bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 **Title page**

2 **Authors**

3 Muhammed Rezwan Kabir[1] ·Heather M. Nonhebel[1*] ·David Backhouse[2] · Gal Winter[1]

4 **Title**

5 Expression of *TaTAR2.3-1B*, *TaYUC9-1* and *TaYUC10* correlates with auxin and starch
6 content of developing wheat grains

7 **Affiliations**

8 [1]School of Science and Technology, University of New England, Armidale, New South
9 Wales 2351, Australia

10 [2]School of Environmental and Rural Science, University of New England, Armidale, New
11 South Wales 2351, Australia

12 Corresponding author: Heather. M. Nonhebel, hnonheb2@une.edu.au

13 **ORCID IDs**

14 MRK 0000-0002-8115-7725, DB 0000-0003-0663-6002, GW 0000-0003-3789-395X, HMN
15 0000-0001-9055-3095

16 **Abstract**

17 The role of auxin in developing grains of wheat (*Triticum aestivum*) is contentious with
18 contradictory reports indicating either positive or negative effects of IAA (indole-3-acetic
19 acid) on grain size. In addition, the contributions to the IAA pool from de novo synthesis via
20 tryptophan, and from hydrolysis of IAA-glucose are unclear. Here we describe the first
21 comprehensive study of tryptophan aminotransferase and indole-3-pyruvate mono-oxygenase
22 expression during wheat grain development from 5 to 20 days after anthesis. A comparison of
23 expression data with measurements of endogenous IAA via combined liquid
24 chromatography-tandem mass spectrometry with heavy isotope labelled internal standards
25 indicates that TaTAR2.3-1B, TaYUC9-A1, TaYUC9-B, TaYUC9-D1, TaYUC10-A and

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

26 TaYUC10-D are primarily responsible for IAA production in developing grains.
27 Furthermore, we show that IAA synthesis is controlled by genes expressed specifically in
28 developing wheat grains as has already been reported in rice (*Oryza sativa*) and maize (*Zea*
29 *mays*). Our results cast doubt on the proposed role of *THOUSAND-GRAIN WEIGHT* gene,
30 *TaTGW6*, in promoting larger grain size via negative effects on grain IAA content. The work
31 on *TaTGW6* has overlooked the contribution of the dominant IAA biosynthesis pathway.
32 Although IAA synthesis occurs primarily in the endosperm of wheat grains, we show that the
33 *TaYUC9-1* group is also strongly expressed in the embryo. Within the endosperm, *TaYUC9-1*
34 expression is highest in aleurone and transfer cells, supporting data from other cereals
35 suggesting that IAA has a key role in differentiation of these tissues.

36 **Keywords**

37 Auxin, grain fill, TaTAR2, TaYUC, TaTGW6, TaTGW-7A

38 **Declarations**

39 **Funding**

40 The research was supported by a Research Training Program (RTP) scholarship provided to
41 Muhammed Rezwan Kabir by the Australian Government.

42 **Conflicts of interest**

43 The authors declare that they have no conflict of interests.

44 **Availability of data and material**

45 Not applicable

46 **Code availability**

47 Not applicable

48 **Author contribution statement**

49 MRK and HMN conceived and designed the research. MRK performed all the experiments.
50 DB contributed to growing wheat plants and preparation of manuscript. GW helped in

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

51 analysing gene expression data. MRK and HMN wrote the manuscript. All authors read and
52 approved the manuscript.

53 **Key message**

54 Expression of *TaTAR2.3*, *TaYUC9-1* and *TaYUC10* coincides with increasing IAA content,
55 grain weight and starch content from 10-15 DAA, highlighting the importance of the
56 TAR/YUCCA pathway in wheat grain filling.

57 **Acknowledgements**

58 The authors are grateful to Kirsten Drew for the analysis of IAA. The authors also
59 acknowledge the Australian Government for providing a Research Training Program (RTP)
60 PhD scholarship to Muhammed Rezwan Kabir.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

61 **Introduction**

62 Global wheat (*Triticum aestivum*) production has reached 760.1 million tonnes (FAOSTAT
63 2020); however, a substantial increase in yield from the existing land area is essential to meet
64 the needs of the rapidly increasing world population. The International Wheat Yield
65 Consortium (WYC) has introduced a strategy to improve yield potential that could accelerate
66 breeding of high yielding wheat varieties (Foulkes et al. 2011). This has identified grain
67 number and grain weight as two key yield-determining factors. The plant hormone auxin or
68 IAA (indole-3-acetic acid) appears to play a major role in both of these aspects of grain
69 development.

70 In particular, inactive alleles of two genes *TaTGW6* and *TaTGW-7A* are reported to
71 improve grain size in wheat via negative effects on the grain IAA content (Hu et al. 2016a;
72 Hu et al. 2016b). *TaTGW6* is homologous to the rice gene *TGW6,* reported by Ishimaru et al.
73 (2013) to encode an IAA-glucose hydrolase. Although the work on wheat did not characterise
74 the gene product of *TaTGW6,* Hu et al. (2016a) reported that an inactive allele, *TaTGW6-c,* as
75 well as a mutant allele, *TaTGW6-b,* were associated with lower IAA content of grains at 20
76 and 30 days after anthesis (DAA) as well as higher grain weight. *TaTGW-7A* is reported as an
77 indole-3-glycerol phosphate synthase (IGPS) like gene (Hu et al. 2016b). Again, there was no
78 characterisation of the gene product. However, the authors reported a similar lower IAA
79 content in grains with the inactive *TaTGW-7Aa* allele as that found in *TaTGW6-b* and
80 *TaTGW6-c* plants.

81 In contrast, the majority of relevant publications report a positive involvement of IAA
82 on grain size in a number of cereals. For example, the shrunken grain phenotype of defective
83 endosperm and defective kernel mutants of maize, *de18* and *dek18* appears to result from low
84 levels of IAA in the developing grains caused by mutations affecting *ZmYUC1* expression
85 (Bernardi et al. 2012; Bernardi et al. 2016). In rice, a mutation in the IAA biosynthesis gene
86 *OsTAR2/FIB* is associated with the *tillering and small grain 1 (tsg1)* phenotype (Guo et al.
87 2019). Finally, in wheat, Shao et al. (2017) showed that overexpression of *TaTAR2.1-3A*
88 increases IAA accumulation in grains and enhances plant height, spike number, biomass and
89 grain yield. Thus the claim that *TaTGW6* and *TaTGW-7A* have a positive effect on grain size
90 via a reduction in IAA content requires careful scrutiny.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

91 In plants, including developing seeds, IAA is produced primarily from tryptophan via
92 the actions of tryptophan aminotransferase of Arabidopsis1 (TAA1) and its related proteins
93 (TAR), and indole-3-pyruvate monooxygenase known as YUCCA (Mashiguchi et al. 2011;
94 Won et al. 2011). In assuming that inactivation of a glucose hydrolase gene can lead to a
95 large reduction in grain IAA content, the *TGW6* work has overlooked this major source of
96 IAA. On the other hand, the same authors suggest that *TaTGW-7A* affects grain IAA content
97 via its effects on tryptophan production. The TAR/YUCCA pathway is known to occur in
98 wheat grains from three studies. Shao et al. (2017) reported a comprehensive study of the
99 *TAR* gene family based on an early version of the wheat genome and showed that of all *TAR*
100 genes, *TaTAR2.3* was maximally expressed in grains. However, they did not investigate the
101 timing of *TaTAR2.3* expression during grain fill or present any data on the IAA content of
102 grains. Li et al. (2014) showed *TaYUC10.3* is highly expressed in developing wheat grains,
103 with expression increasing throughout the grain-fill period but did not investigate any other
104 YUCCA genes from wheat or relate gene expression to the IAA content. Finally, Tuan et al.
105 (2019) described expression changes in "*TaTAR2*" (actually *TaTAR2.3-1D*) and "*TaYUC11*"
106 from microarray data during seed maturation (20-50 DAA) and also measured the IAA
107 content of endosperm and embryo tissue. However, there have been no comprehensive
108 studies of the *YUCCA* gene family in wheat, their expression profile during grain
109 development or correlation with the grain IAA content.

110 The aim of our work was therefore to evaluate the role of the TAR/YUCCA pathway
111 in the regulation of IAA content during grain fill in wheat. The study comprised the first
112 compressive phylogenetic and expression analysis of all *TAR* and *YUCCA* genes in wheat.
113 The expression profile of major IAA biosynthesis genes during early grain development was
114 compared with changes in the grain IAA content measured by the most accurate and specific
115 method available, as well as with changes in the starch content of grains. To inform future
116 studies on the specific signalling role(s) of IAA in wheat grains, we also investigated the
117 location of IAA production via dissection of grains into embryo and endosperm fractions as
118 well as by mining information from published RNA-sequencing (RNA-seq) studies.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

119 ## Materials and methods

120 ### Data mining and Bioinformatic analysis

121 Protein sequences of all *TAR* and *YUCCA* genes from wheat (*Triticum aestivum*) (IWGSC
122 RefSeq v1.1), rice (*Oryza sativa*), brachypodium (*Brachypodium distachyon*) and barley
123 (*Hordeum vulgare*) were downloaded from EnsemblPlants 47 (Kersey et al. 2015) following
124 BlastP searches using the query sequences OsTAR2, OsYUC1 and OsYUC9 from rice.
125 Arabidopsis TAA1 and TAR sequences were downloaded from The Arabidopsis Information
126 Resource (TAIR) (Rhee et al. 2003). Phylogenetic analyses of the protein sequences were
127 carried out in MEGA7.0.26 (Kumar et al. 2016) using the Maximum Likelihood method
128 (Jones et al. 1992). Multiple sequence alignments (MSA) were performed by MUSCLE
129 (Edgar 2004). Bootstrap confidence levels were obtained using 500 replicates (Felsenstein
130 1985). Evolutionary distances were computed using Poisson correction method (Zuckerkandl
131 and Pauling 1965). RNA-seq data available in expVIP (http://www.wheat-expression.com)
132 (Borrill et al. 2016; Ramírez-González et al. 2018) were used for global evaluation of gene
133 expression, as well as for investigating gene expression in dissected aleurone and transfer cell
134 tissues.

135 ### Plant materials

136 Wheat plants (Chinese Spring) were grown in 20 cm × 20 cm pots under natural light at
137 23/14°C (day/night temperatures) in the glasshouse at the University of New England,
138 Armidale, NSW. Pots were fertilized once a week with Thrive® (Yates, 1 g/L) from the
139 initiation of tillering stage. Spikes were tagged when the first spikelet reached anthesis. Grain
140 samples (70–90 mg) were harvested at the same time (4:00 to 5:00 pm) each day 5, 10, 15
141 and 20 DAA, then snap-frozen in liquid nitrogen and stored at –80°C. Some 15 DAA grains
142 were manually dissected into embryos and endosperms. Grain samples from individual spikes
143 were kept separate and all analyses were done on at least three independent biological
144 replicates harvested from different plants, on different days.

145 ### RNA extraction and quantitative analysis

146 Wheat grain samples, harvested and stored as described above, were ground in liquid
147 nitrogen and total RNA was extracted using Trizol (Invitrogen). RNA concentration and

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

148 260/280 ratio were determined using a NanoDrop ND-8000 Spectrophotometer (Thermo
149 Scientific). RNA samples with $A_{260/280}$ of 1.8–2.0 were used for further analysis. The RNA
150 quality was checked by agarose gel electrophoresis for two clear bands of 18S and 28S
151 rRNAs (Nolan et al. 2006). Primers specific for each group of three paralogues on the A, B
152 and D genomes were designed with melting points in the range of 58–60°C and product sizes
153 between 108–251 bp (Supplementary Table S1). Initial amplification was carried out using a
154 One-Step RT-PCR Kit (QIAGEN) in a BIO-RAD T100 Thermal Cycler, with gel analysis to
155 confirm a single product of the expected size. cDNAs were prepared using the SensiFAST
156 cDNA Synthesis Kit (Bioline). Quantitative real-time PCR reactions using cDNA samples as
157 template were carried out using a SensiFAST SYBR No-ROX Kit (Bioline) in a CFX96
158 Touch (BIO-RAD) machine following manufacturer's instructions. Negative control
159 reactions without reverse transcriptase as well as reactions with no template were included.
160 Three biological and three technical replicates were carried out for each primer set. The qRT-
161 PCR program included 40 cycles of 95°C 3 min, 95°C 10 s, 55°C 30 s and 72°C 5 s.
162 Expression was calculated relative to two reference genes, translation elongation factor EF-I
163 alpha Ta53964 (Paolacci et al. 2009) and a cyclin-like protein Ta27922 (Wu et al. 2015).
164 Expression data presented are the average and standard error of biological replicates.
165 Amplified products were sent to the Australian Genome Research Facility for sequencing to
166 confirm identity following the use of the Wizard SV Gel and PCR Clean-Up System
167 (Promega).

168 **IAA extraction and analysis**

169 Wheat grain samples (70–90 mg) were ground in liquid nitrogen; 200 µL of 65% isopropanol
170 /35% 0.2 M pH 7.0 imidazole buffer (Chen et al. 1988) was added with [$^{13}C_6$] IAA internal
171 standard (Cambridge Isotope Laboratories Inc.), and samples were extracted on ice for 1 h.
172 Amounts of standards added varied with the age of samples to ensure that the concentration
173 of standards was similar to that of endogenous IAA; 78 ng of [$^{13}C_6$] IAA was added to 10, 15
174 and 20 DAA samples; 16 ng of [$^{13}C_6$] IAA was added to 5 DAA samples. Blank samples
175 without plant tissue were taken through the entire extraction and analysis protocol to ensure
176 that no contamination from the unlabelled IAA in the laboratory occurred.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

177 Following extraction, samples were diluted with 2 mL deionized water, centrifuged
178 and the supernatant transferred to a glass tube. Sample clean-up followed the solid-phase
179 extraction (SPE) protocol of (Barkawi et al. 2008) with minor modifications. Samples
180 prepared as above were added to discovery® DSC-NH2 SPE 50 mg/mL tubes (Supelco) that
181 had been prewashed sequentially with 500 µL hexane, 500 µL acetonitrile, 500 µL water, 500
182 µL 0.2 M pH 7.0 imidazole buffer and 4.5 mL deionized water. After loading each sample,
183 SPE columns were washed sequentially with 500 µL each of water, hexane, ethyl acetate,
184 acetonitrile, methanol and 600 µL 0.25% phosphoric acid. The IAA was eluted in 1.8 mL
185 0.25% phosphoric acid and the pH of the eluate was adjusted to 3.0–3.5 with 150 µL 0.1 M,
186 pH 6.0 succinate buffer. This fraction was added to Strata-X SPE (8B-S100-UBJ 60 mg/3
187 mL; Phenomenex) that had been prewashed with 1 mL hexane, 1 mL methanol and 2 mL
188 water. Samples loaded on the SPE columns were washed with 3×1 mL water and 100 µL
189 acetonitrile before eluting the IAA in 1 mL acetonitrile.

190 Samples from SPE clean-up were stored at –20ºC. Immediately prior to analysis, they
191 were reduced to dryness under a stream of N_2 and redissolved in 20 µL acetonitrile and 80 µL
192 0.01 M aqueous acetic acid. The analysis of ^{12}C:^{13}C IAA was conducted using a triple
193 quadrupole Liquid Chromatograph Mass Spectrometer (LCMS)-8050, (Shimadzu) with
194 XBridge™ C18 3.5 µm, 2.1×50 mm column (Phenomenex). The chromatography solvent
195 was 20% acetonitrile: 80% 0.01 M acetic acid at a flow rate of 0.2 mL/min. The nebulizing,
196 heating and drying gas flow were 3 L/min, 10 L/min and 10 L/min, respectively. Interface
197 temperature was 300°C, DL was 250°C and the heat block temperature was 400°C. The
198 interface used a capillary voltage of 4 kV. The mass spectrometer was operated in multiple-
199 reaction-monitoring mode (collision energy, 14.0 eV), transitions from *m/z* 174.10 to 130.10
200 for [$^{12}C_6$] and *m/z* 180.20 to 136.15 for [$^{13}C_6$] were monitored. A series of standard mixtures
201 of [$^{13}C_6$] and unlabelled IAA in different ratios 10:1 to 1:10 were also assayed to confirm
202 accuracy of quantitative analysis. Data were obtained from the average of two technical
203 replicates first, then the average and the standard error of three biological replicates from
204 each developmental stage.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

205 **Starch assay**

206 Frozen wheat grain samples were freeze-dried then placed in microfuge tubes with stainless
207 steel beads and ground using a TissueLyser II (QIAGEN) for 3 min at a frequency of 30/s.
208 Starch extraction and analysis followed the methods of Zhao et al. (2010). After the
209 extraction of soluble sugars in 80% ethanol, the remaining starch pellet was hydrolysed
210 sequentially with amylase (Sigma A3403) and amyloglucosidase (Sigma A7095). The
211 resulting glucose was assayed using glucose HK assay reagent (Sigma G3293) in microtitre
212 plates using a SPECTROstar^{Nano} (BMG LABTECH) at 340 nm. Three biological replicates
213 were extracted for each time period; three technical replicates were assayed for each sample.

214 **Results**

215 **Wheat has orthologues of OsTAR1 and OsTAR2 as well as a wheat-specific**
216 **branch of TaTAR2**

217 A BlastP search of the wheat proteome found 15 co-orthologues of Arabidopsis TAR2. Their
218 encoding genes are listed in Table 1, with IWGSC RefSeq v1.1 gene IDs. Twelve of these
219 were previously named by Shao et al. (2017) as *TaTAR2.1* to *TaTAR2.5* with suffixes
220 indicating chromosome/genome. We followed a similar format, naming the three additional
221 genes as group *TaTAR2.6*. Two of these are tandem repeats designated as *TaTAR2.6-1Ba* and
222 *TaTAR2.6-1Bb* whereas the chromosomal location of the third *TaTAR2.6-U* is currently
223 unknown. *TaTAR2.1, TaTAR2.2* and *TaTAR2.3* have one copy in each of the A, B and D
224 genomes. On the other hand, *TaTAR2.4*, *TaTAR2.5* and *TaTAR2.6* do not have a copy on the
225 D genome.

226 The protein phylogenetic tree in Fig. 1 compares wheat TAR sequences with rice and
227 Arabidopsis enzymes that have demonstrated tryptophan aminotransferase activity. The tree
228 also has a branch containing three TaTAR3 proteins orthologous to OsTAR3. These are
229 usually designated as alliinases, and unlikely to be involved in IAA production. Three
230 proteins designated as TaTAR2.3-1A/B/D are co-orthologues of OsTAR1/OsFBL in rice.
231 Barley and Brachypodium also have at least one protein in this clade. A second clade with a
232 high bootstrap value contains OsTAR2/OsFIB as well as six closely related wheat proteins,
233 TaTAR2.1-3A/B/D and TaTAR2.2-1A/B/D and one protein each from barley and

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

234 Brachypodium. The remaining six wheat proteins have been placed in the same clade as
235 OsTAR1 and TaTAR2.3 but with a low bootstrap value indicating ambiguity. This branch
236 has no proteins from the other cereals.

237 Expression of *TaTAR2.3-1B* in grains increases from 5 to 15 DAA

238 We investigated the expression of all *TaTAR2* genes at four times during wheat grain fill, 5,
239 10, 15 and 20 DAA. Due to the large number and high similarity of genes from each genome,
240 six primer sets were designed, each to amplify all genes from sets *TaTAR2.1* to *TaTAR2.6* as
241 shown in Supplementary Table S1. Initial RT-PCR analysis showed efficient amplification of
242 a single band of the expected size for the *TaTAR2.3* group only. In contrast, *TaTAR2.1*,
243 *TaTAR2.2* and *TaTAR2.5* groups showed very low amplification and there was no
244 amplification of *TaTAR2.4* and *TaTAR2.6* groups. We therefore investigated the expression
245 of *TaTAR2.3* by qRT-PCR. A large (44-fold) up-regulation occurred between 5 and 15 DAA
246 after which expression appeared to plateau (Fig. 2a). The location of *TaTAR2.3* expression
247 within grains at 15 DAA was also investigated by dissection of the grains into embryo and
248 endosperm components (Fig. 2b). This revealed endosperm-specific gene activity.
249 Sequencing of the PCR product from 15 DAA samples indicated the B genome copy of
250 *TaTAR2.3* was the primary gene expressed.

251 Additional information on *TaTAR2* expression, from RNA-seq data in the expVIP
252 database is shown in Fig. 3. This confirmed *TaTAR2.3-1B* is the most highly expressed gene
253 in wheat grains, with maximum up-regulation between 10 and 20 DAA. The D genome copy
254 of *TaTAR2.3* was also expressed but at much lower levels. Results from manually dissected
255 tissue layers (Pfeifer et al. 2014), indicated that expression occurred across all parts of the
256 endosperm including starchy endosperm, aleurone layer and transfer cells. In addition,
257 *TaTAR2.5-1B*, *TaTAR2.2-1A* and *TaTAR2.2-1D* are expressed very early in grain
258 development at 2 DAA, followed by a decrease in expression.

259 Comprehensive phylogenetic analysis of YUCCA proteins and systematic
260 naming of wheat YUCCAs

261 The wheat genome contains 35 members of the YUCCA gene family. These are listed in
262 Table 2 along with their chromosomal location and IWGSC RefSeq v1.1 gene IDs. As there

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

263 has been no previous comprehensive study of the YUCCA gene family in wheat, we have
264 named these systematically according to their homology with rice *YUCCA* genes following
265 naming structure in the Wheat Gene Catalogue (https://wheat.pw.usda.gov/GG3/wgc). We
266 suggest renaming the TaYUC10 group, previously designated as *TaYUC10.1, TaYUC10.2*
267 and *TaYUC10.3* by Li et al. (2014) as *TaYUC10-B1, TaYUC10-A* and *TaYUC10-D*,
268 respectively. In addition, Tuan et al. (2019) referred to TraesCS5B02G216000 as *TaYUC11*
269 by homology to Arabidopsis *YUC11*. We suggest it is more logical to refer to this gene as
270 *TaYUC9-B* as it is orthologous to *OsYUC9,* and there are other wheat genes orthologous to
271 *OsYUC11*.

272 Figure 4 shows the phylogenetic relationships between wheat and rice YUCCA
273 proteins. The tree comprises two major clades; the larger clade I contains OsYUC1-8 as well
274 as orthologous wheat proteins. OsYUC2 is missing from the tree as *OsYUC2* appears not to
275 be expressed in rice and it also has no wheat orthologues. Most of the TaYUC sequences in
276 clade I are encoded by each of the A, B and D genomes. However, TaYUC-6 is found in the
277 A and B genomes only. The smaller clade II contains OsYUC9, OsYUC11, OsYUC12 and
278 OsYUC14 (*OsYUC10* and *OsYUC13* being probable pseudogenes) as well as 12 wheat
279 proteins. Wheat has five proteins in the OsYUC9 branch. Three of these, TaYUC9-A1,
280 TaYUC9-B and TaYUC9-D1 are co-orthologous with OsYUC9 and we refer to these are the
281 TaYUC9-1 group. Two less similar proteins in this branch have been named as TaYUC9-A2
282 and TaYUC9-D2. Wheat also has three co-orthologues of OsYUC11; TaYUC11-A1,
283 TaYUC11-D and TaYUC11-A2. In addition, there is a group of four proteins designated as
284 TaYUC10, three of which were previously reported by Li et al. (2014). TaYUC10 proteins
285 have highest amino acid similarity to OsYUC11 but are also similar to OsYUC12/14; this
286 agrees with their positioning in the tree. Due to their importance in grain development, the
287 relationship between proteins in clade II was further investigated with the addition of
288 homologous proteins from Brachypodium and barley (Supplementary Figure S1). This
289 revealed that both species have at least one orthologue of OsYUC9 and OsYUC11 but the
290 OsYUC12/14 branch does not have orthologues in wheat, barley or Brachypodium, all
291 members of the Pooideae. Instead, these cereals have a separate branch including TaYUC10,
292 which has undergone considerable gene duplication within each species, particularly barley.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Expression of *TaYUC9-1* and *TaYUC10* in developing grains increases from 5 to 15 DAA

Figure 3 summarises information from RNA-seq data (expVIP database), on the expression throughout the plant as well as during grain development of *TaYUC* genes. None of the *TaYUC* genes in clade I has significant expression in wheat grains. Some expression occurs in leaves and the spike at anthesis. However, the greatest expression of genes in this clade is in the roots, where *TaYUC3, TaYUC6* and *TaYUC7* are the most highly expressed gene groups, with expression of genes from all three genomes. Conversely, a number of clade II genes show very high expression in grains. Greatest up-regulation is found between 10 and 20 DAA, with *TaYUC9-A1*, *TaYUC9-B*, *TaYUC9-D1*, *TaYUC10-A* and *TaYUC10-D* being the genes with highest activity. A different gene, *TaYUC9-D2* has high expression at 2 DAA in grains, after which it is down-regulated, with no observable expression at 10 DAA. Expression of the *TaYUC10* group is restricted to the grains. However, the *TaYUC9-1* group is also active in vegetative tissue, particularly the stem and roots. *TaYUC9-D2* is the gene with highest activity in leaves and spike.

As gene expression in the developing grains was the primary focus of this study, we carried out a quantitative expression analysis of *TaYUC* genes in clade II from 5 to 20 DAA. Primers were designed to amplify genes from all three genomes as shown in Supplementary Table S1. Initial RT-PCR screening demonstrated successful amplification of the *TaYUC9-1* and *TaYUC10* groups, producing a single product of the correct size. On the other hand, no product was obtained for *TaYUC9-2* and *TaYUC11* confirming that these genes have very little or no expression in grains from 5 to 20 DAA. Investigation by qRT-PCR showed strong up-regulation of the *TaYUC9-1* and *TaYUC10* groups during grain development, as shown in Fig. 2a. The *TaYUC10* genes were up-regulated earlier than the *TaYUC9-1* group, reaching maximum expression at 10 DAA and decreasing activity at 20 DAA. In contrast, the expression of *TaYUC9-1* was primarily up-regulated between 10 and 15 DAA, and showed no reduction in expression at 20 DAA. The location of expression of *TaYUC9-1* and *TaYUC10* groups was also investigated following dissection of wheat grains into the embryos and endosperms. The results in Fig. 2b suggest *TaYUC9-1* genes are expressed in both tissues with higher expression in embryos, whereas *TaYUC10* genes are only expressed in endosperm. Examination of the RNA-seq data from dissected tissue layers provides further

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

324 information on localisation of *TaYUC* gene expression. These data indicate that highest
325 expression of *TaYUC9-1* occurs in the aleurone and transfer cells, whereas *TaYUC10* has the
326 highest expression in the starchy endosperm. Nevertheless, expression of both genes appears
327 to occur throughout the endosperm. No RNA-seq data is available on expression in wheat
328 embryos of the Chinese Spring variety. However, data from the variety Azhurnaya (not
329 shown) confirm our observations of expression of *TaYUC9-1* but not *TaYUC10* in this tissue.

330 Sequencing of the RT-PCR products was used to clarify which copy(s) of the
331 *TaYUC9-1* and *TaYUC10* groups were active. These results indicated maximum expression
332 of *TaYUC9-A1* or *TaYUC9-D1* and *TaYUC10-D* but suggested significant expression of the
333 other genome copies of both genes. Taken with the RNA-seq data, it would appear that
334 *TaYUC9-A1*, *TaYUC9-B*, *TaYUC9-D1, TaYUC10-A* and *TaYUC10-D* are all strongly up-
335 regulated during grain fill.

336 **IAA production coincides with increases in grain weight and starch content**
337 **from 5 to 15 DAA**

338 To investigate the relationship between expression of *TaTAR2* and *TaYUC* genes and the
339 auxin content of the developing grains, we measured IAA by combined liquid
340 chromatography-tandem mass spectrometry in multiple reaction monitoring mode (LC-
341 MS/MS MRM) with [^{13}C] IAA as the internal standard in extracts of grains from 5 to 20
342 DAA. As shown in Fig. 2c, the IAA content increased more than 30-fold during grain
343 development with the largest increase occurring between 5 and 15 DAA. Figures 2d and 2e
344 illustrate the grain fresh weight (FW) and starch content respectively, at the same
345 developmental stages. Both parameters also increased maximally between 5 and 15 DAA,
346 coinciding with the major increase of IAA in developing wheat grains.

347 **Discussion**

348 **Phylogenetic analysis of all *TAR* and *YUCCA* genes in wheat**

349 The TAR/YUCCA pathway has been demonstrated as the main route of IAA synthesis in
350 plants (Mashiguchi et al. 2011; Won et al. 2011). TAR and YUCCA activity has also been
351 reported as the major source of IAA in maize and rice cereal grains (Bernardi et al. 2012;
352 Chourey et al. 2010; Nonhebel and Griffin 2020). However, there has been no comprehensive

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

353 study of these genes in wheat. Additionally, IAA is widely cited as a key signalling
354 component during grain development. Thus, there is a need to understand its production in
355 this important cereal crop. Finally, the frequently cited but anomalous work on the *TGW6*
356 gene (Hu et al. 2016a; Ishimaru et al. 2013) in wheat and rice has effectively ignored this
357 source of IAA in proposing that the activity of an IAA-glucose hydrolase can regulate the
358 IAA content of grains.

359 Here, we present the first complete list and phylogenetic analysis of YUCCA
360 genes/proteins in wheat. We have named these genes based on their homology to rice
361 YUCCAs, rather than in comparison to Arabidopsis proteins (as was done with the previously
362 named TaYUC10 group). YUCCA proteins are quite divergent and there are not
363 unambiguous orthologous relationships between cereal YUCCAs and their homologues in
364 dicots. On the other hand, most rice YUCCAs have conserved homologues in other grasses,
365 with bootstrap values indicating high reliability of relationships (Russell French et al. 2014).
366 It is therefore likely that the sub-functions of YUCCAs may be conserved within the cereals.
367 The systematic naming of all genes in a family rather than *ad hoc* naming of individual genes
368 is important to avoid future confusion. In naming genes, we have followed the convention in
369 the Wheat Gene Catalogue of appending the genome label to each name. In line with this we
370 have renamed the TaYUC10 group, previously described by Li et al. (2014) as *TaYUC10-B1,*
371 *TaYUC10-A* and *TaYUC10-D* with the addition of *TaYUC10-B2*.

372 In clade I, each of the functional rice YUCCA sequences has at least one close
373 homologue in wheat, suggesting conservation of function. Most (OsYUC1, OsYUC3,
374 OsYUC4, OsYUC5 and OsYUC8) have a triplet of co-orthologous proteins. However,
375 OsYUC7 has six co-orthologues in wheat suggesting a gene duplication event in a wheat
376 progenitor. In clade II, the situation is more complex. Both OsYUC9 and OsYUC11 have
377 close homologues in wheat, barley and Brachypodium. This is similar to maize and sorghum
378 (*Sorghum bicolor*) (Russell French et al. 2014). However, the YUC9 group has duplicated in
379 wheat and Brachypodium; the TaYUC9-1 group is most similar to OsYUC9 and its
380 expression profile is also similar to the rice gene whereas the TaYUC9-2 group is more
381 divergent. The TaYUC10 group appears to form a separate Pooid branch that has undergone
382 gene expansion, with multiple close homologues on separate chromosomes in wheat (4A, 5B,
383 5D and 6B) and Brachypodium, as well as apparent tandem repeats in Brachypodium; barley

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

384 has 10 proteins in this group. The phylogeny of the group is somewhat ambiguous with low
385 bootstrap values and different programmes (Maximum Likelihood versus Neighbour Joining)
386 variously placing the group with OsYUC11 or OsYUC12. Careful pairwise comparison of
387 the wheat sequences with both OsYUC11 and OsYUC12 supports the maximum likelihood
388 tree shown, placing the group closest to OsYUC11 rather than OsYUC12 and OsYUC14.
389 This is interesting, as one of us had previously noted OsYUC12 has orthologues in maize and
390 sorghum with similar expression profiles (Russell French et al. 2014).

391 Our phylogenetic analysis of TAR proteins in wheat updates the previous study by
392 Shao et al. (2017), with the addition of three new TaTAR2 proteins, TaTAR2.6-1Ba,
393 TaTAR2.6-1Bb and TaTAR2.6-U. These new proteins, with TaTAR2.4-7A, TaTAR2.5-1A
394 and TaTAR2.5-1B form a separate branch, with no orthologues in rice, barley or
395 Brachypodium. This group of enzymes may have distinct sub-function in wheat.

396 Expression of *TaTAR2.3-1B, TaYUC9-1* and *TaYUC10* correlates with 397 increasing IAA content during grain fill from 5 to 15 DAA

398 Results from both qRT-PCR and analysis of RNA-seq data confirmed the observation of
399 Shao et al. (2017) that *TaTAR2.3-1B* is the most highly expressed tryptophan
400 aminotransferase in developing wheat grains. Strong up-regulation occurred during early
401 grain fill between 5 and 15 DAA coinciding with a similar increase in IAA content of the
402 grains over the same interval. This suggests that TaTAR2.3 is primarily responsible for
403 catalysing the first step in IAA synthesis of developing grains. A similar increase in
404 expression of orthologous genes *ZmTAR1* and *OsTAR1/OsFBL* in developing grains of maize
405 and rice also coincides with the major increase in IAA content, indicating conservation of
406 expression of this clade (Abu-Zaitoon et al. 2012; Chourey et al. 2010). In wheat, expression
407 of *TaTAR2.3-1B* appears to be specific to the developing grains, with no expression
408 detectable in other parts of the plant. This may represent additional subfunctionalisation due
409 to the larger number of TAR genes in wheat compared to other cereals.

410 The expression of *TaTAR2.2-1A, TaTAR2.2-1D* and *TaTAR2.5-1B* at 2 DAA,
411 followed by their rapid down-regulation suggests IAA produced during very early grain
412 development is regulated separately from that during the grain fill period. This is similar to
413 rice where *OsTAR2* appears to be the dominant gene expressed at 1 DAA, whereas *OsTAR1*

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

414 is expressed later (Abu-Zaitoon et al. 2012). The promotive effect of TAR and IAA for grain
415 fill in wheat is supported by the positive effect on grain yield of *TaTAR2.1-3A* overexpression
416 (Shao et al. 2017). In addition, a mutant of the orthologous gene, *tsg1,* has a negative effect
417 on grain size in rice (Guo et al. 2019).

418 Wheat genes positioned with *OsYUC1-8* in clade I are mostly expressed in vegetative
419 and/or floral tissues with little or no expression in grains. *OsYUC1-8* are also active in similar
420 tissues of rice plants with low expression in grains (Abu-Zaitoon et al. 2012; Yamamoto et al.
421 2007; Zhao et al. 2013). On the other hand, clade II genes *TaYUC9-A1*, *TaYUC9-B*, *TaYUC9-*
422 *D1*, *TaYUC10-A* and *TaYUC10-D* are all highly expressed during grain fill. These results
423 confirm data from Li et al. (2014) relating to *TaYUC10-D* (previously *TaYUC10.3*) but
424 expression of *TaYUC9-A1*, *TaYUC9-B* or *TaYUC9-D1* during grain fill has not been
425 described previously. The importance of clade II YUCCA for the production of IAA in cereal
426 grains is demonstrated by similar data from maize and rice (Abu-Zaitoon et al. 2012;
427 Chourey et al. 2010). Interestingly, ZmYUC1 appears to be primarily responsible for IAA
428 production in developing maize grains (Chourey et al. 2010), whereas three genes, *OsYUC9,*
429 *OsYUC11* and *OsYUC12* are all up-regulated in rice grains. These rice genes have differences
430 in their expression profiles, with earlier and more restricted up-regulation of *OsYUC12*
431 (Nonhebel and Griffin 2020). A similar situation may occur in wheat; *TaYUC9-1* genes were
432 maximally up-regulated between 10 and 15 DAA and remained active at 20 DAA, similar to
433 the orthologous *OsYUC9* (Nonhebel and Griffin 2020). Expression of the *TaYUC9-1* group
434 also increased in parallel with *TaTAR2.3-1B* and correlated most closely with the increasing
435 IAA content of grains suggesting that these genes may be primarily responsible for IAA
436 production during grain fill and similar again to the situation in rice. On the other hand,
437 *TaYUC10* was up-regulated earlier than *TaYUC9-1* with a maximum increase between 5 and
438 10 DAA, and a decline in expression at 20 DAA. This expression profile is similar to that of
439 *OsYUC12* in rice. Although *TaYUC10* and *OsYUC12* are not directly orthologous, their
440 expression profile distinct from that of *TaYUC9-1* suggests a requirement for very precise
441 and localised IAA production regulated by separate genes during early grain fill, as suggested
442 by Nonhebel and Griffin (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

443 **Localisation of IAA production at 15 DAA in developing wheat grains**

444 IAA production occurs mostly in endosperm of rice (Abu-Zaitoon et al. 2012; Russell French
445 et al. 2014) and maize (Bernardi et al. 2019; Chourey et al. 2010). Forestan et al. (2010)
446 observed high expression of the auxin transporter gene *ZmPIN1* in the periphery of maize
447 endosperm and the embryo surrounding region. Furthermore, treatment with the auxin
448 transport inhibitor, naphthylphthalamic acid (NPA) during early embryogenesis caused
449 developmental abnormalities in the maize embryo. It is widely assumed, therefore, that the
450 endosperm is the major source of IAA for the early embryo. Our results for wheat indicate
451 that *TaTAR2.3-1B* and the *TaYUC10* group are similarly expressed only in wheat endosperm.
452 However, we found high expression of *TaYUC9-1* in the embryo at 15 DAA, with lower
453 expression in the endosperm. As this result was novel, the experiment was carefully checked
454 and repeated. The observation was also confirmed by RNA-seq data from a different variety
455 of wheat. It is probable therefore, that the embryo is only dependent on IAA imported from
456 the endosperm at the very early stages of development.

457 Several reports suggest IAA may be important in aleurone layer and transfer cell
458 development and crucial for nutrient uptake into the grains. Forestan et al. (2010) showed
459 high immunolocalisation signal of IAA in the aleurone of maize kernels and use of NPA
460 interfered with normal development of the aleurone. Recent work by Bernardi et al. (2019)
461 suggests IAA has a role in the formation of the basal endosperm transfer cell layer (BETL) in
462 maize. This is supported by the poor grain fill of auxin-deficient *de18* and *dek18* mutants of
463 maize (Bernardi et al. 2012; Bernardi et al. 2016). Interestingly, *TaYUC-9.1* genes are most
464 highly expressed in the transfer cells and aleurone layer of wheat grains, with lower
465 expression in the starchy endosperm. This indicates the importance of IAA in these key cell
466 layers is conserved between cereals and warrants further direct investigation.

467 **IAA production by TAR/YUCCA versus grain weight genes *TaTGW6* and**
468 ***TaTGW-7A***

469 Our results have demonstrated high activity of the TAR/YUCCA pathway for IAA
470 biosynthesis in developing wheat grains, coinciding with a large increase in the IAA content
471 of grains. As in maize and rice, IAA in wheat grains accumulates particularly during early
472 grain fill and coincides with the initiation of starch production. In other seeds, it has been

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

473 proposed to initiate the differentiation of endosperm transfer cells and starch production
474 (Bernardi et al. 2019; McAdam et al. 2017; Nonhebel and Griffin 2020). On the other hand, a
475 negative effect of IAA on grain fill has been reported in publications relating to the *TGW6*
476 genes in rice (Ishimaru et al. 2013) and wheat (Hu et al. 2016a), and *TaTGW-7A* in wheat (Hu
477 et al. 2016b). Although our study did not investigate the function of these genes, it does raise
478 some important questions. Most obviously, it is unclear how an inactive IAA glucose-
479 hydrolase gene could have a major effect on the IAA content of grains when the
480 TAR/YUCCA pathway is strongly up-regulated. Although no mutants of these *TaTAR2.3,*
481 *TaYUC9-1* and *TaYUC10* are available in wheat to verify that they regulate IAA content,
482 defective kernel mutants of maize with reduced ZmYUC1 activity demonstrate the
483 importance of this pathway in cereal grain fill. Further, it is clear that there are strong
484 similarities between IAA production in rice, maize and wheat as discussed earlier.

485 In addition, data on the IAA content of wheat grains with different alleles of *TaTGW6*
486 and *TaTGW-7A* do not provide strong support for their role in IAA production either.
487 Measurements of the IAA content of grains reported by Hu et al. (2016a) used an HPLC
488 system with a UV absorbance detector set to 250 nm. This is an inadequate method for
489 determining hormone contents, lacking both sufficient specificity to be sure it is measuring
490 the correct compound as well as internal standard required to account for loss of analyte
491 during sample work-up. In addition, the amounts of IAA found at 20 DAA, the stage at which
492 maximum effect of the inactive genes is reported to occur, were widely variable. Values of
493 IAA content reported in wheat varieties with the active *TaTGW6-a* allele varied from 1300–
494 4700 ng/g FW compared to about 800–1200 ng/g FW in varieties with the inactive or less
495 active b and c alleles (values estimated from graphs in Hu et al. (2016a)). Similar variation
496 was reported in the *TaTGW-7A* paper (Hu et al. 2016b). Large variation was also seen in the
497 mean grain size from wheat varieties with active alleles of both *TaTGW6* and *TaTGW-7A*;
498 this variation bears no relation to the IAA content. Finally, it is worth noting that neither gene
499 in wheat has been experimentally characterised. TaTGW6 is assumed to hydrolyse IAA-
500 glucose based on a report on the rice gene product (Ishimaru et al. 2013). Although the
501 authors refer to *TaTGW-7A* as an “IGPS-like gene” its product is not actually a homologue of
502 IGPS in other plants. In fact, the authors merely state that “the deduced amino acid sequences
503 showed the presence of highly conserved TIM-br_sig_trns and UPF0261 functional domains,
504 forming a TIM barrel fold structure, the same as IGPS”. IGPS is an essential enzyme required

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

505 for tryptophan synthesis; the wheat proteome has three sequences (encoded by
506 TraesCS2A02G335200, TraesCS2D02G329500, TraesCS2B02G348500) with between 54%
507 and 56% amino acid identity to the experimentally characterised Arabidopsis IGPS (Li et al.
508 1995). These genes are expressed in the wheat plant including grains for the purpose of
509 tryptophan production. None of these wheat sequences nor the Arabidopsis IGPS encoded by
510 AT2G04400 has significant homology to *TaTGW-7A,* which is very unlikely to be an indole-
511 3-glycerol phosphate synthase.

512 **Conclusion**

513 In this study, we have demonstrated that auxin biosynthesis genes *TaTAR2.3-1B, TaYUC9-*
514 *A1, TaYUC9-B, TaYUC9-D1, TaYUC10-A* and *TaYUC10-D* are highly active in wheat grains
515 from 10 to 20 DAA. Their expression correlates positively with IAA levels, grain weight and
516 starch synthesis during major grain fill period in wheat. The gene expression data combined
517 with accurate measurements of grain IAA content from 5 to 20 DAA suggests the
518 TAR/YUCCA pathway is a major source of auxin in developing wheat grains. This raises
519 serious questions about how a mutation in *TGW6*, encoding a putative IAA-glucose hydrolase
520 could have the reported effects on grain IAA content. Our data on localisation of auxin
521 biosynthesis gene expression confirm that IAA is produced in wheat endosperm as in other
522 cereals. In addition, the novel finding of *TaYUC9-1* activity in embryos at 15 DAA suggests
523 that embryos have their own source of IAA by this stage of development. Differences in the
524 expression profiles of *TaYUC9-1* and *TaYUC10* groups suggest some sub-functionalisation.
525 The expression of *TaYUC9-1* genes in aleurone and transfer cells corroborates the suggestion
526 by Bernardi et al. (2019) that IAA may be involved in differentiation of these cells in maize.
527 This may partly explain the importance of IAA for grain fill.

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Abu-Zaitoon YM, Bennett K, Normanly J, Nonhebel HM (2012) A large increase in IAA during development of rice grains correlates with the expression of tryptophan aminotransferase *OsTAR1* and a grain-specific *YUCCA*. Physiologia Plant 146:487–499

Barkawi LS, Tam YY, Tillman JA, Pederson B, Calio J, Al-Amier H, Emerick M, Normanly J, Cohen JD (2008) A high-throughput method for the quantitative analysis of indole-3-acetic acid and other auxins from plant tissue. Anal Biochem 372:177–188

Bernardi J, Battaglia R, Bagnaresi P, Lucini L, Marocco A (2019) Transcriptomic and metabolomic analysis of ZmYUC1 mutant reveals the role of auxin during early endosperm formation in maize. Plant Sci 281:133–145

Bernardi J, Lanubile A, Li Q-B, Kumar D, Kladnik A, Cook SD, Ross JJ, Marocco A, Chourey PS (2012) Impaired auxin biosynthesis in the *defective endosperm18* mutant is due to mutational loss of expression in the *ZmYuc1* gene encoding endosperm-specific YUCCA1 protein in maize. Plant Physiol 160:1318–1328

Bernardi J, Li QB, Gao Y, Zhao Y, Battaglia R, Marocco A, Chourey PS (2016) The auxin-deficient *defective kernel18* (*dek18*) mutation alters the expression of seed-specific biosynthetic genes in maize. J Plant Growth Regul 35:770–777

Borrill P, Ramirez-Gonzalez R, Uauy C (2016) expVIP: a customizable RNA-seq data analysis and visualization platform. Plant Physiol 170:2172–2186

Chen KH, Miller AN, Patterson GW, Cohen JD (1988) A rapid and simple procedure for purification of indole-3-acetic acid prior to GC-SIM-MS analysis. Plant Physiol 86:822–825

Chourey PS, Li Q-B, Kumar D (2010) Sugar–hormone cross-talk in seed development: two redundant pathways of IAA biosynthesis are regulated differentially in the invertase-deficient miniature1 (*mn1*) seed mutant in maize. Mol Plant 3:1026–1036

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

FAOSTAT (2020) Food and Agriculture Organisation of the UN (FAO) statistics database: production, trade, supply. http://www.fao.org/faostat/en/#data/QC

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791

Forestan C, Meda S, Varotto S (2010) ZmPIN1-mediated auxin transport is related to cellular differentiation during maize embryogenesis and endosperm development. Plant Physiol 152:1373–1390

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Foulkes MJ, Slafer GA, Davies WJ, Berry PM, Sylvester-Bradley R, Martre P, Calderini DF, Griffiths S, Reynolds MP (2011) Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. J Exp Bot 62:469–486

Guo T, Chen K, Dong NQ, Ye WW, Shan JX, Lin HX (2019) *Tillering and small grain 1* dominates the tryptophan aminotransferase family required for local auxin biosynthesis in rice. J Integr Plant Biol 62:581–600

Hu MJ, Zhang HP, Cao JJ, Zhu XF, Wang SX, Jiang H, Wu ZY, Lu J, Cheng C, Sun GL, Ma CX (2016a) Characterization of an IAA-glucose hydrolase gene *TaTGW6* associated with grain weight in common wheat (*Triticum aestivum* L.). Mol Breed 36:1–11

Hu MJ, Zhang HP, Liu K, Cao JJ, Wang SX, Jiang H, Wu ZY, Lu J, Zhu XF, Xia XC, Sun GL, Ma CX, Chang C (2016b) Cloning and characterization of *TaTGW-7A* gene associated with grain weight in wheat via SLAF-seq-BSA. Front Plant Sci 7:1902

Ishimaru K, Hirotsu N, Madoka Y, Murakami N, Hara N, Onodera H, Kashiwagi T, Ujiie K, Shimizu B, Onishi A, Miyagawa H, Katoh E (2013) Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. Nature Genet 45:707–711

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8:275–282

Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, Christensen M, Davis P, Falin LJ, Grabmueller C, Humphrey J, Kerhornou A, Khobova J, Aranganathan NK, Langridge N, Lowy E, McDowall MD, Maheswari U, Nuhn M, Ong CK, Overduin B, Paulini M, Pedro H, Perry E, Spudich G, Tapanari E, Walts B, Williams G, Tello-Ruiz M, Stein J, Wei S, Ware D, Bolser DM, Howe KL, Kulesha E, Lawson D, Maslen G, Staines DM (2015) Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res 44:D574–D580

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874

Li N, Yin N, Niu Z, Hui W, Song J, Huang C, Wang H, Kong L, Feng D (2014) Isolation and characterization of three *TaYUC10* genes from wheat. Gene 546:187–194

Li J, Chen S, Zhu L, Last RL (1995) Isolation of cDNAs encoding the tryptophan pathway enzyme indole-3-glycerol phosphate synthase from *Arabidopsis thaliana*. Plant Physiol 108:877–878

Mashiguchi K, Tanaka K, Sakai T, Sugawara S, Kawaide H, Natsume M, Hanada A, Yaeno T, Shirasu K, Yao H, McSteen P, Zhao Y, Hayashi KI, Kamiya Y and Kasahara H (2011) The main auxin biosynthesis pathway in Arabidopsis. Proc Natl Acad Sci 108:18512–18517

McAdam EL, Meitzel T, Quittenden LJ, Davidson SE, Dalmais M, Bendahmane AI, Thompson R, Smith JJ, Nichols DS, Urquhart S, Gelinas-Marion A, Aubert G, Ross

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

JJ (2017) Evidence that auxin is required for normal seed size and starch synthesis in pea. New Phytol 216:193–204

Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. Nat Protoc 1:1559–1582

Nonhebel HM, Griffin K (2020) Production and roles of IAA and ABA during development of superior and inferior rice grains. Funct Plant Biol 47:716–726

Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M (2009) Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. BMC Mol Biol 10:11

Pfeifer M, Kugler KG, SandveSR, Zhan B, Rudu H, Hvidsten TR, Mayer KFX, Olsen OA, et al. (2014) International Wheat Genome Sequencing Consortium. Genome interplay in the grain transcriptome of hexaploid bread wheat. Science 345:1250091

Ramírez-González RH, Borrill P, Land D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A, Khedikar Y, Robinson SJ, Cory AT, Florio T, Concia L, Juery C, Schoonbeek H, Steuernagel B, Xiang D, Ridout CJ, Chalhoub B, Mayer KFX, Benhamed M, Latrasse D, Bendahmane A, Wulff BBH, Appels R, Tiwari V, Datla R, Choulet F, Pozniak CJ, Provart NJ, Sharpe AG, Paux E, Spannagl M, Bräutigam A, Uauy C (2018) The transcriptional landscape of polyploid wheat. Science 361:eaar6089

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res 31:224–228.

Russell French S, Abu-Zaitoon Y, Uddin M, Bennett K, Nonhebel H (2014) Auxin and cell wall invertase related signaling during rice grain development. Plants 3:95-112

Shao A, Ma W, Zhao X, Hu M, He X, Teng W, Li H, Tong Y (2017) The auxin biosynthetic *TRYPTOPHAN AMINOTRANSFERASE RELATED TaTAR2.1-3A* increases wheat grain yield. Plant Physiol 174:2274–2288

Tuan PA, Yamasaki Y, Kanno Y, Seo M, Ayele BT (2019) Transcriptomics of cytokinin and auxin metabolism and signaling genes during seed maturation in dormant and non-dormant wheat genotypes. Sci Rep 9:3983

Won C, Shen X, Mashiguchi K, Zheng Z, Dai X, Cheng Y, Kasahara H, Kamiya Y, Chory J, Zhao Y (2011) Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in Arabidopsis. Proc Natl Acad Sci 108:18518–18523

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Wu D, Dong J, Yao Y, Zhao W, Gao X (2015) Identification and evaluation of endogenous control genes for use in quantitative RT-PCR during wheat (*Triticum aestivum* L.) grain filling. Genet Mol Res 14:10530–10542

Yamamoto Y, Kamiya N, Morinaka Y, Matsuoka M, Sazuka T (2007) Auxin biosynthesis by the *YUCCA* genes in rice. Plant Physiol 143:1362–1371

Zhao D, MacKown CT, Starks PJ, Kindiger BK (2010) Rapid analysis of nonstructural carbohydrate components in grass forage using microplate enzymatic assays. Crop Sci 50:1537–1545

Zhao Z, Zhang Y, Liu X, Zhang X, Liu S, Yu X, Ren Y, Zheng X, Zhou K, Jiang L, Guo X, Gai Y, Wu C, Zhai H, Wang H, Wan J (2013) A role for a dioxygenase in auxin metabolism and reproductive development in rice. Dev Cell 27:113–122

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1 List of *TaTAR* genes in the wheat genome

Gene Name	Gene ID (IWGSC RefSeq v1.1)	Chromosome
TaTAR2.1-3A	TraesCS3A02G093000	3A
TaTAR2.1-3B	TraesCS3B02G108200	3B
TaTAR2.1-3D	TraesCS3D02G093300	3D
TaTAR2.2-1A	TraesCS1A02G233800	1A
TaTAR2.2-1B	TraesCS1B02G249600	1B
TaTAR2.2-1D	TraesCS1D02G238100	1D
TaTAR2.3-1A	TraesCS1A02G113400	1A
TaTAR2.3-1B	TraesCS1B02G133500	1B
TaTAR2.3-1D	TraesCS1D02G114800	1D
TaTAR2.4-7A	TraesCS7A02G179400	7A
TaTAR2.5-1A	TraesCS1A02G113600	1A
TaTAR2.5-1B	TraesCS1B02G133900	1B
TaTAR2.6-1Ba	TraesCS1B02G004300	1B
TaTAR2.6-1Bb	TraesCS1B02G004400	1B
TaTAR2.6-U	TraesCSU02G199800	U
TaTAR3.1-3A	TraesCS3A02G244700	3A
TaTAR3.1-3B	TraesCS3B02G275200	3B
TaTAR3.1-3D	TraesCS3D02G246700	3D

U-unknown. The bold gene IDs are three newly reported genes

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 2 List of *TaYUC* genes in the wheat genome

Gene Name	Gene ID (IWGSC RefSeq v1.1)	Chromosome
TaYUC1-A	TraesCS3A02G232600	3A
TaYUC1-B	TraesCS3B02G261900	3B
TaYUC1-D	TraesCS3D02G220500	3D
TaYUC3-A	TraesCS3A02G280900	3A
TaYUC3-B	TraesCS3B02G314700	3B
TaYUC3-D	TraesCS3D02G280900	3D
TaYUC4-A	TraesCS3A02G149500	3A
TaYUC4-B	TraesCS3B02G176800	3B
TaYUC4-D	TraesCS3D02G157600	3D
TaYUC5-A	TraesCS5A02G102700	5A
TaYUC5-B	TraesCS5B02G107000	5B
TaYUC5-D	TraesCS5D02G114500	5D
TaYUC6-A	TraesCS4A02G313200	4A
TaYUC6-B	TraesCS5B02G566700	5B
TaYUC7-A1	TraesCS2A02G011500	2A
TaYUC7-B1	TraesCS2B02G010100	2B
TaYUC7-D1	TraesCS2D02G012100	2D
TaYUC7-A2	TraesCS2A02G533200	2A
TaYUC7-B2	TraesCS2B02G562800	2B
TaYUC7-D2	TraesCS2D02G535100	2D
TaYUC8-A	TraesCS4A02G027500	4A
TaYUC8-B	TraesCS4B02G278300	4B
TaYUC8-D	TraesCS4D02G276600	4D
TaYUC9-A1	TraesCS5A02G217200	5A
TaYUC9-B	TraesCS5B02G216000	5B
TaYUC9-D1	TraesCS5D02G225200	5D
TaYUC9-A2	TraesCS7A02G065600	7A
TaYUC9-D2	TraesCS7D02G060000	7D
TaYUC10-A	TraesCS4A02G325400	4A
TaYUC10-B1	TraesCS5B02G538300	5B
TaYUC10-B2	TraesCS6B02G033700	6B
TaYUC10-D	TraesCS5D02G556600	5D
TaYUC11-A1	TraesCS4A02G373500	4A
TaYUC11-A2	TraesCS7A02G075400	7A
TaYUC11-D	TraesCS7D02G070900	7D

Genes have been numbered by homology to rice genes. Genes in bold were previously reported as *TaYUC10.2*, *TaYUC10.1* and *TaYUC10.3*, respectively by Li et al. (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

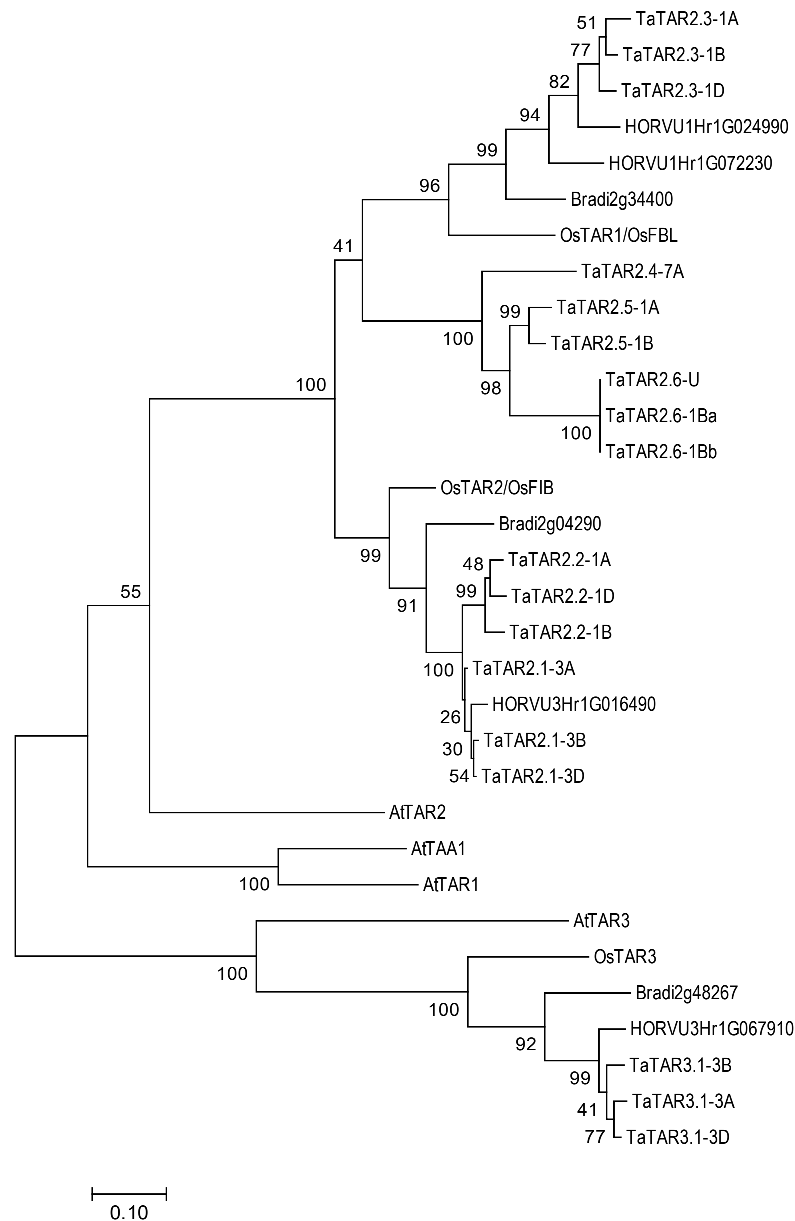

Fig. 1 Phylogenetic tree showing relationships between TAR proteins from *Triticum aestivum* (TaTAR), *Oryza sativa* (OsTAR), *Brachypodium distachyon* (Brad), *Hordeum vulgare* (HORVU) and *Arabidopsis thaliana* (At). The tree was constructed in MEGA7.0.26 (Kumar et al., 2016) using Maximum Likelihood method (Jones et al., 1992). Multiple sequence alignment was performed by MUSCLE (Edgar, 2004). Bootstrap confidence levels were obtained using 500 replicates (Felsenstein, 1985). Evolutionary distances were computed using Poisson correction method (Zuckerkandl and Pauling, 1965). Scale bar=0.10 amino acid substitutions per site

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

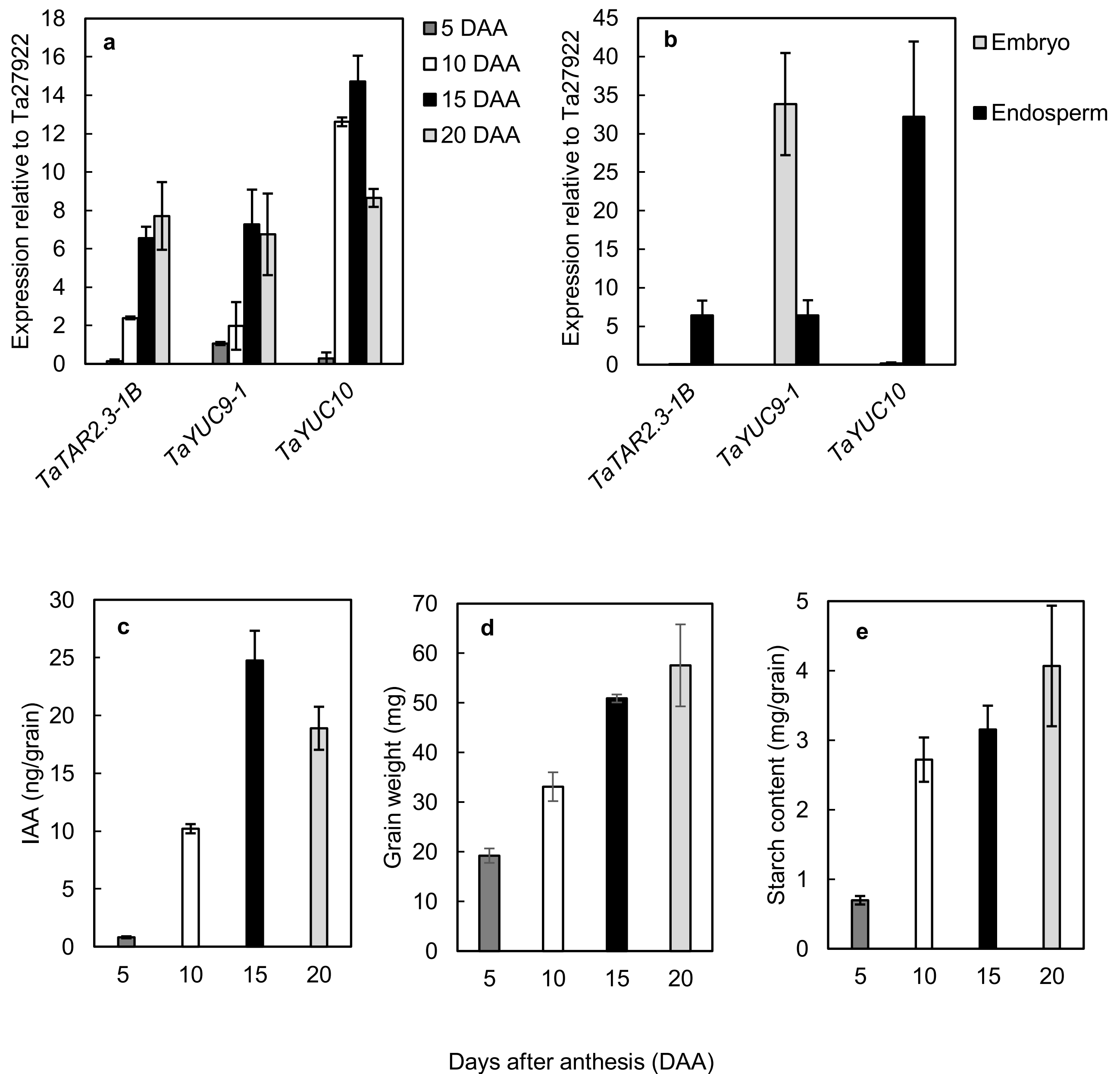

Fig. 2 Quantitative RT-PCR analysis of relative transcript abundance for TaTAR2.3, TaYUC9-1 and TaYUC10 during grain development, whole grains (a), in dissected embryos and endosperms at 15 DAA (b), IAA content in ng per grain measured by LC-MS/MS MRM with [^{13}C] IAA as the internal standard (c), grain fresh weight (d) and starch content (e). Expression relative to reference gene Ta27922 is shown, expression relative to the second reference gene Ta53964 showed the same trend. All data points represent the mean of three biological replicates ± the standard error of the means

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

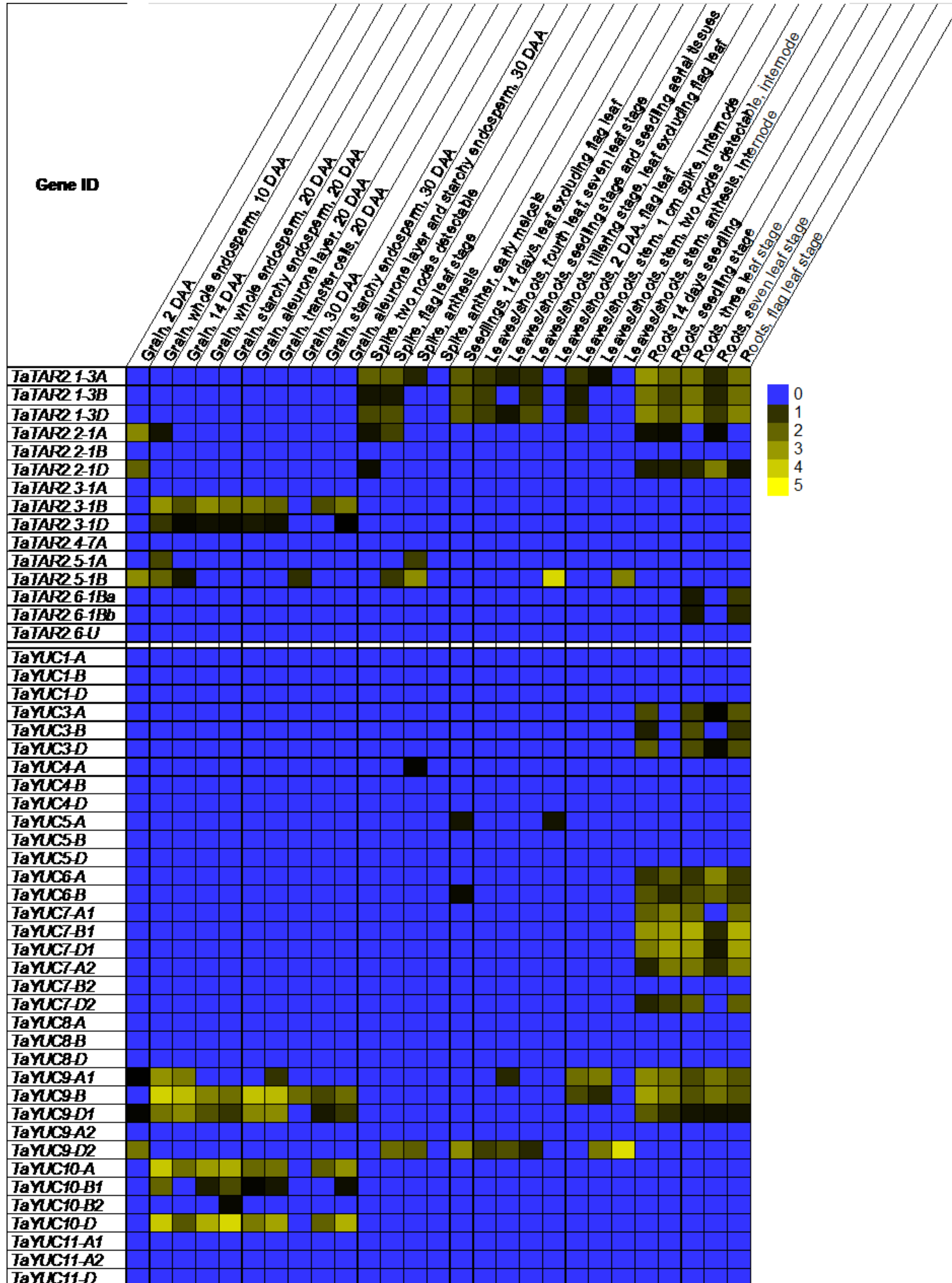

Fig. 3 Heat map depicting expression of *TaTAR2* and *TaYUC* genes from RNA-seq data in expVIP. Data were taken from different studies using the Chinese Spring variety. The relative expression values are normalized as tpm (transcripts per million)

bioRxiv preprint doi: https://doi.org/10.1101/2020.10.12.336560; this version posted October 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

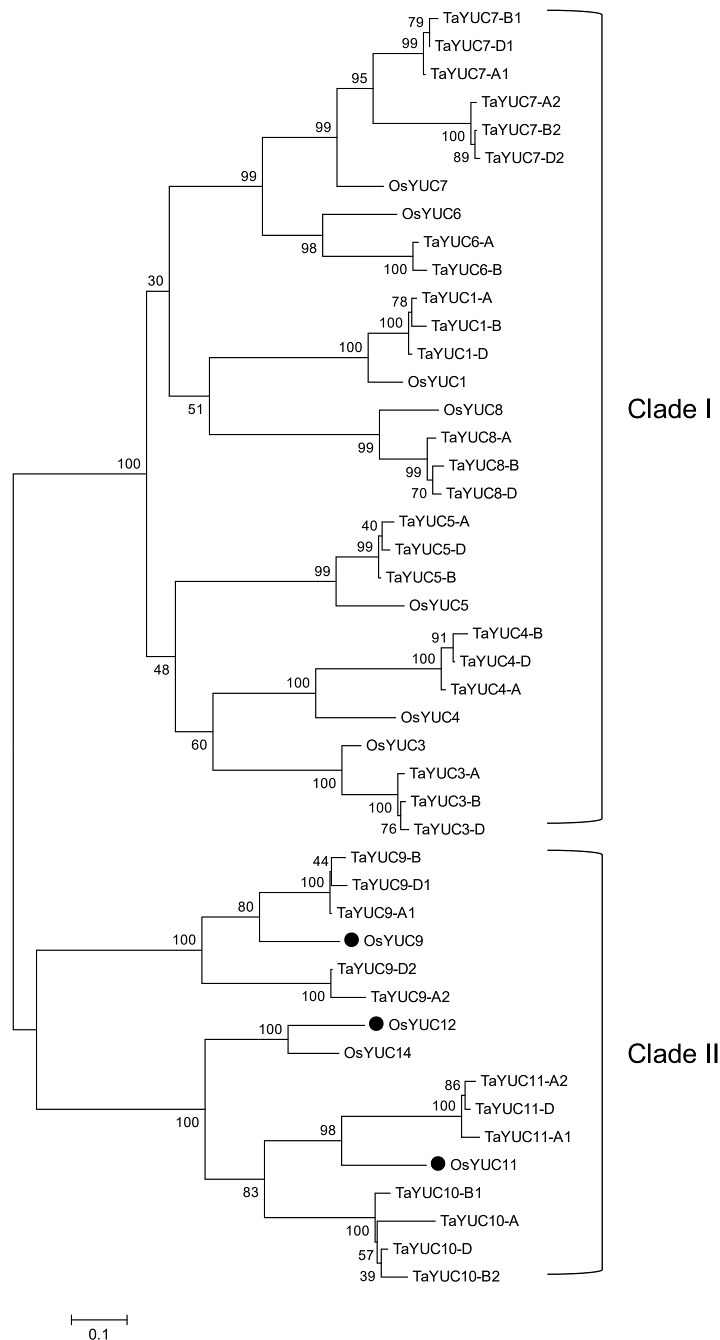

Fig. 4 Phylogenetic tree showing relationships between YUCCA proteins from *Triticum aestivum* (TaYUC) and *Oryza sativa* (OsYUC). The tree was constructed in MEGA7.0.26 (Kumar et al. 2016) using Maximum Likelihood method (Jones et al. 1992). Multiple sequence alignment was performed by MUSCLE (Edgar 2004). Bootstrap confidence levels were obtained using 500 replicates (Felsenstein 1985). Evolutionary distances were computed using Poisson correction method (Zuckerkandl and Pauling 1965). Scale bar=0.10, amino acid substitutions per site. Black dots are encoded by genes expressed in rice grains