Annals of Botany **101**: 863–872, 2008
doi:10.1093/aob/mcm252, available online at www.aob.oxfordjournals.org

OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Detailed Dissection of the Chromosomal Region Containing the *Ph1* Locus in Wheat *Triticum aestivum*: With Deletion Mutants and Expression Profiling

NADIA AL-KAFF, EMILIE KNIGHT, ISABELLE BERTIN, TRACIE FOOTE, NICOLA HART, SIMON GRIFFITHS and GRAHAM MOORE*

John Innes Centre, Norwich Research Park, Colney Lane, Norwich, Norfolk NR4 7UH, UK

Received: 11 May 2007 Returned for revision: 3 August 2007 Accepted: 20 August 2007 Published electronically: 20 October 2007

• *Background and Aims* Understanding *Ph1*, a dominant homoeologous chromosome pairing suppressor locus on the long arm of chromosome 5B in wheat *Triticum aestivum* L., is the core of the investigation in this article. The *Ph1* locus restricts chromosome pairing and recombination at meiosis to true homologues. The importance of wheat as a crop and the need to exploit its wild relatives as donors for economically important traits in wheat breeding programmes is the main drive to uncover the mechanism of the *Ph1* locus and regulate its activity.
• *Methods* Following the molecular genetic characterization of the *Ph1* locus, five additional deletion mutants covering the region have been identified. In addition, more bacterial artificial chromosomes (BACs) were sequenced and analysed to elucidate the complexity of this locus. A semi-quantitative RT–PCR was used to compare the expression profiles of different genes in the 5B region containing the *Ph1* locus with their homoeologues on 5A and 5D. PCR products were cloned and sequenced to identify the gene from which they were derived.
• *Key Results* Deletion mutants and expression profiling of genes in the region containing the *Ph1* locus on 5B has further restricted *Ph1* to a cluster of *cdk-like* genes. Bioinformatic analysis of the *cdk-like* genes revealed their close homology to the checkpoint kinase *Cdk2* from humans. *Cdk2* is involved in the initiation of replication and is required in early meiosis. Expression profiling has revealed that the *cdk-like* gene cluster is unique within the region analysed on 5B in that these genes are transcribed. Deletion of the *cdk-like* locus on 5B results in activation of transcription of functional *cdk-like* copies on 5A and 5D. Thus the *cdk* locus on 5B is dominant to those on 5A and 5D in determining the overall activity, which will be dependent on a complex interplay between transcription from non-functional and functional *cdk-like* genes.
• *Conclusions* The *Ph1* locus has been defined to a *cdk-like* gene cluster related to *Cdk2* in humans, a master checkpoint gene involved in the initiation of replication and required for early meiosis.

Key words: *Triticum aestivum*, *Ph1*, wheat, transcription, CDK, deletion mutants.

INTRODUCTION

Meiosis is a specialized type of cell division in which two rounds of chromosome segregation follow a single round of DNA replication. During the onset of meiosis, homologues must recognize each other and then intimately align (pairing and synapsis). In most eukaryotes this is a prerequisite for genetic recombination and balanced segregation of half-bivalents at anaphase I. Many components of the meiotic recombination machinery are known, especially in yeast, as well as some of the structural components of the synaptonemal complex. However, little is known about how homologous chromosomes recognize each other and become competent to pair. Hexaploid wheat (*Triticum aestivum*; $2n = 6x = 42$; genome AABBDD) possesses three ancestral genomes or seven sets of six related chromosomes. For hexaploid wheat to be highly fertile, only true homologues may pair within each set of the six related chromosomes during meiosis. The major locus controlling this pairing behaviour is *Ph1*, which is a single dominant locus located on chromosome 5B (Riley and Chapman, 1958). Mutants carrying a deletion of the *Ph1* locus exhibit a degree of pairing of related (homoeologous) chromosomes and hence multivalent formation at metaphase I (Sears, 1977; Roberts *et al.*, 1999). So what is the *Ph1* locus? Absence of *Ph1* activity in diploid relatives of wheat suggests that the *Ph1* locus arose on polyploidization (Chapman and Riley, 1970). Studies have shown a lack of phenotypic variation, apart from dosage effects and the failure of ethylmethane sulfonate treatment to yield mutants (Wall *et al.*, 1971). This suggests that *Ph1* is complex in structure. Recently a two-part strategy for the molecular characterization of *Ph1* was used to dissect the complex structure. The first part revealed the gene content of the wheat *Ph1* region using conservation of gene order between the highly repetitive, 16 000 Mb hexaploid wheat genome and the smaller genomes of rice and *Brachypodium sylvaticum*, thus providing markers with which to saturate the wheat region. The second part of the strategy used deletion lines to dissect the *Ph1* locus physically. By these approaches, the *Ph1* locus was localized to a 2·5 Mb region of wheat chromosome 5B containing a structure consisting of a segment of sub-telomeric

* For correspondence. E-mail graham.moore@bbsrc.ac.uk

heterochromatin that inserted into a cluster of *cdc2* (*cdk*)-related genes following polyploidization (Griffiths *et al.*, 2006). Here, further characterization of the *Ph1* region is presented exploiting new deletions and expression analysis studies.

MATERIALS AND METHODS

Plant material and growth conditions

The following hexaploid wheat (*Triticum aestivum*; $2n = 6x = 42$; genome AABBDD) lines were used: euploid Chinese Spring (CS), euploid Paragon, CS *Ph1* mutant line (*ph1b*) (Sears, 1977) and CS nullisomic/tetrasomic lines N5AT5B, N5BT5D and N5DT5A (Sears, 1966). Plants were grown in a controlled-environment growth room at 20 ºC (day) and 15 ºC (night) with 16 h of light and 85% humidity.

Mutagenesis and cytogenetical analysis

The X-ray-irradiated populations were created by X-ray mutagenesis of F_1 hybrid seed between CS and the *ph1b* deletion mutant. Batches of at least 500 seed were subjected to 5, 10, 20, 30, 50 or 100 Gy of X-rays in a Linac-3 machine at Norfolk and Norwich Hospital. Five (M_2) seed from each of the M_1 plants were germinated and DNA extracted from 6-week-old leaf material. PCR amplification was performed on the extracted DNA. The γ-irradiated populations were created by γ-mutagenesis of seeds from hexaploid wheat variety Paragon. The irradiation was performed by the International Atomic Energy Agency (IAEA), Seibersdorf, Austria. Batches of at least 500 seed were subjected to between 150 and 250 Gy of γ-irradiation. Five M_2 seeds were sampled per M_1 surviving plant. The seeds were cut in half with a razor blade and the embryo halves retained. DNA preparations were made from the individual endosperm ends. PCR amplification was performed on the extracted DNA. The embryos identified as carrying informative deletions from these PCR assays were germinated and the chromosome pairing scored at meiosis in the resulting plant and its progeny. Chromosome pairing of control and deletion lines was analysed in 30 pollen mother cells at meiotic prophase I in Feulgen-stained preparations. Wild-type pairing was confirmed by the presence of >19 ring bivalents per cell in both the M_2 and resulting M_3 plants. Line γ250-214 did not yield M_3 progeny, so its pairing phenotype could not be confirmed.

BAC library analysis

The bacterial artificial chromosome (BAC) library (TAACSPALLhA) from *T. aestivum* L. cv. CS was analysed as described by Griffiths *et al.* (2006) and Allouis *et al.* (2003). Sub-cloning from some BACs was carried out using the Invitrogen TOPO® Shotgun Subcloning Kit. DNA was extracted using QIAGEN Miniprep Kit QIAprep® Spin. Both kits were used as described by the manufacturers. The BAC sequence analysis was carried out using contiging and homology searching software (Altschul *et al.*, 1997; Ewing *et al.*, 1998) as described previously by Griffiths *et al.* (2006). Within the BAC sequence, only the genes were accurately annotated. The rest of the sequence was identified as being homologous to repetitive elements but the structure of these elements was not extensively annotated. The *cdk-like* sequences have been submitted to the EMBL database.

Analysis of expression by RT–PCR

Wheat leaf tissues and immature inflorescences from CS and the *ph1b* mutant of wheat were harvested at growth stages 6–7 (Feekes scale) and stored in liquid nitrogen. Seeds were imbibed for 24 h in water immediately prior to RNA extraction. RNA was extracted using an RNeasy plant mini kit (QIAGEN) according to the manufacturer's instructions. PCR amplification was performed on the cDNA using intron-spanning, gene-specific primer pairs or coding region primers, depending upon each gene structure. The primers for the genes were as follows: *FIM*, F-TCAACGTCTGGATCGAGCAC, R-CTCAATAAGAG TGCGGAGG; *Clk* (clv), F-AAGAAGATGTGGTCG CCGGA, R-ACGTCGCTCTTCTCGCTGAT; *pep*, F-GG ACAATTC TATGGCCTTCG, R-ATGGAG GAGCAA TGGCTTGT; *spG1*, F-AGGCAGCTTGCTCGACCATA, R-AGGCAGCTTGCT CGACCATA; *spG2*, F-GCTCTG CTATAT TTGCCTCA, R-GCTCTGCTATATTTGCCT CA; *OshypII*, F-TCAACGTCTGGATCGAGCAC, R-G CTCAATAAGAGTGCGGAGG; *KH hyp4*, F-ATGTGCT GCTGCCCGAGCAA, R-GCGAGGTTGAGGTGGCCGT A; *slp* (sub), F-CTTCAACTCCGAGTCCGGCA, R-CA GAGGAAGTGGAGGTAGTC; *zip* (zinc), F-GACCATCA GCCTGCACAAGA, R-ACGACTGCGATGACACCGAC; *sbp* (selenium), F-TCGCTGTGAAGTGTGGTTGGC, R-CTCACTCTGGCTGCAGCATC; *gtl* (chick), F-CAT CACCGACGTCTACCACA, R-CACATGCTTGCGGTT GGGCA; *h5l* (hap5), F-TCAGGTGTTCTGGGCGGAGC R-TGCCGCAATGTCGGACTTCT; *raf* (ra8), F-GACG CCATCCTCGAGCTCCT, R-GACGCCATCCTCGAG CTCCT; *mic1* (AthypVII), F-CAGCAAGCTCG CCACCTTACC, R-GTATCGTCGATGTCGCGGAGC; *Marcks*, F-GGGAATGGTGATTTGTATGC, R-GTAG CTTGCTATTGTATAGCG; *CCOM*, F-TGTGAACACG CCGACCTCGA, R-GCCGGAAGCAGCCTCTACAG; *WRKY*, F-GCAAACTGTTCGTTGCTTCAG, R-TCGG CACTGTCCTTGGCTGC; and *AthypV*, F-CGCAAG ACAAAAGCCAAAGC, R-AAGCTGCTCAATGTACAG TGTC. *Cdk-like*-specific primers were designed based upon SNP (single nucleotide polymorphism) differences identified within their DNA sequences (primers for: *cdk-like 1B*, F- ATCCACCGACTACCAGGAGAT, R-GCCGTGCTGGCCGTACGGTA; *cdk-like 2B*, F -TCCACCGCGACATCAAGACA, R-CCTTGAGGA CTTCGAAATTC; *cdk-like 3B*, F-GCTCCGCATCATCCA CGGCA, R-GCTGTTGCGTCGTCGTCCCG; *cdk-like 6B*, F-TCTCCGTGGTCATGGAGTGC, R-GCCAGAGCAA CCACGCACGTC; *cdk-like 7B*, F-ATGTCCTTCCACGA GCACCAC, R-CGAAGAGTGGCTTCCCGGAC; *cdk-like 2A*, F-GCCCCTTCGCCGGGAAGCT, R-GTGGATGA

TGTCATGATGCG; *cdk-like 1D*, F-CCGACTTCAA GGTGGACATG, R-TGCCGTGCTGGCCGTACGGA; and *cdk-like 2D*, F-CCCCTTGGCCGGCAAGCTCG, R-CGGC GAGCGGCAAGGACTT). In addition, primers were designed for the *cdk-like* genes based on the conserved domain identified across all the *cdk-like* genes (F-GACTTCAAGGTGGAC ACCAGC, R-GATGCCGGCGCCCTCGTGGGC).

Reverse transcription was carried out using reverse transcription super script RNase H (Invitrogen). PCR products were cloned using a p-GEM T easy vector kit (Promega). PCR products for sequencing were purified using MinElute™ 96 UF or Qiaquick® PCR Purification kits from QIAGEN. All kits were used as described in the manufacturer's instructions. Sequencing was carried out in the Genome Laboratory at the John Innes Centre and BAC sequencing at Washington University-Genome Sequencing Center, St Louis, MO, USA.

RESULTS

Characterization of wheat mutants carrying deletions encompassing the Ph1 region

Prior to the present study, the *Ph1* locus had been defined to a 2·5 Mb region on chromosome 5B of hexaploid wheat (Griffiths *et al.*, 2006). This region had been characterized at the molecular level and found to contain approx. 36 genes including a *cdk-like* gene cluster containing a segment of sub-telomeric heterochromatin as shown in Fig. 1. By exploiting the genomic information gained from the analysis of this region, a systematic screen of mutagenized wheat populations was undertaken to assess the efficiency of generating deletions localized within the 2·5 Mb region. Wheat seeds were treated with either 5, 10, 20, 30, 50 or 100 Gy of X-ray or 150, 200 or 250 Gy of γ-irradiation. The mutagenized populations derived from the seed were then screened using a multiplex PCR-based assay with markers derived from the *Ph1* region. The multiplex PCR assay had been previously described by Griffiths *et al.* (2006). No deletion mutants were identified from screening 500 individuals of M_2 populations irradiated with 5–150 Gy. However, two and three plants carrying a deletion in the *Ph1* region were identified after screening 500 and 900 individuals from the 200 and 250 Gy treatments, respectively. One of the five plants (γ250-19) revealed a wild-type chromosome pairing phenotype and carried the smallest deletion of approx. 500 kb in size (Fig. 1). It eliminated eight genes as being responsible for the *Ph1* phenotype (*hyp5*, *ugg1*, *cmt1*, *hsp20-1*, *at1*, *plp1*, *wdb1* and *hyp6*) (Fig. 1). Another mutant plant (γ250-214) carried a deletion of approx. 1 Mb in size which overlapped with the γ250-19 deletion (Fig. 1). This deletion encompassed a further 11 genes in this region which flanked the sub-telomeric insertion (Fig. 1). However, it was not possible to score the pairing phenotype of this plant accurately because it did not survive. The other three remaining mutants carried larger deletions than those described above, ranging from 1 to >10 Mb. This work clearly demonstrated that wheat seeds treated with 200–250 Gy of γ-irradiation can yield a range of deletion sizes with relatively high frequency. The deletion analysis still defined the *Ph1* locus to the region containing the *cdk-like* (locus) cluster and the segment of heterochromatin (Fig. 1).

Expression profiling in leaf, spike and seed of all genes in the 2·5 Mb Ph1 region of 5B and their homoeologues in the corresponding regions on 5A and 5D

Expression analysis of genes within the 2·5 Mb region containing the *Ph1* locus had previously revealed that the overall level of transcription of the genes within this

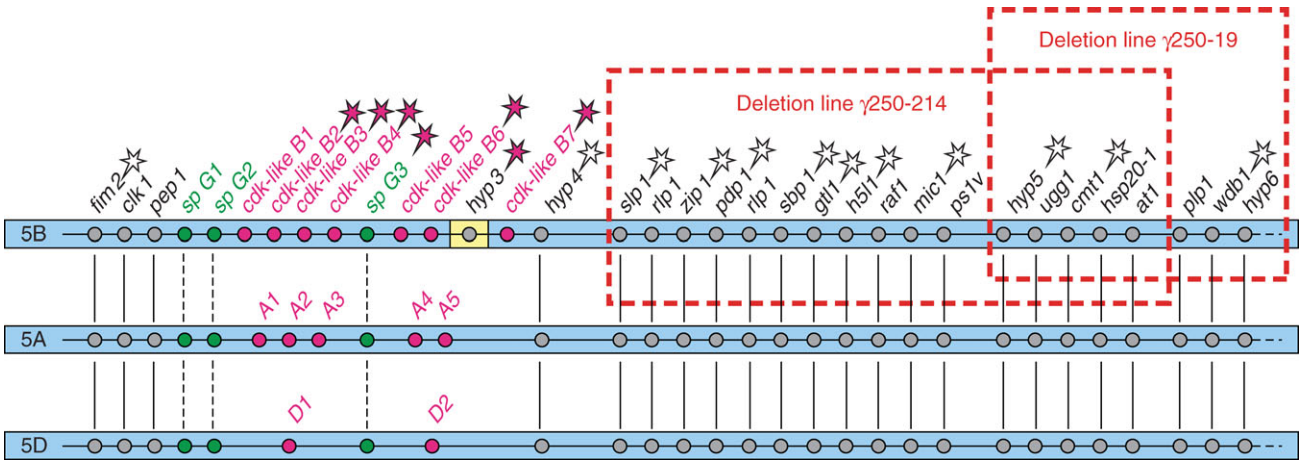

FIG. 1. Schematic diagram of the deletion mutants and annotated genes in the region containing the *Ph1* locus on chromosome 5B compared with the equivalent regions on chromosomes 5A and 5D. The three horizontal blue bars represent part of chromosome 5B and its homoeologous regions on chromosome 5A and 5D. The yellow box within the 5B chromosome bar represents the inserted sub-telomeric repeats. The two rectangles with red dotted lines represent the regions deleted in the two γ-irradiated mutant lines. Grey dots represent genes outside the *Ph1* locus. The magenta dots represent *cdk-like* genes. The green dots represent the *storage protein-like* genes (*SP*) where G1, G2 and G3 indicate that there are differences at the protein level. The stars represent the expression pattern of different genes: uncoloured stars represent genes that are mainly expressed from 5A or 5D, magenta stars represent genes expressed from 5B and genes lacking stars are not expressed. Apart from the *cdk-like* locus, the gene content presented is as described in Griffiths *et al.* (2006). A more detailed analysis of the *cdk-like* loci is provided in Fig. 4.

region showed no significant differences whether the 5B region was present or absent (Griffiths *et al.*, 2006). Recently, more genomic sequences were obtained in this region. A more detailed analysis was therefore undertaken to assess the levels of transcription of the genes on 5B, 5D and 5A (three homoeologues) in the presence of *Ph1*, and any changes in the level of transcription of the genes located on the 5A and 5D genomes when the 5B chromosome region was deleted. The sequence analysis of the coding regions of the homoeologues enabled the design of conserved primers which could be used for all three homoeologues but which would yield PCR products carrying distinct SNPs for each homoeologue (see Materials and Methods). RT–PCR was carried out on the cDNA using these conserved primers. PCR products were cloned and 24 clones were sequenced for each gene. The transcript products were then assigned to the correct homoeologue based on the SNPs within the sequence. The relative contributions made by the homoeologues were calculated as a percentage (Fig. 2). No expression was detected for genes *clk1*, *pep1*, *storage protein-like G1* (*spG1*), *rlp1*, *rdr1* and *raf1* from either wild-type wheat (CS) or the *Ph1* mutant (*ph1b*) (Fig. 1). Analysis of the 2·5 Mb region revealed that the genes accounting for most of the transcription were located in the 5A chromosomal region (Fig. 2). This explains why there was apparently no major effect on the total level of transcription when the 5B region was deleted. However, there was a region where the expression was derived mainly from the 5B genes, namely the region containing the *cdk-like* and the *spG1* clusters and *hyp3* (Fig. 1). The *Ph1* phenotype is specific to 5B and the *Ph1* locus was therefore defined to this region containing specific expression from the 5B genes. In fact there was very little transcription from the genes on 5B flanking the *Ph1* region, with *mic1* being the only other gene which was transcribed to a significant level in the 2·5 Mb chromosomal region. Of the genes transcribed from 5B within the *Ph1* region, *hyp3* was not found in the equivalent region of the related grass *Triticum timopheevi* which has *Ph1* activity, and *hyp1* and *hyp2* (Griffiths *et al.*, 2006) are now known to encode storage-like proteins in wheat (Kawaura *et al.*, 2005). Interestingly, when the 5B region was deleted, the loss of transcription of these genes was compensated by the homoeologues on the remaining genomes (5A and 5D) so that there was no apparent difference in the overall level of transcription in the presence or absence of *Ph1* (Figs 2 and 3). In some cases, the compensation was provided by homoeologous genes lying on other chromosomes such as observed for transcription in both spike and seeds for *hyp3* and in leaf for *Slp1* genes. Thus, only the *cdk-like* gene cluster remained for more detailed analysis with respect to the *Ph1* function.

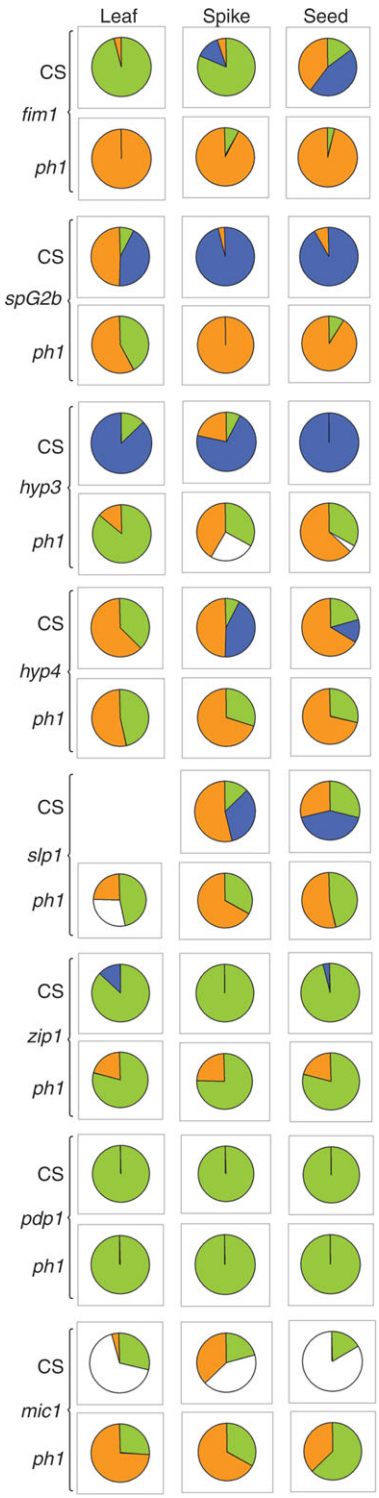

FIG. 2. Expression pattern of genes within the 2·5 Mb region covering the *Ph1* locus on 5B and equivalent regions on 5A and 5D. Total RNA was extracted from leaf, spike and seeds of Chinese Spring and *Ph1* mutant (*ph1b*). For each gene, RT–PCR products were cloned and 24 clones were sequenced. SNPs enabled a semi-quantification of the expression from the genes on chromosomes B, A and D. Results are presented in the form of pie charts where the sizes of the blue, green and orange sectors correspond to the proportion of transcription contributed by genes on chromosomes 5B, A and D, respectively. A white sector indicates that the sequenced transcript did not correspond to any of these genes.

Structural complexity of the cdk cluster

Before undertaking a detailed analysis of the expression of the *cdk-like* genes in euploid wheat (CS) in comparison with the *Ph1* mutant (*ph1b*), the CS BAC library was screened with specific probes for *cdk-like* and *storage protein-like* genes. DNA was prepared from >60 BACs

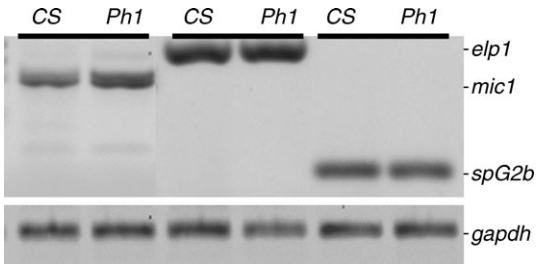

FIG. 3. Transcription of specific genes in the presence or absence of the *Ph1* locus. Semi-quantitated RT–PCR products were amplified from total RNA extracted from the spikes of Chinese Spring (CS) and the *Ph1* mutant. Primers were designed in the conserved regions of *elp1*, *mic1* and the *storage protein-like* gene *spG3b*. The glyceraldehyde-3-phosphate dehydrogenase (*gapdh*) gene was used as a semi-quantitative positive control.

identified. Southern blot analyses were carried out, with blots being probed with the conserved regions of the *cdk-like* and the *storage protein-like* genes. Primers were designed in the most conserved domains of all known *cdk-like* genes already identified in the *Ph1* region and corresponding regions in 5A and 5D. These primers were used to amplify DNA from euploid wheat, *ph1b*, CS nullisomic/tetrasomic lines N5AT5B, N5BT5D and N5DT5A, and all the BACs identified as containing *cdk-like* genes. The PCR products were then cloned and sequenced. Sequence analysis revealed that the *cdk* loci on 5B and 5A were more complex than initially thought. There were higher numbers of both *cdk-like* and *storage protein-like* genes within this region. BAC and genomic analyses showed that seven, five and two *cdk-like* genes were present in the corresponding regions of 5B, 5A and 5D chromosomes, respectively (Figs 1 and 4). The number of *storage protein-like* genes was also more than initially predicted, and at least four genes on each chromosome were identified (Figs 1 and 4). The main finding was that there had been gene duplication of both *cdk-like* and *storage protein-like* genes (Figs 1 and 4). This made BAC analysis more complex because of similar size fragments containing *cdk-like* or *storage protein-like* genes, requiring them to be distinguished by sequencing. Interestingly, the gene duplication differed between 5B, 5A and 5D, indicating that these clusters arose at different times.

Sequence analysis of cdk-like genes in the clusters on 5B, 5A and 5D

From the BAC and genomic analyses, the complete sequence of the seven *cdk-like* genes on 5B, five *cdk-like* genes on 5A and the two *cdk-like* genes on 5D was available (Fig. 5). Within the 5B and 5A *cdk-like* loci, there were also fragments of *storage-like protein* genes present. The sequence analysis of the three *cdk-like* genes (*A3*, *A4* and *A5*) within 5A BAC 1898B19 revealed that in contrast to the flanking regions, the sequence of the three *cdk-like* genes differed every time these *cdk-like* genes were sequenced either from sub-clones, PCR products or the BAC itself (data not shown). Therefore, although these 5A *cdk-like* genes apparently carry frameshifts and stop codons, these may reflect sequencing errors. Moreover, the translation of the sequence of the *cdk-like* genes on 5A suggests that they differ at the protein level in comparison with other *cdk-like* genes in 5B and 5D loci. Again this may reflect the problems in sequencing these specific genes (Figs 4 and 5). The sub-telomeric heterochromatin segment inserted between the two *cdk-like* genes, *cdk-like B6* and *B7* (Fig. 4). *Cdk-like B7* was found to be a pseudo-gene at the protein level and *cdk-like B6* was fused in-frame with a small non-autonomous CACTA transposon (Fig. 6). DNA and protein sequence analyses showed that *cdk-like D1* and *D2* are functional genes and are related to *cdk-like B4* (Figs 5 and 6). The *cdk-like* genes are related to the protein kinase gene *Cdk2* from humans (Fig. 7; Marston and Amon 2004). Of the 15 invariant amino acids found in the conserved sub-domains of these protein kinases, 11 of these amino acids were found in the products of *Ph1*'s *cdk-like* genes (Fig. 7).

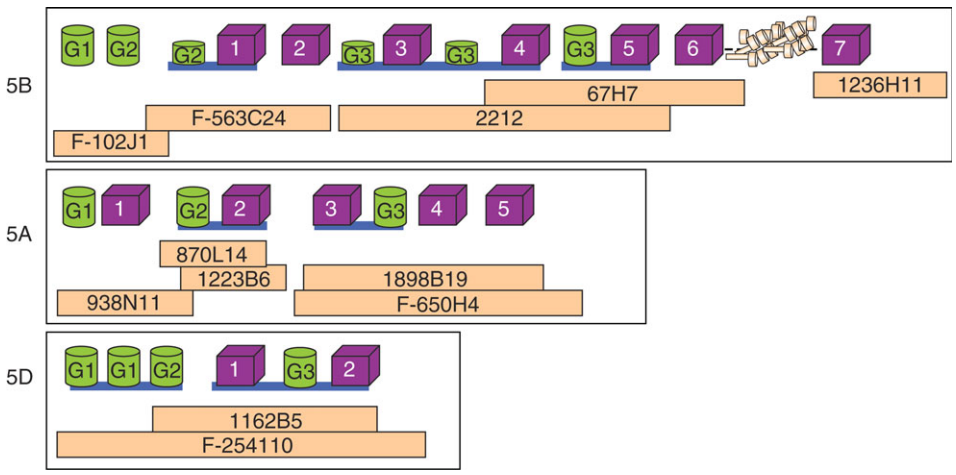

FIG. 4. Detailed schematic diagram of the *cdk* cluster on chromosome 5B and homoeologous regions on chromosomes 5A and 5D. The homoeologous regions of 5B, 5A and 5D are defined in three white boxes. Pale pink blocks represent the BACs. Blocks in magenta represent the *cdk-like* genes, large green cylinders represent *storage protein-like* genes (*sp*) where G1, G2 and G3 indicate that there are differences at the protein level; smaller green cylinders represent fragments of *storage protein-like* genes. The solid blue horizontal line indicates that the genes are found on the same BAC sequence contig. Sub-telomeric heterochromatin is shown inserted between *cdk-like* genes 6 and 7 on 5B.

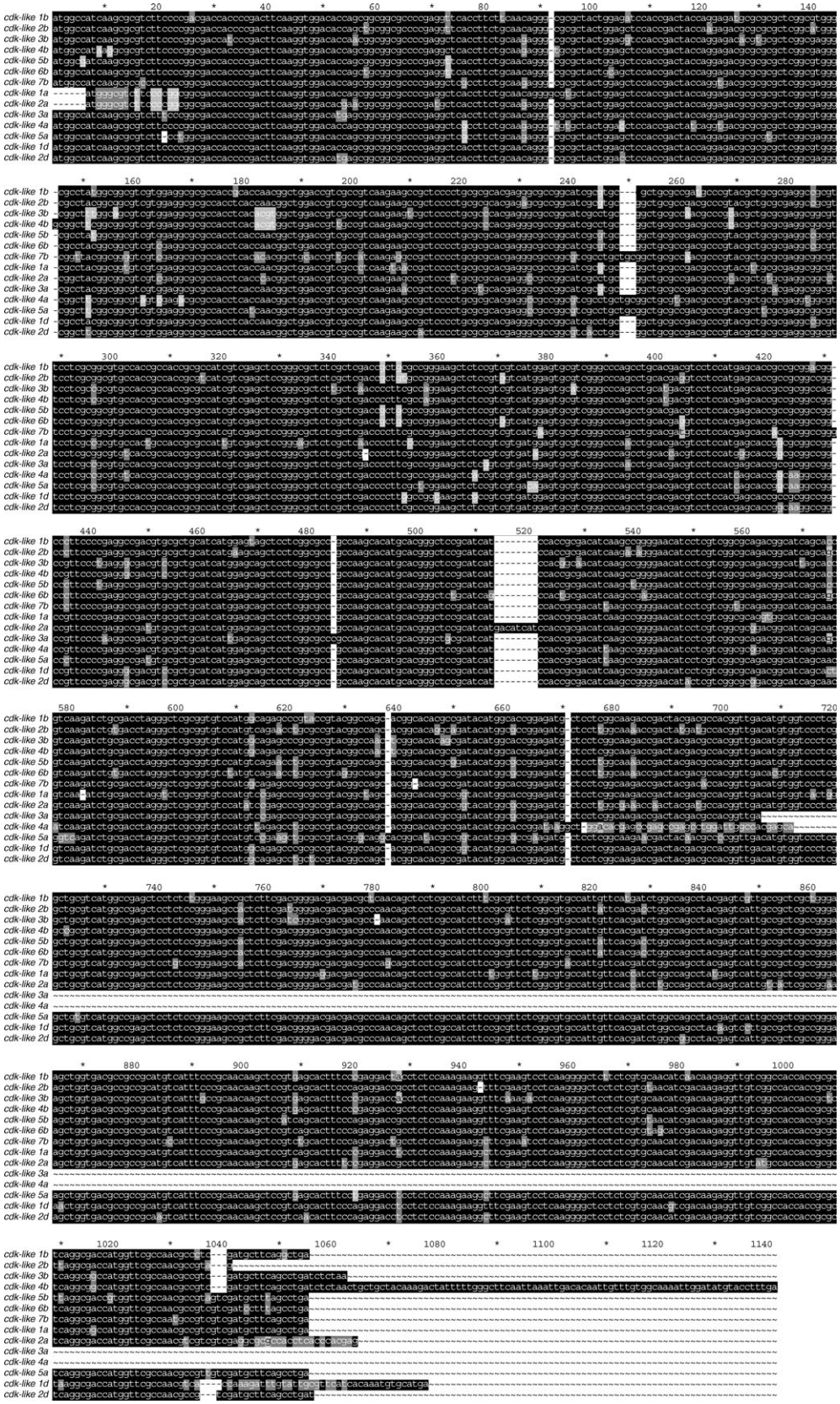

FIG. 5. Sequence alignment of *cdk-like* genes on 5B, 5A and 5D. The sequences are in order from top to bottom: *cdk-like B1–B7*, *cdk-like A1–A5* and *cdk-like D1* and *D2*. The CACTA transposon has been removed from the sequence of *cdk-like B7* for this comparison. As explained within the text, 5A *cdk-like* gene sequences may contain errors.

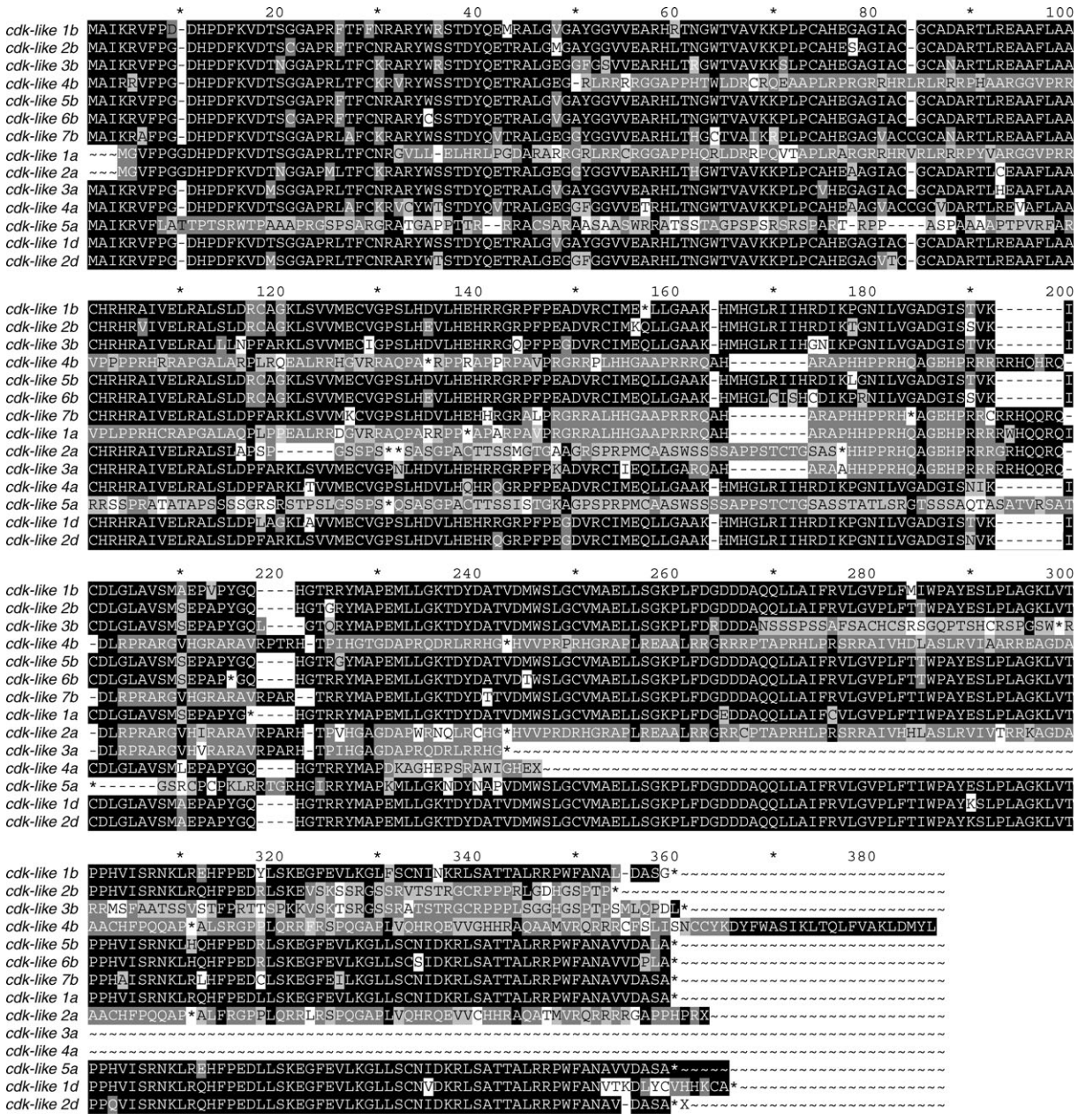

FIG. 6. Protein alignment of *cdk-like* genes on 5B, 5A and 5D. The protein sequences are in order from top to bottom: *cdk-like B1–B7*, *cdk-like A1–5* and *cdk-like D1* and *D2*.

Transcription from the cdk-like gene clusters in the 5B Ph1 and 5A and 5D equivalent regions

Conserved primers were designed in the N-terminal part of the coding region of the *cdk-like* genes. The use of conserved primers provided the opportunity to semi-quantitate the contribution each gene in the cluster made to the total level of transcription, by sequencing RT–PCR products and then assigning the product to a particular gene based on SNP detection. Different sets of conserved primers were designed and tested on cDNA derived from spike RNA of euploid and *Ph1* mutants as well as from genomic DNA of these plants. The PCR products (200 bp) were then cloned and a minimum of 300 clones sequenced. Results were inconclusive in terms of quantification because, in addition to the known *cdk-like* genes on 5B and 5D chromosomes, there were other unknown variants being generated in both euploid and *Ph1* mutant samples. Within the euploid sample, 56 % of the clones were derived from 5B *cdk-like* genes, 24 % from 5D *cdk-like* genes and 20 % were unknown *cdk-like* genes, while within the *Ph1* mutant sample, 68 % were unknown *cdk-like* genes and 36 % derived from 5D *cdk-like* genes. These unknown *cdk-like* variants were mostly absent in the PCR products generated by the conserved primers of

```
                *         20          *         40          *         60          *         80          *        100          *
CDK2: ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~MENFQKVEKIGEGTYGVV--YKARNKLTGE---VVALKK-IRXDTETEGV-----PSTAIREISLLKELN----H
D2  : MAIKRVFPGDHPDFKVDTSGGAPRLTFCNRARYWSSTDYQETRALGVGAYGGV--VEARHLTNG---WTVAVKKPLPCAHEGAGIACGCADARTLREAAFLA--ACHR-H

           120          *        140          *        160          *        180          *        200          *        220
CDK2: PNIVKL--L------DVIHTE-NK--------LYLVFEFLHQDLKKFMDASALTGIPLPL----IKSYLFQLLQGLAFCHSHRVLHRDLKPQNLLINTEG--AIK--LAD
D2  : RAIVELRALSLDPLA--------------G-KLAVVHECVGPSLHDVLHEHR-RGRP--FPEGDVRCIMEQLLGAAKHMHGLRIIHRDIKPGNILVGADGISTVK--ICD

                *        240          *        260          *        280          *        300          *        320          *
CDK2: FGLARAF--GVP-VR-------T--YTHEVVTLWYRAPEILLG-CKYYSTAVDIWSLGCIFAEMVTRRALFPGDSEIDQLFRIFRTLGTPDEVVWPGVTSMP-DYKPS--
D2  : LGLAVSMAEPAPYGQHG--------------TRRYMAPEMLLGK-TDYDATVDMWSLGCVMAELLSGKPLFDGDDDAQQLLAIFRVLGVPLFTIWPAYKSLP--LAGKLV

           340          *        360          *        380          *        400
CDK2: FPKW--ARQDFSKVVPP-----LD--------EDGRSLLSQMLHYDPNKRISAKAALAHPFFQDVTKPVPHLRL~~~~
D2  : TPPHVISRNKLRQHFP---EDLLS--------KEGFEVLKGLLSCNVDKRLSATTALRRPWFANVTKDLYCVHHKCA*
```

FIG. 7. Protein alignment of *cdk-like* D2 with kinase *Cdk2* from humans. Protein alignment of the meiotic checkpoint gene in humans (Debondt *et al.*, 1993) with the *cdk-like* protein *D2* gene. The shaded areas represent the overall homology to these two known genes. Eleven of 15 functional domains are conserved between kinases (underlined).

DNA derived from the nullisomic/tetrasomic line N5AT5B. This suggests that the *cdk-like* unknown variants are being derived from the 5A *cdk-like* genes (unpublished). As described previously, the BAC sequence analysis indicated that variants were being generated from products derived from three *cdk-like* genes (*A3*, *A4* and *A5*). Despite this issue, the expression analysis did indicate that *cdk-like B2*, *B3* and *B7* were being transcribed from the 5B region. To overcome these problems, specific primers were designed from the sequence of the *cdk-like* genes on chromosomes 5B (*cdk-like B4*, *B6* and *B7*), 5A (*cdk-like A2* and *A4*) and 5D (*cdk-like D1* and *D2*). The expected PCR products were between 250 and 500 bp (Fig. 8). The specific primers were first tested on genomic DNA extracted from the nullisomic/tetrasomic lines N5AT5B, N5BT5D and N5DT5A. All primers were found to be specific at the chromosome level. Primers were then used to study the expression of these genes using RT–PCR on spike tissues from both the CS and *Ph1* mutant plants. The PCR products were cloned and 24 clones were sequenced to confirm that the products were in fact derived from the *cdk-like* gene for which the specific primers were designed. The *cdk-like A5* was found to be expressed in both CS and the *Ph1* mutant, whereas *cdk-like D2* was only expressed in the *Ph1* mutant (Fig. 8). Interestingly, *cdk-like B6* and *cdk-like B7* (the CACTA insertion and pseudo-gene, respectively) (Fig. 6) were expressed in CS (Fig. 8). Thus, five *cdk-like* genes were expressed from the 5B locus, namely *cdk-like B2*, *B3*, *B4*, *B6* and *B7*, and one gene from the 5A locus, namely *A5*, while *cdk-like D2* and *cdk-like A* variants were mainly expressed in the absence of the 5B region. Thus, as in the case of the *storage protein-like* gene (*spG1*) and *mic1* (Fig. 2), the transcription from the homoeologues on the other genomes is being suppressed by the presence of the 5B copies. When the 5B copy is deleted, the transcription is activated on the other genomes to compensate.

DISCUSSION

Previously, the *Ph1* locus had been defined to a 2·5 Mb region containing a *cdk-like* gene cluster within which a segment of sub-telomeric heterochromatin had inserted following wheat's polyploidization (Griffiths *et al.*, 2006). Here the identification of mutants carrying further deletions encompassing the 2·5 Mb *Ph1* region has been described. In particular, it has been shown that it is possible to generate deletions of 500 kb upwards in hexaploid wheat with reasonable frequency after γ-irradiation. The ability to generate deletions with such a high frequency enables a complex locus such as *Ph1* to be defined. This approach could make it easier to fine map physically those genes controlling agronomic traits lying in regions of reduced

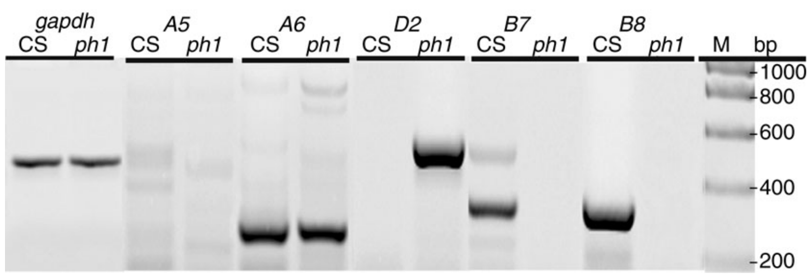

FIG. 8. Expression pattern of *cdk-like* genes from spikes of wild-type wheat (Chinese Spring; CS) and the *Ph1* mutant. Total RNA was extracted from spikes of CS and the *Ph1* mutant and used in a semi-quantitative RT–PCR amplification. All PCR components in the master-mix remained constant except for the primers. The primers were designed specifically for each *cdk-like* gene in the three homoeologous regions from B, A and D genomes. The glyceraldehyde-3-phosphate dehydrogenase (*gapdh*) gene was used as a semi-quantitative positive control. A DNA marker from Invitrogen was used to size the PCR product.

recombination. The characterization of the new deletions covering the 2·5 Mb region presented here is consistent with the original proposal that defined the *Ph1* locus to a region containing a *cdk-like* gene cluster. The original analysis indicated that this region of 5B contained a segment of sub-telomeric heterochromatin inserted within this *cdk-like* cluster, making it distinct from the corresponding *cdk-like* gene clusters on 5A and 5D. The more detailed BAC library analysis in the present study, covering the *cdk* loci on 5B, 5A and 5D, reveals not only that the region is distinct in possessing a segment of heterochromatin, but also that the *cdk-like* cluster on 5B is itself unique compared with the equivalent clusters on 5A and 5D. There are in fact seven *cdk-like* genes on 5B compared with at least five on 5A and two on 5D. The sequence analysis data suggest that these structures arose through tandem duplication events. The *Ph1* phenotype is specific to chromosome 5B. The expression profiling reveals that the region defined as containing the *Ph1* locus also contains the only cluster of genes expressed specifically from 5B in the whole region analysed. The genes in the flanking regions are mostly transcribed from genes located on 5A. Five of the *cdk-like* genes are transcribed, including unusually the two pseudogenes *cdk-like B6* and *cdk-like B7*. None of the other pseudogenes identified are expressed. It is possible that the transcription of these two pseudogenes may therefore be connected to the presence of the sub-telomeric heterochromatin segment which inserted between them. The transcripts generated from these two pseudogenes could be non-coding and may be involved with the functional *cdk-like* transcripts in a small RNA-based interference mechanism. Hexaploid wheat may possess a mechanism for regulating and sensing the total level of transcripts in the cell which can be derived from the homoeologues. In cases where most of the transcription is derived from the genes located on 5B (*cdk-like* genes, *storage protein-like* gene clusters and *mic1*), the deletion of these genes does not necessarily affect the total level of transcription as the genes located on 5A or 5D can compensate for the loss of transcription from the 5B copies. However, the expression studies do indicate that the *cdk* locus on 5B is dominant in terms of transcription control over corresponding *cdk* loci on 5A and 5D. The increase in expression from the 5A or 5D *cdk-like* loci may explain the *Ph1* effects observed on chromosome synapsis (Holm, 1986, 1988). During early meiosis, chromosomes can engage in multiple associations during synapsis. Later in meiosis, these associations are resolved in the presence of *Ph1*, while many associations are still maintained in the absence of *Ph1*. Disruption of chromosome pairing usually causes a lengthening in the time taken to complete meiotic prophase I. Yet despite the disruption to chromosome pairing, meiotic prophase I is not longer in the absence of *Ph1* than in its presence (Bennett *et al.*, 1974). This implies that in the absence of *Ph1*, something is induced which over-rides the checkpoint mechanism and speeds up the completion of meiotic prophase I. The induction of expression of the 5A *cdk-like* genes may be responsible for this phenotype and the consequential inability to resolve incorrect associations within the timeframe available. Equally, it has been observed that increasing the dosage of the 5B *Ph1* locus results in disruption of chromosome pairing (Feldman, 1966). Increasing the dosage of the 5B *cdk-like* locus may result in oversuppression of the *cdk-like* loci on 5A and 5D, resulting in aberrant chromosome pairing.

Finally, detailed bioinformatics has revealed that the kinase *Cdk2* from humans shows the closest homology to the genes within the wheat *cdk* locus. Interestingly, this kinase is required for meiosis and has a number of phosphorylation targets (Marston and Amon, 2004; Cohen *et al.*, 2006). It co-localizes with mismatch repair proteins to recombination nodules and to the telomere regions during early meiosis (Ashley *et al.*, 2000). Disruption of *Cdk2* expression in mouse has also been shown to result in increased 'non-homologous' synapsis of chromosomes at meiosis (Cohen *et al.*, 2006). An increased presence of multiple associations of chromosomes is also observed at synapsis in the *Ph1* mutant (Holm, 1988). *Cdk2* has been proposed to be involved in licensing origins of replication through histone phosphorylation and chromatin remodelling (Alexandrow and Hamlin, 2005). The control of such biological processes during pre-meiotic replication and at the start of meiosis would also be entirely consistent with some of the cell biology and cytogenetic observations associated with *Ph1*, namely the initiation and co-ordination of chromatin remodelling during the onset of meiosis (Aragon *et al.*, 1997; Maestra *et al.*, 2001; Prieto *et al.*, 2004, 2005). In hexaploid wheat, prior to meiosis, the centromeres pair and then form seven clusters corresponding to the seven sets of related chromosomes. At the onset of meiosis, the telomeres cluster to form a telomere bouquet (Martinez-Perez *et al.*, 2003). Within this telomere bouquet, the telomere regions recognize each other and initiate intimate alignment (Prieto *et al.*, 2004). This alignment then propagates along the homologous chromosomes from the telomere regions. At the onset of meiosis, the chromatin of chromosomes is remodelled to enable the homologues to pair (Prieto *et al.*, 2004). *Ph1* is involved in the initiation of this remodelling, and in co-ordinating the remodelling of chromatin on both homologues so that they are in the same conformation at the onset of pairing. In the absence of *Ph1*, the chromatin remodelling can be initiated asynchronously and prematurely so that homologues are in different conformational states (Prieto *et al.*, 2004). Thus, a homologue would be just as likely to align intimately with a related chromosome as with its true homologue. The *Ph1* locus in wheat is also able to block recombination from occurring between similar but distinct chromosome segments located within otherwise identical chromosomes (Dubcovsky *et al.*, 1995; Luo *et al.*, 1996). If chromatin remodelling is required for chromosome pairing and recombination, then the lack of recombination between these regions in the presence of *Ph1* may simply reflect a failure to remodel the regions at the onset of meiosis. Thus, a *cdk-like* gene complex related to *Cdk2* which partially suppresses the activity of corresponding loci on 5A and 5D could explain the multitude of different effects attributed to *Ph1*.

ACKNOWLEDGMENTS

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK. We would like to thank Dr John Doonan, Cell and Developmental Biology Department at the John Innes Centre, for his helpful discussion on *cdk* gene analysis, and the International Atomic Energy Agency, Vienna, Austria for irradiating seeds to create the new *Ph1* mutants.

LITERATURE CITED

Alexandrow MG, Hamlin JL. 2005. Chromatin decondensation in S-phase involves recruitment of Cdk2 by Cdc45 and histone H1 phosporylation. *Journal of Cell Biology* **168**: 875–886.

Allouis S, Moore G, Bellec A, Sharp R, Faivre Rampant P, Mortimer K, *et al*. 2003. Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm 'Chinese Spring'. *Cereal Research Communications* **31**: 331–338.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389–3402.

Aragon-Alcaide L, Miller T, Schwarzacher T, Reader S, Moore G. 1997. Centromere behaviour in wheat with high and low homoeologous chromosomal pairing. *Chromosoma* **106**: 327–333.

Ashley T, Walpita D, de Rooij DG. 2001. Localisation of two mammalian cyclin dependent kinases during mammalian meiosis. *Journal of Cell Science* **114**: 685–693.

Bennett MD, Dover GA, Riley R. 1974. Meiotic duration in wheat genotypes with and without homoeologous meiotic chromosome pairing. *Proceedings of the Royal Society B: Biological Sciences* **187**: 191–207.

Chapman V, Riley R. 1970. Homoeologous meiotic chromosome pairing in *Triticum aestivum* in which chromosome 5B is replaced by an alien homoeologue. *Nature* **226**: 376–377.

Cohen PE, Pollack SE, Pollard JW. 2006. Genetic analysis of chromosome pairing, recombination and cell cycle control during first meiotic prophase in mammals. *Endocrine Reviews* **27**: 398–426.

DeBondt HL, Rosenblatt J, Jancarik J., Jones HD, Morgan DO, Kim SH. 1993. Crystal structure of cyclin-dependent kinase 2. *Nature* **363**: 595–602.

Dubcovsky J, Luo MC, Dvorak J. 1995. Differentiation between homoeologous chromosome 1A of wheat and 1Am of *Triticum monococcum* and its recognition by the wheat *Ph1* locus. *Proceedings of the National Academy of Sciences, USA* **92**: 6645–6649.

Ewing B, Hillier L, Wendl M, Green P. 1998. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8**: 175–185.

Feldman M. 1966. The effect of chromosomes 5B, 5D, and 5A on chromosomal pairing in *Triticum aestivum*. *Proceedings of the National Academy of Sciences, USA* **55**: 1447–1453.

Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G. 2006. Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* **439**: 749–752.

Holm PB. 1986. Chromosome pairing and chiasma formation in allohexaploid wheat, *Triticum aestivum*, analysed by spreading of meiotic nuclei. *Carlsberg Research Communications* **51**: 239–294.

Holm PB. 1988. Chromosome pairing and synaptonemal complex formation in hexaploid wheat nullisomic for chromosome 5B. *Carlsberg Research Communications* **53**: 91–110.

Kawaura K, Mochida K, Ogihara Y. 2005. Expression profile of two storage-protein gene families in hexaploid wheat revealed by large-scale analysis of expressed sequence tags. *Plant Physiology* **139**: 1870–1880.

Luo MC, Dubcovsky J, Dvorak J. 1996. Recognition of homeology by the wheat *Ph1* locus. *Genetics* **144**: 1195–1203.

Marston AL, Amon A. 2004. Meiosis: cell-cycle controls shuffle and deal. *Nature Reviews Molecular Cell Biology* **5**: 983–997.

Martinez-Perez, Shaw P, Aragon-Alcaide L, Moore G. 2003. Chromosomes form into seven groups in hexaploid and tetraploid wheat as a prelude to meiosis. *Plant Journal* **36**: 21–29.

Maestra B, Hans de Jong J, Shepherd K, Naranjo T. 2002. Chromosome arrangement and behaviour of the rye homologous telosomes at the onset of meiosis in disomic wheat 5RL addition lines with and without the *Ph1* locus. *Chromosome Research* **10**: 655–667.

Riley R, Chapman V. 1958. Genetic control of cytologically diploid behaviour of hexaploid wheat. *Nature* **182**: 713–715.

Prieto P, Shaw P, Moore G. 2004. Homologue recognition during meiosis is associated with a change in chromatin conformation. *Nature Cell Biology* **6**: 906–908.

Prieto P, Moore G, Reader S. 2005. Control of conformational changes associated with homologue recognition during meiosis. *Theoretical and Applied Genetics* **111**: 505–510.

Roberts MA, Reader SM, Dalgliesh C, Miller TE, Foote TN, Fish LJ, *et al*. 1999. Induction and characterisation of *Ph1* wheat mutants. *Genetics* **153**: 1909–1918.

Sears ER. 1966. Nullisomic–tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR eds. *Chromosome manipulations and plant genetics*. Edinburgh: Oliver and Boyd, 22–45.

Sears ER. 1977. An induced mutant with homoeologous pairing in common wheat. *Canadian Journal of Genetics and Cytology* **19**: 585–593.

Wall AM, Riley R, Gale MD. 1971. The position of a locus on chromosome 5B of *Triticum aestivum* affecting homoeologous meiotic pairing. *Genetical Research Cambridge* **18**: 329–339.