard, M. and Feuillet, C. (2009) Fine mapping and marker development for the crossability gene *SKr* on chromosome 5BS of hexaploid wheat (*Triticum aestivum* L.). *Genetics*, **183**, 469–481.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

Argueso, C.T., Ferreira, F.J. and Kieber, J.J. (2009) Environmental perception avenues: the interaction of cytokinin and environmental response pathways. *Plant Cell. Environ.* **32**, 1147–1160.

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takahashi, T., Nishimura, A., Angeles, E.R., Qian, Q., Kitano, H. and Matsuoka, M. (2005) Cytokinin oxidase regulates rice grain production. *Science*, **309**, 741–745.

Bartrina, I., Otto, E., Strnad, M., Werner, T. and Schmülling, T. (2011) Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. *Plant Cell*, **23**, 69–80.

Baum, M., Grando, S., Backes, G., Jahoor, A., Sabbagh, A. and Ceccarelli, S. (2003) QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' × '*H. spontaneum* 14-1'. *Theor. Appl. Genet.* **107**, 1215–1225.

Beales, J., Turner, A., Griffiths, S., Snape, J.W. and Laurie, D.A. (2007) A pseudo-response regulator is missexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **115**, 721–733.

Bennetzen, J.L., Ma, J. and Devos, K.M. (2005) Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* **95**, 127–132.

Bilyeu, K.D., Cole, J.L., Laskey, J.G., Riekhof, W.R., Esparza, T.J., Kramer, M.D. and Morris, R.O. (2001) Molecular and biochemical characterization of a cytokinin oxidase from maize. *Plant Physiol.* **125**, 378–386.

Bolot, S., Abrouk, M., Masood-Quraishi, U., Stein, N., Messing, J., Feuillet, C. and Salse, J. (2009) The 'inner circle' of the cereal genome. *Curr. Opin. Plant Biol.* **12**, 119–125.

Brownlee, B.G., Hall, R.H. and Whitty, C.D. (1975) 3-Methyl-2-butenal: an enzymatic degradation product of the cytokinin N^6-(Δ^2-isopentenyl)adenine. *Can. J. Biochem.* **53**, 37–41.

Bruggmann, R., Bharti, A.K., Gundlach, H., Lai, J., Young, S., Pontaroli, A.C., Wei, F., Haberer, G., Fuks, G., Du, C., Raymond, C., Estep,, M.C., Liu, R., Bennetzen, J.L., Chan, A.P., Rabinowicz, P.D., Quackenbush, J., Barbazuk, W.B., Wing, R.A., Birren, B., Nusbaum, C., Rounsley, S., Mayer, K.F. and Messing, J. (2006) Uneven chromosome contraction and expansion in the maize genome. *Genome Res.* **16**, 1241–1251.

Brugière, N., Jiao, S., Hantke, S., Zinselmeier, C., Roessler, J.A., Niu, X., Jones, R.J. and Habben, J.E. (2003) Cytokinin oxidase gene expression in maize is localized to the vasculature, and is induced by cytokinins, abscisic acid, and abiotic stress. *Plant Physiol.* **132**, 1228–1240.

Chiapparino, E., .Donini, P., Reeves, J., Tuberosa, R. and O'Sullivan, D.M. (2006) Distribution of *ß*-amalayse I haplotypes among European cultivated barleys. *Mol. Breed.* **18**, 341–354.

Chutimanitsakun, Y., Nipper, R.W., Cuesta-Marcos, A., Cistué, L., Corey, A., Filichkina, T., Johnson, E.A. and Hayes, P.M. (2011) Construction and application for QTL analysis of a restriction site associated DNA (RAD) linkage map in barley. *BMC Genomics*, **12**, e4.

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S. *et al.* (2009) Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics*, **10**, e582.

Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A. and Greenland, A.J. (2007) Control of flowering time in temperate cereals: genes, domestication and sustainable productivity. *J. Exp. Bot.* **58**, 1231–1244.

Cockram, J., White, J., Zuluaga, D.L., Smith, D., Comadran, J., Macaulay, M., Luo, Z., Kearsey, M.J., Werner, P., Harrap, D. *et al.* (2010a) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl Acad. Sci. USA*, **107**, 21611–21616.

Cockram, J., Howells, R. and O'Sullivan, D.M. (2010b) Segmental duplication harbouring group IV *CONSTANS*-like genes in cereals. *Genome*, **53**, 231–240.

Comai, L., Young, K., Till, B., Reynolds, S., Green, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R. and Henikoff, S. (2003) Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *Plant J.* **37**, 778–786.

Costa, J.M., Corey, A., Hayes, P.M., Jobet, C., Kleinhofs, A., Kopisch-Obusch, A., Kramer, S.F., Kudrna, D., Li, M., Riera-Lizarazu, O. *et al.* (2001) Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. *Theor. Appl. Genet.* **103**, 415–424.

D'Aloia, M., Bonhomme, D., Bouché, F., Tamseddak, K., Ormenese, S., Torti, S., Coupland, G. and Périlleux, C. (2011) Cytokinin promotes flowering of *Arabidopsis* via transcriptional activation of the *FT* paralogue *TSF*. *Plant J.* **65**, 972–979.

Devos, K.M. (2005) Updating the 'crop circle'. *Curr. Opin. Plant Biol.* **8**, 155–162.

Felsenstein, J. (1989) PHYLIP—phylogeny inference package (version 3.2). *Cladistics*, **5**, 164–166.

Feng, D.S., Wang, H.G., Zhang, X.S., Kong, L.R., Tian, J.C. and Li, X.F. (2008) Using an inverse PCR method to clone the wheat cytokinin oxidase/dehydrogenase gene *TaCKX1*. *Plant Mol. Biol. Rep.* **26**, 143–155.

Foote, T., Griffiths, S., Allouis, S. and Moore, G. (2004) Construction and analysis of a BAC library in the grass *Brachypodium sylvaticum*: its use as a tool to bridge the gap between rice and wheat in elucidating gene content. *Funct. Integr. Genomics*, **4**, 26–33.

Gale, M.D. and Devos, K.M. (1998) Comparative genetics in the grasses. *Proc. Natl Acad. Sci. USA*, **95**, 1971–1974.

Galuszka, P., Frébortova, J., Werner, T., Yamada, M., Strnad, M., Schmülling, T. and Frébort, I. (2004) Cytokinin oxidase/dehydrogenase genes in barley and wheat: cloning and heterologous expression. *Eur. J. Biochem.* **271**, 3990–4002.

Galuszka, P., Frébortova, J., Luhova, L., Bilyeu, K.D., English, J.T. and Frébort, I. (2005) Tissue localization of cytokinin dehydrogenase in maize: possible involvement of quinone species generated from plant phenolics by other enzymatic systems in the catalytic reaction. *Plant Cell Physiol.* **46**, 716–728.

Goff, S.A., Cone, K.C. and Chandler, V.L. (1992) Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes Dev.* **6**, 864–875.

Guivarc'h, A., Rembur, J., Gotez, M., Roitsch, T., Noin, M., Schülling, T. and Chriqui, D. (2002) Local expression of the *ipt* gene in transgenic tobacco (*Nicotiana tabacum* L. cv. SR1) axillary buds establishes a role for cytokinins in tuberization and sink formation. *J. Exp. Bot.* **53**, 621–629.

Houba-Hérin, N., Pethe, C., d'Alayer, J. and Laloue, M. (1999) Cytolinin oxidase from *Zea mays*: purification, cDNA cloning and expression in moss protoplasts. *Plant J.* **17**, 615–626.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R. and Gornicki, P. (2002) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. *Plant Mol. Biol.* **48**, 805–820.

International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, **11**, 763–768.

Kowalska, M., Galuszka, P., Frébortova, J., Šebela, M., Béres, T., Hluska, T., Šmehilová, M., Bilyeu, K.D. and Frébot, I. (2010) Vacuolar and cytosolic cytokinin dehydrogenases of *Arabidopsis thaliana*: heterologous expression, purification and properties. *Phytochemistry*, **71**, 1970–1978.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645.

Kudo, T., Kiba, T. and Sakakibara, H. (2010) Metabolism and long-distance translocation of cytokinins. *J. Integr. Biol* **52**, 53–60.

Kuiper, D. (1993) Sink strength: established and regulated by plant growth regulators. *Plant Cell Environ.* **16**, 1025–1026.

Malito, E., Coda, A., Bilyeu, K.D., Fraaije, M.W. and Mattevi, A. (2004) Structure of michaelis and product complexes of plant cytokinin dehydrogenase: implications for flavoenzyme catalysis. *J. Mol. Biol.* **341**, 1237–1249.

McCallum, C.M., Comai, L., Greene, E.A. and Henikoff, S. (2000) Targeted screening for induced mutations. *Nat. Biotechnol.* **18**, 455–457.

Paterson, A.H., Bowers, J.E. and Chapman, B.A. (2004) Ancient polyploidization predating the divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl Acad. Sci. USA*, **101**, 9903–9908.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T. *et al.* (2009) The *Sorghum bicolour* genome and the diversification of grasses. *Nature*, **457**, 551–556.

Sabot, F., Guyot, R., Wicker, T., Chantret, N., Aubin, B., Chalhoub, B., Leroy, P., Sourdille, P. and Bernard, M. (2005) Updating of transposable element annotations from large wheat genomic sequences reveals diverse activities and gene associations. *Mol. Genet. Genomics*, **274**, 119–130.

Sakakibara, H. (2006) Cytokinins: activity, biosynthesis and translocation. *Annu. Rev. Plant Biol.* **57**, 431–449.

Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U.M., Calcagno, T., Cooke, R., Delseny, M. and Feuillet, C. (2008) Identification and characterization of shared duplications between rice and wheat provide new insights into grass genome evolution. *Plant Cell*, **20**, 11–24.

Schmülling, T., Werner, T., Riefler, M., Krupkova, E. and Bartrina y Manns, I. (2003) Structure and function of cytokinin oxidase/dehydrogenase genes in maize, rice, Arabidopsis and other species. *J. Plant. Res.* **116**, 241–252.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A. *et al.* (2009) The B73 maize genome: complexity, diversity and dynamics. *Science*, **326**, 1112–1115.

Sears, E. (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In *Chromosome Manipulation and Plant Genetics* (Riley, R. and Lewis, K.R., eds), pp. 29–45. London: Oliver & Boyd.

Smehilova, M., Galuszka, P., Bilyeu, K.D., Jaworek, P., Kowalska, M., Sebelam, M., Sedlarova, M., English, J.T. and Frebort, I. (2009) Subcellular localization and biochemical comparison of cytosolic and secreted cytokinin dehydrogenase enzymes from maize. *J. Exp. Bot.* **60**, 2701–2712.

Somers, D., Isaac, P. and Edwards, K. (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **109**, 1105–1114.

Thiel, T., Graner, A., Waugh, R., Grosse, I., Close, T.J. and Stein, N. (2009) Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. *BMC Evol. Biol.* **9**, e209.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Turner, A., Beales, J., Faure, S., Dunford, R.P. and Laurie, D.A. (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science*, **310**, 1031–1034.

Van Ooijen, J.W. and Voorrips, R.E. (2004) *JoinMap 3.0: Software for the Calculation of Genetic Linkage Maps in Experimental Populations*. Wageningen, The Netherlands: CPRO-DLO.

Vaseva-Gemisheva, I., Lee, D., Alexieva, V. and Karanov, E. (2004) Cytokinin oxidase/dehydrogenase in *Pisum sativum* plants during vegetative development: influence of UV-B irradiation and high temperature on the enzymatic activity. *Plant Growth Regul.* **46**, 199–208.

Vercruyssen, L., Gonzalez, N., Werner, T., Schmülling, T. and Inzé, D. (2011) Combining enhanced root and shot growth reveals crosstalk between pathways that control plant organ size in Arabidopsis. *Plant Physiol.* **155**, 1339–1352.

Vyroubalová, Š., Václavíková, K., Turečková, V., Novák, O., Smehilová, M., Hluska, T., Ohnoutková, L., Frébort, I. and Galuszka, P. (2009) Characterization of new maize genes putatively involved in CK metabolism

and their expression during osmotic stress in relation with cytokinin levels. *Plant Physiol.* **151**, 433–477.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.

Wei, F., Coe, E., Nelson, W., Bharti, A.K., Engler, F., Butler, E., Kim, H., Goicoechea, J.L., Chen, M., Lee, S. *et al.* (2007) Physical and genetic structure of the maize genome reflects its complex evolutionary history. *PLoS Genet.* **3**, e123.

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H. and Schmülling, T. (2003) Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell*, **15**, 2532–2550.

Werner, T., Holst, K., Pörs, Y., Guivarc'h, A., Mustroph, A., Chriqui, D., Grimm, B. and Schmülling, T. (2008) Cytokinin deficiency causes distinct changes of sink and source parameters in tobacco shoots and roots. *J. Exp. Bot.* **59**, 2659–2672.

Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C. *et al.* (2005) The genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* **3**, e38.

Zalewski, W., Galuszka, P., Gasparis, S., Orczyk, W. and Nadolska-Orczyk, A. (2010) Silencing of the *HvCKX1* gene decreases the cytokinin oxidase/dehydrogenase level in barley and leads to higher plant productivity. *J. Exp. Bot.* **61**, 1839–1851.

Zhang, J., Liu, W., Yang, X., Gao, A., Li, X., Wu, X. and Li, L. (2010) Isolation and characterization of two putative cytokinin oxidase genes related to grain number per spike in wheat. *Mol. Biol. Rep.* **38**, 2337–2347.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Re-analysis of the foxtail millet *SiCKX2* gene prediction.

Figure S2 Alignment of the fully sequenced EST clone TaE5040F02 (*TaCKX2.3*, JF293079) with *TaCKX2.1* (FJ648070) and *TaCKX2.2* (GU084177) cDNAs.

Figure S3 Alignment of rice, wheat and barley cytokinin oxidase/dehydrogenases 2 proteins.

Figure S4 Phylogenetic analysis of Poaceae cytokinin oxidase/dehydrogenases proteins based on separate protein alignments of the FAD and cytokinin domains.

Table S1 Genomic barley *cytokinin oxidase/dehydrogenases* sequences generated.

Table S2 *Cytokinin oxidase/dehydrogenases* homologues identified by BLASTn interrogation of TIGR wheat and barley Transcript Assemblies.

Table S3 Primers and further information for the genetic mapping of barley *cytokinin oxidase/dehydrogenases* genes.

Table S4 Micro-colinearity around *cytokinin oxidase/dehydrogenases* genes in sequenced grass species.

Table S5 Number of colinear genes identified between the *cytokinin oxidase/dehydrogenases 2* regions of sequenced cereal species.

Table S6 Wheat and barley *cytokinin oxidase/dehydrogenases 2* sequencing primers.

Table S7 Details of SNPs used to map *cytokinin oxidase/dehydrogenases 2* genes in wheat.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.