Genetic Characterization and Genome-Wide Association Mapping for Stem Rust Resistance in Spring Bread Wheat

Elias Shewabez (elias.shewabez@gmail.com)
Addis Ababa University Faculty of Science: Addis Ababa University College of Natural Sciences
https://orcid.org/0000-0001-7184-6554
Endashaw Bekele
Addis Ababa University Faculty of Science: Addis Ababa University College of Natural Sciences
Admas Alemu
Swedish University of Agricultural Sciences University Library Uppsala: Sveriges lantbruksuniversitet Biblioteket Uppsala
Laura Mugnai
University of Florence Faculty of Agriculture: Universita degli Studi di Firenze Scuola di Agraria
Wuletaw Tadesse
ICARDA: International Center for Agricultural Research in the Dry Areas

Research article

Keywords: markers, Puccinia graminis f. sp. tritici, QTL

Posted Date: January 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-147919/v1

Version of Record: A version of this preprint was published at BMC Genomic Data on February 14th, 2022. See the published version at https://doi.org/10.1186/s12863-022-01030-4.

Genetic characterization and genome-wide association mapping for stem rust resistance in spring bread wheat

Elias Shewabez[a,*], Endashaw Bekele[a], Admas Alemu[b], Laura Mugnai[c], Wuletaw Tadesse[d]

[a]Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

[b]Department of Biology, Debre Tabor University, P.O. Box 272. Debre Tabor, Ethiopia

[c]Department of Agriculture, Food, Environment and Forestry, University of Florence, Piazzale delle Cascine 18 - 50144 Firenze (FI), Italy

[d]International Center for Agricultural Research in the Dry Areas (ICARDA), P.O. Box 6299, Rabat, Morocco

[*]**Corresponding Author**: Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, P.O. Box 1176, Addis Ababa,Ethiopia.

ABSTRACT

Background

Emerging wheat stem rust races has become a major threat to global wheat production. Finding additional loci responsible for resistance to these races and incorporating them into currently cultivated varieties is the most economical and environmentally sound strategy to combat this problem. Thus, this study aimed to characterize the genetic diversity of wheat and to identify the genetic loci conferring resistance to stem rust of wheat. To accomplish this study, 245 elite lines introduced from the International Center for Agricultural Research in the Dry Areas (ICARDA) were tested under natural stem rust pressure in the field at the Debre Zeit Agricultural Research Center, Ethiopia. The SNP marker data was retrieved from a 15K SNP wheat array. Association analysis was undertaken between SNP markers and best linear unbiased prediction (BLUP) value of the stem rust coefficient of infection (CI) using a mixed linear model.

Results

Phenotypic analysis revealed 46% of lines had a coefficient of infection (CI) between 0 to 19. An average 0.38 in Nei's gene diversity, 0.20 in polymorphism information content, and 0.71 in major allele frequency of the whole genome were identified.A total of 46 marker-trait

associations (MTAs) that were encompassed within 13 quantitative trait loci (QTL) on chromosomes 1B, 3A, 3B, 4A, 4B, and 5A were found for CI. Four major QTLs with $-\log_{10}$ (p) ≥ 3 (EWYP1B.1, EWYP1B.3, EWYP1B.4, and EWYP1B.5) were identified on chromosome 1B.

Conclusions

This study contributes several novel markers associated with stem rust resistance. These can be further facilitating durable rust resistance development through marker-assisted selection. The resistant wheat genotypes identified in this study are recommended to be used in the national wheat breeding programs to improve stem rust resistance.

Key words: markers; *Puccinia graminis* f. sp. *tritici*; QTL

Background

Wheat (*Triticum aestivum* L.*)* is a leading crop in economic value and area of production worldwide [1][2]that provides nearly 20% of daily world human caloric requirements [3]. Out of the total world wheat imports, nearly 77% are by developing countries [4]. Demand is expected to increase by 60 percent by year 2050 [5]. Challenges to meet this demand are negative impacts due to: climatic change, drought, soil degradation, increasing cost of fertilizer, increasing demand for biofuel, and the occurrences of new and emerging diseases and pests. Of these factors, emerging fungal pathogens are the most widespread and economically important diseases of wheat worldwide [6].

Wheat stem (black) rust, caused by *Puccinia graminis* Pers. f. sp. *tritici*, Eriks. & E. Henn (*Pgt*), has been recognized as a major threat to global food security (Gao et al., 2017; Li et al., 2019). The concern of this disease has highly increased after 1998 once the outbreak of a novel virulent race known as Ug99 had emerged from Uganda. Since then, the race produced 13 variants that are spreading in East Africa [9] [10]. This race can infect 90% of the wheat varieties grown worldwide [11] and yield losses can reach up to 100% in susceptible cultivars under conducive environmental conditions [12]. Races other than Ug99 were also reported in different parts of Western Europe. In 2013, a the stem rust epidemic arose in Germany and spread to Denmark, Sweden, and the UK [13]. In 2016/2017, Italy chronicled two epidemics of wheat stem rust caused by race TTRTF, which destroyed tens of thousands of hectare of cultivated wheat[14]. All these reports indicats that the disease is re-emerging as a threat to wheat production.

Ethiopia is a "hot-spot" area for the development and evolution of new *Pgt* races [15]. Many new variants of *Pgt* were first identified in this country and spread to different parts of the world. Rust of wheat, races TTKSK, TKTTF, TRTTF, JRCQC, and TTTTF are the stem rust of wheat races that are threatening wheat production currently in Ethiopia[16]. In 2013/2014, severe stem rust epidemics were caused by *Pgt* race TKTTF (not a member of Ug99 lineage) resulting in nearly total yield loss on widely grown wheat cultivars. Since then, this race spread widely and was found in 10 different countries including western Europe[17].

To overcome this problem, host plant resistance developed through molecular marker technology is the most sustainable, cost-effective and environmentally friendly approach for controlling rust diseases [7][18]. Accordingly, many molecular markers linked with *Pgt* resistance were discovered throughout the wheat genome during the past couple of decades using genome-wide association mapping (GWAS). Crossa et al. (2007) were the first researcher who identified many *Pgt* resistant quantitative trait loci (QTL) and yield-related loci[19]. Following their study, many other related studies have been undertaken [20][21][22][23]. More than 80 genes conferring resistances to *Pgt* have been cataloged in common wheat and wheat relatives[23]. However, few genes are effective against all pathogen strains. Rather, frequent co-evolution of host and

pathogen remains a big challenge in the durability of the released resistant cultivars [24]. The narrow genetic diversity of currently cultivated wheat cultivars [25][21] and climate change [12] are the major contributors to this problem. Thus, continuous search for additional source of resistant genes and marker-assisted gene pyramiding is indispensable to produce durable resistant varieties. Therefore, the aims of this study were to characterize the genetic diversity of wheat and to identify novel loci associated with resistance to stem rust of wheat through GWAS.

Results

1.1. Phenotypic variability and estimates of heritability

Adult plant stage disease phenotyping was scored for two parameters, (DS and IR) across two cropping seasons (2018 and 2019). Their phenotypic performances varied from 10 to 80 in DS, and RMR to S for IR, with 3 to 80 in CI. Among the lines, 21 (8.5%) accessions had DS = 0% to 15%, 115 (46.7%) accessions with DS = 15 to 30%, and 110 (44.8%) with DS = > 30% (Table 1, Fig 1A). Likewise, best linear unbiased estimates (BLUP) values of DS, IR, and CI were calculated from adjusted means of each accession across two years, and they were summarized in Figure 1.

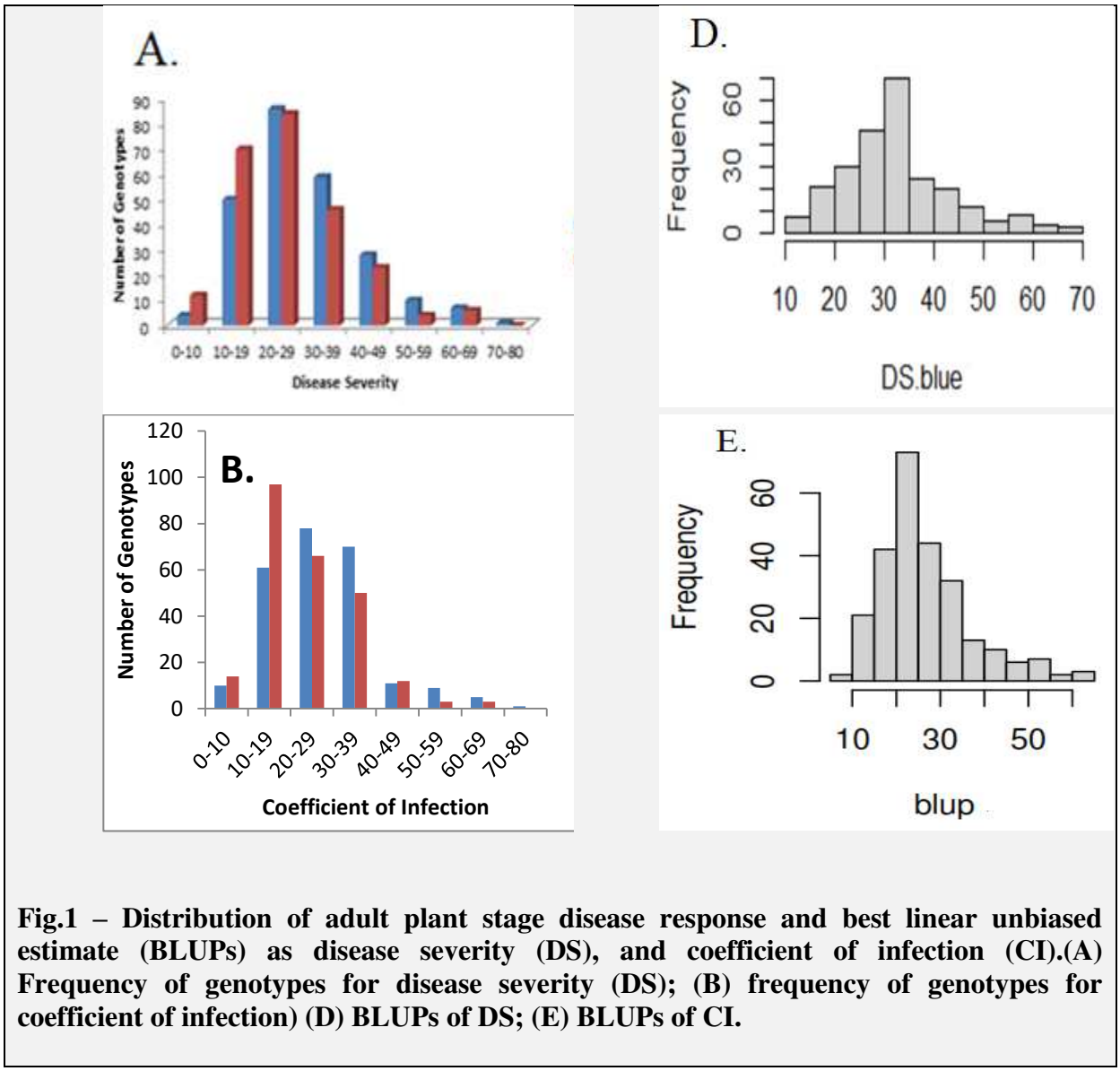

Fig.1 – Distribution of adult plant stage disease response and best linear unbiased estimate (BLUPs) as disease severity (DS), and coefficient of infection (CI).(A) Frequency of genotypes for disease severity (DS); (B) frequency of genotypes for coefficient of infection) (D) BLUPs of DS; (E) BLUPs of CI.

The data of DS and IR were combined to define the disease response as the coefficient of infection (CI). Nearly 71% of lines had less than 30 (Fig 1B).). Of these, the top twenty resistance lines were presented in Table 1, and theythat ranged between the average CI values of 4.5% for pedigree SERI.1B//KAUZ/HEVO/3/AMAD/4/CHAM-6/FLORKWA-2 to 12% for pedigree SERI.1B//KAUZ/HEVO/3/AMAD/4/WEAVER/JACANA. Additional genotypes scored between 6% to 80% of CI were also presented in additional file 1. On the other hand, all local controls (Digelu, Kubssa, Hidasse, Honqolo, and Ogolcho) were susceptible with average CI ranging from 60% for HIDASSE to 80% for OGOLCHO and HONQOLO.

Table 1. Lists of top resistant lines and their pedigree during 2017/2018 main season at Debre-Zeit Agricultural Research Center, Ethiopia

No	Pedigree	Disease Severity and response to Sr	
		2018	2019
1	SERI.1B//KAUZ/HEVO/3/AMAD/4/CHAM-6/FLORKWA-2	10RMR	15MR
2	SERI.1B//KAUZ/HEVO/3/AMAD/4/MO88/MILAN	15MR	10MR
3	SERI.1B//KAUZ/HEVO/3/AMAD/4/TNMU/MILAN/5/WATAN-12	15MR	10MR
4	PBW343*2/KUKUN//22SAWSN - 97	10MRMS	15MR
5	SERI.1B//KAUZ/HEVO/3/AMAD/4/ESDA/SHWA//BCN	10MR	10MR
6	SERI.1B*2/3/KAUZ*2/BOW//KAUZ/4/SHIHAB-7	10MR	10MR
7	CROC-1/AE.SQUARROSA (224)//OPATA/3/FLAG-7	10MR	10MR
8	TRACHA-2/SHUHA-3/3/SHUHA-7//SERI 82/SHUHA'S'	15MRMS	10MRMS
9	SERI.1B//KAUZ/HEVO/3/AMAD/4/PFAU/MILAN	15MR	10MR
10	WATAN-7/SEKHRAH-2	10MR	15MRMS
11	WEAVER/WL 3928//SW 89.3064/3/SOMAMA-3	15MRMS	10MR
12	SERI.1B//KAUZ/HEVO/3/AMAD/4/SHUHA-7//SERI 82/SHUHA'S'	15MR	15MS
13	KAUZ'S'/SERI/3/TEVEE'S'//CROW/VEE'S'	15MRMS	15MR
14	ATTILA*2/CROW/3/VEE#5/SARA//DUCULA	15MR	15MR
15	TILILA/MUBASHIIR-1	15MR	15MR
16	QAFZAH-27/SEKSAKA-6	15MR	15MR
17	SERI.1B*2/3/KAUZ*2/BOW//KAUZ/4/SHIHAB-7	15MR	15MR
18	STAR*3/LOTUS-5/3/CHUM//7*BCN/4/FLAG-2	15MR	10MR
19	HADIAH-14/ANGI-2	10MRMS	15MRMS
20	SERI.1B//KAUZ/HEVO/3/AMAD/4/WEAVER/JACANA	15MR	15MR

Results of ANOVA revealed highly significant variation among genotypes (P <0.001) and genotype x year interactions (P <0.001) for all parameters. Heritability (H) values for DS were (79%), and IR (72%), suggesting that all parameters had a strong genetic basis (Table 2).

Table 2 – Mean response, variance component estimates and heritability for IR, DS, and CI variables

Subject	DS (%)	IR(0-1)	CI
Range	10-80	0.3-1	3-80
Mean	33.90	0.8	28.27
BLUEs	32.60	0.8	26.902
σ^2_G	385.3***	0.013***	374.5***
σ^2_E	21.05*	0.000^{ns}	15.7^{ns}
σ^2_{GxE}	20.70**	0.005^{ns}	24.1**
σ^2 error	166.10	0.004	200.3
H	78.95	72.2	75.7
CV	37.93	8.52	50.06

Disease severity (DS); infection response (IR); coefficient of infection (CI); BLUEs, best linear unbiased estimate; σ^2_G estimate of genotypic variance; σ^2_E estimate of environmental variance; σ^2_{GxE} is the genotype by environment interaction variance, σ^2 error is the residual error variance; heritability (H); *, **, ***, ns, significance at P < 0.05, P < 0.01, P < 0.001, and not significant, respectively.

Population structure was inferred through Bayesian clustering model, principal component analysis (PCA), and neighbor-joining (NJ) tree. Bayesian clustering model applied on STRUCTURE software and subsequent application of STRUCTURE HARVESTER showed a delta K peak value of two (Fig. 2A). As a result, accessions were classified into two subpopulations composed of 106 lines in subpopulation 1 and 139 lines in subpopulation 2 (Fig.2B). The scree plot of PCA showed that weak kinship existed among the lines. For the first 10 principal components (PC), variances of SNP markers were changed from 7.5% (PC1) to 2% (PC10) while for the PC after PC10, the variances were between 2% and 0% (Fig 2C).

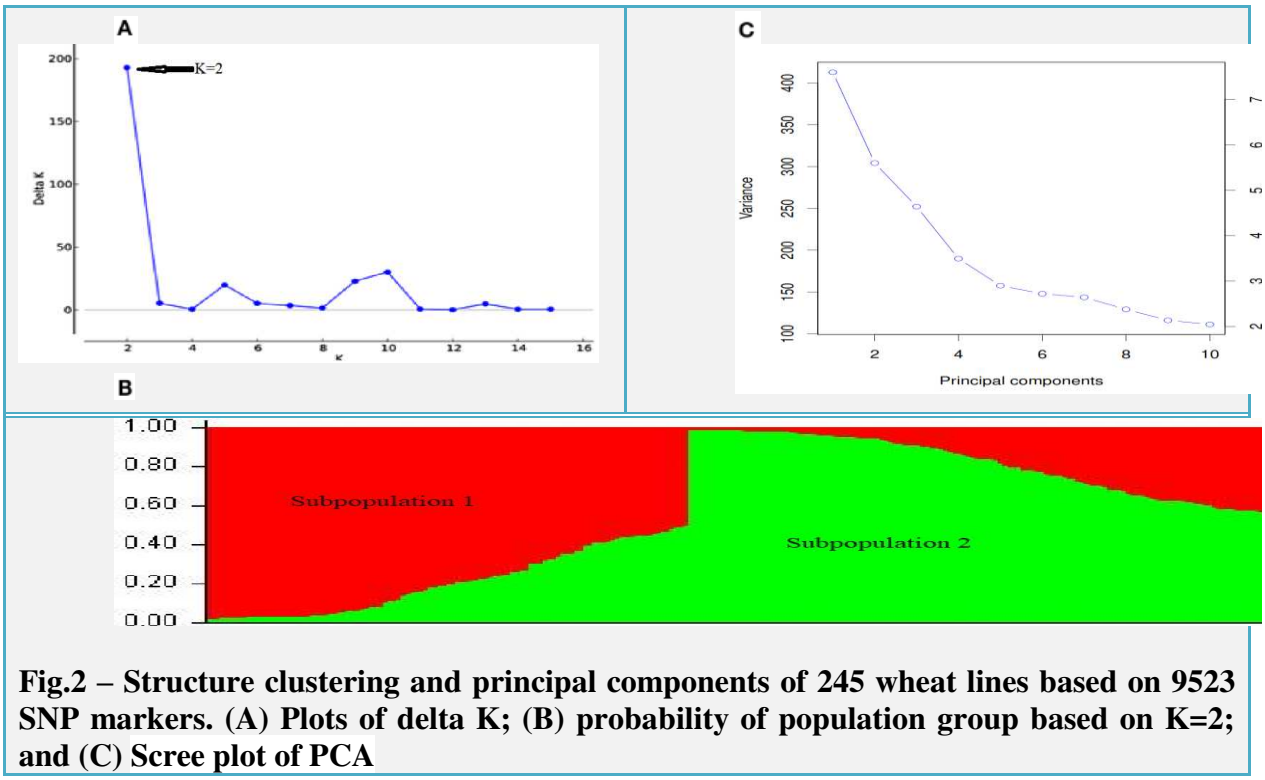

Fig.2 – Structure clustering and principal components of 245 wheat lines based on 9523 SNP markers. (A) Plots of delta K; (B) probability of population group based on K=2; and (C) Scree plot of PCA

Phylogenetic tree analysis of the genetic relationship between the populations was carried out based on distance-based neighbor-joining tree on TASSEL software v5.2.35 followed by web-based visualization software iTOL. The resulting dendrogram showing three phylogenetic groups is color-coded with a STRUCTURE probability distribution. As a result, 78 (56%) of the lines in the first group was composed of a subpopulation 1, whereas 61 lines (44%) were categorized in subpopulation 2. The second group was composed mainly of from subpopulation 2 which consists of49(70%) lines; whereas 21(30%) lines were classified in subpopulation 1. The third

group was composed of 58 (76%) lines from subpopulation 2 and 21(30%) lines from subpopulation 1(Fig. 3).

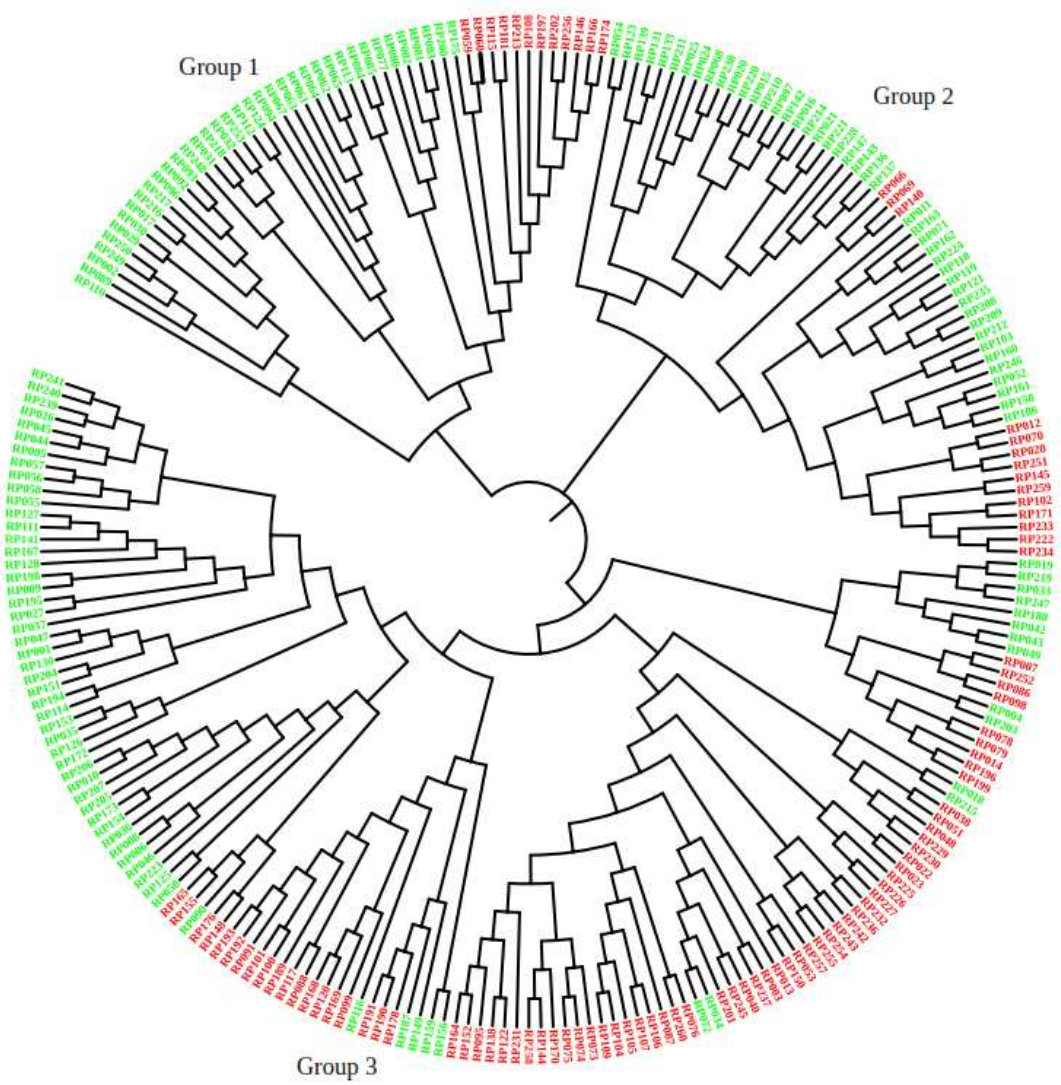

Fig. 3 – Neighbor joining tree based on Nei's genetic distance color coded with STRUCTURE probability distribution

1.2. Genetic data and linkage disequilibrium

After filtering, 9523 SNP markers were retained from 245 lines. Of these, 50% were located on the B sub-genome, 39% on A sub-genome, and 11% were on D sub-genome. The maximum number (930) of SNP markers was observed on chromosome 2B and the minimum number (48) SNP markers were observed on the 4D chromosome (Fig 4D). Major allele frequency, heterozygosity, genome-wide polymorphic information content (PIC) and gene diversity ranged

from 0.5020 to 0.9469 , 0.0 to 0.5673, 0.01 to 0.7034, and 0.1005 to 0.7463 with average value of 0.7150 and 0.0064, 0.2030 and 0.3835, respectively. The result also showed that 7142 (75%) of SNP markers were moderately informative with PIC values between 0.25 and 0.50, 2285 (24%) SNP markers were least informative with PICs value of less than 0.2, and only 90 (1%) SNP markers were highly informative with PICs value greater than 0.50.

Linkage disequilibrium decay value-based on SNP markers of each chromosome was calculated as the Pearson correlation coefficient (r^2) between marker pairs as a function of genetic distance (cM). In consequence, the LOESS curve intercepted the line of critical value at 6 cM in A genome, at 8 cM in B genome, and at 5 cM in D genome indicating that all markers within these ranges were considered as a single locus (Fig 4A, 4B, 4C).

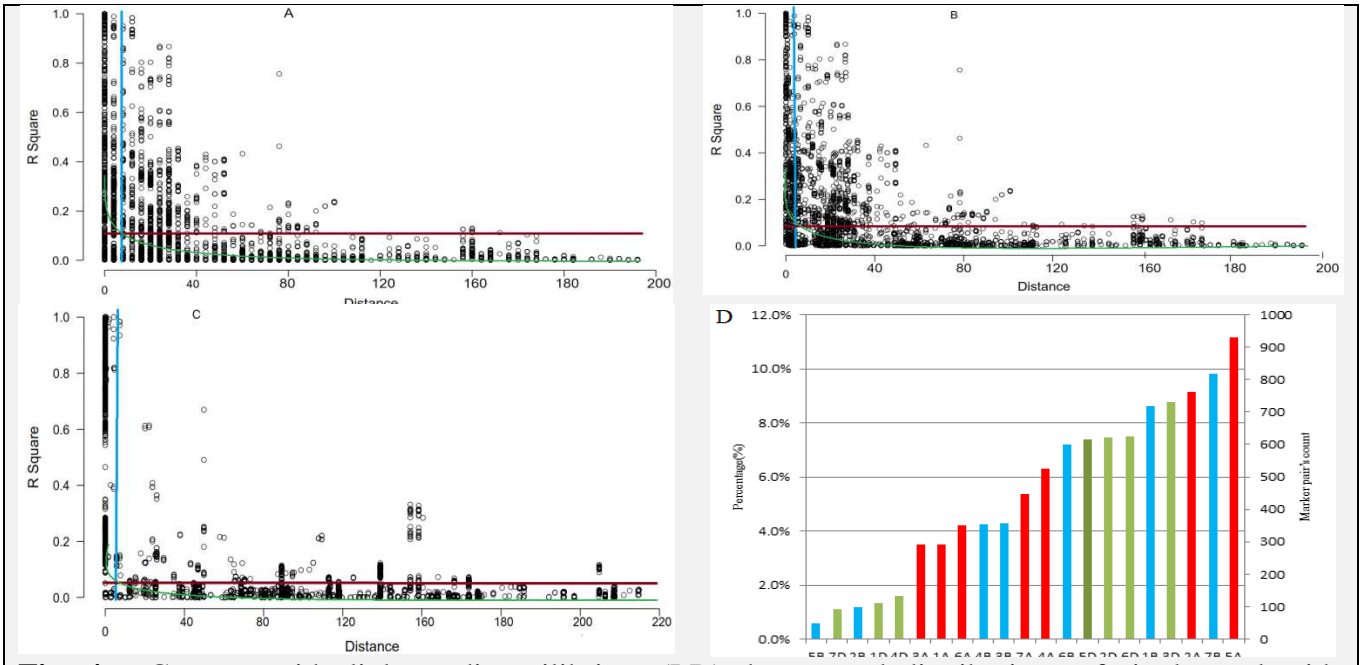

Fig. 4 – Genome wide linkage disequilibrium (LD) decays and distributions of single nucleotide polymorphisms (SNPs) based on 15K genotyping results. Intra-chromosomal pairs in the genetic distance (cM) are plotted against the LD estimate (R^2) for pairs of single-nucleotide polymorphisms.

1.3. Marker-trait associations for stem rust

In this study, a mixed linear model (MLM) that uses genetic marker-based kinship matrix (K) jointly with population structure in the form of principal component analysis (PCA+K) was used to detect marker-trait associations (MTAs). The association was conducted between BLUP values of coefficient of infection (CI) and SNP marker data retained after removing monomorphic markers (missing rate ≥ 10% and MAF > 0.05). This approach gave a minimum

deviation of the observed p-values from the expected values presented as quantile-quantile (Q-Q) plots, and Manhattan plots based on a mixed linear model (Fig. 5).

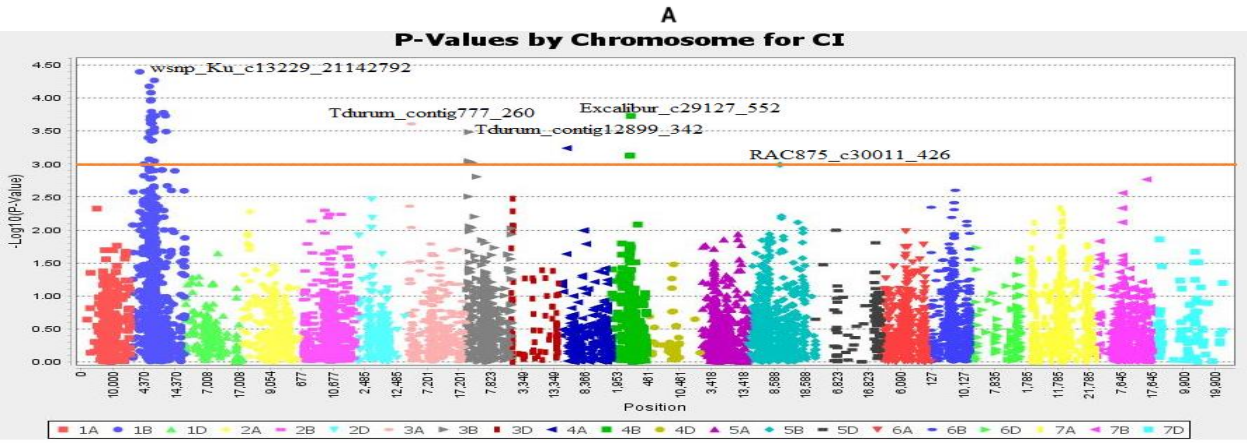

Fig. 5 – Manhattan plots based on the mixed linear model along each of 21 chromosomes showing significantly associated markers. The dots above solid orange horizontal lines represent genome-wide significant markers for threshold p value of 0.001 (MTAs at $-\log_{10} p \geq 3$). The position of each marker was based on the wheat consensus SNP map (Wang et al., 2014).

Based on significant threshold ($-\log_{10}(p) \geq 3$), 18 MTAs were identified using MLM. . In our study, all significant markers were identified from B sub-genome. The highest number of MTAs was located on 1B sub-genome (36 MTAs: 7 loci) accounting for 78% of the total detected significant SNP markers followed by chromosome 3B (4 MTAs, 2 loci) and chromosome 4B (3 MTAs: 1 loci). The remaining three markers namely, (Tdurum_contig777_260, EWYP3A), (Tdurum_contig59603_74, EWYP4A), and (RAC875_c30011_426, EWYP5B) were identified on 3A chromosome at (20.74 cM), 4A chromosome at 26.5 cM and 5B chromosome at 104.55 cM, respectively (Table 3, Fig. 5). The three uppermost QTLs explained phenotypic variance (PV) were EWYP1B.4 at 68.04 cM (8.8%), EWYP1B.5 at 76.89 cM (8.37 %) and EWYP1B.3 at 60.62 cM (8.16%). Other QTLs that accounted for PV ranging from 7.48 % to 4.76 % were comprised EWYP1B.1 (30.34 cM), EWYP1B.2 (43.86 cM), EWYP1B.7 (112.07 cM), EWYP3A (20.74 cM), EWYP4A (26.5 cM), EWYP3B.1 (9.7 cM), EWYP3B.2 (20.14 cM), EWYP5B (104.55 cM), EWYP4B (62.22 cM) (Table 3, Fig. 5). Details of markers distribution in each accession are presented on additional file 1.

Table 3 – Lists of identified SNPs/QTLs for adult plant resistance (APR) to wheat stem rust

QTLs	*SNPs*	Markers name	*Allele*	*Chr*[a]	*Pos*[b]	*MAF*[c]	*F*	$-\log_{10} P$	$R^2 x100$
EWYP1B.1	*IWA6489.1*	*wsnp_Ku_c13229_21142792*	*T/C*	*1B*	*30.34*	*0.213*	*0.671*	*4.398*	*7.14*
EWYPIB.2	*IWB29508*	*Excalibur_c95327_51*	*A/G*	*1B*	*43.86*	*0.419*	*0.378*	*3*	*5.91*
EWYP1B.3	*IWB29475*	*Excalibur_c94756_540*	*T/C*	*1B*	*57.6*	*0.415*	*0.382*	*3.399*	*6.63*
	IWB1569	*BobWhite_c22266_315*	*C/T*	*1B*	*60.62*	*0.213*	*0.671*	*4.178*	*8.16*
	IWB43258	*Kukri_c26168_423*	*C/A*	*1B*	*60.62*	*0.462*	*0.335*	*3.072*	*6.06*
	IWB23446	*Excalibur_c20228_135*	*A/G*	*1B*	*64.1*	*0.095*	*1.023*	*3.717*	*7.27*
	IWB54643	*RAC875_c18282_1390*	*T/C*	*1B*	*64.1*	*0.091*	*1.041*	*3.676*	*7.15*
	IWA5592	*wsnp_Ex_rep_c69266_68192766*	*T/C*	*1B*	*64.1*	*0.095*	*1.023*	*3.668*	*7.13*
	IWB6405	*BS00011450_51*	*T/C*	*1B*	*64.1*	*0.138*	*0.859*	*3.473*	*6.74*
	IWA131	*wsnp_BE443531B_Ta_1_1*	*C/T*	*1B*	*64.32*	*0.115*	*0.941*	*3.49*	*6.77*
	IWB48689	*Kukri_rep_c101799_95*	*C/A*	*1B*	*64.46*	*0.115*	*0.941*	*3.49*	*6.77*
	IWA3631	*wsnp_Ex_c38116_45719983*	*C/T*	*1B*	*64.89*	*0.079*	*1.102*	*3.958*	*7.72*
	IWB51549	*Ra_c23839_884*	*T/C*	*1B*	*65.42*	*0.146*	*0.835*	*3.614*	*7.02*
	IWB31732	*GENE-0193_197*	*A/G*	*1B*	*65.42*	*0.146*	*0.835*	*3.472*	*6.74*
	IWB58051	*RAC875_c44575_561*	*C/T*	*1B*	*66.07*	*0.075*	*1.124*	*3.974*	*8.4*
	IWA6890	*wsnp_Ku_c30982_40765254*	*T/G*	*1B*	*66.07*	*0.075*	*1.124*	*4.082*	*7.97*
EWYP1B.4	*IWB461*	*BobWhite_c1318_691*	*T/C*	*1B*	*66.73*	*0.15*	*0.823*	*3.618*	*7.03*
	IWB37720	*JD_c64600_281*	*G/A*	*1B*	*67.14*	*0.893*	*0.049*	*3.718*	*7.48*
	IWA106	*wsnp_BE442716B_Ta_2_1*	*T/G*	*1B*	*67.14*	*0.103*	*0.988*	*3.68*	*7.15*
	IWB38394	*Ku_c13515_171*	*C/T*	*1B*	*67.14*	*0.103*	*0.988*	*3.68*	*7.15*
	IWB35871	*IACX2701*	*C/T*	*1B*	*67.38*	*0.099*	*1.005*	*3.796*	*7.43*
	IWB56778	*RAC875_c32894_1038*	*C/T*	*1B*	*68.04*	*0.15*	*0.823*	*3.354*	*8.83*
	IWB10444	*BS00070139_51*	*A/C*	*1B*	*68.04*	*0.15*	*0.823*	*3.618*	*7.03*
	IWB27852	*Excalibur_c59016_839*	*A/G*	*1B*	*68.04*	*0.15*	*0.823*	*3.618*	*7.03*
	IWB59327	*RAC875_c5796_424*	*A/G*	*1B*	*68.04*	*0.15*	*0.823*	*3.618*	*7.03*
	IWB8148	*BS00038929_51*	*T/C*	*1B*	*68.04*	*0.15*	*0.823*	*3.618*	*7.03*
	IWB74145	*tplb0023b14_704*	*G/T*	*1B*	*70.08*	*0.34*	*0.469*	*3.357*	*6.59*
	IWB47566	*Kukri_c73734_175*	*C/T*	*1B*	*76.89*	*0.213*	*0.671*	*4.268*	*8.37*
EWYP1B.5	*IWB60433*	*RAC875_c7674_634*	*G/A*	*1B*	*76.89*	*0.277*	*0.558*	*3.546*	*7.4*
	IWB6504	*BS00011973_51*	*T/G*	*1B*	*76.89*	*0.281*	*0.552*	*3.485*	*6.78*
	IWA775	*wsnp_CAP11_c543_375403*	*A/G*	*1B*	*76.89*	*0.217*	*0.663*	*3.045*	*5.89*
	IWA5228	*wsnp_Ex_rep_c66389_64588992*	*A/G*	*1B*	*79.77*	*0.111*	*0.956*	*3.731*	*7.28*
	IWB74900	*tplb0048b10_1365*	*A/G*	*1B*	*79.77*	*0.111*	*0.956*	*3.731*	*7.28*
	IWB10621	*BS00072791_51*	*A/G*	*1B*	*105.83*	*0.277*	*0.558*	*3.779*	*7.61*
EWYP1B.6	*IWB70380*	*Tdurum_contig32775_78*	*A/G*	*1B*	*112.07*	*0.336*	*0.474*	*3.732*	*7.26*
EWYP1B.7	*IWB66198*	*Tdurum_contig10036_977*	*C/A*	*1B*	*114.13*	*0.095*	*1.023*	*3.492*	*6.81*
EWYP3A	*IWB73429*	*Tdurum_contig777_260*	*G/A*	*3A*	*20.74*	*0.34*	*0.469*	*3.609*	*7.12*
EWYP3B.1	*IWB67769*	*Tdurum_contig12899_342*	*T/C*	*3B*	*9.7*	*0.379*	*0.421*	*3.484*	*6.9*
	IWB23457	*Excalibur_c20277_483*	*A/G*	*3B*	*9.7*	*0.372*	*0.43*	*3.042*	*5.88*
	IWB67389	*Tdurum_contig12008_803*	*T/C*	*3B*	*9.7*	*0.395*	*0.403*	*3.017*	*5.83*
EWYP3B.2	*IWB75222*	*tplb0059m03_622*	*C/T*	*3B*	*20.14*	*0.435*	*0.362*	*3.022*	*5.95*
EWYP4A	*IWB72664*	*Tdurum_contig59603_74*	*G/A*	*4A*	*26.5*	*0.15*	*0.823*	*3.24*	*6.4*
EWYP4B	*IWB48189*	*Kukri_c8973_1986*	*T/C*	*4B*	*62.22*	*0.099*	*1.005*	*3.126*	*4.76*
	IWB53758	*RAC875_c13639_2159*	*T/C*	*4B*	*62.92*	*0.099*	*1.005*	*3.126*	*4.76*
	IWB24798	*Excalibur_c29127_552*	*G/A*	*4B*	*64.58*	*0.095*	*1.023*	*3.73*	*7.26*
EWYP5B	*IWB56412*	*RAC875_c30011_426*	*C/T*	*5B*	*104.55*	*7.085*	*0.001*	*2.991*	*5.77*

QTLs , Quantitative Trait Loci; *SNPs* , Single Nucleotide Polymorphisms; *Chr*[a] , Chromosome;

Pos[b] , Allele Position (Wang et al., 2014); *MAF*[c] , Minor Allele Frequency

Discussion

Stem rust of wheat disease has been increasing in severity and incidence and now poses a serious threat to global wheat production [8]**.** To overcome this threat, efforts are going on worldwide to monitor rust diseases, identify rust pathotypes, and to evaluate wheat germplasm for rust resistance [26]. As part of this global effort, this study was designed to characterize the genetic diversity of breeding lines and to search for new sources of resistant genes for *Pgt* races. To this end, 245 elite bread wheat lines were evaluated in the field to identify known and/or unknown resistance (R) genes to wheat stem rust. Considerable variations in disease response between susceptible controls and elite lines were observed. In addition, the study revealed that potentially novel resistance QTL were effective against *Pgt* races in 2018/2019 cropping season in Ethiopia. Therefore, the existing lines could be potential sources to identifying previously uncharacterized resistance genes with minor effects for durable stem rust resistance in the wheat breeding programs.

Field evaluation of wheat germplasm for resistance to stem rust

Disease response characterization under high disease pressure in field conditions remain the best stem rust management strategy in breeding for developing stem rust-resistant cultivars [27]. Ethiopia is considered as a hot-spot area for development of *Pgt* race diversity and frequent disease epidemics. Studies carried out in Ethiopia showed that most previously identified races such as TTKSK, TKTTF, TTTTF, TRTTF, RRTTF, and others were virulent on most varieties grown in the country[28]. Accordingly, many field evaluation studies for *Pgt* response have been carried out in different wheat growing-regions of the country [21][29][30]. In the current study, two phenotypic parameters, namely DS and IR were considered. Most elite breeding lines skewed toward moderate resistance although some differences were observed between subspecies. All parameters showed moderate to high heritability, with significant variation among lines and genotype x year interactions indicating that most of the variation observed between lines would be due to genetic base and not due to the environment. Based on the suggestions made by previous studies, CI was considered as a major trait in the GWAS analysis to define resistance gene in stem rust [31][22]. Thus, the lines were classified as follows:

140(57%) (CI = 0-19); 57(23%) (CI = 21-39), 48 (719%) were (CI > 40) (Fig. 1B, 1E). Based on the previous study for resistance characterization of *Pgt* made on Ethiopian bread wheat cultivars, 13 cultivars (Shorima, Hulluka, Hoggana, Galil, Senkegna, Hidase, Tay, Dinkinesh, Gassay, Millenium, Densa, Guna and Shina) showed CI values between 0 and 20 and were designated as having a high level of slow rusting[28].

Population structure and genetic diversity

Systematic characterization of population structure and genetic diversity provides a foundation for efficient exploitation of genetic resources and can enhance breeding for durable stem rust resistance in wheat. For the population structure study, three different approaches were applied. Firstly, two sub-populations with high admixtures were identified from the STRUCTURE software. Secondly, the population grouped into three sub-groupings based on the Neighbor-Joining algorithm, and color-coded using STRUCTURE. Thirdly, population structures inferred from PCA resulted in only about 10% variation for the first two components. All three clustering approaches infer the absence of a clear distinction between the sub-groups. This might be due to the complex evolutionary and breeding history of the panels. For the genetic diversity study, the lines were subjected to polymorphic information content (PIC), and, genome-wide gene diversity analysis. The mean PIC was 0.25, and the mean genome-wide gene diversity value was 0.3. Based on previous studies on common wheat in Ethiopia, an average PIC of 2_3 and average gene diversity of 3_4 were reported in various wheat collections [32][21][33]. Overall, our results revealed the influence of selection pressures on our breeding lines genetic diversity. Likewise, the fact that the world wheat diversity is currently shrinking magnifies the importance of broadening the genetic base of resistance wheat development schemes.

Linkage disequilibrium in elite breeding lines

The self-pollinating nature of wheat contributes to a have a high level of linkage disequilibrium (LD), and consequently this markers possible to detect QTL using a small number of markers. In the present study, the fastest LD decayed was observed for D sub-genomes. This finding is in agreement with the study of Sehal *et al.* (2017) [34]. In general, LD decay patterns variation

observed in different studies of bread wheat including the present study were because of differences in markers type, population structures, and thresholds of R^2 values.

Genome-wide marker trait association analysis

Adult-plant stage resistance (APR) to stem rust in wheat does usually require more than one gene with minor effect for effective response. To unravel the genetic basis of this resistance, association mapping using a genome-wide association study (GWAS) approach has become the leading method [35]. However, false-positive associations caused by population structure and family relatedness are major constraints in reliable interpretation and usage of GWAS results. To minimize the impact of these limitations, a mixed linear model (MLM) that uses a genetic marker-based kinship matrix (K) jointly with population structure is a widely used approach to QTL identification of APR to stem rust in wheat. From a total of 55 loci conferring resistance to *Pgt*, 50 are expressed as all-stage resistance (seedling resistance), and five that confer only adult-plant resistance (APR) [36]. To date, approximately 27 stem rust resistances (*Sr*) are effective or partially effective against the *Ug99* race group[37]. Of these, *Sr2, Sr13, Sr22, Sr25, Sr26, Sr35, Sr39,* and *Sr40* were reported to be the most effective against *Ug99*[38].

By using MLM, 46 SNP markers (13 QTLs) explaining 58.8% of phenotypic variation on chromosomes 1B, 3A, 3B, 4A, 4B, and 5A were found to be significantly associated for APR to stem rust of wheat (Table 2). Of these, only four QTLs/markers (EWYP1B.1/ wsnp_Ku_c13229_21142792, EWYP1B.3/BobWhite_c22266_315, EWYP1B.4/ wsnp_Ku_c30982_40765254, and EWYP1B.5/Kukri_c73734_175) were identified with $-\log_{10}(p) \geq 3$. Five QTL were located near genomic regions with previously reported *Sr* genes [37]. The new QTL might have the potential to enhance marker-assisted selection and/or genomic selection for *Pgt* resistance in combination with other minor genes.

A total of seven QTL (36 markers) out of 9523 polymorphic SNPs associated with *Sr* resistance were detected on chromosome 1B. These markers were spread from Tdurum_contig12899_3 (mapped at 9.7cM) to Tdurum_contig10036_977 (mapped at 114cM) presented on 68% of genotypes to 91% genotypes (additional file 1). On this chromosome, three stem rust resistance genes (*Sr14, Sr31,* and *Sr58*) were cataloged previously[39]. However, none of the presently identified markers were reported so far in association with wheat stem rust

disease. On the other hand, four out of thirty-six markers were co-located with formerly reported loci with wheat diseases other than stem rust, e.g. wsnp_BE442716B_Ta_2_1 (IWA106) and wsnp_Ex_rep_c69266_68192766 (IWA5592) for stripe rust resistance [33]; wsnp_Ex_c38116_45719983(WA3631) for Fusarium head blight [40], and BS00070139_51(IWA3631) for crown rot resistance [41]. This information allows researchers to design and deploy a breeding program for multiple what diseases resistance at the same time.

Merely 10 markers representing 6 loci were found in a region outside of 1B chromosome. On chromosome 3A, *Sr27* transferred from *Secale cereale* and *Sr35* transferred from *Triticum monococcum* were reported as all-stage resistance [36]. In the present study, marker Tdurum_contig777_260 (IWB73429) designated as EWYP3A at 20.74 cM was detected in this sub-genome in association with stem rust resistance of wheat. The closest APR *Sr* gene to this marker was *Sr27*. However, since the gene was not originally derived from hexaploid wheat, it is less likely that this region harbors *Sr27* genes. Thus, the locus is most likely harboring unknown *Sr* genes that are closely linked to the respective regions. On the short arm of chromosome 3B, *Sr2* originated from *Triticum dicoccum* and *Sr12* originated from *Triticum turgidum* ssp. *durum* were identified previously. In this chromosome, three markers (Tdurum_contig12899_342, Excalibur_c20277_483, and Tdurum_contig12008_803) designed as QTL EWYP3B.1 (mapped at 9.7 cM) were identified to be associated with *Sr* resistance. The nearest *Sr* gene to this region was *Sr2* which has been extensively used in breeding as a source of durable and broad-spectrum adult plant resistance. The Genotypes having these markers are most likely carring *Sr2* genes.The marker was present on 25% of genotypes for Tdurum_contig12899_342 to 39% of genotypes for Tdurum_contig12008_803. Of these, 62(25%), 84(34%), and 88(36%) genotypes showed CI between 8 to 36 for Tdurum_contig12899_342, Excalibur_c20277_483, and Tdurum_contig12008_803 markers, respectively (additional file 1). Marker RAC875_c30011_426 (IWB56412) found on chromosome 5B was responsible for 5.7% of the phenotypic variation. This gene was originally derived from *Triticum aestivum* and identified with close proximity to *Sr56*. Thus, the region is most likely possessing *Sr56* conferring APR to wheat stem rust. Other QTLs designed as EWYP3B.2, EWYP4A, and EWYP4B on chromosome 3B, 4A, and 4B, respectively were not reported in any literature and no close *Sr* gene or QTLs were reported hitherto

Conclusions

The present study characterized the genetic diversity of elite ICARDA breeding lines and performed GWAS based on field stem rust evaluation and 15k genotypic data. As a result, substantial genetic variability and field disease response to *Pgt* was observed among the lines. In addition, the study detected several potentially novel loci associated with *Pgt* resistance. These markers and/or QTL could provide useful genetic information used to unlock the genetic basis of resistance to *Pgt*. Furthermore, these results will accelerate the introgression of these resistance genes in the wheat breeding program through the marker-assisted selection and/or genomic selection. However, before implementing the above-mentioned applications, further studies are needed to validate the detected markers through fine mapping and confirm the location of reported QTL. These resistant lines could also be used as crossing parents for stem-rust-resistant breeding program and/or for direct release by national wheat development program after adaptation trails.

Materials and methods

Plant materials, field stem rust trials, and disease pathotyping

A set of 245 elite breeding lines was obtained from the International Center for Agricultural Research in the Dry Areas (ICARDA) shuttle breeding program.

Field screening was conducted in Ethiopia for two consecutive cropping seasons (2018 and 2019) at the Debre-Zeit Agricultural Research Center (DARC) located at 08° 44' N latitude and 38° 58' E longitude of 1900m.a.s.l with annual average temperature l of 19°C and 851 mm rainfall.,

The experiment was conducted using an augmented design with five local lines as control,, namely ‘Digelu’, ‘Kubssa’, ‘Hidasse’, ‘Honqolo’, and ‘Ogolcho’. Each line was sown in a $5gm^{-2}$ planting density, 1m long single row, and 30cm inter-row spacing. The borders of each block

were surrounded by susceptible local spreader wheat varieties to promote natural stem rust infection.

Stem rust phenotyping was conducted based on disease severity (DS) and infection response (IR) under natural disease pressure [42]. Both parameters were recorded three times for each line in each year. The highest value (most susceptible) in each year per line was used for GWAS analysis, and the coefficient of infection (CI) was determined by combining the two parameters (DS and IR) into a single value. The CI was calculated by multiplying the DS by a constant value for each IR where: immune = 0.0, R = 0.2, RMR = 0.3, MR = 0.4, MRMS= 0.6, MS = 0.8, MSS = 0. 9, and S = 1.0, (Yu et al., 2011).

Statistical analysis of phenotypic data

Analysis of variance (ANOVA) was performed for DS, IR, and CI using the 'nlme' package of the R 4.0.2 software (Pinheiro et al., 2020) by fitting the value of DS, IR and CI as function of Line, Years, and a combination of Line and years. Broad-sense heritability (H^2) was calculated using the following formula:

$$H^2 = \frac{\sigma^2 \text{ G}}{\sigma^2 \text{ G} + (\sigma^2 \text{G}XE)/n + (\sigma^2 \, error)/n}$$

Where $\sigma^2_{\text{ G}}$ is the genotypic variance, $\sigma^2_{\text{ E}}$ is the environment variance, σ^2_{GXE} is the genotype by environment interaction variance and σ^2 error is the residual error variance and n is the number of years.

In order to reduce false positive associations, best linear unbiased predictors (BLUPs) for CI were calculated using a mixed model in lme4 package implemented in R 4.0.2, R software [44] according to the following model, where, *y* is the response variable.

$$y = lmer\,(Trait \sim (1|Genetype) + (1|Year))$$

Population structure and genetic diversity

The optimal sub-population was estimated based on three different approaches. Firstly, the population structure (Q matrix) was estimated using 100 unlinked SNPs (markers located more than 10 cM apart) distributed across the wheat genome based on a Bayesian model using STRUCTURE 2.3.1 software [45][46]. To execute this, three independent runs were performed for each hypothetical K value (2 to 15), with the length of the burn-in period of 10,000 steps followed by 100,000 Monte Carlo Markov Chain (MCMC). The results obtained from this procedure was entered into in a web-based informatics tool "Structure Harvester" [47] to define the optimal K value based on ΔK method [48]. Each genotype was assigned to one subpopulation based on its membership probability. The optimal subpopulation was determined based on marker-based kinship matrix (K matrix) based on scaled identity-by-state method using the whole set of SNP markers from TASSEL 5 software [49]. Principal components analysis (PCA) of genetic relatedness was performed with the same software and added to the regression model as a covariant.

Genetic diversity was discovered based on polymorphic information content (PIC), heterozygosity, and Nei's gene diversity using the whole set of SNP markers from PowerMarker 3.25 software [50] Lastly, phylogenetic analysis based on distance-based neighbor-joining was calculated with TASSEL 5 software and visualized through web-based program iTOL (v 4.3.2) [51].

Genotyping, Linkage Disequilibrium, and Genome-Wide Association

Lines were genotyped using one-week old seedlings following the protocol described by [52] using Cetyeltrimethylammonium bromide (CTAB). Genotyping was performed by Illumina's iSelect 15K single nucleotide polymorphism (SNP) wheat array and called by GenomeStudio V2011.1 software. The resulting 13006 SNPs were further screened using those only with minor allelic frequency (MAF) >5%, and missing data percentage of <10% (five lines were eliminated).Of these, 9523 SNP markers for 245 lines that were distributed over the whole genome were used for all subsequent statistical analysis.

The resulting SNP data were subjected to linkage disequilibrium (LD) analysis as squared allelic frequency correlations (R^2) between each pair implemented in TASSEL v5.2 [49] and GAPIT (Genomic Association and Prediction Integrated Tool) R package [53]. Critical value of R^2 was defined by using 95% of the square root of transformed R^2 data of unlinked markers as the threshold. Anything, beyond that LD is probably caused by a real physical linkage.

Marker-trait association analysis (MTAs) between the BLUP value of CI, and SNPs genotypic data were analyzed using mixed liner model (MLM) in TASSEL 5.2 software, [49]. Genetic marker was used as a covariance matrix between individuals together with population structure and kinship matrix as K + PCA model described by Pressoir *et al*. [54]. The general equations for MLM are: $y = X\alpha + Q\delta + K\mu + e$; where y = phenotypic values, X is SNP marker genotypes, α is a vector containing fixed effects as a result of the genotype, Q was population structure as PCA, $\boldsymbol{\delta}$ is a vector containing fixed effects resulting from population structure, K is the relative kinship matrix, μ is a vector of random additive genetic effects, and is a vector of residuals. Marker trait associations were declared significant at a threshold value of $-\log_{10}(p) \geq 3$ (corresponding $p\ value \leq 0.001$) [55].

Availability of data and materials

The data sets supporting the results of this article are included in this manuscript and its additional information files

Abbreviations

BLUP: Best linear unbiased prediction
CV: Coefficient of variation

GWAS: Genome-wide association study

LD, Linkage disequilibrium

MAF: Minor allele frequency

PCA: Principal component analysis

PIC: Polymorphism information content

Pgt: wheat stem rust fungus *Puccinia graminis*

QQ: quantile-quantile

QTL: Quantitative trait locus

SNP: Single nucleotide polymorphism

References

1. Giraldo P, Benavente E, Manzano-Agugliaro F, Gimenez E. Worldwide research trends on wheat and barley: A bibliometric comparative analysis. Agronomy. 2019;9.

2. Marocco E, Milo M. Food Outlook. 2019.

3. Lucas H. Breakout session P1.1 National Food Security-The Wheat Initiative-an International Research Initiative for Wheat Improvement. Second Glob Conf Agric Res Dev. 2012; September:1–3. www.wheatinitiative.org.

4. Enghiad A, Ufer D, Countryman AM, Thilmany DD. An Overview of Global Wheat Market Fundamentals in an Era of Climate Concerns. Int J Agron. 2017;2017.

5. Schütz H, Jansen M, Verhoff MA. Vom alkohol zum liquid ecstasy (GHB) - Ein überblick über alte und moderne K.-o.-Mittel - Teil 3: γ-Hydroxybuttersäure (GHB, “liquid ecstasy”). Arch Kriminol. 2011;228:151–9.

6. Figueroa M, Hammond-Kosack KE, Solomon PS. A review of wheat diseases—a field perspective. Mol Plant Pathol. 2018;19:1523–36.

7. Gao L, Rouse MN, Mihalyov PD, Bulli P, Pumphrey MO, Anderson JA. Genetic characterization of stem rust resistance in a global spring wheat germplasm collection. Crop Sci. 2017;57:2575–89.

8. Li F, Upadhyaya NM, Sperschneider J, Matny O, Nguyen-Phuc H, Mago R, et al. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. Nat Commun. 2019;10. doi:10.1038/s41467-019-12927-7.

9. Fetch T, Zegeye T, Park RF, Hodson D, Wanyera R. Detection of wheat stem rust races TTHSK and PTKTK in the Ug99 race group in Kenya in 2014. Plant Dis. 2016;100:1495.

10. Newcomb M, Olivera PD, Rouse MN, Szabo LJ, Johnson J, Gale S, et al. Kenyan isolates of Puccinia graminis f. sp. tritici from 2008 to 2014: Virulence to SrTmp in the Ug99 race group and implications for breeding programs. Phytopathology. 2016;106:729–36.

11. Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, Njau P, et al. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu Rev Phytopathol. 2011;49:465–81.

12. Prank M, Kenaley SC, Bergstrom GC, Acevedo M, Mahowald NM. Climate change impacts the spread potential of wheat stem rust, a significant crop disease. Environ Res Lett. 2019;14.

13. Saunders DGO, Pretorius ZA, Hovmøller MS. Tackling the re-emergence of wheat stem rust in Western Europe. Commun Biol. 2019;2:9–11. doi:10.1038/s42003-019-0294-9.

14. Olivera PD, Sikharulidze Z, Dumbadze R, Szabo LJ, Newcomb M, Natsarishvili K, et al. Presence of a Sexual Population of *Puccinia graminis* f. sp. *tritici* in Georgia Provides a Hotspot for Genotypic and Phenotypic Diversity. Phytopathology®. 2019;109:2152–60. doi:10.1094/PHYTO-06-19-0186-R.

15. Leppik EE. Gene Centers of Plants as Sources of Disease Resistance. Annu Rev Phytopathol. 1970;8:323–44.

16. Hundie B. Evaluation of Advanced Bread Wheat Lines for Field and Seedling Resistance to Stem Rust (<i>Puccinia graminis</i> f. sp. <i>tritici</i>). Am J Biol Environ Stat. 2018;4:74.

17. Rahmatov M, Pretorius ZA, Bhavani S. Sources of Stem Rust Resistance in Wheat-Alien Introgression Lines. 2016. doi:10.1094/PDIS-12-15-1448-RE.

18. Aktar-Uz-Zaman M, Tuhina-Khatun M, Hanafi MM, Sahebi M. Genetic analysis of rust resistance genes in global wheat cultivars: an overview. Biotechnol Biotechnol Equip. 2017;31:431–45. doi:10.1080/13102818.2017.1304180.

19. Crossa J, Burgueño J, Dreisigacker S, Vargas M, Herrera-Foessel SA, Lillemo M, et al. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics. 2007;177:1889–913.

20. Yu LX, Lorenz A, Rutkoski J, Singh RP, Bhavani S, Huerta-Espino J, et al. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theor Appl Genet. 2011;123:1257–68.

21. Muleta KT, Rouse MN, Rynearson S, Chen X, Buta BG, Pumphrey MO. Characterization of molecular diversity and genome-wide mapping of loci associated with resistance to stripe rust and stem rust in Ethiopian bread wheat accessions. BMC Plant Biol. 2017;17.

22. Gao L, Turner MK, Chao S, Kolmer J, Anderson JA. Genome Wide Association Study of Seedling and Adult Plant Leaf Rust Resistance in Elite Spring Wheat Breeding Lines. PLoS One. 2016;11:e0148671. doi:10.1371/journal.pone.0148671.

23. Leonova IN, Skolotneva ES, Orlova EA, Orlovskaya OA, Salina EA. Detection of genomic regions associated with resistance to stem rust in russian spring wheat varieties and breeding germplasm. Int J Mol Sci. 2020;21:1–13.

24. Tadesse W, Halila H, Jamal M, Hanafi S, Assefa S, Oweis T, et al. Role of Sustainable Wheat Production to Ensure Food Security in the CWANA region. J Exp Biol Agric Sci. 2017;5 Spl-1-SAFSAW:15–32.

25. Novoselović D, Bentley AR, Šimek R, Dvojković K, Sorrells ME, Gosman N, et al. Characterizing croatian wheat germplasm diversity and structure in a european context by DArT markers. Front Plant Sci. 2016;7.

26. Bhardwaj SC, Singh GP, Gangwar OP, Prasad P, Kumar S. Status of wheat rust research and progress in rust management-indian context. Agronomy. 2019;9:1–14.

27. Singh RP, Hodson DP, Jin Y, Lagudah ES, Ayliffe MA, Bhavani S, et al. Emergence and spread of new races of wheat stem rust fungus: Continued threat to food security and prospects of genetic control. Phytopathology. 2015;105:872–84.

28. Mitiku M, Bacha Hei N, Abera M. Characterization of Slow Rusting Resistance Against Stem Rust (Puccinia graminis f. sp. tritici) in Selected Bread Wheat Cultivars of Ethiopia. Adv Crop Sci Technol. 2018;06.

29. Abebe T, Dawit W, Woldeab G. Physiological Races and Virulence Diversity of Puccinia graminis pers. f. sp. tritici eriks. & e. Henn. on Wheat in Tigray Region of Ethiopia. Int J Phytopathol. 2013;2:01–7.

30. Hei N, Shimelis HA, Laing M, Admassu B. Assessment of ethiopian wheat lines for slow rusting resistance to stem rust of wheat caused by puccinia graminis f.sp. tritici. J Phytopathol. 2015;163:353–63.

31. Yu L-X, Morgounov A, Wanyera R, Keser M, Singh SK, Sorrells M. Identification of Ug99 stem rust resistance loci in winter wheat germplasm using genome-wide association analysis. Theor Appl Genet. 2012;125:749–58. doi:10.1007/s00122-012-1867-x.

32. Tadesse W, Suleiman S, Tahir I, Sanchez-Garcia M, Jighly A, Hagras A, et al. Heat-tolerant QTLs associated with grain yield and its components in spring bread wheat under heat-stressed environments of Sudan and Egypt. Crop Sci. 2019;59:199–211.

33. Zegeye H, Rasheed A, Makdis F, Badebo A, Ogbonnaya FC. Genome-Wide Association Mapping for Seedling and Adult Plant Resistance to Stripe Rust in Synthetic Hexaploid Wheat. PLoS One. 2014;9:e105593. doi:10.1371/journal.pone.0105593.

34. Sehgal D, Autrique E, Singh R, Ellis M, Singh S, Dreisigacker S. Identification of genomic regions for grain yield and yield stability and their epistatic interactions. Sci Rep. 2017;7:1–12. doi:10.1038/srep41578.

35. Alqudah AM, Sallam A, Stephen Baenziger P, Börner A. GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. Journal of Advanced Research. 2020;22:119–35.

36. Park RF. Wheat: Biotrophic Pathogen Resistance. In: Reference Module in Food Science. Elsevier; 2016.

37. Yu LX, Barbier H, Rouse MN, Singh S, Singh RP, Bhavani S, et al. A consensus map for Ug99 stem rust resistance loci in wheat. Theor Appl Genet. 2014;127:1561–81. doi:10.1007/s00122-014-2326-7.

38. Aktar-Uz-Zaman M, Tuhina-Khatun M, Hanafi MM, Sahebi M. Genetic analysis of rust resistance genes in global wheat cultivars: an overview. Biotechnol Biotechnol Equip. 2017;31:431–45. doi:10.1080/13102818.2017.1304180.

39. Wheat: Biotrophic Pathogen Resistance - ScienceDirect. https://www.sciencedirect.com/science/article/pii/B9780081005965002079. Accessed 8 Sep 2020.

40. Aviles AC, Harrison SA, Arceneaux KJ, Brown-Guidera G, Mason RE, Baisakh N. Identification of qtls for resistance to fusarium head blight using a doubled haploid population derived from southeastern united states soft red winter wheat varieties ags 2060 and ags 2035. Genes (Basel). 2020;11:1–18.

41. Rahman M, Davies P, Bansal U, Pasam R, Hayden M, Trethowan R. Marker-assisted recurrent selection improves the crown rot resistance of bread wheat. Mol Breed. 2020;40.

42. Roelfs AP, Singh RP, Saari EE. Concepts and methods of disease management. 1992.

43. Linear T, Mixed N, Models E, Fit D, Hmisc S. Package ‘ nlme .’ 2020.

44. Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, et al. Linear mixed-effects model using “Eigen” and S4, R Package Version 1.1- 23, https://github.com/lme4/lme4/. 2020.

45. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. 2000. http://www.stats.ox.ac.uk/pritch/home.html. Accessed 9 Sep 2020.

46. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. Mol Ecol Notes. 2007;7:574–8.

47. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4:359–61. doi:10.1007/s12686-011-9548-7.

48. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol Ecol. 2005;14:2611–20.

49. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23:2633–5.

50. Liu K, Muse S V. PowerMaker: An integrated analysis environment for genetic maker analysis. Bioinformatics. 2005;21:2128–9.

51. Letunic I, Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. Bioinformatics. 2007;23:127–8.

52. Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc. 2006;1:2320–5. doi:10.1038/nprot.2006.384.

53. Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, et al. GAPIT: Genome association and prediction integrated tool. Bioinformatics. 2012;28:2397–9. doi:10.1093/bioinformatics/bts444.

54. Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2006;38:203–8. doi:10.1038/ng1702.

55. Sukumaran S, Reynolds MP, Sansaloni C. Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. Front Plant Sci. 2018;9 February.

Acknowledgements

This project is a part of ICARDA's shuttle breeding program, which has been designed for performing 15K SNP chip genotyping of the GWAS panel. I thank Professor Endashaw Bekele (Addis Ababa University) for reviewing an earlier draft of this manuscript. I acknowledge Dr. Wuletaw Tadesse (ICARDA) for provision of the lines and for facilitation of this experiment. I'm also very grateful to Dr. Thomas Fetch (Agriculture and Agri-Food Canada) for assisting in reviewing the paper. I acknowledge Debre Zeit Agricultural Research Center for facilitating field experiments

Contributions

ES, WT, and EB conceived and designed the study. WT designed the study, and provided the germplasm and the genotypic data. EB provided guidance throughout the project development. ES writing of this manuscript. ES and AA performed data analysis. LM edited the manuscript. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Ethics declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Figures

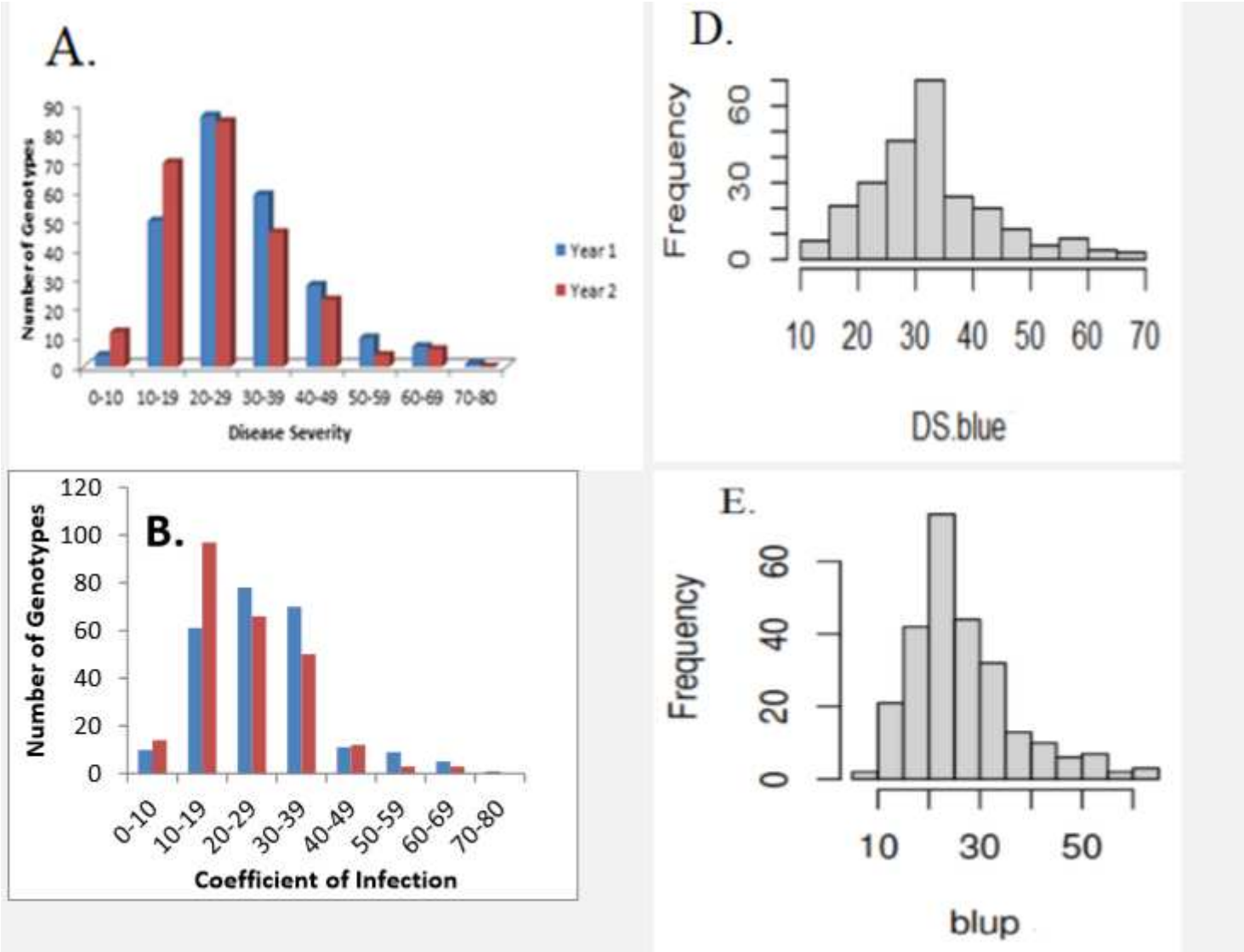

Figure 1

Distribution of adult plant stage disease response and best linear unbiased estimate (BLUPs) as disease severity (DS), and coefficient of infection (CI).(A) Frequency of genotypes for disease severity (DS); (B) frequency of genotypes for coefficient of infection) (D) BLUPs of DS; (E) BLUPs of CI.

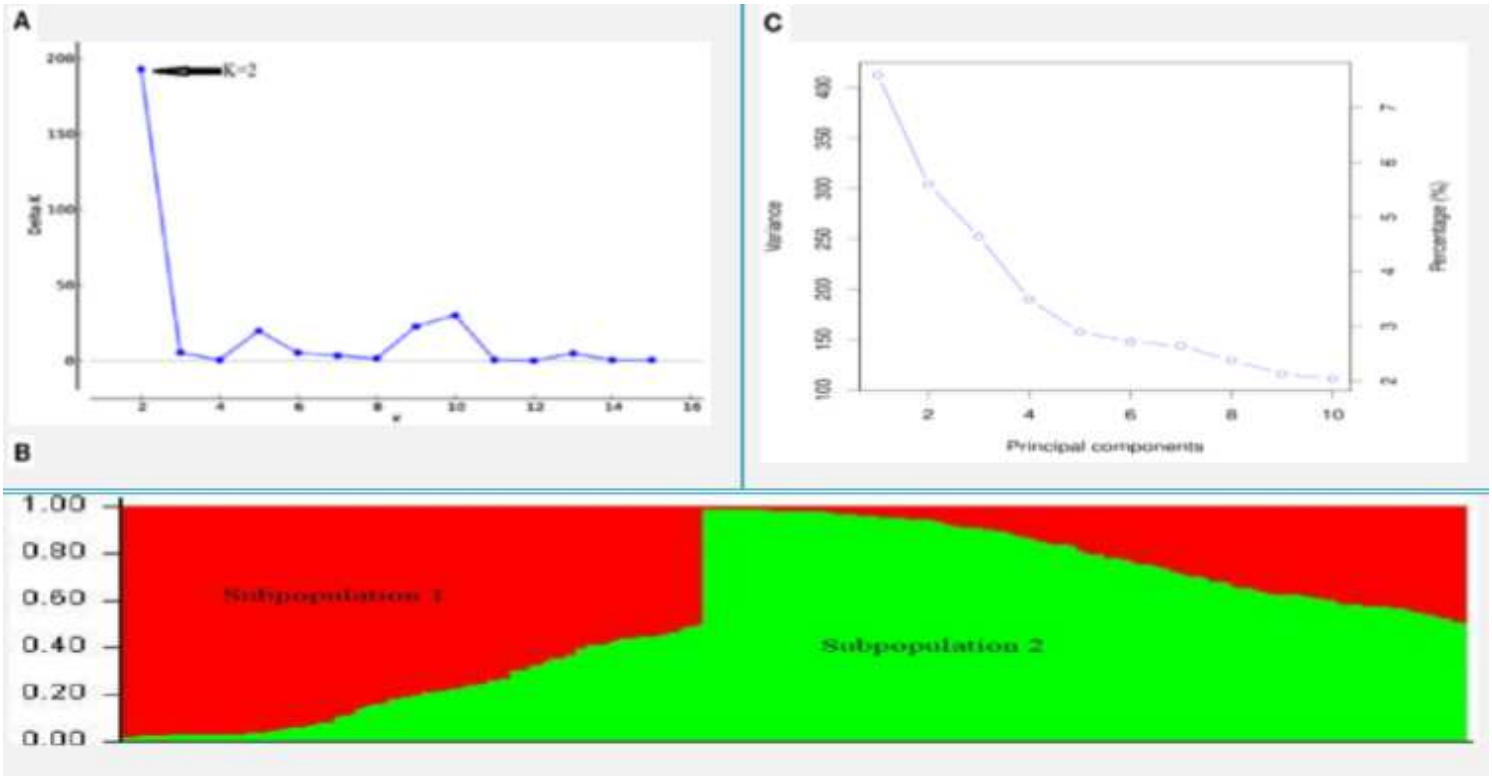

Figure 2

Structure clustering and principal components of 245 wheat lines based on 9523 SNP markers. (A) Plots of delta K; (B) probability of population group based on K=2; and (C) Scree plot of PCA

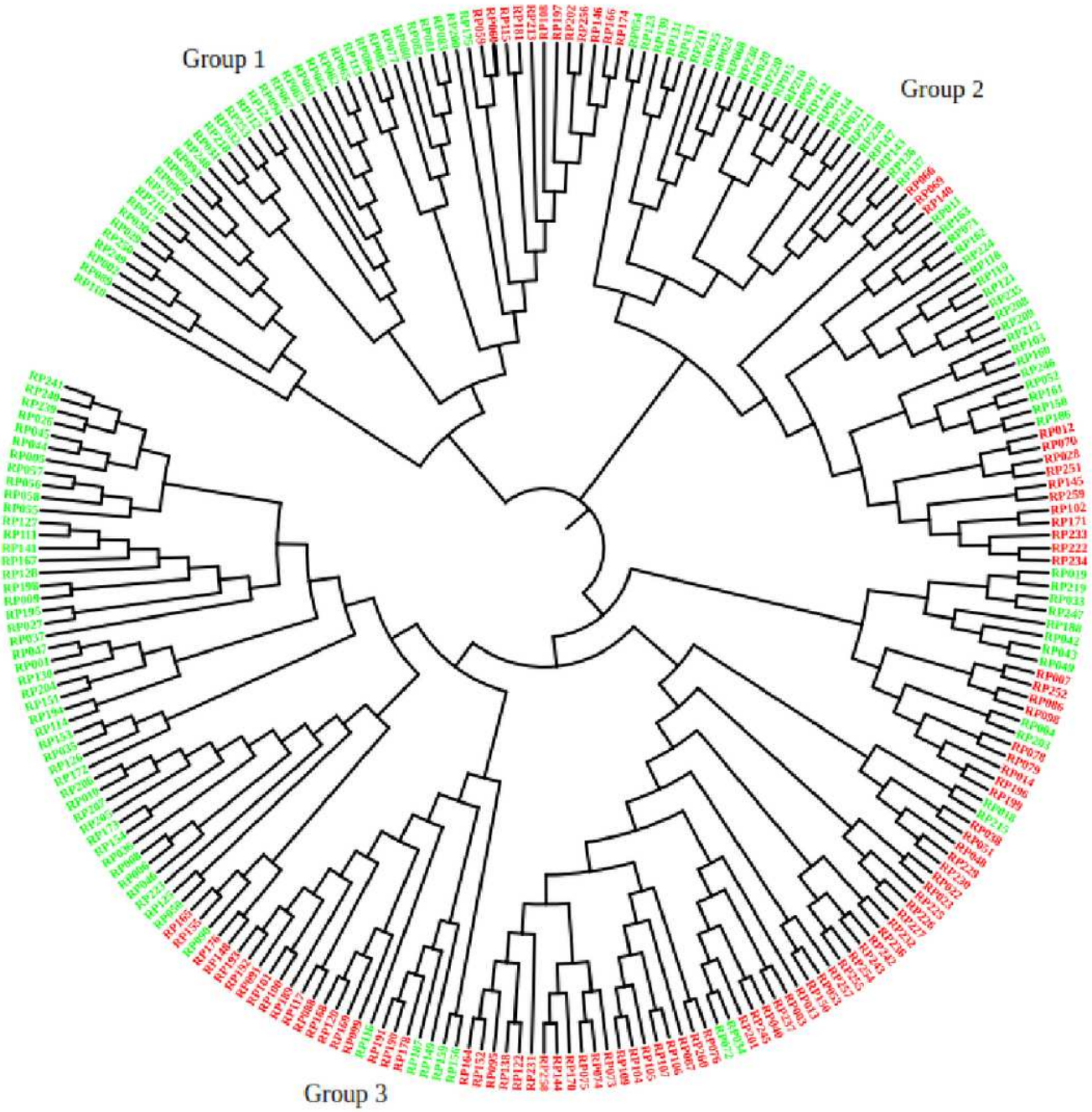

Figure 3

Neighbor joining tree based on Nei's genetic distance color coded with STRUCTURE probability distribution

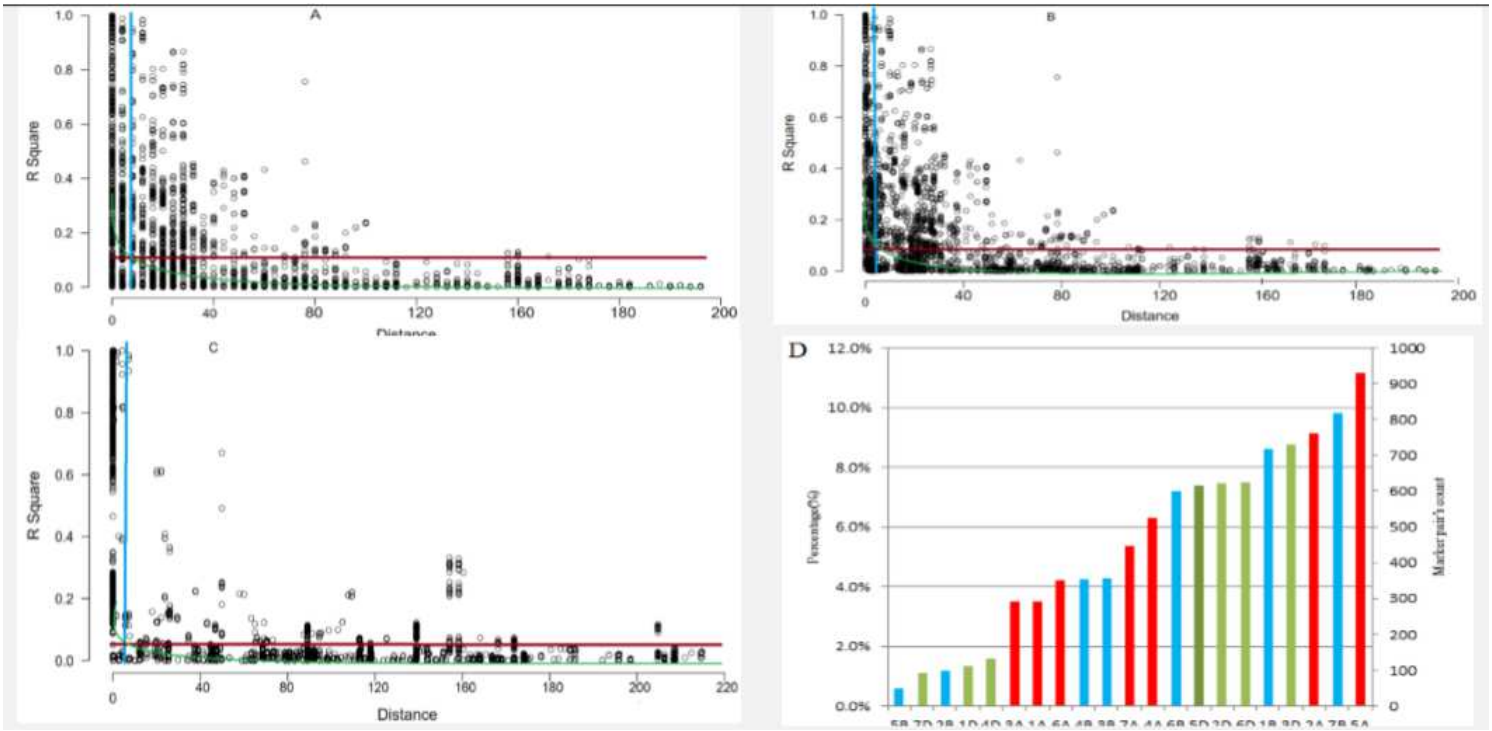

Figure 4

Genome wide linkage disequilibrium (LD) decays and distributions of single nucleotide polymorphisms (SNPs) based on 15K genotyping results. Intra-chromosomal pairs in the genetic distance (cM) are plotted against the LD estimate (R 2) for pairs of single-nucleotide polymorphisms.

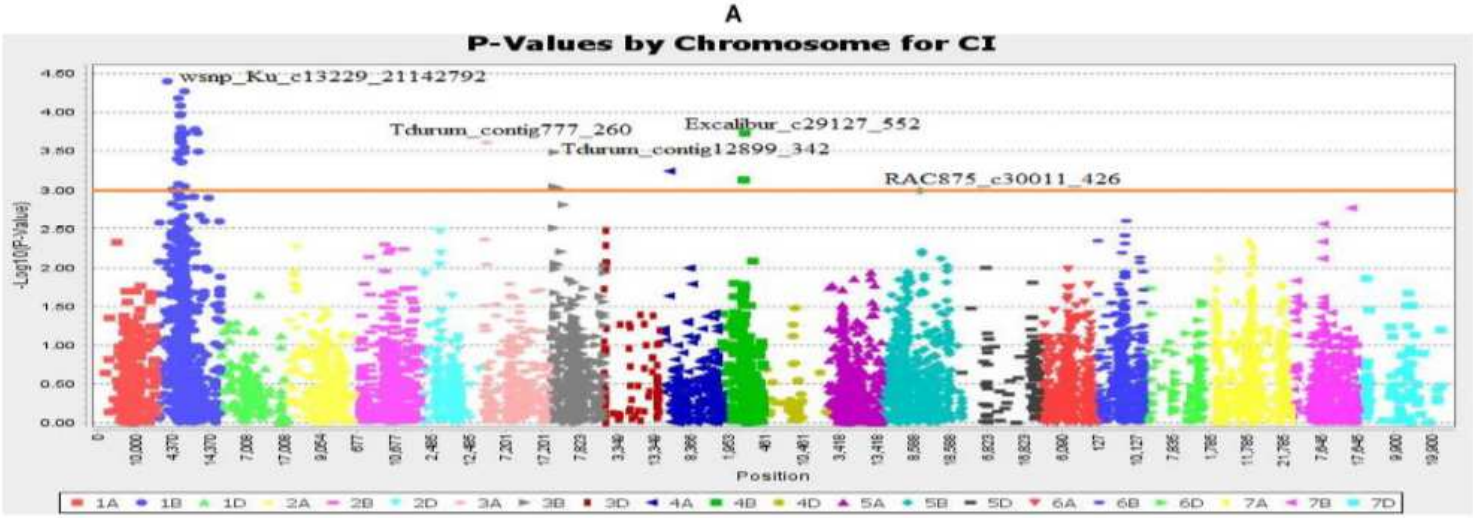

Figure 5

Manhattan plots based on the mixed linear model along each of 21 chromosomes showing significantly associated markers. The dots above solid orange horizontal lines represent genome-wide significant markers for threshold p value of 0.001 (MTAs at $-\log10\ p \geq 3$). The position of each marker was based on the wheat consensus SNP map (Wang et al., 2014).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.pdf