



Genome-Wide Association Study of Phytic Acid in Wheat Grain Unravels Markers for Improving Biofortification

Zhengyu Wen, Philomin Juliana, Harshaant S. Dhugga, Mario Pacheco, Ulises I. Martínez, Agustin Aguilar, Maria I. Ibba, Velu Govindan, Ravi P. Singh and Kanwarpal S. Dhugga*

International Maize and Wheat Improvement Center (CIMMYT), El Batán, Mexico

OPEN ACCESS

Edited by:

Francesco Sestili,
University of Tuscia, Italy

Reviewed by:

Rupesh Kailasrao Deshmukh,
National Agri-Food Biotechnology
Institute, India
Parveen Chhuneja,
Punjab Agricultural University, India

*Correspondence:

Kanwarpal S. Dhugga
K.DHUGGA@cgiar.org

Specialty section:

This article was submitted to
Plant Biotechnology,
a section of the journal
Frontiers in Plant Science

Received: 06 December 2021

Accepted: 10 January 2022

Published: 15 February 2022

Citation:

Wen Z, Juliana P, Dhugga HS, Pacheco M, Martínez UI, Aguilar A, Ibba MI, Govindan V, Singh RP and Dhugga KS (2022) Genome-Wide Association Study of Phytic Acid in Wheat Grain Unravels Markers for Improving Biofortification. *Front. Plant Sci.* 13:830147. doi: 10.3389/fpls.2022.830147

Biofortification of cereal grains offers a lasting solution to combat micronutrient deficiency in developing countries where it poses developmental risks to children. Breeding efforts thus far have been directed toward increasing the grain concentrations of iron (Fe) and zinc (Zn) ions. Phytic acid (PA) chelates these metal ions, reducing their bioavailability in the digestive tract. We present a high-throughput assay for quantification of PA and its application in screening a breeding population. After extraction in 96-well megatiter plates, PA content was determined from the phosphate released after treatment with a commercially available phytase enzyme. In a set of 330 breeding lines of wheat grown in the field over 3 years as part of a HarvestPlus breeding program for high grain Fe and Zn, our assay unraveled variation for PA that ranged from 0.90 to 1.72% with a mean of 1.24%. PA content was not associated with grain yield. High yielding lines were further screened for low molar PA/Fe and PA/Zn ratios for increased metal ion bioavailability, demonstrating the utility of our assay. Genome-wide association study revealed 21 genetic associations, six of which were consistent across years. Five of these associations mapped to chromosomes 1A, 2A, 2D, 5A, and 7D. Additivity over four of these haplotypes accounted for an ~10% reduction in PA. Our study demonstrates it is possible to scale up assays to directly select for low grain PA in forward breeding programs.

Keywords: biofortification, grain nutrition, iron, zinc, phytic acid, GWAS

INTRODUCTION

Biofortification of cereal grains is an effective route to mitigate malnourishment of underprivileged populations in developing countries (Nestel et al., 2006; Bouis and Saltzman, 2017). Most breeding efforts toward this goal have been focused on increasing the concentrations of micronutrients like iron (Fe) and zinc (Zn) in the grain (Velu et al., 2016, 2018; Cu et al., 2020). Yet, the bioavailability of these ions is known to be reduced by phytic acid (PA), which also accumulates in the grain (Turnlund et al., 1984; Hallberg et al., 1989; Roohani et al., 2013). Once chelated by PA, these metal micronutrient ions pass through the digestive tract without being absorbed (Roohani et al., 2013).

Antinutritional effects of PA have been studied in both humans and model animals. Addition of PA at a PA/Zn molar ratio of 15:1 to a formulated diet caused a 50% reduction in Zn absorption in a group of young men (Turnlund et al., 1984). Another study reported a decrease of 82% in Fe absorption upon dietary PA supplementation at a PA/Fe molar ratio of ~5:1 (Hallberg et al., 1989). Although it is difficult to determine the lowest limit of a healthy dietary PA/Zn or PA/Fe ratio, mainly because of the interference from other dietary components, it is well documented that the negative effect of PA on the absorption of micronutrient ions is significant (Yu et al., 2010). In developing countries, where cereals serve as staple food, higher dietary PA poses an even greater threat to micronutrient sufficiency than in developed countries (Ma et al., 2007; Al Hasan et al., 2016). Excreted PA from monogastric livestock pollutes the environment. Because monogastric animals (poultry and swine) cannot metabolize PA, excreted PA in the feces contributes to freshwater eutrophication (Brinch-Pedersen et al., 2002). Breeding for low PA crop varieties offers durable solution to these concerns.

Grains from low phytic acid (*lpa*) mutants contain significantly lower PA than the wildtype in maize (Raboy et al., 2000; Shi et al., 2003; Shi et al., 2005). Nutritional benefits of *lpa* were demonstrated in an animal experiment where rats fed with *lpa* maize meal absorbed twice the Zn as compared to those fed normal meal (Lonnerdal et al., 2011). But undesirable pleiotropic effects, for example, reduced seed germination and low grain yield, associated with the *lpa* mutant have limited its breeding application (Raboy et al., 2000; Pilu et al., 2005). Alternatively, a number of transgenic approaches silencing different steps of the PA biosynthetic pathway resulted in lower grain PA content (Kuwano et al., 2006; Shi et al., 2007; Ali et al., 2013). However, consumer opposition to genetically modified crops makes it an uphill task for these products to be made available for the target geographies, particularly when stewardship remains a concern (Wulff and Dhugga, 2018). It is desirable, thus, to explore genetic determinants of seed PA concentration, so that its content in the grain could be reduced within the constraints of natural variation by forward breeding.

Recently, a genome-wide association study (GWAS) was reported on exploring the genetics of the PA content in the rice grain (Perera et al., 2019). With only 69 accessions and inconsistent planting conditions (field and glasshouse), the results were inconclusive. Clearly, the throughput of PA quantification has been an impediment in studying the genetics of its accumulation in the grain (Schlemmer et al., 2009; Sivakumaran and Kothalawala, 2018).

Currently, there are several PA quantification methods available (Gao et al., 2007; Sivakumaran and Kothalawala, 2018). A commonly used method involves forming a complex of PA in the extracted sample with ferric salts, centrifugation to remove the precipitate, and then measuring residual ferric in the supernatant colorimetrically after mixing with sulfosalicylic acid, which forms a colored complex with ferric ions (Gao et al., 2007). Alternatively, the precipitate can be resuspended and further analyzed to directly measure PA. Other methods involve high performance liquid chromatography, ion exchange

chromatography, or ^{31}P nuclear magnetic resonance. These low throughput methods are not suitable for screening large breeding populations.

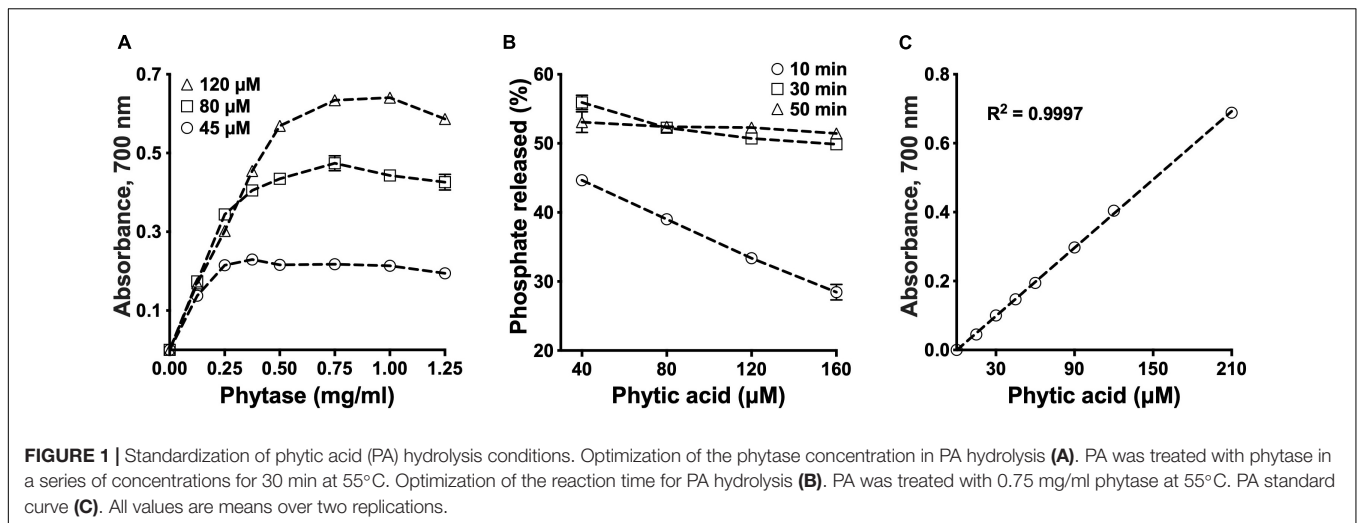
In this report we present an enzyme-base colorimetric method for high-throughput PA quantification. We used this method to measure the grain PA content of more than 300 field-grown wheat lines over 3 years. GWAS revealed several markers significantly associated with the PA content. Selection for as few as two of these markers could significantly reduce the grain PA content. Our study paves the way to breed for low grain PA by direct forward selection in breeding populations, at least in the later stages of a breeding program.

RESULTS AND DISCUSSION

Development of High-Throughput Assay for Phytic Acid Quantification

We used a commercially available phytase, an enzyme that cleaves phosphate groups from PA, to develop a high-throughput assay for PA determination. Phosphate release peaked at a phytase concentration of 0.75 mg/ml (Figure 1A). At 55°C, maximal amount of phosphate was released after 30 min of incubation (Figure 1B). No more than 52% of the total expected phosphate could be released regardless of the initial amount of PA in the reaction, the incubation time and the amount of enzyme used (Figure 1B and Supplementary Figure 1). Absorbance (OD₇₀₀) was linear over a range of 0–210 μM ($R^2 = 1$), which is equivalent to 0–16.8 nmol or 0–11 μg PA in each 100 μl reaction (80 μl PA standard + 20 μl phytase, Figure 1C). Despite minor variation for the background phosphate and the phytase protein in different batches of the commercial enzyme, no differences in enzyme activity were observed as long as the protein concentration was adjusted to 0.75 mg/ml in the reaction (Supplementary Figure 2). The phytase solution was stable for up to 6 weeks at 4°C after the initial preparation (Supplementary Figure 3).

We did not detect any additional release of phosphate by alkaline phosphatase after the phytase treatment (Supplementary Figure 4). In contrast, McKie and McCleary (2016) reported sequential hydrolysis using phytase and then alkaline phosphatase released all the phosphate from PA. Phytase from wheat bran used in our assay is classified as a 4-phytase (1D-numbering system, enzyme commission number 3.1.3.26), which first hydrolyzes the phosphate group at the D-4 position of the *myo*-inositol ring, resulting in *myo*-inositol 1,2,3,5,6-pentakisphosphate (Konietzny and Greiner, 2002). Then, a series of lower *myo*-inositol phosphates are produced, with *myo*-inositol triphosphate as the main product, which is in accordance with the 50% PA hydrolysis rate we observed (Tomlinson and Ballou, 1962). However, lower *myo*-inositol phosphates with three or fewer phosphate residues cannot be hydrolyzed by most plant alkaline phosphatases (Konietzny and Greiner, 2002). Alkaline phosphatase used in our study from calf intestine (enzyme commission number 3.1.3.1) is capable of hydrolyzing *myo*-inositol monophosphate (*mIMP*). However,



given the extremely low, if any, occurrence of *m*IMP in the PA hydrolysis products after phytase treatment, further release of phosphate was unlikely (Tomlinson and Ballou, 1962). In addition, phosphate, produced during each step of PA hydrolysis (averaged at 0.2 mM in our experiment) can strongly inhibit alkaline phosphatase ($K_i = 0.03$ mM) (Kaufman and Kleinberg, 1975). Unlike the wheat bran phytase we used in our study, other phytases, for example fungal phytases, can cleave a larger number of phosphate groups from the PA (Rao et al., 2009). The difference between our results and those of McKie and McCleary (2016) could possibly be attributed to the different phytases used in each study. As long as the phytase enzyme consistently cleaves the same number of phosphate groups from PA in different samples, PA quantification would be consistent.

Free phosphate concentration in wheat seed extracts was negligible as compared to that released by phytase hydrolysis (Supplementary Figure 5A). A comparison of the PA content measured from 330 wheat lines with or without correcting for free phosphate suggested that it was unnecessary to measure the background phosphate for GWAS (Supplementary Figure 5B), which would help reduce the number of assays by half.

A single-enzyme based, high-throughput assay we report here is reliable in quantifying grain PA content (see Supplementary Figure 6. Schematic Diagram of High-throughput Phytic Acid Quantification Method). We compared our method with a commercially available, low-throughput assay and obtained similar results ($R^2 = 0.97$, Supplementary Figure 7; McKie and McCleary, 2016).

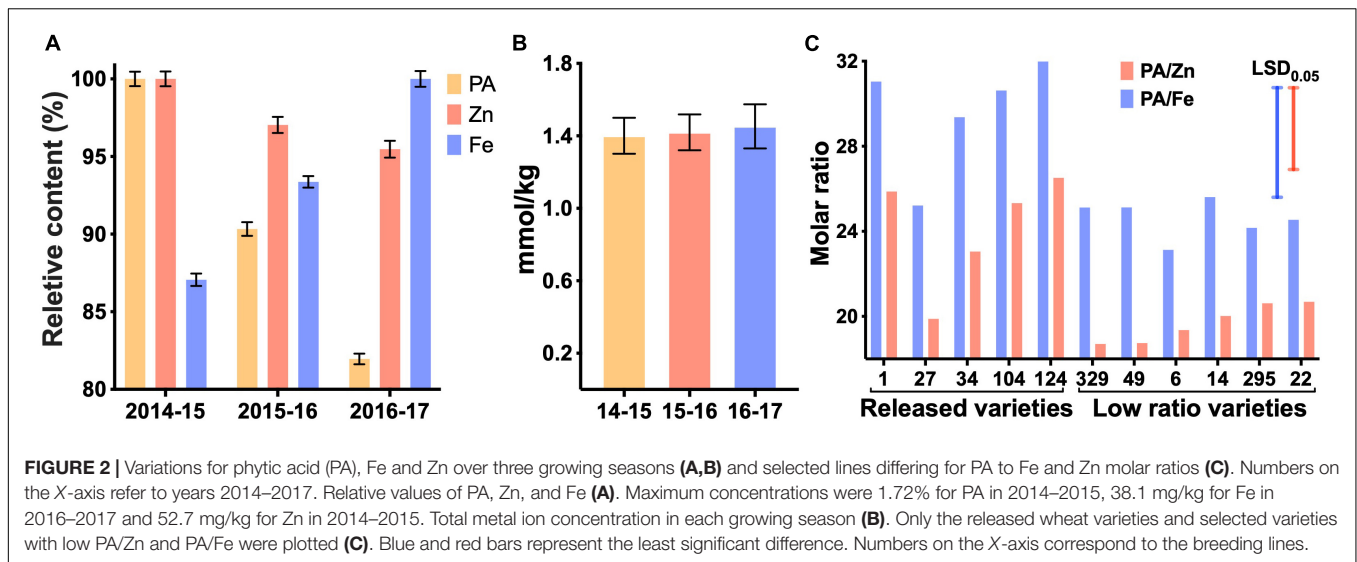
Variation for Phytic Acid in Wheat Genotypes

We measured grain PA in the wheat lines from the HarvestPlus Association Mapping (HPAM) population grown at CIMMYT over 3 years (Velu et al., 2016, 2018). The PA content ranged from 1.06 to 1.72% (w/w) in 2014–2015, from 0.95 to 1.51% in 2015–2016 and from 0.90 to 1.35% in 2016–2017 (Supplementary Figure 8). The mean PA content

was highest in the 2014–2015 season and lowest in the 2016–2017 season, indicating a role of environmental factors in PA accumulation (Figure 2A). Zinc content also followed the same pattern as PA but iron content exhibited an opposite pattern. Whereas it is difficult to pinpoint the environmental factors underlying the observed variation (Supplementary Figure 9), the data suggest that divalent cations, at least iron and zinc together, maintain a homeostatic molar concentration in the grain in fluctuating environments (Figure 2B).

The HPAM panel used in our study was previously used in the CIMMYT biofortification breeding program, the aim of which was to breed lines with increased grain Fe and Zn contents (Velu et al., 2016, 2018). Several biofortified varieties have been released from this program in India (WB-02, PBW01Zn, and HUW711), Pakistan (Zincol-16 and Akbar-19), Bangladesh (BARI-Gom 33) and Nepal (Zinc Gahun 1, Zinc Gahun 2, Bheri-Ganga, Himganga, and Khumal-Shakti) (Singh et al., 2017). However, as important factors in determining the bioavailability of Fe and Zn, the molar ratios of PA to Fe and Zn were not available at the time these selections were made. As a result, some of the released varieties with higher grain Fe and Zn did not have an optimal PA/Zn or PA/Fe molar ratios (Figure 2C).

Plotting of standardized variables of PA/Fe and PA/Zn ratios against grain yield allowed identification of high yielding lines with relatively low grain PA (Figure 3). A lack of correlation ($R^2 = 0.01$) between these two traits makes it possible to advance the recombinants with high yield and low grain PA. Varieties in the bottom-right quadrant (shaded area) should be prioritized during selection for their high bioavailability of Zn and Fe and grain yield. Among the five released breeding lines, only two met these criteria (Figure 3, green diamond). As other agronomic traits must be considered before varietal release, and the grain yield data used for this analysis were from small plots, our assay could be deployed later in the breeding cycle to advance those lines that have desired combinations of high grain metal ions and low PA.



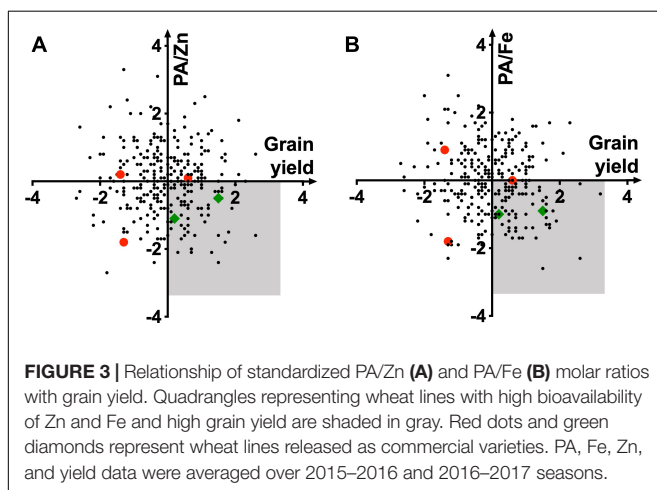
Genome-Wide Association Study for Phytic Acid Content

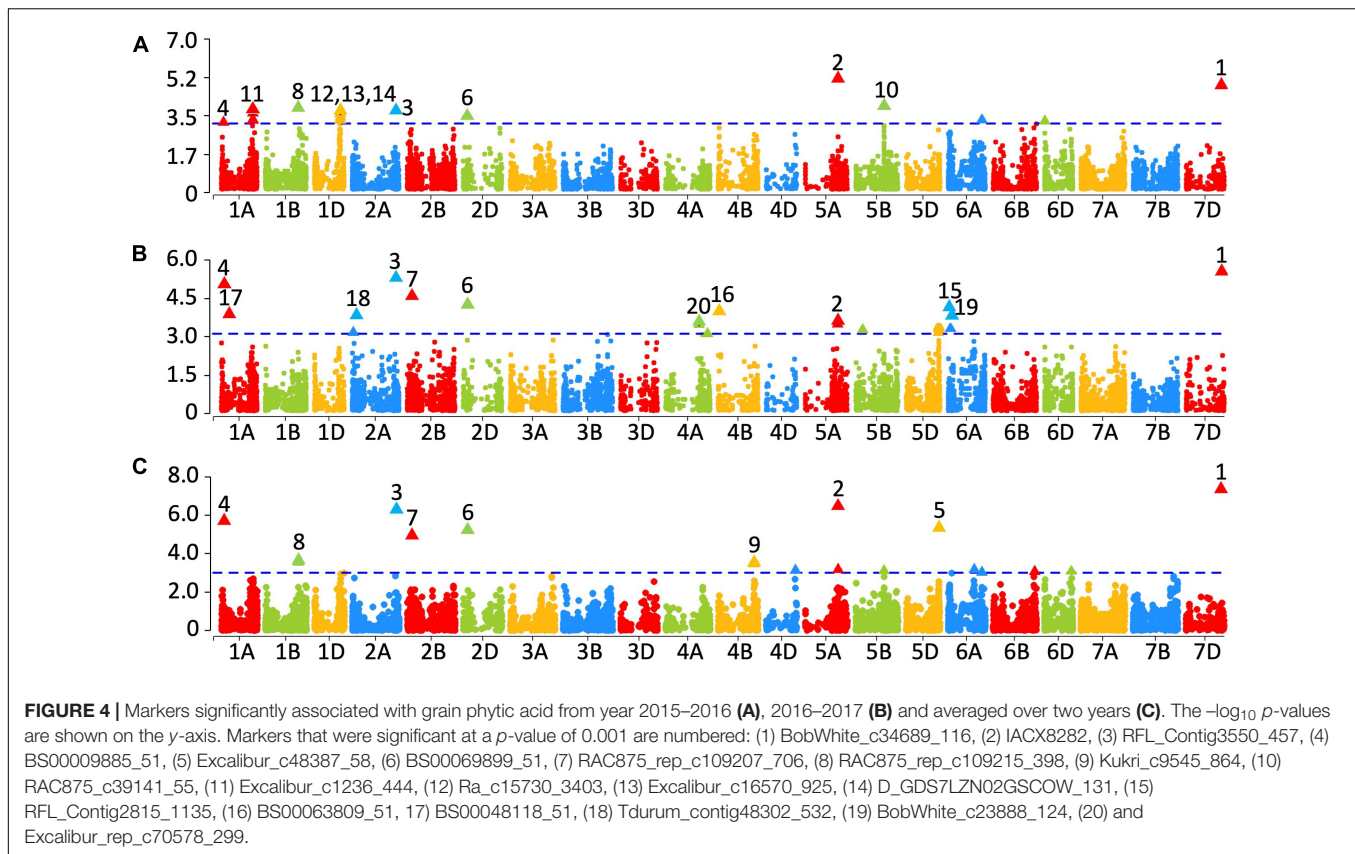
Genome-wide association study for PA content was performed using the mixed linear model with the PA content evaluations in all 3 years and the best linear unbiased estimates (BLUES) obtained from the two evaluations (2015–2016 and 2016–2017, **Figure 4**). The *p*-values for the significance of marker-trait associations, the additive effect of the markers on the PA content and the percentage variation in PA content explained by the markers are shown in **Supplementary Table 1**. In the 2015–2016 dataset, 23 markers were significantly associated with PA content at a *p*-value threshold of 0.001 (**Figure 4A** and **Supplementary Table 1**). Among those, the most significant marker, *w SNP_BE444579B-Ta_2_1* on chromosome 3B (physical location unaligned in Refseq v1.0), followed by the marker *IACX8282* on chromosome 5AL were also significant after Bonferroni correction (α level of 0.20). In addition, *BS00009885_51* on chromosome 1AS, four markers

on chromosome 1AL between base pair positions 529666321 and 530236376, *RAC875_rep_c109215_398* on chromosome 1BL, eight markers on chromosome 1DL between 430565264 and 433276844 bps, *RFL_Contig3550_457* on chromosome 2A, *BS00069899_51* on chromosome 2DS, *Ku_c69970_624* on chromosome 3AL, *RAC875_c39141_55* on chromosome 5BL, *IAAV5761* on chromosome 6AL, *RAC875_c82406_177* on chromosome 6DS and *BobWhite_c34689_116* on chromosome 7DL were also associated with PA content in the 2015–2016 dataset.

In the 2016–2017 dataset, 20 markers were significantly associated with PA content (**Figure 4B** and **Supplementary Table 1**). Among them, the most significant markers were *BobWhite_c34689_116* on chromosome 7DL, *RFL_Contig3550_457* on chromosome 2A and *BS00009885_51* on chromosome 1A. Furthermore, *BS00048118_51* on chromosome 1AS, *BS000716_30_51* and *Tdurum_contig48302_532* on chromosome 2AS, *RAC875_rep_c109207_706* on chromosome 2BS, *BS00069899_51* on chromosome 2DS, *w SNP_BE444579B-Ta_2_1* on chromosome 3B (physical location unaligned in Refseq v1.0), *Excalibur_rep_c70578_299* on chromosome 4AL, *BS00063809_51* on chromosome 4BS, *IACX8282* on chromosome 5AL, four markers on chromosome 5DL between 543272435 and 548368912 bps and three markers on chromosome 6AS including *RFL_Contig2815_1135*, *Kukri_c10226_1815*, and *BobWhite_c23888_124* were also significantly associated with the PA content in the 2016–2017 dataset. No significant associations that were consistent with the 2015–2017 data sets were found in the 2014–2015 data set. It is possible that the environmental conditions that led to high grain PA in that year also contributed to the noise level, masking the genetic signal (**Supplementary Figure 9**).

In the dataset with the BLUES obtained from the two evaluations, 17 markers were significantly associated with PA content (**Figure 4C** and **Supplementary Table 1**), among which *w SNP_BE444579B-Ta_2_1* on





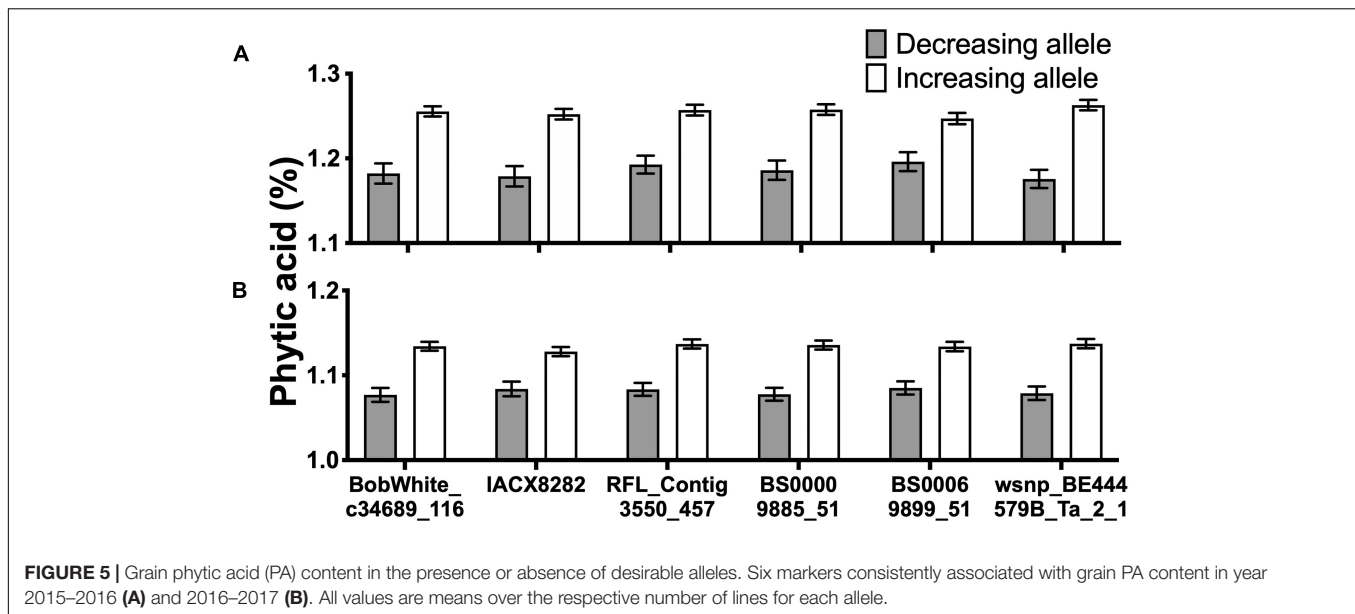
chromosome 3B (physical location unaligned in Refseq v1.0), BobWhite_c34689_116 on chromosome 7DL, IACX8282 on chromosome 5AL, RFL_Contig3550_457 on chromosome 2A, BS00009885_51 on chromosome 1A, Excalibur_c48387_58 on chromosome 5DL, BS00069899_51 on chromosome 2DS and RAC875_rep_c109207_706 on chromosome 2BS were also significant. In addition, RAC875_rep_c109215_398 on chromosome 1BL, IAAV2328 on chromosome 5AL, Kukri_c9545_864 on chromosome 4BL, IACX11246 on chromosome 4DL, IAAV2328 on chromosome 5AL, RAC875_c39141_55 on chromosome 5BL, Kukri_c1281_515 and IAAV5761 on chromosome 6AL, BS00067417_51 on chromosome 6BL, BS00109999_51 on chromosome 6DL and Kukri_c35508_426 on chromosome 7D were also significantly associated with PA content in the BLUEs dataset.

Overall, six consistent and significant associations for the PA content across two of the 3 years were obtained (Figure 5 and Supplementary Table 2). Significant phenotypic variation in PA content could be attributed to the alleles at these six markers (Figure 5). Of the 330 lines, 296 had the allelic fingerprints (increasing allele, decreasing allele, and heterozygote) of these six markers (Supplementary Figure 10). Approximately half the lines, 153, had no favorable allele (less PA), 104 lines had 1–5, and 39 lines had all six 6 favorable alleles (Supplementary Table 2). The effect of four of the six favorable alleles on grain PA was additive, showing a reduction from 1.21% with no favorable allele down to 1.13% (Figure 6), which could potentially

increase the bioavailability of Fe and Zn by ~7%. These results highlight the value of PA assays in lowering its content through forward selection.

Genes within 1 Mbp upstream and downstream of the six markers are listed in Supplementary Table 3. An ABC-2 type transporter gene (TraesCS5A01G344500) on chromosome 5A, which is located ~0.42 Mbp from marker IACX8282, presents a candidate for genetic manipulation. A similar gene has been previously shown to control the PA level in the maize grain and soybean seed (Shi et al., 2007). Silencing this gene specifically in the embryo caused a 75% reduction in the seed PA content without any noticeable pleiotropic effects (Shi et al., 2007). Two other genes (TraesCS7D01G454300 and TraesCS7D01G454500) annotated as dolichyl-diphosphooligosaccharide-protein glycosyltransferase were respectively located 1.3 and 130 Kbp distal to marker BobWhite_c34689_116 on chromosome 7D. These genes were upregulated in low PA soybean mutants, suggesting their possible role in PA formation (Yuan et al., 2017). All these genes can be silenced by gene editing. Knocking out the ABC transporter gene in particular might lead to a greater reduction in grain PA than the allelic variant in our study with the assumption that the desirable allelic variant has reduced function.

The value of our assay in screening breeding populations for improved biofortification can be judged from a reevaluation of the original selections that were made based only on high grain mineral ion concentrations (Velu et al., 2016, 2018). Fewer than



1/3rd of the 330 lines contained four or more desirable alleles for low grain PA (Figure 6), the frequency that is also reflected in the commercial varieties that have been released in different countries (Figure 2C). Only one of the five released varieties had desirable PA to mineral ion ratios (Figure 2C). In the future, aside from their usefulness in direct selection for low grain PA, the desirable haplotypes could be used to prescreen the parental lines to set up crosses in the breeding programs targeted toward improving grain biofortification (Figures 4–6).

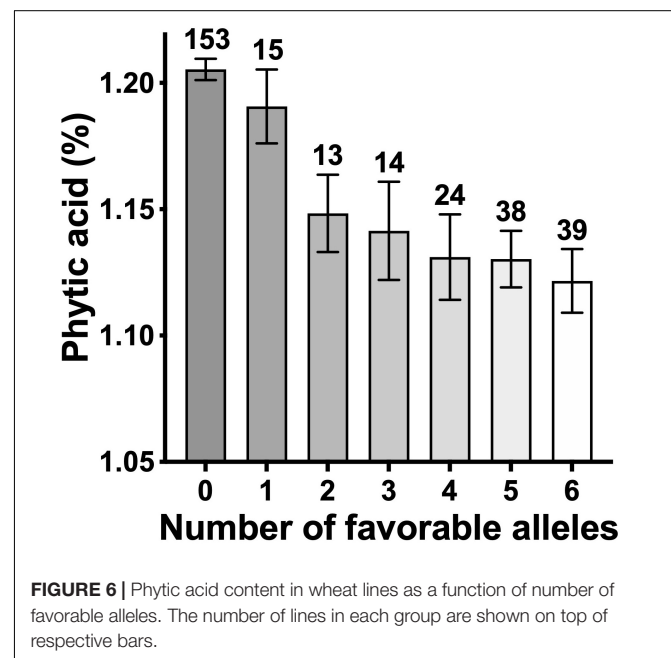
We have presented a high-throughput, reliable and easy-to-use PA quantification method that is scalable for phenotyping large breeding populations. Markers identified in our wheat grain PA GWAS could assist in screening late-stage breeding populations for the identification of low grain PA wheat varieties. Further, our method is widely applicable to other crops.

EXPERIMENTAL PROCEDURES

Wheat Population and Field Experiment

The HarvestPlus Association Mapping (HPAM) panel consisted of 330 wheat lines from CIMMYT's biofortification breeding program. This genetically diverse panel could be divided into five sub-populations that were derived from: (1) landraces; (2) *Triticum durum*-based synthetic hexaploid derivatives; (3) *Triticum dicoccon*-based synthetic hexaploid derivatives; (4) *Triticum spelta* derivatives; and (5) CIMMYT's pre-breeding derivatives of diverse progenitors including *Triticum polonicum* (Velu et al., 2018). The panel was grown in a randomized complete block design in two replications with a plot size of 2 m² at CIMMYT's experimental station in Ciudad Obregon, Sonora, Mexico (27°24'N, 109°56'W) during three successive crop seasons (2014–15, 2015–16, and 2016–17). Trials were irrigated five times throughout the crop cycle and fertilized at a rate of 200:50 (N:P) kg.ha⁻¹, of which 50:50 was applied in

pre-sowing and 150:00 at tillering stage. Diseases and pests were controlled chemically, whereas weeds were controlled manually and chemically according to CIMMYT's standard protocols. Zinc was applied at a rate of 25 kg.ha⁻¹ as ZnSO₄.7H₂O over three crop cycles to reduce soil heterogeneity for this micronutrient. Soil analysis of the experimental area showed an average Zn concentration of 1.2 ppm at soil depth of 0–30 cm, and 0.86 ppm at a soil depth of 30–60 cm. The average Fe concentration in the soil was 5.0 and 6.1 ppm, at 0–30 and 30–60 cm soil depth, respectively. Rest of the details were as previously reported (Cu et al., 2020).



Wheat head samples were collected from five plants per genotype in each plot at physiological maturity in each of the seasons. All samples were oven dried at 40°C for 4 days, threshed, grain separated and subjected to subsequent analyses. Iron and zinc were determined as previously described (Cu et al., 2020).

Phytic Acid Extraction From Wheat Seeds

Wheat grains (10 grains per line) were dried at 60°C for 3 days and ground using a GenoGrinder (model 2010, SPEX Sample Prep, United States) with small grinding vial set (6751, SPEX Sample Prep) for 15 s four times at 1750 rpm. Fifty mg of each flour sample were weighted into individual microtiter tube in a 96-tube rack format, 600 µl of 0.6 M HCl added using a Liquidator96 96-well pipettor (LIQ-96-200, METTLER TOLEDO, United States). After sealing the tubes, PA was extracted at room temperature on a shaker at 150 rpm overnight. The sample racks were centrifuged at 4800 g for 10 min in a Megafuge 40R centrifuge (Thermo Scientific, United States, rotor 75003607), followed by transferring 50 µl supernatant into new microtiter tubes. The clear PA extracts were neutralized by mixing with 50 µl 0.6 M NaOH and then diluted with 900 µl 0.2 M sodium acetate (NaOAc), pH 5.5.

Phytase Working Solution

Phytase (P1259, Sigma-Aldrich, United States) was dissolved in 0.2 M NaOAc, pH 5.5, to give a concentration of 10 mg/ml. Actual phytase concentration was quantified using a bicinchoninic acid protein determination kit (BCA1 and B9643, Sigma-Aldrich, United States) and a microplate reader (FLUOstar Omega, BMG Labtech, Germany). The phytase working solution was prepared by adjusting the actual phytase concentration to 4 mg/ml using the NaOAc buffer.

Phytic Acid and Phosphate Standards

Phytic Acid solution (593648, Sigma-Aldrich, United States) was used to prepare a series of PA standard solutions in the NaOAc buffer. Phosphate standards were prepared by dissolving KH_2PO_4 in the NaOAc buffer.

Phytic Acid Hydrolysis and Quantification

High-throughput assays were performed essentially as described for other traits previously (Silva et al., 2017, 2018). PA in 80 µl diluted extract (and PA standards and phosphate standards) was hydrolyzed by mixing with 20 µl phytase working solution and incubated at 55°C for 30 min. Then each reaction was mixed with the 100 µl color reagent and incubated at 37°C for 30 min. The color reagent is consisted of the four solutions in a ration of A:B:C:D = 1:1:1:2 (A. 2.5% (w/w) $(\text{NH}_4)_2\text{MoO}_4$, B. 3 M H_2SO_4 , C. 10% (w/w) ascorbic acid and D. H_2O). Then the reactions were centrifuged at 4700 rpm for 10 min. 100 µl supernatant of each reaction was transferred to a 96-well plate and absorbance was measured at a wavelength of 700 nm. The wavelength was determined by scanning through the absorbance of phosphate standards.

Genotyping Data, Population Structure, and Kinship Analysis

The 330 lines in the HPAM panel were genotyped using the Illumina iSelect 90 K Infinium single nucleotide polymorphism array. Quality control of the genotyping data was done by filtering markers with greater than 70% missing data, less than 5% minor allele frequency and greater than 10% heterozygosity. We then obtained the physical positions of the markers in the reference genome of wheat (Refseq v1.0) available at <https://triticeaetoolbox.org/wheat/maps> (Appels et al., 2018). Plot of SNP densities in different chromosomes or the number of SNPs within a 10 Mb window revealed moderate to good coverage in the telomeric regions, and a relatively lower coverage in the centromeric regions (**Supplementary Figure 11**). The highest number of markers were in the B genome (4,323 markers, 42.6%), followed by the A genome (4,038 markers, 39.8%) and the D genome (1,797 markers, 17.7%).

Population structure of the lines in the panel was done using the principal component analysis in TASSEL (Trait Analysis by aSSociation Evolution and Linkage) version 5. Moderate population structure was observed and the first two principal components explained 10.9 and 7.2% of the variation, respectively (**Supplementary Figure 12**). In addition, we also obtained the kinship between the lines using the centered identity-by-state method in TASSEL version 5 (Endelman and Jannink, 2012). Several lines formed clusters with a relationship of 0.4–0.6, while only a few lines had a relationship of over 0.7 (**Supplementary Figure 13**).

Genome-Wide Association Mapping and Allelic Fingerprinting

We performed genome-wide association mapping for PA content in TASSEL version 5 with the mixed linear model and used the optimum level of compression and the “population parameters previously determined” options (Yu et al., 2006; Zhang et al., 2010). The first two principal components that accounted for the population structure were used as fixed effects and the kinship matrix among the lines was used as the random effect in the mixed linear model (Price et al., 2006). We obtained the marker *p*-values, additive effects and percentage variation explained by each marker and then plotted the Manhattan plots with the $-\log_{10} p$ -values of the markers using the “R” package CMplot (Lilin-Yin, 2018). To declare significance of the markers, we used a *p*-value threshold of 0.001 and also corrected for multiple testing using the Bonferroni method at an α level of 0.20. The markers that were consistently associated with PA content in the two evaluations were then identified and the alleles associated with increasing and decreasing the PA content at these markers were fingerprinted in all the lines. We then determined the relationships between the number of PA favorable alleles (PA decreasing alleles) at the consistent markers and the PA content and visualized them using box plots, created using the R package, “ggplot2” (Ginestet, 2011). Finally, we also obtained the genes that were 100 Mb upstream and downstream of the consistently significant markers using Jbrowse

available at <https://wheat-urgi.versailles.inra.fr/Tools/JBrowse> and explored the genes in the interval to identify potential PA-associated candidate genes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in the **Supplementary Material**.

AUTHOR CONTRIBUTIONS

ZW, HSD, and KSD developed and refined the assay. ZW oversaw screening of the grain samples. MP, UIM, and AA carried out the assays in the population and MII compared different methods. VG and RPS conducted the field experiments and provided the grain samples. PJ carried out GWAS analysis. ZW and KSD wrote

and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the CGIAR Research Program on Wheat (WHEAT), HarvestPlus, Bill & Melinda Gates Foundation, and Foreign Commonwealth and Development Office of the United Kingdom.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.830147/full#supplementary-material>

REFERENCES

- Al Hasan, S. M., Hassan, M., Saha, S., Islam, M., Billah, M., and Islam, S. (2016). Dietary phytate intake inhibits the bioavailability of iron and calcium in the diets of pregnant women in rural Bangladesh: a cross-sectional study. *BMC Nutr.* 2:24.
- Ali, N., Paul, S., Gayen, D., Sarkar, S. N., Datta, K., and Datta, S. K. (2013). Development of low phytate rice by RNAi mediated seed-specific silencing of inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene (IPK1). *PLoS One* 8:e68161. doi: 10.1371/journal.pone.0068161
- Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191
- Bouis, H. E., and Saltzman, A. (2017). Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. *Glob. Food Sec.* 12, 49–58. doi: 10.1016/j.gfs.2017.01.009
- Brinch-Pedersen, H., Sorensen, L. D., and Holm, P. B. (2002). Engineering crop plants: getting a handle on phosphate. *Trends Plant Sci.* 7, 118–125. doi: 10.1016/s1360-1385(01)02222-1
- Cu, S. T., Guild, G., Nicolson, A., Velu, G., Singh, R., and Stangoulis, J. (2020). Genetic dissection of zinc, iron, copper, manganese and phosphorus in wheat (*Triticum aestivum* L.) grain and rachis at two developmental stages. *Plant Sci.* 291:110338. doi: 10.1016/j.plantsci.2019.110338
- Endelman, J. B., and Jannink, J. L. (2012). Shrinkage estimation of the realized relationship matrix. *G3* 2, 1405–1413. doi: 10.1534/g3.112.004259
- Gao, Y., Shang, C., Maroof, M. A. S., Biyashev, R. M., Grabau, E. A., Kwanyuen, P., et al. (2007). A modified colorimetric method for phytic acid analysis in soybean. *Crop Sci.* 47, 1797–1803. doi: 10.2135/cropsci2007.03.0122
- Ginestet, C. (2011). ggplot2: Elegant graphics for data analysis. *J. R. Stat. Soc. Series* 174, 245–245. doi: 10.1111/j.1467-985x.2010.00676_9.x
- Hallberg, L., Brune, M., and Rossander, L. (1989). Iron absorption in man: ascorbic acid and dose-dependent inhibition by phytate. *Am. J. Clin. Nutr.* 49, 140–144. doi: 10.1093/ajcn/49.1.140
- Kaufman, H. W., and Kleinberg, I. (1975). Hydrolysis of phytate and its inositol phosphate intermediates by an acid and an alkaline phosphatase. *Arch. Oral Biol.* 20, 157–160. doi: 10.1016/0003-9969(75)90002-3
- Konietzny, U., and Greiner, R. (2002). Molecular and catalytic properties of phytate-degrading enzymes (phytases). *Int. J. Food Sci. Technol.* 37, 791–812. doi: 10.1371/journal.pone.0094403
- Kuwano, M., Ohyama, A., Tanaka, Y., Mimura, T., Takaiwa, F., and Yoshida, K. T. (2006). Molecular breeding for transgenic rice with low-phytic-acid phenotype through manipulating myo-inositol 3-phosphate synthase gene. *Mol. Breed.* 18, 263–272. doi: 10.1007/s11032-006-9038-x
- Lilin-Yin (2018). *Package 'CMAPlot' Version 3.4.0*.
- Lonnerdal, B., Mendoza, C., Brown, K. H., Rutger, J. N., and Raboy, V. (2011). Zinc absorption from low phytic acid genotypes of maize (*Zea mays* L.), barley (*Hordeum vulgare* L.), and rice (*Oryza sativa* L.) assessed in a suckling rat pup model. *J. Agric. Food Chem.* 59, 4755–4762. doi: 10.1021/jf1043663
- Ma, G., Li, Y., Jin, Y., Zhai, F., Kok, F. J., and Yang, X. (2007). Phytate intake and molar ratios of phytate to zinc, iron and calcium in the diets of people in China. *Eur. J. Clin. Nutr.* 61, 368–374. doi: 10.1038/sj.ejcn.1602513
- McKie, V. A., and McCleary, B. V. (2016). A novel and rapid colorimetric method for measuring total phosphorus and phytic acid in foods and animal feeds. *J. AOAC Int.* 99, 738–743. doi: 10.5740/jaoacint.16-0029
- Nestel, P., Bouis, H. E., Meenakshi, J. V., and Pfeiffer, W. (2006). Biofortification of staple food crops. *J. Nutr.* 136, 1064–1067.
- Perera, I., Fukushima, A., Arai, M., Yamada, K., Nagasaka, S., Seneweera, S., et al. (2019). Identification of low phytic acid and high Zn bioavailable rice (*Oryza sativa* L.) from 69 accessions of the world rice core collection. *J. Cereal Sci.* 85, 206–213. doi: 10.1016/j.jcs.2018.12.010
- Pilu, R., Landoni, M., Cassani, E., Doria, E., and Nielsen, E. (2005). The maize lpa241 mutation causes a remarkable variability of expression and some pleiotropic effects. *Crop Sci.* 45, 2096–2105. doi: 10.2135/cropsci2004.0651
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847
- Raboy, V., Gerbasi, P. F., Young, K. A., Stoneberg, S. D., Pickett, S. G., Bauman, A. T., et al. (2000). Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiol.* 124, 355–368. doi: 10.1104/pp.124.1.355
- Rao, D. E. C. S., Rao, K. V., Reddy, T. P., and Reddy, V. D. (2009). Molecular characterization, physicochemical properties, known and potential applications of phytases: an overview. *Crit. Rev. Biotechnol.* 29, 182–198. doi: 10.1080/0738850902919571
- Roohani, N., Hurrell, R., Kelishadi, R., and Schulin, R. (2013). Zinc and its importance for human health: an integrative review. *J. Res. Med. Sci.* 18, 144–157.
- Schlemmer, U., Frolich, W., Prieto, R. M., and Grases, F. (2009). Phytate in foods and significance for humans: food sources, intake, processing, bioavailability, protective role and analysis. *Mol. Nutr. Food Res.* 53, S330–S375. doi: 10.1002/mnfr.200900099
- Shi, J. R., Wang, H. Y., Schellin, K., Li, B. L., Faller, M., Stoop, J. M., et al. (2007). Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. *Nat. Biotechnol.* 25, 930–937. doi: 10.1038/nbt1322
- Shi, J., Wang, H., Hazebroek, J., Ertl, D. S., and Harp, T. (2005). The maize low-phytic acid 3 encodes a myo-inositol kinase that plays a role in phytic acid biosynthesis in developing seeds. *Plant J.* 42, 708–719. doi: 10.1111/j.1365-313X.2005.02412.x

- Shi, J., Wang, H., Wu, Y., Hazebroek, J., Meeley, R. B., and Ertl, D. S. (2003). The maize low-phytic acid mutant *lpa2* is caused by mutation in an inositol phosphate kinase gene. *Plant Physiol.* 131, 507–515. doi: 10.1104/pp.014258
- Silva, I. T., Abbaraju, H. K. R., Fallis, L. P., Liu, H. J., Lee, M., and Dhugga, K. S. (2017). Biochemical and genetic analyses of N metabolism in maize testcross seedlings: 1. Leaves. *Theor. Appl. Genet.* 130, 1453–1466. doi: 10.1007/s00122-017-2900-x
- Silva, I. T., Abbaraju, H. K. R., Fallis, L. P., Liu, H. J., Lee, M., and Dhugga, K. S. (2018). Biochemical and genetic analyses of N metabolism in maize testcross seedlings: 2. Roots. *Theor. Appl. Genet.* 131, 1191–1205. doi: 10.1007/s00122-018-3071-0
- Singh, R., Govindan, V., Andersson, M. S., Bouis, H., and Jamora, N. (2017). *Zinc-Biofortified Wheat: Harnessing Genetic Diversity for Improved Nutritional Quality*. Bonn: Global Crop Diversity Trust.
- Sivakumaran, K., and Kothalawala, S. (2018). An overview of the analytical methods for food phytates. *Int. J. Chem. Stud.* 6, 2016–2020.
- Tomlinson, R. V., and Ballou, C. E. (1962). Myoinositol polyphosphate intermediates in the dephosphorylation of phytic acid by phytase. *Biochemistry* 1, 166–171. doi: 10.1021/bi00907a025
- Turnlund, J. R., King, J. C., Keyes, W. R., Gong, B., and Michel, M. C. (1984). A stable isotope study of zinc absorption in young men: effects of phytate and alpha-cellulose. *Am. J. Clin. Nutr.* 40, 1071–1077. doi: 10.1093/ajcn/40.5.1071
- Velu, G., Crossa, J., Singh, R. P., Hao, Y. F., Dreisigacker, S., Perez-Rodriguez, P., et al. (2016). Genomic prediction for grain zinc and iron concentrations in spring wheat. *Theor. Appl. Genet.* 129, 1595–1605. doi: 10.1007/s00122-016-2726-y
- Velu, G., Singh, R. P., Crespo-Herrera, L., Juliana, P., Dreisigacker, S., Valluru, R., et al. (2018). Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Sci. Rep.* 8:13526. doi: 10.1038/s41598-018-31951-z
- Wulff, B. B. H., and Dhugga, K. S. (2018). Wheat—the cereal abandoned by GM Genetic modification of wheat for disease resistance could help stabilize food production. *Science* 361, 451–452.
- Yu, J. M., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Yu, Y., Lu, L., Wang, R. L., Xi, L., Luo, X. G., and Liu, B. (2010). Effects of zinc source and phytate on zinc absorption by in situ ligated intestinal loops of broilers. *Poult. Sci.* 89, 2157–2165. doi: 10.3382/ps.2009-00486
- Yuan, F. J., Yu, X. M., Dong, D. K., Yang, Q. H., Fu, X. J., Zhu, S. L., et al. (2017). Whole genome-wide transcript profiling to identify differentially expressed genes associated with seed field emergence in two soybean low phytate mutants. *Bmc Plant Biol.* 17:16. doi: 10.1186/s12870-016-0953-7
- Zhang, Z. W., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–U118. doi: 10.1038/ng.546

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Wen, Juliana, Dhugga, Pacheco, Martínez, Aguilar, Iba, Govindan, Singh and Dhugga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.