

1 **Title page**

2

3 **Genome mapping of quantitative trait loci (QTL) controlling domestication traits of intermediate wheatgrass**
4 **(*Thinopyrum intermedium*)**

5

6 Steve Larson^{1*}, Lee DeHaan², Jesse Poland³, Xiaofei Zhang⁴, Kevin Dorn³, Traci Kantarski⁵, James Anderson⁶,
7 Jeremy Schmutz^{7,8}, Jane Grimwood⁸, Jerry Jenkins⁸, Shengqiang Shu⁷, Jared Crain³, Matthew Robbins¹, Kevin
8 Jensen¹

9

10 ¹United States Department of Agriculture, Agriculture Research Service, Forage and Range Research, Utah State
11 University, Logan, UT 84322, United States

12 ²The Land Institute, 2440 E. Water Well Rd, Salina, KS 67401, United States

13 ³Department of Plant Pathology, Kansas State University, 4024 Throckmorton, Manhattan, KS 66506, United States

14 ⁴Department of Horticultural Science, North Carolina State University, 212 Kilgore Hall, PO Box 7609, 2721
15 Founders Drive, Raleigh, NC, 27607, United States

16 ⁵American Association for the Advancement of Science, Science and Technology Policy Fellow at the United States
17 Department of Agriculture, Animal and Plant Health Inspection Service, 4700 River Road, Riverdale, MD 20737,
18 United States

19 ⁶Department of Agronomy and Plant Genetics, University of Minnesota, 1991 Buford Circle, St. Paul, MN 55108,
20 United States

21 ⁷Department of Energy, Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, United States

22 ⁸Hudson Alpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, United States

23

24 *Corresponding author: Steve.Larson@usda.gov (tel. 01 4357971703)

25

26 **ORCID list:**

27 Steve Larson (0000-0003-2742-2134), Lee DeHaan (0000-0002-6368-5241), Jesse Poland (0000-0002-7856-1399),
28 Xiaofei Zhang (0000-0003-4516-9179), Kevin Dorn (0000-0003-2722-6221), Traci Kantarski (0000-0002-8584-
29 2965), James Anderson (0000-0003-4655-6517), Jeremy Schmutz (0000-0001-8062-9172), Jane Grimwood (0000-
30 0002-8356-8325), Jerry Jenkins (0000-0002-7943-3997), Shengqiang Shu (0000-0002-4336-8994), Jared Crain
31 (0000-0001-9484-8325), Matthew Robbins (0000-0002-5467-4452), Kevin Jensen (0000-0001-5560-0394)

32

33 **Abstract**

34 Allohexaploid ($2n=6x=42$) intermediate wheatgrass (*Thinopyrum intermedium*), abbreviated IWG, is an outcrossing
35 perennial grass belonging to the tertiary gene pool of wheat. Perenniality would be valuable option for grain
36 production, but attempts to introgress this complex trait from wheat-*Thinopyrum* hybrids have not been
37 commercially successful. Efforts to breed IWG itself as a dual-purpose forage and grain crop have demonstrated
38 useful progress and applications, but grain yields are significantly less than wheat. Therefore, genetic and physical
39 maps have been developed to accelerate domestication of IWG. Herein, these maps were used to identify
40 quantitative trait loci (QTLs) and candidate genes associated with IWG grain production traits in a family of 266
41 full-sib progenies derived from two heterozygous parents, M26 and M35. Transgressive segregation was observed
42 for 17 traits related to seed size, shattering, threshing, inflorescence capacity, fertility, stem size, and flowering time.
43 A total of 111 QTLs were detected in 36 different regions using 3,826 genotype-by-sequence (GBS) markers in 21
44 linkage groups. The most prominent QTL had a LOD score of 15 with synergistic effects of 29% and 22% over the
45 family means for seed retention and percent naked seed, respectively. Many QTLs aligned to one or more IWG gene
46 models corresponding to 42 possible domestication orthogenes including the wheat *Q* and *RHT* genes. A cluster of
47 seed-size and fertility QTLs showed possible alignment to a putative *Z* self-incompatibility gene, which could have
48 detrimental grain-yield effects when genetic variability is low. These findings elucidate pathways and possible
49 hurdles in the domestication of IWG.

50

51 **Keywords:**

52 domestication, genetic mapping, genome sequence, perennial grains, QTL, *Thinopyrum intermedium*,

53 **Introduction**

54 Intermediate wheatgrass (*Thinopyrum intermedium*) is a cool-season perennial Triticeae grass, native to parts of
55 Eastern Europe and western Asia, that has been widely used for soil conservation and forage production in North
56 America and other temperate regions of the World (Jensen et al. 2016; Zair et al. 2018). Intermediate wheatgrass
57 (IWG) ranks among the highest-yielding forage and biomass crops that can be grown across regions of the Upper
58 Midwest, Great Plains and Intermountain regions of Canada, the United States, and other places where it is
59 cultivated (Harmony 2015; Kenneth and Kevin 2001; Larson et al. 2017; Lee et al. 2009; Monono et al. 2013;
60 Pearson et al. 2015; Robins 2010; Wang et al. 2014). At least 15 IWG cultivars have been released for purposes of
61 soil conservation and forage production in Canada and the United States, with some of the first cultivars selected
62 directly from plant introductions from Russia in 1932 (Jensen et al. 2016; Knowles 1977; Pearson et al. 2015).
63 However, selection for better fertility and seed yields was critical for development of the first widely successful
64 North American cultivar ‘Oahe’ (Knowles 1977; Ross 1963), which is still grown today. Modern IWG forage
65 cultivars have also undergone selection for disease resistance, forage quality, forage yield and other traits (Jensen et
66 al. 2016; Krupinsky and Berdahl 2000; Vogel et al. 2005).

67 Species of the genus *Thinopyrum*, including IWG, are considered the closest perennial relatives of common
68 wheat (*Triticum aestivum*) and comprise a useful part of its tertiary gene pool (Ceoloni et al. 2015; Lang et al. 2018;
69 Li et al. 2017; Liu et al. 2017; Uzma et al. 2015; Zair et al. 2018). Common wheat and IWG have similar
70 allohexaploid genomes ($2n=6x=42$). However, unlike wheat, IWG is usually self-incompatible meaning that most
71 populations are highly heterogeneous and most individuals are highly heterozygous (Jensen et al. 2016; Jensen et al.
72 1990; Kantarski et al. 2017; Zhang et al. 2016). Wheat-*Thinopyrum* hybrids have also been utilized for the
73 development of perennial grain crops (Curwen-McAdams and Jones 2017; Hayes et al. 2018), but commercial
74 production of these plant materials has not yet been realized because it has been difficult to introgress and stabilize
75 this complex trait in wheat. However, parallel efforts to directly domesticate IWG itself as a perennial grain crop
76 (Cox et al. 2006; Cox et al. 2010; Wagoner 1990) have led to small-scale production and utilization of IWG grain or
77 flour for baking, beverages, and other edible food products (DeHaan and Ismail 2017). Although current grain yields
78 are modest, IWG has potential to be used as a multipurpose perennial forage and grain crop (Bell et al. 2015; Cattani
79 and Asselin 2017; Jungers et al. 2017; Ryan et al. 2018). These burgeoning efforts to domesticate IWG (Cattani
80 2017; DeHaan et al. 2018; Zhang et al. 2016) aim to diversify and enhance the quantity and quality of food products
81 (DeHaan and Ismail 2017; Marti et al. 2016) while providing improved soil conservation, water quality, carbon
82 sequestration, nutrient management, and other ecosystem services (Culman et al. 2013). These goals have been
83 described as the “ecological intensification of agriculture” (DeHaan et al. 2018).

84 A high-density genotype-by-sequencing (GBS) consensus linkage map (Kantarski et al. 2017) and draft
85 genome sequence were developed to accelerate the domestication of IWG (DeHaan et al. 2018). The IWG
86 consensus map had three homoeologous sets of seven linkage groups (LG01 to LG21) with 10,029 GBS markers
87 showing colinear alignments to the seven chromosomes of diploid barley (Kantarski et al. 2017). The consensus
88 map was constructed using seven full-sib populations derived from 13 heterozygous individuals from the third (C3)
89 and fourth (C4) cycle of selection (DeHaan et al. 2018) including one population derived by self-pollination of one

90 outstanding C3 individual, C3_3471; a biparental population derived from two C4 parents, M26 and M35,
91 descended from C3_3471 and two other C3 grandparents; and another biparental population derived from two C4
92 parents, C4_2856 and C4_5353, also descended from C3_3471 and three other C3 grandparents (Kantarski et al.
93 2017). The C3_3471 individual was identified in the third cycle of selection as the first predominantly free-threshing
94 and non-shattering IWG plant, which also had exceptionally long and heavy seeds. The IWG draft genome sequence
95 was developed by sequencing a haploid twin-seedling (Namikawa and Kawakami 1934) from C4_5353, which was
96 also a parent of two mapping populations (Kantarski et al. 2017). The draft sequence includes 21 recognized
97 chromosome sequences ranging in size from 250.8 to 802.6 megabases (MB), one unrecognized chromosome of
98 849.9 MB, and 1,977 scaffolds ranging in size from 0.1 to 8.8 MB with a total of 159,905 annotated gene models.
99 The 21 chromosome sequences were numbered CHR01 to CHR21 according to alignments with the 21 LGs of the
100 IWG GBS consensus map (Kantarski et al. 2017). The IWG consensus map has been used to identify QTLs and
101 markers associated with seed size in two biparental populations, M26 x M35 and C3-2331 x C3-2595, and one
102 association mapping (AM) population (Zhang et al. 2017). However, the M26 x M35 population was developed
103 mainly to investigate the genetic control of the non-shattering and free-threshing seed traits of C3_3471.

104 The inflorescence of IWG is fundamentally like wheat in that they both have one sessile spikelet at each
105 rachis node and multiple florets per spikelet with up to one seed per floret and variable levels of fertility. Seed
106 disarticulation in wild IWG plants normally occurs in two different ways: 1) abscission above the junction of the
107 rachis and spikelet base producing a complex diaspore with a relatively long wedge-shaped rachis internode
108 extending away from the spikelet similar to that produced by the brittle rachis of barley (Pourkheirandish et al.
109 2015) and certain types of wheat that produce wedge-shaped diaspores (Li and Gill 2006), and 2) abscission below
110 the junction of the rachilla and floret base, within the spikelet, producing a less complex diaspore that includes a
111 short rachilla internode extending up from the base of the floret adjacent to the palea. Threshing of semi-
112 domesticated IWG forage varieties normally produces seeds with the lemma and palea attached to the pericarp, like
113 those of hulled barleys. However, free-threshing forms of IWG, such as C3_3471, and other domesticated forms of
114 wheat and barley produce naked grains comprised of the caryopsis detached from the lemma, palea, and spikelet
115 glumes. At least two different genes have been identified, the barley *nud* gene (Taketa et al. 2008) and wheat *Q* gene
116 (Simons et al. 2006), which are responsible for the free-threshing trait of naked barleys and common wheat,
117 respectively.

118 The domestication and improvement of IWG as a perennial grain crop currently focuses on increasing seed
119 size, fertility, inflorescence capacity, stem and inflorescence compactness, seed retention in the field, percentage of
120 naked seeds after threshing, and uniformity of maturity (DeHaan et al. 2018; DeHaan et al. 2016; Zhang et al. 2016).
121 One of the primary objectives of the study herein was to identify and map QTLs controlling seed shattering,
122 percentage of naked seeds after threshing and other traits related to reproductive fertility, inflorescence capacity,
123 stem and inflorescence compactness, seed retention in the field, and uniformity of maturity in the M26 x M35
124 population using the GBS consensus map (Kantarski et al. 2017). Another major objective of this study was to
125 compare M26 x M35 QTLs, C3-2331 x C3-2595 QTLs, and AM markers to the annotated IWG draft genome
126 sequence and identify possible candidate genes (CGs) controlling relevant domestication and improvement traits

127 (Doebley et al. 2006; Kovach et al. 2007; Lenser and Theißen 2013; Meyer and Purugganan 2013; Tang et al. 2010)
128 such as the percentage of free-threshing naked seeds (Simons et al. 2006; Taketa et al. 2008), seed shattering (Doust
129 et al. 2014; Li and Gill 2006; Pourkheirandish et al. 2015), and grain yield (Nadolska-Orczyk et al. 2017) traits in
130 wheat, barley, rice, and other grain crops.

131

132 **Materials and Methods**

133 **Plant materials and Field Evaluations**

134 The terms “parent”, “hybrid”, or “progeny” are used to describe genetically unique individuals, generally referred to
135 as “genets”, depending on context, whereas the terms “propagule” or “plot” may refer to different clonal replicates
136 of the same genet. A total of 266 full-sib progenies, two C4 parents (M26 and M35), and two known C3
137 grandparents (C3_3471 and C3_3941) comprising one of seven families used to construct a high-density linkage
138 map (Kantarski et al. 2017) were clonally propagated into replicated field plots in Kansas (KS) and Utah (UT).
139 Propagules were planted in grids with 1 m spacing in UT (1.0 m² plots) and 3 ft spacing in KS (0.81 m² plots).
140 Although these propagules eventually grow large enough to cover most plots, tillage between propagules was used
141 to control weeds and help maintain plot integrity. The C4 M26 genet was the maternal seed parent for 128
142 progenies, whereas the C4 M35 genet was the maternal seed parent for the other 138 progenies. The M26 genet
143 originates from a cross of C3_3471 as the seed parent and C3_3941 as the pollen parent in the third cycle of
144 selection for grain production traits (DeHaan et al. 2018; Zhang et al. 2016). The M35 genet originates from a cross
145 of C3_3941 as the seed parent and an unknown C3 pollen parent. Two propagules from each of 221 progenies, two
146 C4 parents, and two known C3 grandparents were transplanted at the KS location (38.771517° N / -97.569408° W)
147 on November 28, 2012. Three clones from each of 253 progenies, two parents, and two known grandparents were
148 transplanted into a field in UT (41.695957° N / -111.831358° W) on May 6, 2013.

149 A total of seventeen seed production traits were measured on each plot once per harvest year for at least
150 two years (2014 and 2015) at both locations (KS and UT) and most traits were also evaluated at the KS location in
151 2013. The KS location was planted early enough in 2012 to enable sufficient plant growth, vernalization, and
152 flowering to evaluate seed production traits in 2013. Conversely, the UT location was planted in the spring of 2013,
153 which did not allow sufficient growth, vernalization, flowering, and set seed to evaluate seed production traits in
154 2013.

155 Four traits measured in the field include crown circumference (CRCI), number of inflorescences per crown
156 (INCR), stem length (STLE), and one Zadoks (Zadoks et al. 1974) maturity rating (ZAMA) on each plot, once per
157 year. The CRCI was and INCR traits evaluated by measuring the distance (cm) around the outermost stems at the
158 base of each propagule, just above the soil surface, and then counting the number of inflorescences in each plot in
159 2014 at both locations. These traits were not evaluated in 2015 in KS because it was too difficult to discern the then
160 natural edges of different propagules. However, in UT, propagules were just beginning to grow into cultivated
161 spaces between rows in 2015, so we devised a faster and possibly more effective way to estimate CRCI and INCR.
162 First, we counted the number of stems through a 10-cm wide section of the widest undisturbed portion of each
163 crown using an open-ended rectangular quadrat with marked distances along each arm to measure the maximum

164 diameter (d) of the crown. Estimates of CRCI, from UT in 2015, were then determined based on the formula for the
165 circumference of a circle ($C = \pi \cdot d$). Estimates of INCR, from UT in 2015, were then determined using the formula
166 for the area of a circular crown [$A = \pi \cdot (0.5 \cdot d)^2$] multiplied by the tiller density [$S/(10 \cdot d)$]. The total (stretched)
167 STLE was measured (cm) from the soil surface to the upper tip of the inflorescence. The peduncle and inflorescence
168 from ten of the tallest culms were harvested from each plot for subsequent measurements of 13 seed and
169 inflorescence traits in the laboratory.

170 Five stem and inflorescence traits including the stem width (STWI), number of spikelets per inflorescence
171 (SPIN), inflorescence length (INLE), seed shattering (SESH), and number florets per spikelet (FLSP) were
172 measured on a subset of three of 10 harvested culms from each plot. The INLE was measured from the rachis node
173 subtending the lowest spikelet to uppermost point of the inflorescence. The SESH at the KS location was determined
174 based on the percentage of disarticulation after bending spikelets to 90° angle from the rachis and dropping
175 inflorescences from a height of 25 cm. The SESH at the UT location was determined by the percentage of
176 disarticulation after three repeated strikes of each stem to a table surface. The STWI was measured on the thinnest
177 part of each stem just below the inflorescence.

178 For UT field evaluations, six seed traits were measured after removing seed from ten harvested culms using
179 a LD 180 laboratory thresher (Wintersteiger Inc., Salt Lake City, Utah, USA). Seeds and chaff were further
180 separated using a General Seed Blower (Seedburo Equipment Co., Des Plaines, Illinois, USA). The percentage of
181 seeds threshed out naked (SENA), seed area (SEAR), seed length (SELE), seed width (SEWI), and total number of
182 seeds were determined using SmartGrain phenotyping software (Tanabata et al. 2012) from images of cleaned seeds
183 from all ten inflorescences taken from a scanning device with a blue-paper background. The SENA was determined
184 by first counting the number of hulled seeds and then counting the number naked seeds based on color recognition
185 of lighter colored hulls and darker colored pericarps, respectively. Measurements of SEAR, SELE, and SEWI were
186 also taken from the darker colored naked seeds, excluding broken parts less than 2.0 mm long clumps of seeds that
187 exceeded 8.0 mm length based on visual validation of SmartGrain image annotations. The total seed yield per
188 inflorescence (SYIN) and average seed mass (SEMA) were determined by dividing the mass of all cleaned seeds
189 (mg) by the number of inflorescences harvested (10) and the total number of seeds, respectively.

190 For KS field evaluations, six seed traits were measured after removing seed from ten harvested culms using
191 a spike-tooth small bundle thresher. Additional cleaning was performed by hand sieving with a 12/64 inch round
192 hole sieve and aspirating with an STS-WM2 Air Separator (Seed Tech Systems, Wilton, CA, USA). The percentage
193 of seed threshed out naked (SENA) was visually estimate for each sample. Approximately 20 naked seeds from each
194 sample were photographed and the images were analyzed in ImageJ (Schneider et al. 2012). Image analysis
195 provided seed area (SEAR), seed length (SELE), seed width (SEWI), and number of seeds. This subsample was
196 weighed and the seed number data was used to calculate average seed mass (SEMA). The total seed yield per
197 inflorescence (SYIN) was determined by dividing the mass of all cleaned seeds by the number of inflorescences
198 harvested (10).

199 Two other seed- and inflorescence-related traits, the number of seeds per spikelet (SESP) and the number
200 of seeds per floret (SEFL), were calculated based on measurements of the average number of seeds per
201 inflorescence, SPIN, and FLSP as described above.

202

203 **Data Analyses**

204 Least square means (LSMEANS) for each genet and least significant differences (LSD) among genets were
205 determined using the SAS version 9.4 (SAS Institute, Cary, North Carolina, USA) MIXED procedure with genets,
206 years, locations, and the three-way interaction term as fixed effects and replications as random effects. The overall
207 means and LSD categories ($p \leq 0.05$) for the grandparents and parents were determined using the repeated option to
208 model covariance structure between years. Segregation of the trait values used for QTL analysis of the full-sib
209 progeny, from LSMEANS procedure, was measured in part by relative standard deviations (standard deviation
210 among trait values / mean over all trait values) which reflects the magnitude of trait variability, relative to the overall
211 mean, among 68% of the most typical individuals in a normally distributed population. The broad-sense heritability
212 (H) estimates and standard errors were determined on a single plot basis and entry (genet) mean basis for
213 randomized complete block designs in multiple environments using a SAS MIXED procedure with environments
214 (five location x year combinations), replications within environments, genet, and genet x environment as random
215 effects (Holland et al. 2010). Pearson correlation tests were performed for 136 pairwise comparisons of the 17 traits,
216 based on individual plot measurements, using environments (locations and years) as a grouping factor for the
217 statsBy function of the R (R Core Team 2017) psych package (Revelle 2018). Significance thresholds for the trait
218 correlation tests were adjusted using a Bonferroni correction to control for multiple testing.

219 Two different approaches for QTL detection based on models for a two-way pseudo-testcross (TWPT) and
220 cross-pollinated (CP) plants, more fully described in following paragraphs, were performed using MapQTL version
221 6 (Van Ooijen 2009). All of the map files and locus data used for these QTL analyses were based on the integrated
222 GBS consensus map (Kantarski et al. 2017) with a total of 3156 markers from the M26xM36 family. This map
223 included 1699 markers that were heterozygous in M26 (progeny genotypes *lm* or *ll*), 1087 markers that were
224 heterozygous in M35 (progeny genotypes *nn* or *nm*), and another 1070 markers there were heterozygous in both
225 M26 and M35 (progeny genotypes *hh*, *hk*, or *kk*). The quantitative trait data were based on LSMEANS trait
226 estimates for each genet, within and among environments, as described above. All TWPT and CP QTL analysis
227 were initially performed using the same single-QTL interval mapping procedure to identify possible QTL markers
228 that were subsequently used as cofactors in the first round of restricted multiple-QTL model (rMQM) mapping (Van
229 Ooijen 2009). A second round of rMQM mapping was also performed for all TWPT and CP QTL analyses using
230 cofactors for QTLs that were significant after the first round of rMQM mapping. All QTL analyses were performed
231 using a maximum likelihood mixture model, with up to 20 iterations, to determine the LOD likelihood ratio statistic
232 (Van Ooijen 2009). A permutation test for each CP and TWPT analysis was used to determine significance
233 thresholds for the LOD statistic corresponding to a genome-wide P-value of 0.05 (5%) to identify putative QTLs and
234 cofactors (Van Ooijen 2009). Only those QTLs exceeding the 5% genome-wide LOD threshold in the second round
235 of rMQM mapping were considered significant.

236 The two-way pseudo-testcross (TWPT) QTL analysis utilized a doubled haploid model with separate maps
 237 for each parent (Van Ooijen 2009), M26 and M35. Briefly, the *lm* and *ll* genotypes from 1699 M26 markers were
 238 changed to *A* and *B*, respectively; and the *nm* and *np* genotypes from 1087 M35 markers were changed to *B* and *A*,
 239 respectively. The remaining 370 markers with genotypes *hh*, *hk*, or *kk* were deleted. A map with 42 LGs including
 240 1699 M26 markers in 21 LGs and 1087 M35 makers in another 21 LGs was assembled using marker positions from
 241 the integrated consensus map (Kantarski et al. 2017). Only one genotypic effect, the difference between *A* or *B*
 242 marker alleles from one parent, is fitted at any given map position in the TWPT approach. The TWPT design is a
 243 simplified model that has possible theoretical advantages because it enables rMQM mapping of each LG from each
 244 parent using QTL marker cofactors from other LGs of both parents, including one homologous LG from the other
 245 parent (Van Ooijen 2009). Another practical advantage of the TWPT design is that it is relatively easy to identify
 246 useful QTL cofactors when the parental LGs are separated since some QTLs may not be heterozygous both parents.
 247 Thus, to identify TWPT rMQM cofactors we initially selected one marker with the highest LOD score overall
 248 environments for each M26 or M35 LG having at least one significant TWTP QTL. However, the automatic
 249 cofactor selection procedure of MapQTL 6 was also used to identify the final set of TWPT rMQM cofactors.

250 The second approach for QTL analysis utilized an integrated map containing 3156 markers in 21 LGs for
 251 both M26 and M35 parents (Kantarski et al. 2017), which was constructed using the cross-pollinator (CP) model
 252 (Van Ooijen 2006). In full-sib CP families, one or more QTLs may be heterozygous in one or both parents with up
 253 to four possible alleles per QTL. The CP QTL approach always fits four possible QTL alleles designated *a* and *b*
 254 corresponding to marker alleles *l* and *m*, respectively, of the first parent (M26) and QTL alleles *c* and *d*
 255 corresponding to marker alleles *n* and *p*, respectively, of the second parent (M35). The CP QTL analysis (Van
 256 Ooijen 2009) has theoretical and practical advantages in that three possible genotypic effects are fitted including the
 257 difference between *a* and *b* QTL alleles of the first parent (α), the difference between *c* and *d* QTL alleles of the
 258 second parent (γ), and the intralocus interaction (τ) as deviations from the overall mean (μ). Thus, in a cross of *ab* x
 259 *cd* QTL alleles the expected progeny phenotypes are modeled (Van Ooijen 2009) as follows:

$$260 \quad ac = \mu - \alpha - \gamma - \tau$$

$$261 \quad ad = \mu - \alpha + \gamma + \tau$$

$$262 \quad bc = \mu + \alpha - \gamma + \tau$$

$$263 \quad bd = \mu + \alpha + \gamma - \tau$$

264 If the parents (M26 and M35) are heterozygous for the same two QTL alleles, *a* and *b*, then τ would represent a
 265 dominance deviation term. However, this is never assumed to be the case because MapQTL 6 CP model always fits
 266 separate effects, α and γ , for both parents. Since more than 72% of the markers in the integrated CP map were not
 267 informative in one parent or the other parent (Kantarski et al. 2017), the interval mapping procedure can assign
 268 relatively high LOD scores to M26 or M35 marker loci that may not be directly associated with the QTL. Thus, in
 269 addition to the TWPT rMQM cofactors, the Kruskal-Wallis procedure of MapQTL 6 was used to help identify
 270 additional CP rMQM cofactors for each CP QTL. However, we did not consider more than one informative marker
 271 per parent per LG for use as a possible CP rMQM cofactor.

272

273 **Comparative mapping**

274 The two-LOD drop off intervals for M26 x M35 QTLs with the highest LOD values for each trait and each LG,
275 including parent-specific TWPT QTLs, were graphed onto the integrated M26 x M35 GBS consensus map
276 (Kantarski et al. 2017) using the R (R Core Team 2017) LinkageMapView package (Ouellette et al. 2018). Another
277 44 SEMA, SEAR, SELE, and SEWI QTLs, detected over one or more years in the C3-2331 x C3-2595 (UMN)
278 family (Zhang et al. 2017), were included on this graph for comparison to M26 x M35 QTLs detected in this study.
279 The probability (P_{xy}) that one of the x most significant QTL markers was the same as one of the y most significant
280 association-mapping markers, with a total of s shared markers, was calculated as one minus the probability of not
281 having any of the same markers, which equals the products of one minus the chance of drawing one of y numbers x
282 times (where the total number of possible makers, s , decreases from $i = 0$ to $x-1$ without replacement) *or* drawing
283 one of x numbers y times (where the total number of possible makers, s , decreases from $i = 0$ to $y-1$ without
284 replacement) as follows:

292
$$1 - \left(\prod_{i=0}^{x-1} \left(1 - \left(\frac{y}{s-i} \right) \right) \right) \text{ or } \prod_{i=0}^{y-1} \left(1 - \left(\frac{x}{s-i} \right) \right) = P_{xy}.$$

285 Sequences of GBS markers (Kantarski et al. 2017), including 51 significant AM markers (Zhang et al. 2017), were
286 aligned to the pre-publication “*Thinopyrum intermedium* C4-5353-T1 Annotated Standard Draft” sequence,
287 available on the JGI Genome Portal (<https://genome.jgi.doe.gov/portal>), using a Basic Linear Alignment Search
288 Tool (BLAST) with an Expect-value (E) threshold $\leq 1e-10$ (Altschul et al. 1990). Graphical comparisons of the 21
289 LGs (LG01 to LG21) and 21 corresponding chromosome sequences (CHR01 to CHR21), based on GBS marker
290 alignments, were performed using LinkageMapView by normalizing LG lengths, measured in centiMorgans (cM),
291 and CHR lengths, measured in nucleotide mega-bases (MB).

293 Seed-yield and domestication orthogenes of wheat, barley, rice, maize, and other plants (Doebley et al.
294 2006; Doust et al. 2014; Hackauf and Wehling 2005; Lenser and Theißen 2013; Li and Gill 2006; Meyer and
295 Purugganan 2013; Nadolska-Orczyk et al. 2017; Pourkheirandish et al. 2015; Simons et al. 2006; Taketa et al. 2008;
296 Tang et al. 2010) including *Arabidopsis* (Balanzà et al. 2016) and *Lolium* (Manzanares et al. 2016; Shinozuka et al.
297 2010) were aligned to the IWG draft genome sequence using BLAST or BLASTX (Altschul et al. 1997) with a
298 minimum significance threshold of $E \leq 1e-20$. At least three BLAST or BLASTX hits corresponding to three
299 possible orthogenes on three homoeologous chromosomes of IWG (Kantarski et al. 2017) were considered for each
300 CG. If the three most significant IWG BLAST hits were located on chromosome sequences with known orthology to
301 chromosomes harboring barley, *Lolium*, rice, or wheat CGs (Klaas et al. 2011; La Rota and Sorrells 2004;
302 Thorogood et al. 2017; Tulpan and Leger 2017) then only these hits were considered as possible orthogenes.
303 Additional IWG BLAST hits having similar E values were considered when queried using *Arabidopsis* genes
304 allowing for ancient duplications and many possible genome rearrangements that have occurred between dicots,
305 such as *Arabidopsis*, and monocots such as IWG. Only those IWG BLAST hits corresponding to annotated IWG
306 gene models, supported by IWG transcripts, were considered as possible IWG CGs.

307

308 **Results**

309 **Phenotypic and genotypic variation**

310 Significant differences ($p \leq 0.05$) were detected among the C3 grandparents and C4 parents for all 17 traits, with up
311 to four levels of difference for SEWI, SYIN, and ZAMA (Table 1). Moreover, significant effects ($p \leq 0.05$) of genet,
312 year, location, and three-way interaction term were detected for all traits except FLSP, SEFL and INLE in the
313 overall analysis of the grandparents, parents, and 266 full-sib progenies. In most cases, those effects were highly
314 significant ($p \leq 0.0001$), therefore all further analyses were conducted for each location x year environment. Only
315 the year effect was not significant for INLE and only the location effect was not significant for FLSP and SEFL. The
316 standard deviations among the progeny trait estimates were within 4% to 8% of the overall trait means for SELE,
317 SEWI, and ZAMA (Table 2). Conversely, the standard deviations varied from 38% to 86% of the trait means for
318 SEFL, SESH, SESP, and SYIN (Table 2). However, broad-sense heritability (H) estimates for traits with lower
319 variability tended to be greater than H estimates for traits with high variability (Table 2). The relative standard
320 deviations for SENA varied from 17% to 94% among environments, in part because the overall means of this trait
321 also varied widely among environments (Table 2). These data, Tables 1 and 2, demonstrate significant genetic
322 variation among the C3 grandparents, C4 parents, and full-sib progeny, respectively, for all 17 traits.

323 Significant correlations ($p < 0.05$) were detected for 100 of the 136 possible pairwise comparisons among
324 17 traits, controlling for multiple testing (Table 3). Although there were up to 12 possible plot measurements taken
325 for each of the 270 genets (two replications in KS evaluated three years and three replications in UT evaluated over
326 two years), with 3240 possible measurements per trait, the maximum number of observations was limited to 2570
327 measurements because not all genets were present in all five replications. Relatively strong positive correlations
328 were observed among the four seed-size traits (SEAR, SELE, SEMA, and SEWI) and between the two fertility
329 traits, SEFL and SESP (Table 3). Seed yield per inflorescence (SYIN) is a complex trait that showed relatively
330 strong and significant correlations with fertility traits, SEFL and SESP, and moderate correlations with seed size
331 (SEAR, SELE, SEMA, and SEWI), inflorescence capacity (SPIN and FLSP), and seed disarticulation (SENA and
332 SESH) traits (Table 3). The percentage of naked seeds (SENA) was negatively correlated SEWI, SESH, and FLSP
333 (Table 3), which was especially true in the KS evaluations where these correlation coefficients were -0.25, -0.34,
334 and -0.12, respectively.

335 336 **QTL analysis**

337 Permutation tests showed that the minimum LOD thresholds required to control for a 5% genome-wide error rate (P
338 < 0.05) were 3.5 and 4.7 or less for the TWPT and CP QTL analyses, respectively, for each trait. Using these LOD
339 thresholds for all 17 traits, there were a total of 210 significant QTLs detected among six QTL analyses for each of
340 five different environments and averages over all five environments (Table 4; Supplemental Document 1). A subset
341 of 56 QTLs were significant in only one analysis, whereas 55 QTLs were significant among two or more analyses
342 (Supplemental Document 1). A total of 57 QTLs were significant in the analysis of trait averages over all five
343 environments but only one QTL, for SELE on LG06, was significant across all six analyses (Table 4). Considering
344 the most significant QTL with maximum LOD on each LG and each trait, there were at least 111 distinct QTLs

345 including 19 M26 and 20 M35 QTLs detected using the TWPT model and 72 QTLs detected using the CP model
346 (Tables 4 and 5; Supplemental Document 1). Only 3 distinct QTLs were detected for CRCI and up to 10 distinct
347 QTLs were detected for ZAMA (Table 5). There were at least two significant QTLs on each of the 21 LGs, with up
348 to 12 significant QTLs on LG06 (Table 5).

349 For most traits except INLE, the total magnitudes of maximum QTL effects (Max-total) and overall
350 average QTL effects (Avg-total), within and among five environments, were not the same for both M26 and M35
351 parents (Table 6). The Max-total and Avg-total M35- γ effects were greater than corresponding M26- α effects for
352 SEMA, FLSP, SEFL, SESP, STWI, ZAMA, CIRCI, and INCR (Table 6). Conversely, the Max-total and Avg-total
353 M26- α effects were greater than corresponding M35- γ effects for SEAR, SENA, SYIN, and STLE (Table 6). The
354 Avg-total M26- α effects were also greater than corresponding M35- γ effects for SELE, SESH and SPIN, but this
355 was not true for the Max-total effects of these four traits (Table 6). A more detailed examination of SENA QTL
356 effects indicates the magnitude of M26- α QTL effects were equal or greater than the M35- γ effects for six of nine
357 SENA QTLs including the LG02 and LG11 SENA QTLs, which had relatively large M26- α effects (Table 6). Thus,
358 the Avg-total M26- α SENA effects (28% of population mean) was about 133% greater corresponding M35- γ effects
359 (12% of population mean), with a total estimated effect of 52% of the SENA mean (Table 6). The LG02 SENA QTL
360 was most significant in the KS15 analysis (Table 4), where progenies with LG02 SENA QTL genotypes *ac* and *bd*
361 were 39% greater than the population mean or 25% less than the population mean, respectively (Table 6;
362 Supplemental Document 1). If the LG02 M26 *a* and *b* alleles are additive, then it could be inferred that the SENA
363 means for progeny with an *aa* genotype would be 42% over the population mean (Table 2), with a M26- α SENA
364 effect of 21% of population mean (Table 6). However, there are no progeny or parents that are homozygous for the
365 four possible alleles (*a*, *b*, *c*, or *d*) of any QTLs in the M26 x M35 family so it is not possible to determine the effects
366 possible QTL genotypes. Nevertheless, the Avg-total SENA QTL effects, 68% of population mean (Table 6), were
367 enough to account for the range of 23.2% to 83.5% SENA variation (55.3% average) among progeny (Table 2).
368 Moreover, the Avg-total M26- α effects, 28% of population mean (Table 6), would also be enough to account for
369 differences of 28.7% and 94.1% SENA between the parents of M26, C3_3941 and C3_3471 (Table 1), if they were
370 homozygous for small- and large-effect alleles, respectively, and if the M26 *a* and *b* alleles are additive.

371 The highest LOD score for all 17 traits was for the LG11 SESH QTL (Table 5), which had relatively strong
372 but opposite effects on both SESH and SENA (Table 6) meaning that it has a desirable or synergistic effects of
373 decreasing seed shattering and increasing the percentage of free-threshing (naked) seed. In fact, this LG11 QTL was
374 the only QTL that had synergistic M26- α SESH and SENA QTL effects and synergistic M35- γ SESH and SENA
375 QTL effects (Table 6). Only one other QTL, on LG10, had synergistic SESH and SENA M26- α effects (Table 6).
376 The total magnitude of the LG10 and LG11 SESH and SENA M26- α effects, 23% and 26% respectively (Table 6),
377 would be enough to account for most SESH and SENA differences between the parents of M26, C3_3941 and
378 C3_3471 (Table 1), if they were homozygous for desirable and undesirable alleles, respectively, and if the M26 *a*
379 and *b* alleles are additive. Two other QTLS on LG12 and LG14 had synergistic SESH and SENA M35- γ effects.
380 Thus, at four QTLs on LG10, LG11, LG12, and LG14 had synergistic M26- α or M35- γ effects on SESH and SENA
381 (Table 6), which account for negative correlations between these traits (Table 3). Only one QTL, on LG06, had the

382 same positive or negative effect on SESH and SENA (Table 6). In general, the relative directions of QTL effects for
383 different traits (Table 6) were consistent with trait correlations (Table 3). For example, nine of the 13 LGs associated
384 with seed size traits (SEAR, SELE, SEMA, and SEWI) had the same directional effects for α or γ on more than one
385 of these traits, which was consistent with positive correlations among these traits (Table 3). Likewise, six of the 10
386 LGs associated with SEFL and SESP had the same directional α or γ effects for both traits, consistent with the
387 positive correlation between these two traits (Table 3).

388 The percent variation explained (PVE) by M26 was substantially greater than M35 for SEAR, SELE,
389 SENA, SPIN, FLSP, and STLE (Table 7). Conversely, the PVE by M35 was substantially greater than M26 for
390 SEMA, SEWI, SEFL, SESP, INLE, STWI, ZAMA, CRCI, and INCR (Table 7). Thus, the relative PVE by M26 and
391 M35 (Table 7) closely mirrors the relative magnitudes of the average M26- α and M35- γ QTL effects among all five
392 environments (Table 6). However, the combined PVE by M26 and M35 in the CP approach was greater than the
393 combined PVE by M26 and M35 in the TWPT approach for SEMA, SEWI, SPIN, ZAMA, and CRCI (Table 7),
394 presumably because the CP model includes effects of intralocus interactions (τ).

395

396 **Alignment of IWG QTL intervals and association mapping (AM) loci to IWG draft genome sequence**

397 A subset of 3,608 (93.6%) of the 3,856 GBS markers comprising the M26 x M26 linkage map had
398 significant matches to the IWG physical map. Of these 3,608 matches, 2,977 (82.5%) showed synteny with
399 homologous chromosome sequences such as LG01 and CHR01 (Figure 1), 239 (6.6%) mapped to different members
400 of the same homoeologous group (HG) such as LG01 and CHR02 or CHR03, 213 (5.8%) mapped to different
401 genetic and physical HGs, and 179 (5.0%) matched unmapped scaffolds. Substantial collinearity among GBS
402 markers of the 21 linkage groups and corresponding chromosome sequences was also discernable (Figure 1). A
403 subset of 1,072 (29.7%) of the 3,608 M26 x M35 GBS marker alignments were located within 1,072 (0.7%) of the
404 159,905 IWG gene models. A slightly higher portion of 37 (33%) of the 111 most significant M26 x M35 QTL
405 markers were located within IWG gene models. A total of 9,403 (93.7%) of the 10,029 GBS markers from the IWG
406 consensus map (Kantarski et al. 2017) showed significant matches to the draft genome sequence, including 3,223
407 (34.3%) located within annotated IWG gene models.

408 A total of 93 GBS markers had the highest LOD score for the 111 most significant M26 x M35 QTLs and
409 the two-LOD drop off intervals for many of the M26 x M35 QTLs were identical or very similar (Supplemental
410 Document 1). The 111 M26 x M35 QTLs collapsed into 72 QTL intervals spanning different regions or different
411 lengths of the linkage map (Figure 1). However, these 72 QTL intervals covered no more than 36 non-overlapping
412 regions of the 21 integrated LGs. One region of LG06 had significant effects on 11 different traits including all seed-
413 size, fertility, and seed harvest traits (Figure 1).

414 A total of 44 SEMA, SEAR, SELE, and SEWI QTLs from the biparental UMN C3-2331 x C3-2595
415 population (Zhang et al. 2017) were mapped to 22 QTL intervals corresponding to 21 non-overlapping regions of
416 the IWG consensus map (Figure 1). Similarly, 25 M26 x M35 seed-size QTLs were mapped to 20 QTL intervals
417 corresponding to only 14 non-overlapping regions of the same IWG consensus map (Figure 1). A total of 69 seed
418 QTLs from both families, M26 x M35 and C3-2331 x C3-2595, mapped to only 26 non-overlapping regions on 20

419 LGs with correspondence between families on LG06, LG08, LG09, LG11, LG14, LG15, LG17, and LG20 (Figure
420 1). A total of 51 AM markers mapped to 18 of the 21 LGs, excluding LG06, LG07 and LG19 (Zhang et al. 2017).
421 No significant seed-size QTLs were detected on LG05, in the M26 x M35 or C3-2331 x C3-2595 families, but there
422 were eight AM markers on this LG. Five LGs (LG02, LG06, LG15, LG17, and LG20) had significant QTL effects
423 on all four seed-size traits (SEMA, SEAR, SELE, and SEWI), but only LG06 had significant QTL effects on all four
424 seed-size traits in both M26 x M35 and C3-2331 x C3-2595 families. Only LG03, LG05, LG13 and LG20 had
425 significant AM markers for all four seed-size traits. Only LG20 had significant QTL and AM markers for all four
426 seed-size traits.

427 Two of the 25 most-significant M26 x M35 and 44 most-significant C3-2331 x C3-2595 UMN seed-size
428 QTL markers, TP678810 and TP693406, were also among 51 most significant seed-size AM markers. The
429 TP678810 GBS polymorphism was the most significant marker for the LG17 SELE QTL in the M26 x M35 family
430 (Supplemental Document 1) and it was also associated with SEMA, SEAR, and SELE in the UMN AM population
431 (Zhang et al. 2017). The TP693406 polymorphism was the most significant marker for a SEWI QTL
432 (*Ti_QSws.umn_4.1*) in the C3-2331 x C3-2595 family and the SELE in the AM population (Zhang et al. 2017). The
433 coincidence of these two markers, ranking among the 69 most significant QTL markers and 51 most significant AM
434 markers, raised questions about the probability of this happening by chance considering that there were 3,856 M26 x
435 M26 markers, 2,167 C3-2331 x C3-2595 markers, and 4,873 AM markers used in these experiments (Zhang et al.
436 2017). The probability of having at least one identical marker between the M26 x M35 QTL experiment and AM
437 experiment was calculated as 13.9% using the formula for P_{xy} where only $x=12$ of the most significant M26 x M35
438 QTL markers, $y=14$ AM markers, and $s=1,139$ markers were common to both experiments. The probability of
439 having at least one identical marker in the C3-2331 x C3-2595 QTL experiment and AM experiment was calculated
440 as 18.9% using the formula for P_{xy} where only $x=23$ of the most significant C3-2331 x C3-2595 markers, $y=10$ of
441 the most significant AM markers, and $s=1,115$ markers were common to both experiments. The probability of both
442 occurrences was 2.6%, which suggests that these coincidences occurred more frequently than expected by chance
443 alone. These two shared markers, LG17 TP678810 and LG04 TP693406, were located within two IWG gene
444 models, Thintv21245331m.g and Thintv21054937m.g, respectively. The LG17 Thintv21245331m.g IWG gene
445 model belongs to a *GDSL lipase (GLIP)* gene and the LG04 Thintv21054937m.g IWG gene model encodes a
446 putative *xylogalacturonan beta-1,3-xylosyltransferase, xylogalacturonan deficient (XGD)*, gene (Supplemental
447 Document 2). A total of 22 (43%) of the 51 most significant AM markers (Zhang et al. 2017) were located in
448 annotated IWG gene models whereas only 34.3% of all 9,403 GBS alignments to the IWG draft genome sequence
449 were located in structurally annotated gene model, suggesting that some of the most significant AM markers may be
450 located in functionally relevant genes.

451 Alignments of 42 domestication genes of wheat, barley, rice, and other plants (Table 8) to the draft genome
452 sequence of IWG detected significant homology to 142 annotated gene models of IWG (Figure 1; Supplemental
453 Document 2). Most of these genes showed evidence of orthology based on functional annotations and known
454 patterns of synteny among species (Supplemental Document 2). Moreover, alignments of IWG chromosome

455 sequences to the IWG QTL map, based on GBS markers, revealed 98 CGs with possible alignments to one or more
456 relevant IWG QTLs (Table 8, Figure 1; Supplemental Document 2).

457 **Discussion**

458 **IWG QTLs correspond to candidate genes (CGs) on seven homoeologous groups (HGs)**

459 One of the ultimate goals of many QTL studies is to identify genes and mutations that underlie functionally
460 important traits, initiated here by the identification of IWG loci corresponding to domestication orthogenes of
461 closely related species such as wheat, barley, and rice. The development of an annotated draft genome sequence for
462 IWG greatly facilitated this effort with identification of 142 possible domestication orthogenes, including 98
463 possible CG-QTL alignments. Discussion of these alignments was organized according to the seven homoeologous
464 groups of wheat, barley, and IWG (Kantarski et al. 2017), which also have well-defined relationships to the 10
465 chromosomes of maize and 12 chromosomes of rice (La Rota and Sorrells 2004; Tulpan and Leger 2017).

466 Homoeologous group 1 was interesting in part because it produced the highest SENA LOD score (Table 5).
467 Two possible orthologs of the *Seedstick* gene of *Arabidopsis* (Balanzà et al. 2016), designated *STKa* and *STKb*, were
468 found on each of the three IWG HG1 chromosomes (Figure 1). The *Seedstick* gene encodes a MADS-box
469 transcription factor required for seed abscission in *Arabidopsis*, with seeds remaining attached to the funiculus when
470 after fruits dehisce in *stk* mutants (Balanzà et al. 2016). Although patterns of synteny between *Arabidopsis* and IWG
471 are not easy to discern (Tulpan and Leger 2017), the functional annotation of the putative IWG *STK* orthogenes
472 match that of *Arabidopsis* (Supplemental Document 2). The IWG *STKa* and *STKb* loci on CHR02 align to a SENA
473 QTL on LG02, which had the most consistent SENA effects across environments (Table 4) and the highest LOD
474 score (Table 5). Orthologs of the rice *Grain weight chromosome-5 QTL (GW5)* gene (Shomura et al. 2008; Weng et
475 al. 2008) aligned to seed-size QTL on LG01, LG02, and LG03 (Table 8). One of three possible IWG orthologs of
476 maize the *Barren inflorescence 4 (Bif4)* gene (Galli et al. 2015) located on CHR01, CHR02, and CHR03 aligned to a
477 SPIN QTL on LG03. The maize *Bif1* and *Bif4* genes encode AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA)
478 proteins required for early steps of inflorescence development (Galli et al. 2015).

479 Homoeologous group 2, particularly LG06, produced the highest M26 x M35 LOD scores for SEMA,
480 SEAR, SELE, and SYIN. The LG06 SELE QTL had the second highest LOD score of 12.0 in this experiment
481 (Table 5) and it was the only QTL that was significant across all five environments (Table 6). Homoeologous group
482 2 was also remarkable in that it showed significant M26 x M35 QTL effects for all traits except CRCI and INCR
483 (Table 5) and the greatest overall number of seed-size QTLs (22) in the M26 x M35 family (four QTLs), C3-2331 x
484 C3-2595 UMN family (7 QTLs), and AM population (11). The *Grain incomplete filling (GIF)* cell-wall invertase
485 gene of rice (Li et al. 2013; Wang et al. 2008; Yan et al. 2011), *grain length chromosome-7 QTL (GL7)* gene (Wang
486 et al. 2015), *Sucrose synthase 2 (SUS2)* gene (Hou et al. 2014; Jiang et al. 2011), and *xylogalacturonan-deficient*
487 (*XGD*) gene of *Arabidopsis* (Jensen et al. 2008) were present on CHR04, CHR05, and CHR06 with alignments to
488 seed-size and seed yield traits on LG04 and LG06 (Figure 1, Table 8). A *GL7* locus on LG06 is an interesting
489 candidate for the major-effect SELE QTL associated with this linkage group (Figure 1, Tables 5 and 6). The *XGD*
490 locus on CHR04 was of particular interest because it contained the most significant SEWI QTL marker in the IWG
491 C3-2331 x C3-2595 family, TP693406, which was also associated with SELE in the AM population (Zhang et al.

2017). The *xgd1* mutant of *Arabidopsis* has decreased levels of xylose and pectic xylogalacturonan, which are important components of cell walls and reproductive tissues that may affect process of plant development such as pectin degradation during fruit ripening (Jensen et al. 2008). The *Clustered primary branch 1 (CPB1)* gene encodes a cytochrome P450 protein involved in brassinosteroid biosynthesis, which is associated with inflorescence architecture, seed size, and plant height traits including the *DWARF11 (D11)* phenotype of rice (Wu et al. 2016). Three *CPB1* loci were found on CHR05, CHR06, and one scaffold that presumably should map to CHR04 where it might align to inflorescence and seed-size QTLs on LG04. The CHR05 and CHR06 *CPB1* genes aligned to STLE and seed-size QTLs on LG05 and LG06, respectively. Candidate genes closely associated with the barley gibberellin-insensitive *Semidwarf 3 (SDW3)* gene (Vu et al. 2010) were present on CHR04, CHR05, and CHR06 including one that also aligned to the relatively broad STLE QTL interval on LG05 (Figure 1). The *Six-rowed spike 1 (Vrs1)* encodes a homeobox-leucine zipper protein that reverts the rudimentary lateral spikelets of two-row barleys into fully developed spikelets of the six-row barleys (Komatsuda et al. 2007). Homoeologous *Vrs1* genes were present on CHR04, CHR05, and CHR06 with two copies of this gene on CHR06. The *Vrs1* genes on CHR04 and CHR06 align to SPIN and fertility (SEFL and SESP) QTLs on LG04 and LG06, respectively (Figure 1, Table 8). The barley *ZEOCRITON (ZEO)* spike density gene, located on the long arm of barley chromosome 2H (Houston et al. 2013), was present on the long arm of IWG CHR04, the short arm of IWG CHR06, and a paralogous locus on IWG CHR21 (Figure 1). The IWG *ZEO* loci did not align to INLE QTLs, but there was an INLE QTL aligned to the long arm of IWG CHR05, which may not be fully and correctly assembled.

Homoeologous group 3 had the highest LOD scores for SEWI, SPIN, FLSP, and INLE (Table 5). A total of 17 loci corresponding to six HG3 CGs (Table 8) were syntenous and colinear among CHR07, CHR08, and CHR09 (Figure 1). Mutations of the maize *Barren inflorescence 1 (BIF1)* gene (Galli et al. 2015) and *Barren stalk 1 (BA1)* gene (Gallavotti et al. 2004) reduce the number of spikelets due to defects in auxin signaling. The barley chromosome-3H *Six-rowed spike 4 (VRS4)* gene, an ortholog of the maize *RAMOSA2* inflorescence architecture gene, is associated with spikelet fertility and determinacy (Koppolu et al. 2013). The *BA1*, *BIF1* and *VRS4* genes aligned within or near SPIN and FLSP QTL intervals on LG07 and LG09. The *Grain number chromosome-1 QTL (GNI)* gene (Ashikari et al. 2005; Zhang et al. 2012a) is a cytokinin oxidase gene associated with a major grain-number QTL and grain weight in rice. The *GNI* gene was present on CHR07 and two CHR08 loci aligned to SEAR and SPIN QTLs on LG07 and a SEWI QTL on LG08. The *Brassinosteroid-insensitive 1 (BR1)* gene is responsible for the *uzu* semidwarfing mutation (Chono et al. 2003), which has pleiotropic effects on spike length (Chen et al. 2016) and aligned to INLE and STLE QTLs on LG07 and LG09. The barley *DENSO (Semidwarf 1)* gene encodes a gibberellic acid (GA)-20 oxidase enzyme required for GA biosynthesis (Jia et al. 2009), which can have pleiotropic effects on heading date and possibly grain size but not spike length (Kuczyńska et al. 2014). The *DENSO* gene aligned within or near STLE and seed-size QTLs on LG08 and LG09. The rice *Grain size chromosome-5 QTL (GS5)* serine carboxypeptidase gene plays a major role in regulating grain size and weight in rice and possibly wheat (Li et al. 2011; Ma et al. 2016). The *GS5* gene maps to wheat 3A, 3B, and 3D (Ma et al. 2016) and IWG CHR07, CHR08, and CHR09 (Figure 1), which align to seed size QTLs on LG07, LG08, and LG09.

528 Homoeologous group 4 produced a total of 23 M26 x M35 QTLs, only slightly fewer than HG5, and
529 produced the highest LOD scores for STLE, SESH, ZAMA, and INCR (Table 5). This HG was exceptional in that it
530 produced the highest LOD score for any trait, which was associated with the LG11 SESH QTL (Table 5). Moreover,
531 significant QTL effects for SESH, SENA, SEFL, and ZAMA were detected on all three HG4 linkage groups (LG10,
532 LG11, and LG12). Orthologs of the wheat ‘green revolution’ *Reduced height (RHT)* gene (Peng et al. 1999) were
533 found on CHR10 and CHR12 including one that aligns to the high-LOD STLE QTL on LG10, which also overlaps
534 with other QTLs controlling other traits (Figure 9). Despite efforts to find seed-shattering CGs from wheat, barley or
535 rice that correspond to IWG HG4, the only relevant gene that we found was the *Arabidopsis Seuss (SEU)* gene
536 (Balanzà et al. 2016), which aligned to the SESH and SENA QTLs on LG10 and the SESH QTL on LG12. Many
537 genes controlling photoperiod and flowering date have been identified in temperate cereals including wheat and
538 barley (Cockram et al. 2007). The *vernalization 2 (VRN2)* gene is a major determinant of flowering time in wheat
539 (Yan et al. 2004) and barley (Dubcovsky et al. 2005; Karsai et al. 2005; Laurie et al. 1995). The *VRN2* gene is
540 located on barley chromosome 4HL and a region of wheat chromosome 4AL that was translocated to 5AL (Devos et
541 al. 1995; Dubcovsky et al. 1998). The *VRN2* gene was located on IWG CHR10 (HG4) and CHR15 (HG5), which
542 was not unexpected because LG12 (HG4) and LG15 (HG5) of IWG display a reciprocal translocation (Kantarski et
543 al. 2017) similar to the 4AL/5AL translocation of *Triticum monococcum*, *T. aestivum* and other Triticeae species
544 (Devos et al. 1995; Dubcovsky et al. 1998; Larson et al. 2012). A third *VRN2* gene was found on an unmapped
545 scaffold_264 that probably belongs to CHR11, which is colinear with barley chromosome 4H (Kantarski et al.
546 2017). However, none of the ZAMA QTLs aligned to the CHR10-, CHR15-, or predicted CHR11-*VRN2* loci. Two
547 *Phytochrome* genes designated *PHYA* and *PHYB* (Mathews and Sharrock 1996) map to CHR10, CHR11, and
548 CHR12 with alignments to ZAMA QTLs on LG10 and LG11. The LG10 ZAMA QTL could be a pleiotropic effect
549 of the LG10 STLE QTL, but there was no alternative explanation for the alignment of the LG11 ZAMA QTL to the
550 CHR10 *PHYB* gene.

551 Homoeologous group 5 was exceptional in that it displayed the greatest number of QTLs, 25 in total, and
552 that it displayed significant effects for SEFL and SESP on all three HG5 linkage groups including LG13, LG14, and
553 LG15 (Table 5). The *Q* gene on wheat chromosome 5AL confers the free-threshing trait of domesticated wheat and
554 has pleiotropic effects on glume shape, glume tenacity, rachis fragility, spike length, plant height and heading date
555 (Faris and Gill 2002; Simons et al. 2006). Orthologs of *Q* were found on the long arm of CHR13, CHR14, and
556 CHR15 with alignment to INLE, SESH, and SENA QTLs on LG14 and possibly another SENA QTL on LG15
557 (Figure 1). Other potentially important CGs located near *Q* include the *Phytochrome C (PHYC)* and *Vernalization 1*
558 (Yan et al. 2003) genes, which aligned to a ZAMA QTL on LG14 (Figure 1). Orthologs of the *CONSTANS 3 (CO3)*
559 *CONSTANS*-like flowering gene (Griffiths et al. 2003) also align to ZAMA QTLs on LG14 and LG15. Homologs
560 of the rice *Grain length chromosome-3 QTL (GL3)* gene (Zhang et al. 2012b) were present at two loci on CHR13,
561 CHR14, and CHR15 with alignments to seed-size QTLs on all three corresponding linkage groups (Figure 1). Two
562 closely-linked homologs of the rice *Thousand-grain weight 6 (TGW6)* gene (Ishimaru et al. 2013), located on
563 CHR15, also aligned to seed-size QTLs on LG15. Homologs of the maize *Barren inflorescence 2 (BIF2)* gene

564 (McSteen et al. 2007) and the *Dense and erect panicle 1 (DEP1)* gene of rice and barley (Huang et al. 2009; Wendt
565 et al. 2016) show alignments to INLE and SPIN QTLs on LG13 and LG14.

566 Homoeologous group 6 was tied with HG2 for having the highest SEAR LOD score and all three HG6
567 linkage groups (LG16, LG17, and LG18) had significant ZAMA effects (Table 5). Three *CONSTANS*-like flowering
568 orthogenes designated *CO2*, *CO5* and *CO7* (Griffiths et al. 2003) were each present on CHR16, CHR17, and
569 CHR18 except that *CO5* was not present on CHR16. At least one of two of these *CONSTANS* genes aligned to each
570 of the ZAMA QTLs on LG16 and LG17, but none aligned to the ZAMA QTL on the distal long arm of LG18.
571 Orthologs of the *Growth-regulating factor 4 (GRF4)* gene (Sun et al. 2016), *Grain weight chromosome-2 QTL*
572 (*GW2*) gene (Song et al. 2007; Su et al. 2011), and a *GDSL lipase (GLIP)* gene aligned to seed-size QTLs on LG16,
573 LG17, and LG18. The rice *GRF4* gene encodes a growth and cytokinin-biosynthesis regulator that increases grain
574 size and inflorescence length and decreases seed shattering (Sun et al. 2016). The *GLIP* locus on LG17 contained
575 one of the most significant SELE QTL markers in the M26 x M35 family, TP678810, which was also associated
576 with SEMA, SEAR, and SELE in the UMN AM population (Zhang et al. 2017). Some *GLIP* genes (Jiang et al.
577 2012) have roles in seed size and metabolism (Clauß et al. 2011; Huang et al. 2015; Lai et al. 2017; Ma et al. 2018;
578 Tiwari et al. 2016).

579 Homoeologous group 7 was remarkable in that it was associated with a total of 21 seed-size QTLs, only
580 one less than HG2, and the highest number of M26 x M35 seed-size QTLs (Table 5). Compared to the other HGs,
581 more seed-size CGs and CG-QTL alignments were also found in HG7. Six relevant seed-size CGs including the rice
582 chromosome-8 *Fertilization independent seed (FEI)* gene (Kapazoglou et al. 2010; Nallamilli et al. 2013), rice
583 *Grain width chromosome-8 QTL (GW8)* gene (Wang et al. 2012), *Glucan, water-dikinase (GWD)* gene (Ral et al.
584 2012; Shu and Rasmussen 2014), *Sucrose synthase 1 (SUS1)* gene (Hou et al. 2014; Jiang et al. 2011), *Thousand-*
585 *grain weight 6 (TGW6)* gene (Ishimaru et al. 2013), and wheat *Thousand-grain weight chromosome-7A (TGW7A)*
586 gene (Hu et al. 2016) all aligned to seed-size QTLs in HG7 (Table 8). The maize *teosinte glume architecture 1*
587 (*tga1*) gene (Wang et al. 2005) and barley *nud* gene (Taketa et al. 2008), responsible for naked grains, were both
588 found on LG19, LG20, and LG21 of HG7 but no significant IWG SENA QTLs were detected in HG7.

589

590 **The S and Z self-incompatibility (SI) genes**

591 The *S* and *Z* SI genes have been shown to affect seed size and fertility traits in perennial ryegrass (*Lolium perenne*)
592 if there is insufficient genetic variability in populations such as a full-sib mapping family (Studer et al. 2008), which
593 raises concerns about possible effects of these genes in the full-sib IWG QTL mapping populations. The *S* and *Z* SI
594 genes could also be a yield-limiting factor in the broader IWG grain breeding populations considering that they
595 purportedly incurred population bottlenecks as few as 14 individuals (Wagoner 1990; Zhang et al. 2016). The *S* and
596 *Z* genes are located on chromosomes that correspond to homoeologous groups 1 and 2, respectively, in both *Lolium*
597 and *Secale* (Hackauf and Wehling 2005; Manzanares et al. 2016; Shinozuka et al. 2010; Thorogood et al. 2017).
598 Seed fertility QTLs (SEFL and SESP) were detected on 10 linkage groups (Table 5) including LG01 (HG1) and
599 LG06 (HG2). In fact, LG06 displayed the highest LOD scores for several seed-size (SEMA, SEAR, and SELE),
600 seed-fertility (SEFL and SESP), and seed yield (SYIN) traits. Identification of *S* and *Z* genes has been difficult

601 (Thorogood et al. 2017), but it is thought that a *domain of unknown function* (*DUF247*) gene is the pollen
602 component of the *S* locus on HG1 (Manzanares et al. 2016). Another paralogous *DUF247* gene (Shinozuka et al.
603 2010) and closely linked *ubiquitin-specific protease* (*USP*) gene (Hackauf and Wehling 2005) are considered the
604 best candidates for the *Z* locus (Thorogood et al. 2017). Putative orthologs of the *S* (*DUF247*) gene were found on
605 IWG CHR01, CHR02, and CHR03 of HG1 with alignments to seed-size QTLs (Figure 1). Moreover, two markers
606 with significant seed-mass effects in the AM population (Zhang et al. 2017) were located about 5 MB and 68 KB
607 from the LG01 and LG02 *S* loci, respectfully (Figure 1). However, none of the putative IWG *S* genes aligned to
608 SESP or SEFL fertility QTLs (Figure 1). Tightly-linked *DUF247* and *USP* genes on IWG CHR01 (HG1) and
609 CHR06 (HG2) were homologous to tightly-linked *DUF247* and *USP* genes corresponding to the *Z* ortholoci on HG2
610 of *Secale* (Hackauf and Wehling 2005) and *Lolium* (Shinozuka et al. 2010). The putative *Z* locus on IWG CHR06
611 shows possible alignment near seed-size, fertility, and seed-yield QTLs on LG06 and it is presumably orthologous to
612 the *Z* locus on HG2 of *Secale* (Hackauf and Wehling 2005) and *Lolium* (Shinozuka et al. 2010). Linkage group 6
613 (LG06) was the only LG that had significant QTLs for all four seed-size traits in both M26 x M35 and C3-2331 x
614 C3-2595 families and LG06 also had the highest LOD scores for SEFL and SESP (Table 5). However, it was also
615 interesting that CHR06 was one of only three chromosomes that did not have any significant seed-size AM effects
616 (Zhang et al. 2017) indicating that diversity of the putative IWG CHR07 *Z* locus was not limiting seed size in the
617 AM population. Tightly linked *DUF247* and *USP* genes on IWG CHR01 aligned to SESP and SYIN QTLs on LG01
618 (Figure 1), but it is not absolutely clear if this CHR01 (HG1) IWG locus is orthologous to the HG2 *Z* locus of *Secale*
619 and *Lolium* (Hackauf and Wehling 2005; Shinozuka et al. 2010). Further research is needed to determine if the *S* or
620 *Z* SI genes are affecting grain size and grain yields in the IWG grain breeding populations and future experimental
621 research should be designed to consider this question.

622

623 **Prospects for gene identification**

624 Recent advancements in DNA sequencing and genomics have enabled scientists to pinpoint genes and chromosome
625 regions that distinguish crops from their wild progenitors, with promising applications to a multitude of additional
626 crop species (Tang et al. 2010). Development of high-density GBS linkage maps (Kantarski et al. 2017) and a draft
627 genome sequence will facilitate identification of CGs associated with IWG domestication QTLs, but it is also
628 recognized that most of these QTLs span large regions of the genome containing many other genes (Figure 1).
629 Additional sequencing of M26, M36, C4-3471, and other IWG reference plants will facilitate identification of
630 mutations in these CGs, some of which may have large or dramatic effects (Tang et al. 2010). Moreover, relatively
631 low levels of linkage disequilibrium and high-levels of outcrossing coupled with recent advances in DNA
632 sequencing and genotyping will facilitate genetic testing of CG variants by association mapping in IWG (Zhang et
633 al. 2017). This approach was used to identify or test CGs of wheat (Hou et al. 2014; Hu et al. 2016; Jiang et al. 2011;
634 Ma et al. 2016; Su et al. 2011; Zhang et al. 2012a), barley (Shu and Rasmussen 2014), maize (Gallavotti et al. 2004),
635 and rice (Li et al. 2011; Wang et al. 2015; Weng et al. 2008; Zhang et al. 2012a) examined in this study (Table 8).
636 However, it is likely that many of the IWG QTLs and AM effects are caused by other genes not examined in this
637 study or perhaps not recognized in any other crop species. Positional cloning was used to identify many CGs of the
638 wheat (Faris et al. 2003; Ishimaru et al. 2013; Yan et al. 2003; Yan et al. 2004), barley (Komatsuda et al. 2007; Vu

639 et al. 2010), rice (Ashikari et al. 2005; Huang et al. 2009; Li et al. 2011; Song et al. 2007; Sun et al. 2016; Wang et
640 al. 2008; Wang et al. 2012; Wang et al. 2015; Weng et al. 2008; Wu et al. 2016; Zhang et al. 2012b), and *Lolium*
641 (Manzanares et al. 2016; Shinozuka et al. 2010) examined in this study (Table 8). Positional cloning has been
642 particularly successful in rice in part because of its relatively small genome size, but recent advances in DNA
643 sequencing will accelerate progress in large-genome species including wheat, barley, and IWG.

644 Comparisons of QTL results from the M26 x M35 and C3-2331 x C3-2595 families identified possible
645 targets for positional cloning. The M26 x M35 family was specifically developed to examine the unique
646 combination seed-retention and naked-seed traits found in C3_3471, by crossing C3_3471 with a more primitive
647 plant, C3_3941, and testing for segregation of QTLs from the F1 hybrid parent, M26. Results of this experiment
648 indicate that this unique combination of SESH and SENA traits, observed in C3_3471 (parent of M26), involved a
649 combination of at least two QTLs located on LG10 and LG11, which had synergistic M26- α effects on both traits
650 (Table 6). The LG11 SESH QTL had the highest LOD score for any trait (Table 5) and it may be a reasonable target
651 for positional cloning especially if this QTL could be isolated in an otherwise isogenic or neutral background. Other
652 possible targets for positional cloning are the SENA QTL on LG02 and seed-size and seed-yield QTLs on LG06, but
653 there are also some potentially good CGs for these QTLs that should be considered as discussed above. The LG02
654 SENA QTL on LG02 had relatively large effects as percentage of the mean (Table 6), meaning that it may have high
655 breeding value. However, the LOD score of the LG02 SENA QTL was not so exceptional (Table 5), meaning that it
656 may be difficult to obtain the mapping precision required for positional cloning unless measurements can be
657 improved, or the experimental complexity can be reduced. Relatively strong seed-size and seed-yield QTLs were
658 associated with LG06 in both M26 x M35 and C3-2331 x C3-2595 families (Figure 1, Table 5) and the LG06 SELE
659 QTL was the only QTL that was significant across all five environments (Table 4), but additional research is needed
660 to determine if these LG06 QTLs were caused by the *Z SI* gene or some other candidate gene such as *GIF*, *SDW3*,
661 *SUS2*, or *GL7* (Figure 1) before positional cloning should be considered.

662

663 **Implications for IWG domestication and improvement**

664 Transgressive genetic variation, where progeny (Table 2) exceed parents (Table 1), and 111 significant QTLs were
665 observed for 17 traits related to seed size, reproductive fertility, inflorescence capacity, stem and inflorescence
666 compactness, seed retention in the field, percentage of naked seeds after threshing, and maturity in the M26 x M35
667 family. Theoretically, fixation of the optimum QTL genotypes could improve trait means in the M26 x M35 family
668 (Table 2) by an average of 12% to 118% across all five environments (Table 6) but this would require a uniform
669 hybrid with one copy of the best M26 QTL allele and one copy of the best M35 allele for each QTL. It may be
670 possible to make greater improvements by fixing the best M26 allele or the best M36 QTL allele, as homozygous
671 genotypes, but the breeding values of these genotypes are unknown because all progeny contained one M26 allele (*a*
672 or *b*) and one M35 allele (*c* or *d*). Thus, marker-assisted selection for most of these traits may be challenging
673 considering the number and complexity of QTL effects and the fact that the M26 x M35 family is a small sample of
674 the genetic variation present in genetically heterogeneous IWG grain breeding populations (Zhang et al. 2016).
675 However, comparisons of QTLs and traits from the M26 x M35 family to those of fully domesticated grain crops,
676 such as wheat, will elucidate pathways for successful IWG domestication.

677 Grain weights of wheat normally range from about 27 to 60 mg per grain among wheat landraces and
678 cultivars (Abbo et al. 2014; Gegas et al. 2010; Ma et al. 2016) with up to 65 seeds per spike (Wang et al. 2010).
679 Seed yield potential of wheat normally ranges from about 1,500 to 2,400 mg per spike, but up 5,000 mg per spike
680 has been reported for Tibetan Triple-Spikelet wheats that have up to 121 seeds per spike (Yang et al. 2005). With up
681 to 37 spikelets per inflorescence and four seeds per spikelet observed on some IWG genets, the total number of
682 seeds possible on each spike (148) exceeds that of wheat or barley. With nearly 9 mg per seed on some M26 x M35
683 IWG genets (Table 2) and up to 12 or 14 mg per seed on others IWG plants (Cattani and Asselin 2018; Zhang et al.
684 2017), the maximum theoretical seed yield per spike would be nearly 2,100 mg. The best M26 x M35 genets
685 produced no more than 525 mg of seed per spike, but up to and 932 mg of seed per spike has been reported for some
686 IWG plants (Cattani 2017). With an average of about seven (Table 2) to nine (Cattani and Asselin 2018) florets per
687 spikelet, up to 37 spikelets per head (Table 2), and grain weights of 12 to 14 mg the maximum theoretical seed yield
688 per spike for IWG would be from 3,100 to 4,600 mg if every floret produced seed. However, we never observed
689 more than 49% floret fertility in the M26 x M35 family (Table 2). Fertility is considered one of the limiting factors
690 in grass seed production (Armstead et al. 2008) and we observed a relatively large number of QTLs with relatively
691 large QTL effects for SEFL and SESP traits in the M26 x M35 family (Tables 5 and 6). These findings indicate that
692 selection for greater fertility may be a promising avenue to improve IWG seed yields, with potential to match the
693 seed-yield per spike of wheat, but domesticated crops often have fewer and larger fruits or grains compared to their
694 progenitors (Doebley et al. 2006).

695 Seed retention, free-threshing (naked) seed, and increased seed size are all key domestication or
696 improvement traits of most grain crops (Abbo et al. 2014; Doebley et al. 2006; Gegas et al. 2010; Kovach et al.
697 2007; Lenser and Theißen 2013; Liu et al. 2016; Meyer and Purugganan 2013) including IWG. Compared to other
698 traits evaluated in this experiment, IWG seed shattering (SESH) was controlled by a relatively small number of
699 QTLs with relatively large QTL effects (Table 6), including the highest LOD score in this study (Table 5), which is
700 similar to observations in other grasses and grain crops (Doebley et al. 2006; Doust et al. 2014; Kovach et al. 2007;
701 Larson and Kellogg 2009; Pourkheirandish et al. 2015; Simons et al. 2006). Conversely, genetic factors controlling
702 seed-threshing (SENA), seed-size (SEMA, SEAR, SELE, and SEWI), and other traits such as maturity (ZAMA)
703 were substantially more complex. Compared to other traits, SENA and ZAMA were controlled by relatively large
704 numbers of QTLs (Tables 5 and 6) that explained more phenotypic variation in the M26 x M35 family (Table 7).
705 Although we only detected four SEMA QTLs, there were a total of at least 12 loci affecting at least one of the four
706 seed-size traits (SEMA, SEAR, SELE, and SEWI). Interestingly, the relative magnitudes of Avg-total QTL effects
707 (Table 6) and PVE (Table 7) for SENA and SEMA were similar to the expected selection gains of 181% and 60%
708 over five cycles of selection for these two traits, respectively (DeHaan et al. 2018). The average seed weights of
709 IWG forage varieties, ranging from about 2 to 8 mg per seed (Berdahl and Frank 1998; Schulz-Schaeffer and Haller
710 1987), overlaps with the associated with IWG plants selected for grain production, which range from 3 to 14 mg per
711 seed as reported in this and other studies (Cattani and Asselin 2018; Zhang et al. 2017). Similar overlap in seed sizes
712 exists among wild and domestic forms of barley, wheat, and rice (Abbo et al. 2014; Fuller 2007; Gegas et al. 2010),
713 but kernel weight has reportedly increased 10-fold during the domestication of maize (Liu et al. 2016). Evidence

714 from this and other studies demonstrate that IWG seed size can be improved with some rapid initial gains (Cattani
715 and Asselin 2018; DeHaan et al. 2018; Zhang et al. 2017), but additional improvements in seed size may be slow
716 and incremental or require new discoveries and breakthroughs. Conversely, relatively strong QTL effects for SENA
717 and SESH, discussed above, suggest that free-threshing (naked) seed and strong seed retention are attainable
718 domestication traits (Abbo et al. 2014; Doebley et al. 2006) that will clearly distinguish wild and domestic forms of
719 IWG.

720

721 **Conclusions**

722 A total of 111 QTLs were detected for 17 variable traits in the M26 x M35 family including several large-effect
723 QTLs responsible for critical IWG domestication and improvement traits related to fertility, inflorescence
724 architecture, plant height, seed retention, seed size and seed-threshing. The magnitude of M26 x M35 QTL effects,
725 heritabilities and range of phenotypic variation observed in this and other studies (Cattani and Asselin 2018; Zhang
726 et al. 2017) demonstrate potential to fix critical domestication traits, including seed retention and free-threshing
727 (naked) seed, and improve other important grain production traits in IWG. With up to four possible alleles for each
728 M26 x M35 QTL and even greater complexity in genetically heterogeneous breeding populations, the prospects for
729 marker-assisted selection are not certain. However, identification of genes and loci directly associated with critical
730 IWG domestication and improvement traits will enable better management and utilization of IWG germplasm. A
731 total 42 domestication orthogenes, including the wheat free-threshing *Q* (Simons et al. 2006) and *reduced-height*
732 *green revolution* (Peng et al. 1999) genes, aligned to one or more relevant QTLs in this experiment. Closely-linked
733 *DUF247* and *USP* genes on IWG CHR06 corresponding to the *Z* self-incompatibility locus of *Secale* and *Lolium*
734 (Hackauf and Wehling 2005; Shinozuka et al. 2010) showed possible alignment to seed-size, fertility, and seed-yield
735 QTLs on IWG LG06 suggesting that diversity of SI genes may be a limiting factor for seed production in the full-sib
736 IWG families. A large-effect QTL, with a LOD score of nearly 15 for seed shattering, had synergistic effects
737 resulting in greater seed retention and more free-threshing (naked) seed in the M26 x M35 family. Although no
738 candidate genes were associated with this large-effect QTL, recent advancements in genome sequencing and
739 genotyping provide useful approaches to pinpoint genes or loci responsible for domestication of other crops (Tang et
740 al. 2010) such as IWG. Herein, two DNA markers with relatively strong effects on IWG seed size across
741 independent QTL and association mapping experiments (Zhang et al. 2017) were located directly within IWG *GDSL*
742 *lipase* (Clauß et al. 2011; Huang et al. 2015; Jiang et al. 2012; Lai et al. 2017; Ma et al. 2018; Tiwari et al. 2016) and
743 *Xylogalacturonan (XGA) xylosyltransferase* (Jensen et al. 2008; Zhang et al. 2015) orthogenes, which may be
744 plausible candidate genes for these IWG seed-size QTLs. Results of this study demonstrated the increasing power of
745 high-density genotyping (Kantarski et al. 2017), genome sequencing, and QTL mapping to elucidate pathways for
746 the domestication and improvement of a new and genetically-complex perennial grain crop.

747

748 **Author Contribution Statement**

749 SL and LD conceived M26 x M35 QTL experiment and conducted field evaluations. SL performed QTL analyses
750 and wrote manuscript. JP and TK genotyped the M26 x M35 plants. XZ and JA conducted C3-2331 x C3-2595 QTL

751 and AM experiments and provided unpublished data from these experiments. TK, SL, XZ, LD, JA, and JP
752 developed linkage maps used for all QTL and AM analyses. JP, LD, KD, and SL conceived the IWG genome
753 sequencing project as described in JGI Proposal Id. 1997. JS, JG, JJ, and SS assembled and annotated the IWG draft
754 genome sequence. JC and MR assisted with data management and data analysis. KJ provided guidance and support
755 for field and trait evaluations in Utah. All authors read and reviewed the manuscript.

756

757 **Acknowledgments**

758 This research was supported by The Malone Family Land Preservation Foundation. The work conducted by the US
759 Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy
760 under Contract No. DE-AC02-05CH11231. KD is supported by USDA-NIFA Post Doctoral Fellowships grant no.
761 2017-67012-26129/project accession no. 1011622 “Exploring the Genomic Landscape of Perenniality within the
762 Triticeae”. The authors wish to thank Martin Mascher and LiangLiang Gao for advice and contributions to the
763 continuing development of the *Thinopyrum intermedium* genome assembly.

764

765 **Conflict of Interest**

766 On behalf of all authors, the corresponding author states that there is no conflict of interest.

767

768 **References**

769 (!!! INVALID CITATION !!! (Manzanares et al. 2016; Thorogood et al. 2017)).

770 Abbo S, Pinhasi van-Oss R, Gopher A, Saranga Y, Ofner I, Peleg Z (2014) Plant domestication versus crop
771 evolution: a conceptual framework for cereals and grain legumes. Trends Plant Sci 19:351-360.
772 doi:<https://doi.org/10.1016/j.tplants.2013.12.002>

773 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol
774 215:403-410. doi:10.1016/S0022-2836(05)80360-2

775 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-
776 BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.
777 doi:10.1093/nar/25.17.3389

778 Armstead IP, Turner LB, Marshall AH, Humphreys MO, King IP, Thorogood D (2008) Identifying genetic
779 components controlling fertility in the outcrossing grass species perennial ryegrass (*Lolium perenne*) by
780 quantitative trait loci analysis and comparative genetics. New Phytol 178:559-571. doi:10.1111/j.1469-
781 8137.2008.02413.x

782 Ashikari M et al. (2005) Cytokinin oxidase regulates rice grain production. Science 309:741-745.
783 doi:10.1126/science.1113373

784 Balanzà V, Roig-Villanova I, Di Marzo M, Masiero S, Colombo L (2016) Seed abscission and fruit dehiscence
785 required for seed dispersal rely on similar genetic networks. Development 143:3372-3381.
786 doi:10.1242/dev.135202

787 Bell LW, Harrison MT, Kirkegaard JA (2015) Dual-purpose cropping – capitalising on potential grain crop grazing
788 to enhance mixed-farming profitability. *Crop and Pasture Science* 66:i-iv. doi:10.1071/CPv66n4_FO
789 Berdahl JD, Frank AB (1998) Seed maturity in four cool-season forage grasses. *Agron J* 90:483-488
790 Cattani D, Asselin S (2018) Has Selection for Grain Yield Altered Intermediate Wheatgrass? *Sustainability* 10:688
791 Cattani DJ (2017) Selection of a perennial grain for seed productivity across years: Intermediate wheatgrass as a test
792 species. *Can J Plant Sci* 97:516-524. doi:10.1139/cjps-2016-0280
793 Cattani DJ, Asselin SR (2017) Extending the growing season: forage seed production and perennial grains. *Can J*
794 *Plant Sci* 98:235-246. doi:10.1139/cjps-2017-0212
795 Ceoloni C, Kuzmanovic L, Forte P, Virili ME, Bitti A (2015) Wheat-perennial Triticeae introgressions: Major
796 achievements and prospects. In: *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and*
797 *Genomics*. pp 273-313. doi:10.1007/978-3-319-23494-6-11
798 Chen G, Li H, Wei Y, Zheng Y-L, Zhou M, Liu C (2016) Pleiotropic effects of the semi-dwarfing gene *uzu* in
799 barley. *Euphytica* 209:749-755. doi:10.1007/s10681-016-1668-4
800 Chono M et al. (2003) A semidwarf phenotype of barley *uzu* results from a nucleotide substitution in the gene
801 encoding a putative brassinosteroid receptor. *Plant Physiol* 133:1209-1219. doi:10.1104/pp.103.026195
802 Clauß K et al. (2011) Overexpression of sinapine esterase *BnSCE3* in oilseed rape seeds triggers global changes in
803 seed metabolism. *Plant Physiol* 155:1127-1145. doi:10.1104/pp.110.169821
804 Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ (2007) Control of flowering time
805 in temperate cereals: genes, domestication, and sustainable productivity. *J Exp Bot* 58:1231-1244
806 Cox TS, Glover JD, van Tassel DL, Cox CM, DeHaan LR (2006) Prospects for developing perennial grain crops.
807 *Bioscience* 56:649-659
808 Cox TS, Van Tassel DL, Cox C, DeHaan L (2010) Progress in breeding perennial grains. *Crop and Pasture Science*
809 61:513-521. doi:10.1071/CP09201
810 Culman SW, Snapp SS, Ollenburger M, Basso B, DeHaan LR (2013) Soil and water quality rapidly responds to the
811 perennial grain *Kernza* wheatgrass. *Agron J* 105:735-744. doi:10.2134/agronj2012.0273
812 Curwen-McAdams C, Jones SS (2017) Breeding perennial grain crops based on wheat. *Crop Sci* 57:1172-1188.
813 doi:10.2135/cropsci2016.10.0869
814 DeHaan L, Christians M, Crain J, Poland J (2018) Development and evolution of an intermediate wheatgrass
815 domestication program. *Sustainability* 10:1499
816 DeHaan LR, Ismail BP (2017) Perennial cereals provide ecosystem benefits. *Cereal Foods World* 62:278-281.
817 doi:10.1094/CFW-62-6-0278
818 DeHaan LR et al. (2016) A pipeline strategy for grain crop domestication. *Crop Sci* 56:917-930.
819 doi:10.2135/cropsci2015.06.0356
820 Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD (1995) Structural evolution of wheat chromosomes 4A,
821 5A, and 7B and its impact on recombination. *Theor Appl Genet* 91:282-288. doi:10.1007/BF00220890
822 Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature* 386:485-488.
823 doi:10.1038/386485a0

824 Doebley JF, Gaut BS, Smith BD (2006) The Molecular Genetics of Crop Domestication. *Cell* 127:1309-1321.
825 doi:<https://doi.org/10.1016/j.cell.2006.12.006>

826 Doust AN, Mauro-Herrera M, Francis AD, Shand LC (2014) Morphological Diversity and Genetic Regulation of
827 Inflorescence Abscission Zones in Grasses. *Am J Bot* 101:1759-1769

828 Dubcovsky J, Chen CL, Yan LL (2005) Molecular characterization of the allelic variation at the *VRN-H2*
829 vernalization locus in barley. *Mol Breed* 15:395-407. doi:10.1007/s11032-005-0084-6

830 Dubcovsky J, Lijavetzky D, Appendino L, Tranquilli G (1998) Comparative RFLP mapping of *Triticum*
831 *monococcum* genes controlling vernalization requirement. *Theor Appl Genet* 97:968-975.
832 doi:10.1007/s001220050978

833 Faris JD, Fellers JP, Brooks SA, Gill BS (2003) A bacterial artificial chromosome contig spanning the major
834 domestication locus Q in wheat and identification of a candidate gene. *Genetics* 164:311-321

835 Faris JD, Gill BS (2002) Genomic targeting and high-resolution mapping of the domestication gene *Q* in wheat.
836 *Genome* 45:706-718

837 Fuller DQ (2007) Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical
838 Insights from the Old World. *Ann Bot* 100:903-924. doi:10.1093/aob/mcm048

839 Gallavotti A et al. (2004) The role of *barren stalk1* in the architecture of maize. *Nature* 432:630.
840 doi:10.1038/nature03148

841 Galli M et al. (2015) Auxin signaling modules regulate maize inflorescence architecture. *Proc Natl Acad Sci USA*
842 112:13372-13377. doi:10.1073/pnas.1516473112

843 Gegas VC et al. (2010) A Genetic Framework for Grain Size and Shape Variation in Wheat. *The Plant Cell* 22:1046-
844 1056. doi:10.1105/tpc.110.074153

845 Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of *CONSTANS*-like gene families in barley,
846 rice, and *Arabidopsis*. *Plant Physiol* 131:1855-1867. doi:10.1104/pp.102.016188

847 Hackauf B, Wehling P (2005) Approaching the self-incompatibility locus Z in rye (*Secale cereale* L.) via
848 comparative genetics. *Theor Appl Genet* 110:832-845. doi:10.1007/s00122-004-1869-4

849 Harmoney KR (2015) Cool-season grass biomass in the southern mixed-grass prairie region of the USA. *Bioenergy*
850 *Research* 8:203-210. doi:10.1007/s12155-014-9514-9

851 Hayes R et al. (2018) The performance of early-generation perennial winter cereals at 21 sites across four
852 continents. *Sustainability* 10:1124

853 Holland JB, Nyquist WE, Cervantes-Martínez CT (2010) Estimating and Interpreting Heritability for Plant
854 Breeding: An Update. *Plant Breeding Reviews*. doi:10.1002/9780470650202.ch2

855 Hou J, Jiang Q, Hao C, Wang Y, Zhang H, Zhang X (2014) Global selection on sucrose synthase haplotypes during
856 a century of wheat breeding. *Plant Physiol* 164:1918-1929. doi:10.1104/pp.113.232454

857 Houston K et al. (2013) Variation in the interaction between alleles of *HvAPETALA2* and microRNA172 determines
858 the density of grains on the barley inflorescence. *Proc Natl Acad Sci USA* 110:16675-16680.
859 doi:10.1073/pnas.1311681110

860 Hu M-J et al. (2016) Cloning and characterization of *TaTGW-7A* gene associated with grain weight in wheat via
861 SLAF-seq-BSA. *Front Plant Sci* 7:1902. doi:10.3389/fpls.2016.01902

862 Huang L-M, Lai C-P, Chen L-FO, Chan M-T, Shaw J-F (2015) *Arabidopsis SFAR4* is a novel GDSL-type esterase
863 involved in fatty acid degradation and glucose tolerance. *Bot Stud* 56:33. doi:10.1186/s40529-015-0114-6

864 Huang X et al. (2009) Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet* 41:494-497.
865 doi:10.1038/ng.352

866 Ishimaru K et al. (2013) Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and
867 increases yield. *Nat Genet* 45:707. doi:10.1038/ng.2612

868 Jensen JK et al. (2008) Identification of a Xylogalacturonan Xylosyltransferase Involved in Pectin Biosynthesis in
869 *Arabidopsis*. *The Plant Cell* 20:1289-1302. doi:10.1105/tpc.107.050906

870 Jensen KB, Yan X, Larson SR, Wang RRC, Robins JG, McIntosh RA (2016) Agronomic and genetic diversity in
871 intermediate wheatgrass (*Thinopyrum intermedium*). *Plant Breeding* 135:751-758. doi:10.1111/pbr.12420

872 Jensen KB, Zhang YF, Dewey DR (1990) Mode of Pollination of Perennial Species of the Triticeae in Relation to
873 Genomically Defined Genera. *Can J Plant Sci* 70:215-225

874 Jia QJ, Zhang JJ, Westcott S, Zhang XQ, Bellgard M, Lance R, Li CD (2009) GA-20 oxidase as a candidate for the
875 semidwarf gene *sdw1/denso* in barley. *Functional & Integrative Genomics* 9:255-262

876 Jiang Q, Hou J, Hao C, Wang L, Ge H, Dong Y, Zhang X (2011) The wheat (*T. aestivum*) sucrose synthase 2 gene
877 (*TaSus2*) active in endosperm development is associated with yield traits. *Functional & Integrative*
878 *Genomics* 11:49-61. doi:10.1007/s10142-010-0188-x

879 Jiang Y, Chen R, Dong J, Xu Z, Gao X (2012) Analysis of GDSL lipase (GLIP) family genes in rice (*Oryza sativa*).
880 *Plant OMICS* 5:351-358

881 Jungers JM, DeHaan LR, Betts KJ, Sheaffer CC, Wyse DL (2017) Intermediate wheatgrass grain and forage yield
882 responses to nitrogen fertilization. *Agron J* 109:462-472. doi:10.2134/agronj2016.07.0438

883 Kantarski T, Larson S, Zhang X, DeHaan L, Borevitz J, Anderson J, Poland J (2017) Development of the first
884 consensus genetic map of intermediate wheatgrass (*Thinopyrum intermedium*) using genotyping-by-
885 sequencing. *Theor Appl Genet* 130:137-150. doi:10.1007/s00122-016-2799-7

886 Kapazoglou A et al. (2010) Epigenetic chromatin modifiers in barley: IV. The study of barley *Polycomb* group
887 (*PcG*) genes during seed development and in response to external ABA. *BMC Plant Biol* 10:73.
888 doi:10.1186/1471-2229-10-73

889 Karsai I et al. (2005) The *Vrn-H2* locus is a major determinant of flowering time in a facultative x winter growth
890 habit barley (*Hordeum vulgare* L.) mapping population. *Theor Appl Genet* 110:1458-1466.
891 doi:10.1007/s00122-005-1979-7

892 Kenneth PV, Kevin JJ (2001) Adaptation of perennial Triticeae to the eastern Central Great Plains. *Journal of Range*
893 *Management* 54:674-679. doi:10.2307/4003670

894 Klaas M et al. (2011) Progress towards elucidating the mechanisms of self-incompatibility in the grasses: further
895 insights from studies in *Lolium*. *Ann Bot* 108:677-685. doi:10.1093/aob/mcr186

896 Knowles RP (1977) Recurrent mass selection for improved seed yields in intermediate wheatgrass. *Crop Sci* 17:51-
897 54. doi:10.2135/cropsci1977.0011183X001700010015x

898 Komatsuda T et al. (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class
899 homeobox gene. *Proc Natl Acad Sci USA* 104:1424-1429

900 Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering
901 during rice domestication. *Science* 312:1392-1396. doi:10.1126/science.1126410

902 Koppolu R et al. (2013) *Six-rowed spike4 (Vrs4)* controls spikelet determinacy and row-type in barley. *Proc Natl*
903 *Acad Sci USA* 110:13198-13203. doi:10.1073/pnas.1221950110

904 Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice domestication. *Trends Genet*
905 23:578-587. doi:<https://doi.org/10.1016/j.tig.2007.08.012>

906 Krupinsky JM, Berdahl JD (2000) Selecting resistance to *Bipolaris sorokiniana* and *Fusarium graminearum* in
907 intermediate wheatgrass. *Plant Dis* 84:1299-1302. doi:10.1094/PDIS.2000.84.12.1299

908 Kuczyńska A, Mikołajczak K, Ćwiek H (2014) Pleiotropic effects of the *sdw1* locus in barley populations
909 representing different rounds of recombination. *Electron J Biotechnol* 17:217-223.
910 doi:10.1016/j.ejbt.2014.07.005

911 Kumar S, You FM, Cloutier S (2012) Genome wide SNP discovery in flax through next generation sequencing of
912 reduced representation libraries. *BMC Genomics* 13:684. doi:10.1186/1471-2164-13-684

913 La Rota M, Sorrells ME (2004) Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity
914 of genome relationships between rice and wheat. *Funct Integr Genomics* 4:34-46. doi:10.1007/s10142-003-
915 0098-2

916 Lai Y et al. (2017) Association mapping of grain weight, length and width in barley (*Hordeum vulgare*) breeding
917 germplasm. *Int J Agric Biol* 19:1175-1186. doi:10.17957/IJAB/15.0406

918 Lang T et al. (2018) Precise identification of wheat – *Thinopyrum intermedium* translocation chromosomes carrying
919 resistance to wheat stripe rust in line Z4 and its derived progenies. *Genome* 61:177-185. doi:10.1139/gen-
920 2017-0229

921 Larson S et al. (2017) Development and testing of cool-season grass species, varieties and hybrids for biomass
922 feedstock production in western North America. *Agronomy* 7:3

923 Larson SR, Kellogg EA (2009) Genetic dissection of seed production traits and identification of a major-effect seed
924 retention QTL in hybrid *Leymus* (Triticeae) wildryes. *Crop Sci* 49:29-40. doi:10.2135/cropsci2008.05.0277

925 Larson SR et al. (2012) *Leymus* EST linkage maps identify 4NsL-5NsL reciprocal translocation, wheat-*Leymus*
926 chromosome introgressions, and functionally important gene loci. *Theor Appl Genet* 124:189-206.
927 doi:10.1007/s00122-011-1698-1

928 Laurie DA, Pratchett N, Snape JW, Bezant JH (1995) RFLP mapping of five major genes and eight quantitative trait
929 loci controlling flowering time in a winter x spring barley (*Hordeum vulgare* L.) cross. *Genome* 38:575-
930 585

931 Lee D, Owens VN, Boe A, Koo BC (2009) Biomass and seed yields of big bluestem, switchgrass, and intermediate
932 wheatgrass in response to manure and harvest timing at two topographic positions. *Glob Change Biol*
933 *Bioenergy* 1:171-179. doi:10.1111/j.1757-1707.2009.01008.x

934 Lenser T, Theißen G (2013) Molecular mechanisms involved in convergent crop domestication. *Trends Plant Sci*
935 18:704-714. doi:<https://doi.org/10.1016/j.tplants.2013.08.007>

936 Li B et al. (2013) Constitutive expression of cell wall invertase genes increases grain yield and starch content in
937 maize. *Plant Biotechnol J* 11:1080-1091. doi:10.1111/pbi.12102

938 Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. *Science* 311:1936-1939

939 Li J et al. (2017) Introduction of *Thinopyrum intermedium* ssp. *trichophorum* chromosomes to wheat by trigeneric
940 hybridization involving *Triticum*, *Secale* and *Thinopyrum* genera. *Planta* 245:1121-1135.
941 doi:10.1007/s00425-017-2669-9

942 Li W, Gill BS (2006) Multiple genetic pathways for seed shattering in the grasses. *Funct Integr Genomics* 6:300-
943 309. doi:10.1007/s10142-005-0015-y

944 Li Y et al. (2011) Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat*
945 *Genet* 43:1266. doi:10.1038/ng.977

946 Liu H et al. (2017) Production and molecular cytogenetic characterization of a durum wheat-*Thinopyrum elongatum*
947 7E disomic addition line with resistance to *Fusarium* head blight. *Cytogenet Genome Res* 153:165-173

948 Liu Z, Garcia A, McMullen MD, Flint-Garcia SA (2016) Genetic analysis of kernel traits in maize-teosinte
949 introgression populations. *G3: Genes, Genomes, Genetics* 6:2523-2530. doi:10.1534/g3.116.030155

950 Ma L, Li T, Hao C, Wang Y, Chen X, Zhang X (2016) *TaGS5-3A*, a grain size gene selected during wheat
951 improvement for larger kernel and yield. *Plant Biotechnol J* 14:1269-1280. doi:10.1111/pbi.12492

952 Ma R, Yuan H, An J, Hao X, Li H (2018) A *Gossypium hirsutum* GDSL lipase/hydrolase gene (*GhGLIP*) appears to
953 be involved in promoting seed growth in *Arabidopsis*. *PLoS ONE* 13. doi:10.1371/journal.pone.0195556

954 Manzanares C et al. (2016) A gene encoding a DUF247 domain protein cosegregates with the *S* self-incompatibility
955 locus in perennial ryegrass. *Mol Biol Evol* 33:870-884. doi:10.1093/molbev/msv335

956 Marti A, Bock JE, Pagani MA, Ismail B, Seetharaman K (2016) Structural characterization of proteins in wheat
957 flour doughs enriched with intermediate wheatgrass (*Thinopyrum intermedium*) flour. *Food Chem* 194:994-
958 1002. doi:10.1016/j.foodchem.2015.08.082

959 Mathews S, Sharrock RA (1996) The phytochrome gene family in grasses (Poaceae): A phylogeny and evidence that
960 grasses have a subset of the loci found in dicot angiosperms. *Mol Biol Evol* 13:1141-1150.
961 doi:10.1093/oxfordjournals.molbev.a025677

962 McSteen P, Malcomber S, Skirpan A, Lunde C, Wu X, Kellogg E, Hake S (2007) *barren inflorescence2* Encodes a
963 co-ortholog of the PINOID serine/threonine kinase and is required for organogenesis during inflorescence
964 and vegetative development in maize. *Plant Physiol* 144:1000-1011. doi:10.1104/pp.107.098558

965 Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. *Nature*
966 *Reviews Genetics* 14:840. doi:10.1038/nrg3605

967 <https://www.nature.com/articles/nrg3605#supplementary-information>

968 Monono EM, Nyren PE, Berti MT, Pryor SW (2013) Variability in biomass yield, chemical composition, and
969 ethanol potential of individual and mixed herbaceous biomass species grown in North Dakota. *Ind Crops*
970 *Prod* 41:331-339. doi:10.1016/j.indcrop.2012.04.051

971 Nadolska-Orczyk A, Rajchel IK, Orczyk W, Gasparis S (2017) Major genes determining yield-related traits in wheat
972 and barley. *Theor Appl Genet* 130:1081-1098. doi:10.1007/s00122-017-2880-x

973 Nallamilli BRR, Zhang J, Mujahid H, Malone BM, Bridges SM, Peng Z (2013) Polycomb group gene *OsFIE2*
974 regulates rice (*Oryza sativa*) seed development and grain filling via a mechanism distinct from *Arabidopsis*.
975 *PLoS Genet* 9:e1003322. doi:10.1371/journal.pgen.1003322

976 Namikawa S, Kawakami Z (1934) On the occurrence of the haploid, triploid, and tetraploid plants in twin seedlings
977 of common wheat. *Proc Imp Acad Jap* 10:668-671

978 Ouellette LA, Reid RW, Blanchard SG, Brouwer CR (2018) LinkageMapView—rendering high-resolution linkage
979 and QTL maps. *Bioinformatics* 34:306-307. doi:10.1093/bioinformatics/btx576

980 Pearson CH, Larson SR, Keske CMH, Jensen KB (2015) Native grasses for biomass production at high elevations. .
981 In: Cruz VMV, Dierig DA (eds) *Handbook of plant breeding*. vol [9]. Springer, New York, pp 101-132

982 Peng JR et al. (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400:256-261

983 Pourkheirandish M et al. (2015) Evolution of the grain dispersal system in barley. *Cell* 162:527-539.
984 doi:10.1016/j.cell.2015.07.002

985 R Core Team (2017) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
986 Computing. Vienna, Austria. URL <https://www.R-project.org/>

987 Ral JP et al. (2012) Down-regulation of Glucan, Water-Dikinase activity in wheat endosperm increases vegetative
988 biomass and yield. *Plant Biotechnol J* 10:871-882. doi:10.1111/j.1467-7652.2012.00711.x

989 Revelle W (2018) *psych: Procedures for psychological, psychometric, and personality research*. Northwestern
990 University. Evanston, Illinois, USA. URL <https://CRAN.R-project.org/package=psych>

991 Robins JG (2010) Cool-season grasses produce more total biomass across the growing season than do warm-season
992 grasses when managed with an applied irrigation gradient. *Biomass & Bioenergy* 34:500-505.
993 doi:10.1016/j.biombioe.2009.12.015

994 Ross JG (1963) Registration of Oahe Intermediate Wheatgrass (Reg. No. 5). *Crop Sci* 3:373-373.
995 doi:10.2135/cropsci1963.0011183X000300040046x

996 Ryan MR, Crews TE, Culman SW, DeHaan LR, Hayes RC, Jungers JM, Bakker MG (2018) Managing for
997 multifunctionality in perennial grain crops. *Bioscience* 68:294-304. doi:10.1093/biosci/biy014

998 Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*
999 9:671-675

1000 Schulz-Schaeffer J, Haller SE (1987) Registration of Montana-2 Perennial \times Agrotriticum *Intermedioidurum*
1001 *Khizhnyak*. *Crop Sci* 27:822-823. doi:10.2135/cropsci1987.0011183X002700040058x

1002 Shinozuka H, Cogan NO, Smith KF, Spangenberg GC, Forster JW (2010) Fine-scale comparative genetic and
1003 physical mapping supports map-based cloning strategies for the self-incompatibility loci of perennial
1004 ryegrass (*Lolium perenne* L.). *Plant Mol Biol* 72:343-355. doi:10.1007/s11103-009-9574-y

1005 Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with
1006 grain size increased yields during rice domestication. *Nat Genet* 40:1023-1028. doi:10.1038/ng.169
1007 Shu X, Rasmussen S (2014) Quantification of amylose, amylopectin, and β -glucan in search for genes controlling
1008 the three major quality traits in barley by genome-wide association studies. *Front Plant Sci* 5:197.
1009 doi:10.3389/fpls.2014.00197
1010 Simons KJ, Fellers JP, Trick HN, Zhang ZC, Tai YS, Gill BS, Faris JD (2006) Molecular characterization of the
1011 major wheat domestication gene *Q*. *Genetics* 172:547-555. doi:10.1534/genetics.105.044727
1012 Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously
1013 unknown RING-type E3 ubiquitin ligase. *Nat Genet* 39:623-630. doi:10.1038/ng2014
1014 Studer B, Jensen LB, Hentrup S, Brazauskas G, Kölliker R, Lübberstedt T (2008) Genetic characterisation of seed
1015 yield and fertility traits in perennial ryegrass (*Lolium perenne* L.). *Theor Appl Genet* 117:781-791.
1016 doi:10.1007/s00122-008-0819-y
1017 Su Z, Hao C, Wang L, Dong Y, Zhang X (2011) Identification and development of a functional marker of *TaGW2*
1018 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theor Appl Genet* 122:211-223.
1019 doi:10.1007/s00122-010-1437-z
1020 Sun PY et al. (2016) *OsGRF4* controls grain shape, panicle length and seed shattering in rice. *J Integr Plant Biol*
1021 58:836-847. doi:10.1111/jipb.12473
1022 Taketa S et al. (2008) Barley grain with adhering hulls is controlled by an ERF family transcription factor gene
1023 regulating a lipid biosynthesis pathway. *Proc Natl Acad Sci USA* 105:4062-4067.
1024 doi:10.1073/pnas.0711034105
1025 Tanabata T, Shibaya T, Hori K, Ebana K, Yano M (2012) SmartGrain: High-throughput phenotyping software for
1026 measuring seed shape through image analysis. *Plant Physiol* 160:1871-1880. doi:10.1104/pp.112.205120
1027 Tang H, Sezen U, Paterson AH (2010) Domestication and plant genomes. *Curr Opin Plant Biol* 13:160-166.
1028 doi:<https://doi.org/10.1016/j.pbi.2009.10.008>
1029 Thorogood D et al. (2017) A novel multivariate approach to phenotyping and association mapping of multi-locus
1030 gametophytic self-incompatibility reveals *S*, *Z*, and other loci in a perennial ryegrass (Poaceae) population.
1031 *Front Plant Sci* 8. doi:10.3389/fpls.2017.01331
1032 Tiwari GJ, Chiang MY, De Silva JR, Song BK, Lau YL, Rahman S (2016) Lipase genes expressed in rice bran:
1033 LOC_Os11g43510 encodes a novel rice lipase. *J Cereal Sci* 71:43-52. doi:10.1016/j.jcs.2016.07.008
1034 Tulpan D, Leger S (2017) The Plant Orthology Browser: An Orthology and Gene-Order Visualizer for Plant
1035 Comparative Genomics. *The Plant Genome* 10. doi:10.3835/plantgenome2016.08.0078
1036 Uzma, Kubra G, Gul A, Mujeeb-Kazi A (2015) Use of alien diversity to combat some major biotic stresses in
1037 *Triticum aestivum* L. In: Hakeem KR (ed) *Crop Production and Global Environmental Issues*. pp 319-347.
1038 doi:10.1007/978-3-319-23162-4_14
1039 Van Ooijen JW (2006) JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations.
1040 Kyazma B.V., Wageningen, Netherlands.

1041 Van Ooijen JW (2009) MapQTL 6, Software for the mapping of quantitative trait loci in experimental populations
1042 of diploid species. Kyazma B.V., Wageningen, Netherlands.

1043 Vogel KP, Tober D, Reece PE, Baltsensperger DD, Schuman G, Nicholson RA (2005) Registration of ‘Haymaker’
1044 Intermediate Wheatgrass Registration by CSSA. *Crop Sci* 45:415-416. doi:10.2135/cropsci2005.0415

1045 Vu GT, Wicker T, Buchmann JP, Chandler PM, Matsumoto T, Graner A, Stein N (2010) Fine mapping and syntenic
1046 integration of the semi-dwarfing gene *sdw3* of barley. *Funct Integr Genomics* 10:509-521.
1047 doi:10.1007/s10142-010-0173-4

1048 Wagoner P (1990) Perennial grain new use for intermediate wheatgrass. *Journal of Soil and Water Conservation*
1049 45:81-82

1050 Wang E et al. (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication.
1051 *Nat Genet* 40:1370-1374. doi:10.1038/ng.220

1052 Wang GJ et al. (2014) Establishment and yield of perennial grass monocultures and binary mixtures for bioenergy in
1053 North Dakota. *Agron J* 106:1605-1613. doi:10.2134/agronj14.0068

1054 Wang H et al. (2005) The origin of the naked grains of maize. *Nature* 436:714-719. doi:10.1038/nature03863

1055 Wang J, Liao X, Li Y, Zhou R, Yang X, Gao L, Jia J (2010) Fine mapping a domestication-related QTL for spike-
1056 related traits in a synthetic wheat. *Genome* 53:798-804. doi:10.1139/g10-066

1057 Wang S et al. (2012) Control of grain size, shape and quality by *OsSPL16* in rice. *Nat Genet* 44:950-954.
1058 doi:10.1038/ng.2327

1059 Wang Y et al. (2015) Copy number variation at the *GL7* locus contributes to grain size diversity in rice. *Nat Genet*
1060 47:944-948. doi:10.1038/ng.3346

1061 Wendt T et al. (2016) *HvDep1* is a positive regulator of culm elongation and grain size in barley and impacts yield
1062 in an environment-dependent manner. *PLoS One* 11. doi:10.1371/journal.pone.0168924

1063 Weng J et al. (2008) Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and
1064 weight. *Cell Res* 18:1199-1209. doi:10.1038/cr.2008.307

1065 Win KT et al. (2017) A single base change explains the independent origin of and selection for the nonshattering
1066 gene in African rice domestication. *New Phytol* 213:1925-1935. doi:10.1111/nph.14290

1067 Wu YZ, Fu YC, Zhao SS, Gu P, Zhu ZF, Sun CQ, Tan LB (2016) *CLUSTERED PRIMARY BRANCH 1*, a new allele
1068 of *DWARF11*, controls panicle architecture and seed size in rice. *Plant Biotechnol J* 14:377-386.
1069 doi:10.1111/pbi.12391

1070 Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat
1071 vernalization gene *VRN1*. *Proc Natl Acad Sci USA* 100:6263-6268. doi:10.1073/pnas.0937399100

1072 Yan LL et al. (2004) The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science*
1073 303:1640-1644. doi:10.1126/science.1094305

1074 Yan S et al. (2011) Seed size is determined by the combinations of the genes controlling different seed
1075 characteristics in rice. *Theor Appl Genet* 123:1173-1181. doi:10.1007/s00122-011-1657-x

1076 Yang W-Y, Lu B-R, Hu X-R, Yu Y, Zhang Y (2005) Inheritance of the Triple-spikelet Character in a Tibetan
1077 Landrace of Common Wheat. *Genet Resour Crop Evol* 52:847-851. doi:10.1007/s10722-003-6089-2

1078 Zadoks JC, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. *Weed Res* 14:415-421

1079 Zair W, Maxted N, Amri A (2018) Setting conservation priorities for crop wild relatives in the Fertile Crescent.

1080 *Genet Resour Crop Evol* 65:855-863. doi:10.1007/s10722-017-0576-3

1081 Zhang L, Zhao YL, Gao LF, Zhao GY, Zhou RH, Zhang BS, Jia JZ (2012a) *TaCKX6-D1*, the ortholog of rice

1082 *OsCKX2*, is associated with grain weight in hexaploid wheat. *New Phytol* 195:574-584.

1083 doi:10.1111/j.1469-8137.2012.04194.x

1084 Zhang X et al. (2017) Uncovering the genetic architecture of seed weight and size in intermediate wheatgrass

1085 through linkage and association mapping. *Plant Genome* 10:0. doi:10.3835/plantgenome2017.03.0022

1086 Zhang X et al. (2012b) Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a

1087 significant yield increase in rice. *Proc Natl Acad Sci USA* 109:21534-21539.

1088 doi:10.1073/pnas.1219776110

1089 Zhang XF et al. (2016) Establishment and optimization of genomic selection to accelerate the domestication and

1090 improvement of intermediate wheatgrass. *Plant Genome* 9. doi:10.3835/plantgenome2015.07.0059

1091 Zhang Y et al. (2015) Establishment of a 100-seed weight quantitative trait locus–allele matrix of the germplasm

1092 population for optimal recombination design in soybean breeding programmes. *J Exp Bot* 66:6311-6325.

1093 doi:10.1093/jxb/erv342

1094

1095 **Figure caption**

1096

1097 **Figure 1.** Alignment of quantitative trait loci (QTLs), association-mapping markers, and possible candidate genes to
1098 21 linkage groups (LG01 – LG21) and 21 chromosome sequences (CHR01 – CHR21) corresponding to seven
1099 homoeologous groups (HG1 – HG7) of allotetraploid ($2n=6x=42$) intermediate wheatgrass. The 2-LOD drop-off
1100 intervals for 111 M26 x M35 QTLs (Tables 4 – 6) and 44 C3-2331 x C3-2595 (UMN) seed-size QTLs (Zhang et al.
1101 2017) are indicated by filled and shaded box plots, respectively, on the right side of each linkage group. The physical
1102 position of UMN seed-size AM markers (Zhang et al. 2017) and possible candidate genes (Table 8) are shown in
1103 bold and italic text, respectively, on the left side of each chromosome. Graphed portions of each linkage group and
1104 chromosome sequence are scaled in centiMorgans (cM) in nucleotide megabases (MB), respectively, and
1105 normalized to comparable lengths.

Table 1 Trait means with least significant differences (LSD) for cycle-3 (C3) grandparents and cycle-4 (C4) parents of the full-sib M26 x M36 population

Trait description	Trait	Units	C3 grandparents		C4 parents		LSD ^c	LSD levels
			C3_3471	C3_3941	M26 ^a	M35 ^b		
Seed mass	SEMA	mg	7.09 A	5.09 B	5.55 B	5.44 B	0.51	AB
Seed area	SEAR	mm ²	5.56 A	4.50 C	4.56 C	4.79 B	0.18	ABC
Seed length	SELE	mm	5.29 A	4.22 D	4.61 B	4.36 C	0.11	ABCD
Seed width	SEWI	mm	1.35 C	1.38 B	1.27 D	1.43 A	0.03	ABCD
Seed shattering	SESH	%	5.0 C	30.0 B	28.8 B	59.4 A	0.6	ABC
Seed nakedness	SENA	%	94.1 A	28.7 C	71.3 B	30.3 C	2.7	ABC
Spiklets inflorescence ⁻¹	SPIN	no.	22.8 A	21.6 B	22.9 A	19.0 C	0.9	ABC
Florets spikelet ⁻¹	FLSP	no.	6.63 B	6.33 B	6.75 B	7.67 A	0.49	AB
Seeds floret ⁻¹	SEFL	no.	0.27 A	0.06 C	0.26 A	0.18 B	0.05	ABC
Seeds spikelet ⁻¹	SESP	no.	1.95 A	0.60 C	1.87 A	1.52 B	0.25	ABC
Seed yield inflorescence ⁻¹	SYIN	mg	323 A	72 D	237 B	149 C	21	ABCD
Inflorescence length	INLE	cm	28.0 B	27.6 B	30.3 A	26.0 C	1.2	ABC
Stem length	STLE	cm	114 B	128 A	128 A	124 A	6	AB
Stem width	STWI	mm	2.20 B	2.10 B	2.32 A	2.33 A	0.10	AB
Zadok's maturity	ZAMA	0-99	61.7 A	56.3 D	57.6 C	58.7 B	0.7	ABCD
Crown circumference	CRCI	cm	147 AB	145 AB	137 B	156 A	19	AB
Inflorescences crown ⁻¹	INCR	no.	219 A	197 AB	212 AB	184 B	31	AB

^a Hybrid of C3_3471 x C3_3941 grandparents

^b Hybrid of C3_3941 x unknown grandparents

^c Least significant differences, $p \leq 0.05$

Table 2 Trait means, relative standard deviations (RSD), ranges, and broad sense heritabilities with standard errors ($H \pm SE$) for 266 M26 x M35 progeny across five location (KS or UT) x year (2013, 2014, and 2015) environments and overall average across five environments (AVG)

Trait	----- Means \pm RSD (and ranges) -----						----- $H \pm SE$ -----	
	KS13	KS14	KS15	UT14	UT15	AVG	plot basis	genet basis
SEMA	4.57 \pm 20% (2.43 - 8.95)	5.33 \pm 14% (3.44 - 7.71)	5.16 \pm 17% (2.09 - 8.70)	5.59 \pm 17% (2.81 - 8.33)	5.27 \pm 13% (3.27 - 7.83)	5.12 \pm 13% (3.39 - 7.53)	0.43 \pm 0.03	0.82 \pm 0.02
SEAR	3.92 \pm 12% (2.74 - 5.46)	5.58 \pm 10% (4.22 - 7.21)	4.70 \pm 11% (3.05 - 6.28)	5.32 \pm 11% (3.66 - 7.42)	4.47 \pm 10% (3.33 - 6.26)	4.60 \pm 9% (3.41 - 6.09)	0.41 \pm 0.03	0.80 \pm 0.02
SELE	4.22 \pm 7% (3.56 - 4.99)	5.04 \pm 6% (4.20 - 5.92)	5.01 \pm 7% (4.22 - 5.88)	4.79 \pm 6% (4.07 - 5.71)	4.18 \pm 6% (3.51 - 4.90)	4.47 \pm 6% (3.87 - 5.21)	0.54 \pm 0.03	0.87 \pm 0.02
SEWI	1.22 \pm 7% (1.02 - 1.46)	1.45 \pm 6% (1.20 - 1.67)	1.38 \pm 7% (1.01 - 1.59)	1.45 \pm 6% (1.18 - 1.68)	1.35 \pm 5% (1.17 - 1.58)	1.34 \pm 5% (1.12 - 1.51)	0.42 \pm 0.03	0.82 \pm 0.02
SESH	---	29.7 \pm 70% (0.0 - 84.7)	40.0 \pm 49% (0.0 - 87.2)	51.8 \pm 53% (0.0 - 86.8)	12.9 \pm 78% (0.0 - 52.0)	33.7 \pm 52% (0.0 - 67.5)	0.38 \pm 0.03	0.71 \pm 0.3
SENA	67.5 \pm 20% (0.0 - 100.0)	29.6 \pm 83% (0.0 - 87.0)	10.1 \pm 94% (0.0 - 55.0)	77.5 \pm 19% (21.7 - 98.0)	66.1 \pm 17% (26.9 - 92.3)	55.3 \pm 20% (23.1 - 83.5)	0.30 \pm 0.03	0.71 \pm 0.03
SPIN	19.9 \pm 11% (8.1 - 26.6)	20.8 \pm 13% (10.6 - 28.3)	21.8 \pm 11% (14.3 - 28.6)	19.9 \pm 15% (11.6 - 37.1)	18.0 \pm 13% (8.5 - 24.5)	19.7 \pm 11% (11.2 - 29.2)	0.30 \pm 0.03	0.75 \pm 0.02
FLSP	8.12 \pm 17% (3.51 - 12.5)	6.72 \pm 13% (4.51 - 10.2)	6.86 \pm 11% (5.01 - 8.51)	---	---	7.24 \pm 11% (4.68 - 9.92)	0.22 \pm 0.03	0.56 \pm 0.03
SEFL	0.12 \pm 47% (0.00 - 0.47)	0.10 \pm 71% (0.00 - 0.31)	0.22 \pm 44% (0.00 - 0.49)	---	---	0.15 \pm 47% (0.00 - 0.33)	0.43 \pm 0.03	0.77 \pm 0.03
SESP	0.98 \pm 79% (0.00 - 4.4)	0.63 \pm 56% (0.00 - 1.8)	1.46 \pm 44% (0.00 - 3.49)	2.08 \pm 33% (0.03 - 3.70)	1.16 \pm 43% (0.00 - 3.10)	1.17 \pm 43% (0.00 - 2.32)	0.44 \pm 0.03	0.77 \pm 0.02
SYIN	95 \pm 86% (0 - 454)	75 \pm 64% (2 - 247)	170 \pm 48% (2 - 437)	235 \pm 38% (2 - 525)	111 \pm 49% (0 - 274)	120 \pm 53% (0 - 326)	0.50 \pm 0.03	0.84 \pm 0.02
INLE	---	28.4 \pm 15% (18.8 - 41.8)	31.0 \pm 12% (22.8 - 41.8)	25.6 \pm 12% (18.2 - 35.4)	23.4 \pm 14% (16.2 - 33.0)	27.3 \pm 11% (20.1 - 37.3)	0.56 \pm 0.03	0.86 \pm 0.01
STLE	41 \pm 10% (30 - 53)	132 \pm 7% (93 - 153)	141 \pm 6% (108 - 165)	170 \pm 11% (76 - 200)	182 \pm 9% (117 - 219)	121 \pm 11% (26 - 149)	0.47 \pm 0.03	0.82 \pm 0.02
STWI	2.43 \pm 14% (1.33 - 3.34)	1.97 \pm 15% (1.27 - 4.50)	2.18 \pm 13% (1.37 - 2.84)	2.09 \pm 10% (1.51 - 2.68)	1.67 \pm 13% (0.91 - 2.30)	2.10 \pm 10% (1.39 - 2.71)	0.33 \pm 0.03	0.75 \pm 0.02
ZAMA	65.8 \pm 7% (32.0 - 71.9)	59.3 \pm 5% (52.4 - 68.9)	58.7 \pm 4% (53.1 - 65.9)	52.3 \pm 4% (48.1 - 57.0)	47.1 \pm 8% (38.8 - 57.9)	56.6 \pm 5% (44.7 - 63.2)	0.48 \pm 0.03	0.82 \pm 0.02
CRCI	---	131 \pm 15% (81 - 179)	---	127 \pm 23% (59 - 234)	165 \pm 18% (53 - 256)	138 \pm 16% (80 - 196)	0.24 \pm 0.03	0.58 \pm 0.04
INCR	---	160 \pm 24% (37 - 299)	180 \pm 25% (60 - 324)	121 \pm 35% (8 - 280)	244 \pm 35% (7 - 563)	176 \pm 24% (43 - 322)	0.23 \pm 0.03	0.63 \pm 0.03

Table 3 Trait correlations among parents and progeny of M26 x M35 family

	SEMA	SEAR	SELE	SEWI	SESH	SENA	SPIN	FLSP	SEFL	SESP	SYIN	INLE	STLE	STWI	ZAMA	CRCI	INCR
SEMA	.	***	***	***	***	---	---	***	***	***	***	***	***	***	***	---	---
SEAR	0.67	.	***	***	***	---	---	***	***	***	***	***	***	***	***	***	---
SELE	0.65	0.86	.	***	***	---	---	***	***	***	***	***	*	***	***	***	---
SEWI	0.52	0.8	0.42	.	***	***	---	***	***	***	***	***	***	***	***	---	---
SESH	0.24	0.24	0.17	0.26	.	***	---	**	***	***	***	***	***	***	***	*	***
SENA	---	---	---	-0.13	-0.2	.	---	**	---	---	---	---	---	---	---	---	---
SPIN	---	---	---	---	---	---	.	***	---	---	***	***	***	***	***	---	*
FLSP	0.16	0.22	0.13	0.26	0.13	-0.12	0.24	.	---	***	***	***	***	***	***	***	***
SEFL	0.2	0.22	0.23	0.14	0.38	---	---	---	.	***	***	---	***	***	***	---	***
SESP	0.15	0.2	0.16	0.17	0.37	---	---	0.23	0.95	.	***	*	***	***	***	---	***
SYIN	0.44	0.4	0.38	0.29	0.38	---	0.27	0.26	0.82	0.87	.	***	***	***	***	---	***
INLE	0.23	0.27	0.28	0.18	0.11	---	0.44	0.38	---	0.08	0.27	.	***	***	---	---	***
STLE	0.22	0.18	0.08	0.22	0.36	---	0.17	0.17	0.12	0.21	0.26	0.23	.	**	***	***	***
STWI	0.27	0.3	0.21	0.31	0.21	---	0.33	0.49	0.18	0.29	0.39	0.48	0.21	.	***	---	---
ZAMA	0.21	0.15	0.12	0.13	0.28	---	0.11	0.27	0.34	0.32	0.35	---	0.25	0.27	.	***	***
CRCI	---	-0.11	-0.11	---	0.09	---	---	-0.22	---	---	---	---	0.23	---	0.16	.	***
INCR	---	---	---	---	0.14	---	0.08	-0.2	0.21	0.12	0.11	-0.1	0.3	---	0.23	0.55	.

*, **, and *** Significant at $P \leq 0.05$, $P \leq 0.01$ and $P \leq 0.001$ controlling for multiple tests with Bonferroni correction, respectively

^a Based on standard normal deviations of SENA and SESH.

Table 4 Linkage groups (LGs) with significant QTL effects in M26 x M35 population across five location (KS or UT) x year (2013, 2014, and 2015) environments and overall average across five environments (AVG)

Trait	KS13 (LG)	KS14 (LG)	KS15 (LG)	UT14 (LG)	UT15 (LG)	AVG (LG)	Total number of QTLs
SEMA	1*	15*	---	6*	---	6, 11*	5 (4*)
SEAR	6*, 20	7	7*	9	17*,	7, 9*, 17, 20*, 19*	11 (6*)
SELE	6	6, 14*, 21	6, 21, 19	6, 9*	6, 17*, 20*, 21*	6*, 17, 20, 21, 19*	18 (7*)
SEWI	6*, 20	9*, 15*	11*,	8*, 16	15 17*	11, 16*, 20*	12 (8*)
SESH	---	11, 12*	6*, 11	11	14*	10*, 11*	8 (5*)
SENA	2, 15*	2, 10*, 11*, 14*, 15	2*, 10, 11	6, 10	2, 6, 12*	2, 6*, 5*, 10, 18*	20 (9*)
SPIN	3, 6*, 7	10*	3, 10, 13	---	4*, 7, 9*, 14*, 13*	3*, 7*	14 (8*)
FLSP	---	---	9*	---	---	4*, 8*, 9, 16*	5 (4*)
SEFL	6, 12*, 14*, 13*, 15*	6	9*, 10*, 11*	---	---	6*, 9	11 (8*)
SESP	6, 14*, 15*	6*	1*, 6, 7, 10*, 11*	7*, 10	6, 7, 14, 13*	7, 10, 14, 15	19 (8*)
SYIN	6, 12*, 15*	6*, 13*, 21*	1*	4*	---	3*, 6, 4	11 (8*)
INLE	---	9, 12, 13	9, 14*	2*, 9, 12*, 14	5*, 7*, 9, 12, 14, 13	9*, 12, 14, 13*	19 (7*)
STLE	19*,	5*, 9 16	16*	5, 10 13*	9*, 10*, 11*, 16	5, 8*	14 (8*)
STWI	15*	---	6, 9	6, 9	4	6*, 4*, 9*	9 (4*)
ZAMA	15	6, 10*, 14*, 18*, 11, 3	14, 11, 15*	6, 11	6*, 11, 17*, 16*, 3*	6, 10, 14, 11*, 3, 12*, 15	24 (10*)
CRCI	---	---	---	12*, 13	---	13*, 19*	4 (3*)
INCR	---	---	---	10*, 13*, 17*	15*	10, 15	6 (4*)
Total number of QTLs	25 (15*)	32 (17*)	31 (15*)	28 (12*)	37 (22*)	57 (30*)	210 (111*)

* Most significant QTL for each linkage group by trait combination.

Table 5 Peak LOD scores for each linkage group (LG) and homoeologous group (HG) of M26 x M36 family

LG	HG	SEMA	SEAR	SELE	SEWI	SESH	SENA	SPIN	FLSP	SEFL	SESP	SYIN	INLE	STLE	STWI	ZAMA	CRCI	INCR
1	1	3.5 ^b	---	---	---	---	---	---	---	---	5.3 ^c	5.7 ^c	---	---	---	---	---	---
2	1	---	---	---	---	---	8.0 ^c	---	---	---	---	---	5.7 ^c	---	---	---	---	---
3	1	---	---	---	---	---	---	5.2 ^c	---	---	---	3.5 ^a	---	---	---	6.8 ^b	---	---
4	2	---	---	---	---	---	---	5.8 ^c	6.1 ^c	---	---	5.6 ^c	---	---	6.5 ^c	---	---	---
5	2	---	---	---	---	---	6.6 ^c	---	---	---	---	---	5.6	6.3	---	---	---	---
6	2	6.8 ^c	6.0 ^c	12.0 ^c	5.1 ^c	5.0 ^c	5.6 ^a	5.6 ^c	---	8.1 ^c	6.6 ^c	8.4 ^c	---	---	5.1 ^c	5.8 ^c	---	---
7	3	---	5.2 ^c	---	---	---	---	6.6 ^a	---	---	5.3 ^b	---	3.6 ^a	---	---	---	---	---
8	3	---	---	---	7.2 ^c	---	---	---	6.2 ^c	---	---	---	---	5.9 ^c	---	---	---	---
9	3	---	5.0 ^c	4.8 ^c	3.7 ^a	---	---	4.2 ^b	7.5 ^c	6.5 ^c	---	---	13.0 ^c	5.7 ^b	4.4 ^b	---	---	---
10	4	---	---	---	---	3.9 ^a	7.6 ^c	6.4 ^c	---	5.0 ^c	5.8 ^c	---	---	9.4 ^a	---	6.3 ^c	---	7.2 ^c
11	4	3.8 ^b	---	---	6.7 ^c	15.0 ^c	6.9 ^c	---	---	3.9 ^a	5.3 ^c	---	---	6.0 ^c	---	7.8 ^c	---	---
12	4	---	---	---	---	4.0 ^b	5.5 ^c	---	---	5.1 ^c	---	3.8 ^b	5.3 ^c	---	---	5.5 ^c	3.8 ^b	---
13	5	---	---	---	---	---	---	5.3 ^a	---	3.6 ^b	3.7 ^b	3.9 ^a	5.6 ^c	5.6 ^c	---	---	6.3 ^c	5.3 ^c
14	5	---	---	5.1 ^b	---	3.7 ^b	5.0 ^c	5.5 ^c	---	6.9 ^c	6.6 ^c	---	5.4 ^c	---	---	5.9 ^c	---	---
15	5	4.5 ^a	---	---	5.4 ^a	---	5.4 ^c	---	---	6.0 ^c	5.7 ^c	7.1 ^c	---	---	3.6 ^b	7.4 ^c	---	4.1 ^b
16	6	---	---	---	4.6 ^a	---	---	---	---	---	---	---	---	5.2 ^a	---	6.3 ^c	---	---
17	6	---	6.0 ^c	4.2 ^b	3.9 ^b	---	---	---	---	---	---	---	---	---	---	6.0 ^c	---	3.7 ^b
18	6	---	---	---	---	---	3.5 ^a	---	---	---	---	---	---	---	---	6.8 ^c	---	---
19	7	---	4.2 ^a	4.0 ^a	---	---	---	---	5.6 ^c	---	---	---	---	4.4 ^a	---	---	6.0 ^c	---
20	7	---	5.1 ^b	5.4 ^c	4.9 ^c	---	---	---	---	---	---	---	---	---	---	---	---	---
21	7	---	---	6.3 ^c	---	---	---	---	---	---	---	3.6 ^a	---	---	---	---	---	---
M26 ^a	1	1	1	3	1	2	2	0	1	0	3	1	3	0	0	0	0	0
M35 ^b	2	1	2	1	2	0	1	0	1	2	1	0	1	2	1	1	1	2
CP ^c	1	4	4	4	2	7	5	4	6	6	4	6	4	2	9	2	2	2
Total	4	6	7	8	5	9	8	4	8	8	8	8	7	8	4	10	3	4

^a M26 two-way pseudo-testercross (TWPT) QTL

^b M35 TWPT QTL

^c Cross-pollinators (CP) QTL

Table 6 Estimates of maximum QTL effects (α, γ, τ) based on six QTL analyses of the average trait values within or among five environments (Tables 4 and 5), total magnitudes of maximum (Max) QTL effects, and total magnitudes of average (Avg) QTL effects over five environments reported as a percentage of population mean where α is difference between QTL alleles of the M26 parent, γ is difference between QTL alleles of the M35 parent, and τ is intralocus interaction

LG	HG	SEMA	SEAR	SELE	SEWI	SESH	SENA	SPIN	FLSP	SEFL	SESP	SYIN	INLE	STLE	STWI	ZAMA	CRCI	INCR
1	1	0 ^a ,5,0 ^a	---	---	---	---	---	---	---	---	13,8,3	10,6,1	---	---	---	---	---	---
2	1	---	---	---	---	---	-21,-11,-7	---	---	---	---	---	-3,-1,0	---	---	---	---	---
3	1	---	---	---	---	---	---	3,0,1	---	---	---	-12,0 ^a ,0 ^a	---	---	---	0 ^a ,3,0 ^a	---	---
4	2	---	---	---	---	---	---	-1,2,-3	-2,2,-2	---	---	-5,-1,-11	---	---	-2,2,-2	---	---	---
5	2	---	---	---	---	---	-1,3,3	---	---	---	---	---	-3,-3,-1	-1,-2,1	---	---	---	---
6	2	-2,-2,-3	-4,0,-3	-2,-1,0	-1,1,-2	-7,-7,-14	-5,0 ^a ,0 ^a	-2,3,-3	---	-9,-9,-13	-10,-10,-15	-12,-12,-15	---	---	-2,-2,-1	-2,-1,-2	---	---
7	3	---	-3,0,0	---	---	---	---	3,0 ^a ,0 ^a	---	---	0 ^a , -10,0 ^a	---	3,0 ^a ,0 ^a	---	---	---	---	---
8	3	---	---	---	0,0,2	---	---	---	-1,5,-3	---	---	---	---	-2,0,2	---	---	---	---
9	3	---	-2,1,-1	-2,0,0	-1,0 ^a ,0 ^a	---	---	0 ^a ,3,0 ^a	-2,2,-2	9,-8,6	---	---	-3,5,0	0 ^a ,2,0 ^a	0 ^a ,3,0 ^a	---	---	---
10	4	---	---	---	---	6,0 ^a ,0 ^a	-10,10,3	0,4,-2	---	2,12,-5	5,14,-7	---	---	-3,0 ^a ,0 ^a	---	-1,1,-1	---	0,9,-1
11	4	0 ^a ,3,0 ^a	---	---	2,2,0	17,7,5	-16,-4,-2	---	---	-12,0 ^a ,0 ^a	13,7,2	---	---	2,0,2	---	1,0,-1	---	---
12	4	---	---	---	---	0 ^a , -16,0 ^a	2,4,3	---	---	-8,-8,-3	---	0 ^a , -22,0 ^a	1,-3,0	---	---	-1,0,-1	0 ^a ,6,0 ^a	---
13	5	---	---	---	---	---	---	3,0 ^a ,0 ^a	---	0 ^a ,19,0 ^a	0 ^a ,10,0 ^a	17,0 ^a ,0 ^a	3,1,-1	-2,0,-2	---	---	-2,-1,-4	-6,1,-5
14	5	---	---	0 ^a ,2,0 ^a	---	0 ^a ,3,0 ^a	-3,-14,3	1,3,0	---	-5,-15,-1	-6,-16,-4	---	1,4,0	---	---	0,-1,0	---	---
15	5	4,0 ^a ,0 ^a	---	---	2,0 ^a ,0 ^a	---	-5,1,1	---	---	-2,-14,7	-2,-15,8	-1,-14,9	---	---	0 ^a , -4,0 ^a	0,-1,0	---	0 ^a , -9, τ ^a
16	6	---	---	---	-1,0 ^a ,0 ^a	---	---	---	-3,1,1	---	---	---	---	2,0 ^a ,0 ^a	---	-2,2,0	---	---
17	6	---	2,-3,1	0 ^a , -2,0 ^a	0 ^a , -1,0 ^a	---	---	---	---	---	---	---	---	---	---	-1,2,1	---	0 ^a , -8, τ ^a
18	6	---	---	---	---	---	-4,0 ^a ,0 ^a	---	---	---	---	---	---	---	---	0,-1,0	---	---
19	7	---	2,0 ^a ,0 ^a	1,0 ^a ,0 ^a	---	---	---	---	---	---	---	---	---	3,0 ^a ,0 ^a	---	---	3,2,-3	---
20	7	---	0 ^a ,3,0 ^a	0,2,0	0,1,1	---	---	---	---	---	---	---	---	---	---	---	---	---
21	7	---	---	-1,2,-1	---	---	---	---	---	---	---	-17,0 ^a ,0 ^a	---	---	---	---	---	---
Max		7,11,3	13,6,6	7,8,1	8,5,5	29,33,19	67,47,22	14,16,9	8,10,7	46,86,37	48,90,39	73,55,37	17,17,3	15,5,8	3,10,3	7,12,6	5,10,8	6,27,6
(total)		(21)	(25)	(16)	(18)	(81)	(136)	(39)	(25)	(169)	(177)	(164)	(36)	(27)	(17)	(25)	(22)	(40)
Avg		5,8,3	8,5,2	6,4,1	5,5,3	30,26,12	28,12,11	14,12,6	7,11,6	33,50,31	32,48,38	32,27,24	14,14,4	8,5,6	3,7,3	5,6,3	5,8,8	2,15,7
(total)		(15)	(16)	(12)	(12)	(68)	(52)	(31)	(24)	(114)	(118)	(82)	(32)	(19)	(14)	(15)	(20)	(24)

^a Non-estimable α , γ , or τ effect for QTLs detected using two-way pseudo-testcross model

Table 7 Total percent variation explained (PVE) by M26 x M35 QTL markers using a two-way pseudo-testcross (TWPT) or a cross-pollinators (CP) across five location (KS or UT) x year (2013, 2014, and 2015) environments and overall averages across five environments (AVG) including AVGs by parent (M26 or M35)

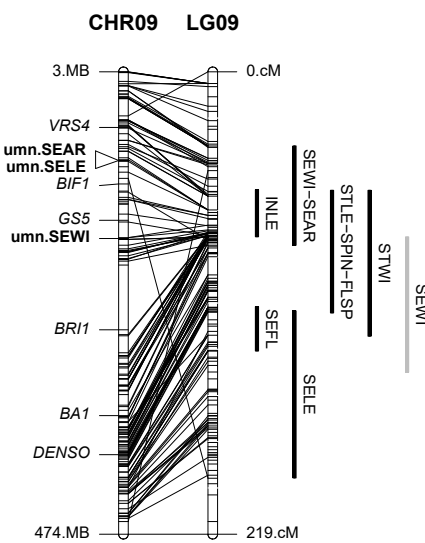
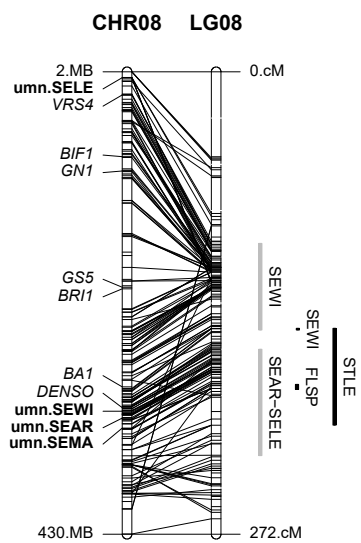
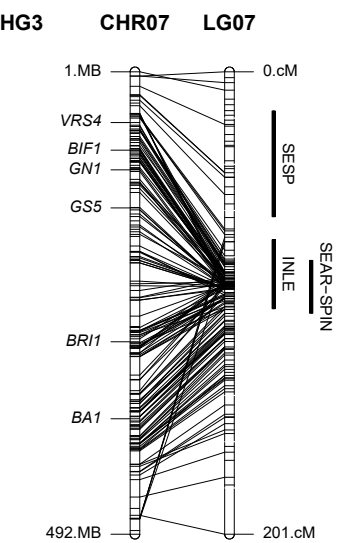
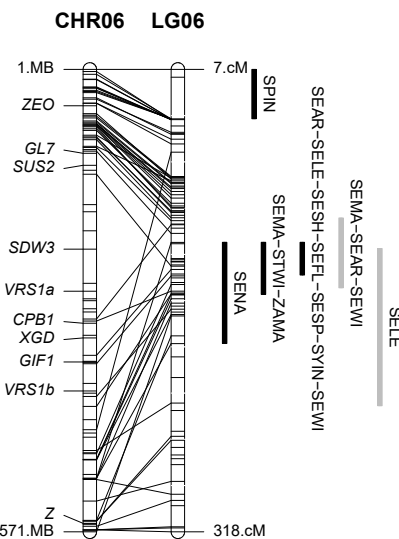
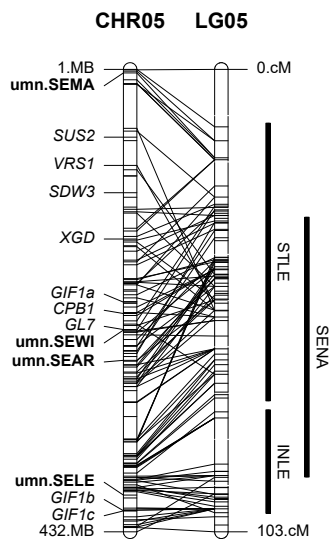
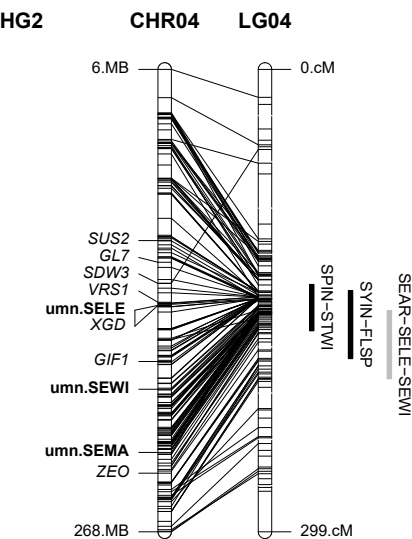
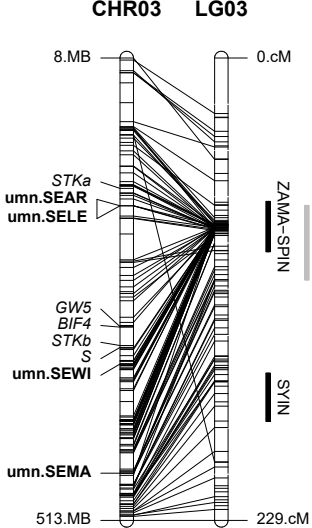
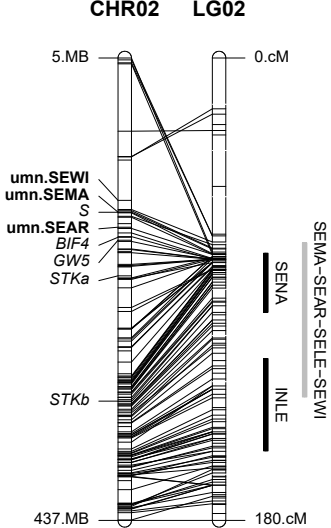
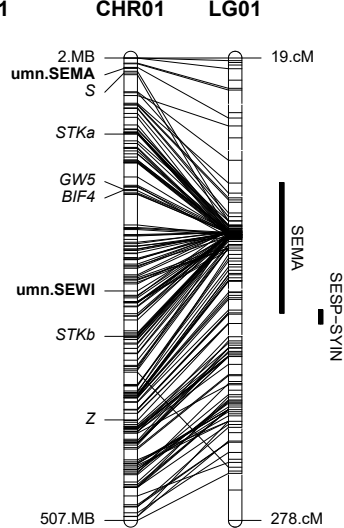
Trait	Model	Number markers	KS13 PVE	KS14 PVE	KS15 PVE	UT14 PVE	UT15 PVE	AVG PVE	AVG M26 PVE	AVG M35 PVE
SEMA	TWPT	5	14.3	16.2	8.0	11.0	11.3	15.1	4.8	11.5
	CP	4	16.6	19.6	12.7	18.1	12.9	20.8	---	---
SEAR	TWPT	10	18.1	21.4	22.7	24.2	17.9	31.5	21.1	15.6
	CP	6	28.1	22.8	18.8	23.9	18.4	29.3	---	---
SELE	TWPT	11	28.4	35.7	36.4	33.3	28.0	39.8	26.6	21.0
	CP	7	29.8	34.7	29.5	31.7	27.7	37.5	---	---
SEWI	TWPT	13	17.6	26.4	19.4	21.1	27.1	28.6	15.2	18.2
	CP	8	27.9	30.0	25.2	27.3	29.6	33.9	---	---
SESH	TWPT	7	---	28.74	16.4	26.4	16.1	29.6	16.8	15.8
	CP	5	---	35.4	26.5	25.7	16.2	29.5	---	---
SENA	TWPT	16	26.3	47.3	46.5	26.3	26.1	44.8	35.4	17.9
	CP	9	28.8	47.7	40.9	28.8	34.7	46.8	---	---
SPIN	TWPT	13	14.4	25.0	22.6	12.3	36.5	27.1	22.6	9.1
	CP	8	22.5	35.6	24.9	13.5	38.2	31.7	---	---
FLSP	TWPT	8	17.8	18.9	26.0	---	---	28.3	16.8	12.0
	CP	4	16.6	16.8	25.3	---	---	27.4	---	---
SEFL	TWPT	14	33.1	15.5	36.4	---	---	34.1	15.2	23.9
	CP	8	35.5	20.2	33.9	---	---	36.8	---	---
SESP	TWPT	15	28.2	17.4	30.1	25.7	32.5	29.7	10.6	20.8
	CP	8	32.7	22.7	29.0	23.6	31.0	30.1	---	---
SYIN	TWPT	12	27.7	31.3	24.3	20.5	22.7	26.9	14.3	13.9
	CP	8	32.7	39.1	26.9	21.9	21.3	27.8	---	---
INLE	TWPT	13	---	31.2	27.0	38.0	32.6	36.9	16.1	25.5
	CP	7	---	30.3	24.7	33.6	34.4	35.4	---	---
STLE	TWPT	12	17.7	25.4	28.6	23.5	37.2	23.2	14.5	12.5
	CP	8	26.3	27.3	34.9	30.5	28.9	38.0	---	---
STWI	TWPT	6	19.7	13.3	19.5	21.5	16.2	25.5	12.5	15.8
	CP	4	22.6	12.3	20.4	20.9	18.7	25.4	---	---
ZAMA	TWPT	19	27.6	41.4	40.4	42.74	44.5	45.2	19.5	30.4
	CP	10	34.2	49.2	45.3	44.0	46.7	49.6	---	---
CRCI	TWPT	5	---	9.6	---	8.7	8.9	8.4	1.9	6.8
	CP	3	---	14.7	---	14.2	13.6	19.2	---	---
INCR	TWPT	6	---	6.5	6.6	22.6	9.9	13.8	3.3	11.6
	CP	4	---	8.5	6.2	24.6	7.6	14.5	---	---

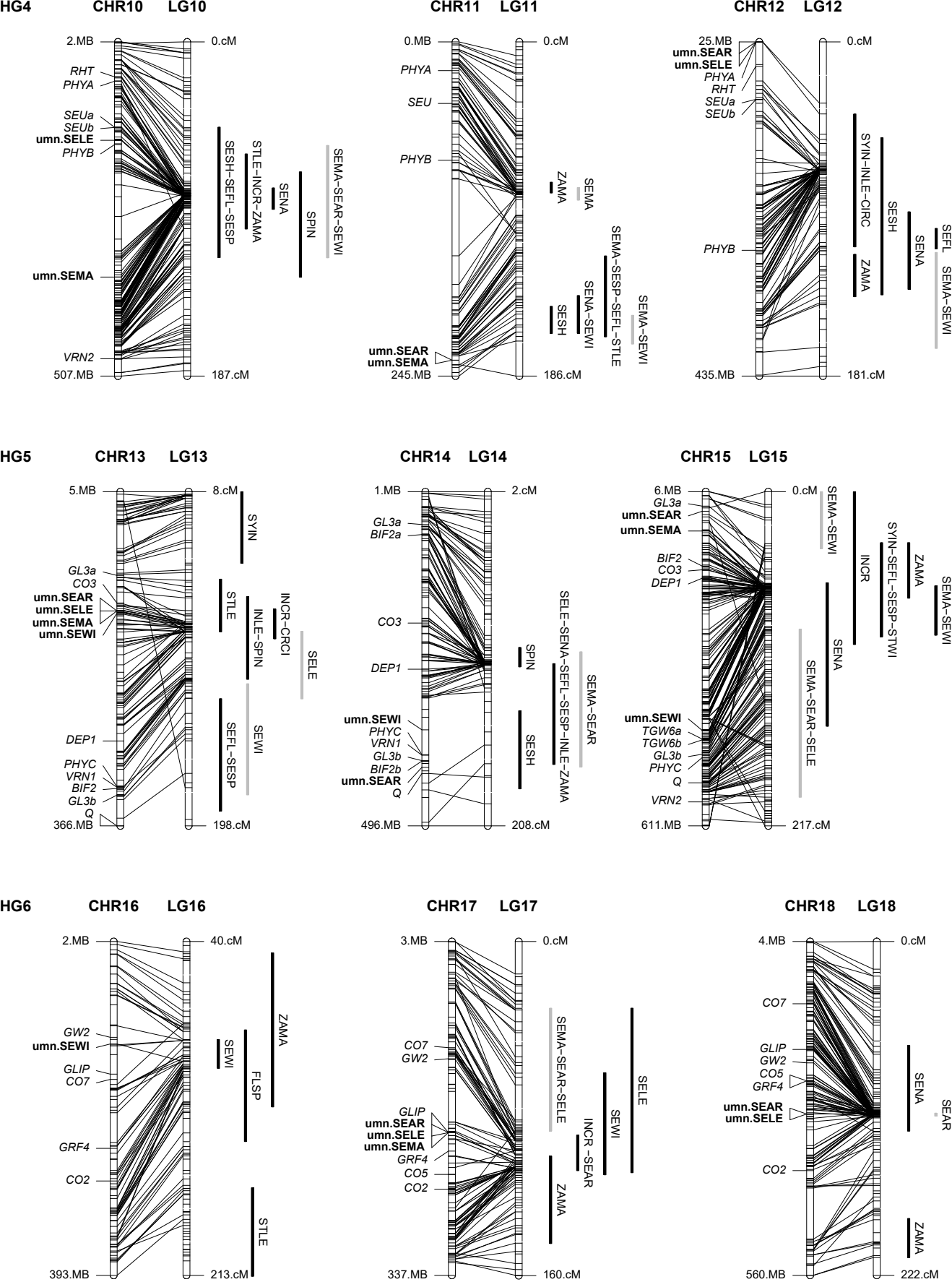
Table 8 Description of 42 candidate genes located in 87 different chromosome regions containing one or more relevant quantitative trait loci (QTL) identified in the intermediate wheatgrass M26 x M35 or C3_2332 x C3_2595 (UMN) families, listed by homoeologous group (HG)

HG	Candidate gene	Chromosome (QTL)
1	<i>Barren inflorescence 4 (BIF4)</i>	CHR03 (SPIN)
1	<i>Grain weight Chr5 QTL (GW5)</i>	CHR01 (SEMA), CHR02 (SEMA, SEAR, SELE, SEWI), CHR03 (SELE)
1	<i>Seedstick (STK)</i>	CHR02 ^a (SENA)
1	<i>Self-incompatibility (S)</i>	CHR01 (SEMA), CHR02 (SEMA, SEAR, SELE, SEWI)
2	<i>Clustered primary branch 1 (CPB1)</i>	CHR05 (STLE), CHR06 (SEMA, SEAR, SELE, SEWI, SYIN)
2	<i>Grain incomplete filling (GIF1)</i>	CHR04 (SEAR, SELE, SEWI, SYIN), CHR06 (SEMA, SEAR, SELE, SEWI, SYIN)
2	<i>Grain length Chr7 QTL (GL7)</i>	CHR04 (SEAR, SELE, SEWI, SYIN), CHR06 (SEMA, SEAR, SELE, SEWI, SYIN)
2	<i>Self-incompatibility (Z)</i>	CHR01 ^b (SESP, SYIN), CHR06 (SEMA, SEAR, SELE, SEWI, SEFL, SESP, SYIN)
2	<i>Semidwarf 3 (SDW3)</i>	CHR05 (STLE)
2	<i>Six-rowed spike 1 (VRS1)</i>	CHR04 (SPIN, FLSP), CHR06 ^a (SEFL, SESP)
2	<i>Sucrose synthase 2 (SUS2)</i>	CHR04 (SEAR, SELE, SEWI, SYIN), CHR06 (SEMA, SEAR, SELE, SEWI, SYIN)
2	<i>Xylogalacturonan deficient (XGD)</i>	CHR04 (SEAR, SELE, SEWI, SYIN), CHR06 (SEMA, SEAR, SELE, SEWI, SYIN)
3	<i>Barren inflorescence 1 (BIF1)</i>	CHR07 (SPIN), CHR09 (SPIN, FLSP)
3	<i>Barren stalk 1 (BA1)</i>	CHR07 (SPIN), CHR09 (SPIN, FLSP)
3	<i>Brassinosteroid-insensitive 1 (BR1)</i>	CHR07 (INLE), CHR09 (INLE, STLE)
3	<i>Grain number Chr1 QTL (GN1)</i>	CHR07 (SEAR, SPIN), CHR08 ^a (SEWI)
3	<i>Grain size Chr5 QTL (GS5)</i>	CHR07 (SEAR), CHR08 (SEWI), CHR09 (SEAR, SEWI)
3	<i>Semidwarf 1 (DENSO)</i>	CHR08 (STLE), CHR09 (STLE)
3	<i>Six-rowed spike 4 (VRS4)</i>	CHR07 (SPIN), CHR09 (SPIN)
4	<i>Phytochrome A (PHYA)</i>	CHR10 (ZAMA)
4	<i>Phytochrome B (PHYB)</i>	CHR10 (ZAMA), CHR11 (ZAMA)
4	<i>Reduced height (RHT)</i>	CHR10 (STLE)
4	<i>Seuss (SEU)</i>	CHR10 ^a (SESH, SENA), CHR12 ^a (SESH)
5	<i>Barren inflorescence 2 (BIF2)</i>	CHR14 (SPIN)
5	<i>Constans 3 (CO3)</i>	CHR14 (ZAMA), CHR15 (ZAMA)
5	<i>Dense and erect panicle 1 (DEP1)</i>	CHR13 (INLE, SPIN), CHR14 (INLE, SPIN)
5	<i>Grain length Chr3 QTL (GL3)</i>	CHR13 ^a (SEWI), CHR14 ^a (SEMA, SEAR, SELE), CHR15 ^a (SEMA, SEAR, SELE, SEWI)
5	<i>Phytochrome C (PHYC)</i>	CHR14 (ZAMA)
5	<i>Q free-threshing (Q)</i>	CHR14 (INLE, SESH, SENA), CHR15 (SENA)
5	<i>Vernalization 1 (VRN1)</i>	CHR14 (ZAMA)
6	<i>Constans 2 (CO2)</i>	CHR16 (ZAMA), CHR17 (ZAMA)
6	<i>Constans 5 (CO5)</i>	CHR17 (ZAMA)
6	<i>Constans 7 (CO7)</i>	CHR16 (ZAMA), CHR17 (ZAMA)
6	<i>GDSL lipase (GLIP)</i>	CHR16 (SEWI), CHR17 (SEMA, SEAR, SELE, SEWI), CHR18 (SEAR)
6	<i>Grain weight Chr2 QTL (GW2)</i>	CHR16 (SEWI), CHR17 (SEMA, SEAR, SELE, SEWI), CHR18 (SEAR)
6	<i>Growth-regulating factor 4 (GRF4)</i>	CHR16 (SEWI), CHR17 (SEMA, SEAR, SELE, SEWI), CHR18 (SEAR, SENA)
7	<i>Fertilization independent seed (FEI)</i>	CHR19 ^a (SEAR, SELE), CHR20 (SEMA, SEAR, SELE, SEWI), CHR21 (SELE)
7	<i>Glucan, water-dikinase (GWD)</i>	CHR19 (SEAR, SELE), CHR20 (SEAR, SELE), CHR21 (SELE)
7	<i>Grain width Chr8 QTL (GW8)</i>	CHR19 (SEAR, SELE), CHR20 (SEAR, SELE), CHR21 (SELE)
7	<i>Sucrose synthase 1 (SUS1)</i>	CHR19 ^a (SEAR, SELE), CHR20 ^a (SEMA, SEAR, SELE, SEWI), CHR21 (SELE)
7	<i>Thousand-grain weight 6 (TGW6)</i>	CHR15 ^{a,b} (SEMA, SEAR, SELE, SEWI), CHR19 (SEAR, SELE), CHR20 (SEAR, SELE, SEWI), CHR21 (SELE, SYIN)
7	<i>Thousand-grain weight Chr7A (TGW7A)</i>	CHR19 (SEAR, SELE), CHR20 (SEAR, SELE), CHR21 (SELE)

^a Multiple gene loci present on same chromosome

^b Paralogous gene locus

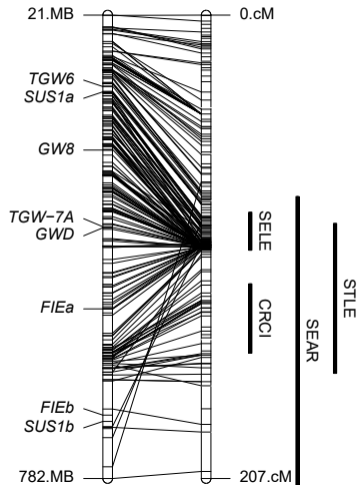




HG7

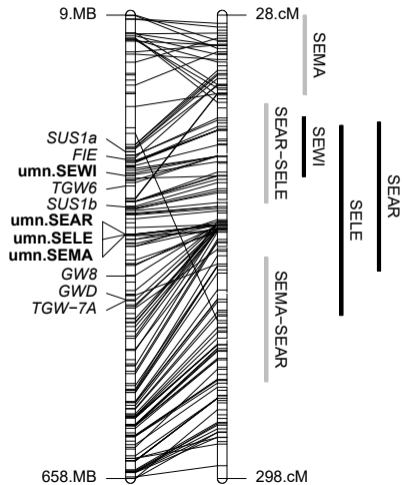
CHR19

LG19



CHR20

LG20



CHR21

LG21

