1    **Applied phenomics and genomics for improving barley yellow dwarf resistance in winter**

2    **wheat**

3

4    Paula Silva[1,2], Byron Evers[1], Alexandria Kieffaber[1], Xu Wang[3], Richard Brown[1], Liangliang

5    Gao[1], Allan Fritz[4], Jared Crain[1], Jesse Poland[1,5,§]

6

7    [1] Department of Plant Pathology, College of Agriculture, Kansas State University, Manhattan,

8    Kansas, 66506

9    [2] Programa Nacional de Cultivos de Secano, Instituto Nacional de Investigación Agropecuaria

10   (INIA), Estación Experimental La Estanzuela, Colonia, Uruguay, 70006

11   [3] Department of Agricultural and Biological Engineering, University of Florida, IFAS Gulf Coast

12   Research and Education Center, Wimauma, Florida, 33598

13   [4] Department of Agronomy, College of Agriculture, Kansas State University, Manhattan, Kansas,

14   66506

15   [5] King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

16   [§] corresponding author: jpoland@ksu.edu, jesse.poland@kaust.edu.sa

17

18   ORCID

19   Paula Silva:  0000-0003-2655-2949

20   Byron Evers:  0000-0003-1840-5842

21   Xu Wang:  0000-0002-7144-6865

22   Liangliang Gao: 0000-0002-8562-0027

23   Allan Fritz:  0000-0003-2574-8675

24   Jared Crain:  0000-0001-9484-8325

25   Jesse Poland:  0000-0002-7856-1399

26

27    Running title – Phenomics and genomics for BYD

28

29    Keywords: Barley yellow dwarf (BYD), High-throughput Phenotyping (HTP), *Triticum*

30    *aestivum*, Virus, Resistance, Tolerance, Genome-wide Association Mapping (GWAS), Genomic

31    Selection (GS)

32

**Abstract**

Barley yellow dwarf (BYD) is one of the major viral diseases of cereals. Phenotyping BYD in wheat is extremely challenging due to similarities to other biotic and abiotic stresses. Breeding for resistance is additionally challenging as the wheat primary germplasm pool lacks genetic resistance, with most of the few resistance genes named to date originating from a wild relative species. The objectives of this study were to, i) evaluate the use of high-throughput phenotyping (HTP) from unmanned aerial systems to improve BYD assessment and selection, ii) identify genomic regions associated with BYD resistance, and iii) evaluate genomic prediction models ability to predict BYD resistance. Up to 107 wheat lines were phenotyped during each of five field seasons under both insecticide treated and untreated plots. Across all seasons, BYD severity was lower with the insecticide treatment and plant height (PTHTM) and grain yield (GY) showed increased values relative to untreated entries. Only 9.2% of the lines were positive for the presence of the translocated segment carrying resistance gene *Bdv2* on chromosome 7DL. Despite the low frequency, this region was identified through association mapping. Furthermore, we mapped a potentially novel genomic region for resistance on chromosome 5AS. Given the variable heritability of the trait (0.211 – 0.806), we obtained relatively good predictive ability for BYD severity ranging between 0.06 – 0.26. Including *Bdv2* on the predictive model had a large effect for predicting BYD but almost no effect for PTHTM and GY. This study was the first attempt to characterize BYD using field-HTP and apply GS to predict the disease severity. These methods have the potential to improve BYD characterization and identifying new sources of resistance will be crucial for delivering BYD resistant germplasm.

3

**Introduction**

Wheat (*Triticum aestivum* L.) is one of the most essential food crops in the world and is constantly threatened by several biotic stresses (Savary *et al.* 2019). Among the most important viral stresses is barley yellow dwarf (BYD). This disease is widespread across the world, caused by viruses and transmitted by aphids (Shah *et al*. 2012), and can cause significant yield reductions in susceptible cultivars. In Kansas, BYD is the fourth most significant wheat disease in terms of average estimated yield losses with an average yield loss of approximately 1% estimated over the past 20 years (Hollandbeck *et al*. 2019), equivalent to a loss of more than $10 million per year. However, yield losses are highly variable ranging from 5% to 80% in a single field depending on the environment, management practices, the host, and the genetic background, (Miller and Rasochová 1997; Perry *et al*. 2000; Gaunce and Bockus 2015). Moreover, the wide host range and the complex lifestyle of its vectors make BYD extremely difficult to manage, and different management strategies (e.g., planting date and control of vector populations) are inconsistent depending on climate and location (Bockus *et al*. 2016). Thus, in many production environments, particularly in the Central and Eastern regions of Kansas, BYD is often the most economically impactful disease.

Barley yellow dwarf disease symptoms are highly variable depending on the crop, variety, time, and developmental stage when the infection occurs, aphid pressure, and environmental conditions (Shah *et al*. 2012; Choudhury *et al*. 2019b). BYD characterization in the field is extremely challenging as the symptoms can easily be confused with other viral disease symptoms such as wheat streak mosaic virus symptoms, nutrient deficiencies, or environmental stresses like waterlogging (Shah *et al*. 2012). Typical BYD symptoms can be observed at all levels of plant organization – leaf, roots, and flowers. Leaf discoloration in shades of yellow, red, or purple, specifically starting at the tip of the leaf and spreading from the margins toward the base is common as well as a reduction in chlorophyll content (Jensen and Van Sambeek 1972; D'arcy 1995). Often the entire plant visually appears stunted or dwarfed from a reduction in biomass by reducing tiller numbers. Spike grain yield is decreased through a reduction in kernels per spike and kernel weight which also affects grain quality (Riedell *et al*. 2003; Choudhury *et al*. 2019b). Quality can be further reduced by a reduction in starch content (Peiris

4

84    *et al.* 2019). Below ground effects of BYD have also been reported including reduced root

85    growth (Riedell *et al.* 2003).

86

87    Currently, there is no simple solution to control BYD (Walls *et al.* 2019), however, the use of

88    genetic resistance and tolerance is the most appealing and cost-effective option to control this

89    disease (Comeau and Haber 2002; Choudhury *et al.* 2017; 2019b). Resistance and tolerance

90    could be different genetic mechanisms, namely stopping virus replication and minimizing

91    disease symptoms respectively, but within this paper all mention of resistance includes both

92    genetic resistance and tolerance. Breeding strategies involving genetic resistance can target

93    either the aphids or the virus. Resistance to aphids can be achieved by three different strategies,

94    antixenosis, antibiosis, or tolerance (Girvin *et al.* 2017). To date, most breeding efforts have

95    been directed to the identification of viral tolerance, also known as 'field resistance', that refers

96    to the ability of the plant to yield under BYD infection and is associated with a reduction of

97    symptoms of infection independent of the virus titer (Foresman *et al.* 2016). Field resistance has

98    been reported to be polygenic, falling under the quantitative resistance class, where several genes

99    with very small effects control the resistance response (Qualset *et al.* 1973, Cisar *et al.* 1982;

100    Ayala *et al.* 2002; Choudhury *et al.* 2019a; c).

101

102    Presently, no major gene conferring immunity or a strong resistant phenotype to BYD has been

103    identified in bread wheat, and only four resistance genes have been described for BYD. Located

104    on chromosome 7DS, *Bdv1* is the only gene described from the primary pool of wheat and was

105    originally identified in the wheat cultivar 'Anza' (Qualset *et al.* 1984; Singh *et al.* 1993). This

106    gene provides resistance to some but not all the viruses that cause BYD (Ayala-Navarrete and

107    Larkin 2011). The other three named genes were all introduced into wheat through wide

108    crossing from intermediate wheatgrass (*Thinopyrum intermedium*) (Ayala *et al.* 2001; Zhang *et*

109    *al.* 2009). *Bdv2* and *Bdv3* are both located on a translocation segment on wheat chromosome

110    7DL (Brettell *et al.* 1988; Sharma *et al.* 1995), while *Bdv4* is located on a translocation segment

111    on chromosome 2D (Larkin *et al.* 1995; Lin *et al.* 2007). *Bdv2* was the first gene successfully

112    introgressed in wheat breeding programs from the tertiary gene pool for BYD resistance (Banks

113    *et al.* 1995) and deployed into varieties.

114

115    In addition to the four known resistance genes, other genomic regions associated with BYD

116    resistance have been identified through genetic mapping. These regions have been described on

117    nearly all wheat chromosomes but have not been genetically characterized (Ayala *et al*. 2002;

118    Jarošová *et al*. 2016; Choudhury *et al*. 2019a; b; c). Moreover, two recent studies have reported

119    that some of these new genomic regions display additive effects (Choudhury *et al*. 2019a; b).

120    Additive genetic effects had already been reported in lines combining *Bdv2* and *Bdv4* (Jahier *et*

121    *al*. 2009).

122

123    Taken together, research indicates that resistance genes to BYD in wheat are rare. With a lack of

124    major genes and difficulty to characterize resistance in the wheat pool likely due to the polygenic

125    nature of many small effect loci, identifying resistance has been limited. Nevertheless, breeding

126    programs have devoted large efforts for breeding BYD resistance due to the economic

127    importance of this disease, with some of the greatest success coming from wide crosses to the

128    tertiary gene pool.

129

130    Breeding for BYD resistance can be improved by applying strategies for more effective

131    evaluation and utilization of the identified resistance. To get a better understanding of BYD and

132    its quantitative nature, consistent and high-throughput methods are needed for the identification

133    of resistant wheat lines for large-scale selection in breeding programs (Aradottir and Crespo-

134    Herrera 2021). Effective selection on the quantitative resistance with low heritability can be

135    aided by the high-throughput genotyping, high-throughput phenotyping (HTP), or a combination

136    of both.

137

138    Access to high-density genetic markers at a very low-cost, owing to the rapid developments in

139    DNA sequencing, have enabled breeding programs to apply molecular breeding for quantitative

140    traits. Genomic selection (GS) is a powerful tool to breed for quantitative traits with complex

141    genetic architecture and low heritability (e.g., yield, quality, and diseases such as Fusarium head

142    blight), because it has greater power to capture loci with small effect compared with other

143    marker-assisted selection strategies (Meuwissen *et al.* 2001; Poland and Rutkoski 2016). In

144    addition to molecular data, HTP using unmanned aerial systems (UAS), or ground-based sensors

145    is providing high density phenotypic data that can be incorporated into breeding programs to

6

146  increase genetic gain (Haghighattalab *et al*. 2016; Crain *et al.* 2018; Wang *et al.* 2020). Using

147  precision phenotyping for disease scoring can improve the capacity for rapid and non-biased

148  evaluation of large field-scale numbers of entries (Poland and Nelson 2011). Taken together

149  improvements in genomics and phenomics have the potential to aid breeding progress for BYD

150  resistance.

151

152  In an effort to accelerate the development of resistant lines, we combined high throughput

153  genotyping and phenotyping to assess BYD severity in a large panel of elite wheat lines. We

154  evaluated the potential of HTP data to accurately assess BYD severity as well as identify genetic

155  regions associated with BYD resistance and inform whole genome prediction to identify resistant

156  lines.

157

158  **Materials and Methods**

159

160  **Plant Material**

161  A total of 381 different wheat genotypes were characterized for BYD resistance, including 30

162  wheat cultivars and 351 advanced breeding lines in field nurseries over five years (Table S1). In

163  each nursery, an unbalanced set of 52 – 107 wheat entries were evaluated including both

164  cultivars and breeding lines (Table 1). The BYD susceptible cultivar 'Art' and BYD resistant

165  cultivar 'Everest' were included in all the nurseries (seasons) as checks.

166

167  **Field Experiments**

168  Nurseries for BYD field-screening were conducted during five consecutive wheat seasons (2015

169  – 2016 to 2019 – 2020) (Table 1). Seasons 2015 – 16 and 2016 – 17 were conducted at Kansas

170  State University (KSU) Rocky Ford experimental station (39°13'45.60" N, 96°34'41.21" W),

171  while the 2017 – 18, 2018 – 19, and 2019 – 20 nurseries were planted at KSU Ashland Bottoms

172  experimental station (39°07'53.76" N, 96°37'05.20" W). The nurseries were established for

173  natural infections by planting about three weeks earlier than the normal planting window in mid-

174  September. The susceptible cultivar 'Art' was planted as a spreader plot in the borders and as a

175  control check plot also with the resistant cultivar 'Everest'. The experimental unit was 1.5m ×

176  2.4m with a six-row plot on 20cm row spacing.

7

177

178    A split-plot field design with two or three replications was used where the main plot was

179    insecticide treatment, and the split plot was the wheat genotype. Three replications were used

180    for proof of concept during the first two seasons but then two replications were chosen as a

181    balance of space and number of entries for the following seasons. For the treated replications the

182    seed were treated at planting with Gaucho XT (combination of insecticide and fungicide) at a

183    rate of 0.22 ml/100g of seed, followed with foliar insecticide applications starting from

184    approximately 2 – 3 weeks after planting through heading. Depending on field conditions, spray

185    treatments were conducted every 14 – 21 days if average air temperatures remained above 10°C.

186    Foliar insecticides were applied to the treated replications in a spray volume of 280.5 L/ha using

187    a Bowman MudMaster plot sprayer equipped with TeeJet Turbo TwinJet tips. Insecticide

188    applications consisted of a rotation of Warrior II, Lorsban, and Mustang Max at rates of

189    0.14L/ha, 1.17L/ha, and 0.29L/ha, respectively. For the control insecticide treatment (untreated),

190    the seed were treated with Raxil MD (fungicide) at a rate of 0.28ml/100g of seed, and no foliar

191    insecticide applications were applied. Foliar fungicide Nexicor was applied to the whole

192    experiment at a rate of 0.73L/ha, at both planting and heading, to control all other diseases so the

193    main disease pressure was focused on BYD.

194

195    **Phenotypic Data**

196    Individual plots were assessed for i) BYD severity characterized as the typical visual symptoms

197    of yellowing or purpling on leaves using a 0 – 100% visual scale, determined directly after spike

198    emergence by recording the proportion of the plot exhibiting the symptoms (Table 1), ii) manual

199    plant height ($PTHT_M$, meters), and iii) grain yield (GY, tons/ha). Experimental plots were

200    harvested using a Kincaid 8XP plot combine (Kincaid Manufacturing., Haven, KS, USA). Grain

201    weight, grain moisture and test weight measurements for each plot was recorded using a Harvest

202    Master Classic GrainGage and Mirus harvest software (Juniper Systems, Logan, UT, USA).

203    Visual phenotypic assessment was recorded using the Field Book phenoapp (Rife and Poland

204    2014).

205

**High-Throughput Phenotyping**

To compliment the manually recorded phenotypic data, we applied HTP using a ground-based proximal sensing platform or an UAS (Table 2). Seasons 2015 – 16 and 2016 – 17 were characterized by the ground platform as described in Barker *et al.* (2016) and Wang *et al.* (2018). For the other three seasons, we used a quadcopter DJI Matrice 100 (DJI, Shenzhen, China) carrying a MicaSense RedEdge-M multispectral camera (MicaSense Inc., United States). The HTP data was collected on multiple dates throughout the growth cycle from stem elongation to ripening (GS 30 – 90; Zadoks *et al.* 1974) (Table 2). Flight plans were created using CSIRO mission planner application and missions were executed using the Litchi Mobile App (VC Technology Ltd., UK, https://uavmissionplanner.netlify.app/) for DJI Matrice100. The aerial image overlap rate between two geospatially adjacent images was set to 80% both sequentially and laterally to ensure optimal orthomosaic photo stitching quality. All UAS flights were set at 20m above ground level at 2m/s and conducted within two hours of solar noon. To improve the geospatial accuracy of orthomosaic images, white square tiles with a dimension of 0.30m × 0.30m were used as ground control points and uniformly distributed in the field experiment before image acquisition and surveyed to cm-level resolution using the Emlid REACH RS+ Real-Time Kinematic Global Navigation Satellite System unit (Emlid Ltd., HongKong, China).

An automated image processing pipeline (Wang *et al.* 2020) was used to generate the orthomosaics and extract plot-level plant height (PTHT$_D$ (m), Singh *et al.* 2019) and the normalized difference vegetation index (NDVI) (Rouse *et al.* 1974), calculated as:

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \qquad\qquad\qquad \text{[Eq. 1]}$$

where NIR and Red are the near-infrared and red bands of the multispectral images and NDVI is the output image. Both traits were selected based on potential BYD characterization where the most typical BYD symptoms include chlorosis and stunting of the plants, thus, influencing NDVI and PTHT.

9

234

235    **Statistical Data Analyses**

236    First, the adjusted mean best linear unbiased estimator (BLUE) was calculated for each entry for

237    all the different traits for each season (Table S1), using the following model:

238

239    $$y_{ijklm} = \mu + G_i + T_j + GT_{ij} + R_{k(j)} + B_{l(kj)} + C_{m(kj)} + e_{ijklm}$$    [Eq. 2]

240

241    where $y_{ijklm}$ is the phenotype for the trait of interest, $\mu$ is the overall mean, $G_i$ is the fixed effect

242    of the $i^{th}$ entry (genotype), $T_j$ is the fixed effect of the $j^{th}$ insecticide treatment, $GT_{ij}$ is the fixed

243    effect of the interaction between the $i^{th}$ entry and the $j^{th}$ insecticide treatment (genotype by

244    treatment effect), $R_{k(j)}$ is the random effect of the $k^{th}$ replication nested within the $j^{th}$

245    insecticide treatment and distributed as iid $R_{k(j)} \sim N(0, \sigma_R^2)$, $B_{l(kj)}$ is the random effect of the

246    $l^{th}$ row nested within the $k^{th}$ replication and $j^{th}$ treatment distributed as iid $B_{l(kj)} \sim N(0, \sigma_B^2)$,

247    $C_{m(kj)}$ is the random effect of the $m^{th}$ column nested within the $k^{th}$ replication and $j^{th}$ treatment

248    and assumed distributed as iid $C_{m(kj)} \sim N(0, \sigma_C^2)$, and $e_{ijklm}$ is the residual for the $ijklm^{th}$ plot

249    and distributed as iid $e_{ijklm} \sim N(0, \sigma_e^2)$. The 'lme4' R package (Bates $et\ al$. 2014) was used for

250    fitting the models.

251

252    The BLUEs were used to inspect trait distributions and to calculate Pearson correlations between

253    all traits. In addition, BLUE values were used to calculate the reduction in GY for each entry as

254    the difference of GY between the untreated and insecticide treated main plots. This variable

255    reflects the level of BYD resistance of each entry, and it was used to perform GWAS and GS

256    analyses.

257

258    For NDVI and PTHT$_D$, the plot-level observed values extracted for the different phenotypic dates

259    were fitted to a logistic non-linear regression model (Fox and Weisberg 2011) as,

260

261    $$y = \frac{\theta_1}{1 + e^{-(\theta_2 + \theta_3 x)}} + \varepsilon$$    [Eq. 3]

262

10

263    where is $y$ the phenotype for the trait of interest at the time-point $x$ measured as days after

264    January 1, $\theta_1$ is the maximum value (upper asymptote) represented by the final PTHT or

265    maximum achieved NDVI, $\theta_2$ is the inflection point that represents the greatest rate of change in

266    the growth curve, either senescence for NDVI or height of growth, $\theta_3$ is the lag phase or onset of

267    senescence or growth rate from time $x$ where x is the calendar day of the year since January 1,

268    and $\varepsilon$ is the residual error (Figure S1). The "nlme" R package was used for model fitting

269    (Pinheiro *et al*. 2015). The model parameters obtained for each trait ($\theta_{1NDVI}, \theta_{2NDVI}, \theta_{3NDVI},$

270    $\theta_{1PTHT_D}, \theta_{2PTHT_D},$ and $\theta_{3PTHT_D}$) were used in addition to the other phenotypic traits to calculate

271    BLUEs, distributions, correlations, and BLUPs.

272

273    Secondly, we used a mixed linear model to calculate the best linear unbiased predictors (BLUPs)

274    for each entry in each nursery (season) (Table S1), using the same model as described in

275    equation 2 but defining $G_i, T_j,$ and $GT_{ij}$ as random effects. BLUPs were used because of the

276    unbalanced nature of the data (not all lines were evaluated in all the seasons). The BLUPs

277    calculated for each season were then combined for GWAS and GS. Furthermore, we calculated

278    broad-sense heritability on a line-mean basis by splitting the data based on whole plot treatment

279    for insecticide treatments as:

280

281    $H^2 = \dfrac{\sigma_G^2}{\sigma_G^2 + \dfrac{\sigma_e^2}{r}}$                     [Eq. 4]

282

283    where $\sigma_G^2$ is the genotypic variance, $\sigma_e^2$ is the residual error variance, and $r$ is the number of

284    replications.

285

286    **Genotypic Data**

287    A total of 346 wheat entries were genotyped using genotyping-by-sequencing (GBS) (Poland *et*

288    *al*. 2012) and sequenced on an Illumina Hi Seq2000. Single nucleotide polymorphisms (SNPs)

289    were called using Tassel GBSv2 pipeline (Glaubitz *et al*. 2014) and anchored to the Chinese

290    Spring genome assembly v1.0 (Appels *et al*. 2018). SNP markers with minor allele frequency <

291    0.01, missing data > 85%, or heterozygosity > 15% were removed from the analysis. After

292    filtering, we retained 29,480 SNPs markers that were used to investigate the population structure

293    through principal component analysis (PCA), genome-wide association analysis (GWAS), and

294    GS.  In addition, GBS data was used to run a bioinformatics pipeline to predict the presence or

295    absence of the translocated segment on chromomere 7DL carrying the *Bdv2* gene for each entry

296    (Table S1).  The prediction was done based on a modified alien predict pipeline (Gao *et al*.

297    2021).  Briefly, alien or wheat specific tags were counted in the 7DL region and tabulated using

298    a training set of cultivars or lines that are known to be *Bdv2* positive and negative. A simple

299    classification was done based on alien to wheat tag counts ratios.

300

**Genome-Wide Association Analysis**

302    The GWAS analysis was performed with a mixed linear model implemented in the 'GAPIT' R

303    package (Lipka *et al*. 2012) that includes principal components to account for population

304    structure as fixed effects and the individuals to explains familial relatedness as random effects,

305

306    $y = X\beta + Zu_i + e$                                                   [Eq. 5]

307

308    where *y* is the vector of phenotypic BLUPs, $X$ and $Z$ are the incidence matrix of $\beta$ and $u_i$,

309    respectively, with $u_i$ assumed $\sim \mathrm{N}\,(0, 2K_i\sigma_i^2)$ where $K$ is the individual kinship matrix, and $e$ is

310    the vector of random residual effects with $\sim \mathrm{N}\,(0, \mathrm{I}\sigma_e^2)$, where $I$ is the identity matrix and $\sigma_e^2$ is

311    the unknown residual variance.  The false discovery rate correction with an experimental

312    significance level value of 0.01 was used to assess marker-trait associations. Manhattan plots

313    were generated with 'CMplot' package in R software (Yin 2020).  PCA using GBS-SNPs was

314    performed in R language. Eigenvalues and eigenvectors were computed with 'e' function using

315    'A.mat' function and the 'mean' imputation method of 'rrBLUP' package (Endelman, 2011).  To

316    declare a quantitative trait locus (QTL) we considered only the regions having several SNP

317    markers in linkage disequilibrium, clearly showing a peak.  We did not consider regions with a

318    single SNP above the significant threshold as a QTL.

319

**Genomic Selection**

321    Using data from the five seasons, GS models using the genomic best linear unbiased predictor

322    (G-BLUP) were developed to assess predictive ability.  A five-fold cross-validation method was

323    used to assess model accuracy where the data set was split into five sets based on season, with

324    four seasons forming the training set and the fifth season serving as prediction set. This process

325    was repeated until all seasons were predicted. Along with predicting all other seasons from each

326    season, a model was evaluated with a leave-two-out cross-validation strategy. This strategy was

327    used to get a better mix of years with and without disease incidence, where the training

328    population consisted of three seasons, and the remaining two seasons were predicted from the

329    combined training population. The GS model was fitted with the training population using

330    'rrBLUP' *kin.blup* function (Endelman 2011), the GS model equation was,

331

332    $y = Wg + \varepsilon$                                        [Eq. 6]

333

334    where $y$ is a vector of phenotypic BLUPs, $W$ is the design matrix of $g$, $g$ is the vector of

335    genotypic values $\sim$ N $(0, K\sigma_g^2)$ and $\varepsilon$ is the vector of residual errors (Endelman 2011).

336    Predictive ability was assessed using Pearson's correlation ($r$) between the predicted value (G-

337    BLUP) and the BLUP for the respective phenotype. In addition, for both GS strategies we also

338    tested the effect of adding the genotype of the *Bdv2* loci as a fixed effect cofactor, using the

339    model,

340

341    $y = \mu + X\beta + Wg + \varepsilon$                               [Eq. 7]

342

343    which combines parameters described in equation 6 and $X$ is the matrix ($n \ x \ 1$) of individual

344    observation for presence or absence of *Bdv2* and $\beta$ is the fixed effect for the *Bdv2* measurements.

345

346    **Results**

347

348    **Phenotypic Data**

349    We analyzed five years of BYD field-screening nurseries (seasons 2015-16 to 2019-20)

350    characterizing a total of 381 wheat lines. The disease pressure and the expression of BYD

351    associated symptoms varied each season, however, we were able to observe a significant effect

352    of the insecticide treatment in all seasons (Figure 1). Across all seasons, BYD symptoms were

353    lower on the insecticide treated plots and both PTHT$_M$ and GY increased compared to the non-

354    treated control. Season 2016-17 had the most conducive conditions for BYD screening, resulting

355    in high average severity and a larger difference between mean values for the treated vs untreated

356    plots for all the collected traits (Figure 1).  There was general consistency in order across all

357    seasons with the susceptible check 'Art' ranked among the highest in BYD severity (Figure S2).

358

359    Phenotypic correlations between the traits showed a negative correlation between BYD and GY

360    for all the seasons and a negative or no correlation between BYD and $PTHT_M$ (Figure S3).  The

361    same correlation trends were observed under insecticide treated and untreated plots.  Broad-sense

362    heritability was moderate to high for all the traits, ranging between 0.21 and 0.79 for the

363    insecticide treated plots and between 0.41 and 0.84 for the untreated plots.  Across all traits, the

364    untreated insecticide replications showed higher $H^2$ values, with season $2016 - 17$ showing the

365    highest values (Figure 2).

366

367    For the HTP data collected (Table 2), we obtained three different parameters ($\theta_1$, $\theta_2$, and $\theta_3$) for

368    both $PTHT_D$ and NDVI after fitting a logistic regression model using the data collected during

369    the experiments (2015-16 season data was not included due to lack of data quality) (Figure S1).

370    Correlations between these parameters and the phenotypic traits collected manually were

371    different for all the traits (Figure S3).  For the insecticide untreated plots, BYD resulted in a

372    negative correlation with $\theta_{2NDVI}$ and a positive correlation with $\theta_{3NDVI}$, in most of the field

373    seasons.  We did not find a clear correlation pattern between BYD and $PTHT_D$.  For $PTHT_M$ we

374    detected a positive correlation with $\theta_{1PTHT_D}$ across all seasons, and for GY we observed a

375    positive correlation with  $\theta_{1NDVI}$ and $\theta_{2NDVI}$, and a negative correlation with $\theta_{3NDVI}$ (Figure S3).

376

**Prediction of *Bdv2* Resistance Gene**

378    We used GBS data to genotype the *Bdv2* resistance gene located on a translocation segment from

379    intermediate wheatgrass on chromosome 7DL of bread wheat.  In total, 33 of the 346 wheat lines

380    carried the *Th. intermedium* chromosomal translocation with *Bdv2* (Table S1).  Interestingly, 28

381    of these *Bdv2* lines belonged to the same breeding cycle, entering the advanced yield nursery

382    stage of the KSU breeding program in the $2017 - 18$ season.  Furthermore, only 7 pedigrees are

383    represented within the 28 *Bdv2* entries, meaning that these lines are highly related.  The

384    remaining 5 *Bdv2* lines were distributed in $2015 - 16$ (n=3), $2018 - 19$ (n=1), and $2019 - 20$

385    (n=1), and none of the lines from the season $2016 - 17$ had the presence of *Bdv2* (Table S1).

14

386

**Population Structure**

388 We studied the population structure of 346 wheat lines using 29,480 GBS-derived SNP markers.

389 The PCA did not reveal a strong pattern of population structure (Figure 3). Moreover, the

390 variation explained by the first two principal components (4 and 3%, respectively) also supports

391 the hypothesis of minimal population structure within a single breeding program. We observed

392 that most of the wheat cultivars released by KSU breeding program were located outside the

393 cluster grouping all the breeding lines (Figure 3A). Lines with the presence of *Bdv2* clustered

394 together (Figure 3B), likely due to a related pedigree to the original source, and we did not

395 identify any evident pattern for BYD severity associated with the population structure (Figure

396 3C).

397

**Genome-Wide Association Analysis**

399 To investigate the genetic architecture of BYD we performed GWAS analyses for all collected

400 traits using the BLUP values for 346 lines and 29,480 SNP markers. The first two principal

401 components from PCA and the kinship matrix were included in the mixed model to account for

402 population structure and genetic relatedness. We found significant marker-trait associations for

403 BYD severity on chromosomes 5AS, 7AL, and 7DL (Figure 4A). The highest peak was

404 observed on the proximal end of chromosome 7DL, located at 571 Mbp – 637 Mbp. To test the

405 hypothesis that this association was explained by the resistance gene *Bdv2* (located on

406 chromosome 7DL), we investigated the haplotypes defined by the 16 SNP markers associated

407 with BYD severity and were able to identify two haplotypes that exactly matched the presence or

408 absence of *Bdv2* (Fig 4A). This same region was mapped using BYD severity and the presence

409 or absence of *Bdv2* as a fixed covariate (Figure 4B). This analysis (Figure 4B) also detected a

410 peak on chromosome 7AL. Lastly, we explored the effect of *Bdv2* on both BYD BLUEs and

411 BLUPs, and we observed that the presence of *Bdv2* had a positive effect in reducing the disease

412 severity by approximantly 10% (Figure 5A). The significant peak on chromosome 5AS, located

413 at 46 Mbp – 103 Mbp, was explained by 10 SNP markers, comprising two main haplotypes, one

414 of them associated with reduced BYD severity (Fig 5B). When we combined the different 5AS

415 haplotypes with *Bdv2*, we observed that the presence of *Bdv2* had a positive effect, reducing the

416 levels of BYD when combined with both 5AS haplotypes (Figure 5C), and suggesting an

417     additive effect.  Compared to the associations found for *Bdv2* (Fig 4B), we did not find any

418     strong evidence of marker trait associations for the other evaluated traits (Figure S4).

419

420     **Genomic Selection**

421     To evaluate the potential of GS to predict BYD disease severity, we fit several GS models to the

422     phenotypic BLUPs of BYD, $PTH_M$, and reduction in GY.  Across all traits, to determine

423     predictive ability we used a five-fold cross validation where prediction ability ranged from -0.08-

424     0.26.  There was relatively good predictive ability for BYD severity ranging between 0.06 –

425     0.26, in comparison with $PTHT_M$ and reduction in GY resulting in a lower range from 0.02 – 017

426     and -0.08 – 0.2, respectively (Figure 6).  Evaluating the conformation of the training population,

427     we observed that when including 2016-17 season, prediction abilities were the highest for BYD

428     but the lowest for the other two traits, implying that season 2016 – 17 was either a good season

429     to train the prediction models or a difficult season to predict based on available data.

430

431     To further investigate the power of GS, we developed models using a leave-two-out strategy,

432     where two seasons were excluded from the training population and used as the testing

433     population.  We fitted GS models for all possible two-season combinations.  This strategy

434     resulted in slightly smaller training populations which decreased overall predictive ability

435     (Figure 6).  This result was evident for BYD predictions were excluding two seasons had a larger

436     negative impact.

437

438     Lastly, we evaluated the effect of adding information about the genotype of the *Bdv2* resistance

439     gene as a phenotypic fixed covariate into the GS models.  There were differences in the effect of

440     *Bdv2* on the predictive ability across BYD severity, $PTHT_M$, and GY, showing a large effect for

441     predicting BYD but almost no effect for $PTHT_M$ and reduction in GY (Figure 6).  The improved

442     predictive ability for BYD was clearly reflected with the decrease of prediction ability obtained

443     when season 2017 – 18 was excluded from the training population since most of the lines with

444     the presence of *Bdv2* were evaluated in that season.

445

16

**Discussion**

**Phenotypic Data**

The success of breeding for BYD resistance is highly impacted by the ability to precisely characterize breeding material and disease symptoms. Even though BYD is spread worldwide, its incidence in a given year depends on several factors such as aphid pressure, planting date, and environmental conditions (e.g., temperature, rainfall, frost, etc.). In this study, we evaluated winter wheat advanced breeding lines during five seasons implementing a rigorous field-testing approach, that ultimately enabled us to consistently have plots contrasting with BYD infection and uninfected or low incident plots. Moreover, by using large yield-size plots we were able to calculate the reduction in GY and use this parameter as an estimate of field resistance.

The expression of BYD symptoms, however, was highly inconsistent during the different seasons. Seasons 2015-16 and 2016-17 showed the best expression of the disease symptoms, supported by the wide range of BYD severity between treated and untreated replications (Figure 1). Interestingly, both these seasons were conducted in the same experimental field (Table 1), suggesting that this location could favor the expression of BYD. Moreover, weather conditions were variable for all the seasons, suggesting that these had a huge impact on the disease occurrence. While temperature records were similar for all the seasons, precipitation records did show some differences. Season 2017 – 18 was dryer than normal, with 34% less precipitation than the 30 years historical average (1981 – 2010). On the other hand, season 2018-19 was wetter than normal, with 58% more precipitation than the 30 years historical average (Table S2).

**High-Throughput Phenotyping**

Evaluating BYD resistance using visual phenotypic selection can be challenging due to the complex nature of the disease and rater variability (Poland and Nelson 2011). The use of HTP with UAS is gaining popularity within breeding programs because it further improves selection based on classical phenotyping. Accurate phenotyping is crucial for understanding the genetic basis of quantitative and complex traits like BYD. In this study, we used HTP to complement the visual BYD scoring. This tool improved our capacity for rapid, non-destructive, and non-biased evaluation of large field-scale numbers of entries for BYD resistance. We were able to

477    determine strong correlation patterns between visual BYD severity and HTP derived parameters

478    (Figure S3).  However, none of the traits collected with UAS had a common genetic base with

479    BYD severity (Figure 4 and Figure S4).  Disease scoring using HTP is scaling fast among

480    breeding programs; however, how to effectively use this data remains challenging.  Some studies

481    have shown that data collected with sensor-based tools can be substituted to improve classical

482    disease visual evaluation (Sankaran *et al*. 2010; Kumar *et al*. 2016; Zheng *et al*. 2018); however,

483    to the best of our knowledge this study is the first attempt to characterize BYD in wheat using

484    HTP.

485

486    **Genome-Wide Association Analysis**

487    Using GWAS we detected QTLs on chromosomes 5AS, 7AL, and 7DL for BYD severity BLUPs

488    values.  Using GBS tags that mapped to known alien fragments, we confirmed *Bdv2* resistance

489    gene was located at 7DL, and confirmed that the 7DL QTL was explained by the presence of the

490    *Bdv2* resistance gene.  Even though only 33 wheat lines were positive for the presence of *Bdv2*,

491    we still had enough power to detect its effect, suggesting that *Bdv2* has a strong effect on BYD

492    under Kansas field conditions (Figure 5).  The associations on chromosome 7AL, observed for

493    both BYD severity and *Bdv2*, suggest that the SNP markers on the 7AL peak may be miss-

494    anchored markers that should have mapped to 7DL.  The relatively high heritability values

495    obtained for the untreated replications (Figure 2) allowed us to detect a minor QTL on 5AS.

496    Marza *et al*. (2005) reported a QTL at 38cM on the short arm of chromosome 5A associated with

497    yellowing symptoms caused by BYD, and it is possible that this is the same region yet more data

498    is needed to confirm if these QTLs are the same. The only other study reporting GWAS for BYD

499    in wheat was able to identify several markers associated with BYD resistance on chromosomes

500    2A, 2B, 6A, and 7A (Choudhury *et al*. 2019b).  However, most of the association were explained

501    by individual SNP markers, and to date do not have any definitive biological link.  GWAS

502    results for the other traits used in this study did not discover genomic regions associated with the

503    traits (Figure S4).  Taken together, these results suggest that BYD resistance is not controlled by

504    any large effect loci that could easily be incorporated into the breeding program, thus GS could

505    be an efficient way to enhance BYD resistance.

506

**Genomic Selection**

507
508     We evaluated several different GS models to identify the best approach for predicting BYD

509     (Figure 6). Overall, we observed some trends including i) incorporating years with consistent

510     BYD disease data in the training population increased the model predictive ability, ii) predicting

511     years with high disease pressure is difficult, iii) using major effect QTL, such as *Bdv2*, had

512     increased prediction performance, suggesting that it is responsible for much of the predictive

513     power. These results suggest that GS based on G-BLUP with *Bdv2* as fixed effects would lead to

514     the greatest genetic gain for BYD breeding. Using selected major QTL as a fixed effect to

515     improve GS models was suggested in a simulation study (Bernardo 2014) and demonstrated with

516     empirical studies (Rutkoski *et al*. 2014). Nonetheless, using *Bdv2* as a fixed effect in our GS

517     strategies did not consistently improve the predictive ability for $PTH_M$ or reduction in GY (Rice

518     and Lipka 2019). However, there was not a consistent distribution of *Bdv2* allele across the

519     cohorts. BYD predictions were low compared to other disease (reviewed by Poland and

520     Rutkoski 2016). However, since this is the first report of GS for BYD resistance in wheat, we do

521     not have similar results to make better comparisons. BYD has traditionally been reported to

522     have low $H^2$ (Tola and Kronstad 1984; Choudhury *et al*., 2019b) and in this study, even with

523     well managemed plots that often had $H^2$ approaching 0.8, we still had difficulty reproducing

524     these results year to year as evidence of the challenge of studying this pathosystem. Moreover,

525     the correlation between HTP parameters and BYD phenotypes was interesting, but not sufficient

526     to be useful in combination with GS in the germplasm tested.

527

**Conclusions**

528
529     We were able to show that *Bdv2* has a major effect controlling BYD resistance in the KSU

530     breeding germplasm. Apart from the known *Bdv2* and a potentially novel 5AS region, we did

531     not find evidence of other regions controlling BYD resistance supporting the hypothesis of

532     limited resistance available in the current wheat gene pool and the highly polygenic nature of the

533     trait. Moreover, our study was the first attempt to characterize and improve BYD field-

534     phenotyping using HTP and apply GS to predict the disease. HTP traits showed strong

535     correlation patterns with BYD severity, however, none of these parameters shared a common

536     genetic architecture with BYD severity. The GS predictive ability results that we found in this

537     study open the door for further improvement and testing GS implementation for breeding for

538  BYD resistance. Continuing the improvement of BYD characterization and the search of new

539  sources of resistance using species related to wheat, will be crucial to broadening the resistant

540  genes available to introgress into wheat germplasm.

541

542

543  **Data Availability Statement**

544  Supplemental material, including raw and analyzed phenotypic data, genotypic data,

545  supplementary tables and figures, and basic plot scripts are available at Dyrad

546  doi:10.5061/dryad.ncjsxkswd (temporary link: https://datadryad.org/stash/share/xkKdr62QYB-

547  YA93mkzq18_4yFUgwdnvv0uZUyuEHpAI) and GitHub

548  https://github.com/umngao/wsm1_bdv2

549  **Supplementary data description**

550  **Table S1 –** List of wheat entries phenotypically evaluated in the study. The table includes the

551  type of entry (cultivar or breeding line), the season that the entry was evaluated, the result for the

552  prediction of the presence/absence of the segment carrying the resistance gene *Bdv2*, and the best

553  linear unbiased predictors (BLUPs) for all the phenotypic traits collected.

554  **Table S2 –** Precipitation (inches) during the five field seasons in Riley County, KS, where

555  Rocky Ford and Ashland Bottoms experimental units are located. Normal temperature is defined

556  as a 30–year average from 1981 – 2010. Data was obtained from Kansas State University

557  (http://climate.k-state.edu/precip/county/)

558  **Figure S1 –** Growth trajectories and adjustment of the non-linear regression model of wheat

559  lines for A-B) normalized difference vegetation index (NDVI) and C-D) digital plant height

560  (meters). The data used correspond to season 2016 – 17 phenotypic data. Calendar days is the

561  number of days starting at January 1, 2017.

562  **Figure S2 –** Boxplots showing the phenotypic response of the wheat checks 'Art' (susceptible)

563  and 'Everest' (tolerant) for A) barley yellow dwarf (BYD) disease severity (%), B) manual plant

564  height ($PTHT_M$) (m) and C) grain yield (GY) (tons/ha). Adjusted phenotypic values are shown

565  for both insecticide treatment replications (treated and untreated).

566  **Figure S3 –** Scatterplots showing distribution and Pearson's correlation values for the

567  phenotypic traits studied during all the field seasons under two insecticide treatments (treated

20

568    and untreated). A-B) season 2016 – 17, C-D) season 2017 – 18, E-F) season 2018 – 19, and G-H)

569    season 2019 – 20.

570    **Figure S4 –** Manhattan plots showing genome-wide association analysis (GWAS) results for the

571    phenotypic traits collected during the study.

572

573    **Acknowledgements**

579

580

**Literature Cited**

Appels, R., K. Eversole, C. Feuillet, B. Keller, J, Rogers, *et al*., 2018 Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science, 361(6403) http://dx.doi. org/10.1126/ science.aar7191

Aradottir, G. I., and L. Crespo-Herrera, 2021 Host plant resistance in wheat to barley yellow dwarf viruses and their aphid vectors: a review. Current Opinion in Insect Science, 2021, 45:59–68 https://doi.org/10.1016/j.cois.2021.01.002

Ayala-Navarrete, L., and P. Larkin, 2011 Wheat virus diseases: breeding for resistance and tolerance. World wheat book: a history of wheat breeding, Lavoisier, Paris, 2, 1073-107.

Ayala, L., M. Henry, M. Van Ginkel, R. Singh, B. Keller, *et al*., 2002 Identification of QTLs for BYDV tolerance in bread wheat. Euphytica, 128(2), 249-259 https://doi.org/10.1023/A:1020883414410

Ayala, L., M. Van Ginkel, M. Khairallah, B. Keller, and M. Henry, 2001 Expression of *Thinopyrum intermedium*-derived Barley yellow dwarf virus resistance in elite bread wheat backgrounds. Phytopathology, 91(1), 55–62 https://doi.org/10.1094/PHYTO.2001.91.1.55

Banks, P. M., J. L. Davidson, H. Bariana, and P. J. Larkin, 1995 Effects of barley yellow dwarf virus on the yield of winter wheat. Australian Journal of Agricultural Research, 46(5), 935-946

Barker III, J., N. Zhang, J. Sharon, R. Steeves, X. Wang, *et al*., 2016 Development of a field-based high-throughput mobile phenotyping platform. Computers and Electronics in Agriculture, 122, 74-85 https://doi.org/10.1016/j.compag.2016.01.017

Bates, D., M. Mächler, B. Bolker, and S. Walker, 2014 Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823

Bernardo, R., 2014 Genome-wide selection when major genes are known. Crop Science, 54(1), 68-75 https://doi.org/10.2135/cropsci2013.05.0315

Bockus, W. W., E. D. De Wolf, and T. Todd, 2016 Management strategies for barley yellow dwarf on winter wheat in Kansas. Plant Health Progress, 17(2), 122–127 https://doi.org/10.1094/PHP-RS-15-0050

610  Brettell, R. I. S., P. M. Banks, Y. Cauderon, X. Chen, z. Cheng, *et al.*, 1988 A single wheatgrass
611      chromosome reduces the concentration of barley yellow dwarf virus in wheat. Annals of
612      Applied Biology, 113(3), 599–603 https://doi.org/10.1111/j.1744-7348.1988.tb03337.x
613  Choudhury, S., H. Hu, Y. Fan, P. Larkin, M. Hayden, *et al.*, 2019a Identification of new QTL
614      contributing to barley yellow dwarf virus-PAV (BYDV-PAV) resistance in wheat. Plant
615      Disease, 103(11), 2798–2803 https://doi.org/10.1094/PDIS-02-19-0271-RE
616  Choudhury, S., P. Larkin, H. Meinke, M. D. Hasanuzzaman, P. Johnson, and M. Zhou, 2019b
617      Barley yellow dwarf virus infection affects physiology, morphology, grain yield and flour
618      pasting properties of wheat. Crop and Pasture Science, 70(1), 16-25
619      https://doi.org/10.1071/CP1836
620  Choudhury, S., P. Larkin, R. Xu, M. Hayden, K. Forrest, *et al.*, 2019c Genome wide association
621      study reveals novel QTL for barley yellow dwarf virus resistance in wheat. BMC
622      genomics, 20(1), 1-8 https://doi.org/10.1186/s12864-019-6249-1
623  Choudhury, S., H. Hu, H. Meinke, S. Shabala, G. Westmore, *et al.*, 2017 Barley yellow dwarf
624      viruses: infection mechanisms and breeding strategies. Euphytica, Vol. 213, pp. 1–22
625      https://doi.org/10.1007/s10681-017-1955-8
626  Cisar, G., C. M. Brown, and H. Jedlinski, 1982 Diallel analyses for tolerance in winter wheat to
627      the barley yellow dwarf virus 1. Crop Science, 22(2), 328-333
628  Comeau, A., and S. Haber, 2002 Breeding for BYDV tolerance in wheat as a basis for a multiple
629      stress tolerance strategy. Barley Yellow Dwarf Disease: Recent Advances and Future
630      Strategies, 82
631  Crain, J., S. Mondal, J. Rutkoski, R. P. Singh, and J. Poland, 2018 Combining High-Throughput
632      Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in
633      Wheat Breeding. The Plant Genome, 11(1), 170043
634      https://doi.org/10.3835/plantgenome2017.05.0043
635  D'arcy, C. J., 1995 Symptomatology and host range of barley yellow dwarf. Barley yellow
636      dwarf, 40, 9-28
637  Endelman, J. B., 2011 Ridge Regression and Other Kernels for Genomic Selection with R
638      Package rrBLUP. The Plant Genome, 4(3), 250–255.
639      https://doi.org/10.3835/plantgenome2011.08.0024

640  Foresman, B. J., R. E. Oliver, E. W. Jackson, S. Chao, M. P. Arruda, *et al*., 2016 Genome-wide
641      association mapping of barley yellow dwarf virus tolerance in spring oat (*Avena sativa*
642      L.). PloS one, 11(5), e0155376 https://doi.org/10.1371/journal.pone.0155376

643  Fox, J., and S. Weisberg, 2011 Nonlinear regression and nonlinear least squares in R. In: J. Fox
644      and S. Weisberg, An R companion to applied regression. 2nd ed. SAGE Publications, Los
645      Angeles, CA

646  Gao, L., D. H. Koo, P. Juliana, T. Rife, D. Singh, *et al*., 2021 The *Aegilops ventricosa* 2N$^v$ S
647      segment in bread wheat: cytology, genomics, and breeding. Theoretical and Applied
648      Genetics, 134(2), 529-542. https://doi.org/10.1007/s00122-020-03712-y

649  Gaunce, G. M., and W. W. Bockus, 2015 Estimating Yield Losses Due to Barley Yellow Dwarf
650      on Winter Wheat in Kansas Using Disease Phenotypic Data. Plant Health Progress,
651      16(1), 1–6 https://doi.org/10.1094/php-rs-14-0039

652  Girvin, J., R. J. Whitworth, L. M. A. Rojas, and C. M. Smith, 2017 Resistance of select winter
653      wheat (*Triticum aestivum*) cultivars to *Rhopalosiphum padi* (Hemiptera: *Aphididae*).
654      Journal of Economic Entomology, 110(4), 1886–1889 https://doi.org/10.1093/jee/tox164

655  Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. Elshire, *et al*., 2014 TASSEL-GBS: A
656      High-Capacity Genotyping by Sequencing Analysis Pipeline. PLoS ONE, 9(2), e90346
657      https://doi.org/10.1371/journal.pone.0090346

658  Haghighattalab, A., L. G. Pérez, S. Mondal, D. Singh, D. Schinstock, *et al*., 2016 Application of
659      unmanned aerial systems for high throughput phenotyping of large wheat breeding
660      nurseries. Plant Methods, 12(1), 1–15 https://doi.org/10.1186/s13007-016-0134-6

661  Hollandbeck, G. F., E. DeWolf, and T. Todd, 2019 Kansas cooperative plant disease survey
662      report. Preliminary 2019 Kansas wheat disease loss estimates. October 2, 2019
663      https://agriculture.ks.gov/divisionsprograms/plant-protect-weed-control/reports-and-
664      publications

665  Jahier, J., F. Chain, D. Barloy, A. M. Tanguy, J. Lemoine J, *et al*., 2009 Effect of combining two
666      genes for partial resistance to Barley yellow dwarf virus-PAV (BYDV- PAV) derived
667      from *Thinopyrum intermedium* in wheat. Plant Pathol 2009, 58:807-814
668      https://doi.org/10.1111/j.1365-3059.2009.02084.x

669   Jarošová, J., E. Beoni, and J. K. Kundu, 2016 Barley yellow dwarf virus resistance in cereals:

670       Approaches, strategies, and prospects. Field Crops Research, 198, 200–214

671       https://doi.org/10.1016/j.fcr.2016.08.030

672   Jensen, S. G., and J. W. Van Sambeek, 1972 Differential Effects of Barley Yellow Dwarf Virus

673       on the Physiology of Tissues of Hard Red Spring Wheat. Phytopathology, 62(2), 290

674       https://doi.org/10.1094/phyto-62-290

675   Kumar, S., M. S. Röder, R. P. Singh, S. Kumar, R. Chand, *et al*., 2016 Mapping of spot blotch

676       disease resistance using NDVI as a substitute to visual observation in wheat (*Triticum*

677       *aestivum* L.). Molecular Breeding, 36(7), 1–11 https://doi.org/10.1007/s11032-016-0515-

678       6

679   Larkin, P. J., P. M. Banks, E. S. Lagudah, R. Appels, C. Xiao, *et al*., 1995 Disomic *Thinopyrum*

680       *intermedium* addition lines in wheat with barley yellow dwarf virus resistance and with

681       rust resistances. Genome, 38(2), 385–394 https://doi.org/10.1139/g95-050

682   Lin, Z. S., Z. F. Cui, X. Y. Zeng, Y. Z. Ma, Z. Y. Zhang, *et al*., 2007 Analysis of wheat-

683       *Thinopyrum intermedium* derivatives with BYDV-resistance: Analysis of wheat-

684       *Th.intermedium* derivatives. Euphytica, 158(1–2), 109–118

685       https://doi.org/10.1007/s10681-007-9435-1

686   Lipka, A. E., F. Tian, Q. Wang, J. Peiffer, M. Li, *et al*., 2012 GAPIT: genome association and

687       prediction integrated tool. Bioinformatics, 28(18), 2397–2399

688       https://doi.org/10.1093/bioinformatics/bts444

689   Marza, F., G. H. Bai, B. F. Carver, and W. C. Zhou, 2005 Quantitative trait loci for yield and

690       related traits in the wheat population Ning7840× Clark. Theoretical and Applied

691       Genetics, 112(4), 688-698 https://doi.org/10.1007/s00122-005-0172-3

692   Meuwissen, T. H., B. J. Hayes, and M. E. Goddard, 2001 Prediction of total genetic value using

693       genome-wide dense marker maps. Genetics, 157(4), 1819-1829

694       https://doi.org/10.1093/genetics/157.4.1819

695   Miller, W. A., and L. Rasochová, 1997 Barley yellow dwarf viruses. Annual Review of

696       Phytopathology, Vol. 35, pp. 167–190 https://doi.org/10.1146/annurev.phyto.35.1.167

697   Mondal, S., J. E. Rutkoski, G. Velu, P. K. Singh, L. A. Crespo-Herrera, *et al*., 2016 Harnessing

698       diversity in wheat to enhance grain yield, climate resilience, disease and insect pest

699  resistance and nutrition through conventional and modern breeding approaches. Frontiers

700  in plant science, 7, 991 https://doi.org/10.3389/fpls.2016.00991

701  Perry, K. L., F. L. Kolb, B. Sammons, C. Lawson, G. Cisar, *et al.*, 2000 Yield effects of Barley

702  yellow dwarf virus in soft red winter wheat. Phytopathology, 90(9), 1043–1048

703  https://doi.org/10.1094/PHYTO.2000.90.9.1043

704  Peiris, K. H., R. L. Bowden, T. C. Todd, W. W. Bockus, M. A. Davis, *et al.*, 2019 Effects of

705  barley yellow dwarf disease on wheat grain quality traits. Cereal Chemistry, 96(4), 754-

706  764 https://doi.org/10.1002/cche.10177

707  Pinheiro, J., D. Bates, S. DebRoy, and D. Sarkar, 2015 R Core Team. 2015. *nlme*: linear

708  and nonlinear mixed effects models. R package version 3.1-120. R package version, 3-1.

709  Poland, J., 2015 Breeding-assisted genomics. Current Opinion in Plant Biology, Vol. 24, pp.

710  119–124 https://doi.org/10.1016/j.pbi.2015.02.009

711  Poland, J. A., and R. J. Nelson, 2011 In the eye of the beholder: The effect of rater variability

712  and different rating scales on QTL mapping. Phytopathology, 101(2), 290–298

713  https://doi.org/10.1094/PHYTO-03-10-0087

714  Poland, J. A., P. J. Brown, M. E. Sorrells, and J. L. Jannink, 2012 Development of high-density

715  genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing

716  approach. PloS one, 7(2), e32253 https://doi.org/10.1371/journal.pone.0032253

717  Poland, J., and J. Rutkoski, 2016 Advances and Challenges in Genomic Selection for Disease

718  Resistance. Annual Review of Phytopathology, 54(1), 79–98

719  https://doi.org/10.1146/annurev-phyto-080615-100056

720  Qualset, C. O., J. C. Williams, M. A. Topcu, and H. E. Vogt, 1973 The barley yellow-dwarf

721  virus in wheat: importance, sources of resistance, and heritability. In Proceedings, Fourth

722  International Wheat Genetics Symposium, Columbia, Missouri (pp. 465-470)

723  Rice, B., and A. E. Lipka, 2019 Evaluation of RR-BLUP genomic selection models that

724  incorporate peak genome-wide association study signals in maize and sorghum. Plant

725  Genome 12:180052 https://doi.org/10.3835/plantgenome2018.07.0052

726  Riedell, W. E., R. W. Kieckhefer, M. A. C. Langham, L. S. and Hesler, 2003 Root and Shoot

727  Responses to Bird Cherry-Oat Aphids and Barley yellow dwarf virus in Spring Wheat.

728  Crop Science, 43(4), 1380–1386 https://doi.org/10.2135/cropsci2003.1380

729   Rife, T. W., and J. A. Poland, 2014 Field Book: An Open-Source Application for Field Data
730        Collection on Android. Crop Science, 54(4), 1624–1627
731        https://doi.org/10.2135/cropsci2013.08.0579

732   Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering, 1974 Monitoring vegetation systems
733        in the Great Plains with ERTS. NASA special publication, 351(1974), 309

734   Rutkoski, J. E., J. A. Poland, R. P. Singh, J. Huerta-Espino, S. Bhavani, *et al*., 2014 Genomic
735        Selection for Quantitative Adult Plant Stem Rust Resistance in Wheat. The Plant
736        Genome, 7(3) https://doi.org/10.3835/plantgenome2014.02.0006

737   Sankaran, S., A. Mishra, R. Ehsani, and C. Davis, 2010 A review of advanced techniques for
738        detecting plant diseases. Computers and Electronics in Agriculture, 72(1), 1–13
739        https://doi.org/10.1016/j.compag.2010.02.007

740   Savary, S., L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, *et al*., 2019 The global
741        burden of pathogens and pests on major food crops. Nature ecology and evolution, 3(3),
742        430-439 https://doi.org/10.1038/s41559-018-0793-y

743   Shah, S. J. A., M. Bashir, and N. Manzoor, 2012 A review on barley yellow dwarf virus. In Crop
744        Production for Agricultural Improvement, Vol. 9789400741164, pp. 747–782
745        https://doi.org/10.1007/978-94-007-4116-4_29

746   Sharma, H., H. Ohm, L. Goulart, R. Lister, R. Appels, *et al*., 1995 Introgression and
747        characterization of barley yellow dwarf virus resistance from *Thinopyrum intermedium*
748        into wheat. Genome, 38(2), 406–413 https://doi.org/10.1139/g95-052

749   Singh, R. P., P.A. Burnett, M. Albarran, and S. Rajaram, 1993 *Bdv1*: a gene for tolerance to
750        barley yellow dwarf virus in bread wheats. Crop Science, 33(2), 231-234
751        https://doi.org/10.2135/cropsci1993.0011183X003300020002x

752   Singh, D., X. Wang, U. Kumar, L. Gao, M. Noor, *et al*., 2019 High-throughput phenotyping
753        enabled genetic dissection of crop lodging in wheat. Frontiers in plant science, 10, 394
754        https://doi.org/10.3389/fpls.2019.00394

755   Tola, J. E., and W. E. Kronstad, 1984The genetics of resistance to barley yellow dwarf virus in
756        wheat. In Barley Yellow Dwarf Workshop, Mexico, DF (Mexico), 6-8 Dec 1983.
757        CIMMYT

758 Walls, J., E. Rajotte, and C. Rosa, 2019 The past, present, and future of barley yellow dwarf
759      management. Agriculture (Switzerland), 9(1), 1–16
760      https://doi.org/10.3390/agriculture9010023
761 Wang, X., P. Silva, N. M. Bello, D. Singh, B. Evers, *et al.*, 2020 Improved Accuracy of High-
762      Throughput Phenotyping from Unmanned Aerial Systems by Extracting Traits Directly
763      from Orthorectified Images. Frontiers in Plant Science, 11, 1616
764      https://doi.org/10.3389/fpls.2020.587093
765 Wang, X., D. Singh, S. Marla, G. Morris, and J. Poland, 2018 Field-based high-throughput
766      phenotyping of plant height in sorghum using different sensing technologies. Plant
767      Methods, 14(1), 53 https://doi.org/10.1186/s13007-018-0324-5
768 Zadoks, J. C., T. T. Chang, and C. F. Konzak, 1974 A decimal code for the growth stages of
769      cereals. Weed research 14.6 (1974): 415-421
770 Yin, LiLin, 2020 CMplot: Circle Manhattan Plot. R package version 3.6.2 https://CRAN.R-
771      project.org/package=CMplot
772 Yu, J., G. Pressoir, W. H. Briggs, I. V. Bi, M. Yamasaki, *et al.*, 2006 A unified mixed-model
773      method for association mapping that accounts for multiple levels of relatedness. Nature
774      genetics, 38(2), 203-208 https://doi.org/10.1038/ng1702
775 Zhang, Z., Z. Lin, and Z. Xin, 2009 Research progress in BYDV resistance genes derived from
776      wheat and its wild relatives. Journal of Genetics and Genomics, 36(9), 567–573
777      https://doi.org/10.1016/S1673-8527(08)60148-4
778 Zheng, Q., W. Huang, X. Cui, Y. Dong, Y. Shi, *et al.*, 2018 Identification of Wheat Yellow Rust
779      Using Optimal Three-Band Spectral Indices in Different Growth Stages. Sensors, 19(1),
780      35 https://doi.org/10.3390/s19010035
781

**Table 1 –** Field experimental details for the five wheat nurseries

| Season | 2015 – 2016 | 2016 – 2017 | 2017 – 2018 | 2018 – 2019 | 2019 – 2020 |
|---|---|---|---|---|---|
| Location | Rocky Ford farm | | Ashland Bottoms farm | | |
| | 39˚13'45.60" N, 96˚34'41.21" W | | 39˚07'53.76" N, 96˚37'05.20" W | | |
| Planting Date | Sep. 17, 2015 | Sep. 12, 2016 | Sep. 19, 2017 | Sep. 17, 2018 | Sep. 17, 2019 |
| Number of Entries | 68 | 52 | 81 | 81 | 107 |
| Number of Plots | 504 | 360 | 400 | 392 | 684 |
| Field Design | split-plot with insecticide treatment as major factor effect and wheat genotype as secondary factor | | | | |
| Replications | 3 | 3 | 2 | 2 | 2 |
| Plot Size | 6 rows plots - 1.5 m × 2.4 m | | | | |
| BYD Evaluation | April 28, 2016 | May 12, 2017 | May 19, 2018 | May 13, 2019 | May 19, 2020 |
| Harvesting Date | June 20, 2016 | June 19, 2017 | June 23, 2018 | June 28, 2019 | June 25, 2020 |

**Table 2 –** Dates of high-throughput phenotypic data collection and details of image acquisition in the five wheat nurseries screened for BYD, Kansas, USA (2015-2020).

| Season | 2015 – 2016 | 2016 – 2017 | 2017 – 2018 | 2018 – 2019 | 2019 – 2020 |
|---|---|---|---|---|---|
| UAS Platform | PheMU | | DJI Matrice 100 | | |
| Imaging Sensor | multiple digital single-lens reflex (DSLR) cameras | | MicaSense RedEdge-M | | |
| Flight/Pass speed | 0.3–0.5 m/s | | 2 m/s | | |
| Flight Dates | 2016-03-31 2016-04-07 2016-04-14 2016-05-06 | 2017-03-28 2017-04-13 2017-05-01 2017-05-09 2017-05-21 2017-05-23 2017-05-30 2017-06-05 2017-06-13 | 2018-03-30 2018-04-04 2018-04-12 2018-04-19 2018-04-23 2018-05-16 2018-06-13 | 2019-04-01 2019-04-09 2019-04-19 2019-04-26 2019-05-02 2019-05-10 2019-05-15 2019-05-23 2019-05-31 2019-06-05 2019-06-12 2019-06-17 | 2020-03-20 2020-04-11 2020-04-23 2020-05-03 2020-05-19 2020-06-05 2020-06-11 |
| Flight/Pass altitude | 0.5 m above the canopy | | 20 m AGL | | |
| In-Air Flight Duration | NA | | ~ 11–14 min | | |

30
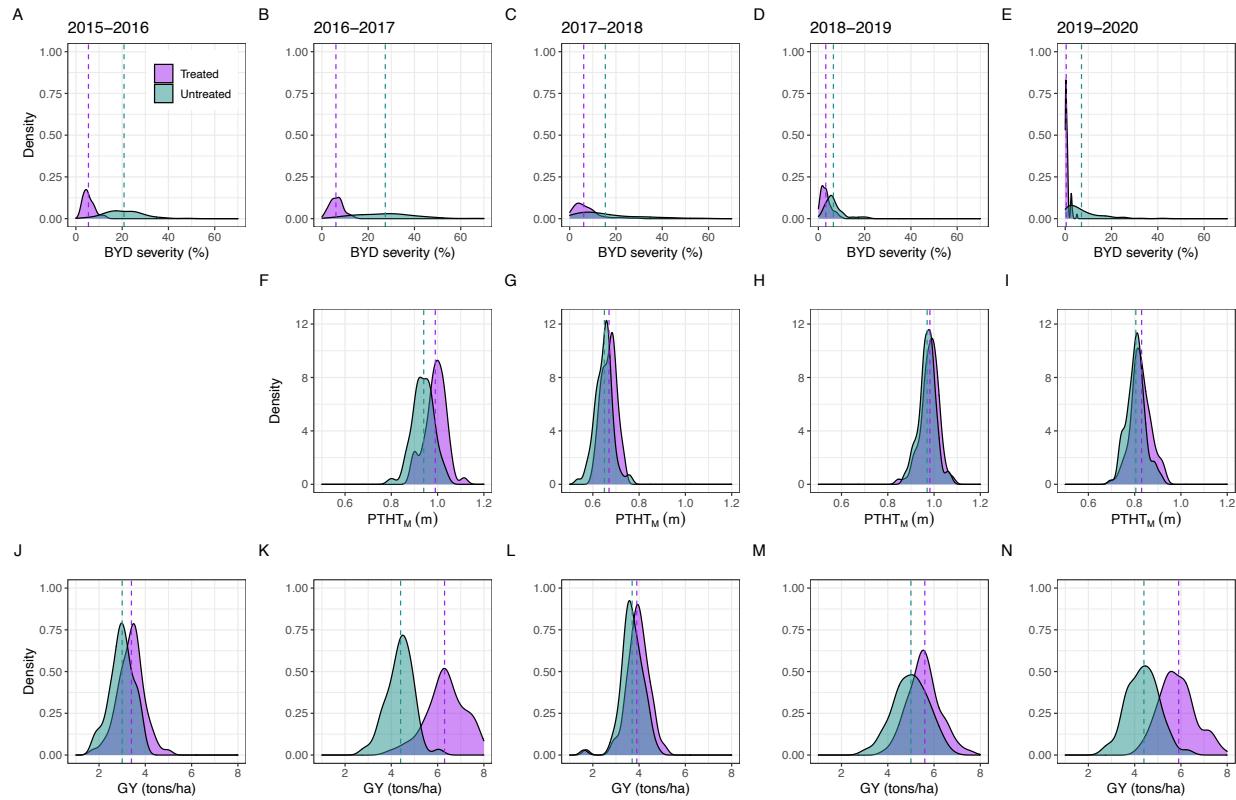
**Figure 1** Adjusted phenotypic values for the traits collected manually for five different field seasons (2015-2016 to 2019-2020). A-E): barley yellow dwarf (BYD) severity (%) characterized as the typical visual symptoms of yellowing/purpling on leaves using a 0 – 100% visual scale, F-I) manual plant height/stunting (PTHT$_M$) (meters), note that the trait was not recorded for the 2015 – 2016 season, and J-N) grain yield (GY) (tons/ha). Insecticide-treated and untreated replications are represented by purple and green, respectively. The dashed line represents the mean value for the trait in each treatment
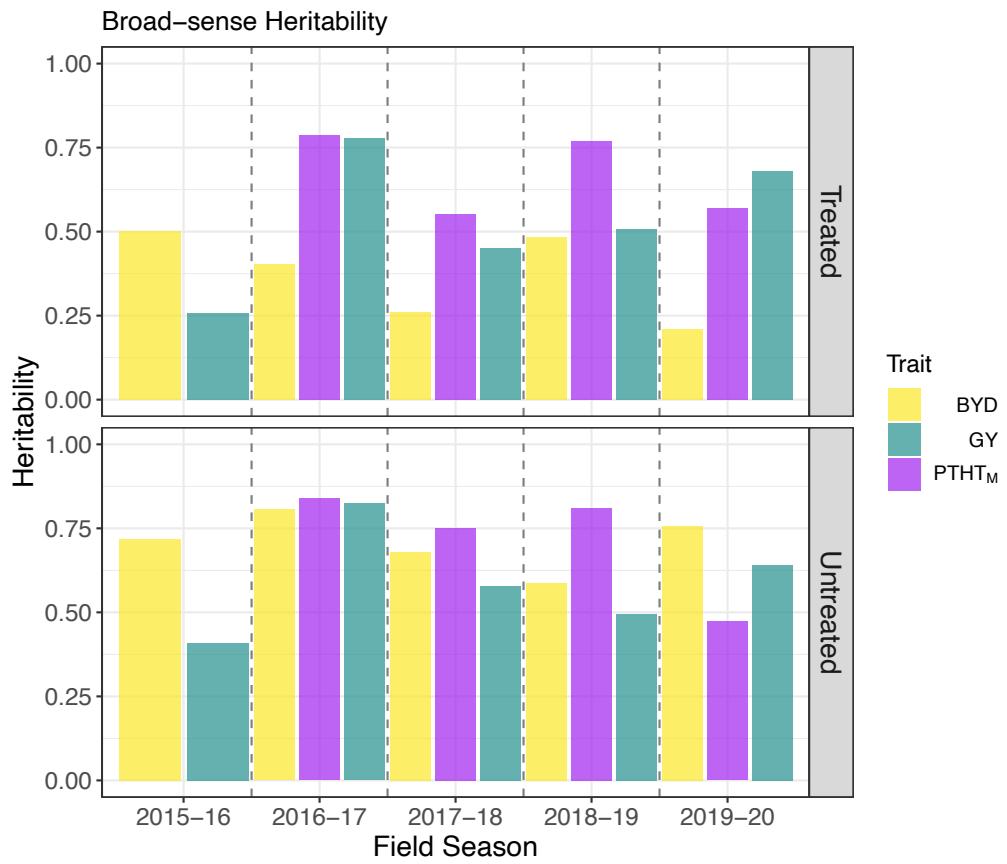
**Figure 2** Broad-sense heritability of wheat phenotypic traits collected manually, including visual barley yellow dwarf (BYD) score, plant height (PTHT$_M$) and grain yield (GY) during five different field seasons under two insecticide treatments.
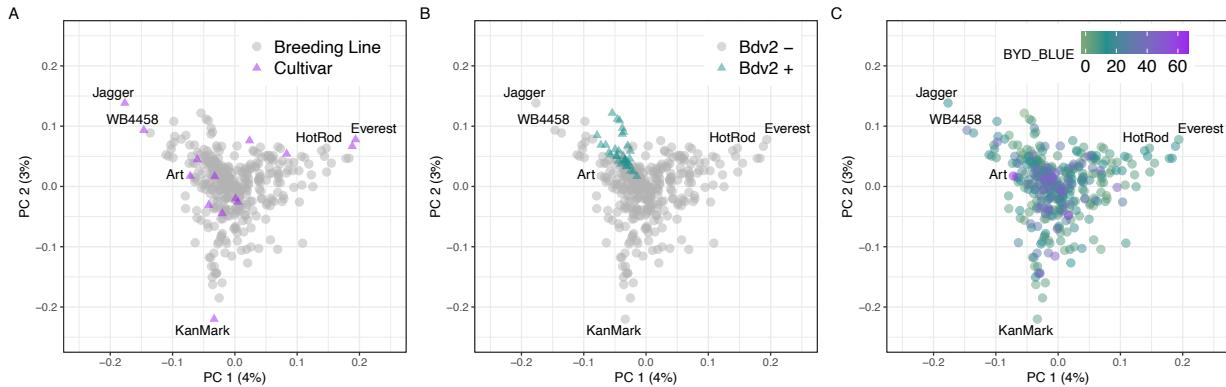
**Figure 3** Scatterplot of the first two principal component axis, made from principal component analysis on the marker matrix, n = 357 wheat lines, markers = 29,480. Each data point represents an individual wheat line that is color-coded by A) breeding status, B) prediction of *Bdv2* presence/absence, and C) adjusted mean for BYD severity (BYD BLUE) scored visually. Total variance explained by each principal component (PC) is listed on the axis.

33
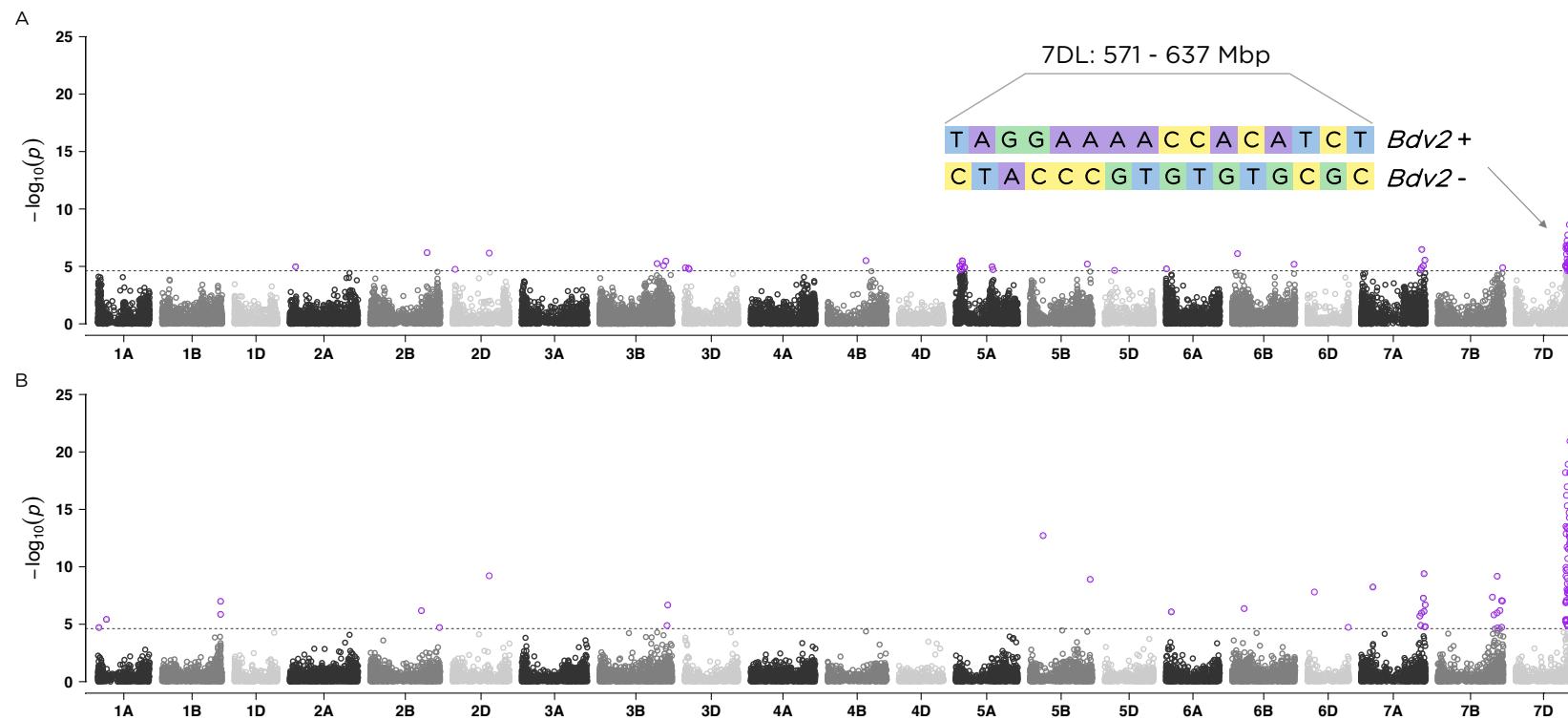
**Figure 4** Manhattan plots showing the marker-trait associations using 346 wheat accessions and 29,480 SNP markers obtained with genotyping-by-sequencing (GBS) for A) BYD severity and B) presence/absence of *Bdv2* resistance gene. The 21 labeled wheat chromosomes with physical positions are on the x-axis and y-axis is the –log10 of the p-value for each SNP marker. Horizontal dashed lines represent the false discovery rate threshold at 0.01 level and data points highlighted in purple and above the threshold represent SNPs significantly associated with the trait. In panel a, the length of the region and the haplotypes defined by the significant SNP markers is displayed.
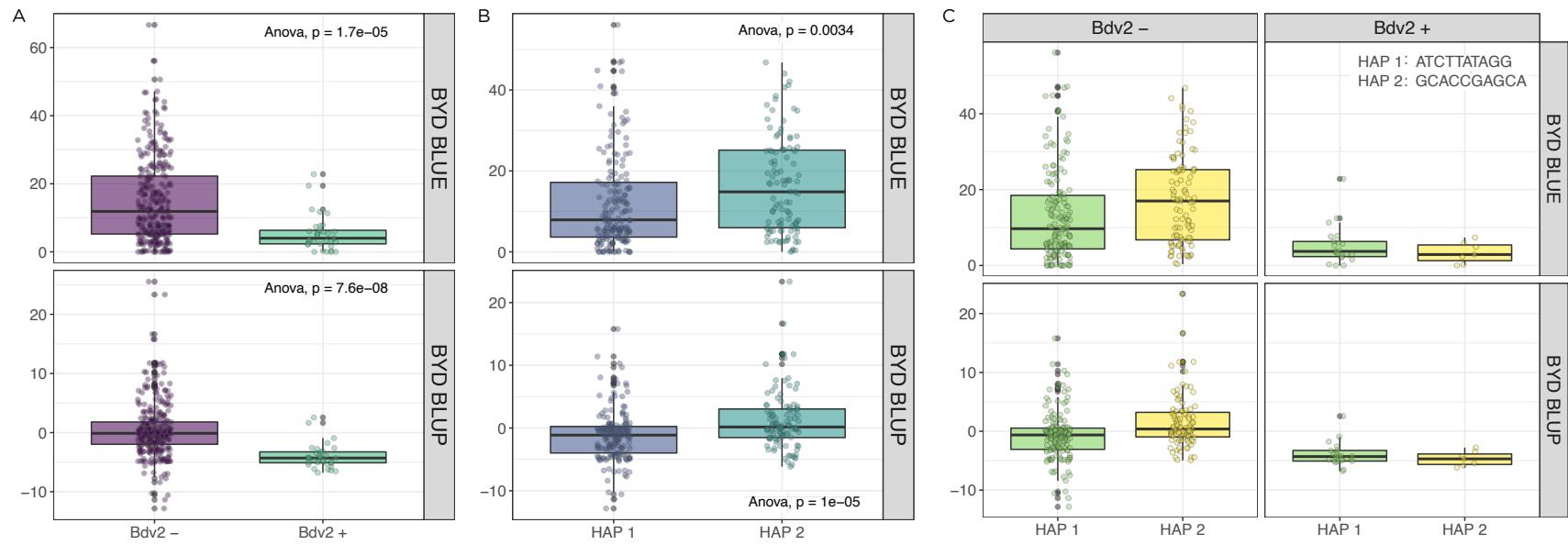
**Figure 5** Measurement of barley yellow dwarf disease severity in wheat based on certain haplotype effects were panel A) represents the translocation segment carrying the resistance gene Bdv2, B) displays the two haplotypes for the significant region on chromosome 5AS, and C) shows the combination of Bdv2 resistance gene and 5A haplotype. Boxplots showing the significant reduction of BYD disease severity by averaging the phenotypic best linear unbiased estimated (BLUE) or best linear unbiased predicted (BLUP) values for the lines.
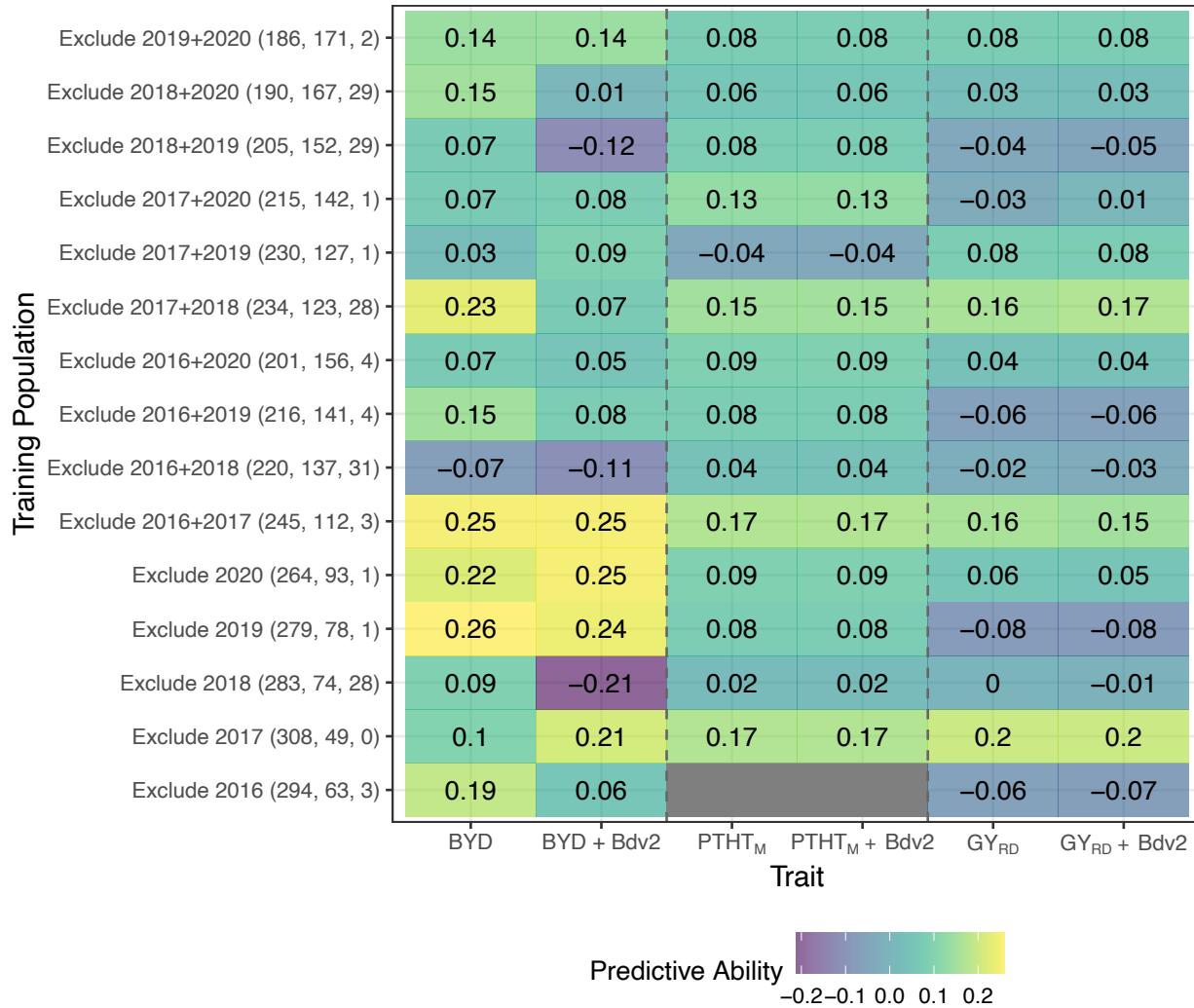
**Figure 6** Genomic selection model predictive ability where each column represents one trait, and each row shows the conformation of the training population including size of training and testing population and number of lines with presence of *Bdv2* resistance gene. The value in each cell represents the predictive ability which is the correlation between the GS predicted value (GBLUP) and the phenotypic best linear unbiased predictor (BLUP).