

1       **Exploiting the *ZIP4* homologue within the wheat *Ph1* locus has identified two**  
2       **lines exhibiting homoeologous crossover in wheat-wild relative hybrids**

3

4       María-Dolores Rey<sup>1, 2</sup>, Azahara C Martín<sup>1, 2</sup>, Janet Higgins<sup>3</sup>, David Swarbreck<sup>3</sup>,  
5       Cristobal Uauy<sup>2</sup>, Peter Shaw<sup>2</sup> & Graham Moore<sup>2\*</sup>

6       <sup>1</sup>Joint first author

7       <sup>2</sup>John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

8       <sup>3</sup>Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK

9

10      \***Corresponding author:** Graham Moore (email: [graham.moore@jic.ac.uk](mailto:graham.moore@jic.ac.uk))

11      <http://orcid.org/0000-0002-0205-7436>

12

13      Total word count: 5207

14      Title: 20

15      Summary: 167

16      Introduction: 956

17      Materials and Methods: 1187

18      Results and discussion: 1571

19      Number of figures: 3

20      Supplementary figures: 3

21      Supplementary tables: 2

22

23      **Acknowledgements**

24      This work was supported by the UK Biotechnology and Biological Research Council

25      (BBSRC), through three grants (Grant BB/J004588/1; Grant BB/M009599/1; Grant

26 BB/J007188/1); and by a Marie Curie Fellowship Grant (H2020-MSCA-IF-2015-  
27 703117).

28

29 **Key message**

30 Exploiting the *ZIP4* homologue within the wheat *Phl* locus has identified two wheat  
31 mutants through a non-GM route, which can be exploited as an alternative to the  
32 Chinese Spring *ph1b* mutant in wheat introgression strategies.

33

34 **Abstract**

35 Despite possessing related ancestral genomes, hexaploid wheat behaves as a diploid  
36 during meiosis. The wheat *Phl* locus promotes accurate synapsis and crossover of  
37 homologous chromosomes. Interspecific hybrids between wheat and wild relatives are  
38 exploited by breeders to introgress important traits from wild relatives into wheat,  
39 although in hybrids between hexaploid wheat and wild relatives, which possess only  
40 homoeologues, crossovers do not take place during meiosis at metaphase I. However,  
41 in hybrids between *Phl* deletion mutants and wild relatives, crossovers do take place.  
42 A single *Phl* deletion (*ph1b*) mutant has been exploited for the last 40 years for this  
43 activity. We show here that selection of chemical induced mutant lines possessing  
44 mutations in *TaZIP4-B2* exhibit high levels of homoeologous crossovers when  
45 crossed with a wild relative. Exploitation of *Tazip4-B2* mutants rather than mutants  
46 with whole *Phl* locus deletions may improve introgression of wild relative  
47 chromosome segments into wheat. Such mutant lines may be more stable over  
48 multiple generations, as multivalents causing accumulation of chromosome  
49 translocations are less frequent.

50

51 **Keywords:** wheat, *Ph1*, *ZIP4*, homoeologues, synapsis, crossover

52 **Introduction**

53 Hexaploid bread wheat (*Triticum aestivum*) is composed of three related ancestral  
54 genomes (A, B and D), each containing 7 identical (homologous) chromosome pairs  
55 (homologue pairs 1-7). Each homologue pair has a corresponding related  
56 homoeologous chromosome pair, possessing similar gene order and content, within  
57 each of the other two genomes. Despite the similarity between homoeologues, wheat  
58 behaves as a diploid at meiosis, with synapsis and crossovers (COs) only occurring  
59 between homologous chromosomes, rather than between homoeologous  
60 chromosomes (for example 1A only pairs with 1A but not with either 1B or 1D,  
61 holding true for all seven chromosome groups). This diploid-like behaviour is  
62 predominantly controlled by *Ph1*, a dominant locus on chromosome 5B (Riley and  
63 Chapman 1958; Sears and Okamoto 1958). Sexual hybridisation between wheat and  
64 wild relatives (for example rye or *Aegilops variabilis*) produces interspecific hybrids,  
65 containing haploid sets of wheat and wild relative homoeologous chromosomes, but  
66 exhibiting virtually no COs during meiosis. However, in *Ph1* deleted wheat-rye  
67 hybrids, an average of 7 COs per cell is observed (Sears 1977; Dhaliwal et al. 1977).  
68 Such interspecific hybrids have been used to introgress important traits from wild  
69 relatives into wheat.

70

71 Synapsis is a process early in meiosis by which homologues intimately align with  
72 each other, forming bivalents held together by the proteinaceous structure or  
73 synaptonemal complex (SC). Ultimately the SC is degraded, so that the bivalents are  
74 then held together by chiasmata or COs at metaphase I, allowing their correct  
75 segregation. Recently, *Ph1* has been shown to have a dual effect on synapsis and CO

76 formation in wheat (Martín et al. 2014; Martín et al. 2017). The effect on synapsis  
77 occurs during the telomere bouquet stage, when *Phl* promotes more efficient  
78 homologous synapsis, thereby reducing the chance of homoeologous synapsis (Martín  
79 et al. 2017). The effect on CO formation happens later in meiosis, when *Phl* prevents  
80 MLH1 sites (Double Holliday Junctions marked to become COs) on synapsed  
81 homoeologues from becoming COs. The *Phl* locus was defined as a deletion  
82 phenotype, first described by scoring meiosis in wheat hybrids lacking the whole 5B  
83 chromosome (Riley and Chapman 1958; Sears and Okamoto 1958). Exploitation of  
84 smaller chromosome 5B deletions, later characterised and defined the *Phl* locus to a  
85 region on chromosome 5B, containing a duplicated chromosome 3B segment carrying  
86 heterochromatin and *TaZIP4-B2* (originally termed *Hyp3*, UniProtKB - Q2L3T5),  
87 inserted into a cluster of *CDK2-like* genes interspersed with methyl transferase genes  
88 (originally termed *SpG*, UniProtKB- Q2L3W3) (Griffiths et al. 2006; Al-Kaff et al.  
89 2008; Martín et al. 2017). Ethylmethane sulphonate (EMS) treatment failed to  
90 produce mutants exhibiting the full *Phl* deletion phenotype (Griffiths et al. 2006).  
91 This observation, combined with the dual activity of the *Phl* locus on synapsis and  
92 COs, suggested that the *Phl* phenotype was likely to result from the activity of more  
93 than a single gene (Griffiths et al. 2006).

94

95 A single *Phl* deletion mutant, developed in hexaploid wheat variety Chinese Spring  
96 (CS) (CS *ph1b*), has been used by breeding programmes worldwide to introgress wild  
97 relative chromosome segments into wheat. However, the CS *ph1b* mutant is reported  
98 to accumulate extensive rearrangements reducing fertility (Sánchez-Morán et al.  
99 2001), due to homoeologous synapsis and COs as visualised by the occurrence of  
100 multivalents at metaphase I during meiosis. It would therefore be useful to identify

101 novel wheat *Ph1* mutant lines, with reduced homoeologous synapsis and CO at  
102 meiosis, but which do exhibit homoeologous COs in hybrids with wild-relatives. As  
103 described previously, the *Ph1* locus, which affects both synapsis and CO, is a  
104 complex cluster of *CDK2*-like and methyl transferase genes containing a *ZIP4*  
105 paralogue. It has been proposed that *Ph1*'s effect on synapsis is connected to altered  
106 Histone H1 *CDK2*-dependent phosphorylation in the presence and absence of *Ph1*.  
107 Altered phosphorylation affects chromatin structure and delays premeiotic replication,  
108 subsequently affecting homologue synapsis, thus allowing homoeologous synapsis to  
109 take place (Greer et al. 2012; Martín et al. 2017). Lines carrying mutations in the *Ph1*  
110 *CDK2*-like homologue in Arabidopsis, also exhibit reduced synapsis under specific  
111 conditions, suggesting a role for these genes in efficient synapsis (Zheng et al. 2014).  
112 We have also previously proposed that the effect of *CDK2*-like genes on chromatin  
113 structure not only affects synapsis, but might also affect the resolution of Double  
114 Holliday Junctions (marked by MLH1) as COs (Greer et al. 2012). Okadaic acid  
115 treatment affects chromatin structure and can induce homoeologous CO in wheat-wild  
116 relative hybrids (Knight et al. 2008). However, given the locus contains multiple  
117 copies of the *CDK2*-like and methyltransferase genes, it would be a complex and  
118 laborious study to identify EMS mutants within these genes and combine possible  
119 mutations. The transfer of multiple mutated genes into different elite genetic  
120 backgrounds by breeders, for subsequent crossing with wild relatives, would also be  
121 laborious. Moreover, as previously indicated, we want to identify wheat mutant lines  
122 which exhibit reduced homoeologous synapsis or multivalents at metaphase I, so the  
123 *CDK2*-like genes would not be the initial candidates for such an approach.

124

125 Although there are *ZIP4* homologues on group 3 chromosomes, the *ZIP4* paralogue  
126 (*TaZIP4-B2*) within the *Ph1* locus on chromosome 5B is single copy, compared to the  
127 *CDK2*-like and methyl transferase gene cluster. Moreover, *ZIP4* has been shown to  
128 have a major effect on homologous COs, but not on synapsis, in both Arabidopsis and  
129 rice (Chelysheva et al. 2007; Shen et al. 2012). Knockouts of this gene in diploids  
130 usually result in sterility, as elimination of homologous COs leads to metaphase I  
131 pairing failure and incorrect segregation. We therefore assessed whether the selection  
132 and scoring of *Tazip4-B2* EMS mutants would identify wheat lines with minimal  
133 homoeologous synapsis and CO (as observed by the occurrence of multivalents at  
134 metaphase I), but which exhibit homoeologous COs in hybrids with wild-relatives.  
135 We describe the identification of two such lines through this approach.

136

## 137 **Materials and methods**

### 138 **Plant material**

139 Plant material used in this study included: wild-type hexaploid wheat (*Triticum*  
140 *aestivum* cv. Chinese Spring and cv. Cadenza); a Chinese Spring mutant lacking the  
141 *Ph1* locus (*ph1b*); two *Tazip4-B2* mutant Cadenza lines (Cadenza1691 and  
142 Cadenza0348); hexaploid wheat- *Aegilops variabilis* hybrids - crosses between  
143 hexaploid wheat (*T. aestivum* cv. Cadenza) and *Ae. variabilis* ( $2n=4x=28$ ) - using  
144 either wild type Cadenza or *Tazip4-B2* mutant lines.

145

146 For meiotic studies, the seedlings were vernalised for 3 weeks at 8°C and then  
147 transferred to a controlled environmental room until meiosis under the following  
148 growth conditions: 16 h: 8 h, light: dark photoperiod at 20°C day and 15°C night,  
149 with 70% humidity. Tillers were harvested after 6 to 7 weeks, when the flag leaf was

150 starting to emerge, and anthers collected. For each dissected floret, one of the three  
151 synchronised anthers was squashed in 45% acetic acid in water and assigned to each  
152 meiotic stage by observation under a LEICA DM2000 microscope  
153 (LeicaMicrosystems, <http://www.leica-microsystems.com/>). The two remaining  
154 anthers were either fixed in 100% ethanol: acetic acid 3:1 (v/v) for cytological  
155 analysis of meiocytes or frozen in liquid nitrogen and stored at -80°C for RNA-seq  
156 analysis.

157

## 158 **RNA-seq experiments**

### 159 *Sample preparation*

160 Anthers from wild-type wheat (WT) and wheat lacking the *Ph1* locus (*ph1b* deletion)  
161 were harvested as described in the plant material section. Anthers at late leptotene-  
162 early zygotene stage were later harvested into RNA later (Ambion, Austin, TX). The  
163 anthers from three plants of each genotype were pooled in a 1.5 ml Eppendorf until  
164 reaching 200 to 400 anthers. Once sufficient anthers had been collected, the material  
165 was disrupted using a pestle, centrifuged to eliminate the RNA, and then  
166 homogenized using QIAshredder spin columns (Qiagen, Hilden, Germany). RNA  
167 extraction was performed using a miRNeasy Micro Kit (Qiagen, Hilden, Germany)  
168 according to the manufacturer's instructions. This protocol allows purification of a  
169 separate miRNA-enriched fraction (used for further analysis) and the total RNA  
170 fraction (□200 nt) used in this study. This process was repeated to obtain three  
171 biological samples of each genotype.

172

### 173 *RNA-seq library preparation and sequencing*

174 1 µg of RNA was purified to extract mRNA with a poly-A pull down using biotin  
175 beads. A total of six libraries were constructed using the NEXTflex™ Rapid  
176 Directional RNA-Seq Kit (Bioo Scientific Corporation, Austin, Texas, USA) with the  
177 NEXTflex™ DNA Barcodes – 48 (Bioo Scientific Corporation, Austin, Texas, USA)  
178 diluted to 6 µM. The library preparation involved an initial QC of the RNA using  
179 Qubit DNA (Life technologies, CA, Carlsbad) and RNA (Life technologies, CA,  
180 Carlsbad) assays as well as a quality check using the PerkinElmer GX with the RNA  
181 assay (PerkinElmer Life and Analytical Sciences, Inc., Waltham, MA, USA). The  
182 constructed stranded RNA libraries were normalised and equimolar pooled into one  
183 final pool of 5.5 nM using elution buffer (Qiagen, Hilden, Germany, Hilden,  
184 Germany). The library pool was diluted to 2 nM with NaOH, and 5 µl were  
185 transferred into 995 µl HT1 (Illumina) to give a final concentration of 10 pM. 120 µl  
186 of the diluted library pool were then transferred into a 200 µl strip tube, spiked with  
187 1% PhiX Control v3, and placed on ice before loading onto the Illumina cBot. The  
188 flow cell was clustered using HiSeq PE Cluster Kit v4, utilising the Illumina  
189 PE\_HiSeq\_Cluster\_Kit\_V4\_cBot\_recipe\_V9.0 method on the Illumina cBot.  
190 Following the clustering procedure, the flow cell was loaded onto the Illumina  
191 HiSeq2500 instrument following the manufacturer's instructions. The sequencing  
192 chemistry used was HiSeq SBS Kit v4 with HiSeq Control Software 2.2.58 and RTA  
193 1.18.64. The library pool was run in a single lane for 125 cycles of each paired end  
194 read. Reads in bcl format were demultiplexed based on the 6 bp Illumina index by  
195 CASAVA 1.8, allowing for a one base-pair mismatch per library, and converted to  
196 FASTQ format by bcl2fastq.

197

198 *RNA-seq Data Processing*

199 The raw reads were processed using SortMeRNA v2.0 (Kopylova et al. 2012) to  
200 remove rRNA reads. The non rRNA reads were then trimmed using Trim Galore  
201 v0.4.1 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)) to remove  
202 adaptor sequences and low quality reads (-q 20 --length 80 --stringency 3). 273,739  
203 transcripts (*Triticum\_aestivum\_CS42\_TGACv1\_scaffold.annotation*) were quantified  
204 using kallisto v0.43.0 (Bray et al. 2016). The index was built using a k-mer length of  
205 31, then Kallisto quant was run using the following options -b 100 --rf-stranded.  
206 Transcript abundance was obtained as Transcripts Per Million (TPM) for each gene.

207

### 208 **Cytological analysis and image processing**

209 A total of five plants per line were examined. Anthers from *Tazip4-B2* Cadenza  
210 mutant lines, wild type Cadenza, *Tazip4-B2* mutant line-*Ae variabilis* hybrids and  
211 Cadenza-*Ae. variabilis* hybrids were harvested as described in the plant material  
212 section. Cytological analysis of meiocytes was performed using Feulgen reagent as  
213 previously described (Sharma and Sharma 1980). Images were collected using a  
214 LEICA DM2000 microscope (LeicaMicrosystems, [http://www.leica-](http://www.leica-microsystems.com/)  
215 [microsystems.com/](http://www.leica-microsystems.com/)), equipped with a Leica DFC450 camera and controlled by LAS  
216 v4.4 system software (Leica Biosystems, Wetzlar, Germany). Images were processed  
217 using Adobe Photoshop CS5 (Adobe Systems Incorporated, US) extended version  
218 12.0 x 64.

219

### 220 **Nucleotide analysis**

221 The regions of each mutation in both mutant Cadenza lines were sequenced to  
222 confirm the existence of either missense or nonsense mutations in Cadenza1961 and  
223 Cadeza0348 respectively. Wheat leaf tissues from wild type Cadenza and mutant

224 Cadenza lines were harvested at growth stages 3-4 (Feekes scale). DNA was extracted  
225 using the CTAB method (Murray and Thomson 1980). Two pairs of primers were  
226 designed using the Primer3plus software (Untergasser et al. 2007) based on the *ZIP4*  
227 sequence called TRIAE\_CS42\_5BL\_TGACv1\_404600\_AA1305800  
228 ([http://plants.ensembl.org/Triticum\\_aestivum/Info/Index](http://plants.ensembl.org/Triticum_aestivum/Info/Index)). The primers used were:  
229 forward primer: 5'GCCGCCATGACGATCTCCGAG3' and reverse primer:  
230 5'GGACGCGAGGGACGCGAG3' for Cadenza1691 and forward primer:  
231 5'GTGTTTCCTAATGCTCACAACCTC3' and reserve primer:  
232 5'ACCAGACATACTTGTGCTTGGT3' for Cadenza0348. PCR amplification was  
233 performed using MyFi Polymerase (Bioline Tauton, MA, USA), according to the  
234 manufacturer's instructions. The primers were amplified as follows: 3 min 95 °C, 35  
235 cycles of 15s at 95 °C, 15s at 58 °C and 30s at 72 °C. PCR products were resolved on  
236 2% agarose gels in 1xTBE and stained with ethidium bromide and visualised under  
237 UV light. PCR products were purified using Qiaquick PCR Purification kits (Qiagen,  
238 Hilden, Germany) and cloned using a p-GEM T easy vector kit (Promega, Madison,  
239 Wisconsin, USA). The ligation mixture was used to transform *Escherichia coli* DH5a,  
240 and transformants were selected on LB agar containing ampicillin (100mg/ml)  
241 (Sigma, St. Louis, MO, USA), Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG, 100  
242 mM) (Sigma, St. Louis, MO, USA), and 5-bromo-4-chloro-3-indolyl  $\beta$ -D-  
243 galactopyranoside (X-Gal, 20 mg/ml) (Sigma, St. Louis, MO, USA). The PCR  
244 fragments were isolated using QIAprep Spin Miniprep kit (Qiagen, Hilden,  
245 Germany). All kits were used as described in the manufacturer's instructions.  
246 Sequencing was carried out by the Eurofins Company. Alignment of sequences was  
247 carried out using Clustal Omega software (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).  
248

## 249 **Statistical analyses**

250 Statistical analyses were performed using STATISTIX 10.0 software (Analytical  
251 Software, Tallahassee, FL, USA). *Tazip4-B2* mutant line-*Ae. variabilis* hybrids were  
252 analysed by the Kruskal–Wallis test (nonparametric ANOVA) to compare the  
253 chiasma frequency among lines followed by Dunn’s Test,  $P < 0.05$ . *Tazip4-B2* mutant  
254 Cadenza lines were analysed by analysis of variance (ANOVA) based on randomised  
255 blocks. Several data transformations were applied to meet the requirement of  
256 homogeneity of variances. These included: exponential (univalents), cosine (rod  
257 bivalents and chiasma frequency) and sine (ring bivalents) transformations. Means  
258 were separated using the Least Significant Difference (LSD) test and with a  
259 probability level of 0.05.

260

## 261 **Results and Discussion**

### 262 ***TaZIP4-B2* expression**

263 In hexaploid wheat, *ZIP4* homologues are located within the *Ph1* locus on 5B, and  
264 also on chromosomes 3A, 3B and 3D. Before undertaking the targeted induced lesion  
265 in genomes (tilling) mutant analysis, we assessed the expression of *TaZIP4-B2* to  
266 confirm: that the *TaZIP4-B2* gene within the *Ph1* locus was expressed during meiosis;  
267 that it had a higher level of expression than those present on chromosome group 3;  
268 and finally, that deletion of *Ph1* significantly reduced overall *ZIP4* expression. At the  
269 coding DNA sequence and amino acid level, *TaZIP4-B2* (AA1305800.1) showed  
270 95.3% and 89.2% similarity to *TaZIP4-B1* (AA0809860.1), 94.1% and 87.5% to  
271 *TaZIP4-A1* (AA0645950.1), and 94.4% and 87.2% to *TaZIP4-D1* (AA0884100.1)  
272 respectively (Supplementary Fig. 1). To compare the relative expression of these *ZIP4*  
273 homologues on chromosomes 3A, 3B, 3D and 5B, RNA samples were collected from

274 anthers of hexaploid wheat (*Triticum aestivum* cv. Chinese Spring) (WT) and the *Ph1*  
275 deletion mutant (*ph1b*) at late leptotene-early zygotene stage, and six libraries were  
276 prepared for the RNA-seq study. RNA-seq analysis showed that *TaZIP4-B2* exhibited  
277 a higher level of expression than the *ZIP4* homologues on homoeologous group 3  
278 chromosomes (Fig. 1; Supplementary Fig. 1). Moreover, *TaZIP4-B2* also showed  
279 three splice variants (Supplementary Fig. 2) in contrast to homoeologous group 3  
280 chromosome *ZIP4* homologues. One of these splice variants (splice variant 1)  
281 accounted for 97% of the *TaZIP4-B2* transcripts. As expected, when the *Ph1* locus  
282 was deleted, the level of expression of *TaZIP4-B2* was also eliminated ( $p < 0.05$ ), but  
283 there was no apparent increase in the transcription of the *ZIP4* homologues on  
284 homoeologous group 3 chromosomes, to compensate for the absence of *ZIP4* on  
285 chromosome 5B ( $p > 0.05$ ) (Fig. 1; Supplementary Table 1). Thus, RNA-seq data  
286 revealed that the expression of *ZIP4* was derived mainly from the gene (*TaZIP4-B2*)  
287 on chromosome 5B, within the *Ph1* locus.

288

### 289 ***TaZIP4-B2* suppresses homoeologous COs in wheat-*Ae. variabilis* hybrids**

290 The protein-coding sequences of 1,200 EMS mutant lines (Rakszegi et al. 2010) from  
291 hexaploid variety “Cadenza” have been recently sequenced using exome-capture, and  
292 displayed to allow the identification of millions of mutations in the sequenced genes  
293 ([www.wheat-tilling.com](http://www.wheat-tilling.com)) (Krasileva et al. 2017). The mutations identified are  
294 accessible using the wheat survey sequence (Mayer et al. 2014) via a database, which  
295 includes their location within the gene and the predicted effect that each variant has  
296 on its protein. Simply searching this database reveals those plants possessing  
297 mutations in the target genes, as well as a list of all mutations possessed by the plant  
298 (Krasileva et al. 2017). We selected seven of the 1,200 EMS mutant lines, which

299 possessed potentially interesting mutations within *TaZIP4-B2*  
300 (Traes\_5BL\_9663AB85C.1) (Supplementary Table 2). Five of these mutant lines  
301 exhibited regular pairing at meiotic metaphase I, so were not taken further. However,  
302 two of the mutant lines (Cadenza1691 and Cadenza0348) showed reduced number of  
303 COs, suggesting that their *Tazip4-B2* mutations exhibit a phenotype. Both lines were  
304 selected for wide crossing studies with wild relatives to score the effect of their  
305 *Tazip4-B2* mutations on homoeologous CO frequency in the wheat *Tazip4-B2* mutant-  
306 wild relative hybrids, as compared to wheat wild type-wild relative hybrids.  
307 Mutations within *TaZIP4-B2* were verified by sequencing, and primers were designed  
308 to the mutated regions to follow mutated genes during crossing (Supplementary Fig. 3  
309 and Materials and Methods). *Tazip4-B2*, within the Cadenza1691 mutant line,  
310 possessed a missense mutation within one of the tetratricopeptide repeats (C to T  
311 change leading to an A167V), shown to be important for ZIP4 function (Perry et al.  
312 2005). *Tazip4-B2*, within the Cadenza0348 mutant line, possessed a nonsense  
313 mutation (a premature stop codon: G to A change leading to W612\*) (Supplementary  
314 Fig. 3). In addition to these *Tazip4-B2* mutations, the two mutant lines also possessed  
315 mutations (mostly missense, but also splice or stop codons) within the coding  
316 sequences of 106 other shared genes. Sixteen of these genes, including *TaZIP4-B2*,  
317 were located on chromosome 5B. However, none of the genes apart from *TaZIP4-B2*  
318 were located within the 2.5 MB *Ph1* region defined in our previous study (Griffiths et  
319 al. 2006; Al-Kaff et al. 2008).

320

321 Compared to the chromosome 5B deletion mutant - wild relative hybrid, no other  
322 wheat chromosome deletion mutants have previously been reported as exhibiting a  
323 similar level of homoeologous CO formation at metaphase I (Riley and Chapman

324 1958; Sears 1977). For example, the 3D locus *Ph2*, exhibits a four-fold lower level of  
325 induction compared to *Ph1* (Prieto et al. 2005). Equally, deleting regions of  
326 chromosome 5B apart from the 2.5 MB *Ph1* region did not result in homoeologous  
327 CO formation at metaphase I when the lines were crossed with wild relatives (Roberts  
328 et al. 1999; Griffiths et al. 2006; Al-Kaff et al. 2008). Sears (1977) used such crosses  
329 between hexaploid wheat cv. Chinese Spring, both in the presence and absence (*ph1b*  
330 deletion) of *Ph1*, and the wild relative tetraploid *Aegilops kotschy* (also termed *Ae.*  
331 *variabilis*), to show that homoeologous COs are induced when the *Ph1* locus is  
332 deleted. Interspecific hybrids of the *ph1b* mutant and wild relatives have been  
333 subsequently used in plant breeding programmes for introgression purposes (Sears  
334 1977). Sears (1977) observed one rod bivalent at metaphase I in the presence of *Ph1*,  
335 and 6.35-7.28 rod bivalents in these *Ph1* absent hybrids. Chiasma frequency in the  
336 hexaploid wheat-*Ae. kotschy* or hexaploid wheat-*Ae. variabilis* hybrids, was between  
337 1-3 in the presence of *Ph1*, and 11-14 in the absence of *Ph1* (Sears 1977; Farooq et al.  
338 1990; Fernández-Calvín and Orellana 1991; Kousaka and Endo 2012).

339

340 In this study, both *Tazip4-B2* mutant Cadenza lines, as well as a wild type Cadenza  
341 (*TaZIP4-B2*), were crossed with *Ae. variabilis*. The frequency of univalents,  
342 bivalents, multivalents and total chiasma frequency was scored at meiotic metaphase I  
343 in the resulting F<sub>1</sub> hybrid (Fig. 2). In these hybrids, there were similar numbers of rod  
344 bivalents to that reported by Sears (1977), with 6.75 (SE 0.17) (Cadenza1691) and  
345 6.64 (SE 0.18) (Cadenza0348) rod bivalents at metaphase I in the *Tazip4-B2* mutants,  
346 and 1.48 rod bivalents (SE 0.12) at metaphase I in the wild type Cadenza. Moreover,  
347 the chiasma mean frequency was 1.48 (SE 0.12) in the presence of *Ph1* (*TaZIP4-B2*),  
348 and 12.21 (SE 0.19) and 12.23 (SE 0.20) in the Cadenza1691-*Ae. variabilis* and

349 Cadenza0348-*Ae. variabilis* hybrids, respectively. The observed chiasma frequencies  
350 at metaphase I, in the two *Tazip4-B2* mutant line-*Ae. variabilis* hybrids, are similar to  
351 those previously reported at metaphase I in the *Ph1* deletion mutant (*ph1b*)-*Ae.*  
352 *variabilis* hybrids. Thus, the data indicate that *TaZIP4-B2* within the *Ph1* locus is  
353 likely to be involved in the suppression of homoeologous COs.

354

#### 355 ***Tazip4-B2* mutant Cadenza lines show no multivalents**

356 The frequencies of meiotic associations at metaphase I in hexaploid wheat and the  
357 *Ph1* deletion mutant (*ph1b*) have been reported previously (Martín et al. 2014).  
358 Martín et al. (2014) observed 20 ring bivalents and one rod bivalent, with a chiasma  
359 frequency of 40.97 in the presence of the *Ph1* locus. However, the number of ring  
360 bivalents decreased to 14.83, with a reduced chiasma frequency of 35.78, while the  
361 number of rod bivalents, univalents, trivalents and tetravalents increased to 4.73, 0.80,  
362 0.20 and 0.37, respectively, when the *Ph1* locus was absent. The number of  
363 univalents, bivalents, multivalents and chiasma frequency at meiotic metaphase I was  
364 also scored in both the *Tazip4-B2* mutant lines and in the wild type Cadenza (Fig. 3).  
365 The *Tazip4-B2* mutant lines exhibited a reduction in the number of ring bivalents at  
366 metaphase I, and a slight increase in the number of rod bivalents, from a mean of 1.30  
367 (SE 0.17) in the wild type Cadenza, to 3.29 (SE 0.16) in Cadenza1691 and 3.63 (SE  
368 0.18) in Cadenza0348. This indicates a slight reduction in homologous COs in these  
369 *Tazip4-B2* mutant lines. CO frequency was a mean of 40.50 (SE 0.21) in the wild type  
370 Cadenza, 38.13 (SE 0.20) in Cadenza1691 and 37.30 (SE 0.23) in Cadenza0348.  
371 These observed chiasma frequencies at metaphase I in the two mutant Cadenza lines,  
372 are again similar to those previously reported at metaphase I in wheat in the absence  
373 of the *Ph1* locus. However, no multivalents were observed, and there was no

374 significant increase in the number of univalents at metaphase I in the *Tazip4-B2*  
375 mutant lines, as is normally observed in *Ph1* deletion mutants (Roberts et al. 1999).

376

377 If *Tazip4-B2* mutants had enabled homoeologues to synapse while failing to CO, then  
378 a significant increase in univalents would be expected, but this was not observed. This  
379 suggests that homoeologous synapsis may not be significantly affected by *TaZIP4-B2*.

380 On the other hand, the lack of multivalents at metaphase I suggests that both mutant  
381 lines will exhibit a reduced level of homoeologous exchange or chromosome  
382 translocation to that observed in the CS *ph1b* mutant. The *ph1b* mutant line has been  
383 reported to accumulate extensive background translocations over multiple generations  
384 due to homoeologous synapsis and COs (Sánchez-Morán et al. 2001). Thus, the  
385 apparent lack of multivalents in the *Tazip4-B2* mutant lines could allow their  
386 exploitation for introgression purposes during plant breeding programmes, rather than  
387 the current *ph1b* line.

388

389 In summary, seven lines carrying mutations within the *TaZIP4-B2* gene were  
390 screened for a phenotype with reduced homologous crossover at metaphase I. Of  
391 these, two lines were identified with this phenotype, one carrying a nonsense mutation  
392 within *TaZip4-B2*, and the other carrying a mutation in one of the key functional  
393 domains of *TaZip4-B2*. When crossed with *Ae. variabilis*, both of these lines also  
394 exhibited increased homoeologous crossover at metaphase I in the resulting hybrid,  
395 suggesting that the two phenotypes were linked. Therefore, in this case, lines with  
396 increased homoeologous crossover had been identified without an initial screen for  
397 the desired phenotype. Until now, the only way by which homoeologous crossover at  
398 metaphase I could be increased to this extent in wheat-wild relative hybrids was by

399 deletion of the *Phl* locus, defined as a deletion effect phenotype specific to  
400 chromosome 5B. However, an alternative way of reproducing the *Phl* deletion effect  
401 would be to use EMS treatment to generate nonsense or truncation mutations in the  
402 homoeologous crossover-suppressing gene within the *Phl* locus. Analysis of the  
403 1200-line Tilling population revealed that in any given mutant line, 1.5% of genes  
404 will have a truncation allele and 2% a missense allele (Krasileva et al. 2017). Thus,  
405 the probability of two mutant lines both sharing a truncation or missense mutation by  
406 chance in the same second gene is  $P < 0.0005$ . The probability that two mutated genes  
407 will be located on the same chromosome is extremely low ( $2.4 \times 10^{-5}$ ), and the  
408 probability that they will both be located within the *Phl* region is even lower ( $2.4$   
409  $\times 10^{-7}$ ). Thus, it is extremely unlikely that the increased homoeologous crossover  
410 phenotype found in both the two *Tazip4-B2* mutant lines results from a nonsense  
411 mutation in a further gene independently linked with *Tazip4-B2* within  
412 the *Phl* locus. These mutants were identified through a non-GM route, and can be  
413 exploited as an alternative to the CS *ph1b* mutant. Seeds for both mutants have been  
414 deposited with the Germplasm Resource Unit at the John Innes Centre  
415 ([www.jic.ac.uk/research/germplasm-resources-unit](http://www.jic.ac.uk/research/germplasm-resources-unit)). The accession number of seeds  
416 for Cad0348 is W10336 and for Cad1691 is W10337. The seeds for both lines are  
417 available on request, free of intellectual property restrictions.

418

#### 419 **Author contribution statements**

420 These authors made the following contributions to the manuscript: M-DR, ACM, PS  
421 and GM designed the research; M-DR, ACM and JH performed the research; M-DR,  
422 ACM, JH, PS and GM analysed the data; M-DR, PS and GM wrote the manuscript.

423

424 **Compliance with ethical standards**

425 This research does not involve human participants or animals.

426

427 **Conflict of interests**

428 The authors declare that they have no competing interests.

429

430 **References**

431 Al-Kaff N, Knight E, Bertin I, Foote T, Hart N, Griffiths S, Moore G (2008) Detailed

432 dissection of the chromosomal region containing the *Ph1* locus in wheat *Triticum*

433 *aestivum*: with deletion mutants and expression profiling. *Ann Bot* 101:863-872

434 Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-

435 seq quantification. *Nat Biotechnol* 34:525-527

436 Chelysheva L, Gendrot G, Vezon D, Doutriaux MP, Mercier R, Grelon M (2007)

437 *Zip4/Spo22* is required for class I CO formation but not for Synapsis completion in

438 *Arabidopsis thaliana*. *PLoS Genet* 3:802-813

439 Dhaliwal HS, Gill BS, Waines JG (1977) Analysis of induced homoeologous pairing

440 in a *ph* mutant wheat x rye hybrid. *J Hered* 68:207-209

441 Farooq S, Iqbal N, Shah TM (1990) Promotion of homoeologous chromosome pairing

442 in hybrids of *Triticum aestivum* x *Aegilops variabilis*. *Genome* 33:825-828

443 Fernández-Calvín B, Orellana J (1991) Metaphase I bound arms frequency and

444 genome analysis in wheat-Aegilops hybrids: *Ae. variabilis* - wheat and *Ae.*

445 *kotschyi* - wheat hybrids with low and high homoeologous pairing. *Theor Appl*

446 *Genet* 83:264-272

- 447 Greer E, Martín AC, Pendle A, Colas I, Jones AME, Moore G, Shaw P (2012) The  
448 *Ph1* locus suppresses Cdk2-type activity during premeiosis and meiosis in wheat.  
449 Plant Cell 24:152-162
- 450 Griffiths S, Sharp R, Foote TN, Bertin I, Wanous M, Reader S, Colas I, Moore G  
451 (2006) Molecular characterization of *Ph1* as a major chromosome pairing locus in  
452 polyploid wheat. Nature 439:749-752
- 453 Knight E, Greer E, Draeger T, Thole V, Reader S, Shaw P, Moore G (2010) Inducing  
454 chromosome pairing through premature condensation: analysis of wheat  
455 interspecific hybrids. Funct. Integr. Genomics. 10: 603-608.
- 456 Kopylova E, Laurent N, Touzet H (2012) SortMeRNA: fast and accurate filtering of  
457 ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211-3217
- 458 Kousaka R, Endo TR (2012) Effect of rye B chromosome and its segments on  
459 homoeologous pairing in hybrids between common wheat and *Aegilops variabilis*.  
460 Genes Genet Syst 87:1-7
- 461 Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L,  
462 Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P, Fosker C, Ayling S,  
463 Phillips AL, Uauy C, Dubcovsky J (2017) Uncovering hidden variation in  
464 polyploid wheat. Proc Natl Acad Sci USA 114:E913-E921
- 465 Martín AC, Shaw P, Phillips D, Reader S, Moore G (2014) Licensing MLH1 sites for  
466 crossover during meiosis. Nat Commun 5:1-5
- 467 Martín AC, Rey MD, Shaw P, Moore G (2017) Dual effect of the wheat *Ph1* locus on  
468 chromosome synapsis and crossover. Chromosoma DOI: 10.1007/s00412-017-  
469 0630-0

- 470 Marcussen T, Sandve SR, Heier L, Wulff BB, Steuernagel B, Mayer KF, Olsen OA  
471 (2014) A chromosome-based draft sequence of the hexaploid bread wheat  
472 (*Triticum aestivum*) genome. *Science* 345:1251788
- 473 Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant  
474 DNA. *Nucleic Acids Res* 8:4321–4326
- 475 Perry P, Kleckner N, Börner GV (2005) Bioinformatic analyses implicate the  
476 collaborating meiotic crossover/chiasma proteins Zip2, Zip3, and Spo22/Zip4 in  
477 ubiquitin labeling. *Proc Natl Acad Sci USA* 102:17594-17599
- 478 Prieto P, Moore G, Reader S (2005) Control of conformation change associations  
479 with homologue recognition during meiosis. *Theor Appl Genet* 111:505-510
- 480 Rakszegi M, Kisgyörgy BN, Tearall K, Shewry PR, Láng L, Phillips A, Bedó Z  
481 (2010) Diversity of agronomic and morphological traits in a mutant population of  
482 bread wheat studied in the Healthgrain program. *Euphytica* 174:409-421
- 483 Riley R, Chapman V (1958) Genetic control of the cytologically diploid behaviour of  
484 hexaploid wheat. *Nature* 182:713-715
- 485 Roberts MA, Reader SM, Dalgliesh C, Miller TE, Foote TN, Fish LJ, Snape JW,  
486 Moore G (1999) Induction and characterisation of the *Ph1* wheat mutants. *Genetics*  
487 153:1909-1918
- 488 Sánchez-Morán E, Benavente E, Orellana J (2001) Analysis of karyotypic stability of  
489 homoeologous-pairing (*ph*) mutants in allopolyploid wheats. *Chromosoma*  
490 110:371-377
- 491 Sears ER, Okamoto M (1958) Intergenomic chromosome relationships in hexaploid  
492 wheat. *Proc Xth Internat Congr Genet, Montreal* 2:258–259
- 493 Sears ER (1977) Induced mutant with homoeologous pairing in common wheat. *Can J*  
494 *Genet Cytol* 19:585-593

495 Sharma AK, Sharma A (2014) Chromosome techniques: theory and practice.  
496 Butterworth-Heinemann  
497 Shen Y, Tang D, Wang K, Wang M, Huang J, Luo W, Luo Q, Hong L, Li M, Cheng  
498 Z (2012) ZIP4 in homologous chromosome synapsis and crossover formation in  
499 rice meiosis. J Cell Sci 125:2581-2591  
500 Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA (2007)  
501 Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35:W71-  
502 W74  
503 Zheng T, Nibau C, Phillips DW, Jenkins G, Armstrong SL, Doonan JH (2014)  
504 CDKG1 protein kinase and male meiosis at high ambient temperature in  
505 *Arabidopsis thaliana*. Proc Natl Acad Sci USA 111: 2182-2187.

506

507

## 508 **Figures**

509 **Fig. 1** Relative expression of *ZIP4* homologs in *Triticum aestivum* cv. Chinese Spring  
510 in presence (WT) and in absence (*ph1b* deletion) of the *Ph1* locus obtained by RNA-  
511 seq analysis.

512

513 **Fig. 2** Homoeologous CO frequency at metaphase I is increased in *Tazip4-B2* mutant  
514 line-*Ae. variabilis* hybrids (**b** and **c**), in comparison to the wild type Cadenza-*Ae.*  
515 *variabilis* hybrid (**a**). The number of univalents, bivalents, multivalents and chiasma  
516 frequency were scored at meiotic metaphase I in Cadenza0000 (*TaZIP4-B2*) x *Ae.*  
517 *variabilis* hybrids, and in Cadenza1691 (*Tazip4-B2*) x *Ae. variabilis* hybrids and  
518 Cadenza0348 (*Tazip4-B2*) x *Ae. variabilis* hybrids. The same letter indicates no

519 differences between *TaZIP4-B2* (a) and *Tazip4-B2* hybrids (b and c) in metaphase I at  
520  $P < 0.05$ . *Scale bar* represents 10  $\mu\text{m}$  for all panels.

521

522 **Fig. 3** Homologous CO frequency is reduced in *Tazip4-B2* mutant Cadenza lines (b  
523 and c), in comparison to wild type Cadenza (a). The number of univalents, bivalents  
524 and chiasma frequency were scored at meiotic metaphase I in Cadenza0000 (*TaZIP4-*  
525 *B2*), and in Cadenza1691 (*Tazip4-B2*) and Cadenza0348 (*Tazip4-B2*). Asterisks  
526 indicate the presence of rod bivalents in both mutant Cadenza lines. The same letter  
527 indicates no differences between *TaZIP4-B2* (a) and *Tazip4-B2* wheat (b and c) in  
528 metaphase I at  $P < 0.05$ . *Scale bar* represents 10  $\mu\text{m}$  for all panels.

529

### 530 **Supplementary material**

531 **Supplementary Fig. 1** Alignment of coding DNA (a) and amino acid sequences (b)  
532 from *TaZIP4-B2* (TRIAE\_CS42\_5BL\_TGACv1\_404600\_AA1305800), *TaZIP4-B1*  
533 (TRIAE\_CS42\_3B\_TGACv1\_225572\_AA0809860), *TaZIP4-A1*  
534 (TRIAE\_CS42\_3AL\_TGACv1\_195180\_AA0645950) and *TaZIP4-D1*  
535 (TRIAE\_CS42\_3DL\_TGACv1\_251716\_AA0884100) CS+ refers to wild type and  
536 CS- to the *ph1b* mutant.

537

538 **Supplementary Fig. 2** Alignment of *TaZIP4-B2* splice variants in wild type Cadenza.

539

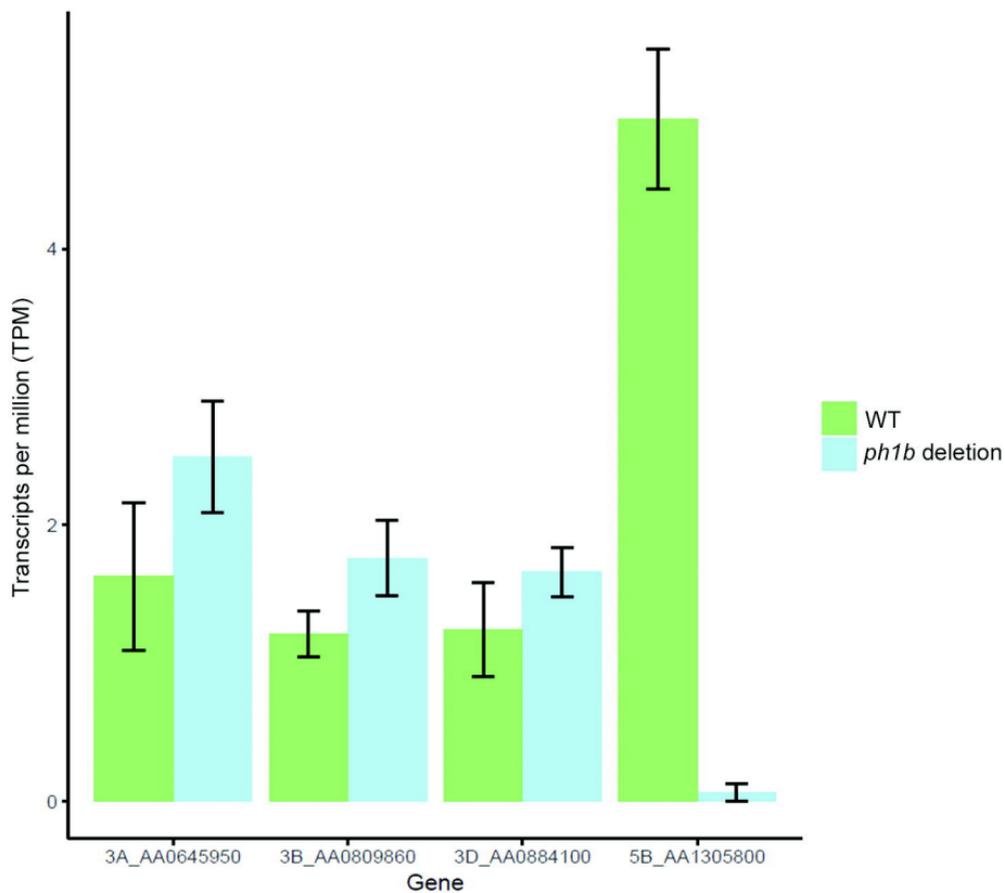
540 **Supplementary Fig. 3** Alignment of *TaZIP4-B2* coding DNA sequences from wild  
541 type Cadenza (WTC) and mutant Cadenza lines (MCL). Both missense and nonsense  
542 mutants are highlighted. Primer sequences used to follow mutated genes during  
543 crossing are underlined. All primers are shown in direction 5'  $\rightarrow$  3'.

544

545 **Supplementary Table 1** Detailed information of all transcripts obtained by Kallisto  
546 and the statistical analysis of transcripts per million (TPM) in wheat chromosomes  
547 3AL, 3B, 3DL and 5BL, both in presence (WT) and in absence (*ph1b* deletion) of the  
548 *Ph1* locus. Data represent mean values  $\pm$  standard error (SE) from RNA samples  
549 collected at late leptotene-early zygotene stage in WT and in *ph1b* deletion.

550

551 **Supplementary Table 2** Detailed information on the seven EMS mutant lines  
552 selected as possessing potentially interesting mutations within *TaZIP4-B2*  
553 (Traes\_5BL\_9663AB85C.1). The two mutant lines (Cadenza1691 and Cadenza0348)  
554 which showed reduced number of COs in Cadenza mutant lines are indicated in bold.



Line	No. of cell examined	Univalents	Rod bivalents	Ring bivalents	Trivalents	Tetraploids	Pentavalents	Chiasma frequency
		Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)	Mean $\pm$ SE (Range)
<b>Cad0000 x <i>Ae. variabilis</i> hybrids</b>	128	32.05 $\pm$ 0.24 (25-35)	1.48 $\pm$ 0.12 (0-5)	-	-	-	-	1.48 $\pm$ 0.12 <sup>b</sup> (0-5)
<b>Cad1691 x <i>Ae. variabilis</i> hybrids</b>	117	14.74 $\pm$ 0.29 (7-26)	6.75 $\pm$ 0.17 (3-11)	1.26 $\pm$ 0.08 (0-4)	1.05 $\pm$ 0.08 (0-4)	0.22 $\pm$ 0.04 (0-2)	0.03 $\pm$ 0.02 (0-1)	12.21 $\pm$ 0.19 <sup>a</sup> (8-18)
<b>Cad0348 x <i>Ae. variabilis</i> hybrids</b>	102	14.63 $\pm$ 0.28 (6-21)	6.64 $\pm$ 0.18 (3-10)	1.12 $\pm$ 0.10 (0-4)	1.27 $\pm$ 0.11 (0-4)	0.21 $\pm$ 0.04 (0-1)	0.05 $\pm$ 0.02 (0-1)	12.23 $\pm$ 0.20 <sup>a</sup> (7-19)
<i>p-value</i>								<b>0.0000</b>



Line	No. of cell examined	Univalents Mean $\pm$ SE (Range)	Rod bivalents Mean $\pm$ SE (Range)	Ring bivalents Mean $\pm$ SE (Range)	Chiasma frequency Mean $\pm$ SE (Range)
<b>Cadenza0000</b>	80	0.18 $\pm$ 0.11 <sup>a</sup> (0-2)	1.30 $\pm$ 0.17 <sup>a</sup> (0-4)	19.60 $\pm$ 0.19 <sup>a</sup> (17-21)	40.50 $\pm$ 0.21 <sup>b</sup> (38-42)
<b>Cadenza1691</b>	96	0.58 $\pm$ 0.10 <sup>a</sup> (0-6)	3.29 $\pm$ 0.16 <sup>c</sup> (0-7)	17.42 $\pm$ 0.17 <sup>c</sup> (13-21)	38.13 $\pm$ 0.20 <sup>a</sup> (31-42)
<b>Cadenza0348</b>	71	1.07 $\pm$ 0.12 <sup>a</sup> (0-6)	3.63 $\pm$ 0.18 <sup>b</sup> (0-7)	16.83 $\pm$ 0.20 <sup>b</sup> (12-21)	37.30 $\pm$ 0.23 <sup>a</sup> (30-42)
<i>p-value</i>		<i>0.2095</i>	<i>0.0000</i>	<i>0.0000</i>	<i>0.0000</i>

