JARQ 58 (3), 151-164 (2024) https://www.jircas.go.jp

Genomic Prediction of Maize Inbred Lines Using a Small Combined Training Population and Evaluation of Untested Germplasms for Resistance to Pythium Root and Stalk Rot

Shohei MITSUHASHI[1]* and Hiroyuki TAMAKI[1, 2]

[1] Institute of Livestock and Grassland Science, National Agriculture and Food Research Organization, Nasushiobara, Japan
[2] Hokkaido Branch, Japan Grassland Agriculture & Forage Seed Association, Ebetsu, Japan

Abstract

Genotypic studies using molecular markers, such as genomic prediction (GP), have been implemented in maize (*Zea mays* L.) breeding, leading to a better understanding of significant traits. Pythium root and stalk rot (RSR) resistance is an important trait in maize bred in Japan. The study aims to (1) develop a GP model for Pythium RSR resistance in maize bred within the Japanese public sector and (2) utilize GP to analyze untested maize germplasms for Pythium RSR resistance via a combined training set derived from different heterotic groups (dent and flint). Through 1,000 repetitions of sampling and five-fold cross-validation, a high average prediction accuracy (r = 0.695, 95% confidence interval: 0.682-0.708) was achieved across populations. Prediction accuracy improved as the number of markers increased, but it eventually reached a plateau that exceeded 1,000 markers. The population component and linkage disequilibrium between markers confirmed previous reports. These findings show the feasibility of GP, even with a small population (N = 41) and marker size (approximately 1,000). Several old inbred lines were identified with lower predicted RSR scores, indicating their potential as breeding materials. This is the first report on the prediction of maize Pythium RSR resistance using GP and emphasizes new possibilities for addressing Pythium RSR resistance in maize breeding.

Discipline: Crop Science
Additional key words: best linear unbiased prediction, *Zea mays* L.

Introduction

Globally, maize (*Zea mays* L.) is an important crop resource for animal feed, food, and biofuel production. It is extremely diverse both phenotypically and genotypically, and several tools have been developed to utilize its diversity in hybrid breeding (Riedelsheimer et al. 2012). Furthermore, the genotypic studies on this crop have made remarkable progress in recent years; genome-wide association studies (GWAS) and marker-assisted selection have contributed to the enhancement of its important traits, which include disease resistance (Kump et al. 2011), flowering time (Chardon et al. 2004), and leaf architecture (Tian et al. 2011). Considering that the cost of molecular genotyping has decreased, genome-wide marker polymorphisms are useful not only for detecting causal genes for agronomic traits via association mapping (e.g., GWAS) but also for predicting the agronomic performance of untested genotypes in major crop species.

Root and stalk rot (RSR) caused by soil-borne disease pathogens of the genus Pythium is an important trait for maize breeding in Japan. The symptoms of this disease include wilting or lodging of whole plants and drooping of ears; consequently, the plants become too soft to be cut during harvest, which causes challenges in processing the crop for forage. Furthermore, the quality and nutritional value of plants are decreased. In our previous study, certain parental inbred lines, which are susceptible when used in F_1 hybrids, are observed to exhibit Pythium RSR resistance in field inoculation tests (Mitsuhashi et al. 2015). This makes it difficult to effectively develop Pythium RSR-resistant hybrids. As a result, genotypic studies of Pythium RSR resistance are crucial; however, in China, only a few studies on quantitative trait locus analyses exist (Duan et al. 2019,

*Corresponding author: shouheim@affrc.go.jp
Received 7 August 2023; accepted 9 December 2023.

Song et al. 2015).

"Genomic prediction (GP)" or "genomic selection (GS)" are breeding techniques in which a part of phenotyping can be substituted with molecular genotyping using predicted breeding values of untested genotypes. The proposal by Meuwissen et al. (2001) has established a remarkable achievement in dairy cattle breeding using genomic breeding values predicted only via genotypic marker information (Hayes et al. 2009b). It has also attracted the attention of crop breeding. Bernardo and Yu (2007) have attempted to employ GP and GS in maize breeding. Our previous computer simulation study has indicated that GP and GS can be powerful tools for maize breeding teams in the Japanese public sector (Tamaki et al. 2012). Nonetheless, the use of GP and GS in practical maize breeding in Japan has not yet been achieved.

The preparation of large-scale training sets is the major difficulty in GP. To address this, some studies have evaluated the effectiveness of GP for small populations of combined, multibreed training sets in dairy cattle breeding (Erbe et al. 2012, Hayes et al. 2009a). In the case of the resistance breeding of maize, Technow et al. (2013) have indicated that combined training sets (the dent and flint inbred lines) provide adequate prediction accuracy by cross-validation even if the training population sizes are very small ($N \geq 25$). The objectives of this study are as follows: (1) to develop GP models for maize Pythium RSR resistance using data obtained in the breeding process within the Japanese public sector and (2) to employ GP to predict the Pythium RSR resistance of untested maize germplasms, using a combined training data set derived from different heterotic groups.

Materials and methods

1. Plant materials

In total, 41 maize inbred lines comprising 18 dent and 23 flint lines were tested as the training dataset for developing the GP model, and 188 inbred lines were considered untested germplasms for predicting the Pythium RSR resistance using this model. All inbred lines were developed at the Nasushiobara, Hokkaido, and Miyakonojo research stations of NARO (The National Agriculture and Food Research Organization, Japan), Prefectural public breeding sections in Nagano, Japan, or Governmental Tokachi Agricultural Experiment Station in Hokkaido, Japan. The Results and Discussion sections describe the details of the inbred lines.

2. Field test and phenotypic data

The breeding values of the 41 inbred lines used in this study were calculated based on a parental-progeny-based best linear unbiased prediction (BLUP) approach. These values were derived from the phenotypic scores of the resistance of the F_1 combinations against Pythium RSR in field tests carried out for 4 years from 2016 to 2019. These breeding values were also utilized as the phenotypic data for the training dataset described below.

Field experiments were conducted at the Institute of Livestock and Grassland Science, NARO (NILGS), Nasushiobara, Tochigi, Japan (36°55′04″N, 139°56′29″E, 320 m above mean sea level), following the customary practices had been utilized in our breeding processes. Maize sudangrass (*Sorghum sudanense* [Piper] Stapf.) annual rotation was implemented in the experimental and adjacent fields throughout the experiments. Before sowing, in early spring of each year, the fields were treated with 50 metric tons per hectare of manure, 600 kilograms per hectare of fertilizer (containing 14% each of N, P_2O_5, and K_2O), and 60 kilograms per hectare of insecticide (diazinon granule). Herbicides (5.0 liters per hectare of alachlor emulsion, 2.0 liters per hectare of atrazine wettable powder, and 1.5 liters per hectare of topramezone) were applied after sowing. All field tests were conducted in a randomized complete block design with two or four replicates. Each entry was grown in single-row plots, measuring 2.4 m or 3.6 m in length, with a spacing of 0.75 m between rows. Each plot consisted of 13 or 19 plants. As our fields were presumed to have a high level of RSR contamination, F_1 hybrids were evaluated under natural infection conditions. Nevertheless, considering disease susceptibility, inbred lines underwent an inoculation test.

At the yellow-ripe stage of each hybrid (approximately 40 days after silking), the plants were cut approximately 5 cm above the ground, and the extent of rotting on the cut surface of the stalks was recorded. Pythium RSR was evaluated based on the infection frequency, represented as the percentage of plants with scores of 2 or higher. The field experiments, RSR inoculation, evaluation, and scoring procedures were consistent with those described in previous studies by the authors (Mitsuhashi et al. 2015, Mitsuhashi & Tamaki 2022). Refer to Table 1 for further details in terms of the scale and overview of the field tests.

3. BLUP to estimate the breeding values of each inbred

To predict the breeding values (i.e., general combining ability [GCA]) of the 41 inbred lines, the following BLUP mixed model matrix equation

Table 1. Scale and outline of the field tests

Experiment number	Number of				Broad-sense heritability (h^2)
	Hybrids	Inbred lines crossed for hybrids	Plants in a single row	Replications	
2016-1	5	7	19	4	0.600
2016-2	23	22	19	2	0.517
2016-3	77	41	13	2	0.562
2016-4	61	44	13	2	0.050 †
2017-1	11	13	19	2	0.824
2017-2	20	25	19	2	0.840
2017-3	136	42	13	2	0.082 †
2017-4	58	35	13	2	0.204
2018-1	7	12	19	4	0.749
2018-2	100	32	19	2	0.278
2018-3	44	29	13	2	0.746
2018-4	47	32	13	2	0.652
2019-1	6	7	19	4	0.668
2019-2	17	24	19	2	0.796
2019-3	100	51	19	2	0.697
2019-4	10	13	13	2	0.845
2019-5	29	28	13	2	0.935

Sowing and observation dates were different for each year or environment.
Some common hybrids and inbred lines were used among the test plots.
A total of 650 F_1 hybrids were tested, derived from crosses among 106 inbred lines.
† These were excluded from the calculation of the mean as outlier values (Smirnov-Grubbs test $P < 0.05$).
The average value of h^2 was 0.661.

described in our previous study was adopted (Mitsuhashi & Tamaki 2022):

$$y = X\beta + Za + \varepsilon \quad (1)$$

where y is the phenotypic values of F_1 hybrids via the field experiments in Section 2, all those with and without the 41 inbred lines as parents, X indicates a design matrix to express which F_1 combination is tested in which field experiments, β represents an unknown vector for environmental values of each experiment, a represents a vector for the GCA of each inbred used in the study, Z indicates a design matrix to express which inbred lines are the parents of each F_1 combination, and ε indicates the residual effect. The solution of the matrix Equation 1 follows the previous study by our team:

$$\begin{pmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{pmatrix} = \begin{pmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + V_G^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{a} \end{pmatrix} \quad (2)$$

$$\begin{pmatrix} \hat{\beta} \\ \hat{a} \end{pmatrix} = \begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z + \left(\frac{1-h^2}{h^2}\right) \end{pmatrix}^{-1} \begin{pmatrix} X'y \\ Z'y \end{pmatrix} \quad (3)$$

where X' is the transposed matrix of X, R is a variance–covariance matrix for residual effects, and V_G is the variance–covariance matrix for the effect of traits. Equation 2 can be transformed into Equation 3, given that σ_a^2 is the additive genetic variances and σ_e^2 are assumed to be $h^2\sigma_y^2$ and $(1-h^2)\,\sigma_y^2$ using the heritability (h^2). Broad-sense heritability was determined as the average of the values obtained via the analysis of variance carried out for each test plot (refer to Table 1). The solution program was implemented using R version 4.0, developed by the R Core Team (2020). The GCA with the lowest value was set as the reference, and the remaining values were adjusted accordingly. The calculated results were then compared with the RSR infection frequency observed in the field inoculation test conducted in 2021, as described earlier.

4. Genomic data and their analysis

All inbred lines were genotyped with the "Maize LD Bead chip" (Illumina Inc., San Diego, USA) containing 3,047 single-nucleotide polymorphisms (SNPs). Markers with more than 5% missing data were removed. "BEAGLE" (Browning et al. 2018), version 5.4,

was employed to impute all remaining missing marker genotypes, which resulted in 2,581 SNPs available for further analysis. To see its pattern, the linkage disequilibrium (LD) was plotted as measured by r^2 with a marker distance in Mbp for all 2,581 markers. To investigate the genetic distinction of different heterotic groups, a principal component analysis using the 2,581 SNP marker profiles of the inbred lines was adopted. These two indicators were calculated using "Tassel" version 5.2.87 (Bradbury et al. 2007).

5. Prediction procedures

All validation and prediction were carried out using parental inbred lines only; phenotypic data and genotypic data were substituted in order to develop the following linear model:

$$y_i = \mu + \sum \beta_j X_{ij} + \varepsilon_i \quad (4)$$

where y_i is the BLUP score of inbred i calculated in Section 3, μ is the overall mean, β_j is the genetic effect of the marker j (j = 1, 2…), X_{ij} is the genotype of marker j for inbred i and is defined by 1 or −1 for contrasting homozygous genotypes and 0 for heterozygous, and ε_i is the error deviation assumed to follow N (0, σ^2). To estimate the genetic effect coefficients β_j, the Ridge regression method was applied. Following this regression method, a prediction model was obtained using the R package "glmnet" (Friedman et al. 2010).

6. Cross-validation and prediction

To assess the accuracy of GP, five-fold cross-validation was applied on the basis of the method described by Zhao et al. (2012). To determine the optimal number of markers for prediction, different marker patterns, which included 250, 500, 1,000, 2,000, and all 2,581 markers, were evaluated. A fixed number of markers were randomly distributed throughout the entire genome, and the dataset derived from the 41 inbred lines was randomly divided into five subsets. Four subsets were combined to form the training dataset for estimating genetic effects, whereas the remaining subset served as the validation dataset. To determine the prediction accuracy (r), the correlation between the BLUP scores of the validation dataset and the calculated scores from the predicted genetic effects of the training model was employed. The process of randomly locating markers was repeated 100 times, the sampling of training and validation sets was repeated 1,000 times for each marker set, and the mean of both prediction accuracy and 95% confidence interval (CI) for each marker set was calculated. Moreover, to predict the Pythium RSR resistance of an additional 188 untested inbred lines, the training model, employing all 41 inbred lines and 2,581 markers, was applied.

Results and discussion

Table 2 provides details of the 41 inbred lines. The breeding values predicted by BLUP were inconsistent with the results of the field inoculation test conducted in 2021, with a correlation coefficient of 0.027 (N = 17, not significant). This inconsistency was particularly evident in the case of 'Na50,' a representative parental inbred line of susceptible F_1 hybrids that exhibited resistance itself, consistent with our previous studies (Mitsuhashi & Tamaki 2022). Therefore, it can be confirmed that the results of field inoculation tests carried out on inbred lines cannot be considered phenotypic values for GP of Pythium RSR. Instead, the BLUP values are more appropriate.

Labor and field capacity are limiting factors in the maize breeding system. Nonetheless, larger population sizes are well-known to lead to higher genetic gains in GP (Lorenz et al. 2011). Thus, to design breeding schemes that enhance efficiency while reducing labor requirements, significant efforts have been made. Tecnow et al. (2013) suggested that small training datasets, combining dent and flint inbred lines, can achieve adequate prediction accuracy for Northern corn leaf blight resistance. The population composition within our training dataset, which depicts a distinct separation between dent and flint inbred lines, is illustrated in Figure 1. This confirms the previous study and is considered appropriate for achieving high prediction accuracy even with a small population size.

LD (r^2) between markers for the 41 maize inbred lines within a distance of less than 0.24 Mbp exceeded 0.30. At approximately 0.60 Mbp, it gradually decreased to approximately 0.20. Beyond 1.0 Mbp, LD continued to decrease slightly; however, at approximately 2.0 Mbp, it remained above 0.10 (Fig. 2). In Technow et al. (2013), the LD continued to decrease at above 1.0 Mbp but remained greater than 0.10 over the entire range of distances (about 5.0 Mbp) considered. Calus and Veerkamp (2007) and Calus et al. (2008) indicated that for effective selection in GP, the average LD between adjacent markers should be ≥ 0.125. Our findings confirm these observations and demonstrate the feasibility of GP using a small training population that is derived from dent and flint inbred lines.

During the repetition of randomly locating each marker set 100 times and performing sampling and validation 1,000 times for each marker set, outliers of prediction accuracy (r) were excluded using the

Table 2. Elite inbred lines and their resistance for Pythium root and stalk rot (RSR)

Inbred name	Group	BLUP scores	Observed values in the field	Predicted values in the model	Developed by
CHU44	F	9.06	10.00	8.84	CAES, Nagano Pref. †
JC-028	D	7.36	4.55	7.51	CAES, Nagano Pref.
CHU68	F	9.60	4.55	9.72	CAES, Nagano Pref.
JC-037	D	4.46	0.00	5.47	CAES, Nagano Pref.
JC-038	D	10.66	-	9.32	CAES, Nagano Pref.
Mi47	F	9.01	0.00	8.47	KARC, NARO ‡
Mi91	D	10.18	-	9.98	KARC, NARO
Mi103	F	9.47	0.00	8.95	KARC, NARO
Mi111	F	3.52	0.00	4.24	KARC, NARO
Mi115	F	7.61	-	8.73	KARC, NARO
N09-07	F	2.55	-	3.25	ILGS, NARO §
N10-01	D	7.17	-	7.22	ILGS, NARO
N10-02	D	4.43	9.55	4.58	ILGS, NARO
N10-08	F	5.95	-	5.99	ILGS, NARO
N10-12	F	0.00	-	1.86	ILGS, NARO
N11-02	F	6.41	-	6.79	ILGS, NARO
N12-01	D	5.25	-	5.36	ILGS, NARO
N12-02	F	10.97	-	11.08	ILGS, NARO
N12-05	F	7.99	34.85	8.26	ILGS, NARO
N12-07	F	6.10	-	6.69	ILGS, NARO
N13-01	D	5.60	-	5.85	ILGS, NARO
N13-05	F	19.72	-	18.76	ILGS, NARO
N13-06	F	20.62	-	19.66	ILGS, NARO
N13-08	F	11.26	-	10.58	ILGS, NARO
N14-01	D	7.29	-	7.21	ILGS, NARO
N14-02	D	3.75	-	4.22	ILGS, NARO
N15-01	D	3.96	-	4.40	ILGS, NARO
N16-03	D	6.85	-	6.50	ILGS, NARO
N16-07	F	10.83	-	11.02	ILGS, NARO
Na50	F	23.22	0.00	21.25	ILGS, NARO
Na65	D	8.64	27.78	8.19	ILGS, NARO
Na71	D	2.91	0.00	3.22	ILGS, NARO
Na83	F	17.28	-	16.24	ILGS, NARO
Na98	D	2.72	0.00	3.56	ILGS, NARO
Na100	D	13.13	-	11.89	ILGS, NARO
Na102	D	0.39	5.00	1.22	ILGS, NARO
Na106	F	17.18	6.25	16.28	ILGS, NARO
Na109	D	7.11	-	7.16	ILGS, NARO
Na111	F	6.38	0.00	6.90	ILGS, NARO
Na112	F	8.62	0.00	9.07	ILGS, NARO
Na113	F	18.64	-	18.34	ILGS, NARO

BLUP scores were the general combining ability (GCA) of each inbred for Pythium RSR (Mitsuhashi & Tamaki 2022). Observed values in the field were RSR infection frequency of field inoculation test in 2021. The correlation coefficient between BLUP scores and observed values was 0.027 ($N = 17$, not significant). Predicted values were in the training model using all 2,581 SNPs.

D: dent, F: flint.
† Chushin Agricultural Experiment Station, Nagano Pref.
‡ Kyushu Okinawa Agricultural Research Center, NARO
§ Institute of Livestock and Grassland Science, NARO

CHU is the registered inbred lines. JC is promising inbred lines.
Mi and Na are the registered or promising inbred lines. N is the superior inbred line before Na was named.

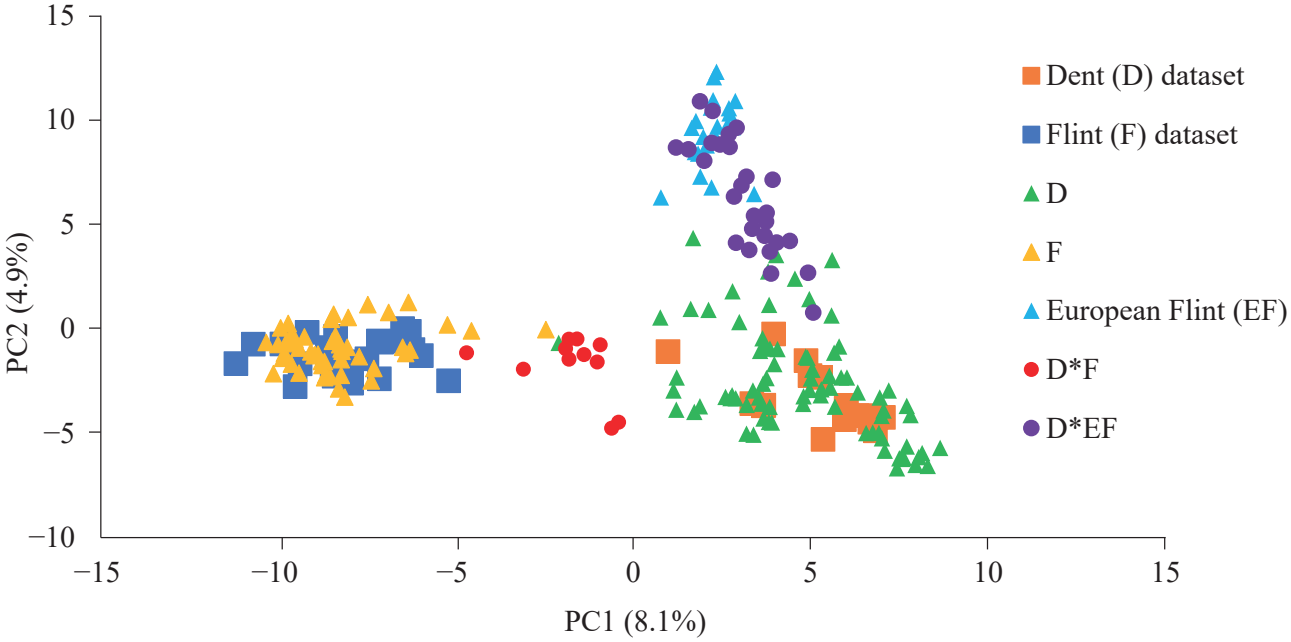

Fig. 1. The plot of principal component (PC) 1 and PC 2 scores based on 2,581 SNP markers of all the 229 inbred lines used in this study
* represents the crosses which is derived from different two groups.

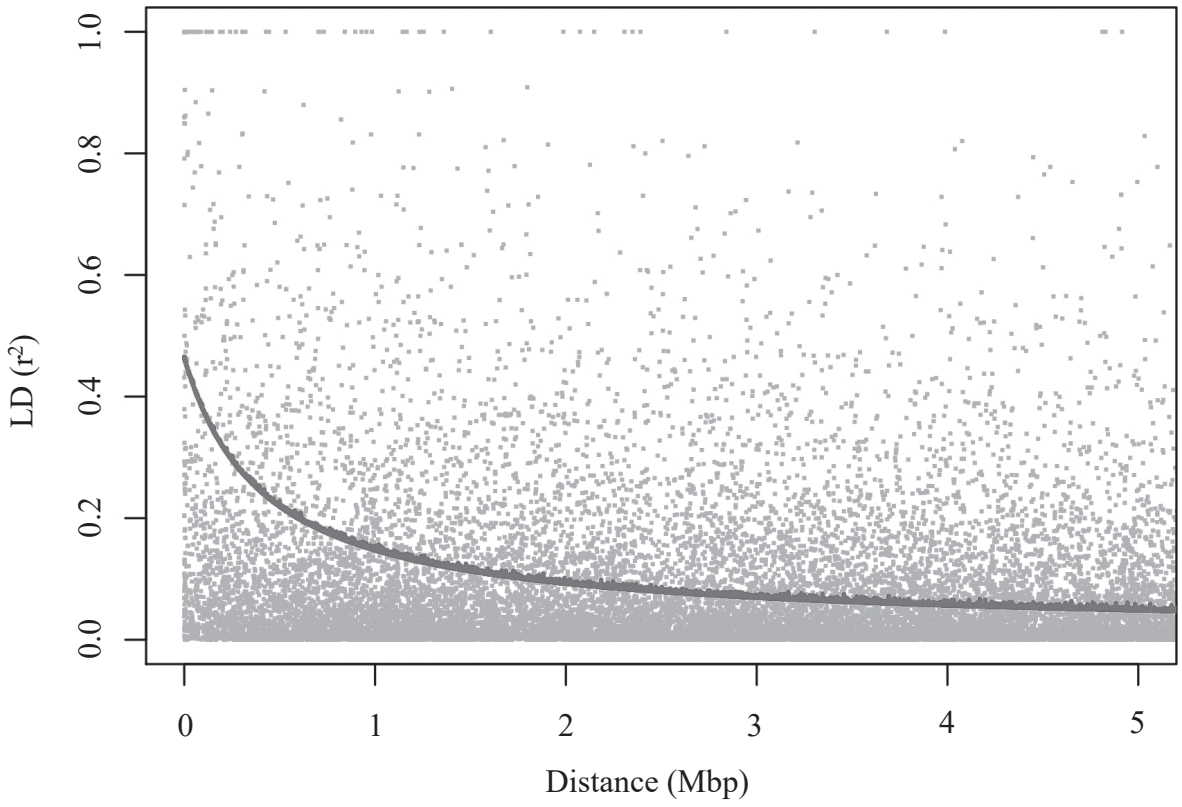

Fig. 2. Linkage disequilibrium (LD: r^2) decay plot of 2,581 markers as a function of physical distance (Mbp) for the 41 maize inbred lines used in this study

Smirnov–Grubbs test ($P < 0.05$) for each of the marker sets. As the number of markers increased, the average prediction accuracy between populations improved. Nevertheless, when the number of markers exceeded 1,000 (r = 0.680, 95% CI: 0.666-0.693), the rate of improvement declined, reaching a plateau at 2,000 markers (r = 0.693, 95% CI: 0.680-0.706). The prediction accuracy using all 2,581 markers was 0.695 (95% CI: 0.682-0.708; Table 3, Fig. 3). Tecnow et al. (2013) reported that the prediction accuracies are 0.576-0.589 (N = 50) and 0.690-0.706 (N = 75) in Northern corn leaf blight. Crossa et al. (2014) showed that the prediction accuracies are 0.588-0.790 in flowering, 0.513-0.572 in anthesis-silking interval, and 0.415-0.525 in grain yield (N = 284). Given the use of a smaller training population (N = 41), it can be confirmed that the GP model present in this study achieves sufficient prediction accuracy compared with previous studies.

Rashid et al. (2020) employed approximately 300,000 SNPs for conducting GWAS on resistance to maize Northern corn leaf blight, whereas Liu et al. (2021) utilized over 200,000 SNPs for resistance to Fusarium ear rot. Nevertheless, GP should be performed using lower-density markers if achieving cost-effectiveness is the aim (Heffner et al. 2010). In our study, the prediction accuracy enhanced as the number of markers increased, but it approached a plateau when the number of markers exceeded 1,000. This observation confirms a previous study by Zhao et al. (2012), in which prediction accuracy plateaus at approximately 800 SNPs. These results indicate that by using a small population size and a moderate number of markers, GP for Pythium RSR resistance can be achieved.

The predicted GCA values for the highest and lowest five inbred lines (dent, flint or European flint, and their crosses) among the 188 untested inbred lines, in terms of resistance to Pythium RSR, based on the prediction model utilized in this study is presented in Table 4. Supplementary Table 1 shows the complete values. The lowest predicted RSR values for both dent and flint inbred lines were similar (4.70 and 4.62), whereas the

Table 3. Prediction accuracy (r) of genomic predictions across populations according to the number of each marker revealed by five-fold cross-validation for Pythium root and stalk rot (RSR)

The number of markers	Prediction accuracy (r)	95% confidence interval
250	0.631	0.616-0.645
500	0.658	0.644-0.672
1,000	0.680	0.666-0.693
2,000	0.693	0.680-0.706
2,581	0.695	0.682-0.708

The process of randomly locating markers was repeated 100 times, and the sampling of training and validation sets was repeated 1,000 times for each marker set.

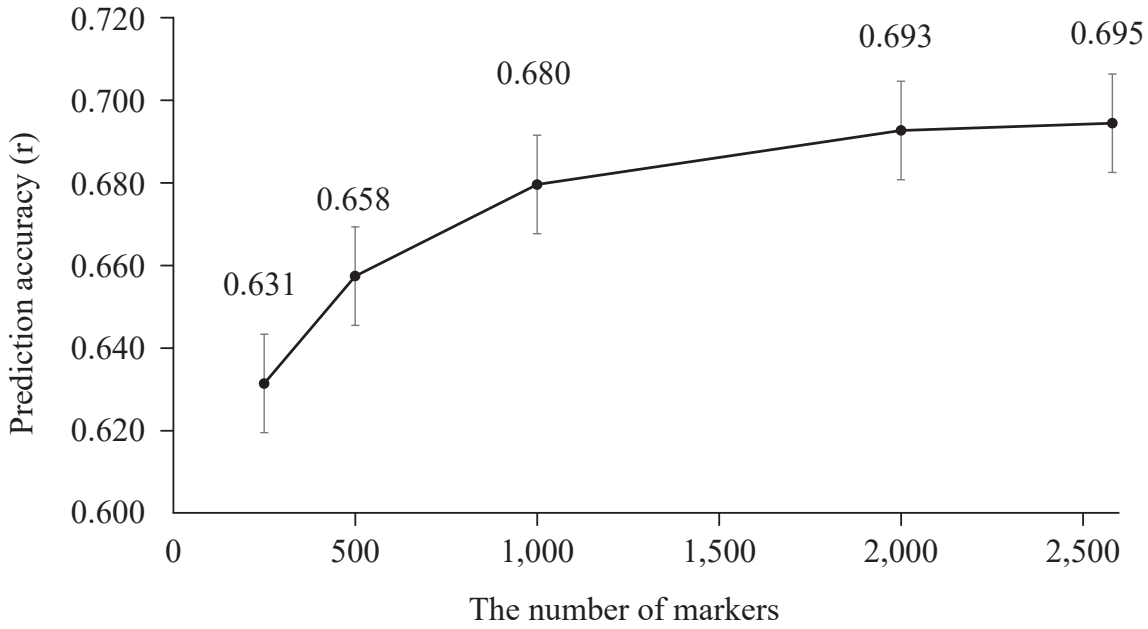

Fig. 3. Transition of the prediction accuracy (r) of genomic predictions across populations revealed by five-fold cross-validation for Pythium root and stalk rot (RSR)
The process of randomly locating markers was repeated 100 times, and the sampling of training and validation sets was repeated 1,000 times for each marker set. Error bars represent 95% confidence intervals.

Table 4. Predicted general combining ability (GCA) for each of the highest and lowest five (dent, flint, European flint, and their crosses) of 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study

Inbred name	Group	Predicted GCA values	Developed by
Na60	D	4.70	ILGS, NARO §
Na78	D	4.78	ILGS, NARO
Ho57	D	5.08	HARC, NARO ¶
Na54	D	5.44	ILGS, NARO
Na29	D	5.46	ILGS, NARO
Na84	F	4.62	ILGS, NARO
IM-459	F	6.79	KARC, NARO ‡
Ho99	EF	7.30	HARC, NARO
J1608	F	7.61	CAES, Nagano Pref. †
TI-083	EF	7.96	HARC, NARO
IM-254	D*F	7.27	KARC, NARO
TI-133	D*EF	7.64	HARC, NARO
TI-114	D*EF	7.69	HARC, NARO
To113	D*EF	7.70	TAES, Hokkaido Govt. ‖
TI-145	D*EF	7.79	HARC, NARO
JC-036	D	9.71	CAES, Nagano Pref.
J1407	D	9.75	CAES, Nagano Pref.
JC-046	D	9.83	CAES, Nagano Pref.
J1383	D	9.99	CAES, Nagano Pref.
J1539	D	10.24	CAES, Nagano Pref.
Na85	F	12.73	ILGS, NARO
IM-430	F	13.00	KARC, NARO
Na28	F	13.32	ILGS, NARO
Na95	F	13.46	ILGS, NARO
Na89	F	13.82	ILGS, NARO
To90	D*EF	9.30	TAES, Hokkaido Govt.
TI-108	D*EF	9.46	HARC, NARO
To38	D*EF	9.55	TAES, Hokkaido Govt.
J1707	D*F	9.80	CAES, Nagano Pref.
Na94	D*F	11.17	ILGS, NARO

The prediction model was derived from the training data sets of the 41 elite inbred lines and 2,581 SNPs.

D: dent, F: flint, EF: European flint. * is derived from crosses between different two groups.

† Chushin Agricultural Experiment Station, Nagano Pref.
‡ Kyushu Okinawa Agricultural Research Center, NARO
§ Institute of Livestock and Grassland Science, NARO
¶ Hokkaido Agricultural Research Center, NARO
‖ Tokachi Agricultural Experiment Station, Hokkaido Govt.

Na, Ho, J, JC, and To are the registered or promising inbred lines.
IM and TI are the superior inbred lines before Mi and Ho were named.

highest values varied (10.24 and 13.82). Inbred lines derived from dent and flint crosses exhibited intermediate values. This finding confirms the results of the PCA (Fig. 1). Some of the flint inbred lines are predominantly derived from regions with colder climates, such as Hokkaido or Northern Europe. Breeding for RSR resistance, which is more prevalent in hot and humid conditions, presents difficulties in such regions. These

results likely reflect the different selection pressures on dent and flint lines. None of the five dent, flint, or crossed inbred lines with the lowest RSR scores have been utilized in recent breeding programs because of their age (developed from the 1990s to early 2000s). These results suggest that such old germplasms can be potentially used as valuable materials for developing Pythium RSR-resistant hybrids.

The geographic diversity of Japan, which spans from Hokkaido in the north to Kyushu in the south, shows different challenges for breeding and selecting maize inbred lines appropriate for each region. Pythium RSR outbreak does not frequently take place in Hokkaido because *Pythium* spp. thrives at higher temperatures (Kageyama 2014). Likewise, Kyushu experiences different cropping systems from Kanto (Nasushiobara). Nevertheless, under specific conditions such as adequate temperature and heavy rainfall during the dough ripening stage, large Pythium RSR outbreaks can occur even in these regions (Deep & Lipps 1996, Reyes-Tena et al. 2018, Yenar et al. 1997). Accurately predicting Pythium RSR resistance using GP can assist in making preliminary selections solely on the basis of genotypic data from the constructed model.

To conclude, this study is the first to predict maize Pythium RSR resistance using GP in the Japanese public sector. The GP model presented in this study achieved adequate prediction accuracy, even though a smaller training population (N = 41) and lower-density markers (approximately 1,000 SNPs) were employed. The results have significant implications, particularly in regions with limited labor and field resources. Nevertheless, the reliability of the predicted data through field tests that involve actual F_1 hybrids derived from the 188 untested inbred lines used in this study should be validated. These findings offer new possibilities for breeding maize with Pythium RSR resistance in Japan.

Acknowledgements

We wish to express our gratitude to the Advanced Analysis Center of NARO, Tsukuba, Japan, for carrying out the genotyping analysis in this study and to Dr. Koya Sugawara for generously adjusting the source materials of toothpick inoculation test in our field. Ms. Yoko Nakada and Etsuko Imayoshi of the Institute of Livestock and Grassland Science, NARO, are also thanked for their assistance during the study, as well as the staff of Nasu Field Technical Team, NARO, for their field management.

References

Bernardo, R. & Yu, J. (2007) Prospects for genome wide selection for quantitative traits in maize. *Crop Sci*., **47**, 1082-1090.

Bradbury, P. J. et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633-2635.

Browning, B. L. et al. (2018) A one-penny imputed genome from next generation reference panels. *Am. J. Hum. Genet*., **103**, 338-348.

Calus, M. P. & Veerkamp, R. F. (2007) Accuracy of breeding values when using & ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP per cM. *J. Anim. Breed. Genet*., **124**, 362-368.

Calus, M. P. et al. (2008) Accuracy of genomic selection using different methods to define haplotypes. *Genetics*, **178**, 553-561.

Chardon, F. et al. (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis & synteny conservation with the rice genome. *Genetics*, **168**, 2169-2185.

Crossa, J. et al. (2014) Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity* (*Edinb*), **112**, 48-60.

Deep, I. W. & Lipps, P. E. (1996) Recovery of *Pythium arrhenomanes* and its virulence to corn. *Crop Prot*., **15**, 85-90.

Duan, C. et al. (2019) Characterization and molecular mapping of two novel genes resistant to Pythium stalk rot in maize. *Phytopathology*, **109**, 804-809.

Erbe, M. et al. (2012) Improving accuracy of genomic predictions within & between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. *J. Dairy Sci*., **95**, 4114-4129.

Friedman, J. et al. (2010) Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw*., **33**, 1-22.

Hayes, B. J. et al. (2009a) Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol*., **41**, 51.

Hayes, B. J. et al. (2009b) Invited review: Genomic selection in dairy cattle: Progress & challenges. *J. Dairy Sci*., **92**, 433-443.

Heffner, E. L. et al. (2010) Plant breeding with genomic selection: gain per unit time and cost. *Crop Sci*., **50**, 1681-1690.

Kageyama, K. (2014) Molecular taxonomy & its application to ecological studies of *Pythium* species. *J. Gen. Plant Pathol*., **80**, 314-326.

Kump, K. L. et al. (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize association mapping population. *Nat. Genet*., **43**, 163-168.

Liu, Y. et al. (2021) Genome-wide association study and genomic prediction of *Fusarium* ear rot resistance in tropical maize germplasm. *Crop J*., **9**, 325-341.

Lorenz, A. J. et al. (2011) Genomic selection in plant breeding: Knowledge & prospects. *Adv. Agron*., **110**, 77-123.

Meuwissen, T. H. et al. (2001) Prediction of total genetic value using genome wide dense marker maps. *Genetics*, **157**, 1819-1829.

Mitsuhashi, S. & Tamaki, H. (2022) Parental-progeny-based best linear unbiased prediction (BLUP) for determining maize single-cross performance and resistance to Pythium root and stalk rot. *Grassl. Sci*., **68**, 226-232.

Mitsuhashi, S. et al. (2015) Evaluation of resistance to Pythium root rot by *Pythium arrhenomanes* in corn by using a toothpick inoculation method. *Grassl. Sci.*, **61**, 181-184.

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rashid, Z. et al. (2020) Genome-wide association studies in tropical maize germplasm reveal novel and known genomic regions for resistance to Northern corn leaf blight. *Sci. Rep.*, **10**, 1-16.

Reyes-Tena, A. et al. (2018) *Pythium arrhenomanes* causal agent of root rot on yellow maize in Mexico. *Australas. Plant Dis. Notes*, **13**, 6.

Riedelsheimer, C. et al. (2012) Genomic & metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet*., **44**, 217-220.

Song, F. J. et al. (2015) Two genes conferring resistance to Pythium stalk rot in maize inbred line Qi319. *Mol. Genet. Genomics*, **290**, 1543-1549.

Tamaki, H. et al. (2012) Preliminary study on 'genome wide selection' for maize (*Zea mays* L.) breeding in Japanese public sectors: Estimated selection response in the first selfing cycle of single-cross progeny. *Grassl. Sci*., **58**, 20-27.

Technow, F. et al. (2013) Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3 Genes|Genomes|Genetics*, **3**, 197-203.

Tian, F. et al. (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.*, **43**, 159-162.

Yanar, Y. et al. (1997) Effect of soil saturation duration and soil water content on root rot of maize caused by *Pythium arrhenomanes*. *Plant Dis.*, **81**, 475-480.

Zhao, Y. et al. (2012) Accuracy of genomic selection in European maize elite breeding populations. *Theor. Appl. Genet.*, **124**, 769-776.

Supplementary Table 1. Predicted general combining ability (GCA) for 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study

Inbred name	Group	Predicted GCA values	Developed by
J1330	D	8.99	CAES, Nagano Pref. †
J1350	D	5.55	CAES, Nagano Pref.
J1383	D	9.99	CAES, Nagano Pref.
J1407	D	9.75	CAES, Nagano Pref.
J1417	D	6.95	CAES, Nagano Pref.
J1539	D	10.24	CAES, Nagano Pref.
J1605	D	7.90	CAES, Nagano Pref.
J1608	F	7.61	CAES, Nagano Pref.
J1559	D	8.13	CAES, Nagano Pref.
J1693	F	10.12	CAES, Nagano Pref.
J1707	D*F	9.80	CAES, Nagano Pref.
J1706	D	7.92	CAES, Nagano Pref.
JC-002	D	7.00	CAES, Nagano Pref.
J1698	D	7.78	CAES, Nagano Pref.
J1785	F	10.20	CAES, Nagano Pref.
JC-009	F	12.46	CAES, Nagano Pref.
JC-014	D	6.61	CAES, Nagano Pref.
JC-026	F	9.88	CAES, Nagano Pref.
JC-036	D	9.71	CAES, Nagano Pref.
JC-053	F	8.50	CAES, Nagano Pref.
JC-050	D	7.38	CAES, Nagano Pref.
JC-046	D	9.83	CAES, Nagano Pref.
JC-054	D	7.76	CAES, Nagano Pref.
JC-064	D	7.64	CAES, Nagano Pref.
JC-034	F	8.75	CAES, Nagano Pref.

(Continued on next page)

Supplementary Table 1. Predicted general combining ability (GCA) for 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study (Continued 1)

IM-239	D*F	8.45	KARC, NARO ‡
IM-248	F	12.41	KARC, NARO
IM-252	F	9.59	KARC, NARO
IM-254	D*F	7.27	KARC, NARO
IM-270	D*F	8.91	KARC, NARO
IM-308	D*F	8.07	KARC, NARO
IM-347	D*F	9.08	KARC, NARO
IM-390	D	5.96	KARC, NARO
IM-402	F	8.91	KARC, NARO
IM-403	F	9.05	KARC, NARO
IM-419	D*F	9.09	KARC, NARO
IM-421	D	7.47	KARC, NARO
IM-422	D	8.13	KARC, NARO
IM-423	D	7.17	KARC, NARO
IM-424	D	8.49	KARC, NARO
IM-426	D	8.73	KARC, NARO
IM-427	D	8.44	KARC, NARO
IM-429	D	7.02	KARC, NARO
IM-430	F	13.00	KARC, NARO
IM-431	F	9.86	KARC, NARO
IM-435	D	6.95	KARC, NARO
IM-436	D	7.10	KARC, NARO
IM-437	D	6.88	KARC, NARO
IM-450	D	7.17	KARC, NARO
IM-452	F	10.25	KARC, NARO
IM-453	F	10.20	KARC, NARO
IM-454	F	9.75	KARC, NARO
IM-455	D	9.13	KARC, NARO
IM-458	F	8.15	KARC, NARO
IM-459	F	6.79	KARC, NARO
IM-460	F	10.55	KARC, NARO
IM-461	F	11.10	KARC, NARO
IM-464	F	10.05	KARC, NARO
IM-465	D	7.92	KARC, NARO
IM-466	D	7.40	KARC, NARO
IM-467	D	7.40	KARC, NARO
IM-468	F	9.77	KARC, NARO
IM-469	F	9.19	KARC, NARO
IM-470	F	8.62	KARC, NARO
IM-472	F	12.22	KARC, NARO
IM-475	D	8.26	KARC, NARO
IM-477	D	7.75	KARC, NARO
Mi83	D	7.63	KARC, NARO
Mi102	F	9.57	KARC, NARO
Mi105	F	10.53	KARC, NARO

(Continued on next page)

Supplementary Table 1. Predicted general combining ability (GCA) for 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study (Continued 2)

Mi107	F	8.12	KARC, NARO
N14-04	F	10.00	ILGS, NARO §
N17-F04	F	9.73	ILGS, NARO
Na2	F	12.48	ILGS, NARO
Na4	F	11.94	ILGS, NARO
Na5	F	11.11	ILGS, NARO
Na6	D	7.86	ILGS, NARO
Na7	D	6.58	ILGS, NARO
Na8	D	7.98	ILGS, NARO
Na9	D	6.10	ILGS, NARO
Na13	D	8.46	ILGS, NARO
Na15	D	7.27	ILGS, NARO
Na17	D	7.94	ILGS, NARO
Na18	D	7.79	ILGS, NARO
Na23	D	7.01	ILGS, NARO
Na25	D	7.99	ILGS, NARO
Na26	F	11.66	ILGS, NARO
Na27	F	10.90	ILGS, NARO
Na28	F	13.32	ILGS, NARO
Na29	D	5.46	ILGS, NARO
Na30	F	11.16	ILGS, NARO
Na32	D	6.89	ILGS, NARO
Na34	D	6.88	ILGS, NARO
Na36	D	6.01	ILGS, NARO
Na38	D	7.59	ILGS, NARO
Na41	D	6.56	ILGS, NARO
Na42	D	5.84	ILGS, NARO
Na43	D	9.51	ILGS, NARO
Na45	D	9.30	ILGS, NARO
Na49	D	5.57	ILGS, NARO
Na51	F	11.30	ILGS, NARO
Na53	D	6.56	ILGS, NARO
Na54	D	5.44	ILGS, NARO
Na55	D	6.69	ILGS, NARO
Na56	D	5.73	ILGS, NARO
Na57	D*F	8.55	ILGS, NARO
Na58	D	7.59	ILGS, NARO
Na60	D	4.70	ILGS, NARO
Na61	D	6.43	ILGS, NARO
Na62	D	6.77	ILGS, NARO
Na64	D	6.73	ILGS, NARO
Na66	F	11.13	ILGS, NARO
Na69	D	8.14	ILGS, NARO
Na70	D	6.75	ILGS, NARO
Na72	D*F	8.31	ILGS, NARO

(Continued on next page)

Supplementary Table 1. Predicted general combining ability (GCA) for 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study (Continued 3)

Na74	D	7.75	ILGS, NARO
Na76	F	10.77	ILGS, NARO
Na77	D	7.24	ILGS, NARO
Na78	D	4.78	ILGS, NARO
Na79	F	9.02	ILGS, NARO
Na80	F	11.81	ILGS, NARO
Na81	D	6.16	ILGS, NARO
Na82	D*F	9.29	ILGS, NARO
Na84	F	4.62	ILGS, NARO
Na85	F	12.73	ILGS, NARO
Na86	D	6.82	ILGS, NARO
Na87	D	8.86	ILGS, NARO
Na88	F	9.40	ILGS, NARO
Na89	F	13.82	ILGS, NARO
Na92	F	9.23	ILGS, NARO
Na93	F	11.57	ILGS, NARO
Na94	D*F	11.17	ILGS, NARO
Na95	F	13.46	ILGS, NARO
Na97	F	9.53	ILGS, NARO
Ho49	EF	8.73	HARC, NARO ¶
Ho52	D	6.33	HARC, NARO
Ho57	D	5.08	HARC, NARO
Ho68	D	7.61	HARC, NARO
Ho87	EF	9.64	HARC, NARO
Ho90	EF	8.64	HARC, NARO
Ho96	EF	9.18	HARC, NARO
Ho99	EF	7.30	HARC, NARO
Ho104	D	7.80	HARC, NARO
Ho106	D	8.09	HARC, NARO
Ho119	EF	9.46	HARC, NARO
Ho120	EF	8.70	HARC, NARO
Ho121	EF	9.65	HARC, NARO
Ho124	EF	9.60	HARC, NARO
Ho126	EF	8.94	HARC, NARO
Ho127	EF	9.53	HARC, NARO
Ho129	EF	8.96	HARC, NARO
Ho130	EF	9.38	HARC, NARO
Ho131	EF	7.98	HARC, NARO
TI-044	D*EF	8.36	HARC, NARO
TI-045	EF	8.77	HARC, NARO
TI-061	D*EF	8.85	HARC, NARO
TI-064	EF	9.39	HARC, NARO
TI-081	D*EF	8.78	HARC, NARO
TI-083	EF	7.96	HARC, NARO
TI-086	D*EF	8.06	HARC, NARO

(Continued on next page)

Supplementary Table 1. Predicted general combining ability (GCA) for 188 inbred lines in the resistance of Pythium root and stalk rot (RSR) based on the prediction model used in this study (Continued 4)

TI-091	EF	8.47	HARC, NARO
TI-092	EF	8.88	HARC, NARO
TI-094	EF	9.20	HARC, NARO
TI-095	D*EF	8.76	HARC, NARO
TI-096	D*EF	8.62	HARC, NARO
TI-097	D*EF	8.19	HARC, NARO
TI-098	D*EF	7.92	HARC, NARO
TI-105	EF	10.04	HARC, NARO
TI-106	EF	9.15	HARC, NARO
TI-107	D*EF	7.84	HARC, NARO
TI-108	D*EF	9.46	HARC, NARO
TI-111	D*EF	8.87	HARC, NARO
TI-114	D*EF	7.69	HARC, NARO
TI-118	D*EF	9.11	HARC, NARO
TI-123	D	7.97	HARC, NARO
TI-126	D*EF	8.70	HARC, NARO
TI-130	D*EF	9.04	HARC, NARO
TI-131	D*EF	8.81	HARC, NARO
TI-132	D*EF	8.46	HARC, NARO
TI-133	D*EF	7.64	HARC, NARO
TI-136	D*EF	9.08	HARC, NARO
TI-137	EF	8.78	HARC, NARO
TI-145	D*EF	7.79	HARC, NARO
To15	D*EF	8.75	TAES, Hokkaido Govt. ‖
To38	D*EF	9.55	TAES, Hokkaido Govt.
To85	D*EF	8.50	TAES, Hokkaido Govt.
To90	D*EF	9.30	TAES, Hokkaido Govt.
To113	D*EF	7.70	TAES, Hokkaido Govt.

The prediction model was derived from the training data sets of the 41 elite inbred lines and 2,581 SNPs.

D: dent, F: flint, EF: European flint. * is derived from crosses between two groups.
† Chushin Agricultural Experiment Station, Nagano Pref.
‡ Kyushu Okinawa Agricultural Research Center, NARO
§ Institute of Livestock and Grassland Science, NARO
¶ Hokkaido Agricultural Research Center, NARO
‖ Tokachi Agricultural Experiment Station, Hokkaido Govt.

J, JC Mi, Na, Ho, and To are the registered or promising inbred lines.
IM, N, and TI are the superior inbred lines before Mi, Na, and Ho were named.