



Article

Combination of Linkage Mapping, GWAS, and GP to Dissect the Genetic Basis of Common Rust Resistance in Tropical Maize Germplasm

Maguta Kibe ^{1,2}, Christine Nyaga ¹, Sudha K. Nair ³, Yoseph Beyene ¹, Biswanath Das ¹, Suresh L. M ¹, Jumbo M. Bright ¹, Dan Makumbi ¹, Johnson Kinyua ², Michael S. Olsen ¹, Boddupalli M. Prasanna ¹ and Manje Gowda ^{1,*}

¹ International Maize and Wheat Improvement Center (CIMMYT), P. O. Box 1041-00621, Nairobi 00100, Kenya; maguta.kibe@gmail.com (M.K.); christinenyaga96@gmail.com (C.N.); Y.Beyene@cgiar.org (Y.B.); b.das@cgiar.org (B.D.); l.m.suresh@cgiar.org (S.L.M.); b.jumbo@cgiar.org (J.M.B.); D.Makumbi@cgiar.org (D.M.); M.Olsen@cgiar.org (M.S.O.); b.m.prasanna@cgiar.org (B.M.P.)

² Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi 00100, Kenya; johnsonkinyua@jkuat.ac.ke

³ International Maize and Wheat Improvement Center (CIMMYT), ICRISAT Campus, Patancheru, Greater Hyderabad 502324, India; Sudha.Nair@cgiar.org

* Correspondence: m.gowda@cgiar.org

Received: 20 August 2020; Accepted: 4 September 2020; Published: 6 September 2020



Abstract: Common rust (CR) caused by *Puccinia sorghii* is one of the destructive fungal foliar diseases of maize and has been reported to cause moderate to high yield losses. Providing CR resistant germplasm has the potential to increase yields. To dissect the genetic architecture of CR resistance in maize, association mapping, in conjunction with linkage mapping, joint linkage association mapping (JLAM), and genomic prediction (GP) was conducted on an association-mapping panel and five F₃ biparental populations using genotyping-by-sequencing (GBS) single-nucleotide polymorphisms (SNPs). Analysis of variance for the biparental populations and the association panel showed significant genotypic and genotype x environment (GXE) interaction variances except for GXE of Pop4. Heritability (h^2) estimates were moderate with 0.37–0.45 for the individual F₃ populations, 0.45 across five populations and 0.65 for the association panel. Genome-wide association study (GWAS) analyses revealed 14 significant marker-trait associations which individually explained 6–10% of the total phenotypic variances. Individual population-based linkage analysis revealed 26 QTLs associated with CR resistance and together explained 14–40% of the total phenotypic variances. Linkage mapping revealed seven QTLs in pop1, nine QTL in pop2, four QTL in pop3, five QTL in pop4, and one QTL in pop5, distributed on all chromosomes except chromosome 10. JLAM for the 921 F₃ families from five populations detected 18 QTLs distributed in all chromosomes except on chromosome 8. These QTLs individually explained 0.3 to 3.1% and together explained 45% of the total phenotypic variance. Among the 18 QTL detected through JLAM, six QTLs, *qCR1-78*, *qCR1-227*, *qCR3-172*, *qCR3-186*, *qCR4-171*, and *qCR7-137* were also detected in linkage mapping. GP within population revealed low to moderate correlations with a range from 0.19 to 0.51. Prediction correlation was high with $r = 0.78$ for combined analysis of the five F₃ populations. Prediction of biparental populations by using association panel as training set reveals positive correlations ranging from 0.05 to 0.22, which encourages to develop an independent but related population as a training set which can be used to predict diverse but related populations. The findings of this study provide valuable information on understanding the genetic basis of CR resistance and the obtained information can be used for developing functional molecular markers for marker-assisted selection and for implementing GP to improve CR resistance in tropical maize.

Keywords: genome-wide association study; genomic prediction; joint linkage association mapping; genotyping by sequencing; resistance; common rust

1. Introduction

Common rust (hereafter CR) caused by *Puccinia sorghi* is one of the most destructive fungal foliar diseases in maize-growing regions predominantly in humid areas and has been reported to cause 12 to 61% yield losses in a favorable environment [1,2]. These yield losses are subject to leaf area infected and host growth stages whereby the former has been estimated to cause about 3–8% yield loss for each 10% of the total leaf area affected [3]. Quantitative resistance is due to partial resistance or adult plant resistance [4]. Numerous studies have suggested that older and mature tissues have more resistance to CR than younger soft tissues [4–6].

Past efforts to control CR through conventional means have been largely unsuccessful and also affected by unpredictable weather, and the use of fungicides leads to environmental effects and increased production costs [7]. Host-plant resistance has been identified as the most reliable and economically viable option among several available options to alleviate plant diseases [7,8]. In the case of CR, researchers have identified both qualitative and quantitative resistance [9,10]. Resistance through R genes has been identified more than 25 dominant race-specific (*Rp*) genes in chromosomes (chr) 3, 4, and 10 [9,11], which mediate the recognition of the pathogen and trigger a hypersensitive reaction to prevent further spread of the pathogen [12]. However, novel *P. sorghi* races can overcome the qualitative resistance in some genotypes and this requires a continuous search for sources of stable and durable resistance (quantitative) in order to manage the disease. Identifying resistant germplasm to CR and incorporation of resistant genes or genomic regions to elite lines and commercial hybrids has the potential to increase yields with lower production costs [13].

Genetic mapping through linkage analyses and GWAS have been used in many studies in plant breeding [12,14–26]. The two approaches exploit the recombination's ability to break up the genome into fragments that can be correlated with phenotypic variation but differ with the type of control they have on the recombination [27]. Linkage mapping in plants utilizes biparental crosses and, thus, is a closed controlled system. This further limits the number of recombinations that can sufficiently shuffle the genome into small fragments and results into QTLs localized in large chromosomal regions [28]. On the other hand, GWAS uses natural populations that mimic historical recombinations and provide a higher resolution as compared to linkage mapping [25]. This approach, however, has no control over relatedness and is prone to spurious associations. Accounting for population structure and kinship relatedness in a mixed model has been the most effective method of reducing false associations in GWAS [29]. Quantitative resistance for CR, a more durable approach, has also been exploited in a few studies through quantitative trait loci (QTL) analysis and genome-wide association studies (GWAS). A total of 41 QTLs have been identified conferring resistance to CR across all maize chromosomes in four separate studies utilizing different maize germplasm [7,10,13,30]. One study by Olukolu et al. [12], carried out GWAS and identified three marker trait associations (MTAs) in chr 2, 3, and 8 using 274 diverse inbred lines and 246,497 SNPs. Another study combined linkage and association mapping on 296 tropical maize inbreds and identified 25 QTLs on chr 1, 3, 5, 6, 8, and 10 associated with CR resistance [31]. These markers and candidate genes were directly or indirectly involved with plant disease responses [12,31].

In cases where the phenotype is strongly correlated with relatedness, population mapping even with the use of mixed models can be severely underpowered [27]. Joint linkage association mapping (JLAM) has the potential to overcome the downsides of the both linkage mapping and association mapping approaches whereby one compliments the other. Association mapping will increase the mapping resolution power while linkage mapping will account for relatedness in cases where Q + K explains most of the phenotypic variance. To our knowledge, no study has been reported utilizing JLAM to identify QTLs conferring resistance to CR in maize.

Another promising genomic tool is the genomic prediction (GP) which has been applied successfully in many plant breeding studies [32,33]. Previous reports have indicated the potential of GP to increase genetic gain and reduce the time taken in breeding programs significantly [33]. GP utilizes genome-wide markers to estimate the effects of all loci and computes a genomic estimated breeding value (GEBVs) [34]. Unlike genetic mapping, GP does not identify significant MTAs first but goes ahead to use all markers available to estimate their effects thus providing a powerful approach to account for any effects that might have been missed by genetic mapping. However, this doesn't mean complete withdrawal of genetic mapping but rather the incorporation of the two in genetic studies as complimentary approaches since each provide significant advantages. The main objectives of this study were to: (1) assess the phenotypic variation for CR by evaluating five segregating populations and an association panel in multiple environments; (2) detect stable and the novel genetic loci significantly associated with CR resistance by using QTL mapping, JLAM, and GWAS; and (3) GP within populations and Improved maize for African Soil (IMAS) panel and among the IMAS and F3 populations in order to understand the genetic architecture of maize CR resistance.

2. Results

Comparable disease pressure was observed in each environment as indicated by significant genotypic variance at each environment for each population (Table S1). Further, significant ($p < 0.05$) Pearson correlations were also observed among phenotypic values determined at different environments for each population (Table S2). This suggested that there was enough CR disease pressure in each environment. Cross-environment analyses revealed normal distributions of CR disease severity (score 1–9) in the F₃ populations and IMAS association panel, ranging from high resistance to moderate susceptibility with scores ranging from 2.0 to 6.0 and with average means ranging from 3.4 to 4.1 for different populations (Table 1 and Figure 1). Analysis of variance for the biparental populations and the association panel showed significant genotypic and genotype x environment (GXE) variances except for GXE of Pop1 and Pop4 (Table 1). Heritability (h^2) estimates were moderate with range of 0.37–0.45 for the individual F₃ populations, 0.45 across the F₃ populations and 0.65 for the IMAS panel.

Table 1. Means, ranges, and components of variance for Common rust in an IMAS association panel, F3 populations and across F3 populations.

Population	Mean (Range)	σ^2_G	$\sigma^2_{G \times E}$	σ^2_e	h^2
IMAS AMP	3.80 (2.40–5.90)	0.07 *	0.03 *	0.13	0.68
CZL0618 × LaPostaSeqC7-F71-1-2-1-1B-Pop1	3.61 (2.00–5.50)	0.024 *	0.01	0.10	0.44
CZL074 × LaPostaSeqC7-F103-1-2-1-1B-Pop2	3.70 (2.50–5.30)	0.02 *	0.02 *	0.12	0.43
CZL0009 × CZL99017-Pop3	3.42 (2.32–5.00)	0.03 *	0.02 *	0.11	0.45
CML505 × CZL99017-Pop4	4.10 (2.76–5.14)	0.01 *	0.01	0.12	0.37
CZL0723 × CZL0724-Pop5	4.02 (2.51–6.02)	0.03 *	0.02 *	0.16	0.38
Across five populations	3.85 (2.02–6.00)	0.04 *	0.10 *	0.15	0.45

* Significance at $p < 0.05$.

An IMAS association mapping panel was earlier used to study the maize lethal necrosis resistance [20], where the population structure was reported in detail. In the IMAS panel, a rapid linkage disequilibrium (LD) decay across physical distance in kb was reported (Figure 2). At LD cut-off points of $r^2 = 0.1$, the average physical distance was 14.97 kilo base pairs. GWAS was performed using a mixed linear model by integrating population structure (PCA) and family relatedness (kinship) within the IMAS panel using 337,110 high quality SNPs. GWAS for CR identified 14 significant marker-trait associations (MTA, significant threshold $p < 9 \times 10^{-6}$). These SNPs were found across all chromosomes except on chr 6, 7, and 8 (Table 2 and Figure 3). The distribution of SNPs across chromosomes and their level of significance for CR are shown in a Manhattan plot (Figure 3A). To test the ability of the model used, a quantile-quantile plot of the observed-log p -value vs. the expected-log p -value was plotted. As shown in Figure 3B, the population structure was controlled well by the mixed linear model (Figure 3). In terms of the percentage of phenotypic variance explained (PVE), the SNPs identified

from GWAS individually explained 6–10% (Table 2). SNP *S5_51353429* had the largest significant MTA while *S3_21856582* explained the most phenotypic variance. Candidate genes were selected around the significant SNPs and identify the putative function of these genes. Six candidate genes were identified in the significant SNP sites or adjacent to these sites (Table 2). We identified two candidate genes each on chromosomes 2, 5, and 10.

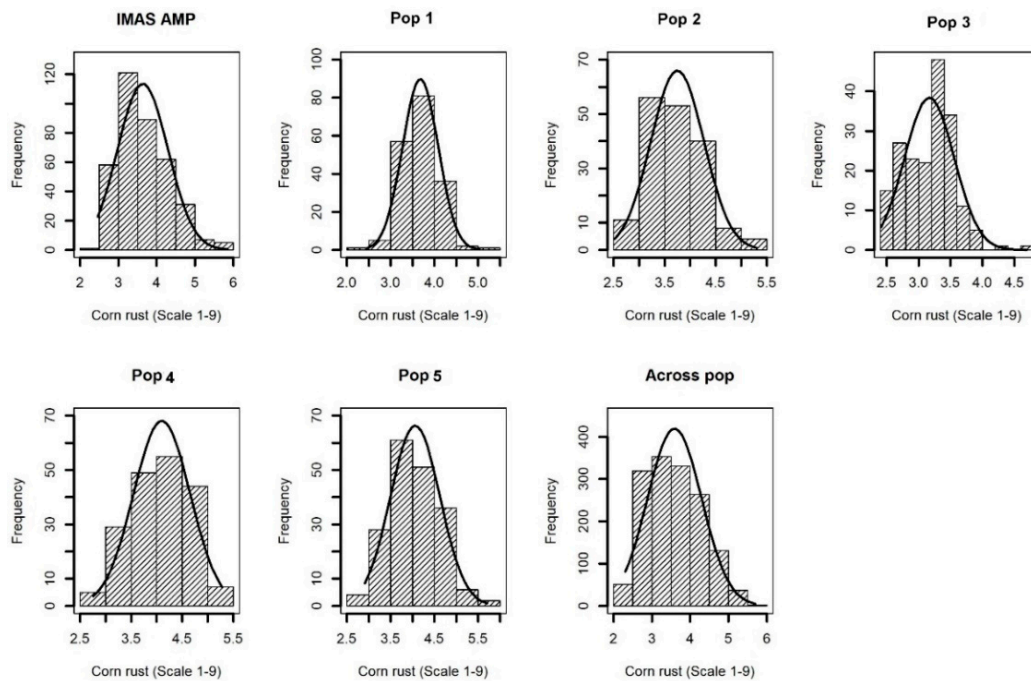


Figure 1. Phenotypic distribution for corn rust or common rust severity in the IMAS association panel, five F_3 populations and across F_3 populations.

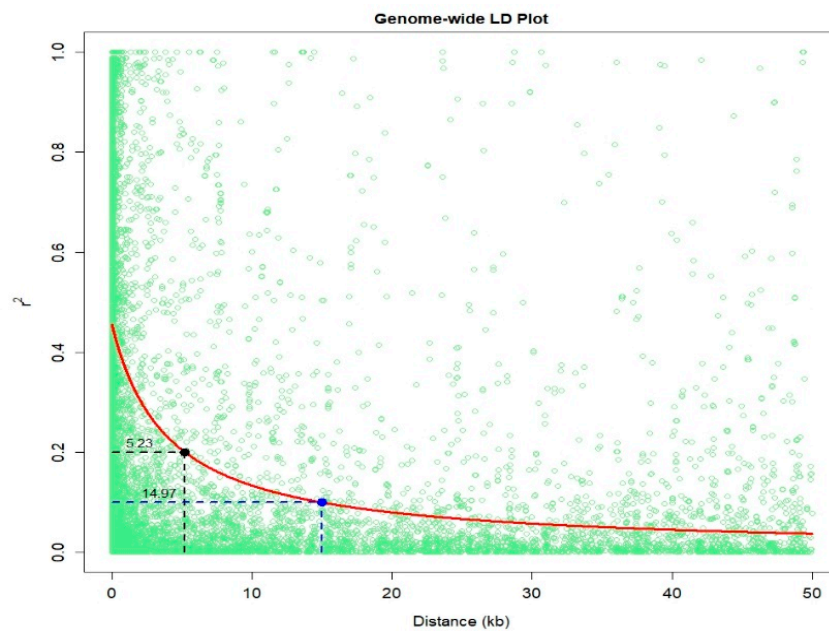


Figure 2. Linkage disequilibrium (LD) plot representing the average genome-wide LD decay in the panels with genome-wide markers. The values on the Y-axis represents the squared correlation coefficient r^2 and the X-axis represents the physical distance in (kb).

Table 2. Chromosomal position and SNPs significantly associated with Common rust disease severity (DS) detected by SNP-based GWAS in the IMAS association mapping panel.

SNP ^a	Chr	MLM-P Value	R ² (%)	MAF	Allele	Putative Candidate Genes	Predicted Function of Candidate Gene
S1_12663024	1	4.86 × 10 ⁻⁶	7.00	0.14	A/G	<i>GRMZM2G480386</i>	uncharacterized
S1_41433126	1	4.69 × 10 ⁻⁶	7.88	0.05	G/A	<i>GRMZM5G886521</i>	uncharacterized
S1_220067760	1	5.86 × 10 ⁻⁷	7.80	0.46	C/T	<i>GRMZM2G564469</i>	uncharacterized
S2_16361185	2	5.23 × 10 ⁻⁶	7.70	0.08	C/T	<i>GRMZM2G086484</i>	Pleckstrin homology (PH) domain superfamily protein
S2_222274747	2	7.65 × 10 ⁻⁶	7.10	0.05	C/T	<i>GRMZM2G009188</i>	11-beta-hydroxysteroid dehydrogenase 1B
S3_21856582	3	3.84 × 10 ⁻⁶	9.54	0.42	A/C	<i>GRMZM2G395983</i>	uncharacterized
S3_34683394	3	4.83 × 10 ⁻⁶	9.02	0.35	A/G	<i>GRMZM5G881063</i>	uncharacterized
S3_147013779	3	3.89 × 10 ⁻⁶	7.70	0.17	G/C	<i>GRMZM2G060540</i>	uncharacterized
S4_130478096	4	6.72 × 10 ⁻⁶	6.93	0.12	A/T	<i>GRMZM5G833902</i>	uncharacterized
S5_10087070	5	8.02 × 10 ⁻⁶	6.25	0.06	A/G	<i>GRMZM2G181002</i>	Phosphotransferases, Serine or threonine-specific kinase subfamily
S5_10089138	5	4.39 × 10 ⁻⁶	6.56	0.07	T/C	<i>GRMZM2G181002</i>	uncharacterized
S5_51353429	5	9.05 × 10 ⁻⁶	7.40	0.19	G/A	<i>GRMZM2G457211</i>	uncharacterized
S10_134585613	10	3.68 × 10 ⁻⁶	6.65	0.12	C/T	<i>GRMZM2G322582</i>	ATP binding protein
S10_134831452	10	4.78 × 10 ⁻⁷	8.22	0.14	A/G	<i>GRMZM2G181030</i>	MYB-related transcription factor family that regulates hypocotyl growth by regulating free auxin levels in a time-of-day specific manner (RVE1)

MAF—Minor allele frequency; the underlined SNP allele is minor allele; R²—proportion of phenotypic variance explained by SNP; ^a The exact physical position of the SNP can be inferred from marker's name, for example, S5_51353429: chromosome 5; 51,353,429 bp (Ref. Gen_v2 of B73).

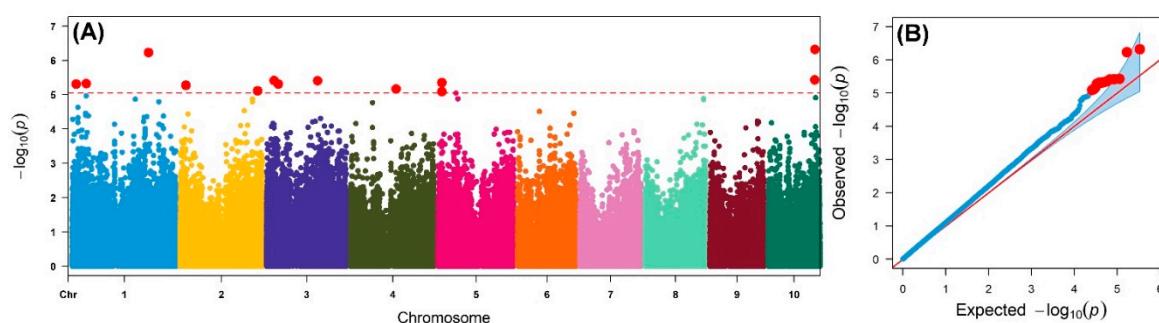


Figure 3. (A) Manhattan and q–q plots for the GWAS of common rust for the IMAS association mapping panel. The dashed horizontal line of the Manhattan plot depicts the significance threshold ($p < 9 \times 10^{-6}$). The X-axis indicates the SNP location along the 10 chromosomes, with chromosomes separated by different colors. (B) The red line on the q–q plot indicates a line of best fit with trait significance determined by a plot of expected- $\log_{10}(p)$ against observed- $\log_{10}(p)$.

Linkage map for each of the five populations was constructed. For each population, the number of progenies or families, markers, map lengths, and average genetic distances between the markers are presented in Table S3. Detection of QTL in the F_3 populations revealed seven QTLs in pop1, nine QTL in pop2, four QTL in pop3, five QTL in pop4 and one QTL in pop5, distributed on all chromosomes except chromosome 10 (Table 3). In pop1, QTLs individually explained 3–7% of phenotypic variance, in pop2, individual QTL explained 2–11% of total variation. In pop3, each QTL explained 5–9% of variation, whereas the range was 3–20% in pop4 and 17% in pop5. The total variance explained by each population ranged from 14 to 40%. Comparison of QTL detected among five populations revealed that QTL *qCR1-78* was detected in both pop1 and pop5 on chromosome 1. Another QTL *qCR3-151* detected in pop2 was within the confidence interval of the QTL *qCR3-113* detected in pop1 (Table 3). QTL *qCR9-117* detected in pop3 overlapped with QTL *qCR9-118* observed in pop2.

All five F_3 populations and IMAS panel was plotted by using first three principal components, which together contributed for 27.5% of the variation. PCA plot showed clear clustering of the five F_3 populations and IMAS panel into five clusters as shown in Figure 4. PCA distribution suggests pop 1 and pop 2 are more related to the IMAS panel compared to pop 3, 4, and 5. Joint linkage association mapping across five F_3 populations constituted 921 families, identified eighteen QTLs distributed in all chromosomes except on chromosome 8 (Table 4). These QTLs individually explained 0.3 to 3.1% and together explained 45% of the total phenotypic variance. Among the 18 QTL detected through JLAM, six QTLs, *qCR1-78*, *qCR1-227*, *qCR3-172*, *qCR3-186*, *qCR4-171*, and *qCR7-137* were also detected in individual population-based linkage mapping (Table 3). Out of these 18 QTLs, one QTL on chromosome 7 (*qCR7-10*) followed by QTL on chromosome 1 (*qCR1-78*) were strongly associated with CR disease severity in terms of p values.

Table 3. Detection of QTL associated with resistance to Common rust, their physical positions and genetic effects in five F3 populations.

QTL Name	Chr	Position (cM)	LOD	PVE (%)	Add	Dom	Total PVE (%)	Left Marker	Right Marker
CZL0618 × LaPostaSeqC7-F71-1-2-1-1B–Pop1									
<i>qCR1-78</i>	1	551	4.41	3.32	−0.20	−0.08		S1_77801418	S1_80167797
<i>qCR1-290</i>	1	799	2.82	3.59	0.18	−0.13		S1_290957469	S1_285979058
<i>qCR2-198</i>	2	187	3.17	6.70	−0.34	−0.53		S2_198394488	S2_230388748
<i>qCR3-113</i>	3	307	2.85	6.50	0.26	−0.24	14.95	S3_224567900	S3_113425715
<i>qCR6-38</i>	6	194	2.79	3.62	0.21	−0.01		S6_37902339	S6_63537451
<i>qCR6-63</i>	6	197	4.51	3.27	−0.21	−0.02		S6_63537451	S6_65299800
<i>qCR6-146</i>	6	446	3.09	7.03	0.30	−0.42		S6_147225115	S6_146382028
CZL074 × LaPostaSeqC7-F103-1-2-1-1B–Pop2									
<i>qCR2-137</i>	2	414	2.54	2.38	0.00	0.18		S2_158674609	S2_136562142
<i>qCR3-8</i>	3	262	3.33	2.97	0.63	0.52		S3_8300745	S3_8888914
<i>qCR3-151</i>	3	405	6.56	6.28	0.17	−0.14		S3_150831482	S3_166811360
<i>qCR4-198</i>	4	351	2.58	5.64	−0.19	−0.05		S4_197820294	S4_200964285
<i>qCR4-198</i>	4	354	5.77	8.85	0.24	−0.01	39.5	S4_200964285	S4_198430250
<i>qCR5-51</i>	5	374	2.93	3.7	−0.17	0.27		S5_186678634	S5_51355494
<i>qCR8-123</i>	8	165	4.69	6.23	0.20	0.00		S8_130213071	S8_123469991
<i>qCR9-118</i>	9	291	2.55	8.46	−0.23	0.03		S9_120748383	S9_118065757
<i>qCR9-12</i>	9	359	5.63	10.95	−0.28	−0.03		S9_12599819	S9_11929364
CZL00009 × CZL99017–Pop3									
<i>qCR1-18</i>	1	394	3	5.21	−0.13	−0.08		S1_19328973	S1_17679542
<i>qCR1-172</i>	1	501	2.81	6.09	0.67	−0.6	12.99	S1_196052894	S1_171534815
<i>qCR2-16</i>	2	93	3.61	5.84	0.7	−0.47		S2_16401968	S2_181538947
<i>qCR9-117</i>	9	300	4.31	8.35	−0.18	−0.04		S9_122035011	S9_116948078
CML505 × CZL99017–Pop4									
<i>qCR1-139</i>	1	182	3.87	20.16	0.32	0.67		S1_139463362	S1_227241027
<i>qCR4-171</i>	4	239	4.83	5.55	−0.23	0.01		S4_171215058	S4_173802342
<i>qCR7-137</i>	7	87	9.79	11.01	−0.34	0.22	23.89	S7_140894965	S7_137169719
<i>qCR9-47</i>	9	410	3.03	2.79	0.11	−0.24		S9_47064183	S9_58143264
<i>qCR9-90</i>	9	432	3.18	3	−0.13	−0.2		S9_90366846	S9_97737243
CZL0723 × CZL0724–Pop5									
<i>qCR1-77</i>	1	89	6.52	17.39	0.29	−0.09	14.28	S1_73375502	S1_77145631

Chr—Chromosome; LOD—logarithm of odds; Add—additive effect; Dom—dominance effect; PVE—phenotypic variance explained; Markers with bold letters are the QTL consistent in at least two across populations.

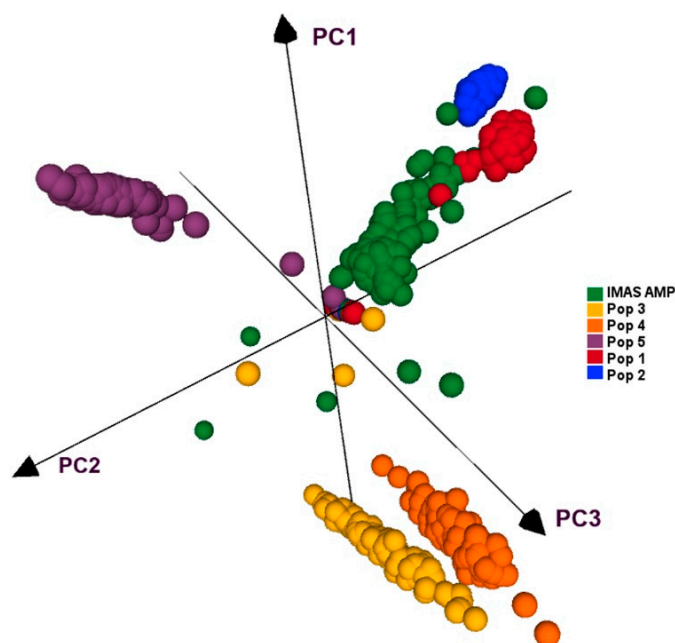


Figure 4. PCA analyses of five F3 populations and the IMAS panel based on GBS markers.

Table 4. Analysis of trait-associated markers, allele substitution (α) effects, and the total phenotypic variance (R^2) of the joint linkage association mapping based on combined 8 F3 populations.

Marker	QTL_Name ^a	Chrom	Pos	α -effect	p Value	PVE (%)	P_G
S1_77801418	<i>qCR1_78</i>	1	77.80	0.16	1.34×10^{-12}	2.9	7.2
S1_227241027	<i>qCR1_227</i>	1	227.24	0.17	3.75×10^{-7}	1.5	3.8
S2_20589802	<i>qCR2_20</i>	2	205.90	0.11	1.51×10^{-3}	0.6	1.5
S3_172332492	<i>qCR3_172</i>	3	172.33	-0.06	1.50×10^{-2}	0.3	0.7
S3_186725598	<i>qCR3_186</i>	3	186.73	0.10	9.88×10^{-3}	0.4	1
S4_828312	<i>qCR4_1</i>	4	0.83	-0.19	6.02×10^{-10}	2.2	5.5
S4_5238963	<i>qCR4_5</i>	4	5.24	0.12	7.45×10^{-8}	1.6	4
S4_171215058	<i>qCR4_171</i>	4	171.22	-0.05	3.79×10^{-2}	0.2	0.5
S5_2363546	<i>qCR5_2</i>	5	2.36	-0.06	2.03×10^{-2}	0.3	0.7
S6_32969273	<i>qCR6_32</i>	6	32.97	0.10	4.49×10^{-5}	0.9	2.2
S6_144280146	<i>qCR6_144</i>	6	144.28	0.06	2.57×10^{-2}	0.3	0.7
S6_154981658	<i>qCR6_155</i>	6	154.98	-0.19	8.37×10^{-4}	0.6	1.5
S7_10651847	<i>qCR7_10</i>	7	10.65	-0.28	1.78×10^{-13}	3.1	7.8
S7_13389227	<i>qCR7_13</i>	7	133.89	0.20	3.25×10^{-11}	2.5	6.2
S7_137335046	<i>qCR7_137</i>	7	137.34	-0.14	3.26×10^{-9}	2	5
S9_134919722	<i>qCR9_135</i>	9	134.92	-0.14	5.08×10^{-5}	0.9	2.2
S10_132612571	<i>qCR10_132</i>	10	132.61	-0.15	2.76×10^{-9}	2	5
S10_133744261	<i>qCR10_133</i>	10	133.74	-0.09	1.79×10^{-3}	0.5	1.2

^a QTL name composed by the trait code followed by the chromosome number in which the QTL was mapped and a physical position of the QTL, Chr—Chromosome, PVE—phenotypic variance explained, P_G —genotypic variance explained, The marker name indicates chromosome number followed by its physical position, for example S6_68207704 represents marker is in chromosome 6 at a position of 68,207,704bp (Ref. Gen_v2 of B73). Markers with bold letters are the QTL consistent with JLAM and linkage mapping.

We used the RR-BLUP model to predict the lines performance within each population. The prediction correlation was highest for pop4 ($r = 0.51$) followed by the IMAS panel ($r = 0.46$) and pop 2 ($r = 0.46$) and was low for pop3 with 0.19 (Figure 5A). However, the prediction through combined analysis of the five F₃ populations reported high improvement in the prediction accuracy with 0.78 for CR (Figure 5A). Prediction of biparental populations by using IMAS panel as training population reveals the correlations of 0.15, 0.05, -0.14, 0.22 and 0.07 for pop1, pop2, pop3, pop4, and pop5, respectively (Figure 5B).

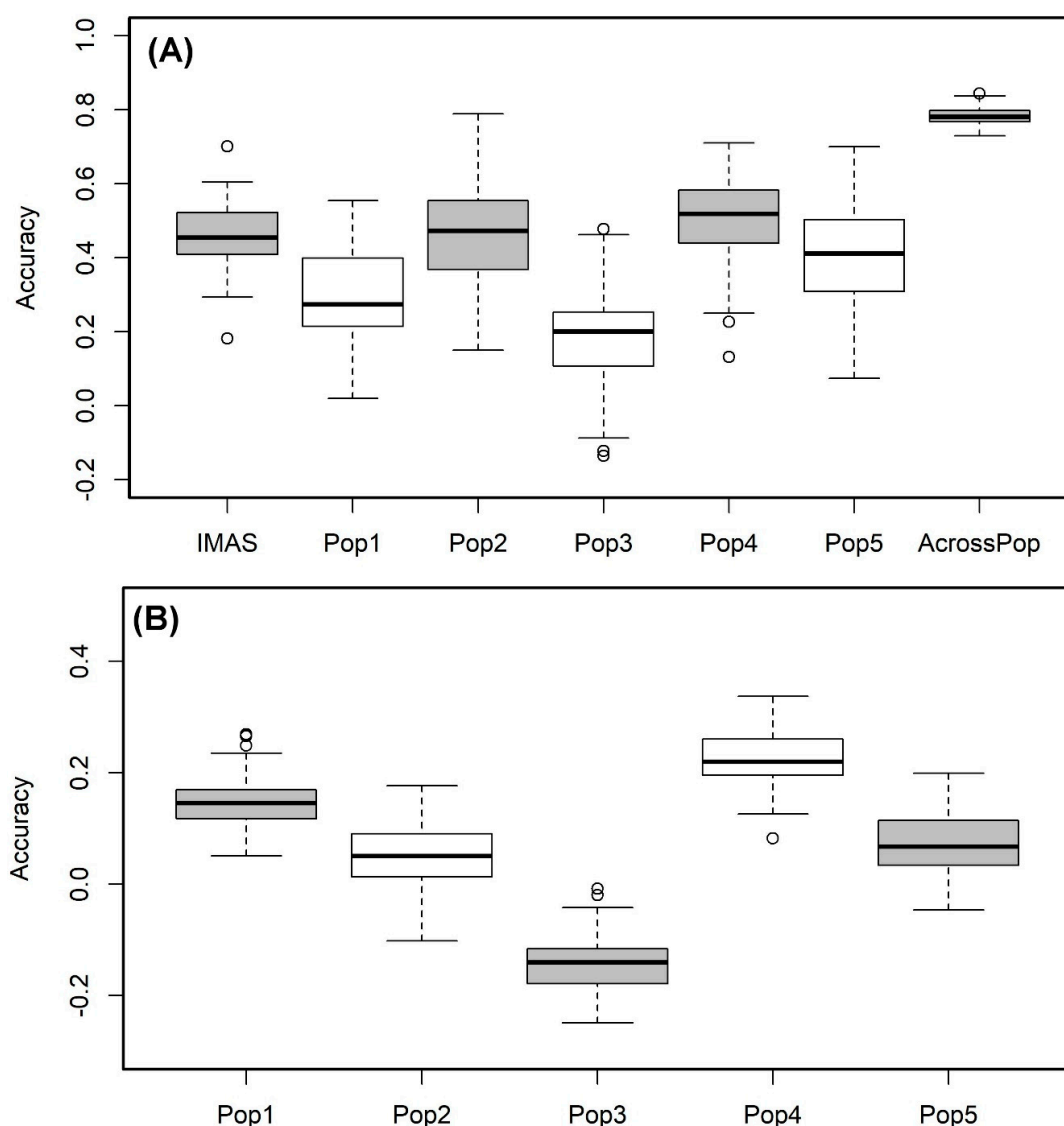


Figure 5. Genome-wide prediction correlations for CR resistance in bi-parental and IMAS AM panel based on two different scenarios. Scenario (A) estimation and prediction within IMAS AM panel and biparental populations and combined F_3 populations; Scenario (B): prediction of each F_3 population using IMAS AM panel as a training set with five-fold cross-validation.

3. Discussion

Over the past years, CR epidemics have mainly resulted from high levels of susceptibility among commercial maize hybrids [3]. Past efforts to exploit genetic resistance for CR have largely been through R genes [9,11]. A major drawback with genetic resistance through R genes is that it is overcome with time [12]. Quantitative resistance is by far a better option compared to single gene based resistance; however, few studies have been carried out on quantitative resistance to CR. This study therefore aimed to detect quantitative resistance to CR through GWAS by using IMAS association mapping panel, and QTL detection, JLAM, and GP using five F_3 populations.

Phenotypic analyses revealed normal distributions in all populations varied from highly resistance to moderate susceptibility, therefore indicating polygenic resistance in the materials used. Genotype and GxE interaction variances were significant in most populations which showed their importance in CR resistance. Heritability estimates were moderate in F_3 populations. High heritability observed in the IMAS association panel may be attributed to the large genetic diversity of the germplasm used.

Previous studies on the population structure of the IMAS association panel used in the present GWAS showed confounding structure in the panel posing a need to account for it [18,35]. The quantile–quantile plot of expected vs. observed log p values indicated that the mixed model used in this study effectively accounted for the population structure and kinship matrix. Ideally, the p values are expected to follow the diagonal plot of the expected vs. observed p values assuming no marker trait associations. In real datasets, the markers are expected to deviate to the left side of the plot as a result of true marker trait associations, however, strong deviations to the left indicate inflated false positives while to the right indicate deflated false negatives [25].

Association mapping resolution also strongly depends on the LD of the particular population under study since it exploits the historical recombinations that occur in natural populations [36]. Rapid LD decay as shown in the IMAS population shows the significant diversity in the panel and its suitability for GWAS. The small mapping distance of 50kb and rapid decay at $r^2 = 0.1$ enable the detection of marker trait associations with small effects that might have otherwise been missed with other mapping approaches.

In the GWAS, 14 significant SNPs were associated with *P. sorghi* resistance across three environments (Table 2, Figure 3), of which some overlapped with previously reported QTL intervals [12,31,37]. Specifically, SNP *S3_21856582* detected in this study is also reported earlier [12] and interestingly it is also overlapped with two biparental based QTL studies [10,13] which suggests the possibility of potential candidate gene/marker for CR resistance across genetic backgrounds and environments. Another two SNPs on chromosome 5, *S5_10087070* and *S5_10089138* are found close to SNP *S5_10,055,423* found in the previous association mapping study in tropical maize germplasm [31]. Other SNPs detected in this study seems to be novel and add for the new source of resistance for CR.

The candidate genes in the present study were identified as encoding transcription factors (TFs), ATP binding and intracellular signaling. A putative candidate gene GRMZM2G181030 has been linked with the MYB superfamily which is one of the three largest TF families in maize [38]. The important roles of MYB superfamily are in developmental processes and defense responses in plants [39]. GRMZM2G322582 was encoded as ATP binding protein which play important roles in membrane transport, cellular motility and regulation of various metabolic processes [40]. GRMZM2G086484 was linked to the Pleckstrin homology (PH) domain superfamily protein which plays a role in recruiting proteins to different membranes, thus targeting them to appropriate cellular compartments or enabling them to interact with other components of the signal transduction pathways [41]. GRMZM2G181002 was annotated in protein kinase superfamily specifically the phosphotransferases, serine and threonine specific kinase superfamily. Serine/threonine kinase receptors play a role in the regulation of cell proliferation, programmed cell death (apoptosis), cell differentiation, and embryonic development [42]. Lastly, GRMZM2G009188 was encoded as 11-beta-hydroxysteroid dehydrogenase 1B whose role is in the regulation of energy metabolism and immune system by locally reactivating glucocorticoids predominantly by catalyzing the reduction of cortisone to cortisol in intact cells that also express hexose-6-phosphate dehydrogenase (H6PDH), which provides cofactor NADPH [43].

The linkage analysis within the individual bi-parental populations was also carried out revealing 23 QTLs associated with CR resistance in maize and were distributed on all chromosomes (Table 3). Linkage groups conferring CR resistance have been previously reported on bin 1.05, 1.06, 2.05, 3.04, 5.02, 6.04, 8.03, 8.05, 10.1, 10.5, and 10.6 [7,10,12,13,31]. The present study's results corroborate these earlier studies, as most of the identified QTLs fell in those bins of linkage groups. Within the current study we also found several QTLs overlapping across populations with linkage mapping, across methods with linkage analyses, JLAM, and GWAS (Table 2, Table 3, Table 4). QTL *qCR1-139* detected on pop4 was co-located with QTL detected in F₃ pop 3 and F₃ pop 5, which suggests one of possible potential candidate for CR resistance in chromosome 1 across different genetic back grounds.

Identification of genomic regions for CR resistance have previously been carried out in biparental-based linkage analysis [7,10,13,30] and natural population-based association mapping [12,31]. The current study consequently sought to exploit JLAM using multiple segregating populations in

conjunction with QTL analysis and GWAS. GWAS and JLAM increased the resolution within the confidence intervals of QTL for CR resistance. *qCR1-139*, which explained >20% of the phenotypic variance, was a locus with major effect based on genetic linkage mapping, and was likely to consist of two QTL as revealed by GWAS and JLAM results. SNP *S1_220067760* identified through GWAS and SNP *S1_227241027* identified through JLAM fall within the confidence interval of QTL *qCR1-139* (Tables 2–4). The significant marker *S2_16361185* detected through GWAS and marker *S2_20589802* (*qCR2-20*) detected through JLAM fall within the confidence interval of another QTL on chromosome 2, *qCR2-16* detected in pop3, which helps to increase the precision of QTL position. Fractionation of single major QTL was also reported previously in maize [44]. Thus, the combination of linkage mapping, JLAM and GWAS approaches helped to further refine the CR resistance loci to an extent that it became possible to separate the effects of two co-segregating QTL (on chromosome 1 and 2). Thus, GWAS and JLAM can serve as a great complementary tool to identify and validate the molecular markers linked to QTL and to be used in applied breeding. Nevertheless, it is warranted to validate through fine mapping and gene cloning coupled with functional genomics studies, in order to clarify further on QTL intervals identified and refined in this study.

Since GS was demonstrated to be useful in plant breeding, there have been many studies that demonstrate the utility of GP in breeding for disease resistance in crops [26,45–49]. Prediction correlations for within association panel and biparental populations is comparable to the other disease reported in maize [18,24,50]. As described in earlier study [51], the benefits of large populations or combined population predictions were higher compared to individual population-based predictions. This corroborates with the present study whereby combined population prediction was higher over individual or smaller population-based predictions (Figure 5A).

The success of GP gains more attention when the training population is related but independent of testing or validating populations. In the current study, since the parental lines of F₃ populations are part of the IMAS association panel, we tried to use IMAS panel as an independent training population and predict the F₃ populations for their performance for CR resistance. We observed, for four out of five populations, prediction correlations of 0.05, 0.07, 0.15, and 0.22 (Figure 4), which is comparable to the prediction correlations observed for maize lethal necrosis [26] under similar testing environments. Overall, results from our study encourages to develop an independent but related diverse population as training population to predict CR resistance. This also supports the use of GEBVs as indirect selection criteria, at least to remove the lines with poor performance for CR resistance. Conducting separate field trials for CR screening in the breeding pipelines always takes additional resources. The model's predictive ability, with cross validation, showed moderate prediction accuracies for CR in pop1 and pop4 were comparable to the correlations observed under similar scenario of predicting biparental populations by using association mapping panel for maize lethal necrosis and maize chlorotic mottle virus resistance [26]. Overall, to have an effective independent training population, relatedness between training and validation population is critical as well as the trait architecture and heritability. Compared to abiotic stress related traits, traits like CR and many other maize diseases are relatively less complex and our results suggests that it is possible to have an independent training population with continuous improvement and used routinely in breeding program.

In conclusion, we used five F₃ populations and one IMAS association mapping panel, together comprising 1300 lines, to unravel the genetic architecture of CR resistance. In this study we identified new QTLs as well as reconfirmed the QTLs reported in earlier studies for CR resistance in tropical and subtropical maize germplasm. Linkage mapping identified 26 QTLs with four major QTLs, which explained >10% of phenotypic variance. The detected QTLs were validated with GWAS, and several SNPs were found overlapping with the identified QTLs through either linkage mapping or JLAM. These genomic regions can serve as potential sources to improve resistance to CR. GP can be used in combined populations to predict the response of the germplasm to CR resistance. Having a common training population derived from intensively phenotyped and genotyped lines with diverse representation from a breeding program holds promise in breeding for CR resistance.

4. Materials and Methods

4.1. Plant Materials and Trial Design

A collection of 380 tropical and subtropical maize inbred lines denoted as the Improved Maize for African Soils (IMAS) panel, representing some of the genetic diversity available in CIMMYT's and several national research programs from Africa and Mexico breeding programs (low N, drought, and biotic stresses) was used in this study [26,35]. In addition five F₃ populations namely Pop1 (CZL0618 × LaPostaSeqC7-F71-1-2-1-1B, *n* = 183), Pop2 (CZL074 × LaPostaSeqC7-F103-1-2-1-1B, *n* = 174), Pop3 (CZL00009 × CZL99017, *n* = 187), Pop4 (CML505 × CZL99017, *n* = 189) and Pop5 (CZL0723 × CZL0724, *n* = 188) were used in this study [52]. These five F₃ populations were used to study the grain yield under optimum and drought stress conditions [52,53]. The IMAS panel was evaluated in three location-year combinations, at Kitale (1.0191° N 35.0023° E, 1900 meters above sea level (masl)) in 2013 and 2014 and at Kakamega (0°17'3.19" N 34°45'8.24" E, 1535 masl) in 2014, and F₃ pop 2 was evaluated in three locations Kakamega, Kitale, and Embu (0°31'52" S 37°27'02" E, 1406 masl) and the remaining four F₃ populations were evaluated in two locations in Kakamega and Embu at 2011 in Kenya (Tables S1 and S2).

All the lines from IMAS panel and five F₃ populations were planted in a 4-m-long single row plots in an alpha lattice design with two replications in each location. Two seeds were planted per hill and thinned to a single plant per hill, three weeks after emergence to ensure uniform plant density. Standard agronomic practices were followed. The chosen locations were natural hotspots for foliar diseases including CR; good disease infection pressure across the trials at each location was observed. The IMAS panel, and F₃ populations were evaluated for their responses to CR in two to three environments (Tables S1 and S2). CR disease severity data was recorded after flowering and scored plot-wise on an ordinal scale of 1 (no rust, highly resistant, without disease symptoms) to 9 (highly susceptible, most severe). On the IMAS panel, in addition to CR severity scoring, data were also collected for anthesis date (AD). The populations used in this study, methods applied, and the different type of analyses used in this study are presented as a general analyses diagram (Figure 6).

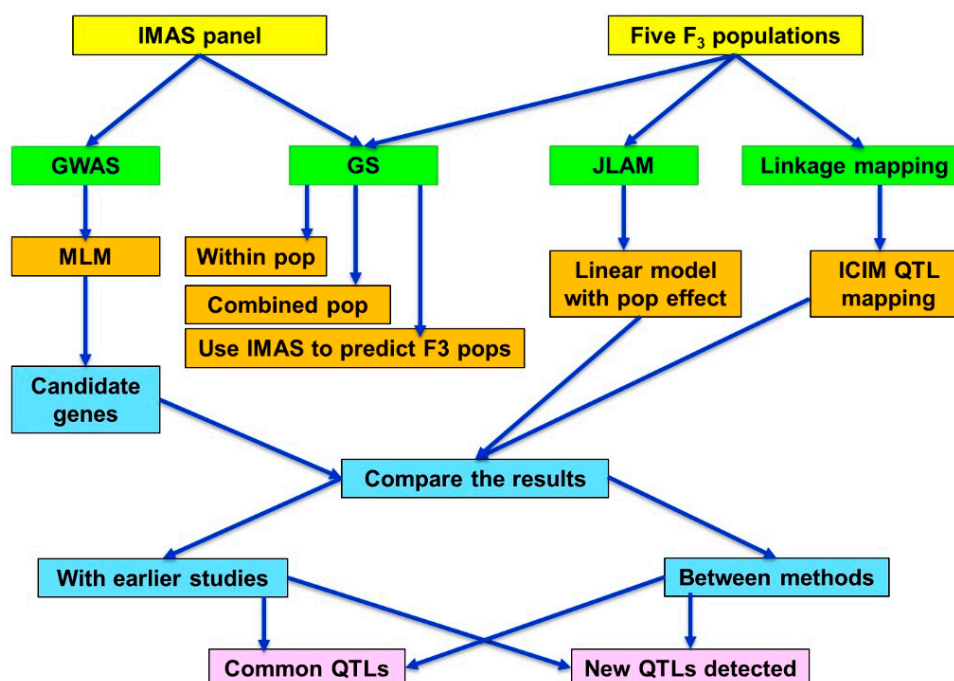


Figure 6. Outline of the general flow chart on population used in this study and the analyses methods applied. GWAS—Genome-wide association study, GS—Genomic selection, JLAM—Joint linkage association mapping, MLM—mixed linear model, ICIM—Inclusive composite interval mapping.

4.2. Phenotypic and Genotypic Data Analysis

For IMAS panel and each F₃ populations, each location-year combinations were treated as an independent environment which resulted into three environments for IMAS panel, and F₃ pop2, whereas the number of environments were two for F₃ pop1, F₃ pop3, F₃ pop4, and F₃ pop5. Analysis of variance for individual and across environments was undertaken using the ASREML-R [54] for each bi-parental population and the IMAS panel. The following statistical model was used to estimate the variance components (Equation (1)):

$$CR_{mnop} = \mu + \text{Genotype}_m + \text{Env}_n + (\text{Genotype} \times \text{Env})_{mn} + \text{Rep}(\text{Env})_{on} + \text{Block}(\text{Rep}.\text{Env})_{pno} + e_{mnop} \quad (1)$$

where CR_{mnop} is the phenotypic performance of the m th genotype at the n th environment in the o th replication of the p th incomplete block, μ is an intercept term, Genotype_m is the genetic effect of the m th genotype, Env_n is the effect of the n th environment, $(\text{Genotype} \times \text{Env})_{mn}$ is the interaction effect between genotype and environment, $\text{Rep}(\text{Env})_{on}$ is the effect of the o th replication at the n th environment, $\text{Block}(\text{Rep}.\text{Env})_{pno}$ is the effect of the p th incomplete block in the o th replication at the n th environment, and e_{mnop} is the residual. The genotypic effect, genotype by environment interaction and effect of incomplete blocks were treated as random effects in order to estimate their variances and residual error variance. Environments and replications were treated as fixed effects. Assuming fixed genotypic effects, a mixed linear model was fitted to obtain the best linear unbiased estimates (BLUEs). Significance of variance components were tested by model comparison with likelihood ratio tests in which the halved p -values were used as an approximation. Broad-sense heritability (H^2) was calculated for all the traits using the following Equation (2):

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{gl}^2}{l} + \frac{\sigma_\epsilon^2}{lr}} \quad (2)$$

where σ_g^2 is the genotype variance; σ_{gl}^2 is the genotype \times environment interaction variance; and σ_ϵ^2 is the error variance, l represents number of environments and r for number of replications. META-R software [55] was used to obtain best linear unbiased prediction (BLUP) for each genotype across environments. Combined analyses of the five F₃ populations were carried out in META-R.

The F₃ populations lines as well as their parents, and the IMAS association panel inbred lines were genotyped with Genotyping by Sequencing (GBS). DNA of all lines was extracted from 3–4 leaves stage seedlings and genotyped using GBS platform at the Institute for Genomic Diversity, Cornell University, Ithaca, USA as per the procedure described in earlier studies [56]. The ~955K GBS SNP datasets were filtered where a minor allele frequency of <0.05, heterozygosity of >5% and missing data rates >10% were excluded from further analysis in TASSEL ver 5.2 [57].

4.3. PCA and Linkage Disequilibrium

The principal components (PC) of the five F₃ populations and IMAS panel were estimated using TASSEL ver 5.2 and visualized using R software version 3.2.5 (<https://www.r-project.org/>) to obtain the explained variance of each PCs. Further, the LD of the IMAS panel was calculated using Tassel version 5.2 to obtain the rate of LD decay in the population. LD decay rate between each pair of SNPs was analyzed with the squared Pearson correlation coefficient (r^2). The rate of LD decay with physical distance was visualized and average pairwise distances at which LD decayed at $r^2 = 0.1$ and 0.2 were calculated in R software. The 'nlin' function in R was used to fit non-linear models into the genome-wide LD data by incorporating r^2 as responses (y axis) and pairwise distances (x axis) as predictors. The average estimator for LD decay was calculated at 'significance' threshold of $r^2 = 0.1$ and $r^2 = 0.2$ cutoff points in relation to distance [58] and a representative scatter plot was drawn as LD between adjacent markers versus chromosome distance (kb).

4.4. Genome-wide Association Analyses

The BLUP values obtained for CR were used in GWAS as phenotypes. The kinship matrix obtained with a centered identity by state (IBS) and the first five PCs which explained maximum variation were used to correct the population structure in a mixed linear model using TASSEL version 5.2. Genome wide scans for marker-trait associations were conducted with mixed linear model. The amount of phenotypic variation explained by the model was assessed using the R^2 statistics, calculated by fitting all significant SNPs simultaneously in a linear model in R. To determine the significance threshold, multiple testing correction was conducted where the total number of tests were estimated based on the average extent of LD at $r^2 = 0.1$ [19]. With respect to the above, significant associations were declared when p values in independent tests were less than 9×10^{-6} . The 50 bp source sequences of the significantly associated SNPs were used to perform BLAST searches against the B73 RefGen v2 genome set in Maize GDB (<http://www.maizegdb.org>). The putative candidate genes identified in Maize GDB were within or adjacent to each associated SNP.

4.5. Detection of QTLs and Joint Linkage Association Mapping

The number of SNPs was further reduced by selecting homozygous and polymorphic markers between the parents in each population and were further filtered based on minimum distance between adjacent SNPs as ≥ 200 kilo base pairs (Kbps). For JLAM, markers from all five F_3 populations were combined, and markers with $<1\%$ missing value and $>5\%$ MAF and heterozygosity of $<5\%$ were retained. Finally, a set of 5000 SNPs that are uniformly distributed across the genome were used for JLAM analyses.

QTL IciMapping version 4.1 [59] was used to construct the linkage map based on data from all five biparental populations. QTL IciMapping was used to remove the highly correlated SNPs that do not provide any additional information by using an inbuilt tool BIN. This resulted in retention of 1130, 1047, 1099, 1122, and 1081 high-quality SNPs in pop1, pop2, pop3, pop4, and pop5, respectively (Table S3). BLUP values across environments for the F_3 populations were used in QTL detection analysis using inclusive composite interval mapping (ICIM). The probability in the stepwise regression was set at 0.01 and the scanning step was 1 cM. For determination of QTL significance, the threshold LOD score was set to >2.5 by using 1000 permutations and a p value ≤ 0.05 . The phenotypic variation explained (PVE) by each QTL and across all QTLs for each trait was estimated. The origin of the favorable allele for CR resistance was identified based on the sign of the additive effects of each QTL.

A set of 5000 high-quality GBS SNPs derived from the genotyped F_3 populations were used in JLAM analyses. BLUPs calculated across populations and environments were used as phenotypes. To best carry out association mapping in multiple biparental populations, a two-step biometric model which incorporates population effect, cofactors and a marker effect across populations was used to detect QTL [60,61]. This model was explained in detail by Liu et al. [61] and Würschum et al. [60]. The first step involved selection of cofactors based on Schwarz Bayesian Criterion (SBC) [62] by including a population effect and cofactors carried out using PROC GLM SELECT implemented in the statistical software SAS 9.4 [63]. In the second step, p values for the F -test were calculated using a full model (including SNP effect) versus a reduced model (without SNP effect). Genome-wide scans for QTLs were implemented in R version 3.2.5 (R Development Core Team, Vienna, Austria; 2015). The Bonferroni–Holm procedure [64] was used to detect markers with significant ($p < 0.05$) main effects and was controlled for multiple testing. The total proportion of PVE by the detected QTLs was calculated by fitting all significant SNPs simultaneously in a linear model to obtain an adjusted R^2 [65].

4.6. Genomic Prediction

GP was carried out with ridge regression BLUP [66] within and across the five F_3 populations as well as in the IMAS panel for CR resistance at five-fold cross-validation. BLUEs across environments for each of the biparental populations and across five F_3 populations were used for the analysis.

For all biparental populations and IMAS panel, same set of high-quality uniformly distributed 5000 SNPs with no missing values and MAF > 0.05 were used. Prediction scenarios used for GP include: ‘within population’ where training and validation sets are derived from within individual biparental population and IMAS panel; a ‘combined population’ prediction approach where data from five F₃ populations were combined and sampled randomly to form validation and training set; and ‘across population’ prediction in which IMAS association panel was used as a training set and each F₃ population as a validation population. For each approach, 100 iterations were done for sampling of the training and validation sets.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1422-0067/21/18/6518/s1>.

Author Contributions: Conceptualization: M.K., B.D., D.M., B.M.P., and M.G.; methodology: B.D., M.G., S.K.N., C.N., and M.K.; software: C.N. and M.K.; validation: M.K., M.S.O., D.M., Y.B., and M.G.; formal analysis: M.K., C.N., and M.G.; investigation: M.G. and J.K.; resources: B.D., S.K.N., B.M.P., M.S.O., and M.G.; data curation: M.K., C.N., and M.G.; writing—original draft preparation: M.K., S.K.N., M.G., and C.N.; writing—review and editing: M.G., B.D., D.M., Y.B., M.S.O., B.M.P., M.K., J.M.B., S.L.M., J.K., and S.K.N.; visualization: M.G. and S.K.N.; supervision: M.G., J.K., M.S.O., S.L.M., and B.M.P.; project administration: M.G. and J.K.; funding acquisition: B.M.P. and M.S.O. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Bill and Melinda Gates Foundation (B&MGF), and the United States Agency for International Development (USAID) through the Stress Tolerant Maize for Africa (STMA, B&MGF Grant # OPP1134248) Project, AGGMW (Accelerating Genetic Gains in Maize and Wheat for Improved Livelihoods, B&MGF Investment ID INV-003439) project, Improved Maize for African Soils (IMAS) Project, and the CGIAR Research Program on Maize (MAIZE). MAIZE receives W1&W2 support from the Governments of Australia, Belgium, Canada, China, France, India, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Sweden, Switzerland, U.K., U.S., and the World Bank.

Acknowledgments: The study was managed by The Global maize Program, CIMMYT. We thank CIMMYT technicians for data collection at experimental sites. Our sincere thanks to CIMMYT technicians in Kenya; and we also thank Edward S. Buckler and the Institute of Genomic Diversity, Cornell University, for the high-density genotyping (GBS) and imputation service.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Dey, U.; Harlapur, S.I.; Dhutraj, D.N.; Suryawanshi, A.P.; Bhattacharjee, R. Integrated disease management strategy of common rust of maize incited by *Puccinia sorghi* Schw. *Afr. J. Microbiol. Res.* **2015**, *9*, 1345–1351.
- Dey, U.; Harlapur, S.I.; Dhutraj, D.N.; Suryawanshi, A.P.; Badgujar, S.L.; Jagtap, G.P.; Kuldhar, D.P. Spatiotemporal yield loss assessment in corn due to common rust caused by *Puccinia sorghi* Schw. *Afr. J. Agric. Res.* **2012**, *7*, 5265–5269. [[CrossRef](#)]
- Crautam, P.; Stein, J. Induction of systemic acquired resistance to *Puccinia sorghi* in corn. *Inter. J. Plant Pathol.* **2011**, *2*, 43–50.
- Chandler, M.A.; Tracy, W.F. Vegetative phase change among sweet corn (*Zea mays* L.) hybrids varying for reaction to common rust (*Puccinia sorghi* Schw.). *Plant Breed.* **2007**, *126*, 569–573. [[CrossRef](#)]
- Abedon, B.G.; Tracy, W.F. Corngrass 1 of maize (*Zea mays* L.) delays development of adult plant resistance to common rust (*Puccinia sorghi* Schw.) and European corn borer (*Ostrinia nubilalis* Hubner). *J. Hered.* **1996**, *87*, 219–223. [[CrossRef](#)]
- Hurkman, M.M. Vegetative Phase Transition and Response to Disease in Maize (*Zea mays* L.). Ph.D. Thesis, University of Wisconsin, Madison, WI, USA, 2003. Available online: <https://elibrary.ru/item.asp?id=5282527> (accessed on 15 August 2020).
- Danson, J.; Lagat, M.; Kimani, M.; Kuria, A. Quantitative trait loci (QTLs) for resistance to gray leaf spot and common rust diseases of maize. *African, J. Biotechnol.* **2008**, *7*, 3247–3254.
- Ayiga-Aluba, J.; Edemal, R.; Tusiime, G.; Asea, G.; Gibson, P. Response to two cycles of S1 recurrent selection for turicum leave blight in an open pollinated maize variety population (Longe 5). *Adv. Appl. Sci. Res.* **2015**, *6*, 4–12.
- Delaney, D.E.; Webb, C.A.; Hulbert, S.H. A novel rust resistance gene in maize showing overdominance. *Mol. plant-microbe Interact.* **1998**, *11*, 242–245. [[CrossRef](#)]

10. Lübberstedt, T.; Klein, D.; Melchinger, A.E. Comparative quantitative trait loci mapping of partial resistance to Puccinia sorghi across four populations of European flint maize. *Phytopathology* **1998**, *88*, 1324–1329. [[CrossRef](#)]
11. Hooker, A.L. Corn and sorghum rusts. In *Diseases, Distribution, Epidemiology, and Control*; Elsevier: Amsterdam, The Netherlands, 1985; pp. 207–236.
12. Olukolu, B.A.; Tracy, W.F.; Wisser, R.; De Vries, B.; Balint-Kurti, P.J. A Genome-Wide Association Study for Partial Resistance to Maize Common Rust. *Phytopathology* **2016**, *106*, 745–751. [[CrossRef](#)]
13. Brown, A.F.; Juvik, J.A.; Pataky, J.K. Quantitative trait loci in sweet corn associated with partial resistance to Stewart's wilt, northern corn leaf blight, and common rust. *Phytopathology* **2001**, *91*, 293–300. [[CrossRef](#)] [[PubMed](#)]
14. Kump, K.L.; Bradbury, P.J.; Wisser, R.J.; Buckler, E.S.; Belcher, A.R.; Oropeza-Rosas, M.A.; Zwonitzer, J.C.; Kresovich, S.; McMullen, M.D.; Ware, D. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* **2011**, *43*, 163. [[CrossRef](#)] [[PubMed](#)]
15. Zila, C.T.; Fernando Samayoa, L.; Santiago, R.; Butrón, A.; Holland, J.B. A Genome-Wide Association Study Reveals Genes Associated with Fusarium Ear Rot Resistance in a Maize Core Diversity Panel. *G3 (Bethesda)* **2013**, *3*, 2095–2104. [[CrossRef](#)] [[PubMed](#)]
16. Laidò, G.; Marone, D.; Russo, M.A.; Colecchia, S.A.; Mastrangelo, A.M.; De Vita, P.; Papa, R. Linkage Disequilibrium and Genome-Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum* L.). *PLoS ONE* **2014**, *9*, e95211. [[CrossRef](#)]
17. Ding, J.; Ali, F.; Chen, G.; Li, H.; Mahuku, G.; Yang, N.; Narro, L.; Magorokosho, C.; Makumbi, D.; Yan, J. Genome-wide association mapping reveals novel sources of resistance to northern corn leaf blight in maize. *BMC Plant Biol.* **2015**, *15*, 206. [[CrossRef](#)]
18. Gowda, M.; Das, B.; Makumbi, D.; Babu, R.; Semagn, K.; Mahuku, G.; Olsen, M.S.; Bright, J.M.; Beyene, Y.; Prasanna, B.M. Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theor. Appl. Genet.* **2015**, *128*, 1957–1968. [[CrossRef](#)]
19. Cui, Z.; Luo, J.; Qi, C.; Ruan, Y.; Li, J.; Zhang, A.; Yang, X.; He, Y. Genome-wide association study (GWAS) reveals the genetic architecture of four husk traits in maize. *BMC Genomics* **2016**, *17*, 946. [[CrossRef](#)]
20. Gustafson, T.J.; de Leon, N.; Kaeppler, S.M.; Tracy, W.F. Genetic Analysis of Sugarcane mosaic virus Resistance in the Wisconsin Diversity Panel of Maize. *Crop Sci.* **2018**, *58*, 1853–1865. [[CrossRef](#)]
21. Liu, N.; Xue, Y.; Guo, Z.; Li, W.; Tang, J. Genome-Wide Association Study Identifies Candidate Genes for Starch Content Regulation in Maize Kernels. *Front. Plant Sci.* **2016**, *7*, 1046. [[CrossRef](#)]
22. Kuki, M.C.; Scapim, C.A.; Rossi, E.S.; Mangolin, C.A.; do Amaral Junior, A.T.; Pinto, R.J.B. Genome wide association study for gray leaf spot resistance in tropical maize core. *PLoS ONE* **2018**, *13*, e0199539. [[CrossRef](#)]
23. Chaikam, V.; Gowda, M.; Nair, S.; Melchinger, A.; Prasanna, B. Genome-wide association study to identify genomic regions influencing spontaneous fertility in maize haploids. *Euphytica* **2019**, *215*, 138. [[CrossRef](#)] [[PubMed](#)]
24. Nyaga, C.; Gowda, M.; Beyene, Y.; Muriithi, W.; Makumbi, D.; Olsen, M.; Suresh, L.M.; Bright, J.; Das, B.; Prasanna, B. Genome-Wide Analyses and Prediction of Resistance to MLN in Large Tropical Maize Germplasm. *Genes* **2019**, *11*, 16. [[CrossRef](#)] [[PubMed](#)]
25. Jaiswal, V.; Gupta, S.; Gahlaut, V.; Muthamilarasan, M.; Bandyopadhyay, T.; Ramchiary, N.; Prasad, M. Genome-Wide Association Study of Major Agronomic Traits in Foxtail Millet (*Setaria italica* L.) Using ddRAD Sequencing. *Sci. Rep.* **2019**, *9*, 5020. [[CrossRef](#)] [[PubMed](#)]
26. Sitonik, C.; Suresh, L.M.; Beyene, Y.; Olsen, M.S.; Makumbi, D.; Oliver, K.; Das, B.; Bright, J.M.; Mugo, S.; Crossa, J.; et al. Genetic architecture of maize chlorotic mottle virus and maize lethal necrosis through GWAS, linkage analysis and genomic prediction in tropical maize germplasm. *Theor. Appl. Genet.* **2019**, *132*, 2381–2399. [[CrossRef](#)] [[PubMed](#)]
27. Myles, S.; Peiffer, J.; Brown, P.J.; Ersoz, E.S.; Zhang, Z.; Costich, D.E.; Buckler, E.S. Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell* **2009**, *21*, 2194–2202. [[CrossRef](#)] [[PubMed](#)]
28. Holland, J.B. Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* **2007**, *10*, 156–161. [[CrossRef](#)] [[PubMed](#)]

29. Kang, H.M.; Zaitlen, N.A.; Wade, C.M.; Kirby, A.; Heckerman, D.; Daly, M.J.; Eskin, E. Efficient control of population structure in model organism association mapping. *Genetics* **2008**, *178*, 1709–1723. [[CrossRef](#)]
30. Kerns, M.R.; Dudley, J.W.; Rufener, G.K. QTL for resistance to common rust and smut in maize. *Maydica* **1999**, *44*, 37–45.
31. Zheng, H.; Chen, J.; Mu, C.; Makumbi, D.; Xu, Y.; Mahuku, G. Combined linkage and association mapping reveal QTL for host plant resistance to common rust (*Puccinia sorghi*) in tropical maize. *BMC Plant Biol.* **2018**, *18*, 310. [[CrossRef](#)]
32. Lorenz, A.J.; Chao, S.; Asoro, F.G.; Heffner, E.L.; Hayashi, T.; Iwata, H.; Smith, K.P.; Sorrells, M.E.; Jannink, J.L. Genomic Selection in Plant Breeding: Knowledge and Prospects. *Adv. Agron.* **2011**, *110*, 77–123.
33. Crossa, J.; Pérez-Rodríguez, P.; Cuevas, J.; Montesinos-López, O.; Jarquín, D.; de los Campos, G.; Burgueño, J.; González-Camacho, J.M.; Pérez-Elizalde, S.; Beyene, Y.; et al. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Sci.* **2017**, *22*, 961–975. [[CrossRef](#)] [[PubMed](#)]
34. Massman, J.M.; Gordillo, A.; Lorenzana, R.E.; Bernardo, R. Genomewide predictions from maize single-cross data. *Theor. Appl. Genet.* **2013**, *126*, 13–22. [[CrossRef](#)] [[PubMed](#)]
35. Ertiro, B.T.; Labuschagne, M.; Olsen, M.; Das, B.; Prasanna, B.M.; Gowda, M. Genetic dissection of nitrogen use efficiency in tropical maize through genome-wide association and genomic prediction. *Front. Plant Sci.* **2020**, *11*, 474. [[CrossRef](#)] [[PubMed](#)]
36. Kaler, A.; Gillman, J.; Beissinger, T.; Purcell, L. Comparing Different Statistical Models and Multiple Testing Corrections for Association Mapping in Soybean and Maize. *Front. Plant Sci.* **2020**, *10*, 1794. [[CrossRef](#)]
37. Wissler, R.J.; Balint-Kurti, P.J.; Nelson, R.J. The genetic architecture of disease resistance in maize: A synthesis of published studies. *Phytopathology* **2006**, *96*, 120–129. [[CrossRef](#)]
38. Jiang, Y.; Zeng, B.; Zhao, H.; Zhang, M.; Xie, S.; Lai, J. Genome-wide Transcription Factor Gene Prediction and their Expressional Tissue-Specificities in Maize, F.J. *Integr. Plant Biol.* **2012**, *54*, 616–630. [[CrossRef](#)]
39. Chen, Y.; Yang, X.; He, K.; Liu, M.; Li, J.; Gao, Z.; Lin, Z.; Zhang, Y.; Wang, X.; Qiu, X.; et al. The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family. *Plant Mol. Biol.* **2006**, *60*, 107–124.
40. Chauhan, J.S.; Mishra, N.K.; Raghava, G.P.S. Identification of ATP binding residues of a protein from its primary sequence. *BMC Bioinformatics* **2009**, *10*, 434. [[CrossRef](#)]
41. Wang, D.S.; Shaw, R.; Winkelmann, J.C.; Shaw, G. Binding of PH domains of β -adrenergic-receptor kinase and β -spectrin to WD40/ β -transducin repeat containing regions of the β -subunit of trimeric G-proteins. *Biochem. Biophys. Res. Commun.* **1994**, *203*, 29–35. [[CrossRef](#)]
42. Stone, J.M.; Walker, J.C. Plant protein kinase families and signal transduction. *Plant Physiol.* **1995**, *108*, 451–457. [[CrossRef](#)]
43. Nashev, L.G.; Chandsawangbhuwana, C.; Balazs, Z.; Atanasov, A.G.; Dick, B.; Frey, F.J.; Baker, M.E.; Odermatt, A. Hexose-6-phosphate Dehydrogenase Modulates 11β -Hydroxysteroid Dehydrogenase Type 1-Dependent Metabolism of 7-keto- and 7β -hydroxy-neurosteroids. *PLoS ONE* **2007**, *2*, e561. [[CrossRef](#)] [[PubMed](#)]
44. Mammadov, J.; Sun, X.; Gao, Y.; Ochsenfeld, C.; Bakker, E.; Ren, R.; Flora, J.; Wang, X.; Kumpatla, S.; Meyer, D. Combining powers of linkage and association mapping for precise dissection of QTL controlling resistance to gray leaf spot disease in maize (*Zea mays* L.). *BMC Genomics* **2015**, *16*, 916. [[CrossRef](#)]
45. Heffner, E.L.; Sorrells, M.E.; Jannink, J. Genomic selection for crop improvement. *Crop Sci.* **2009**, *49*, 1–12. [[CrossRef](#)]
46. Lorenz, A.J.; Smith, K.P.; Jannink, J. Potential and optimization of genomic selection for Fusarium head blight resistance in six-row barley. *Crop Sci.* **2012**, *52*, 1609–1621. [[CrossRef](#)]
47. Rutkoski, J.E.; Poland, J.A.; Singh, R.P.; Huerta-Espino, J.; Bhavani, S.; Barbier, H.; Rouse, M.N.; Jannink, J.; Sorrells, M.E. Genomic selection for quantitative adult plant stem rust resistance in wheat. *Plant Genome* **2014**, *7*, 1–10. [[CrossRef](#)]
48. Rutkoski, J.; Benson, J.; Jia, Y.; Brown-Guedira, G.; Jannink, J.; Sorrells, M. Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Genome* **2012**, *5*, 51–61. [[CrossRef](#)]
49. Poland, J.; Rutkoski, J. Advances and challenges in genomic selection for disease resistance. *Annu. Rev. Phytopathol.* **2016**, *54*, 79–98. [[CrossRef](#)]

50. Beyene, Y.; Gowda, M.; Olsen, M.; Robbins, K.R.; Pérez-Rodríguez, P.; Alvarado, G.; Dreher, K.; Gao, S.Y.; Mugo, S.; Prasanna, B.M.; et al. Empirical Comparison of Tropical Maize Hybrids Selected Through Genomic and Phenotypic Selections. *Front. Plant Sci.* **2019**, *10*, 1–11. [[CrossRef](#)]
51. de Roos, A.P.W.; Hayes, B.J.; Goddard, M.E. Reliability of genomic predictions across multiple populations. *Genetics* **2009**, *183*, 1545–1553. [[CrossRef](#)]
52. Semagn, K.; Beyene, Y.; Warburton, M.; Tarekegne, A.; Mugo, S.; Meisel, B.; Sehabiague, P.; Prasanna, B.M. Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water-stressed and well-watered environments. *BMC Genomics*. **2013**, *14*, 313. [[CrossRef](#)]
53. Zhang, X.; Pérez-Rodríguez, P.; Semagn, K.; Beyene, Y.; Babu, R.; López-Cruz, M.A.; Vicente, F.S.; Olsen, M.; Buckler, E.; Jannink, J.-L.; et al. Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity (Edinb)* **2015**, *114*, 291–299. [[CrossRef](#)] [[PubMed](#)]
54. Gilmour, A.R.; Gogel, B.J.; Cullis, B.R.; Thompson, R. *ASReml User Guide Release 3.0*; VSN International Ltd.: Hemel Hempstead, UK, 2009.
55. Alvarado, G.; López, M.; Vargas, M.; Pacheco, Á.; Rodríguez, F.; Burgueño, J.; Crossa, J. META-R (Multi Environment Trait Analysis with R for Windows)-CIMMYT Research Data & Software Repository Network. (Version 5.0). Available online: <https://hdl.handle.net/11529/10201> (accessed on 14 August 2020).
56. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **2011**, *6*, 1–10. [[CrossRef](#)] [[PubMed](#)]
57. Bradbury, P.; Zhang, Z.; Kroon, D.E.; Casstevens, T.; Ramdoss, Y.; Buckler, E. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* **2007**, *23*, 2633–2635. [[CrossRef](#)] [[PubMed](#)]
58. Hill, W.G.; Weir, B.S. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul. Biol.* **1988**, *33*, 54–78. [[CrossRef](#)]
59. Meng, L.; Li, H.; Zhang, L.; Wang, J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop. J.* **2015**, *3*, 269–283. [[CrossRef](#)]
60. Würschum, T.; Liu, W.; Gowda, M.; Maurer, H.P.; Fischer, S.; Schechert, A.; Reif, J.C. Comparison of biometrical models for joint linkage association mapping. *Heredity (Edinb)*. **2012**, *108*, 332–340. [[CrossRef](#)]
61. Liu, W.; Gowda, M.; Steinhoff, J.; Maurer, H.P.; Würschum, T.; Longin, C.F.H.; Cossic, F.; Reif, J.C. Association mapping in an elite maize breeding population. *Theor. Appl. Genet.* **2011**, *123*, 847–858. [[CrossRef](#)]
62. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [[CrossRef](#)]
63. SAS. *SAS, Version 9.4*; Intelligence Platform: System Administration Guide; SAS Institute Inc.: Cary, NC, USA, 2015.
64. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
65. Utz, H.F.; Melchinger, A.E.; Schön, C.C. Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. *Genetics* **2000**, *154*, 1839–1849.
66. Zhao, Y.; Gowda, M.; Liu, W.; Würschum, T.; Maurer, H.P.; Longin, F.H.; Ranc, N.; Reif, J.C. Accuracy of genomic selection in European maize elite breeding populations. *Theor. Appl. Genet.* **2012**, *124*, 769–776. [[CrossRef](#)] [[PubMed](#)]

