Upcycling yield trial data using a weather-driven crop growth model

Hiroyuki Shimono (shimn@iwate-u.ac.jp)
Iwate University

Akira Abe
Iwate Biotechnology Research Center https://orcid.org/0000-0002-0344-2643

Chyon Kim
Iwate University

Chiashi Sato
Ifuu Rinrin

Hiroyoshi Iwata
Laboratory of Biometry and Bioinformatics, Graduate School of Agricultural and Life Sciences, The University of Tokyo

Article

Keywords: Big data, Genomic prediction, Breeding, Genetic gains, Meta-analysis, Pedigree

Posted Date: April 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1544941/v1

Upcycling yield trial data using a weather-driven crop growth model

Hiroyuki SHIMONO[1,2]*, Akira ABE[3], Chyon Hae KIM[4], Chikashi SATO[5], Hiroyoshi IWATA[6]

[1] Faculty of Agriculture, Iwate University, 3-18-8, Ueda, Morioka, Iwate 020-8550, Japan
[2] Agri-Innovation Center, Iwate University, 3-18-8, Ueda, Morioka, Iwate 020-8550, Japan
[3] Iwate Biotechnology Research Center, 22-174-4, Narita, Kitakami, Iwate 024-0003, Japan
[4] Faculty of Science and Engineering, Iwate University, 3-18-8, Ueda, Morioka, Iwate 020-8550, Japan
[5] Ifuu Rinrin, Takata aza sunabatake 77-9, Rikuzentakata, Iwate 029-2205, Japan
[6] Laboratory of Biometry and Bioinformatics, University of Tokyo
*Corresponding Author: shimn@iwate-u.ac.jp, tel./fax: +81-19-621-6146

Key words

Big data, Genomic prediction, Breeding, Genetic gains, Meta-analysis, Pedigree

Abbreviations

CGM, crop growth model; *DVI*, development index; gBLUP, genomic best linear unbiased prediction; gpBLUP, a combination of gBLUP and pBLUP; GP, genomic prediction; GWAS, genome-wide association study; pBLUP, pedigree-based best linear unbiased prediction; *SPY*, the standardized potential yield; *RMSE*, root-mean-square error; SNP, single-nucleotide polymorphism; SRA, sequence read archive; Y_p, potential yield

Abstract

Numerous breeding efforts have improved crop yields by screening new high-yield cultivars through yield trials in multiple environments. However, the accumulated data from these trails has not been effectively upcycled to guide future breeding programs because of the strength of the genotype by environment (G × E) interaction. Here, we propose a new method that accounts for these interactions by using a weather-driven crop growth model; the recorded yields of each cultivar were expressed using a unique linear regression in response to the potential yield (Y_p) calculated by a weather-driven crop growth model. We call this approach the YpCGM method. We applied this method to 72 510 independent datasets from yield trials of paddy rice that used 237 core cultivars (n = 20 to 6342 field trials) measured at 110 locations in Japan during the 38 years from 1980 to 2017. The 237 core cultivars were selected from a pedigree matrix of 14 032 Japanese cultivars by using the k-medoids method (k = 200 clusters), and the genomic

information for 91 800 single-nucleotide polymorphisms was obtained. The genotypic coefficients of yield-ability and yield-plasticity differed among the 237 cultivars, with values ranging from 2.5 to 7.3 t/ha and from –0.23 to +0.95, respectively. Genomic prediction validated the values of these two parameters by leave-one-out cross-validation based on the pedigree and genome information. Our study represents a novel “big data” method for managing and generalizing G × E interactions using a crop growth model supported by large amounts of field data.

Introduction

Crop yield must improve to ensure food security by meeting the growing demand for food, which is predicted to increase by 70% by 2050 [1]. This must be accomplished under the constraints imposed by climate change. Genetic improvement of crop performance will be one important solution. Recent advances in genotyping technologies have accelerated our ability to screen candidate high-yielding cultivars efficiently using methods such as genome-wide association studies (GWAS) and genomic prediction (GP) [2-10]. Although genotyping of a given cultivar is now relatively easy and affordable using genome-wide DNA sequencing, phenotype analysis based on crop yield must be repeatedly evaluated in season-long field trials across a range of environments. The yield of a given cultivar is not stable across environments, since yield results from the interactions of many physiological processes that respond to environmental fluctuations throughout the growing season. This leads to strong genotype by environment (G × E) interactions, which make it difficult to predict crop yields across a range of environments.

The pioneering study of Finlay and Wilkinson[11] proposed a method for quantifying the G × E interaction by linear regression (FW method). The FW method standardized the observed yield of each cultivar in a given environment against a mean yield of all cultivars in a comparable environment (Supplemental Fig. S1a). The slope of the regression represents the plasticity of the yield response, with the average plasticity equal to 1.0; cultivars with above-average plasticity have a slope of >1.0, and cultivars with below-average plasticity have slope of <1.0. This represents a foundational achievement for evaluating G × E interactions, and it has been widely used in plant science, including genomic analysis [12-18]. However, the FW method can be used only with yield data measured at the same site in a side-by-side yield comparison, in which different cultivars grow together under identical environments. The plasticities of a cultivar that were determined under different environmental conditions are consequently not comparable among experiments.

To solve this problem, we designed this study to develop a new method for characterizing the cultivar-specific yield response standardized by the potential yield (Y_p) by using a weather-driven crop growth model, which we call the "YpCGM method" (Supplemental Fig. S1b). The novelty of this method is that it lets researchers combine yield trial data from different studies to estimate Y_p from weather data. Only two parameters are required: the yield-plasticity (α; dimensionless) and the yield-ability (β; t/ha). These represent the slope and estimated yield at the standardized potential yield (*SPY*; t/ha), respectively, in a cultivar-specific regression:

$$Y_{obs} = \alpha\, Y_p + b \qquad (1)$$
$$= \alpha\,(Y_p - SPY) + \beta$$

where Y_{obs} and Y_p represent the observed and potential yield of a given cultivar in a given year, location, and management regime, and b is the intercept at $Y_p = 0$. Our conversion of equation 1 to a form that uses SPY lets us characterize the yield-ability β, which represents the intercept at $Y_p = SPY$ in x-axis, thus the expected yield at $Y_p = SPY$. SPY can be set by considering the majority of the observed Y_p range among the tested cultivars. This covers the observed range of yield and provides a straightforward understanding of productivity of a give cultivar.

The regression provides a simple expression of a cultivar's yield characteristics of α and β. This differs from previous modelling studies, which required the measurement and parameterization of many physiological processes for each cultivar, including leaf expansion, photosynthesis, biomass production, and carbon allocation to harvestable organs [19-22]. Our method can be applied in a "big data" context using the accumulated yield data from many previous studies, including studies that recorded only yield, without measuring these and other physiological processes. Reuse (i.e., upcycling) of these data for yield phenotyping has great potential to help breeders identify promising cultivars that will let them boost crop yields by using currently available genotyping data.

We estimated the potential yield for a cultivar in each year and at each location by using a weather-driven crop growth model [23] with fixed default parameters. We used the observed cultivation schedule (transplanting, heading, and maturity dates) for each trial as input variables to increase the accuracy of the cultivar-specific coefficients by standardizing the variations in the growth period of each cultivar in each year and location. This method is based on the results of a preliminary study [23] that compared linear regressions of the yield of three cultivars as a function of Y_p using the model, but that did not extract the parameters as the genotypic coefficients described here.

We applied our new method to datasets from 72 510 yield trials under precise management practices of disease and pest control with 237 core cultivars of rice (*Oryza sativa* L.) obtained during the 38 years from 1980 to 2017, measured at 110 public agricultural experimental stations in Japan along a climatic transect from south to north and from west (facing the Sea of Japan) to east (facing the Pacific Ocean), which cover latitudes from 30°N to 43°N. Supplemental Figure S2 shows the locations of the study sites, and Supplemental Table S1 provides geographical details, time periods, and numbers of cultivars.

The 237 core cultivars were selected for their proximity to the central node of a pedigree

matrix, and were classified using the *k*-medoids method (k = 200 clusters) from the 14 032 rice cultivars for which data were available, which ranged from the oldest released cultivar ('Asahimochi' released in 1926) to the most recently released ('Tachiharuka' and 'Yukigozen' released in 2010). Each cultivar provided a dataset with 20 to 6342 phenotype datasets (yield, days to heading and maturity, panicle number, and panicle length). We validated the reliability of our method of calculating Y_p by using leave-one-out cross-validation for the cultivars based on their pedigree and genome. The cultivar-specific parameters were used to analyse the genetic gain and for GWAS to support future breeding studies for the development of cultivars suitable for the world's changing climate.

Results

Y_p captures cultivar-specific yield characteristics

The phenotype data for the 237 cultivars recorded data from 20 to 6342 yield trials under a range of environmental conditions (Supplemental Fig. S3). For example, the popular cultivar 'Koshihikari', which was used as a parent of 853 progeny, provided phenotyping data from 6342 trials, with yields that ranged from 0.9 to 8.7 t/ha (Supplemental Fig. S3vii), with 39 to 117 days to heading (Supplemental Fig. S3xvii), 178 to 684 panicles per m^2 (Supplemental Fig. S3xxvii), and a panicle length of 12.6 to 23.8 cm (Supplemental Fig. S3xxxvii). The other nine cultivars with 'Koshihikari' in these figures, which currently account for more than half of crop production in Japan, show a similar range of variation in their phenotypes (Supplemental Figure S3). Data for all 237 cultivars are listed in Supplemental Tables S2 and S3.

To quantify the genotypic coefficients of yield-ability (β in equation 1, defined as the expected yield at Y_p = 8 t/ha, *SPY*) of a given cultivar and its plasticity (α) using the YpCGM method, we plotted the observed yields per cultivar against Y_p estimated from weather records by the weather-driven crop-growth model. The genotypic coefficients of β varied from 2.5 to 7.3 t/ha among the 237 cultivars (Fig. 1a), and α varied from −0.23 to +0.95 (Fig. 1b). Fig. S4i-x is example of the quantification of coefficients for the 10 major cultivars by YpCGM, and Supplemental Table S2 listed coefficients of all 237 cultivars.

To compare the prediction accuracy of Y_p with estimates based on the observed panicle number and length, which strongly determine yield variation in rice [24], we calculated root-mean-square error (*RMSE*). The *RMSE* of the difference between the observed yield and Y_p ranged from 0.36 to 1.45 t/ha and averaged 0.84 t/ha for the 237 cultivars (Supplemental Table S3). The *RMSE* of Y_p was plotted against that of observed panicle number and length, and we found 1.2% smaller than the value based on the observed panicle number and 4.2% smaller

than that based on length (Fig. 2; Supplemental Fig. S4 xi-xxx). On this basis, Y_p appears to be an appropriate index that accounts for the variation of observed yield of each cultivar in response to environmental changes.

GP for explaining variations of the genotypic coefficients

Using leave-one-out cross-validation, genomic prediction explained the variation of yield-ability (β) based on genome information from 91 800 single-nucleotide polymorphisms (SNPs), with a significant positive correlation ($r = 0.34$, $P < 0.001$) in the genomic best linear unbiased prediction (gBLUP), which was similar the result for the pedigree-based best linear unbiased prediction (pBLUP; $r = 0.34$, $P < 0.001$) and the combination of genomic and pedigree BLUP (gpBLUP) ($r = 0.32$, $P < 0.001$) and the combination BLUP with the pedigree × genotype interaction (g×pBLUP) ($r = 0.33$, $P < 0.001$) (Fig. 3a). The yield-plasticity (α) was predicted with similar accuracy by the pedigree and genome methods, at $r = 0.19 \sim 0.22$, respectively ($P < 0.01$), and the values were lower than the correlations for β. Interestingly, the *b* parameter in the basic form of equation 1 that does not include *SPY* was not well explained by the genome and pedigree information ($r = -0.19 \sim -0.01$). This demonstrates the effectiveness of using β rather than *b* for characterizing the genotypic characteristics.

The heritability of β to explain the variation was significant and high, at 0.556 for genome information and 0.994 for pedigree information. These values were much higher than those of α, at 0.365 and 0.179, respectively. We also observed poor heritability of *b*, at 0.148 and 0.046, respectively.

Discussion

We successfully standardized the cultivar characteristics of rice yield using two genotypic coefficients that describe yield-ability (β) and its plasticity (α), and using Y_p from the weather-driven crop growth model to account for environmental effects on the observed yield data. In this, we used a big data approach that processed data from 72 510 independent rice yield trials (Fig. 1ab, Supplemental Table S2). The advantage of our novel method is that it allowed the use of a large volume of yield data from 20 to 6342 trials per cultivar under different environmental conditions to determine each genotypic coefficient for 237 core cultivars. We confirmed our results using genomic prediction analysis, and in particular revealed the high heritability of β (Fig. 3). The α and β coefficients will guide future breeding programs and help breeders to select candidate donor cultivars predicted to have high yield under future climates.

Our method should also be applicable to the millions of phenotyping records of crop species

such as wheat [7,18], maize [6,3], sorghum [4], chickpea [5] and common bean [8] based on data from field trials under a diverse set of growing conditions. These datasets can be analysed using a crop growth model that was developed previously for each species, such as for wheat [25], maize [26], and rice [27]. Future use of our concept would upcycle previous phenotyping data and increase the efficiency of breeding by calculating the genotypic coefficients for potential yield and yield plasticity.

Genetic gain can be defined as the yield increase over time and can be measured by evaluating the yield performance of rice cultivars released in different years but grown under the same experimental conditions [28-29]. Our novel method let us compare the genetic gain in studies conducted in different years and at different locations with Japan's 237 core cultivars by calculating just two genetic coefficients; we also performed this analysis for the 23 most widely grown cultivars from 1920 to the present (i.e., farmer favourites) (Supplemental Fig. S5). The cultivars that we included in our analysis represent a diverse range of maturity groups that could not be easily compared directly side by side, particularly owing to differences in photoperiod sensitivity [30].

Figure 1c shows that genetic gain in β of yield-ability from 1926 to 2010 was 10.2 kg/ha/year among all 237 core cultivars combined and more than double that, at 21.8 kg/ha/year, among the 23 major cultivars. The magnitude of the latter increase in regions from Hokkaido to Kyushu was similar to the range reported by Anzoua et al. [31] in Hokkaido, at 21 to 29 kg/ha among eight cultivars in eight cultivars that were introduced between 1905 and 1988 in a 2-year trial, and the results of Zhang and Kokubun [32], at 17 kg/ha for 10 cultivars introduced between 1893 and 1991 at three sites with different environments in the Tohou region (calculated from their Fig. 1).

In terms of yield plasticity of α, the genetic gain of the 237 core cultivars was not significant, but interestingly, that of the 23 major cultivars was marginally significant ($P = 0.065$), with a 1.1% annual increase relative to the mean plasticity of the 23 cultivars (Fig. 1d). This suggests that Japanese rice breeders have continuously genetically improved their rice cultivars in terms of potential yield, whereas local farmers have been choosing cultivars not only for higher potential yield, but also with the goal of obtaining higher yield to cope with a changing climate. A similar breeding direction to increase yield plasticity was reported of wheat cultivars in Argentina, Australia, Italy, and the UK [33].

It is worth noting that we observed no significant relationship between the potential yield and yield plasticity of all 237 cultivars combined or of the 23 major cultivars (Fig. 1e). This suggests that the two traits are independently genetically controlled.

Plant growth under favourable environmental conditions generally benefits from increased availability of resources, including temperature [34-35], atmospheric CO_2 [36-39], and solar radiation [40]. We hypothesized that a cultivar with higher yield plasticity may increase yield when grown with no constraints imposed by water availability but under a changing climate in Japan. We tested six scenarios based on different yield-plasticity of α with values ranging from 0.0 to 0.6, in scenarios CV1 to CV6 (i.e., six values within the observed range)(Supplemental Fig S6a). We showed the results assuming a fixed yield-ability, β at 8 t/ha in 6342 field trial environments using cultivar 'Koshihikari' and data from 1980 to 2017. Simulation CV1, with zero plasticity, provided a constant yield of 8 t/ha, and increasing plasticity increased the yield, with the maximum increase in scenario CV6, with $\alpha = 0.6$ (Supplemental Fig. S6b). The absolute yield increased with increasing plasticity throughout the simulation period, and with the greatest gain in the 2010s, at 14%, versus the 1980s, at 8%. This simulation was based on historical weather records and suggested that yield plasticity is a promising future breeding target to accelerate genetic gain in rice yield as an adaptation to climate change in Japan.

Supplemental Figure S6c suggests that environmental conditions have improved for cultivar 'Koshihikari' from 1980 to 2017, leading to a observed yield increase of 13.5 kg/ha/year (n = 6342; Supplemental Fig S11c). This trend agrees with previous results for rice cultivar 'Sasanishiki' [41]. This response in Japan contrasts with results from a tropical climate in the Philippines, where yield has been decreasing [42]. The trend for 'Koshihikari' may have resulted from increasing mean air temperature before heading, which increased by 0.03 °C/year, in the range c.a. 18 to 26°C (Supplemental Fig. S6d) but without a significant trend in solar radiation (Supplemental Fig. S6e). Y_p, which represents the cumulative effect of environmental resources, increased by 21.5 kg/ha/year (Supplemental Fig. S6f) even without taking into account for CO_2 fertilization effect in this model [36-39]. Jagermeyr et al. (2021)[43] predicted increased rice yield from 2069 to 2099 under future climate scenario SSP585, especially at higher latitudes, but a decrease at lower latitudes. Their prediction means that environmental resources will become better for rice at high latitudes. The genetic control of yield plasticity would be a powerful option to compensate for the predicted yield losses at lower latitudes, and could boost global crop productivity under future climatic conditions if this plasticity can strengthen abiotic stress tolerance [44]. This approach, combined with stronger tolerance of abiotic stress to mitigate the effects of unusual climate events, which are expected to increase in frequency and severity, could play a key role in securing global food security for the predicted population of 9 billion people in 2050.

GWAS lets us explore genomic regions related to yield, and the results can provide guidance

for future breeding targets. However, one limitation of GWAS is that it requires a large quantity of yield phenotype data for a large number of cultivars [22]. Our new method may let breeders use the large number of historical records that are available to calculate the two genotypic coefficients that we identified here. In Supplemental Figure S7, we show the result of our GWAS to identify new genomic regions associated with yield for either coefficient of α and β. The analysis revealed several but non-significant peaks, especially β, which had a high heritability (Fig. 3b). We identified seven potential peaks, with one on chromosome 2 (position 15 039 574; −log10(p) = 4.04) is close to the respective locations of *TAC4*, which controls tiller angle by regulating the endogenous auxin content [45]. The additional previously unknown regions for potential yield may exist (Supplemental Figure S7). Although we could not detect significant peaks, probably because yield is regulated by many physiological processes in complex and interacting ways [24], our methodology might be a useful way to do GWAS supported by other forms of genomic analysis such as detection of quantitative trait loci by recycling data from previous research, without requiring time-consuming and labour-intensive field trials.

One limitation of our new approach results from sacrificed statistical power of n in the big data, compiles the variations of 72 510 yield of each cultivar in the linear regression into 237 cultivars as two genotypic coefficients. Using the full dataset of 72 510 individual yield trials would improve the model's ability to account for G × E interactions and for the genotypic characteristics of each cultivar. Recently, several authors tried to improve the accuracy of GP for genomic and phenotyping data in a use of intermediate secondary traits of means of environment during a specific growth period or to improve a model's output [2-3,9,10]. We may extend our approach through the use of the full 72 510 yield trial data to characterize genotypic coefficients of the 237 cultivars.

A second limitation was that we used only a single crop growth model to calculate Y_p. This means that the coefficients include an error component caused by the structure of the model, which, for example, does not account for soil type and fertility [46], atmospheric CO_2 [19], or the temperature of the irrigation water [47]. Another improvement would be to test ensemble multi-model analysis using two or more models with different complexities and different abilities to account for physiological processes so as to increase the accuracy of the Y_p calculation, as done recently in simulations of future crop yield [25-27,43]. Such an approach improves on previous labour-intensive and time-consuming research methods, which previously required months of careful work for each dataset.

Conclusion

We developed a new method for mitigating the effects of the genotype × environment (G × E), YpCGM method: using a weather-driven crop growth model to combine data from independent yield trials under different environmental characteristics with the model's estimated potential yield based on weather data to re-analyse the valuable but under-used data from previous trials. To the best of our knowledge, we are the first to use this approach to mitigate the effects of the G × E interaction and characterise the productivity of cultivars by integrating data from many trials across a range of environments using only two regression coefficients.

Online Methods

Upcycling phenotypic datasets and select core cultivars

We obtained rice yield datasets from a total of 207 331 trials with 8524 cultivars during the 38 years from 1980 to 2017. The studies were conducted at 110 public agricultural experimental stations in Japan by the Institute of Crop Science of the National Agriculture and Food Research Organization, Japan (NARO) (2017 version) [48]. From the database, we selected 237 core cultivars, for a total of 72 510 yield datasets, as follows: (1) From the pedigree network information of 14 032 rice cultivars, we selected 200 cultivars that were central nodes in the classification by the k-medoids method ($k = 200$) (Supplementary Fig. S8). (2) One cultivar has many phenotypic data from each cluster constructed by the hierarchical clustering according to the Wards method based on the pedigree relationship matrix (except for the cluster containing 200 cultivars selected by the k-medoids method described above) were selected. We selected a total of 158 cultivars. (3) Finally, 237 core cultivars with at least 20 trials data per cultivar (to maintain the accuracy of the regression analysis when n was incrementally increased) and seeds available were selected.

We cleaned the data to remove outliers by using the interquartile range 4.0. The 237 cultivars included the oldest released cultivar ('Asahimochi' released in1926) and the most recently released ('Tachiharuka' and 'Yukigozen' released in 2010) (Supplemental Tables S2 and S3). Each cultivar provided trials of $n = 20$ to 6342 yield trials of phenotype datasets (yield, days to heading and maturity, panicle number, and panicle length). Each study used the best management practices for their time in terms of site preparation, fertilization rate, and pest and disease control. All yields were converted to a 14% moisture content before reporting.

Whole genome resequencing and SNP genotyping

We resequenced 166 rice cultivars (Supplemental Table S4). First, we extracted genomic DNA from young leaf tissue of each cultivar with the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). We then quantified the DNA with a Qubit fluorometer (Invitrogen, Waltham, MA, USA). Next, we constructed libraries with the Riptide High Throughput Rapid DNA Library Prep Kit (iGenomX, Carlsbad, CA, USA), following the manufacturer's protocol. We sequenced the Riptide libraries on the Illumina NovaSeq platform (Illumina, San Diego, CA, USA), using 150-bp paired-end reads, and then demultiplexed the results in fgbio software v.1.3.0 (https://github.com/fulcrumgenomics/fgbio). We performed additional sequencing on the Illumina HiSeq platform of cultivars with a low number (< 6 300 000) of sequence reads. The sequence data have been deposited in the DNA Data Bank of Japan Sequence Read Archive (DRA) (Supplemental Table S4). In addition, we obtained FASTQ reads from the DRA for 70 rice cultivars [49-50]. DRA accessions are provided in Supplemental Table S4.

The raw sequence reads were qualified (quality trimming and adapter clipping) in Trimmomatic software v. 0.39 [51] with the options "PE -phred 33 ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:15 MINLEN:75". We retained only paired output reads. After these preprocessing steps, we mapped the remaining reads onto the 'Nipponbare' reference genome IRGSP-1.0, using the *bwa mem* command in BWA [52], with the options "-a -T 0". We obtained coordinate-sorted files in BAM format using the samtools *sort* command in SAMtools v. 1.9 [53]. We obtained BAM files that contained only correctly oriented and properly paired mapped reads by filtering the specified bit in the FLAG field during scanning with the samtools *view* command. The BAM files were filled in as mate coordinates by using the samtools *fixmate* command, and finally we obtained BAM files that contained only paired mapped reads by using the samtools *view* command again.

SNP-based genotype calling can be obtained as a file in variant call format (VCF). We generated the VCF file from the BAM files using the samtools *mpileup* command with the options "-q 10 -Q 13 -a DP,AD,SP,ADF,ADR -B -C 50," and called and filtered the variants using the bcftools *call* command in BCFtools v. 1.9 [53] with the options "-v -m -f GQ,GP" and the bcftools *filter* command with the options "-i 'INFO/MQ> = 10". We then filtered twice using the bcftools *view* command with the options "-i 'F_PASS(FMT/DP> = 5 & FMT/GQ> = 20) > 0.8'" and "-i 'INFO/MQ> = 30". Subsequently, we converted the genotypes with low genotype quality and low depth to "missing" in VCFtools [54] with the options "--minGQ 10 --minDP 2" and filtered using the bcftools *view* command with the options "-e 'MAF[0] < 0.001' -i 'F_PASS(FMT/GT=="het") < 0.1'". We imputed the resulting genotype dataset in

BEAGLE v. 5.1 [55]. Finally, we completed a core genotype dataset comprising 237 cultivars, including 'Nipponbare' and 91 800 SNPs.

Crop growth model

We performed our analysis using a simple crop growth model developed by Masuya and Shimono [23] using the default parameter values. The model integrates daily canopy radiation capture and use as a function of input daily air temperature and solar radiation. In the development sub-model, we predicted the phenology of a developmental index (*DVI*) [56], and combined this model with observed heading and maturity dates. The *DVI* on each date was calculated from the air temperature by using the ratio of the actual temperature to the cumulative effective air temperature (here, defined as >10 °C) for each period between the observed transplanting, heading and maturity dates. This eliminated the need for daylength input data and the need for cultivar-specific parameters that accounted for different phenological responses. This let us standardize the yield per unit area by using the observed solar energy capture even for identical cultivars grown in different years and at different locations, as well as for cultivars with different maturity dates.

We set *SPY* in equation (1) as 8 t/ha because Y_p of 201 in the 237 cultivars (85%) covered 8 t/ha in which Y_p of the 237 cultivars averaged 5.0 t/ha, with a minimum of 0.1 t/ha and a maximum of 9.8 t/ha.

Weather data

The daily air temperature and solar radiation during the study period at the weather station closest to each of the 110 locations that provided yield data were obtained from the MeteoCrop database (https://meteocrop.dc.affrc.go.jp/real/top.php).

GWAS analysis

We performed GWAS using linear mixed models to correct for differences in the population structure and genetic relationships, using the *gaston* package in R software [57]. We used α and β of the 237 cultivars listed in Supplemental Table S2. We removed SNPs that were not biallelic, that were monomorphic, or that had a minor allele frequency of <5%, leaving a total of 77 651 SNPs.

GP analysis: Prediction of genotypic coefficients and calculation of heritability

The accuracy of the predictions based on genomic and pedigree data for the genotypic

coefficients α and β was evaluated based on the genomic (realization) relationship matrix and the pedigree (numerator) relationship matrices. We tested genomic best linear unbiased prediction (gBLUP), pedigree-based best linear unbiased prediction (pBLUP), a combination of the two approaches (gpBLUP), and a prediction based on the genome × pedigree information (g×pBLUP).

The accuracy of the predictions based on genomic and pedigree data for the genotypic coefficients α and β was evaluated. Predictions were performed based on the genomic (realization) relationship matrix and the pedigree (numerator) relationship matrices.

The genomic relationship matrix was calculated using the same set of SNPs used for the GWAS (91,800 SNPs satisfying MAF ≥ 0.5). Specifically, it was calculated as $\mathbf{G} = \mathbf{X}\mathbf{X}'/m$ based on the matrix $\mathbf{X}$ representing the SNP genotypes of the 237 cultivars. where the (i, j) element of $\mathbf{X}$ represents the genotype of the j-th SNP of the i-th cultivar, scored as -1 for reference-type homozygous, 1 for non-reference-type homozygous, and 0 for heterozygous. m represents the number of SNPs. The pedigree relationship matrix was calculated based on the parent information of each cultivar obtained from the rice cultivar database in NARO (https://ineweb.narcc.affrc.go.jp/). The pedigree relationship matrix, $\mathbf{A}$, was calculated based on the pedigree information of 15,145 cultivars, including lines that were included as dummies to account for backcrossing steps.

Prediction models were constructed using the genomic and pedigree relationship matrices. Specifically, the following four BLUP models were constructed.

gBLUP:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{g}_{\mathrm{g}} + \mathbf{e}$$

pBLUP

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{g}_{\mathrm{p}} + \mathbf{e}$$

gpBLUP

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{g}_{\mathrm{g}} + \mathbf{g}_{\mathrm{p}} + \mathbf{e}$$

gxpBLUP

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{g}_{\mathrm{g}} + \mathbf{g}_{\mathrm{p}} + \mathbf{g}_{\mathrm{gxp}} + \mathbf{e}$$

where $\mathbf{y}$ is the vector of values of genetic coefficients whose i-th element is α_i or β_i, $\boldsymbol{\mu}$ is the vector whose elements are the overall mean, $\mathbf{g}_{\mathrm{g}}$ is the vector of genetic effects accounted by the genomic-based relationships, i.e., $\mathbf{g}_{\mathrm{g}} \sim N(\mathbf{0}, \mathbf{G}\sigma_G^2)$, where σ_G^2 is the variance of the

genomic-based effects, $\mathbf{g}_{\mathrm{p}}$ is the vector of genetic effects accounted by the pedigree-based relationships, i.e., $\mathbf{g}_{\mathrm{p}} \sim N(\mathbf{0}, \mathbf{A}\sigma_A^2)$, where σ_A^2 is the variance of the pedigree-based effects, and $\mathbf{g}_{\mathrm{gxp}}$ is the vector of genetic effects accounted by the interaction between genomic- and pedigree-based relationships, i.e., $\boldsymbol{\gamma} \sim N\left(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}\sigma_{GA}^2\right)$, where σ_{GA}^2 is the variance of the interaction effects and $\otimes$ is the element-wise product (Hadamard product) of matrices, $\mathbf{Z}$ is the design matrix for the genetic effects, and $\boldsymbol{\epsilon}$ is the vector of residuals assuming $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma_E^2)$, where σ_E^2 is the residual variance. The gxpBLUP is the full model with interaction terms between genomic and pedigree relationships [58]. The function "lmerUvcov" in the R package "lme4GS" [59] was used to estimate the parameters of the model and to calculate BLUP values.

To evaluate the prediction accuracy of the model, we conducted one-cultivar out cross-validation. Specifically, we estimated the parameters of a prediction model and calculated BLUPs for 236 cultivars, and then predicted the y value for the left-out one cultivar. This step was repeated for all the 237 cultivars.

To evaluate the levels of genetic control of the genotypic coefficients α and β, we calculated heritability using genomic and pedigree relationship matrices. For the genotypic coefficients, one value was obtained for each cultivar and there were no replications. Thus, heritability was calculated simply by taking the ratio of the variation explained by $\mathbf{g}_{\mathrm{g}}$ (i.e., σ_G^2) or $\mathbf{g}_{\mathrm{p}}$ (i.e., σ_A^2) and the variance explained by residuals (i.e., σ_E^2). In the calculation of heritability, the parameters of the gBLUP and pBLUP models were estimated using data from all 237 breeds/strains. The function "marker_h2_means" in the R package "heritability" [60] was used for the calculation.

Ethics statement

The authors declare that all methods were performed in accordance with the relevant guidelines and regulations. The following organizations provided the test materials: NARO and local public agricultural research stations at Hokkaido pref., Aomori Pref., Iwate Pref., Yamagata Pref., Tochigi Pref., Chiba Pref., Nagano Pref., Niigata Pref., Toyama Pref., Fukui Pref., Shiga Pref., Aichi Pref., Hyogo Pref., Miyazaki Pref., and Kagoshima Pref. We obtained permission from all organizations to use these accessions.

Contributions

H.S. conceived the idea for the study. H.S., C.S., A.A., C.H.K. and H.I. performed the data analysis and participated in interpretation of the results and manuscript preparation.

450

451 **References**

452 1. FAO. To feed world population in 2050. FAO Newsletter 44 (2009).

453 2. Technow, F., Messina, C. D., Totir, L. R. & Cooper, M. Integrating crop growth models with
454 whole genome prediction through approximate Bayesian computation. PLOS ONE 10,
455 e0130855 (2015).

456 3. Cooper, M., Technow, F., Messina, C., Gho, C. & Totir, L. R. Use of crop growth models
457 with whole-genome prediction: application to a maize multienvironment trial. Crop Science
458 56, 2141–2156 (2016).

459 4. Yu, X. et al. Genomic prediction contributing to a promising global strategy to turbocharge
460 gene banks. Nature Plants 2, 16150 (2016).

461 5. Roorkiwal, M. et al. Genomic-enabled prediction models using multi-environment trials to
462 estimate the effect of genotype x environment interaction on prediction accuracy in chickpea.
463 Scientific Reports 8, 11701 (2018).

464 6. Millet, E. J. et al. Genomic prediction of maize yield across European environmental
465 conditions. Nature Genetics 51, 952–956 (2019).

466 7. de los Campos, G., Perez-Rodriguez, P., Bogard, M., Gouache, D. & Crossa, J. A data-driven
467 simulation platform to predict cultivars' performances under uncertain weather conditions.
468 Nature Communications 11, 4876 (2020).

469 8. Keller, B. et al. Genomic prediction of agronomic traits in common bean (Phaseolus vulgaris
470 L.) under environmental stress. Frontiers in Plant Science 11 (2020).

471 9. Robert, P., Le Gouis, J., Consortium, T. B. & Rincent, R. Combining crop growth modeling
472 with trait-assisted prediction improved the prediction of genotype by environment
473 Interactions. Frontiers in Plant Science 11 (2020).

474 10. Toda, Y. et al. Predicting biomass of rice with intermediate traits: Modeling method
475 combining crop growth models and genomic prediction models. PLOS ONE 15, e0233951
476 (2020).

477 11. Finlay, K. W. & Wilkinson, G. N. The analysis of adaptation in a plant-breeding programme.
478 Australian Journal of Agricultural Research 14, 742–754 (1963).

479 12. Kraakman, A. T. W., Niks, R. E., Van den Berg, P. M. M. M., Stam, P. & Van Eeuwijk, F.
480 A. Linkage disequilibrium mapping of yield and yield stability in modern spring barley
481 cultivars. Genetics 168, 435–446 (2004).

482 13. Inostroza, L. et al. Association mapping of plant height, yield, and yield stability in
483 recombinant chromosome substitution lines (RCSLs) using Hordeum vulgare subsp.

spontaneum as a source of donor alleles in a Hordeum vulgare subsp. vulgare background. Molecular Breeding 23, 365–376 (2009).
14. Lacaze, X., Hayes, P. M. & Korol, A. Genetics of phenotypic plasticity: QTL analysis in barley, Hordeum vulgare. Heredity 102, 163–173 (2009).
15. Kumagai, E., Homma, K., Kuroda, E. & Shimono, H. Finlay-Wilkinson's regression coefficient as a pre-screening criterion for yield responsiveness to elevated atmospheric CO_2 concentration in crops. Physiologia Plantarum 158, 312–317 (2016).
16. Gage, J. L. et al. The effect of artificial selection on phenotypic plasticity in maize. Nature Communications 8, 1348 (2017).
17. Zhang, H., Berger, J. D. & Herrmann, C. Yield stability and adaptability of canola (*Brassica napus* L.) in multiple environment trials. Euphytica 213, 155 (2017).
18. Xiong, W. et al. Increased ranking change in wheat breeding under climate change. Nature Plants 7, 1207–1212 (2021).
19. Yoshida, H. & Horie, T. A process model for explaining genotypic and environmental variation in growth and yield of rice based on measured plant N accumulation. Field Crops Research 113, 227–237 (2009).
20. Gu, J., Yin, X., Zhang, C., Wang, H. & Struik, P. C. Linking ecophysiological modelling with quantitative genetics to support marker-assisted crop design for improved yields of rice (*Oryza sativa*) under drought stress. Annals of Botany 114, 499–511 (2014).
21. Jones, M. R. et al. Exploring process-level genotypic and environmental effects on sugarcane yield using an international experimental dataset. Field Crops Research 244, 107622 (2019).
22. Kadam, N. N., Jagadish, S. V. K., Struik, P. C., van der Linden, C. G. & Yin, X. Incorporating genome-wide association into eco-physiological simulation to identify markers for improving rice yields. Journal of Experimental Botany 70, 2575–2586 (2019).
23. Masuya, Y. & Shimono, H. Mining a yield-trial database to identify high-yielding cultivars by simulation modeling: A case study for rice. Journal of Agricultural Meteorology 73, 51–58 (2017).
24. Yoshida, S. Fundamentals of rice crop science (IRRI, Los Baños, Philippines, 1981).
25. Wallach, D. et al. Multimodel ensembles improve predictions of crop-environment-management interactions. Global Change Biology 24, 5072–5083 (2018).
26. Bassu, S. et al. How do various maize crop models vary in their responses to climate change factors? Global Change Biology 20, 2301–2320 (2014).
27. Li, T. et al. Uncertainties in predicting rice yield by current crop models under a wide range

518 of climatic conditions. Global Change Biology 21, 1328–1341 (2015).
519 28. Austin, R. B. Yield of wheat in the United Kingdom: Recent advances and prospects. Crop
520 Sci 39, 1604–1610 (1999).
521 29. Peng, S., Cassman, K. G., Virmani, S. S., Sheehy, J. & Khush, G. S. Yield potential trends
522 of tropical rice since the release of IR8 and the challenge of increasing rice yield potential.
523 Crop Sci. 39, 1552–1559 (1999).
524 30. Hamasaki, T. et al. Comparison of temperature-gradient chamber experiments at different
525 latitudes for estimating the effects of global warming on the heading date of paddy rice.
526 Journal of Agricultural Meteorology 66, 193–200 (2010).
527 31. Anzoua, K. G., Junichi, K., Toshihiro, H., Kazuto, I. & Yutaka, J. Genetic improvements
528 for high yield and low soil nitrogen tolerance in rice (*Oryza sativa* L.) under a cold
529 environment. Field Crops Research 116, 38–45 (2010).
530 32. Zhang, W.-H. & Kokubun, M. Historical changes in grain yield and photosynthetic rate of
531 rice cultivars released in the 20th century in Tohoku region. Plant Production Science 7, 36–
532 44 (2004).
533 33. Calderini, D. F. & Slafer, G. A. Has yield stability changed with genetic improvement of
534 wheat yield? Euphytica 107, 51-59 (1999).
535 34. Shimono, H. & Okada, M. Plasticity of rice tiller production is related to genotypic variation
536 in the biomass response to elevated atmospheric CO_2 concentration and low temperatures
537 during vegetative growth. Environmental and Experimental Botany 87, 227–234 (2013).
538 35. Horai, K., Ishii, A. & Shimono, H. Genotypic variation in the response to high water
539 temperature during vegetative growth and the effects on rice productivity under a cool
540 climate. Field Crops Research 162, 12–19 (2014).
541 36. Hasegawa, T. et al. Rice cultivar responses to elevated CO_2 at two free-air CO_2 enrichment
542 (FACE) sites in Japan. Functional Plant Biology 40, 148–159 (2013).
543 37. Bishop, K. A., Betzelberger, A. M., Long, S. P. & Ainsworth, E. A. Is there potential to
544 adapt soybean (*Glycine max* Merr.) to future [CO_2]? An analysis of the yield response of 18
545 genotypes in free-air CO_2 enrichment. Plant, Cell and Environment 38, 1765–1774 (2015).
546 38. Kumagai, E., Aoki, N., Masuya, Y. & Shimono, H. Phenotypic plasticity conditions the
547 response of soybean seed yield to elevated atmospheric CO_2 concentration. Plant Physiology
548 169, 2021–2029 (2015).
549 39. Shimono, H. et al. Prescreening in large populations as a tool for identifying elevated CO_2-
550 responsive genotypes in plants. Functional Plant Biology 46, 1–14 (2019).
551 40. Sinclair, T. R. & Horie, T. Leaf nitrogen, photosynthesis, and crop radiation use efficiency:

552 a review. Crop Sci. 29, 90–98 (1989).
553 41. Shimono, H. Impact of global warming on yield fluctuation in rice in the northern part of
554 Japan. 77, 489–497 (in Japanese with English abstract) (2008).
555 42. Peng, S. et al. Rice yields decline with higher night temperature from global warming.
556 PNAS, Proceedings of the National Academy of Sciences of the United States of America
557 101, 9971–9975 (2004).
558 43. Jagermeyr, J. et al. Climate impacts on global agriculture emerge earlier in new generation
559 of climate and crop models. Nature Food 2, 873–885 (2021).
560 44. Rivero, R. M., Mittler, R., Blumwald, E. & Zandalinas, S. I. Developing climate-resilient
561 crops: improving plant tolerance to stress combination. The Plant Journal 109, 373-389
562 (2022).
563 45. Li, H. et al. TAC4 controls tiller angle by regulating the endogenous auxin content and
564 distribution in rice. Plant Biotechnology Journal 19, 64–73 (2020).
565 46. Hasegawa, T. & Horie, T. in Applications of systems approaches at the field level. (eds.
566 Kropff, M. J. et al.) 243–257 pp (Kluwer Academic, Dordrecht/ Boston/ London, 1997).
567 47. Shimono, H., Hasegawa, T. & Iwama, K. Modeling the effects of water temperature on rice
568 growth and yield under a cool climate: I. Model development. Agron J 99, 1327–1337 (2007).
569 48. Goto, A. et al. Rice yield trial database (1980-2017). CD-ROM. NARO Institute of Crop
570 Science Director of Rice Research Area (2017).
571 49. Yabe S, Yoshida H, Kajiya-Kanegae H, Yamasaki M, Iwata H, Ebana K, et al. Description
572 of grain weight distribution leading to genomic selection for grain-filling characteristics in
573 rice. PLoS ONE 13(11), e0207627 (2018).
574 50. Hamazaki, K., Kajiya-Kanegae, H., Yamasaki, M., Ebana, K., Yabe, S., Nakagawa, H.,
575 Iwata, H. Choosing the optimal population for a genome-wide association study: A
576 simulation of whole-genome sequences from rice. The Plant Genome 13, e20005 (2020).
577 51. Bolger, A. M., Lohse, M., & Usadel, B. Trimmomatic: A flexible trimmer for Illumina
578 Sequence Data. Bioinformatics 30, 2114-2120 (2014).
579 52. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
580 arXiv:1303.3997v2 [q-bio.GN] (2013).
581 53. Danecek, P. et al. The variant call format and VCFtools. Bioinformatics, 27 (15), 2156–
582 2158 (2011).
583 54. Danecek, P. et al. Twelve years of SAMtools and BCFtools. GigaScience 10, giab008
584 (2021).
585 55. Browning, B. L., Zhou, Y., Browning S. R. A one-penny imputed genome from next

generation reference panels. Am J Hum Genet 103(3), 338-348 (2018).
56. Horie, T. & Nakagawa, H. Modelling and prediction of developmental process in rice: I. Structure and method of parameter estimation of a model for simulating developmental process toward heading. Japanese Journal of Crop Science 59, 687-695 (1990)(in Japanese with English abstract)
57. Perdry, H., Dandine-Roulland, C. gaston: Genetic Data Handling (QC, GRM, LD, PCA) & Linear Mixed Models. R package version 1.5.7. (2020) https://CRAN.R-project.org/package=gaston
58. Howard, R. et al. Joint use of genome, pedigree, and their interaction with environment for predicting the performance of wheat lines in new environments. G3 9, 2925–2934 (2019).
59. Caamal-Pat, D. et al. lme4GS: An R-Package for genomic selection. Front. Genet. 12, 680569 (2021).
60. Kruijer, W. et al. Marker-based estimation of heritability in immortal populations. Genetics 199, 379–398 (2015).

Acknowledgements

This work was supported by a Japan Society for the Promotion of Science KAKENHI Grant (Number JP 19H00938). We thank Emeritus Prof. M. Okada of Iwate University for their helpful suggestions. We also thank M. Ozaki, A. Oyamada, Dr. Y. Masuya, K. Shioi and K. Nagai, Iwate University. We also thank K. Ito, Y. Ogasawara and S. Natsume, Iwate Biotechnology Research Center for their technical support. We also thank the following organizations for providing the seeds of rice cultivars: the Genebank at the National Agriculture and Food Research Organization, Japan (NARO), NARO rice breeding sections, and local public agricultural research stations at Hokkaido pref., Aomori Pref., Iwate Pref., Yamagata Pref., Tochigi Pref., Chiba Pref., Nagano Pref., Niigata Pref., Toyama Pref., Fukui Pref., Shiga Pref., Aichi Pref., Hyogo Pref., Miyazaki Pref., and Kagoshima Pref. In this research work we used the NIG supercomputer at ROIS National Institute of Genetics, and the supercomputer of ACCMS, Kyoto University.

Figure captions

Figure 1. Genotypic coefficients for the yield-ability (β) and the yield-plasticity (α) of the 237 core cultivars. Frequency distributions for (a) yield-ability and (b) yield-plasticity. Relationships between the year when a cultivar was released and the (c) yield-ability and (d) yield-plasticity. (e) Relationship between yield-ability and yield-plasticity. Red, green, and blue arrows in (a) and (b) represent 'Koshihikari', a long-popular cultivar; 'Nipponbare', an older popular cultivar that was used for the rice genome project; and 'Hokuriku 193', a high-yield cultivar, respectively. Black and red data points in (c), (d), and (e) indicate data for all 237 cultivars and for the 23 major cultivars, which represent the most widely grown cultivars from 1920 to 2020 (i.e., farmer favourite historical cultivars), respectively. ***$P < 0.001$, +$P < 0.1$, ns not significant. The 23 major cultivars were 'Norin 22' (1941), 'Kimmaze' (1948), 'Hounenwase' (1952), 'Koshihikari' (1953), 'Etsujiwase' (1953), 'Fujiminori' (1958), 'Sasanishiki' (1960), 'Nipponbare' (1961), 'Reimei' (1963), 'Todorokiwase' (1965), 'Toyonishiki' (1966), 'Reiho' (1966), 'Ishihikari' (1968), 'Akihikari' (1974), 'Yukihikari' (1981), 'Akitakomachi' (1982), 'Kirara 397' (1985), 'Hinohikari' (1986), 'Hitomebore' (1988), 'Fukuhibiki' (1988), 'Nanatsuboshi' (1998), 'Masshigura' (1999), 'Hokuriku 193' (2001). Number in the parenthesis is the year released.

Figure 2. Root-mean-square errors (*RMSE*) of yield prediction based on the potential yield (Y_p) calculated by the weather-driven crop growth model, and relative to the observed (a) panicle number (PN) and (b) panicle length (PL) of the 237 core cultivars. *RMSE* of each cultivar was calculated independently from $n = 20$ to 6342 trials.

Figure 3. (a) Comparison of the strength of the ability to predict the yield-ability (β in equation 1), yield-plasticity (α in equation 1) and *b* in equation 1 by pedigree-based best linear unbiased prediction (pBLUP), genome-based best linear unbiased prediction (gBLUP), their combination (gpBLUP), and genome × pedigree information (g×pBLUP). (b) Comparison of the heritability of yield of the 237 core rice cultivars. The correlations between the observed and predicted values were based on leave-one-out cross-validation ($n = 237$).

Figures

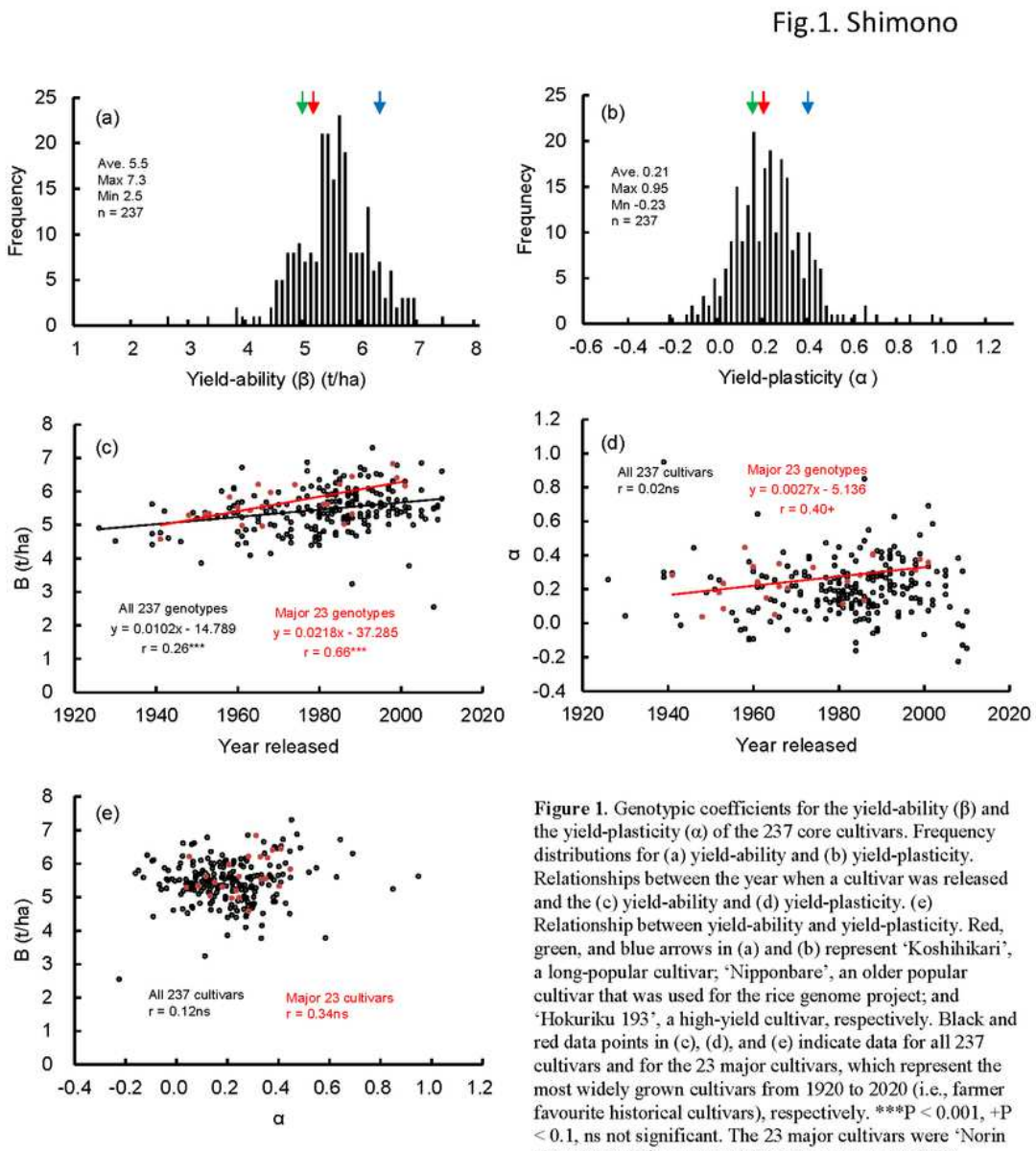

Figure 1. Genotypic coefficients for the yield-ability (β) and the yield-plasticity (α) of the 237 core cultivars. Frequency distributions for (a) yield-ability and (b) yield-plasticity. Relationships between the year when a cultivar was released and the (c) yield-ability and (d) yield-plasticity. (e) Relationship between yield-ability and yield-plasticity. Red, green, and blue arrows in (a) and (b) represent 'Koshihikari', a long-popular cultivar; 'Nipponbare', an older popular cultivar that was used for the rice genome project; and 'Hokuriku 193', a high-yield cultivar, respectively. Black and red data points in (c), (d), and (e) indicate data for all 237 cultivars and for the 23 major cultivars, which represent the most widely grown cultivars from 1920 to 2020 (i.e., farmer favourite historical cultivars), respectively. ***P < 0.001, +P < 0.1, ns not significant. The 23 major cultivars were 'Norin 22' (1941), 'Kimmaze' (1948), 'Hounenwase' (1952), 'Koshihikari' (1953), 'Etsujiwase' (1953), 'Fujiminori' (1958), 'Sasanishiki' (1960), 'Nipponbare' (1961), 'Reimei' (1963), 'Todorokiwase' (1965), 'Toyonishiki' (1966), 'Reiho' (1966), 'Ishihikari' (1968), 'Akihikari' (1974), 'Yukihikari' (1981), 'Akitakomachi' (1982), 'Kirara 397' (1985), 'Hinohikari' (1986), 'Hitomebore' (1988), 'Fukuhibiki' (1988), 'Nanatsuboshi' (1998), 'Masshigura' (1999), 'Hokuriku 193' (2001). Number in the parenthesis is the year released.

Figure 1

See image above for figure legend.

Fig.2. Shimono

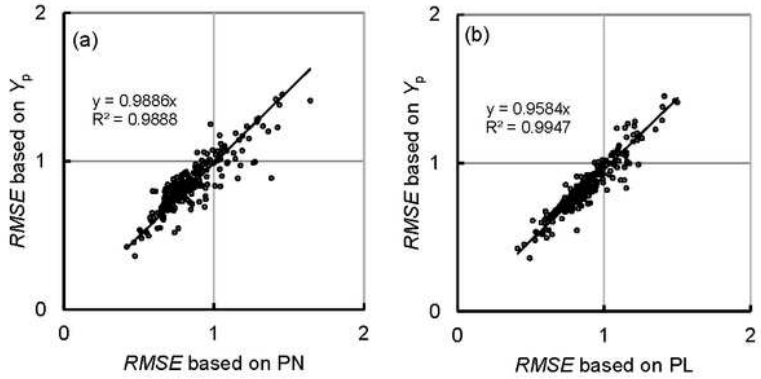

Figure 2. Root-mean-square errors (*RMSE*) of yield prediction based on the potential yield (Y_p) calculated by the weather-driven crop growth model, and relative to the observed (a) panicle number (PN) and (b) panicle length (PL) of the 237 core cultivars. *RMSE* of each cultivar was calculated independently from n = 20 to 6342 trials.

Figure 2

See image above for figure legend.

Fig.3. Shimono

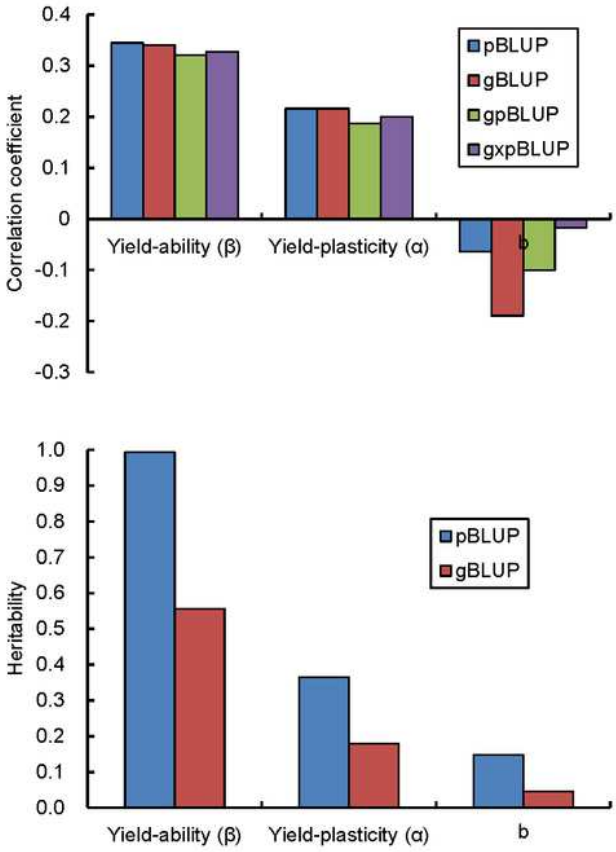

Figure 3. (a) Comparison of the strength of the ability to predict the yield-ability (β in equation 1), yield-plasticity (α in equation 1) and *b* in equation 1 by pedigree-based best linear unbiased prediction (pBLUP), genome-based best linear unbiased prediction (gBLUP), their combination (gpBLUP), and genome × pedigree information (g×pBLUP). (b) Comparison of the heritability of yield of the 237 core rice cultivars. The correlations between the observed and predicted values were based on leave-one-out cross-validation (n = 237).

Figure 3

See image above for figure legend.

Supplementary Files

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

- MergedSuppFigTab.pdf