Molecular Biotechnology
https://doi.org/10.1007/s12033-022-00535-8

ORIGINAL PAPER

Preliminary Dissection of Grain Yield and Related Traits at Differential Nitrogen Levels in Diverse Pre-Breeding Wheat Germplasm Through Association Mapping

Achla Sharma[1] · Mian A. R. Arif[2] · M. Shamshad[1] · Kanwardeep S. Rawale[3] · Anureet Brar[1] · Juan Burgueño[4] · Sajid Shokat[2] · Ravinder Kaur[1] · Parsahnt Vikram[5] · Puja Srivastava[1] · Nitika Sandhu[1] · Jayesh Singh[1] · Satinder Kaur[1] · Parveen Chhuneja[1] · Sukhwinder Singh[3]

Received: 28 June 2021 / Accepted: 14 July 2022

Abstract

Development of nutrient efficient cultivars depends on effective identification and utilization of genetic variation. We characterized a set of 276 pre-breeding lines (PBLs) for several traits at different levels of nitrogen application. These PBLs originate from synthetic wheats and landraces. We witnessed significant variation in various traits among PBLs to different nitrogen doses. There was ~ 4–18% variation range in different agronomic traits in response to nitrogen application, with the highest variation for the biological yield (BY) and the harvest index. Among various agronomic traits measured, plant height, tiller number, and BY showed a positive correlation with nitrogen applications. GWAS analysis detected 182 marker-trait associations (MTAs) (at p-value < 0.001), out of which 8 MTAs on chromosomes 5D, 4A, 6A, 1B, and 5B explained more than 10% phenotypic variance. Out of all, 40 MTAs observed for differential nitrogen application response were contributed by the synthetic derivatives. Moreover, 20 PBLs exhibited significantly higher grain yield than checks and can be selected as potential donors for improved plant nitrogen use efficiency (pNUE).

Keywords Wheat · pNUE · Pre-breeding lines · Nitrogen · Landraces · GWAS · Marker-trait-association

Abbreviations

pNUE	Plant Nitrogen Use Efficiency
MTA	Marker Trait Associations
PBLs	Pre-Breeding lines
GWAS	Genome-wide association analysis

Introduction

Bread wheat is one of the most important cereal crops in the world and accounts for 20% of the daily calories and protein for more than 1/3rd of the world's population [31]. Currently, wheat occupies around 220 million hectaresof area globally, with an annual harvest of 731 million tones [31]. In India, wheat is cultivated on 14% of the global area (29.7 million ha), leading to a harvest of 99.9 million tons of grain with peak productivity of 3.37 t/ha, thus making it the second-highestwheat-producing country after China [27]. The agriculture production among major cereals was doubled mainly because of semi-dwarf and inputs efficient cultivars [25]. The use of nitrogen-based fertilizers has significantly increased in both developing and developed world countries over the years, leading to enhanced production. However, for most crop plants, the nitrogen use efficiency (NUE) is not more than 50% of applied nitrogen [24].

Improving the NUE of crop plants by developing designer crops having high nutrient use efficiency along with high yields is a major goal in both developing and developed countries. Moreover, improving the nitrogen use efficiency (NUE) in crop plants has both economic and environmental benefits. Plant nitrogen use efficiency (pNUE) is a combination of both efficiencies of nitrogen (N) uptake and NUE [23]. Different efforts in improving pNUE in wheat have ranged from exploring variation among wheat genotypes and genetic manipulation of N uptake, nitrate allocation, N

Achla Sharma and Mian A. R. Arif have contributed equally to this work.

✉ Achla Sharma
achla12@gmail.com; achla@pau.edu

✉ Sukhwinder Singh
sukhmatola@hotmail.com

Extended author information available on the last page of the article

Published online: 30 July 2022

metabolism, and its allocation [13, 15, 23, 36]. The potential utility and impact of landraces and exotic material has been exploited to address biotic and abiotic stress tolerance [12, 22]. However, a few explorations to identify natural variation in wheat for NUE have revealed presence of significant variation in historical and current varieties [14]. The variation in distant wild relatives of wheat for improved NUE with suitable agronomic traits such as better root architecture has been identified [17, 21]. Thus, highlighting a potential pathway for the transfer of pNUE from ancestral parents and distant relatives to modern high-yielding semi-dwarf varieties could be beneficial. N uptake variation in wheat has been mainly observed by measuring the amount of N applied and N harvested [5]. But the direct phenotyping of N uptake is expensive, and the study of the N level treatment on different agronomic traits has been performed from the point of view of certain reference varieties that donot give insights into the genetic variation observed by N treatment. Besides, a significant emphasis of N uptake variation in wheat focused on roots and root-associated traits that are difficult to phenotype [28].

The explorations for exotic germplasm for the identification of novel variation require a long-time commitment as breeders have to segregate desirable alleles from the pool of desirable and undesirable alleles. These efforts have yielded success for biotic stress tolerance but little to no success for complex traits like abiotic stress tolerance and NUE [9, 19, 32]. Even though underlying QTLs have been identified for NUE in wheat, their effective use in different breeding programsis limited due to the diverse nature of crosses. Seeds for Discovery program at CIMMYT used a novel approach for developing a set of advanced lines that carried exotic alleles from landraces and wild relatives of wheat in the background of elite germplasm [34]. The genotyping and phenotyping analysis of the advanced material helped in the elimination of undesirable agronomic traits from exotic germplasm while capturing the maximum variation [34].

In the present study, we evaluated the variation in the advanced wheat germplasm of effect of N supply on plant height (PH), fertile tiller number (FTN), spikelet number (SPKL), grain yield (GY), biological yield (BY), harvest index (HI), and thousand kernel weight (TW)) under different nitrogen applications, thus focusing on yield increase as an indicator for pNUE efficiency. Genome-wide association study (GWAS) was used to locate to identify marker-trait associations with agronomic traits. Our study has highlighted the presence of significant variation in the PBLs for N uptake efficiency. Also, the differential response of the agronomic traits across genetically diverse PBLs highlighted their potential use as phenotypic markers for pNUE. Identification of genetic footprints from exotic germplasm regulating N uptake can be helpful to develop breeder-friendlymolecular markers for use in future wheat breeding for improved nitrogen use.

Materials and Methods

Plant Material

The plant material is composed of 276 PBLs created through link top cross (LTP) as reported by Singh et al. [34] at CIMMYT in the "Seeds of Discovery" Project [26]. Table S1 provides the pedigree, genotypic ID, and all other relevant details of all 276 PBLs. Two high-yielding cultivars recommended for cultivation in the region (PBW725 and HD3086) were used as checks for comparison.

Experiment Location and Climatic Conditions

The study was accomplished at Wheat Section, Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, India (30° 54' N latitude, 75° 48' E longitude, and 247 m above sea level), where the soil is classified as the loamy sand having neutral pH ~ 6–8 (Table S2). The conventionalrice-wheat rotation has beenfollowed on thefield forthe last 15 years. The climate of the area is characterized as subtropical and semi-arid with cold winters from November to January and mild climate during February and March., and very hot and dry during summer from April to June, followed by hot and humid conditions from July to September. The daily minimum temperature ranges from 0–4 °C in January, and the maximum temperature during May ranges from 40 to 45 °C. The weather data (including weekly temperature, rainfall, relative humidity, and saturated vapor pressure) during experimental years 2017–18 (S1) and 2018–19 (S2) were recorded and provided in Fig (S1 and S2).

Experimental Design and Phenotypic Data

The study was conducted in six environments (2 years × 3 N doses) in an alpha lattice design with two replications per N dose. A total of two experiments were conducted in 2017–18 (S1) and 2018–19 (S2). The germplasm was sown on the 20th of November and 17th of November in S1 and S2, respectively. Seeds from 2016 to 17 multiplication were sown on 20th Nov. 2017 for the experiment in S1, while seeds from the S1 experiment were sown on 15th Nov. 2018 for a subsequent experiment in S2 using 100 kg seed rate

per hakeeping a row spacing of 20 cm. There were 280 plots per replication per treatment, with a gross plot size of 2.5 m × 2.4 m = 6.0m^{2}and a net plot after discarding the border rows measured 2 m long and 2 m wide with an area of 4 m^2. Each experiment was planted under three nitrogen (N) regimes viz. no N (N0), half N (N1) and full N (N2) representing 0, 60, and 120 kg N ha^{-1}, respectively. The recommended phosphorus and potassium were applied at the time of sowing. For treatments N1 and N2, half of the N was applied at the time of sowing, while the other half was applied in two equal splits, first at the crown root initiation stage and the remaining at the maximum tillering stage. N0 was also treated as a control. Recommended fungicides and insecticides were applied to control stripe rusts, brown rusts, and aphids at jointing, booting, and 10 days after anthesis to prevent diseases and pests. Weeds were controlled manually.

A total of seven traits were measured in all experiments at all three doses: (a) PH was measured in cm using the scale from the ground base of the plant to the end of the spike of five random plants per genotype, (b) FTN (Fertile tiller number) was measured using the quadrant as the number of productive tillers arising from the ground per meter square, (c) SPKLNwas measured by taking five spikes randomly from each plot whose number of spikelets were recorded manually and the average was calculated, (d) TWin grams was recorded by counting 500 seeds and then multiplying it by 2 per replicate, (e) BY in grams was calculated by weighing bundles of each genotype per plot using the beam balance, (f) GY in grams was measured on per plot basis by weighing the grains harvested from respective bundles, and (g) HI was derived from the biological yield and grain yield with the following formulas.

$$Harvest\ Index(HI) = \frac{Grain\ yield\ \left(Kgha^{-1}\right)}{Biological\ yield\ \left(Kgha^{-1}\right)}$$

Molecular Markers Analysis

Marker analysis of all 276 genotypes was achieved by extracting DNA from TC1F$_5$ plants using in-house modified cetyltrimethylammonium bromide (CTAB) method from flag leaves sampled at booting stage and subsequently preserved at −80 °C. Nano-Drop 8000 spectrophotometer V 2.1.0 was employed to quantification the extracted DNA. In the next step, high-quality DNA was genotypically characterized through DArTseq™ technology (http://www.diversityarrays.com/dart-application-dartseq) at Genetic Analysis Service for Agriculture (SAGA) service unit at CIMMYT headquarters (Texcoco, Mexico), which resulted in the generation of 58,378 high-quality SNP markers. After carefully considering call rate (genotyping quality) and reproducibility (marker consistency over replicated assays), this number was trimmed to 9,427 best SNPs. 100 K-marker DArT-seq consensus map available at Diversity Arrays Technology Pty Ltd. (DArT) (http://www.diversityarrays.com/sequence-maps) was used to further assign chromosome numbers, orders, and genetic distances to these SNPs.

Statistical Analysis

Phenotypic data analysis was performed using META-R for an alpha lattice design [3],where individual raw values were transformed into BLUPs for subsequence analyses. Experiments were analyzed individually and in a combined analysis using a linear mixed model. Visualization (circular Manhattan plots (through "CM plot" package)), phenotypic box plots (through "ggplot2" package), correlation and genotype–phenotype network plots (through "qgraph" package) construction were carried out in R software version 4.0.2 software unless otherwise mentioned.

Genome-Wide Association Analysis (GWAS)

GWAS was accomplished using mean BLUP values (phenotypic data) of S1 and S2 for N0, N1, and N2 through freely available software, *TASSEL v5.2.43* [7]. TASSEL is implemented with two models, namely the general linear model (GLM) and the mixed linear model (MLM) for GWAS. We adopted the MLM option that requires either population structure (Q-matrix), which can be calculated from STRUCTURE, or principal component analysis (PCA) matrix (which can be generated in*TASSEL* itself, and kinship (K-matrix) matrix as covariates to avoid false positives. Based on our last report [11], we kept PCA = 5 for final analysis. For convenience, markers that gave a p-value < 0.001 (−log10 value of 3) for a given trait were claimed as significant associations.

Geno-pheno network for better understanding of the relationship between the measured traits and two most significantly associated SNPs per trait was constructed employing the "Spring" layout command [implemented in "qgraph" package (Epskamp et al. 2012)] to construct the correlation network. This command implements force-embedded algorithms (Di Battista et al. 1994) of a correlation matrix that iteratively computes the layout of a given graph in a two-dimensional space. A slightly modified version of the above-mentioned method (Fruchterman and Reingold 1991) is implemented in the "qgraph." This algorithm uses an iterative process to compute a layout in which the length of edges depends on the absolute weight of the edges. Furthermore, this algorithm calculates the minimum space required

to create the graph of closely or distantly related traits to facilitate the decision-making w.r.t closely related traits/parameters.

Results

Phenotypic Variation

Significant differences were observed across PBLs for different agronomic traits for all N treatments in both seasons. For all the traits studied except for SPKL, the season had no significant effect (Fig S3), thus highlighting that the differences observed were strictly due to N application. Among the seven recorded traits, PH and FTN increased at N1and N2 in comparison to N0 (Figure S3, Table S3), where PBLs 249 and 197 showed an increase as compared to PBLs 27 and 79,which showed either no response or negative response to N application for PH and FTN, respectively. Similarly, 70 PBLsshowed a positive response to N application in SPKL, whereas 206 PBLsshowed no or negative response to N application. Although GY increased significantly in N2 treatment in comparison to N0, 165 PBLs exhibited a decrease in GY from N0 to N2.Effect of N application on BY revealed a significant increase from N0 to N1 treatment in 193 PBLs, and from N1 to N2 treatment in 157 PBLs. HI decreased in 181 PBLs from N0 to N1 treatment and in 181 PBLs from N1 to N2 treatment. A direct comparison of HI between N0 and N2 treatment revealed a significant decrease in 171 advanced lines. TW didn't show any significant change in 156 advanced lines from N0 to N1 but decreased at N2. All the traits exhibited high heritabilities (>0.8) in both seasons [except in S1 for SPKL at N0 (0.54) and N1 (0.52), for PH at N1 (0.62),and in S2 for SPKL at N2 (0.77)] (Table S3).

Combined Correlations

Strong correlation ($r \geq 0.84$***) was observed under combined effect of season and nitrogen except for SPKL at N1 ($r=0.72$***) and N2 ($r=0.68$***) and HI at N0 ($r=0.70$***) and N1 ($r=0.68$***) (Table S4-a). Correlation analysis of mean data of the two seasons among various traits across three treatments indicated that PH was in positive correlation at different N doses ($r=0.26$***—0.52***) whereas only FTN of N0 and N2 were in strong positive correlation ($r=0.62$***) (Fig. 1, Table S4-b). SPKL was not correlated at any dose of N. Strong positve correlation between GY at N1 and N2 was detected ($r=0.61$***). BY at N0 ($r=0.23$***) and at N1 and N2 ($r=0.51$***) was also postively correlated. A positive correlation ($r=0.28$***) was also evident for HI at N1 and N2, whereas, a meager positve correlation ($r=0.13$*) was detected for TW at N0 and N1. Across traits, a positive correlation of GY with PH, BY and HI at N0 ($r=0.12$*, 0.55** and 0.55**, respectively), N1 ($r=0.30$***, 0.62*** and 0.54***, respectively) and N2 ($r=0.23$**, 0.71*** and 0.40***) was evident (Fig. 1, Table S2(b)). There was, however, slight but significant negative correlation of BY with HI at N0 ($r=-0.34$**), N1 ($r=-0.26$**) and N2 ($r=-0.27$**). PH and BY were in positive correlation at N1 ($r=0.28$***) and N2 ($r=0.16$**). Slight positive correlation was also evident between PH and SPKLN at N0 ($r=0.16$**) and N1 ($r=0.28$***), but not at N2.

Keeping in view the very strong correlations among various traits across the seasons (Table S4 (a)), mean data of the two seasons are used for further analysis (Fig. 2). The mean trend for all traits was the same as for individual seasons. Mean PH (cm) at N0 was lowest (76.7 ± 0.5) as compared to N1 (89.3 ± 0.5) and N2 (89.4 ± 0.4), with H^2 being 0.95, 0.87, and 0.93, respectively. Likewise, FTN (number of productive tillers arising from the ground per meter square)was 32.3 ± 0.5 ($H^2=0.95$) at N0 that increased to 36.7 ± 0.5 ($H^2=0.95$) and 36.2 ± 0.6 ($H^2=0.97$) at N1 and N2, respectively. On the other hand, SPKL (average of spikelet ount of five random spikes) decreased from 17.9 ± 0.1 ($H^2=0.84$) at N0 to 16.5 ± 0.1 ($H^2=0.79$) and 17.2 ± 0.11 ($H^2=0.76$). An increase in GY (grams) was witnessed from 1435.5 ± 34.0 ($H^2=0.92$) at N0 to 1584.7 ± 39.9 ($H^2=0.93$) at N1 and 1524.0 ± 41.7 ($H^2=0.92$). Similarly, BY increased from 4919.2 ± 92.9 ($H^2=0.94$) at N0 to 6020.6 ± 121.1 ($H^2=0.99$) and 6383.3 ± 0.96 ($H^2=0.96$) at N2. HI, however, witnessed a decreasing trend where it decreased from 0.3 ± 0.0 ($H^2=0.82$) at N0 to 0.27 ± 0.0 ($H^2=0.81$) at N1 and 0.25 ± 0.0 ($H^2=0.94$) at N2. TW(grams) remained similar at N0 and N1 (38.9 ± 0.3) but decreased at N2 (36.8 ± 0.3) where $H^2$was 0.96, 0.94, and 0.93 at N0, N1, and N2, respectively.

Association Mapping

To verify the existing population structure in the PBLs, principal component analysis was carried out in TASSEL with 9427 SNPs distributed across 21 wheat chromosomes with minor allele frequency (MAF) of >0.05 and the missing call rates <0.2. The first four principal components were the most informative, and after that, there was a slow decline until the 10^{th} PC component. Also,our experience with PBLs has shown the existence of ~4 sub-populations [1] in the germplasm when considering either three, four, or five subpopulations in the germplasm. Hence, association analysis was conducted using a mixed linear model (MLM) taken into account PCA = 5, which was computed in *TASSEL* where mean data of S1 and S2 were used to detect marker-trait associations (MTAs) between the phenotype and genotype.

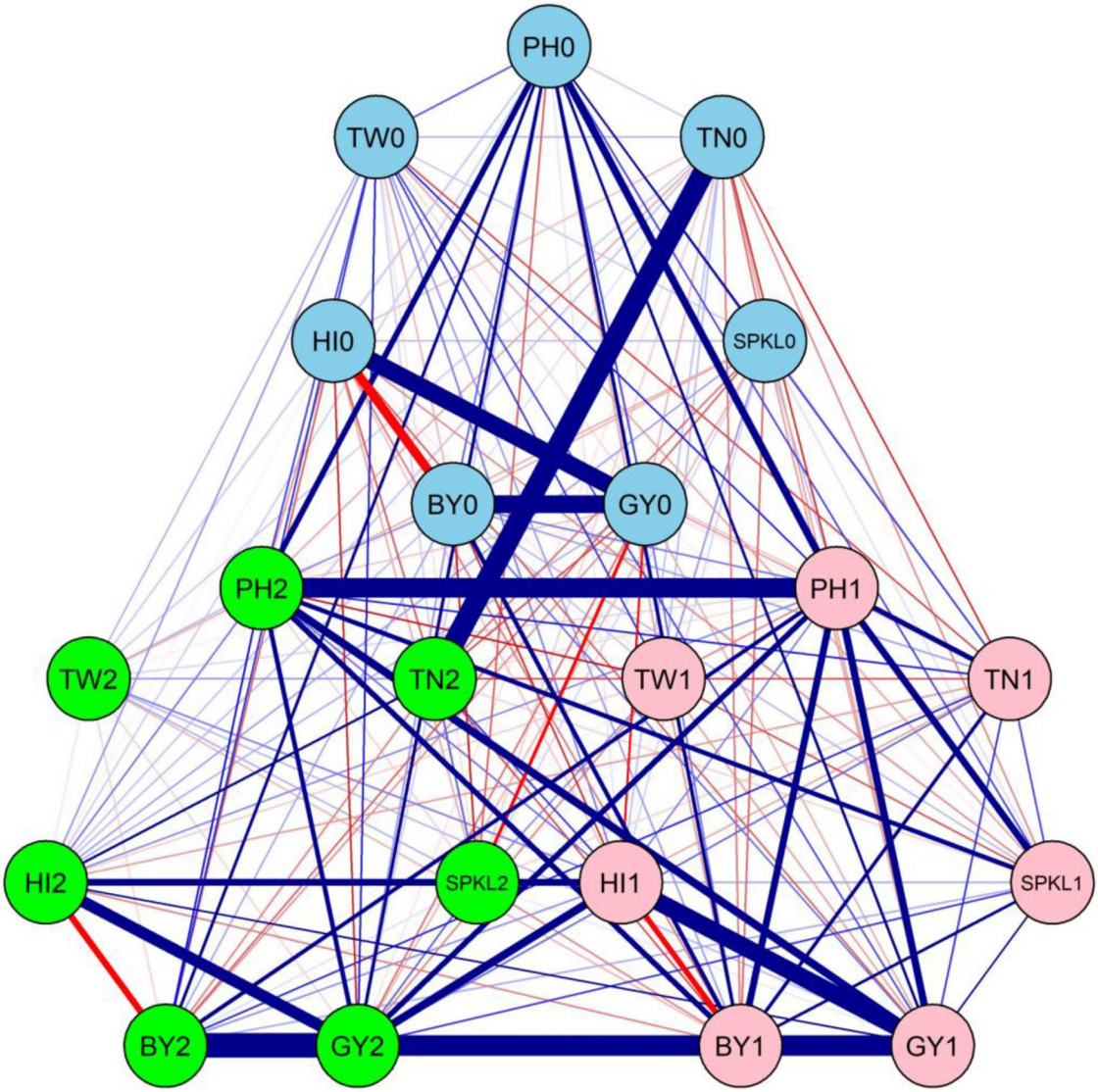


Fig. 1 Correlation network among plant height (PH), tiller number (FTN), spikelet number per spike (SPKL), grain yield (GY), biological yield (BY), harvest index (HI) and thousand kernel weight (TW) at N0 (skyblue: 0 nitrogen), N1 (pink: 60 kg nitrogen) and N2 (green:120 kg nitrogen) of mean data of two seasons in the germplasm. Blue and red indicate positive and negative correlation, respectively, whereas thickness of the line is proportional to the strenght of correlation

The quantil-quantile (Q-Q) plots (Fig S4 (a-g)) of the traits measured provided a glimpse that the expected –log10(p) value was close to the observed distribution that proved that the GWAS through MLM (PC + K) was acceptable to search for markers linked with traits of interest in our germplasm.

At p-value 0.001, a total of 182 marker-trait associations (MTA) were identified for all traits at all treatments, which were distributed on all the 21 chromosomes of the wheat genome (Fig. 3, Table S5). Among them, 24 MTAs explained > 8% phenotypic variation (Table 1). Total 40 MTAs were specific to the exotic gene pool, whereas 142 MTAs could be from elite and exotics. In terms of chromosomes, the highest number of MTAs were detected on chromosome 2B (25 MTAs), followed by chromosome 2A (24 MTAs) and1B (19 MTAs). Chromosomes 1A and 3B carried 13 MTAs each. Chromosome 5A carried 10 MTAs, which was followed by chromosome 3A with 9 MTAs. Chromosomes 2D, 5B, and 6A carried 8 MTAs, and 4A, 7A, and 7B carried 7 MTAs individually. Chromosomes 7D, 6B, and 4B carried 6, 5, and 4 MTAs, respectively, and 5D carried 3 MTAs. Finally, 2 MTAs were detected on chromosomes 1D and 4D, each, whereas 3D and 6D carried single MTA each. Among groups, group 2 chromosomes 57 MTAs followed by group 1 with 34

	PH_N0	TN_N0	SPKL_N0	GY_N0	BY_N0	HI_N0	TW_N0	PH_N1	TN_N1	SPKL_N1	GY_N1	BY_N1	HI_N1	TW_N1	PH_N2	TN_N2	SPKL_N2	GY_N2	BY_N2	HI_N2	TW_N2
Min.	53.5	17.0	13.8	333.6	1322.7	0.1	23.1	66.8	16.6	13.3	347.1	1676.4	0.03	17.5	69.4	14.5	12.6	163.1	973.8	0.04	18.9
Max.	102.4	57.3	21.3	3253.1	8585.4	0.6	58	104.7	62.1	20.6	4043	12919.8	0.7	53.3	108.8	62.5	22.04	4151.5	12871.4	0.7	52.2
Mean ± SE	76.7 ± 0.5	32.3 ± 0.5	17.9 ± 0.1	1435.5 ± 34.0	4919.2 ± 92.9	0.3 ± 0	38.9 ± 0.3	89.3 ± 0.3	36.7 ± 0.5	16.5 ± 0.1	1584.7 ± 39.9	6020.6 ± 121.1	0.27 ± 0	38.9 ± 0.3	89.4 ± 0.4	36.2 ± 0.6	17.2 ± 0.1	1524.0 ± 41.7	6383.3 ± 141.9	0.25 ± 0.0	36.8 ± 0.34
H^2	0.95	0.95	0.84	0.92	0.94	0.82	0.96	0.87	0.95	0.79	0.93	0.99	0.81	0.94	0.93	0.97	0.76	0.92	0.96	0.94	0.93

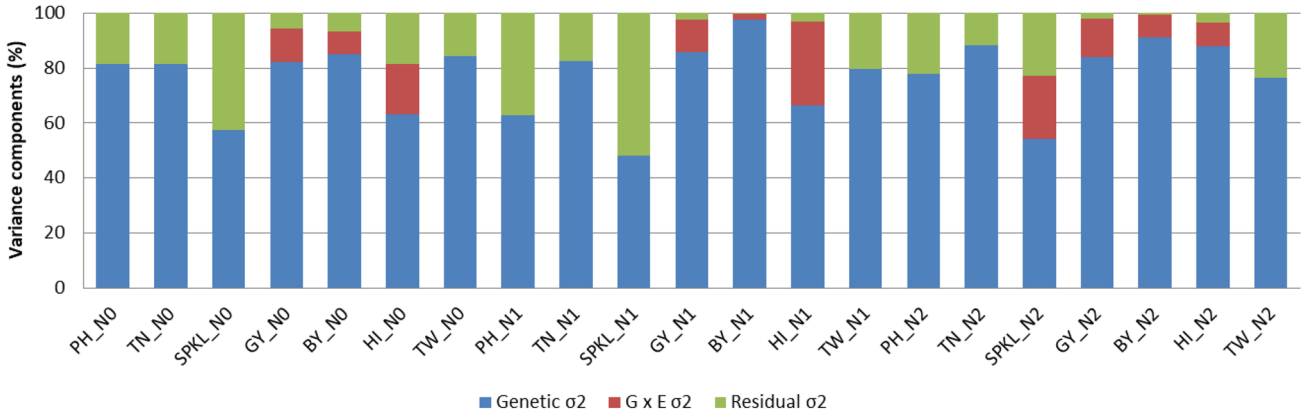


Fig. 2 Minimum (Min.), maximum (Max.), mean ± standard error (SE), heritability (H2) and variance component analysis of mean phenotypic data. PH- plant height, FTN – number of tillers, SPKLN – number of spikelets per spike, GY – grain yield, BY – biological yield, HI – harvest index, TW – thousand kernel weight, N0 – no nitrogen, N1 – nitrogen @ 60 kg/ha and N2 – nitrogen @ 120 kg/ha

MTAs followed by 3D with 23 MTAs. Group 5 chromosomes carried 21 MTAs, whereas group 7 chromosomes carried 20 MTAs. Finally, group 6 chromosomes carried 14 MTAs, and group 4 carried the least number of 13 MTAs. In terms of the genome, the highest number of MTAs were carried by B-genome (81 MTAs) followed by A-genome (78 MTAs) and D-genome (23 MTAs). Trait wise dissection of MTAs is provided below:

Plant Height (PH)

At p-value < 0.001, a total of 19 MTAs were detected for PH (Fig. 3a), where four associations were detected at N0 on chromosomes 3A (3 MTAs) and 5D (Table S5). Out of these 19 MTAs, 4 were specific to the exotic gene pool. A total of 10 MTAs were detected on chromosomes 2A (5 MTAs), 2B, 4B, 6A, 7B, and 7D at N1. The five different chromosomes (1B, 3B, 4A, 5B, and 6B) carried five MTAs at N2. Among all the MTAs observed, the PH, MTA observed with SNP on chromosome 5D explained the highest phenotypic variance. Similarly, for N1 and N2 treatment, the highest phenotypic variance was explained by MTAs observed on chromosomes 2B and 3B, respectively. For the MTAs from the exotic gene pool only, the highest phenotypic variance was explained by chromosomes 3A, 2B, and 6B for N0, N1, and N2 treatment, respectively.

FertileTiller Number (FTN)

For FTN, we detected 24 MTAs distributed on eight wheat chromosomes where 7, 12,and 5 MTAs were detected at N0, N1, and N2 treatments, respectively (Fig. 3b, (Table S5). Out of these 24 MTAs, 2 were specific to the exotic gene pool. At N0, the chromosomes involved were 2B, 3B (3 MTAs), 4A, and 5B (2 MTAs). Likewise, the chromosomes involved at N1 were 2A (7 MTAs), 3B, 5B, 6A, and 7A. At N2, chromosomes 2B, 6A, 7A, and 7B (2 MTAs) were involved. In addition, one SNP (4,990,734|F|0–18:A > G-18:A > G) on chromosome 2B at 67.73 cM was significant at both N0 and N2. Out of all the MTAs observed for FTN, MTAs from 5B explained the highest phenotypic variance for N0 treatment, whereas MTAs on chromosome 2A and 7B explained the highest phenotypic variance observed for N1 and N2 treatment, respectively. For MTAs from the exotic gene pool only, the highest phenotypic variance was explained by chromosomes 5B and 2A for N0, and N1 treatment, respectively.

Spikelet Number (SPKL)

With regards to SPKL, taking into consideration all three treatments, a total of 19 MTAs were detected on 11 different chromosomes with 10, 3, and 6 MTAs detected at N0, N1, and N2 treatments (Fig. 3c, Table S5). Out of the 19 SNPs explaining the significant phenotypic variance, 3 belonged to the exotic gene pool. The 10 MTAs detected at N0 resided

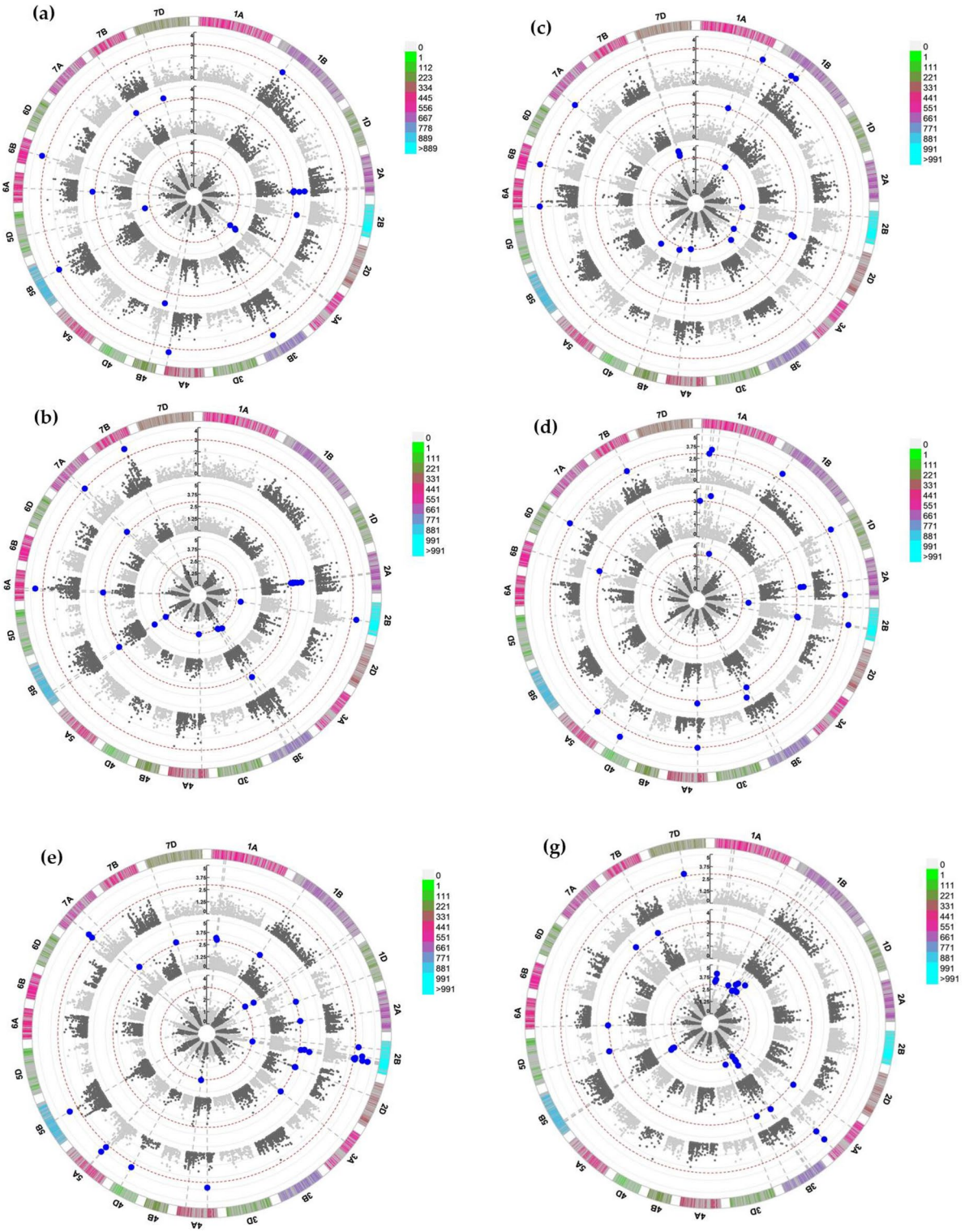


Fig. 3 Genome-wide scan of **a** plant height (PH), **b** Fertile tiller number (FTN), **c** number of spikelets per spike (SPKL), **d** grain yield (GY), **e** biological yield (BY), **f** harvest index (HI) and **g** thousand kernel weight (TW) at 0 N application (N0: inner most circle), 60 kg N application (N1: middle circle) and 120 kg N application (N2: third circle) in the form of circular Manhattan plots where the chromosomes are plotted at the outermost circle where thin dotted red line indicates significance level at p-value < 0.001 ($-\log10 = 3$ or more) beyond which an association is counted as true association (highlighted blue dots). Scale between chromosome 7D and 1A indicates LOD threshold. Colored boxes outside on the top right side indicate SNP density across the genome where green to aqua indicates less dense to dense (for reference see Table S5)

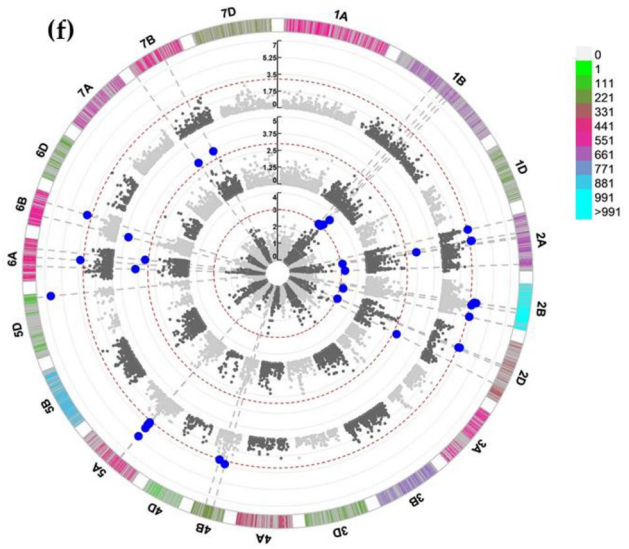


Fig. 3 (continued)

on chromosomes 1B, 2B, 3A (2 MTAs), 4A, 4B, 5A, and 7D, and the three MTAs detected at N1 were located on chromosomes 1A and 2D (2 MTAs), whereas the six MTAs detected at N2 were distributed on five different chromosomes (1A, 2B (2 MTAs), 6A, 6B and 7A). Among the 19 MTAs observed for SPKLN, MTA observed on chromosome 5A explained the highest phenotypic variance for N0 treatment. Similarly, for N1 and N2 treatment, MTAs from chromosome 2D and 1A explained the highest phenotypic variance, respectively. For MTAs from the exotic gene pool only, the highest phenotypic variance was explained by chromosomes 7D and 2A for N0, and N1 treatment, respectively.

Grain Yield (GY)

Considering all three levels of N, twelve different chromosomes carried 24 GY MTAs where two MTAs were associated at N0 and eleven each was associated at N1 and N2 (Fig. 3d). The N0 MTAs resided on chromosomes 1A and 2B. The N1 MTAs resided on chromosomes 1A (2 MTAs), 2A (3 MTAs), 2B (2 MTAs), 3B (2 MTAs), 4A, and 6B. In addition, the MTAs at N2 were exhibited on chromosomes 1A (2 MTAs), 1B, 1D, 2A, 2B, 4A, 4D, 5A, 6D, and 7B (Table S3). For the treatments N0, N1, and N2, the MTA from chromosomes 1A, 3B, and 4A explained the highest phenotypic variance.Similarly, out of the three MTAs specific to the exotic gene pool, 1D specific MTA explained the highest phenotypic variance (~5.5%).

Table 1 Selected significant marker-trait associations explaining significant phenotypic variance (> 8%) for plant height (PH), Fertile tiller number (FTN), grain yield (GY), biological yield (BY), harvest index (HI) and thousand kernel weight (TW) at 0 N application (_N0), 60 kg N application (_N1) and 120 kg N application (_N2) detected through linear model (MLM)

Trait	SNP	Chr	Pos	F	*p*-value	R^2
PH_N0	*1,126,619|F|0–21:A > T-21:A > T*	5D	69.42	13.29	3.87E-04	0.085
PH_N1	*1,107,392|F|0–15:G > A-15:G > A*	2B	78.31	13.43	3.54E-04	0.088
FTN_N0	*7,353,380|F|0–68:A > G-68:A > G*	4A	22.1	11.9	8.54E-04	0.141
FTN_N0	*2,279,108|F|0–12:T > C-12:T > C*	5B	30.8	17.73	3.81E-05	0.0886
FTN_N2	*1,049,818|F|0–60:G > A-60:G > A*	7B	94.01	14.39	2.12E-04	0.0845
GY_N1	*1,031,417|F|0–24:G > A-24:G > A*	4A	32.57	12.84	5.30E-04	0.105
GY_N1	*9,724,050|F|0–29:G > A-29:G > A*	3B	124.26	14.84	1.67E-04	0.083
GY_N2	*1,031,417|F|0–24:G > A-24:G > A*	4A	32.57	11.81	8.75E-04	0.102
BY_N0	*1,107,071|F|0–54:T > C-54:T > C*	2B	78.31	11.39	9.49E-04	0.085
BY_N2	*1,031,417|F|0–24:G > A-24:G > A*	4A	32.57	14.22	2.81E-04	0.122
BY_N2	*1,078,426|F|0–29:A > C-29:A > C*	2B	78.41	15.59	1.17E-04	0.087
BY_N2	*1,215,338|F|0–26:G > T-26:G > T*	2B	78.57	18.01	3.26E-05	0.081
HI_N0	*4,991,461|F|0–41:A > G-41:A > G*	1B	157.86	14.3	2.32E-04	0.108
HI_N1	*1,206,985|F|0–7:A > G-7:A > G*	6A	79.01	12.74	5.69E-04	0.119
HI_N1	*5,412,023|F|0–12:C > G-12:C > G*	6A	36.15	15.52	1.20E-04	0.084
HI_N1	*7,344,586|F|0–20:C > T-20:C > T*	6B	27.33	19.26	1.75E-05	0.081
HI_N2	*2,250,930|F|0–23:A > G-23:A > G*	5D	144.84	28.75	3.38E-07	0.197
HI_N2	*988,747|F|0–7:A > G-7:A > G*	5A	57.82	20.64	8.78E-06	0.083
HI_N2	*4,911,150|F|0–25:C > T-25:C > T*	2B	63.25	14.3	2.24E-04	0.082
TW_N0	*1,078,640|F|0–10:T > C-10:T > C*	5B	31.96	15.2	1.52E-04	0.106
TW_N0	*1,011,620|F|0–63:C > A-63:C > A*	1A	66.22	16.55	7.01E-05	0.085
TW_N1	*1,193,932|F|0–14:C > A-14:C > A*	3B	39.13	12.66	5.33E-04	0.090
TW_N2	*1,046,576|F|0–14:A > G-14:A > G*	3A	125.2	17.51	4.37E-05	0.088
TW_N2	*1,046,522|F|0–48:T > C-48:T > C*	7D	101.77	12.43	5.73E-04	0.081

Biological Yield (BY)

The 32 MTAs of BY resided on 13 different chromosomes, where the most frequent was chromosome 3B,which carried 12 MTAs (Fig. 3e). Among the 32 MTAs, four MTAs were detected at N0 on chromosomes 1B (2 MTAs), 2B, and 4A. The 13 MTAs at N1 were carried by chromosomes 1A (2 MTAs), 1B, 1D, 2A, 2B (3 MTAs), 3A, 7A, and 7D. The chromosomes 2B (8 MTAs), 4A, 4D, 5A (2 MTAs), 5B, and 7A (2 MTAs) harbored the 17 MTAs at N2 (Table S3). Interestingly, SNP 1,078,426|F|0–29:A > C-29:A > C on chromosome 2B at 78.41 cM was significant at both N1 and N2. The MTAs on chromosomes 2B, 7D, and 4A explained the highest phenotypic variance for treatments N0, N1, and N2, respectively. Similarly, three MTAs specific to the exotic gene pool, with MAT from the chromosome 5A, explained the highest phenotypic variance (6.5%).

Harvest Index (HI)

The highest number of MTAs for an individual trait was 35, detected for HI that involved 10 chromosomes (Fig. 3 (f)). For N0, the eight associations were located on chromosomes 2A (4 MTAs), 2A (2 MTAs), 2B, and 2D. The seven MTAs at N1 were located on chromosomes 2A, 2D, 6A (2 MTAs), 6B, and 7B (2 MTAs). The 19 MTAs at N2 treatment resided on chromosomes 2A (3 MTAs), 2B (4 MTAs), 2D (2 MTAs), 4B (2 MTAs), 5A (6 MTAs), 5D, 6A and 6B (Table S3). The MTAs on chromosomes 1B, 6A, and 5D explained the highest phenotypic variance for treatment N0, N1, and N2. Similarly, out of the 15 MTAs specific to the exotic gene pool, MTA on chromosome 6A explained the highest phenotypic variance (~ 11%).

Thousand Kernel Weight (TW)

The 29 TW MTAs were carried by 10 different chromosomes, where the most frequent was chromosome 3B with 6 MTAs (Fig. 3g). The 19 MTAs at N0 were located on chromosomes 1A (4 MTAs), 1B (7 MTAs), 3B (4 MTAs), 3D, and 5B (1B (3 MTAs). Six different chromosomes (3A, 3B, 5D, 6A, 7A, and 7B) carried the seven MTAs detected at N1. Finally, chromosomes 3A (2 MTAs) and 7D were involved at N2 with 3 MTAs (Table S3). The MTAs on chromosomes 5B, 3B, and 3A explained the highest phenotypic variance for treatment N0, N1, and N2, respectively. Also, for the MTAs specific to the exotic gene pool, the highest phenotypic variance was explained by chromosomes 3B and 3A for N0, N1, and N2 treatment, respectively.

Haplotype Analysis

Haplotype block analysis of the LTP revealed a total of 51 haplotype blocks (N0: 9, N1:20, N2:22) significantly associated with haplotype blocks. The most significant haplotype blocks were identified for TW, SPKL, BY, FTN.

FTN under N0 treatment and for GY, BY, HI, SPKL, and PH under N1 and N2 treatments, both except SPKL under N1 treatment only (Table S5). Among these, the haplotype blocks HB2B-33 and HB2B-22 on chromosome 2B and HB5A-16 on chromosome5A associated with TW, SPKL, and BY, respectively, had shown significant association across seasons under N0 treatment (Table S4). Different haplotypes reported to be significantly associated with the same trait such as HB1A-9, HB1B-11, HB2B-28, and HB6B-1 associated with BY under N1 treatment; HB2B-23, HB2B-52, HB4A-9, HB7A-1 associated with HI and HB2B-22, HB3B-3 associated with PH under N2 treatment in both the seasons (Table S5). In contrast, the same haplotype showed significant association with different traits under different treatments across seasons. The haplotype block HB2B-22 on chromosome 2B showed association with SPKL under N0 treatment and with BY, PH under N2 treatment; HB1A-9 on chromosome 1A with BY under N1 treatment (Table S5). The association of haplotype block HB2B-28 on chromosome 2B was conserved across seasons and treatments (Table S5).

Discussion

Crop improvement to increase global production by developing high-yielding varieties has inadvertently relied on their ability to better N utilization efficiency. The advent of semi-dwarf varieties resulted in the more efficient utilization of N-based fertilizer compared to tall varieties, as they were lodging resistant [25]. Over the past few decades, a volume of work has been performed to develop more efficient crop plants that result in an equal or higher yield per unit of the inputs supplied. The development of input-efficient crop plants is dependent on the presence of variation. But in crop plants such as wheat, the efforts are constrained by the limited variation in cultivated germplasm, whereas significant variation for various biotic, abiotic, and other value-added traits has been reported in landraces and wild germplasm of wheat. Thus, with the objective for effective utilization of natural variation in wheat improvement, a set of 2967 PBLs were initially developed by the Seed for discovery program of CIMMYT

that carried ~ 16–40% exotic alleles [34]. In this study, to identify the PBLs with improved pNUE efficiency that is explained by the combination of N uptake and NUE efficiency [23], we focused on characterizing the phenotypic variation present in a subset of initial PBLs for N uptake by focusing on variation present in different agronomic traits as markers. We performed a GWAS analysis to identify the underlying genetic regions that can be potentially converted to breeder-friendly DNA-based markers.

Phenotypic Variation

Evaluation of 276 PBLs for two seasons at three N level treatments highlighted that for most of the agronomic traits except SPKLN, no significant differences were observed across seasons. A significant positive effect of N treatment was observed for agronomic traits such as PH, FTN, BY, GY, HI, whereas a low positive correlation or negative effect was observed for SPKLN and TW. The negative effect observed for SPKLN due to N treatment could be due to the relative availability differences of N and phosphorus (P) as they are essential for vegetative and reproductive growth, respectively [29]. The genotypic differences were observed within PBLs for different agronomic traits for N treatment. Within different PBLs, PH and FTN can be effectively used as positive phenotypic markers for pNUE. The effect of N treatment on GY indicated that higher N treatment doesn't invariably suggest a higher yield, as previously reported [18, 30]. Although statistically significant higher GY treatment was observed for N2 compared to N0, 165 PBLs showed a decrease in GY at N2 compared to N0, suggesting that extra supplied N doesn't directly correlate with GY. Thus, extra N is either unavailable to plant or is being lost due to denitrification, leaching, and volatilization or is being used for other plant developmental processes such as biomass. The contrast for GY in PBLs could be due to the exotic alleles that could be regulating partitioning of the nutrients from vegetative tissues to grains, as BY showed a positive increase across both N treatments [16].

Marker Trait Associations

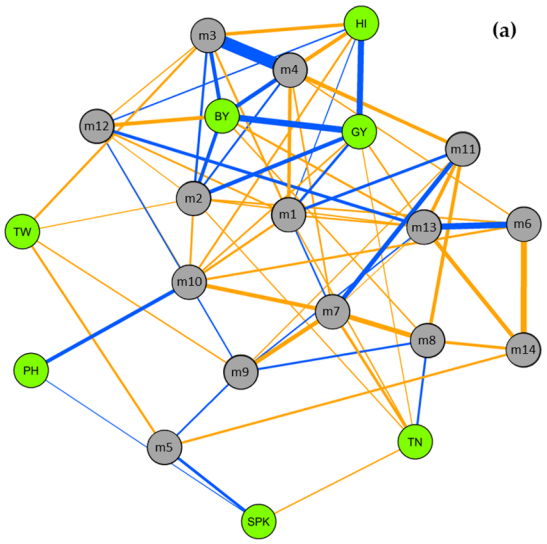
The development of reliable markers for pNUE through breeding and next-generation technologies is critical to developing nutrient efficient cultivars as traditional phenotypic evaluation of segregating populations for NUE is an expensive and time-consuming process. Mapping for NUE using different approaches identified genetic loci on all chromosomes of wheat contributing towards pNUE, highlighting the complex nature of the trait [10, 20, 37]. A total of 13 chromosomes carried the 19 unique MTAs linked to PH at all three N applications where none was common at any other trait that is not surprising because as many as 17 chromosomes are known to carry PH QTLs [19], Zanke et al., 2015). In a similar study, PH was associated with 30 regions, including 14 SNP markers spread over 17 chromosomes [6], where twenty of these coincided with the regions involved in GY. Moreover, 54 MTAs on 15 of the 21 chromosomes have been reported recently [1, 6, 23], which were mainly concentrated on chromosomes 2B, 3D, and 7A. The MTA explaining ~ 9% phenotypic variance is an effective target that explained PH response to N application (PH_N1, *1,107,392|F|0–15:G > A-15:G > A,* 2B). In addition to PH, 24 MTAs linked to FTN, 19 MTAs for SPKLN, 23 MTAs for GY, 30 MTAs for BY, 35 MTAs for HI, and 29 MTAs for TW were discovered. Studies on the effect of N addition towards FTN do not exist, and we are reporting the identification of genetic loci regulating FTN for N treatment. Out of all MTAs observed for FTN.

FTN, chromosome 4A explained more than 14% of observed phenotypic variation. The highest phenotypic variance (~ 12.2% & ~ 10%) for N treatment for BY and GY was explained by an MTA on chromosome 4A (Table 1). Moreover, previous work has reported stable yield-related QTLs on chromosome 4A [32], and thus these MTAs can be converted to reliable breeder-friendly markers. The association analysis of the same set of PBLs revealed > 90% similar results [11]. In addition, for HI for all three treatments, three MTAs were observed that explained more than 10% phenotypic variance (Table 1). Three MTAs, explaining ~ 10%, 9%, and 8% phenotypic variance, for TW for N0, N1, and N2 treatment, were observed at chromosome 5B, 3B, and 3A, respectively.

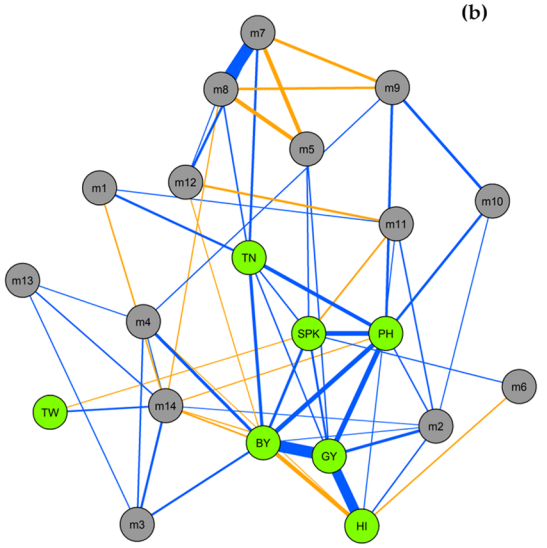
Geno-Pheno Network

We constructed a geno-pheno network for a better understanding of the relationship between the measured traits and the two most significantly associated SNPs per trait. Table 2The traits viz. GY, BY and HI was more closely related as compared to other traits at N0. Hence, selection for markers linked to either of the said traits will not influence other traits (Fig. 4a). At N1, however, except TW, all traits viz. GY, BY, HI, SPKLN, FTN, and PH were tightly knitted together (Fig. 4b). In fact, at N1, FTN was at the center, indicating its pivotal role in GY, which is also evident from correlation analysis [Table S4 (b)]. Table 3At N2, the four traits closer to each other was GY, BY, HI, and PH, indicating that at least some part of GY improvement is coming from increased height alongside BY and PH (Fig. 4c). Table 4We conclude that, like in *Arabidopsis,* where a complex network of additive and epistatic QTLs control resistance to *Ralstoniasolanacearum* under heat stress [4], final GY and related traits are entangled in a complex network of

N0				
Sr #	Trait	SNP	Chr	Pos (cM)
m1	GY	1864355|F|0-68:C>G-68:C>G	2B	1.26
m2	GY	1012827|F|0-63:C>G-63:C>G	1A	135.38
m3	BY	1161065|F|0-56:A>C-56:A>C	1B	286.64
m4	BY	994956|F|0-44:G>A-44:G>A	1B	270.55
m5	SPKL	101276566|F|0-19:T>G-19:T>G	7D	14.6
m6	SPKL	1094596|F|0-45:C>G-45:C>G	5A	60.39
m7	TN	2279108|F|0-12:T>C-12:T>C	5B	30.8
m8	TN	4990734|F|0-18:A>G-18:A>G	2B	67.73
m9	PH	4988938|F|0-8:G>C-8:G>C	3A	48.22
m10	PH	101422385|F|0-7:G>T-7:G>T	3A	43.6
m11	HI	4991461|F|0-41:A>G-41:A>G	1B	157.86
m12	HI	3955235|F|0-7:T>A-7:T>A	2A	122.68
m13	TW	3026224|F|0-25:G>T-25:G>T	1B	140.75
m14	TW	996888|F|0-37:T>G-37:T>G	3B	54.72



N1				
Sr #	Trait	SNP	Chr	Pos (cM)
m1	GY	9724050|F|0-29:G>A-29:G>A	3B	124.26
m2	GY	1071231|F|0-65:A>G-65:A>G	2A	68.56
m3	BY	101351962|F|0-8:G>C-8:G>C	2B	81.25
m4	BY	1124944|F|0-25:G>A-25:G>A	2B	73.56
m5	SPKL	1002792|F|0-48:T>C-48:T>C	2D	29.23
m6	SPKL	3022408|F|0-16:G>C-16:G>C	1A	179.08
m7	TN	3570146|F|0-9:A>G-9:A>G	2A	72.16
m8	TN	4909866|F|0-29:A>T-29:A>T	2A	71.38
m9	PH	5582507|F|0-13:C>G-13:C>G	2A	114.17
m10	PH	100812054|F|0-23:T>C-23:T>C	4B	28.23
m11	HI	7344586|F|0-20:C>T-20:C>T	6B	27.33
m12	HI	5412023|F|0-12:C>G-12:C>G	6A	36.15
m13	TW	1070651|F|0-19:T>C-19:T>C	7A	101.52
m14	TW	1193932|F|0-14:C>A-14:C>A	3B	39.13



N2				
Sr #	Trait	SNP	Chr	Pos (cM)
m1	GY	1667202|F|0-61:A>C-61:A>C	4D	66.81
m2	GY	101080470|F|0-7:G>A-7:G>A	2B	70.73
m3	BY	1215338|F|0-26:G>T-26:G>T	2B	78.57
m4	BY	1220092|F|0-6:G>A-6:G>A	5A	71.16
m5	SPKL	1206667|F|0-9:C>T-9:C>T	6B	32.16
m6	SPKL	3028745|F|0-26:G>C-26:G>C	1B	99.94
m7	TN	1128830|F|0-35:C>A-35:C>A	7B	93.44
m8	TN	1049818|F|0-60:G>A-60:G>A	7B	94.01
m9	PH	101242119|F|0-9:C>A-9:C>A	3B	95.48
m10	PH	101383024|F|0-20:C>T-20:C>T	4A	130.68
m11	HI	2250930|F|0-23:A>G-23:A>G	5D	144.84
m12	HI	988747|F|0-7:A>G-7:A>G	5A	57.82
m13	TW	1046576|F|0-14:A>G-14:A>G	3A	125.2
m14	TW	1046522|F|0-48:T>C-48:T>C	7D	101.77

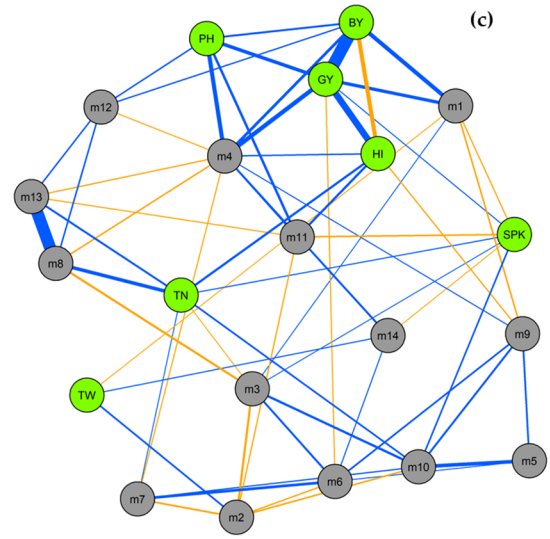


Fig. 4 Phenotype-genotype network at 0 (N0, **a**), 60 kg (N1, **b**) and 120 kg N (N2, **c**) application with two most strongly associated markers (based on p-values) per trait at each N application based on correlation statistics implemented in "qgraph". Blue indicates positive relation and orange indicates negative relation and thickness and closeness is proportional to the strength of the relation. < 10% relationships between markers and phenotypes are not shown

Table 2 Grain yield (GY—kg/ha) of best 20 PBLs at N0 treatment and two standard checks (PBW725 and HD3086)

GID	GY_N0	GY_N1	GY_N2
7,933,014	2427.7	1665.3	1629.6
7,933,035	2307.5	1666.1	1695.9
7,933,094	2547.6	1418.7	936.6
7,933,205	2742.8	4043	2378.9
7,933,503	3253.1	1432.9	828.1
7,933,799	2348.1	1994.7	1252.2
7,933,829	2562.1	1265.6	807.7
7,933,866	2672.3	430.9	622.9
7,933,881	2389.4	2077.7	530.2
7,933,975	2632.2	1756.9	2002.4
7,934,105	2621.3	1243.0	1289.2
7,934,240	2490	1582.5	1695.8
7,934,428	2683.1	1759.3	2066
7,934,479	2723.1	852.1	899.3
7,934,483	2750.1	1250.4	1602.7
7,934,531	2470.5	1254.5	924.4
7,934,841	2437.7	1262.6	1684.3
7,934,904	2753.9	764.6	76.69
7,935,072	2354.6	2007.2	836.81
7,935,102	2661.2	1655.2	1429.6
Population mean	1435.5	1584.7	1524.0
Check 1	1054	3313.6	1677.3
Check 2	1547	2749	1542

Table 3 Grain yield (GY—kg/ha) of best PBLs at 60 kg N application (N1) and two standard checks (PBW725 and HD3086)

GID	GY_N0	GY_N1	GY_N2
7,933,135	1214.1	2901.2	2352.5
7,933,150	1701.3	3066.7	2307.1
7,933,202	1864.3	2903.4	1513.7
7,933,205	2742.8	4043	2378.9
7,933,239	1443.6	2804.6	2100.1
7,933,359	1439	3728	4161.5
7,933,539	1869	3403.2	2752.0
7,933,683	1457.2	3148.5	3469.8
7,933,744	1351.6	3164.4	2968.8
7,933,864	691.4	2986.1	2777.1
7,933,897	1092.9	3477.7	2208.1
7,933,910	1185.4	2829.4	3376.8
7,933,945	1417.1	3065.2	2677.1
7,934,390	1387.5	3735.8	3935.5
7,934,846	850.5	3149.0	1622.4
7,935,221	1665.4	2906.5	2377.9
Population mean	1435.5	1584.7	1524.0
Check 1	1054	3313.6	1677.3
Check 2	1547	2749	1542

Table 4 Grain yield (GY—kg/ha) of best 20 PBLs at 120 kg N application (N2) and two standard checks (PBW725 and HD3086)

GID	GY_N0	GY_N1	GY_N2
7,933,008	1439	1423.2	2947.1
7,933,091	971.7	1666	2509.4
7,933,105	1335.7	2077	2796.8
7,933,318	1381.7	2659.4	3015.5
7,933,359	1439.0	3728	4161.5
7,933,539	1869.8	3403.2	2752
7,933,683	1457.2	3148.5	3469.7
7,933,744	1351.5	3164.4	2968.8
7,933,845	1821.3	2076.7	2818.2
7,933,864	691.42	2986.1	2777.1
7,933,894	1042.3	2478.4	4026.5
7,933,910	1185.4	2829.4	3376.8
7,933,939	632.48	1652.5	2568
7,933,945	1417.1	3065.2	2677.1
7,934,042	1256.6	1584.2	2761.6
7,934,229	880.37	1325.1	2634.2
7,934,242	1810.8	1655.0	2826.1
7,934,390	1387.5	3735.8	3935.5
7,934,683	758.20	2477.9	2538.4
7,934,760	1669	1120.8	2581.5
Population mean	1435.5	1584.7	1524.0
Check 1	1054	3313.6	1677.3
Check 2	1547	2749	1542

associations/QTLs at various N applications in wheat. Network of QTLs colocalizations for 28 traits determining NUE components [8] and spikelet sterility and yield-related traits [2] in wheat have also been reported recently. Therefore, each trait must be addressed in detail separately and in connection with related traits to improve GY at limited N levels.

Donor Lines for pNUE

The availability of ready-to-use donor lines for regional breeding programs is of critical importance. These PBLs have already been screened for different agronomic traits and biotic and abiotic stress tolerance [33–35] and thus have the potential to be released for a targeted environment. Based on their performance at N0, N1, and N2, we have selected the top 20 lines that performed better than the check and thus can be used as donors in different breeding programs and can be further used for understanding the mechanisms of improved N application response (Tables 2–4). In SE-Asia, particularly in India and Pakistan, the N-based fertilizer is the most frequently used fertilizer that costs every farmer ~ $30/ha. Thus, the development and release of plant varieties with improved pNUE can prevent environmental damage due to excessive application of N-based fertilizer

and can further reduce the operational cost for wheat production.

Conclusion

The development of nutrient efficient cultivars is reliant on the availability of the presence of natural genetic variation and simpler phenotypic approaches compatible with breeding pipelines. In this study, we have characterized the presence of genotypic variation in pre-breeding lines (PBLs) for pNUE known to carry ~ 25% exotic alleles. The identification of phenotypic traits for their responsiveness to the N application and underlying genetic loci provides an alternative approach to developing NUE crop varieties.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12033-022-00535-8.

Acknowledgements All authors duly acknowledge the 'Seeds of Discovery' project of CIMMYT. Authors from Punjab Agricultural University are grateful to Newton Fund project, "Indo-UK Centre for the improvement of nitrogen use efficiency in Wheat (INEW)" number # (BT/IN/UK-VNC/43/KV/2015-16) and Cambridge-India Network for Translational Research in Nitrogen (CINTRIN), number# BT/IN/UK-VNC/42/RG/2014-15; funded by Department of Biotechnology, Government of India. The direct and indirect support of other researchers and non-scientific staff is duly acknowledged.

Funding Department of biotechnology,ministry of science and technology,BT/IN/UK-VNC/43/KV/2015-16,Achla Sharma,BT/IN/UK-VNC/42/RG/2014-15,Parveen Chhuneja

Declarations

Conflict of interest Authors declare no conflict of interest.

References

1. Akram, S., Arif, M. A. R., & Hameed, A. (2020). A GBS-based GWAS analysis of adaptability and yield traits in bread wheat (*Triticum aestivum* L.). *Journal of Applied Genetics, 62*, 27–41.
2. Alqudah, A. M., Haile, J. K., Alomari, D. Z., Pozniak, C. J., Kobiljski, B., & Börner, A. (2020). Genome-wide and Snp network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. *Science and Reports, 10*, 1–12.
3. Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., Pérez-Rodríguez, P., & Lopez-Cruz, M. A. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *The Crop Journal, 8*, 745–756.
4. Aoun, N., Desaint, H., Boyrie, L., Bonhomme, M., Deslandes, L., Berthomé, R., & Roux, F. (2020). A complex network of additive and epistatic quantitative trait loci underlies natural variation of *Arabidopsis thaliana* quantitative disease resistance to *Ralstonia solanacearum* under heat stress. *Molecular Plant Pathology, 21*, 1405–1420.
5. Barraclough, P. B., Lopez-Bellido, R., & Hawkesford, M. J. (2014). Genotypic variation in the uptake, partitioning and remobilisation of nitrogen during grain-filling in wheat. *Field Crops Research, 156*, 242–248.
6. Börner, A., Schumann, E., Fürste, A., Cöster, H., Leithold, B., Röder, M., & Weber, W. (2002). Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *TAG. Theoretical and Applied Genetics, 105*, 921–936.
7. Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics, 23*, 2633–2635.
8. Cormier, F., Le Gouis, J., Dubreuil, P., Lafarge, S., & Praud, S. (2014). A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). *TAG. Theoretical and Applied Genetics, 127*, 2679–2693.
9. Cox, T., & Hatchett, J. (1994). Hessian fly-resistance gene H26 transferred from *Triticum tauschii* to common wheat. *Crop Science, 34*, 958–960.
10. Cui, F., Fan, X., Zhao, C., Zhang, W., Chen, M., Ji, J., & Li, J. (2014). A novel genetic map of wheat: Utility for mapping QTL for yield under different nitrogen treatments. *BMC Genetics, 15*, 57.
11. Dababat, A., Arif, M. A. R., Toktay, H., Atiya, O., Shokat, S., Gul, E., Imren, M., & Singh, S. A. (2021). A GWAS to identify the cereal cyst nematode (*Heterodera filipjevi*) resistance loci in diverse wheatprebreeding lines. *Journal of Applied Genetics, 62*, 93–98.
12. Dwivedi, S. L., Scheben, A., Edwards, D., Spillane, C., & Ortiz, R. (2017). Assessing and exploiting functional diversity in germplasm pools to enhance abiotic stress adaptation and yield in cereals and food legumes. *Frontiers in Plant Science, 8*, 1461.
13. Fang, Z., Xia, K., Yang, X., Grotemeyer, M. S., Meier, S., Rentsch, D., Xu, X., & Zhang, M. (2013). Altered expression of the PTR/NRT 1 homologue Os PTR 9 affects nitrogen utilization efficiency, growth and grain yield in rice. *Plant Biotechnology Journal, 11*, 446–458.
14. Guttieri, M. J., Frels, K., Regassa, T., Waters, B. M., & Baenziger, P. S. (2017). Variation for nitrogen use efficiency traits in current and historical great plains hard winter wheat. *Euphytica, 213*, 87.
15. Habash, D., Massiah, A., Rong, H., Wallsgrove, R., & Leigh, R. (2001). The role of cytosolic glutamine synthetase in wheat. *The Annals of Applied Biology, 138*, 83–89.
16. Ram, H., Singh, A., Katoch, M., Kaur, R., Sardar, S., Palia, S., Satyam, R., Sonah, H., & Deshmukh, R. (2021). Ajay Kumar Pandey, Ishaan Gupta, Tilak Raj Sharma, Dissecting the nutrient partitioning mechanism in rice grain using spatially resolved gene expression profiling. *Journal of Experimental Botany, 72*(6), 2212–2230. https://doi.org/10.1093/jxb/eraa536
17. Islam, S., Zhang, J., Zhao, Y., She, M., & Ma, W. (2020). Genetic regulation of the traits contributing to wheat nitrogen use efficiency. *Plant Science, 303*, 110759.
18. Kostić, M. M., Tagarakis, A. C., Ljubičić, N., Blagojević, D., Radulović, M., Ivošević, B., & Rakić, D. (2021). The effect of N fertilizer application timing on wheat yield on chernozem soil. *Agronomy, 11*, 1413. https://doi.org/10.3390/agronomy11071413
19. Lopes, M., et al. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *Journal of Experimental Botany, 66*, 3477–3486.
20. Mahjourimajd, S., Taylor, J., Sznajder, B., Timmins, A., Shahinnia, F., Rengel, Z., Khabaz-Saberi, H., Kuchel, H., Okamoto, M., & Langridge, P. (2016). Genetic basis for variation in wheat grain yield in response to varying nitrogen application. *PLoS ONE, 11*, e0159374.
21. Mehrabi, A. A., Pour-Aboughadareh, A., Mansouri, S., & Hosseini, A. (2020). Genome-wide association analysis of root system

architecture features and agronomic traits in durum wheat. *Molecular Breeding, 40*, 1–17.
22. Mohammadi, R., Sadeghzadeh, B., Ahmadi, H., Bahrami, N., & Amri, A. (2015). Field evaluation of durum wheat landraces for prevailing abiotic and biotic stresses in highland rainfed regions of Iran. *The Crop Journal, 3*, 423–433.
23. Moll, R., Kamprath, E., & Jackson, W. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization 1. *Agronomy Journal, 74*, 562–564.
24. Omara, P., Aula, L., Oyebiyi, F., & Raun, W. R. (2019). World cereal nitrogen use efficiency trends: Review and current knowledge. *Agrosystems Geosciences & Environment, 2*, 180045.
25. Pingali, P. L. (2012). Green revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences, 109*, 12302–12308.
26. Pixley, K. V., Salinas-Garcia, G. E., Hall, A., Kropff, M., Ortiz, C., Bouvet, L., Suhalia, A., Vikram, P., & Singh, S. (2018). CIMMYT's Seeds of Discovery Initiative: Harnessing biodiversity for food security and sustainable development. *Indian Journal of Plant Genetic Resources, 31*, 1–10.
27. Ramadas, S., Kumar, T., & Singh, G. P. (2020). *Wheat production in India: Trends and prospects*. IntechOpen: In Recent Advances in Grain Crops Research.
28. Ranjan, R., Yadav, R., Kumar, A., & Mandal, S. N. (2019). Contributing traits for nitrogen use efficiency in selected wheat genotypes and corollary between screening methodologies. *Acta Agriculturae Scandinavica Section B Soil & Plant Science, 69*, 588–595.
29. Rodríguez, D., Andrade, F., & Goudriaan, J. (1999). Effects of phosphorus nutrition on tiller emergence in wheat. *Plant and Soil, 209*, 283–295.
30. Sandhu, N., Kaur, A., Sethi, M., Kaur, S., Singh, V., Sharma, A., Bentley, A. R., Barsby, T., & Chhuneja, P. (2021). Genetic dissection uncovers genome wide marker-trait associations for plant growth, yield and yield related traits under varying nitrogen levels in nested synthetic wheat introgression libraries. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2021.738710
31. Shiferaw, B., Smale, M., Braun, H. J., Duveiller, E., Reynolds, M., & Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security, 5*, 291–317.
32. Shokat, S., Sehgal, D., Vikram, P., Liu, F., & Singh, S. (2020). Molecular markers associated with agrophysiologicaltraits under terminal drought conditions in bread Wheat. *International Journal of Molecular Sciences, 21*, 3156.
33. Singh, S., Ledesma-Ramírez, L., Solís-Moya, E., Iturriaga, G., Sehgal, D., Reyes-Valdes, M. H., Montero-Tavera, V., Sansaloni, C. P., Burgueño, J., & Ortiz, C. (2019). GWAS to identify genetic loci for resistance to yellow rust in wheat pre-breeding lines derived from diverse exotic crosses. *Frontiers in Plant Science, 10*, 1390.
34. Singh, S., et al. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Science and Reports, 8*, 12527.
35. Singh, S., Sehgal, D., Kumar, S., Rehman-Arif, M. A., Vikram, P., Sansaloni, C., Fuentes-Dávila, G., & Ortiz, C. (2020). GWAS revealed a novel resistance locus on chromosome 4D for the quarantine disease Karnal bunt in diverse wheat pre-breeding germplasm. *Science and Reports, 10*, 1–11.
36. Tsay, Y.-F., Fan, S.-C., Chen, H.-Y., & Chen, K.-E. (2016). *Method for changing nitrogen utilization efficiency in plants*. Google Patents.
37. Zanke, C. D., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., & Neumann, K. (2014). Whole genome association mapping of plant height in winter wheat (Triticum aestivum L.). *PLoS ONE, 9*, e113287.

Authors and Affiliations

Achla Sharma[1] · Mian A. R. Arif[2] · M. Shamshad[1] · Kanwardeep S. Rawale[3] · Anureet Brar[1] · Juan Burgueño[4] · Sajid Shokat[2] · Ravinder Kaur[1] · Parsahnt Vikram[5] · Puja Srivastava[1] · Nitika Sandhu[1] · Jayesh Singh[1] · Satinder Kaur[1] · Parveen Chhuneja[1] · Sukhwinder Singh[3]

Mian A. R. Arif
m.a.rehman.arif@gmail.com

M. Shamshad
Shamshad.rattan@gmail.com

Kanwardeep S. Rawale
kanwar.harry1989@gmail.com

Anureet Brar
brar.13anureet@gmail.com

Juan Burgueño
j.burgueno@cgiar.org

Sajid Shokat
Sajid_agrarian@yahoo.com

Ravinder Kaur
gillravinder1990@gmail.com

Parsahnt Vikram
pvikramseedwheat@gmail.com

Puja Srivastava
pujasri1@gmail.com

Nitika Sandhu
nitikasandhu@pau.edu

Jayesh Singh
jayesh_singh@pau.edu

Satinder Kaur
satinder.biotech@pau.edu

Parveen Chhuneja
pchhuneja@pau.edu

[1] Punjab Agricultural University, Ludhiana, India

[2] Nuclear Institute for Agriculture and Biology, Faisalabad 38000, Pakistan

[3] 222 Mary Jena Lane, Pullman, WA 99163, USA

[4] CIMMYT, Carretera México Veracruz Km. 45, El Batán, 56237 Texcoco, CP, Mexico

[5] International Center for Biosaline Agriculture, Academic City, Dubai, UAE