Genome-Wide Association Mapping of Virulence Genes in Wheat Karnal Bunt Fungus *Tilletia indica* Using Double Digest Restriction-Site Associated DNA-Genotyping by Sequencing Approach

Mohamad Ayham Shakouka[1,2], Malkhan Singh Gurjar[1]*, Rashmi Aggarwal[1], Mahender Singh Saharan[1], Robin Gogoi[1], Naresh Bainsla Kumar[3], Shweta Agarwal[1], Tej Pratap Jitendra Kumar[1], Bassam Bayaa[2] and Fateh Khatib[2]

[1] Division of Plant Pathology, Indian Agricultural Research Institute ICAR-IARI, New Delhi, India, [2] Division of Plant Protection, University of Aleppo, Aleppo, Syria, [3] Division of Genetics, Indian Agricultural Research Institute ICAR-IARI, New Delhi, India

Tilletia indica is a quarantine fungal pathogen that poses a serious biosecurity threat to wheat-exporting countries. Acquiring genetic data for the pathogenicity characters of *T. indica* is still a challenge for wheat breeders and geneticists. In the current study, double digest restriction-site associated-DNA genotyping by sequencing was carried out for 39 *T. indica* isolates collected from different locations in India. The generated libraries upon sequencing were with 3,346,759 raw reads on average, and 151 x 2 nucleotides read length. The obtained bases per read ranged from 87 Mb in Ti 25 to 1,708 Mb in Ti 39, with 505 Mb on average per read. Trait association mapping was performed using 41,473 SNPs, infection phenotyping data, population structure, and Kinship matrix, to find single nucleotide polymorphisms (SNPs) linked to virulence genes. Population structure analysis divided the *T. indica* population in India into three subpopulations with genetic mixing in each subpopulation. However, the division was not in accordance with the degree of virulence. Trait association mapping revealed the presence of 13 SNPs associated with virulence. Using sequences analysis tools, one gene (*g4132*) near a significant SNP was predicted to be an effector, and its relative expression was assessed and found upregulated upon infection.

Keywords: ddRAD-GBS genotyping by sequencing, GWAS-genome wide association study, SNP, association mapping, virulence genes, *T. indica*, population structure

1. INTRODUCTION

The virulence of fungal plant pathogens is complex and difficult to be characterized at the genetic level, due to several factors including the high level of genetic diversity among genotypes in the natural populations, the complexity of inheritance between pathogens' generations, the difficulties in conducting crosses among different genotypes for association mapping, the scarcity of genomic information, and the high cost of genotyping using *de novo* sequencing (Kumar et al., 2016). However, obtaining high density genotyping methods

can circumvent these limitations, especially since, the conventional markers only revealed a weak link between genotypes and virulence traits (Kolmer, 2001, 2013, 2015; Kolmer and Acevedo, 2016). The 6 k Illumina Infinium single nucleotide polymorphism (SNP) assay is considered the first high dense genotyping tool (Song et al., 2015), which has been used in many bi-parental mapping studies and helped in discovering many important agronomic traits. However, fast genotyping of hundreds or thousands of SNPs is not feasible using it (Bello et al., 2014; Brisco et al., 2014; Mukeshimana et al., 2014). Recently, a new high-density genotyping method called genotyping by sequencing (GBS) has become available, in which, next-generation sequencing (NGS) is combined with restriction-associated DNA (RAD) tagging that successfully generated a large number of SNPs per population, which could be used in the mapping of particular genomic regions (Moghaddam et al., 2016), and enabled multiple individuals to be sequenced at the same time (Mascher et al., 2013). This method employs restriction enzymes to capture genomic regions of interest, followed by the adaptors' ligation to the flanking sections of the restriction sites, resulting in barcoded libraries. HiSeq and Illumina GAII platforms were primarily employed in this method (Poland and Rife, 2012). Recently, an enhanced RAD-GBS methodology adopted for the Ion Torrent has just been available (Rothberg et al., 2011). RAD-GBS has been successfully employed in genotyping a variety of plant traits (Leboldus et al., 2015; Gao et al., 2016; Carlsen et al., 2017), generating high-density genotypes datasets by utilizing Illumina NGS technologies (Baird et al., 2008; Baxter et al., 2011; Chutimanitsakun et al., 2011; Elshire et al., 2011; Pfender et al., 2011; Ma et al., 2012; Poland et al., 2012a). Targeting certain genomic regions can result in a significant number of (SNP) markers that can be utilized to create a lower-complexity genome sequence (Elshire et al., 2011; Poland et al., 2012b), and gives a plethora of genetic data that may be used to determine the relationship between desired phenotypes and genetic markers using recombination events based on linkage disequilibrium (LD) accumulated over many generations (Korinsak et al., 2019). Furthermore, this approach has been employed in a variety of investigations, including genetic mapping, genomic selection, elucidating the basis of disease resistance related traits at the genetic level, and genetic variability studies in a variety of taxa (Adhikari and Missaoui, 2019). Using the generated genotyping data in genome-wide association studies (GWASs) identified thousands of genetic loci associated with complex phenotypes and had advantages over biparental mapping in terms of map resolution and time to build a mapping population (Welter et al., 2014). Plants have long employed association mapping using this approach, and fungi have lately been added to the list. In *Puccinia triticina*, 30 isolates collected from different locations were genotyped based on 2,125 SNP markers discovered using the ddRAD approach, phylogenetical analysis grouped the studied isolates in nine distinct clusters (Aoun et al., 2020) with a general correlation between the genotyping data and virulence phenotyping data. Similarly, genome wide association study of 73 isolates of *Magnaporthe oryzae* demonstrated the presence of eight subpopulations, however, their division was unrelated to the degree of virulence. Association mapping analysis revealed that five SNPs were found associated with fungal virulence in the studied population (Korinsak et al., 2019). The availability of such a method can suggest a sufficient number of SNP markers (Poland and Rife, 2012) making the information on pathogenicity profiles of the pathogen more available. A detailed understanding of the genetic makeup and genes governing pathogenicity is essential to provide the required significant data for use in effective management strategies of some pathogens, particularly those for which genomic information is still scarce. *Tilletia indica* is a quarantine fungus causing Karnal bunt of wheat in many tropical countries, a characteristic rotting- fish like odor (due to the presence of trimethylamine) liberates from infected grains. The main economic effect of Karnal bunt disease is a reduction in grain quality rather than quantity. It reduces the viability and weight of the seed, as well as the quality of the flour (Singh, 2005). High virulence and genetic variability were detected among *T. indica* isolates due to the secondary sporidia mating immediately before the infection resulting in genetic recombination (Gurjar et al., 021a; Kumar et al., 2009; Aggarwal et al., 2010). The knowledge about the genomic regions linked with *T. indica* virulence is still limited, and obtaining such information is essential for the identification and characterization of virulence-related traits, as well as their regulatory genes and their corresponding resistance related genes. Determining which virulence genes exist in each population will also aid in the formulation of a breeding program based on the usage of optimal resistance genes. Accordingly, the present study was performed with a view to executing a genome-wide association study in *T. indica* using a ddRADseq-genotyping by sequencing approach to identify SNPs and virulence factors in the population of *T. indica* in India.

2. MATERIALS AND METHODS

2.1. Fungal Population

Infected wheat grains were collected from 33 different locations in India (**Supplementary Table S1**) and bunted kernels were detected by visual examination and then separated. The fungal cultures were established following the Warham method (Warham, 1986). Briefly, 20 μl of the spore suspension from each sample was placed into 2 ml centrifuge tubes, the volume was made up to 2 ml with sterile distilled water and kept overnight inside an incubator at 21°C to moisten the harvested teliospores, and increased the susceptibility of contaminating saprophytes to sterilization. The tubes were then centrifuged at 1,200 g for 3 min, and the supernatants were discarded. For surface sterilization, the pellets were suspended in 5 ml of (0.3–0.5% NaOCl), reversing the tubes several times, and then centrifuged at 1,200 g for 1 min (Bonde et al., 1999). The supernatant was discarded and the pellets were washed two times with 1 ml of sterile distilled water and centrifuged at 1,200 g for 5 min. Finally, 1 ml of sterile distilled water (SDW) was used to resuspend the pellets and 200 μl of the spores suspension was inoculated into 2 % water agar (WA) plates containing streptomycin sulfate (Bulletin, 2007). The water agar plates were kept inside an incubator at 18 ± 2°C with a 12-h light cycle and checked regularly under a microscope

Grades	Description
0	Healthy seed of wheat
1	Germinal tip infection
2	Infection advancing to the kernel groove (1/2th grain)
3	More advanced infection (3/4th of the grain)
4	"Canoe" symptom-the hollowing out of seed interior

until the teliospores germinated and produced promycelium bearing basidiospores. Upon germination, a small bit of media (1 x 1 cm) having a germinated teliospore was transferred and stuck to the underside of a Petri plate cover, which allowed the sporidia to be deposited onto the media surface. The inoculated plates were then kept inside an incubator at 18 ± 2°C with a 12-h light cycle until mycelial mats appeared on the media surface, thereafter it was transferred into a test tube containing PDA media, labeled, and incubated at 18 ± 2°C with a 12 h light cycle, with subculturing on regular time intervals of 3 weeks.

2.2. Assessment of Karnal Bunt Disease and Virulence Analysis

Seeds of susceptible wheat genotype (WH542) were sown in 12-inch pots under the net house conditions during the winter season (November–March). To make the inoculum, sterile distilled water was poured over the fungal slants cultures and scraped gently with a glass rod to liberate allantoid secondary sporidia. To remove media debris, the spore suspension was filtered through a double layer of muslin fabric, and the spore concentration was set at more than 10,000 sporidia per ml. Five healthy plants per pot were inoculated at the booting stage (Z-49 stage) (Zadoks et al., 1974) by syringe inoculation method in the evening hours (Aujla et al., 1987) under conditions of excessive humidity. The inoculated heads were plucked and threshed individually as they reached maturity. The endosperm portion that had developed into a sooty mass of teliospores was utilized to classify the infected grains into five infection categories (**Table 1**).

The coefficient of infection was calculated as per the method of Aujla et al. (1989).

$$coefficient\ of\ infection = (XiYi/N)100 \qquad (1)$$

Where, Xi and Yi refer to numerical ratings and grains counts, respectively and i is a value between 0 and 4 (grade of infection).

2.3. Genomic DNA Extraction and Double Digest Restriction-Site Associated DNA-Genotyping by Sequencing

Liquid potato dextrose broth media was used to establish the fungal colonies. Mycelial mats were gathered once they appeared using sterile filter papers and preserved at −80°C. The CTAB method was used to extract the DNA (Crowhurst et al., 1995). Extracted DNA samples (50 μl) were then incubated for one hour at 37°C with 2 μl of RNase-A. The concentration and

purity of DNA were evaluated using UV absorbance at 260 and 280 nm in a Nano Drop 2000/2000 (Thermo Scientific) as well as electrophoresis on agarose gels with a marker (1 kb DNA ladder, Thermo Fisher Scientific, USA) and visualized with a UV transilluminator. GBS libraries were created following Peterson's approach (Peterson et al., 2012). Briefly, using the restriction enzymes *SphI* and *MluCI* double digestion of genomic DNA (1 μg) of each *T. indica* isolate was conducted. After that, Ampure beads were used to clean the digested products before they were ligated using T4 DNA ligase with the forward and reverse P1 and P2 barcoded adaptors. The ligated products were then pooled, cleaned, and size selected using 2% agarose gel electrophoresis (Thermo Fisher, Waltham, MA, USA). PCR amplification was carried out to enrich and add the Illumina specific adapters and flow cell annealing sequences. The quality was checked on a bioanalyzer high sensitivity DNA kit (Agilent Technologies, Santa Clara, CA, USA), followed by final pooling and sequencing in AgriGenome Labs Pvt Ltd. Upon sequencing, indexes were removed from raw and checked for the presence of rad-tags on both reads using in-house scripts. The dDocent (v.2.6.0) was used to remove any low-quality bases (below a quality score of 20) and adapters (Puritz et al., 2014), a sliding 5 bp window was used to trim the bases when the average quality score drops below 10. After quality filtering, filtered reads took by the dDocent pipeline for alignment using BWA (v-0.7.8). The alignment has been done with filtered reads against the reference genome *T. indica* RAKB–UP–1 ASM222083v1. The SNPs were called with Freebayes (v.1.2.0), and the bi-allelic SNPs were filtered with bcftools (v.1.6) and Agri Genome Inhouse scripts at read depths of 5 and 10. The homozygous polymorphic marker identification was carried out using the bi-allelic genotype at read depth 10. The kinship matrix was generated on TASSEL (v 5.2.31) (Bradbury et al., 2007).

2.4. Population Structure and Phylogenetic Analysis

The number of possible population structures of *T. indica* was analyzed using the Bayesian clustering approach in STRUCTURE version 2.3.4 (Greenbaum et al., 2016). The number of feasible clusters (K) was checked from 1 to 10, with each K value receiving 10 iteration runs. The admixture model was used to run 1,00,000 Markov Chain Monte Carlo (MCMC) repetitions after a burn-in time of 1,00,000 steps. The *ad hoc* statistic Delta K was used to calculate the ideal K value in STRUCTURE HARVESTER version 0.6.94 (Earl and vonHoldt, 2011). STRUCTURE PLOT 2.0 (Ramasamy et al., 2014) was used to create a graph. DARwin version 6.013 (Bradbury et al., 2007) was used to implement the neighbor-joining cluster analysis and principal coordinate analysis (PCoA). The principal component analysis (PCA) was done using SNP related package in R (Team et al., 2013) based on the identified SNPs.

2.5. Virulence Gene Association Mapping

Genome-wide association study (GWAS) was carried out based on the mixed linear model (MLM) to calculate the relationships between markers and phenotyping data using TASSEL software v 5.2.31 (Bradbury et al., 2007). To avoid misleading connections,

MLM uses phenotyping data, population structure (Q), genotyping data, and kinship (K) matrix as covariates during a GWAS. The logarithm of odds was calculated based on the *P*-value estimated using TASSEL software v 5.2.31, and significantly associated SNPs were cut if their LOD score was equal to or greater than 3.0. The proportion of the phenotypic variation explained (PVE) was calculated for each marker using the relevant R2 in TASSEL software v 5.2.31.

2.6. Gene Annotation

The nucleotides sequence was retrieved from a region of 250 Kb upstream and downstream of the SNPs using a customized R script (Team et al., 2013), and used for candidate genes search. Functional annotations of each postulated candidate gene were investigated using the Swiss-Prot database (Boutet et al., 2016). The candidate genes were run with SignalP 4.0 (Petersen et al., 2011) to obtain the initial list of potential secreted proteins (SPs). TMHMM 2.0 (Krogh et al., 2001) was used to predict potential transmembrane domains. Only proteins with zero or one transmembrane domain prediction that overlapped the signal peptide prediction were considered for further investigation. Protein sub-cellular localizations were predicted using WoLFPSORT (Horton et al., 2007).

2.7. *In planta* Relative Expression Analysis of the Predicted Virulence Gene

Under net house conditions, seeds of resistant (HD30) and susceptible (WH542) wheat genotypes were sowed in 12-inch pots (10–15 seeds per pot). Five healthy plants per pot were inoculated at the booting stage (Z-49 stage) (Zadoks et al., 1974) using the syringe method (Aujla et al., 1989). Inoculated heads were harvested at 0, 1, 3, 5, and 13 days after inoculation. Three biological replicates from each treatment were harvested and stored at −80°C. Trizol method (Chomczynski and Mackey, 1995) was used for RNA extraction from the infected spikes. Total obtained RNA was used for cDNA synthesis using the Verso cDNA synthesis kit (Thermo Fisher Scientific). The GenScript Primer Design software (https://www.genscript.com/tools/real-time-pcr-tagmanprimer-design-tool) was used to design the specific primers for the putative effector gene (Forward: 5' AAGAACGCCGAATGGGATTG 3', Reverse: 5' AGAGAGGGCAGACCTTGAAC 3'). *GAPDH* gene was used as a reference gene (Forward: 5' GGGAGCAGAGATGACAACCT 3', Reverse: 5' GAGTCCACCGGTGTCTTCA 3'). GCC Biotech (India) Pvt. Ltd. synthesized the designed primers. The expression of the potential effector was measured using a real-time PCR technique and the fold changes value was calculated using the 2 ("delta delta CT) method (Rao et al., 2013), and statistically analyzed using an unpaired *t*-test (Schober and Vetter, 2019).

3. RESULTS

3.1. Virulence Analysis of *T. indica* Isolates

Successfully, 39 *T. indica* cultures were established and maintained. The virulence analysis on susceptible wheat genotype (WH 542) showed variability in virulence among the isolates. The coefficient of infection (CI) ranged from (3.88) in

Ti 21, Tonk, Rajasthan which was less virulent to (32.5) in Ti36, Taroari, Karnal, Haryana which was highly virulent (**Figure 1**). Among the isolates, 17 isolates showed highly aggressive causes of the disease, most of them were collected from Rajasthan, and one of them, Ti39, was collected during the year 2007 from Meerut, UP, and found still highly virulent (**Supplementary Table S1**).

3.2. Genotyping by Sequencing (dd-RAD-GBS) of *T. indica* Isolates

Double digestion of the extracted genomic DNA from 39 *T. indica* isolates was performed using two restriction enzymes *SphI* and *MluCI*, those generated reduced representation libraries, with 3,346,759 raw reads per sample on average, and 151 x 2 nucleotides read length. The obtained bases per read ranged from 87 Mb in Ti 25 to 1,708 Mb in Ti 39, with 505 Mb on average per read. The sequencing reads were then filtered for the presence of tags on both sides, and reads with tags on both sides ranged from 117,000 reads in Ti 5 to 9,000,000 in Ti 11, with an average of 1,804,287 reads. The data were then cleaned up by removing the low-quality bases and adaptors, resulting in 1,717,146 reads on an average per sample, which has been used for further analysis.

3.3. Mapping and Allelic Variation

The sequencing reads with quality scores ratings greater than 20% were aligned against the reference genome *T. indica* RAKB–UP–1 ASM222083v1 and revealed the presence of 41,473 SNPs among *T. indica* isolates. The most frequent base in the variation sites was guanine (83,993 times), while the lowest was W (either Adenine or Thiamine) with 1,540 times (**Figure 2**). The transition and transversion SNPs analysis showed that transition SNPs were more frequent (64.19%, 22,239 SNPs) than transversion ones (35.80%, 12,402 SNPs). The A/G transitions were detected in 11,160 positions, accounting for the highest frequency (32.21%), while A/T transversions occurred in the lowest frequency (6.26%) and were detected in 2,171 positions (**Figure 3**). The observed heterozygosity ranged from 0.01 to 15.56% with an average of 1.34% (**Figure 4**).

## 3.4. Minor Allele Frequency Distribution

The population's minor allele frequency (MAF) spectrum exhibited a substantial skew toward rare alleles, indicating that there were no recent genetic bottlenecks or mutations and that the detected major alleles for the SNP are preserved (**Figure 5**).

3.5. Population Structure of *T. indica* in India

The number of subpopulations in the 39 *T. indica* isolates tested was determined using the STRUCTURE 2.3.4 software. The population was analyzed using a total of 41,473 SNPs detected in 39 isolates. To determine the appropriate number of subpopulations, the number of clusters (K) was plotted against ΔK (**Figure 6**). The highest ΔK value was found at K = 3, indicating that the investigated isolates belong to three subpopulations. As a result, three clusters were the most likely number of clusters present in this population. Clusters ordered by Q (membership coefficient) are shown in **Figure 7**. Genetic mixing of more than one sub-population with varying proportions was seen in all populations. The three groups consist
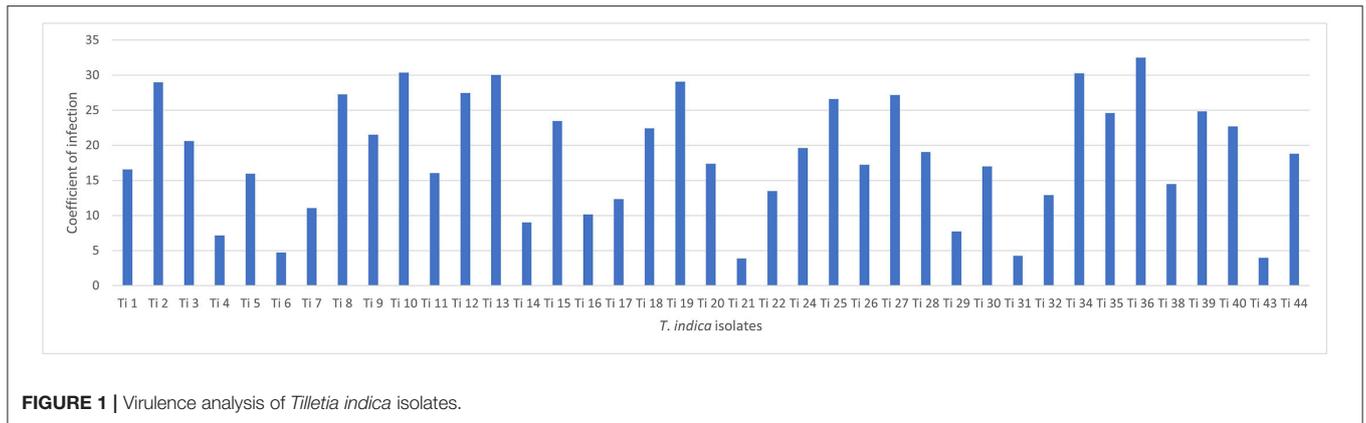
FIGURE 1 | Virulence analysis of *Tilletia indica* isolates.
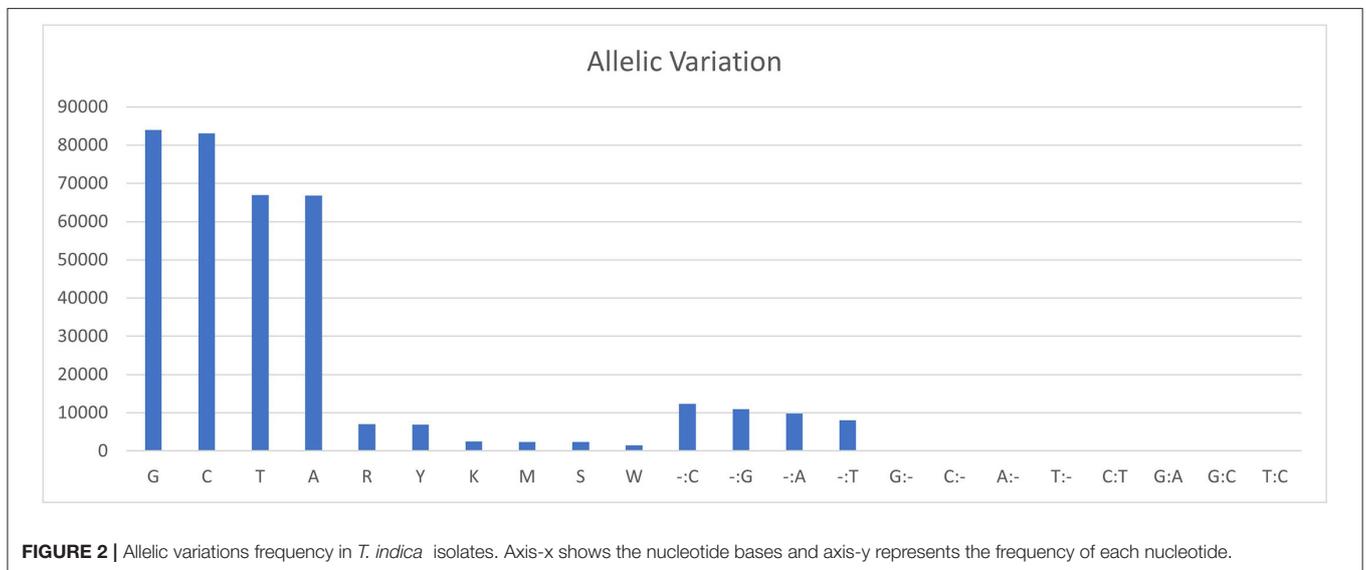
FIGURE 2 | Allelic variations frequency in *T. indica* isolates. Axis-x shows the nucleotide bases and axis-y represents the frequency of each nucleotide.

of the red, blue, and green groups, respectively. PCA showed that the first and the second principal components explained 50.32 and 9.91%, respectively, of the total variability (**Figure 8**). The first quarter accommodated isolates most of them were collected from Uttar Pradesh, India, and were more virulent as compared to the remaining quadrants. The second quarter accommodated the reference genome along with 3 isolates collected from Rajasthan, India, and all of them were plotted away from the origin. The third and fourth quarters accommodated isolates were collected from Haryana, India, and Rajasthan, India. The Multi-Dimensional scaling (MDS) and PcoA (**Figure 9**) showed a similar relationship among *T. indica* isolates.

3.6. Phylogenetic Tree Analysis in 39 Isolates of *T. indica*

Neighbor- joining phylogenetic tree grouped the isolates into three clusters (**Figure 10**), with a variable number of individuals inside each one. The first cluster included only two isolates Ti38 and Ti40 which were quite distinct from other isolates and belong to Uttar Pradesh, India, while the second cluster included isolates

and most of them were collected from Rajasthan, India. The third cluster included all the remaining isolates and most of them were collected from Uttar Pradesh, India.

3.7. Trait Association Mapping

The SNP matrix used for association mapping in this study was generated by genotyping by sequencing (GBS) of 39 *T. indica* isolates, followed by GBS "Discovery Pipeline" using TASSEL (v 5.2.31). A total of 41,473 detected SNPs, saved in HapMap were used as genotypic data for GWAS. In order to find virulence genes, the disease reactions on susceptible wheat genotype (WH 542), principal components analysis (PCA), population structure, and Kinship matrix were used to conduct SNP-trait association analysis and revealed significant single nucleotide polymorphisms associated with the disease virulence at 13 positions in the genome (**Figure 11** and **Table 2**).

3.8. Identification of Candidate Genes

A search within 500 kb ($\pm 250 kb$) around of the peak SNPs identified a total of 43 potential genes near significant SNP (**Supplementary Table S3**). Gene annotation using
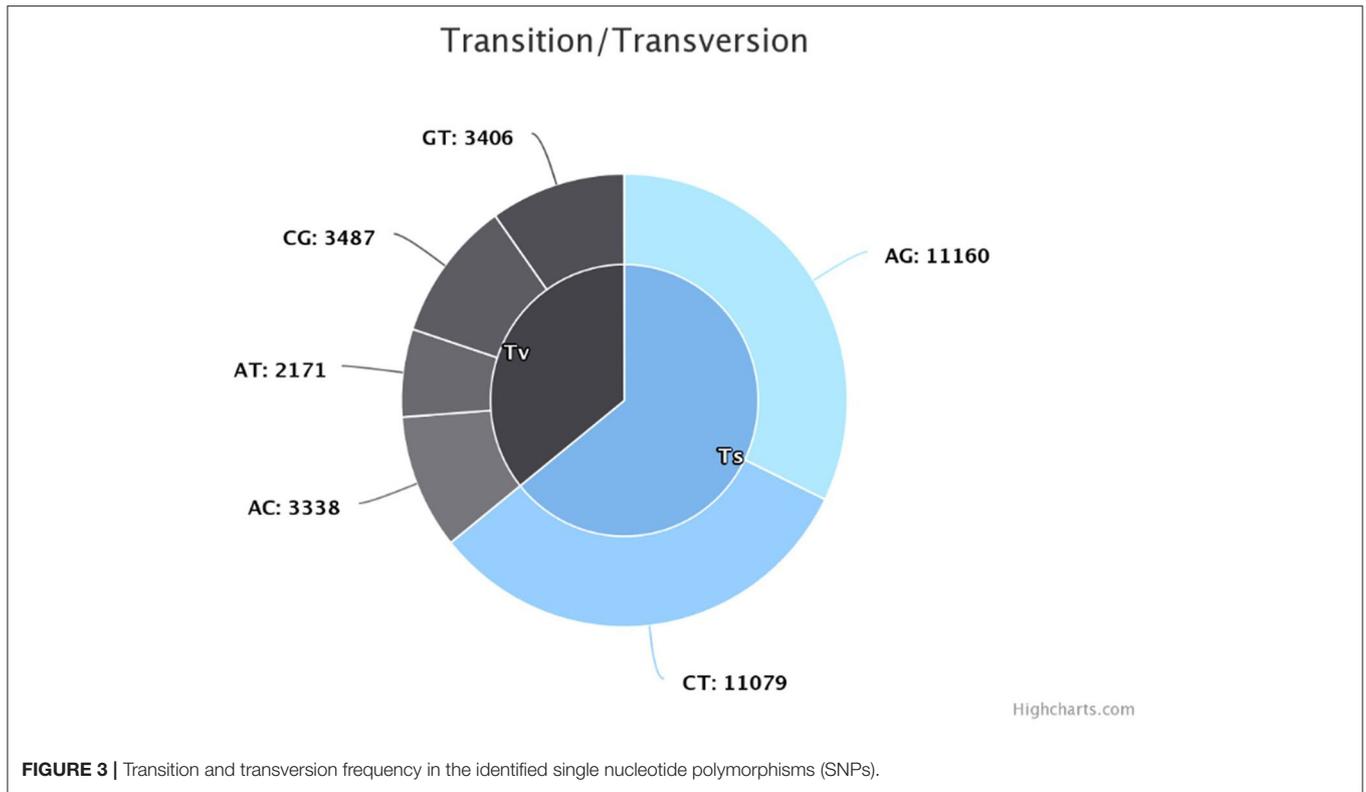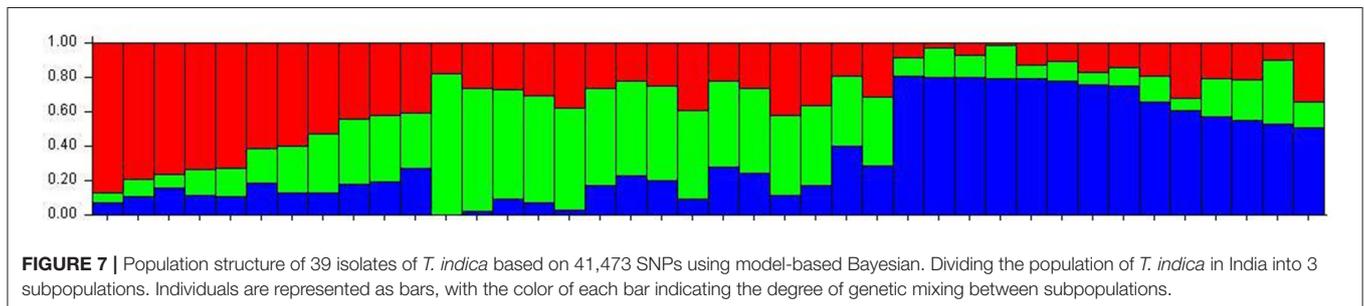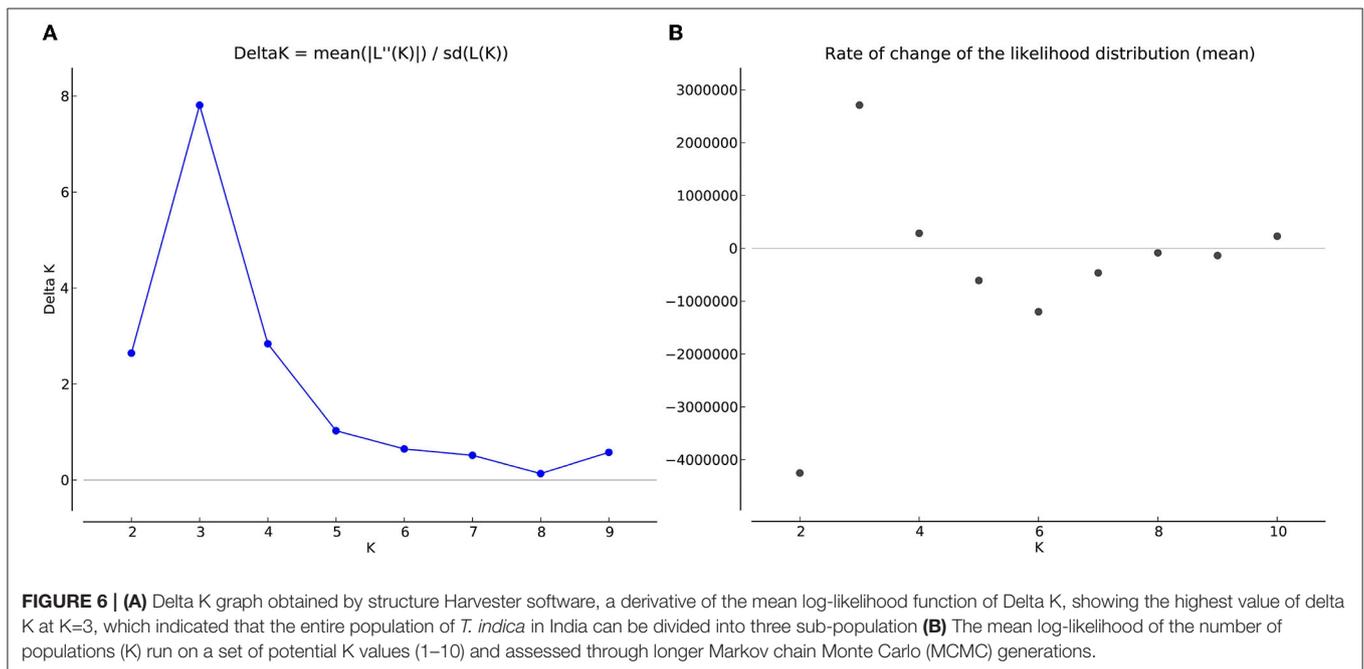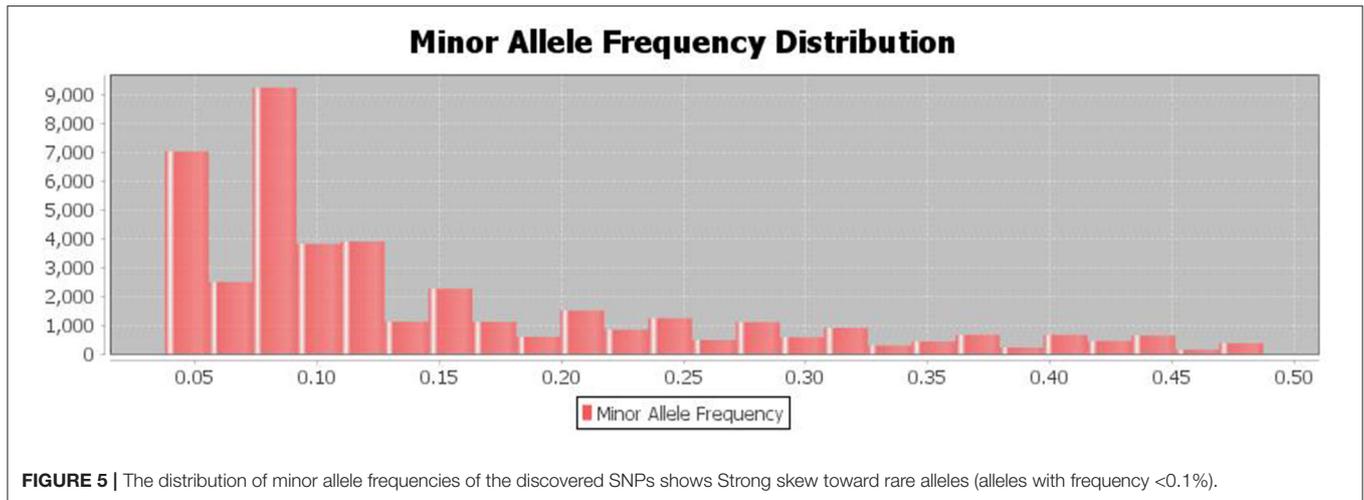
FIGURE 3 | Transition and transversion frequency in the identified single nucleotide polymorphisms (SNPs).
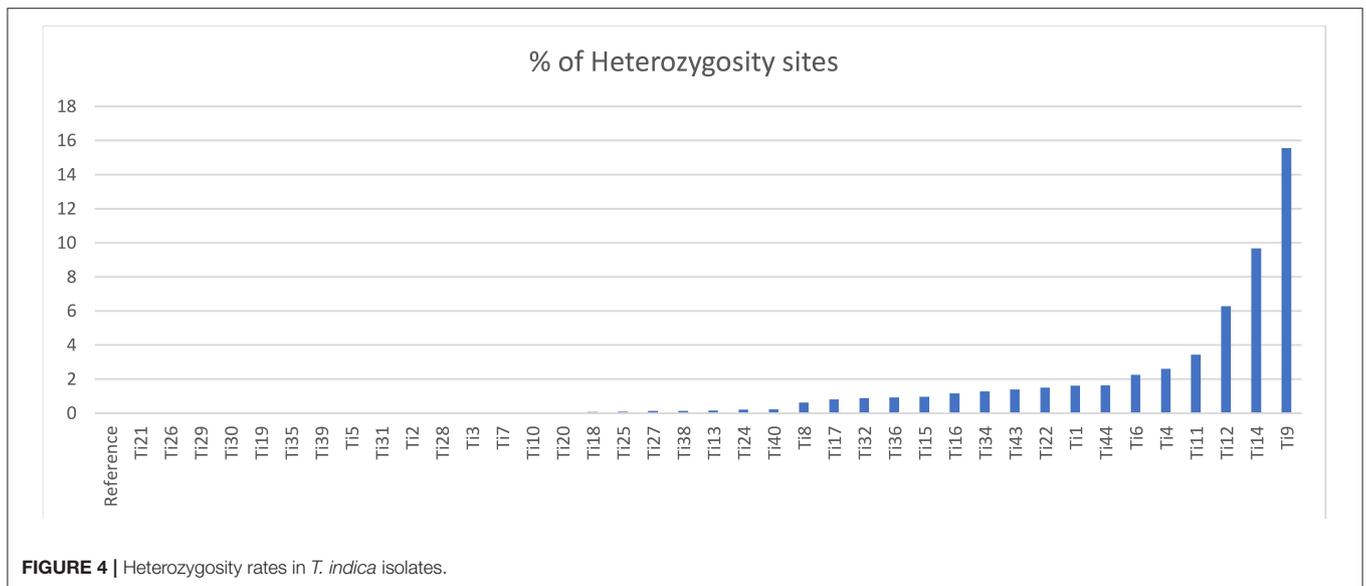
FIGURE 4 | Heterozygosity rates in *T. indica* isolates.

Swiss-Prot showed that most of those genes (22 genes) had an uncharacterized function and only 21 genes had functional annotations. The annotated genes had a wide variety of functions, 12 genes were found to have putative virulence functions. To discover potential effectors, the identified genes were analyzed using sequence-based prediction tools. Out of the 43 genes, 12 genes were predicted to secrete proteins have a signal peptide. The potential transmembrane domains were also predicted and only 7 genes were found to have no or only one transmembrane domain protein overlapping the signal peptide prediction, and these genes were further investigated. Subcellular localization using revealed that 4 proteins were expected to be extracellularly localized. Blast analysis of these genes against pathogen-host interaction (PHI) database showed that, *g4132* gene is predicted to secret protein homology to an effector protein in *Rhizoctonia solani*, two genes *g1133* and *g1132* are homology to virulence genes in *Ustilago maydis*, while *g4133* gene secrets a protein of uncharacterized function with no hit.

FIGURE 5 | The distribution of minor allele frequencies of the discovered SNPs shows Strong skew toward rare alleles (alleles with frequency <0.1%).

FIGURE 6 | (A) Delta K graph obtained by structure Harvester software, a derivative of the mean log-likelihood function of Delta K, showing the highest value of delta K at K=3, which indicated that the entire population of *T. indica* in India can be divided into three sub-population **(B)** The mean log-likelihood of the number of populations (K) run on a set of potential K values (1–10) and assessed through longer Markov chain Monte Carlo (MCMC) generations.

FIGURE 7 | Population structure of 39 isolates of *T. indica* based on 41,473 SNPs using model-based Bayesian. Dividing the population of *T. indica* in India into 3 subpopulations. Individuals are represented as bars, with the color of each bar indicating the degree of genetic mixing between subpopulations.

3.9. *In planta* Relative Expression Analysis of the Predicted Effector

Expression of the putative effector gene (*g4132*) was found upregulated in both genotypes at 3 days post inoculation (dpi). A sharp statistically significant increase ($p < 0.05$) in the susceptible wheat genotype (WH 542) compared to the resistant genotype (HD30) was detected at 7 dpi reaching its maximum expression 108 fold. After 13 days, the expression decreased
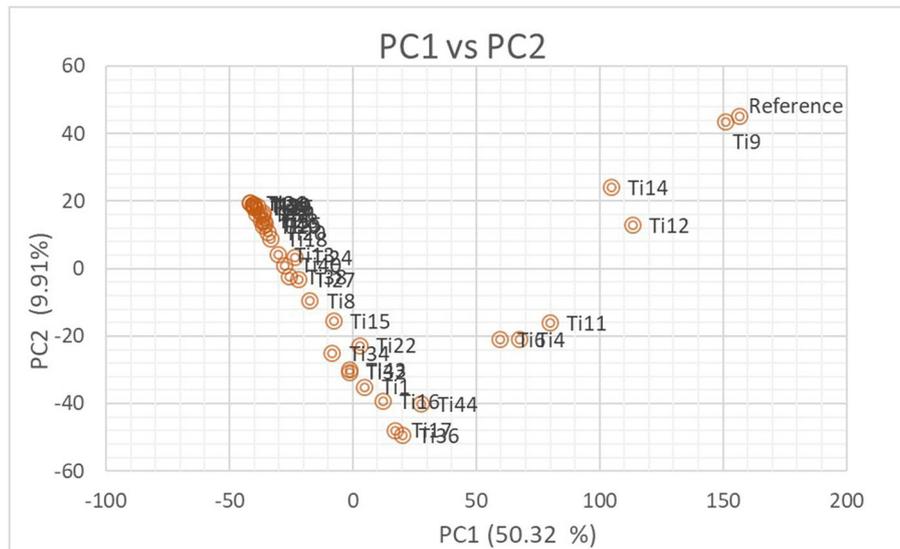
FIGURE 8 | Biplot showing the relationship among *T. indica* isolates in terms of principal components analysis (PCA). Axis-x and axis-y refer to the first and second principal components, which explained 50.32% and 9.91% of the total variation, respectively.
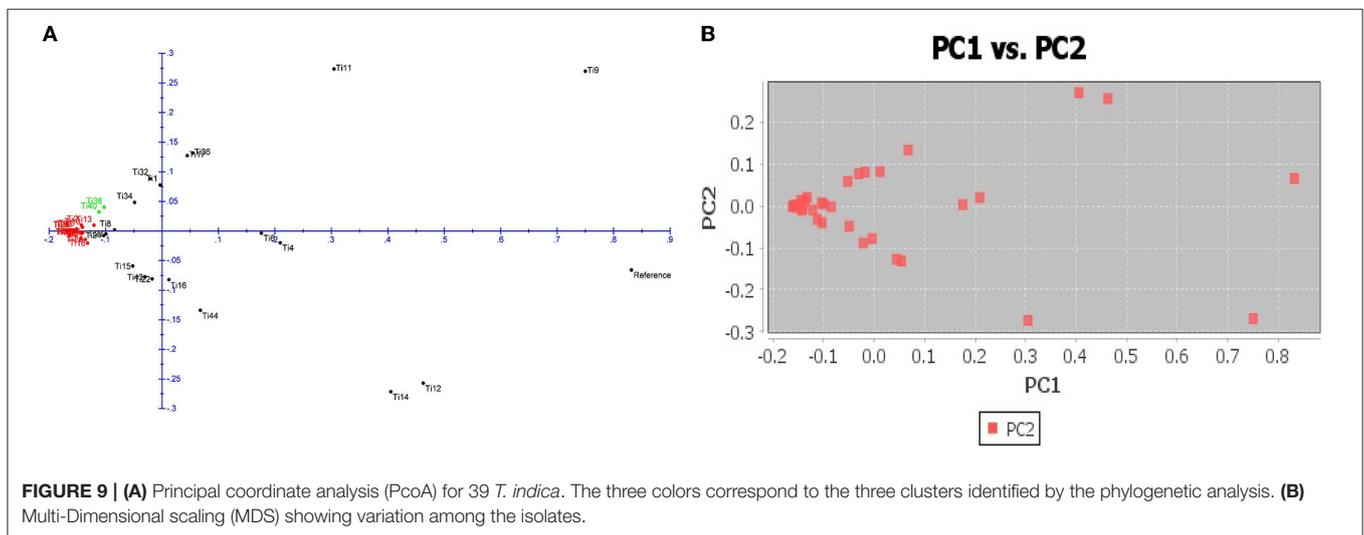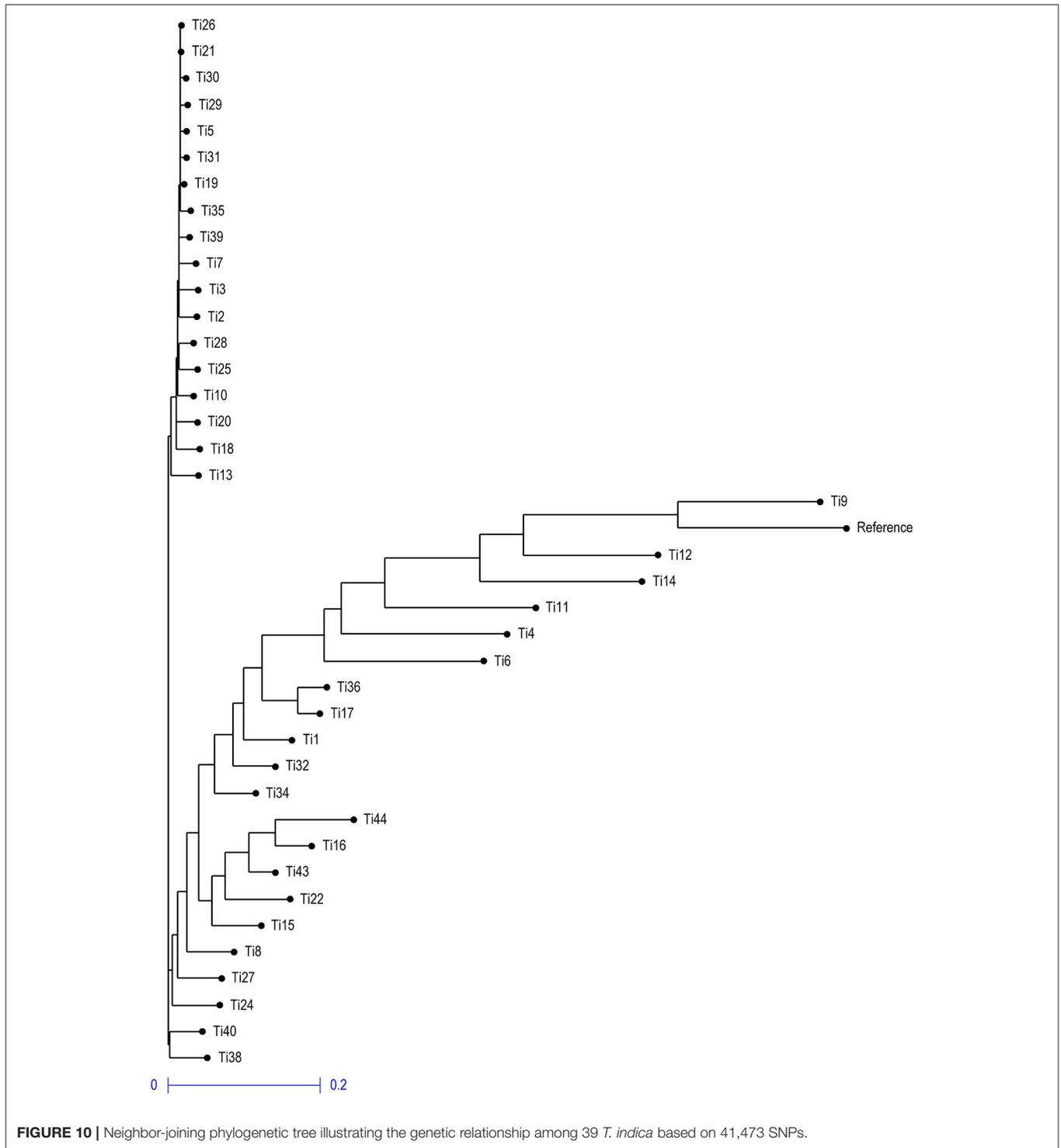

FIGURE 9 | (A) Principal coordinate analysis (PcoA) for 39 *T. indica*. The three colors correspond to the three clusters identified by the phylogenetic analysis. **(B)** Multi-Dimensional scaling (MDS) showing variation among the isolates.

up to 53 fold in the susceptible genotype, while the maximum expression in the resistant genotype (HD30) was 38 fold at 3 dpi (**Figure 12**).
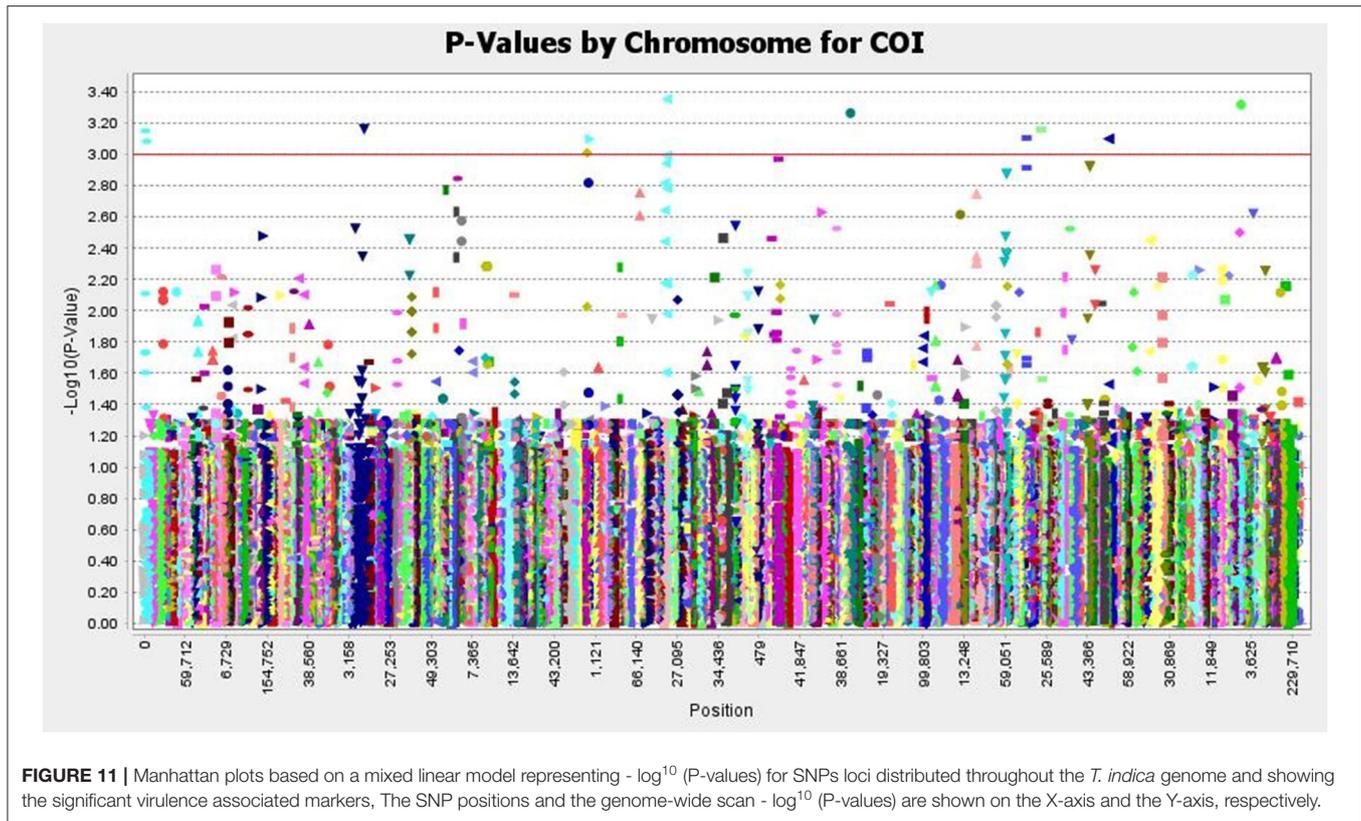
4. DISCUSSION

Karnal bunt is a complicated, economically significant, and difficult-to-manage fungal plant disease that affects numerous wheat-growing regions around the world. Effective disease management strategy necessitates knowledge of the pathogen's pathogenicity profiles. Genotyping by sequencing can yield a large number of SNP markers, which can be used to perform powerful analysis, especially when combined with phenotyping data in GWAS (Kim et al., 2016), which allow for the discovery

of specific markers linked to virulence genes without the need for population mapping studies (Soto-Cerda and Cloutier, 2012). In the present investigation, genotyping by sequencing 39 *T. indica* isolates using the ddRAD technique was carried out to investigate the genetic variation among isolates, and revealed the presence of 41,473 SNPs exhibiting the high genetic variability among isolates. The population of *T. indica* in India was divided into three subpopulations, each with genetic mixing of several genotypes, revealing the presence of common ancestry between the studied isolates or the result could be influenced by the small size of the population used in this study (Saleh et al., 2014). Previous research revealed that the majority of genetic variation is caused by genetic mutations, parasexual recombination in fungus, and natural or artificial selection

FIGURE 10 | Neighbor-joining phylogenetic tree illustrating the genetic relationship among 39 *T. indica* based on 41,473 SNPs.

pressure (Saleh et al., 2014). One of these variables could be to blame for the observed genetic variety among *T. indica* isolates, where substantial levels of nucleotide polymorphisms were discovered in the 39 isolates tested. The cluster pattern, on the other hand, did not correspond to the pathogenicity of each *T. indica* isolate. Similarly, earlier research has found

no link between populations, virulence, and location, which could be due to gene or genotype flow (Rajalakshmi et al., 2017; Korinsak et al., 2019). Our study focuses on gaining a deeper knowledge of *T. indica* pathogenicity. Because *T. indica* is a quarantine fungal pathogen, it is critical to avoid creating new virulent genotypes by artificial sexual reproduction.
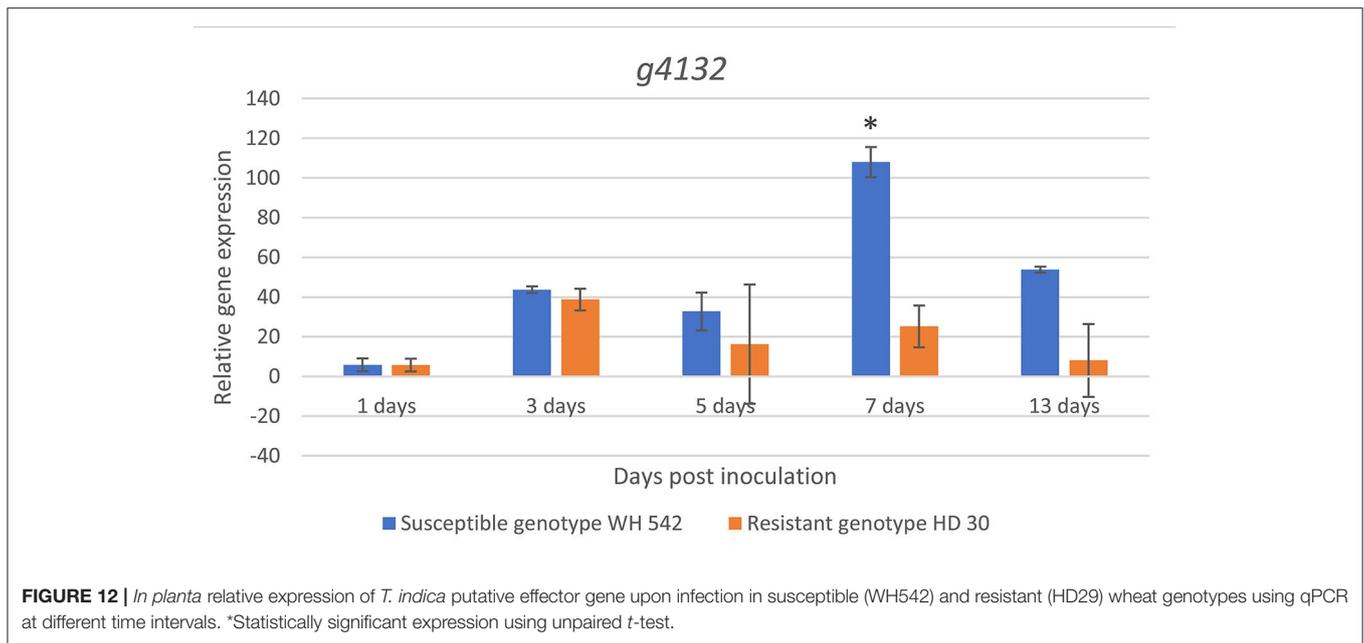
FIGURE 11 | Manhattan plots based on a mixed linear model representing - log[10] (P-values) for SNPs loci distributed throughout the *T. indica* genome and showing the significant virulence associated markers, The SNP positions and the genome-wide scan - log[10] (P-values) are shown on the X-axis and the Y-axis, respectively.

TABLE 2 | The position of significantly associated single nucleotide polymorphisms (SNPs) with virulence of *Tilletia indica* based on association analysis using the mixed linear model (MLM).

Trait (CI*)	Marker	Alleles	Position	MarkerR2
CI	rs.MBSW01000808.1.55604	C/T	5,5604	0.54587
CI	rs.MBSW01001646.1.3308	G/A	3,308	0.53852
CI	rs.MBSW01001051.1.80074	C/G	80,074	0.52753
CI	rs.MBSW01000359.1.360837	G/A	360,837	0.50702
CI	rs.MBSW01001343.1.21803	T/C	21,803	0.50666
CI	rs.MBSW01001343.1.21832	C/T	21,832	0.50666
CI	rs.MBSW01000006.1.6428	G/A	6,428	0.50505
CI	rs.MBSW01001301.1.21986	C/T	21,986	0.49573
CI	rs.MBSW01001419.1.48293	A/G	48,293	0.49462
CI	rs.MBSW01001419.1.48328	G/A	48,328	0.49462
CI	rs.MBSW01000668.1.18323	C/T	18,323	0.49453
CI	rs.MBSW01000006.1.40825	T/G	40,825	0.49141
CI	rs.MBSW01000660.1.17242	A/C	17,242	0.47701

CI, Coefficient of infection.

Thus, rather than using artificial sexual reproduction, we used genome-wide association mapping to find new markers linked to virulence. The trait association mapping study indicated a substantial number of markers ($n = 13$) related to virulence, the location of the retrieved markers on the genome was located, and BLAST analysis validated the homology of the sequences surrounding each SNP (250 Kb upstream and downstream). However, the majority of the markers found were in genes encoding putative proteins whose translation products are still unknown. According to a previous study, the genome of the *T. indica* has more than 10,113 genes, 31.2% of them remaining unannotated (Gurjar et al., 2019). As a result, it is not unexpected that the bulk of the markers found were simply hypothetical protein hits. A total of 12 genes encoding putative virulence factors were identified upon genes annotation. *g1132* and *g1133* genes encode a chitin deacetylase which enables hydrolysis and is involved in the carbohydrate metabolic process. Previously, it was discovered that chitin deacetylase plays a role in virulence *via* direct deacetylation of chitin oligomers, which contributes to lysine motif (LysM)-containing receptor perception in the host, inducing ligand-triggered immunity. This has been observed in a variety of fungal and bacterial pathogens (Gao et al., 2019). *g1135* gene encoding a tRNA 2-thiolation plays a role in pathogenicity by intermediating the induction of some pathogenicity related factors in response to the environmental conditions (San Koh and Sarin, 2018). *g1127* gene encoding a UDP glycosyltransferase (UGT), this enzyme enables the transfer of glucose into glycogen. Recent studies have reported the role of glycogen in pathogen colonization and virulence. The breakdown of glycogen may play an important role in the interactions between the host and pathogens which accumulate glycogen throughout their life cycles (He et al., 2018). *g3612* gene encodes a methyltransferase, which is found to play a role in plant defense reactions suppression, but information on how this is achieved is scarce (Ludwig-Müller et al., 2015). *g4366* gene encoding an RNB domain-containing protein, it was shown that the adaptability of some pathogens during the infection process is depending

FIGURE 12 | *In planta* relative expression of *T. indica* putative effector gene upon infection in susceptible (WH542) and resistant (HD29) wheat genotypes using qPCR at different time intervals. *Statistically significant expression using unpaired *t*-test.

on this enzyme's plasticity. It was also reported to be important in eukaryotic cell adhesion and invasion (Matos et al., 2014). *g4527* gene encoding a RING-type domain-containing protein, which manipulates host immune function through promiscuous interactions with numerous proteins evolved during specialized parasite-host interactions (Zhao et al., 2019). *g8485* gene encoding an extracellular carboxylic ester hydrolase, produced by many bacterial pathogens and have been shown to be important for virulence of some pathogens, suggesting that it may play important roles in nutrient utilization and tissue invasion (Xie et al., 2008). *g740* gene encoding an alpha/beta-hydrolase, it was discovered that infection causes it to be induced in *Sclerotinia sclerotiorum*, playing a role in lipid degradation (Seifbarghi et al., 2017). *g4487* gene encoding a cellulase domain-containing protein, which is an important factor in disease development. Several studies support the involvement of cellulases in pathogenesis (Jones and Ospina-Giraldo, 2011). *g4362* gene encoding a thioredoxin domain-containing proteins have the potential to functionally regulate numerous plant immune signaling proteins (Mata-Pérez and Spoel, 2019). *g6053* gene encoding a peptidase M43 domain-containing protein, which is a degrading enzyme used by several pathogens to breakdown host macromolecules, providing nutrition for the pathogen. Moreover, fungal peptidases can inactivate or modify protein components of the host defense machinery and ultimately suppress defense responses (Krishnan et al., 2018). Expression variation of the candidate effector gene (g4132) showed high upregulation in its expression upon infection during host-pathogen interaction in a susceptible host as compared to a resistant host, starting at 3 days post inoculation (dpi), indicating that the pathogen needs few days to establish a relationship with the host (Singh et al., 2020), the expression level continued to increase reaching the expression peak at 7 dpi. The expression pattern of this gene was in accordance with

earlier investigations of the pathogen infection (Aggarwal et al., 1994). This analysis will open new prospects in understanding the nature of the pathogen host interaction by using gene expression levels in elaborating the used pathogenic mechanism which will aid in the development of management measures.

This is the first study to use association mapping to identify SNP sites related to virulence genes in *T. indica*. The information gained using this approach will be valuable in future studies, due to narrowing down the genomic regions including virulence genes without having to map the population or undergoing functional analysis. Especially that *T. indica* is very slow growing on artificial media and it is pathogenesis mechanisms is still complex due to the long survival of teliospore in soil and genetic recombination immediately before infection which happens upon mating of two compatible sporidia (Gurjar et al., 021b). These findings will improve the knowledge of *T. indica*-wheat molecular interaction, and enhance the breeding program for durable Karnal bunt resistance. To our knowledge, there have been only two other pathogens apart from this pathogen that have been studied using genotyping by sequencing for population analysis and genetic variability, both of them were performed on ascomycetes fungi *Magnaporthe oryzae* and *Verticillium dahliae* (Rafiei et al., 2018; Korinsak et al., 2019). Genotyping by sequencing was carried out on several isolates of *V. dahliae* and yielded 26,748 single nucleotide polymorphisms (SNPs), widely distributed throughout the genome, which can be used in specific diagnostic markers development, showing the usefulness of using genotyping by sequencing as a high density genotyping tool (Milgroom et al., 2014). In our study, genotyping of 39 *T. indica* isolates yielded 41,473 filtered, high-quality single nucleotide polymorphism, widely dispersed across the genome, offering more in-depth genotyping and association mapping analyses. This is the first time such a large number of SNPs have been reported in this field of study, proving the utility

of employing genotyping by sequencing in order to carry out GWAS for the identification of virulence gene markers. This investigation shed light on unique gene functions linked to Karnal bunt virulence that would otherwise go unnoticed by standard experimental analyses. Moreover, it might be possible to identify significantly associated genes with cultivar resistance using this method. Further research is still needed to discover how resistance is conditioned in resistant cultivars and the functions of the identified genes in virulence.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NCBI, Sequence Read Archive (SRA) repository, accession number PRJNA796749 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA796749).

AUTHOR CONTRIBUTIONS

MG and MSh contributed to the conception and design of the study. MSh did all the experiments, the result analysis, and wrote the first draft of the manuscript. MSh, MG, RA, RG, NK, MSa, and FK reviewed the manuscript. MG, BB, and FK were in charge of overall direction and planning of research work. SA and TK provided technical assistance. All authors contributed to the manuscript, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.852727/full#supplementary-material

Supplementary Table S1 | Isolates collection location and virulence analysis.

Supplementary Table S2 | Sequencing clean data summary. Number of reads, %GC, Mean Read length.

Supplementary Table S3 | Genes and Description.

REFERENCES

Adhikari, L., and Missaoui, A. M. (2019). Quantitative trait loci mapping of leaf rust resistance in tetraploid alfalfa. *Physiol. Mol. Plant Pathol.* 106, 238–245. doi: 10.1016/j.pmpp.2019.02.006

Aggarwal, R., Singh, D. V., and Srivastava, K. D. (1994). Host-pathogen interaction in karnal bunt of wheat. *Indian Phytopathol.* 47, 381–385.

Aggarwal, R., Tripathi, A., and Yadav, A. (2010). Pathogenic and genetic variability in *Tilletia indica* monosporidial culture lines using universal rice primer-pcr. *Eur. J. Plant Pathol.* 128, 333–342. doi: 10.1007/s10658-010-9655-4

Aoun, M., Kolmer, J. A., Breiland, M., Richards, J., Brueggeman, R. S., Szabo, L. J., et al. (2020). Genotyping-by-sequencing for the study of genetic diversity in Puccinia triticina. *Plant Dis.* 104, 752–760. doi: 10.1094/PDIS-09-19-1890-RE

Aujla, S. S., Sharma, I., and Singh, B. B. (1987). Physiologic specialization of karnal bunt of wheat. *Indian Phytopathol.* 40, 333–336.

Aujla, S. S., Sharma, I., and Singh, B. B. (1989). Rating scale for identifying wheat varieties resistant to Neovossia indica (mitra) mundkur. *Indian Phytopathol.* 42, 161–162.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3, e3376. doi: 10.1371/journal.pone.0003376

Baxter, S. W., Davey, J. W., Johnston, J. S., Shelton, A. M., Heckel, D. G., Jiggins, C. D., et al. (2011). Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE* 6, e19315. doi: 10.1371/journal.pone.0019315

Bello, M. H., Moghaddam, S. M., Massoudi, M., McClean, P. E., Cregan, P. B., and Miklas, P. N. (2014). Application of in silico bulked segregant analysis for rapid development of markers linked to bean common mosaic virus resistance in common bean. *BMC Genomics.* 15, 903–916. doi: 10.1186/1471-2164-15-903

Bonde, M. R., Nester, S. E., Smilanick, J. L., Frederick, R. D., and Schaad, N. W. (1999). Comparison of effects of acidic electrolysed water and naocl on *Tilletia indica* teliospore germination. *Plant Dis.* 83, 627–632. doi: 10.1094/PDIS.1999.83.7.627

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., Bansal, P., Bridge, A. J., et al. (2016). Uniprotkb/swiss-prot, the manually annotated section of the Uniprot knowledgebase: how to use the entry view. *Methods Mol. Biol.* 1374, 23–54. doi: 10.1007/978-1-4939-3167-5_2

Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., and Buckler, E. (2007). Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brisco, E. I., Porch, T. G., Cregan, P. B., and Kelly, J. D. (2014). Quantitative trait loci associated with resistance to Empoasca in common bean. *Crop Sci* 54, 2509–2519. doi: 10.2135/cropsci2014.02.0159

EPPO. (2007). Tilletia indica. *Bull. OEPP/EPPO Bull.* 37, 503–520. doi: 10.1111/j.1365-2338.2007.01158.x

Carlsen, S. A., Neupane, A., Wyatt, N. A., Richards, J. K., Faris, J. D., Xu, S. S., et al. (2017). Characterizing the Pyrenophora teres f. maculata-barley interaction using pathogen genetics. *G3* 7, 2615–2626. doi: 10.1534/g3.117.043265

Chomczynski, P., and Mackey, K. (1995). Short technical reports. modification of the tri reagent procedure for isolation of RNA from polysaccharide-and proteoglycan-rich sources. *Biotechniques* 19, 942–945.

Chutimanitsakun, Y., Nipper, R. W., Cuesta-Marcos, A., Cistue, L., Filichkina, T., Johnson, E. A., et al. (2011). Construction and application for QTL analysis of a restriction site associated DNA (RAD) linkage map in barley. *BMC Genomics.* 12, 4. doi: 10.1186/1471-2164-12-4

Crowhurst, R. N., King, F. Y., Hawthorne, B. T., Sanderson, F. R., and Choi-Pheng, Y. (1995). Rapd characterization of Fusarium oxysporum associated with wilt of angsana (Pterocarpus indicus) in Singapore. *Mycol. Res.* 99, 14–18. doi: 10.1016/S0953-7562(09)80310-9

Earl, D., and vonHoldt, B. (2011). Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379. doi: 10.1371/journal.pone.0019379

Gao, F., Zhang, B. S., Zhao, J. H., Huang, J. F., Jia, P. S., Wang, S., et al. (2019). Deacetylation of chitin oligomers increases virulence in soil-borne fungal pathogens. *Nat. Plants* 5, 1167–1176. doi: 10.1038/s41477-019-0527-4

Gao, Y., Liu, Z., Faris, J. D., Richards, J., Brueggeman, R. S., Li, X., et al. (2016). Validation of genome-wide association studies as a tool to identify virulence factors in *Parastagonospora nodorum*. *Phytopathology* 106, 1177–1185. doi: 10.1094/PHYTO-02-16-0113-FI

Greenbaum, G., Templeton, A., and Bar-David, S. (2016). Inference and analysis of population structure using genetic data and network theory. *Genetics* 202, 1299–1312. doi: 10.1534/genetics.115.182626

Gurjar, M. S., Aggarwal, R., Jain, S., Sharma, S., Singh, J., Gupta, S., et al. (2021a). Multilocus sequence typing and single nucleotide polymorphism analysis in *Tilletia indica* isolates inciting karnal bunt of wheat. *J. Fungi* 7, 103. doi: 10.3390/jof7020103

Gurjar, M. S., Aggarwal, R., Jogawat, A., Kulshreshtha, D., Sharma, S., Solanke, A. U., et al. (2019). De novo genome sequencing and secretome analysis of *Tilletia indica* inciting karnal bunt of wheat provides pathogenesis-related genes. *3 Biotech* 9, 1–11. doi: 10.1007/s13205-019-1743-3

Gurjar, M. S., Mohan, M., H., S., J., S., et al. (2021b). *Tilletia indica*: biology, variability, detection, genomics and future perspective. *Indian Phytopathol.* , 21–31. doi: 10.1007/s42360-021-00319-1

He, Y., Ahmad, D., Zhang, X., Zhang, Y., Wu, L., Jiang, P., et al. (2018). Genome-wide analysis of family-1 UDP glycosyltransferases (UGT) and identification of UGT genes for FHB resistance in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 18, 1–20. doi: 10.1186/s12870-018-1286-5

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C., et al. (2007). Wolf psort: protein localization predictor. *Nucleic Acids Res.* (suppl_2), W585–W587. doi: 10.1093/nar/gkm259

Jones, R. W., and Ospina-Giraldo, M. (2011). Novel cellulose-binding-domain protein in Phytophthora is cell wall localized. *PLoS ONE* 6, e23555. doi: 10.1371/journal.pone.0023555

Kim, J., Zhang, Y., and Pan, W. (2016). Powerful and adaptive testing for multi-trait and multi-SNP associations with gwas and sequencing data. *Genetics*. 203, 715–731. doi: 10.1534/genetics.115.186502

Kolmer, J. A. (2001). Molecular polymorphism and virulence phenotypes of the wheat leaf rust fungus *Puccinia triticina* in Canada. *Can. J. Bot.* 79, 917–926. doi: 10.1139/b01-075

Kolmer, J. A. (2013). Leaf rust of wheat: pathogen biology, variation and host resistance. *Forests* 4, 70–84. doi: 10.3390/f4010070

Kolmer, J. A. (2015). Collections of *Puccinia triticina* in different provinces of China are highly related for virulence and molecular genotype. *Phytopathology* 105, 700–706. doi: 10.1094/PHYTO-11-14-0293-R

Kolmer, J. A., and Acevedo, M. (2016). Genetically divergent types of the wheat leaf rust fungus Puccinia triticina in Ethiopia, a center of tetraploid wheat diversity. *Phytopathology* 106, 380–385. doi: 10.1094/PHYTO-10-15-0247-R

Korinsak, S., Tangphatsornruang, S., Pootakham, W., Wanchana, S., Plabpla, A., Jantasuriyarat, C., et al. (2019). Genome-wide association mapping of virulence gene in rice blast fungus *Magnaporthe oryzae* using a genotyping by sequencing approach. *Genomics* 111, 661–668. doi: 10.1016/j.ygeno.2018.05.011

Krishnan, P., Ma, X., McDonald, B. A., and Brunner, P. C. (2018). Widespread signatures of selection for secreted peptidases in a fungal plant pathogen. *BMC Evol. Biol.* 18, 1. doi: 10.1186/s12862-018-1123-3

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Kumar, J., Shoran, J., Sharma, I., Aggarwal, R., Saharan, M. S., and Sharma, A. K. (2009). Distribution of heterothallic alleles of *Tilletia indica* in india. *Indian Phytopath* 62, 285–294.

Kumar, V., Sanseau, P., Simola, D. F., Hurle, M. R., and Agarwal, P. (2016). Systematic analysis of drug targets confirms expression in disease-relevant tissues. *Sci. Rep.* 6, 1. doi: 10.1038/srep36205

Leboldus, J. M., Kinzer, K., Richards, J., Ya, Z., Yan, C., Friesen, T. L., et al. (2015). Genotype-by-sequencing of the plant-pathogenic fungi *Pyrenophora teres* and *Sphaerulina musiva* utilizing ion torrent sequence technology. *Mol. Plant Pathol.* 16, 623–632. doi: 10.1111/mpp.12214

Ludwig-Müller, J., Jülke, S., Geiß, K., Richter, F., Mithöfer, A., Šola, I., et al. (2015). A novel methyltransferase from the intracellular pathogen *Plasmodiophora brassicae* methylates salicylic acid. *Mol. Plant Pathol.* 16, 349–364. doi: 10.1111/mpp.12185

Ma, X.-F., Jensen, E., Alexandrov, N., Troukhan, M., Zhang, L., Thomas-Jones, S., et al. (2012). High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid *Miscanthus sinensis*. *PLoS ONE* 7, e33821. doi: 10.1371/journal.pone.0033821

Mascher, M., Wu, S., Amand, P. S., Stein, N., and Poland, J. (2013). Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS ONE* 8, e76925. doi: 10.1371/journal.pone.0076925

Mata-Pérez, C., and Spoel, S. H. (2019). Thioredoxin-mediated redox signalling in plant immunity. *Plant Sci.* 279, 27–33. doi: 10.1016/j.plantsci.2018.05.001

Matos, R. G., Barria, C., Moreira, R. N., Barahona, S., Domingues, S., and Arraiano, C. M. (2014). The importance of proteins of the RNase II/RNB-family in pathogenic bacteria. *Front. Cell. Infect. Microbiol.* 4, 68. doi: 10.3389/fcimb.2014.00068

Milgroom, M. G., Jimenez-Gasco, M, d, M., Olivares García, C., Drott, M. T., and Jimenez-Diaz, R. M. (2014). Recombination between clonal lineages of the asexual fungus *Verticillium dahliae* detected by genotyping by sequencing. *PLoS ONE* 9, e106740. doi: 10.1371/journal.pone.0106740

Moghaddam, S., Mamidi, S., Osorno, J., Lee, R., Brick, M., Kelly, J., et al. (2016). Genome-wide association study identifies candidate loci underlying seven agronomic traits in middle american diversity panel in common bean (*Phaseolus vulgaris* L.). *Plant Genome* 9, 1–21. doi: 10.3835/plantgenome2016.02.0012

Mukeshimana, G., Butare, L., Cregan, P. B., Blair, M. W., and Kelly, J. D. (2014). Quantitative trait loci associated with drought tolerance in common bean. *Crop Sci.* 54, 923–938. doi: 10.2135/cropsci2013.06.0427

Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). Signalp 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7, e37135. doi: 10.1371/journal.pone.0037135

Pfender, W., Saha, M., Johnson, E., and Slabaugh, M. (2011). Mapping with RAD (restriction-site associated dna) markers to rapidly identify QTL for stem rust resistance in *Lolium perenne*. *Theor. Appl. Genet.* 122, 1467–1480. doi: 10.1007/s00122-011-1546-3

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012a). Genomic selection in wheat breeding using genotyping-by sequencing. *Plant Gene* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012b). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7, e32253. doi: 10.1371/journal.pone.0032253

Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5, 95–102. doi: 10.3835/plantgenome2012.05.0005

Puritz, J. B., Hollenbeck, C. M., and Gold, J. R. (2014). ddocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ.* 2, e431. doi: 10.7717/peerj.431

Rafiei, V., Banihashemi, Z., Bautista-Jalon, L. S., del Mar Jiménez-Gasco, M., Turgeon, B. G., and Milgroom, M. G. (2018). Population genetics of *Verticillium dahliae* in iran based on microsatellite and single nucleotide polymorphism markers. *Phytopathology* 108, 780–788. doi: 10.1094/PHYTO-11-17-0360-R

Rajalakshmi, R., Rajalakshmi, S., and Ajay, P. (2017). Evaluation of the genetic diversity and population structure in drumstick (*Moringa oleifera* L.) using SSR markers. *Curr. Sci.* 112, 1250–1256. doi: 10.18520/cs/v112/i06/1250-1256

Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., and Naik, V. G. (2014). Structure plot: a program for drawing elegant structure bar plots in user friendly interface. *Springerplus* 3, 1–3. doi: 10.1186/2193-1801-3-431

Rao, X., Huang, X., Zhou, Z., and Lin, X. (2013). An improvement of the 2 (-delta delta ct) method for quantitative real-time polymerase chain reaction data analysis. *Biostat. Bioinforma. Biomath.* 3, 71.

Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., et al. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475, 348–352. doi: 10.1038/nature10242

Saleh, D., Milazzo, J., Adreit, H., Fournier, E., and Tharreau, D. (2014). South-east asia is the center of origin diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae. New Phytol.* 201, 1440–1456. doi: 10.1111/nph.12627

San Koh, C., and Sarin, L. P. (2018). Transfer rna modification and infection-implications for pathogenicity and host responses. *Biochim. Biophys. Acta* 1861, 419–432. doi: 10.1016/j.bbagrm.2018.01.015

Schober, P., and Vetter, T. R. (2019). Two-sample unpaired t tests in medical research. *Anesth. Analgesia* 129, 911. doi: 10.1213/ANE.00000000000 04373

Seifbarghi, S., Borhan, M. H., Wei, Y., Coutu, C., Robinson, S. J., and Hegedus, D. D. (2017). Changes in the Sclerotinia sclerotiorum transcriptome during infection of Brassica napus. *BMC Genomics* 18, 1. doi: 10.1186/s12864-017-3642-5

Singh, D. V. (2005). Karnal bunt of wheat: a global perspective. *Indian Phytopathol.* 58, 1–9.

Singh, J., Aggarwal, R., Gurjar, M., Sharma, S., Jain, S., and Saharan, M. (2020). Identification and expression analysis of pathogenicity-related genes in *Tilletia indica* inciting karnal bunt of wheat. *Austr. Plant Pathol.* 49, 393–402. doi: 10.1007/s13313-020-00711-x

Song, Q., Jia, G., Hyten, D. L., Jenkins, J., Hwang, E.-Y., Schroeder, S. G., et al. (2015). SNP assay development for linkage map construction, anchoring whole-genome sequence, and other genetic and genomic applications in common bean. *G3* 5, 2285–2290. doi: 10.1534/g3.115.0 20594

Soto-Cerda, B. J., and Cloutier, S. (2012). Association mapping in plant genomes. *Genetic Diversity in Plants, pages* 29–54.

R Core Team (2013). *R: A Language and Environment for Statistical Computing.* Vienna: R Core Team.

Warham, E. J. (1986). Karnal bunt disease of wheat: a literature review. *Trop. Pest Manag.* 32, 229–242. doi: 10.1080/09670878609 371068

Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., et al. (2014). The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42, D1001–D1006. doi: 10.1093/nar/gkt1229

Xie, G., Liu, M., Zhu, H., and Lei, B. (2008). Esterase see of Streptococcus equissp.equiis a novel nonspecific carboxylic ester hydrolase. *FEMS Microbiol. Lett.* 289, 181–186. doi: 10.1111/j.1574-6968.2008.01377.x

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed. Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zhao, C., Rispe, C., and Nabity, P. D. (2019). Secretory ring finger proteins function as effectors in a grapevine galling insect. *BMC Genomics* 20, 1. doi: 10.1186/s12864-019-6313-x