Supplementary information

Multi-environment genomic selection in rice elite breeding lines

Nguyen V. H.[1,2,3], Morantte R. I. Z.[3], Lopena V.[3], Verdeprado H.[3], Murori R.[3], Ndayiragije A.[3], Katiyar S. [3], Islam M. R. [3], Juma R.[3], Galvez H.[4], Glaszmann J-C.[1,2], Cobb J. N. [3,5], Bartholomé J.[2,6,7*]

1 CIRAD, UMR AGAP Institut, F-34398, Montpellier, France

2 UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

3 Rice Breeding Platform, International Rice Research Institute, DAPO Box7777, Metro Manila, Philippines

4 University of the Philippines Los Baños, College, Los Baños, Laguna 4031, Philippines

5 RiceTec. Inc, PO Box 1305, Alvin, TX 77512

6 CIRAD, UMR AGAP Institut, Cali, Colombia

7 Alliance Bioversity-CIAT, Cali, Colombia

*Corresponding author: jerome.bartholome@cirad.fr

This additional file is containing the supplementary figures from the main text of the paper

Fig. S1 Physical positions of the 882 SNPs on the 12 chromosomes of rice using Nipponbare reference genome.

Fig. S2 Graphical representation of the allocation of ECP across 15 environments. Blue vertical lines represent the ECP-environment combinations that were observed, and the white lines (blanks) correspond to unobserved combinations.

Fig. S3 Schematic representation of the developmental stages of ECP in each environment. The main three developmental stages of each environment were identified in the following: the vegetative stage: the interval time from the transplanting date up to the date of 50% flowering of the first ECP; the flowering stage: being the interval time between the first ECP and the last ECP are flowered of 50%; maturity stage: the interval time from the date of 50% flowering of the last ECP up to harvesting.

Fig. S4 Unweighted neighbor-joining trees between the ECP and Oryza sativa groups from 3K-RG (A), or the indica subgroups (B). The NJ trees were constructed using 837 common SNPs by combining the SNPs data of 3K-RG (Wang et al., 2018) and the ECP from the 1K-RiCA assay.

Fig. S5 The scatter plot, histogram, and correlation between the DTF, HT, and YLD traits in the single environments. The names of environments are abbreviated with the format of country-location-year and season. At BD (Bangladesh), the trial environments were conducted at GZ (Gazipur), and NM (Nizmawna). At IN (India) conducted at HY (Hyderabad), CU (Cuttack), MA (Maruteru), and RP (Raipur). At KE (Kenya) conducted at AH (Ahero), and MW (Mwea). At MZ (Mozambique) conducted at CK (Chokwe), and MP (Maputo). At PH (Philippines) conducted at LB (Los Banos). At TZ (Tanzania) conducted at DK (Dakawa). 18W: Wet season in 2018. 19D: Dry season in 2019. 19W: Wet season in 2019. 20W: Wet season in 2020. *, **, ***: statistically significant at 5% ($p<0.05$), 1% ($p<0.01$), and 0.1% ($p<0.001$).

Fig. S6 The scatter plot, histogram, and correlation between environments for DTF trait (A), HT trait (B) and YLD trait (C). The names of environments are abbreviated with the format of country-location-year and season. At BD (Bangladesh), the trial environments were conducted at GZ (Gazipur), and NM (Nizmawna). At IN (India) conducted at HY (Hyderabad), CU (Cuttack), MA (Maruteru), and RP (Raipur). At KE (Kenya) conducted at AH (Ahero), and MW (Mwea). At MZ (Mozambique) conducted at CK (Chokwe), and MP (Maputo). At PH (Philippines) conducted at LB (Los Banos). At TZ (Tanzania) conducted at DK (Dakawa). 18W: Wet season in 2018. 19D: Dry season in 2019. 19W: Wet season in 2019. 20W: Wet season in 2020. *, **, ***: statistically significant at 5% ($p<0.05$), 1% ($p<0.01$), and 0.1% ($p<0.001$).

Fig. S7 Comparison of predictive abilities between the three cross-validation scenarios. (A) models with genotypes and environmental covariates main effects (GW) and their interactions (GW-GxW); (B) models combined all main effects (GEW) and their interactions (GEW-GxE-GxW).

Fig. S1 Physical positions of the 882 SNPs on the 12 chromosomes of rice using Nipponbare reference genome.

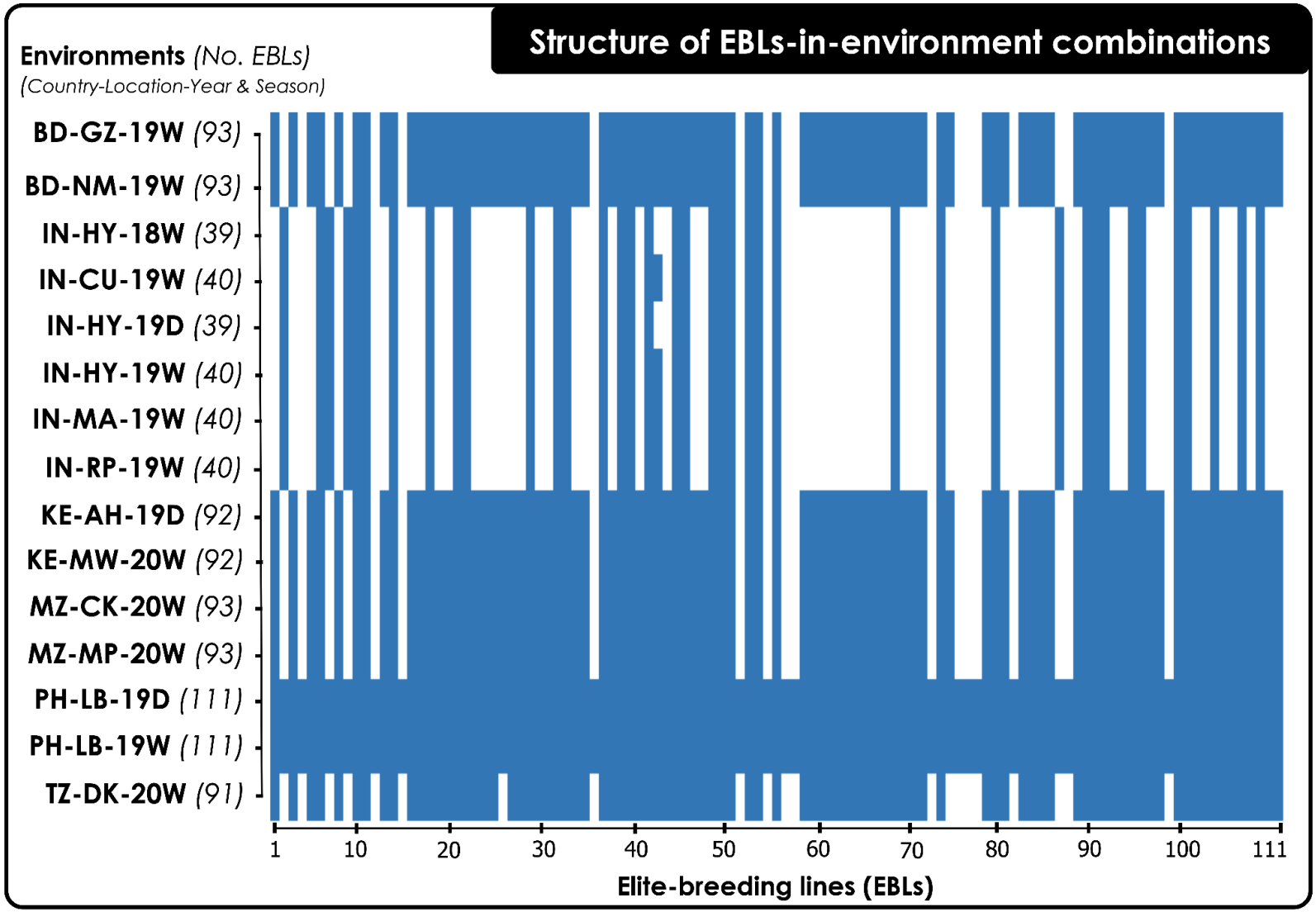
Structure of EBLs-in-environment combinations
Environments (No. EBLs)
(Country-Location-Year & Season)
BD-GZ-19W (93)
BD-NM-19W (93)
IN-HY-18W (39)
IN-CU-19W (40)
IN-HY-19D (39)
IN-HY-19W (40)
IN-MA-19W (40)
IN-RP-19W (40)
KE-AH-19D (92)
KE-MW-20W (92)
MZ-CK-20W (93)
MZ-MP-20W (93)
PH-LB-19D (111)
PH-LB-19W (111)
TZ-DK-20W (91)
1
10
20
30
40
50
60
70
80
90
100
111
Elite-breeding lines (EBLs)

Fig. S2 Graphical representation of the allocation of ECP across 15 environments. Blue vertical lines represent the ECP-environment combinations that were observed, and the white lines (blanks) correspond to unobserved combinations.

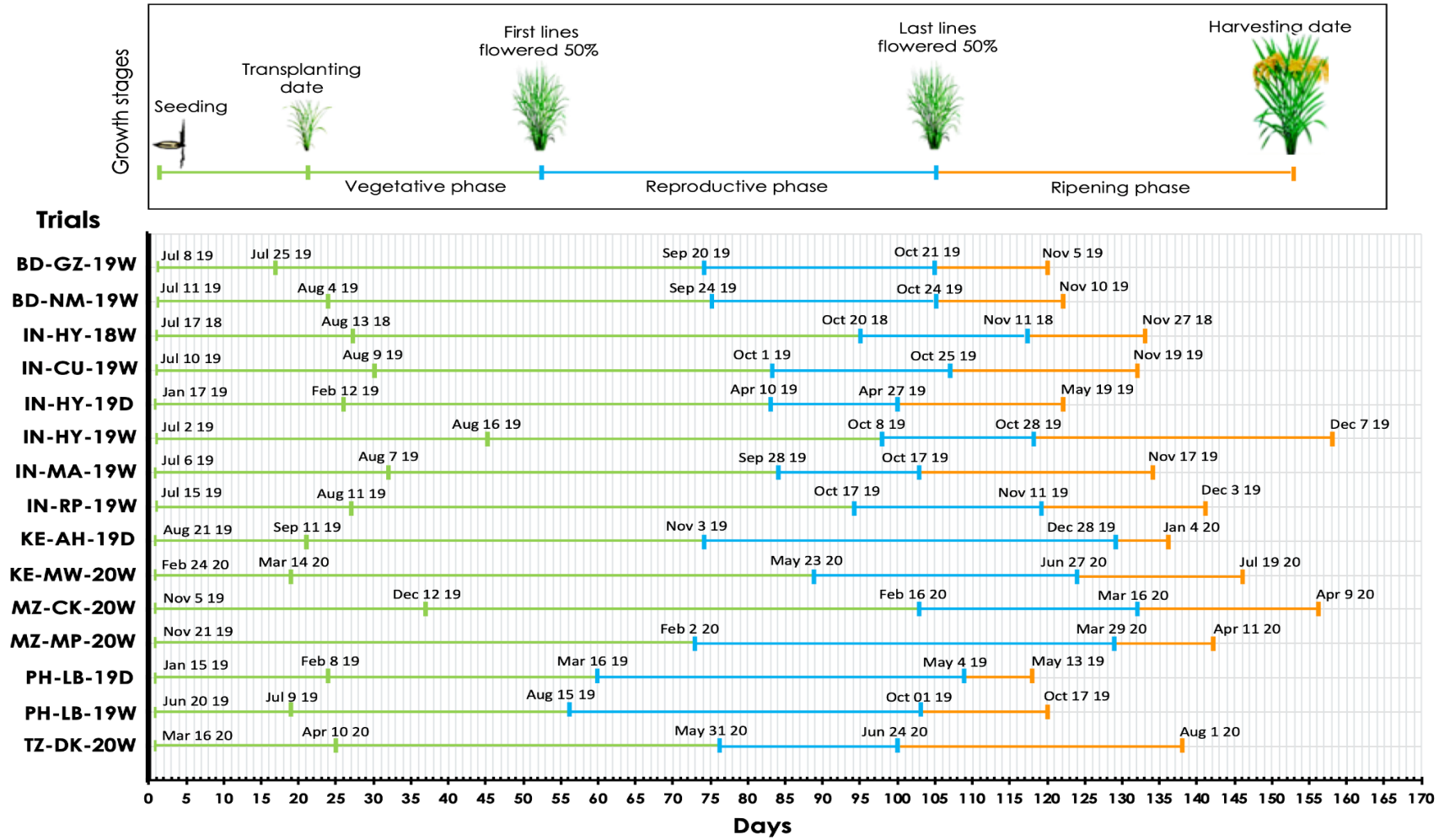
Growth stages
Seeding
Transplanting date
First lines flowered 50%
Last lines flowered 50%
Harvesting date
Vegetative phase
Reproductive phase
Ripening phase
Trials
BD-GZ-19W
Jul 8 19
Jul 25 19
Sep 20 19
Oct 21 19
Nov 5 19
BD-NM-19W
Jul 11 19
Aug 4 19
Sep 24 19
Oct 24 19
Nov 10 19
IN-HY-18W
Jul 17 18
Aug 13 18
Oct 20 18
Nov 11 18
Nov 27 18
IN-CU-19W
Jul 10 19
Aug 9 19
Oct 1 19
Oct 25 19
Nov 19 19
IN-HY-19D
Jan 17 19
Feb 12 19
Apr 10 19
Apr 27 19
May 19 19
IN-HY-19W
Jul 2 19
Aug 16 19
Oct 8 19
Oct 28 19
Dec 7 19
IN-MA-19W
Jul 6 19
Aug 7 19
Sep 28 19
Oct 17 19
Nov 17 19
IN-RP-19W
Jul 15 19
Aug 11 19
Oct 17 19
Nov 11 19
Dec 3 19
KE-AH-19D
Aug 21 19
Sep 11 19
Nov 3 19
Dec 28 19
Jan 4 20
KE-MW-20W
Feb 24 20
Mar 14 20
May 23 20
Jun 27 20
Jul 19 20
MZ-CK-20W
Nov 5 19
Dec 12 19
Feb 16 20
Mar 16 20
Apr 9 20
MZ-MP-20W
Nov 21 19
Feb 2 20
Mar 29 20
Apr 11 20
PH-LB-19D
Jan 15 19
Feb 8 19
Mar 16 19
May 4 19
May 13 19
PH-LB-19W
Jun 20 19
Jul 9 19
Aug 15 19
Oct 01 19
Oct 17 19
TZ-DK-20W
Mar 16 20
Apr 10 20
May 31 20
Jun 24 20
Aug 1 20
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170
Days

Fig. S3 Schematic representation of the developmental stages of ECP in each environment. The main three developmental stages of each environment were identified in the following: the vegetative stage: the interval time from the transplanting date up to the date of 50% flowering of the first ECP; the flowering stage: being the interval time between the first ECP and the last ECP are flowered of 50%; maturity stage: the interval time from the date of 50% flowering of the last ECP up to harvesting.

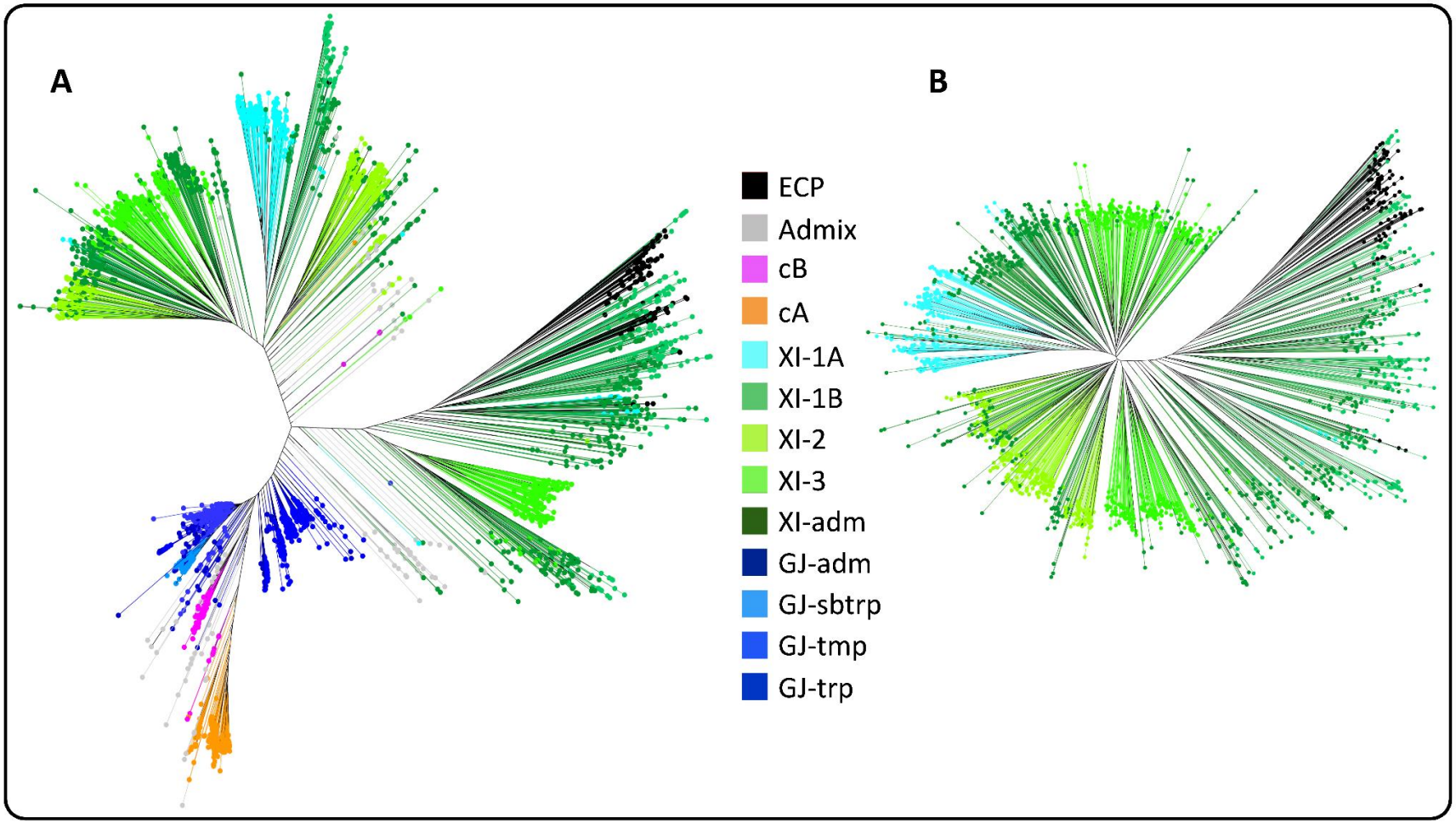

Fig. S4 Unweighted neighbor-joining trees between the ECP and *Oryza sativa* groups from 3K-RG (A), or the *indica* subgroups (B). The NJ trees were constructed using 837 common SNPs by combining the SNPs data of 3K-RG (Wang *et al.*, 2018) and the ECP from the 1K-RiCA assay.

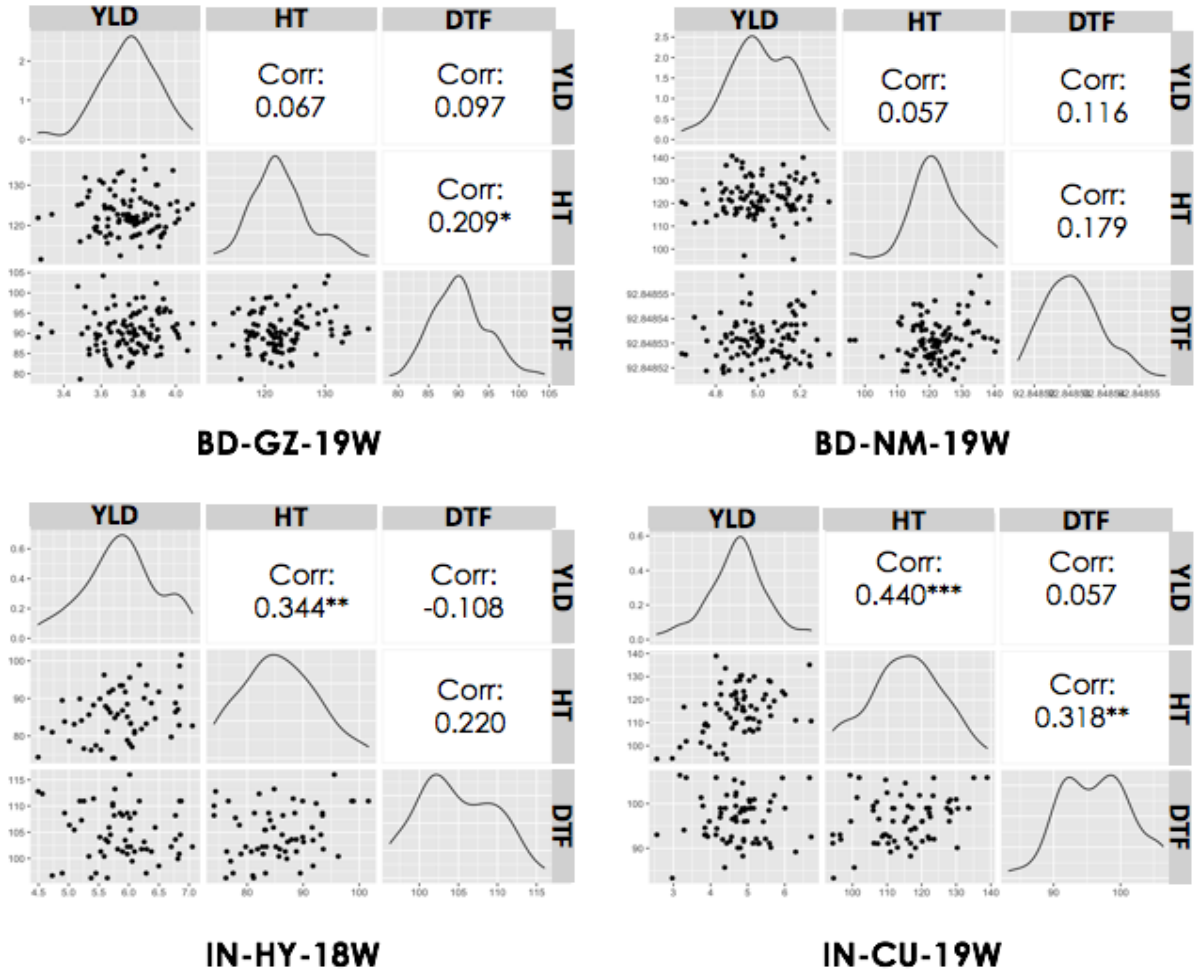

Fig. S5 The scatter plot, histogram, and correlation between the DTF, HT, and YLD traits in the single environments. The names of environments are abbreviated with the format of country-location-year and season. At BD (Bangladesh), the trial environments were conducted at GZ (Gazipur), and NM (Nizmawna). At IN (India) conducted at HY (Hyderabad), CU (Cuttack), MA (Maruteru), and RP (Raipur). At KE (Kenya) conducted at AH (Ahero), and MW (Mwea). At MZ (Mozambique) conducted at CK (Chokwe), and MP (Maputo). At PH (Philippines) conducted at LB (Los Banos). At TZ (Tanzania) conducted at DK (Dakawa). 18W: Wet season in 2018. 19D: Dry season in 2019. 19W: Wet season in 2019. 20W: Wet season in 2020. *, **, ***: statistically significant at 5% ($p<0.05$), 1% ($p<0.01$), and 0.1% ($p<0.001$).

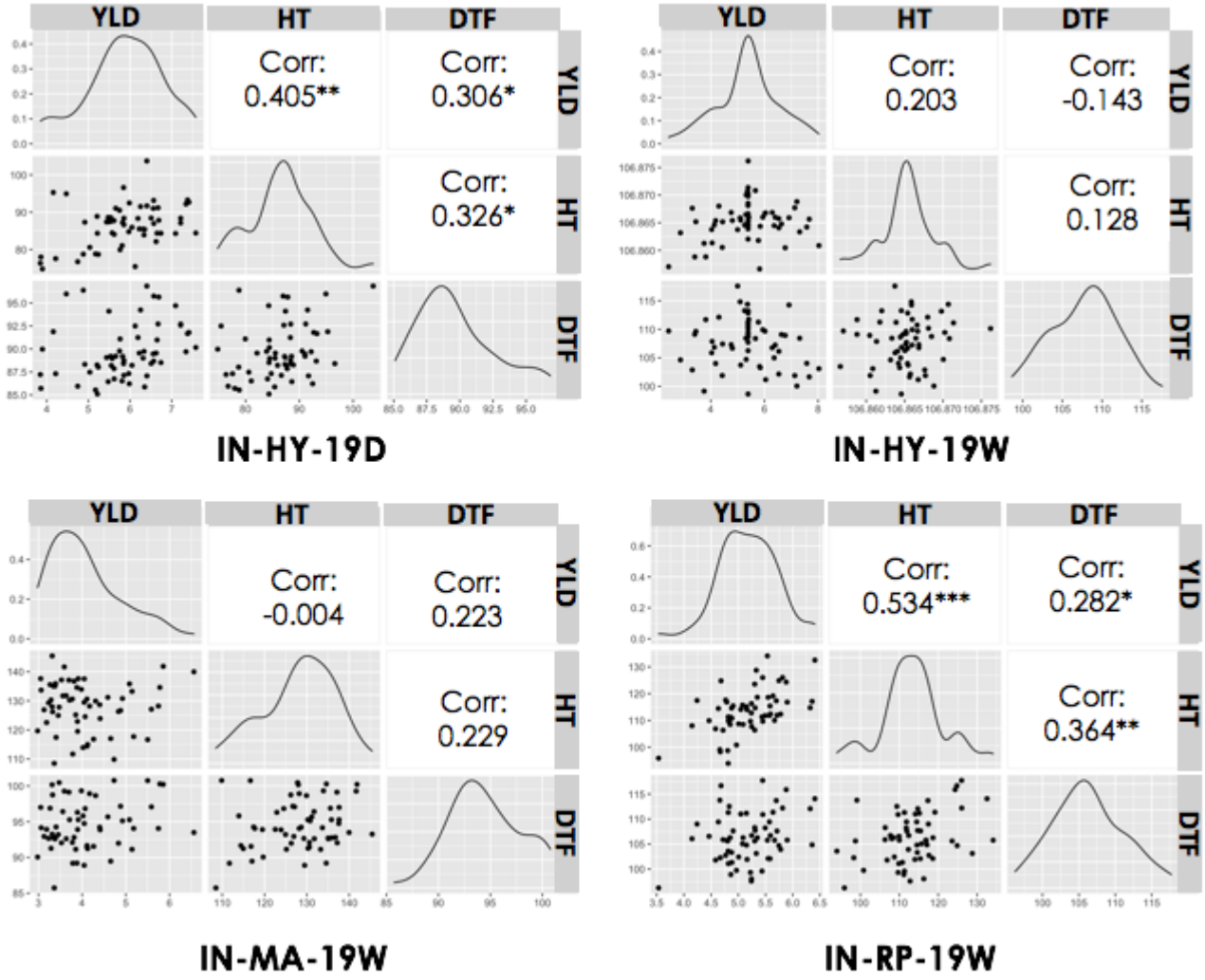

Fig. S5 Continued.

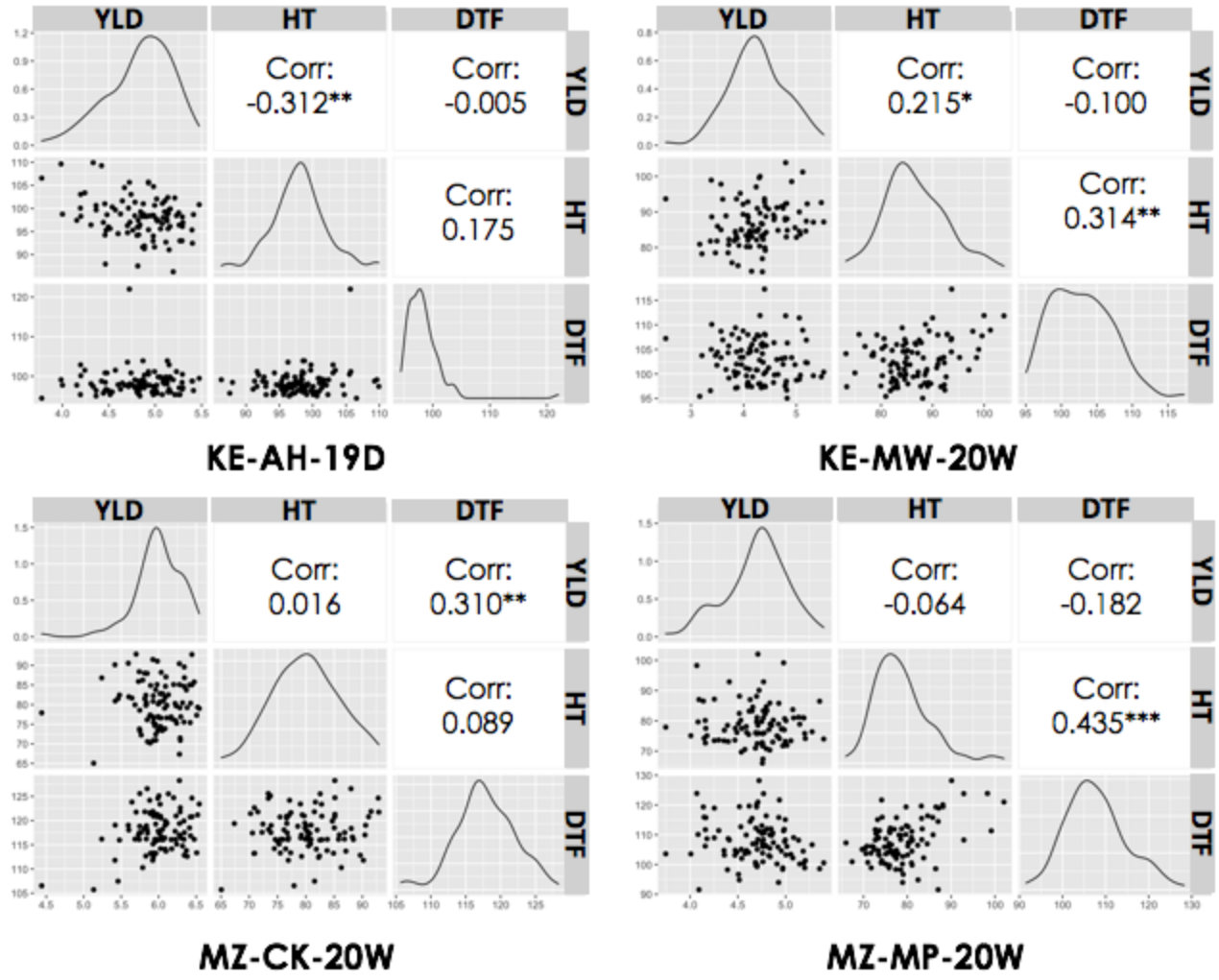

Fig. S5 Continued.

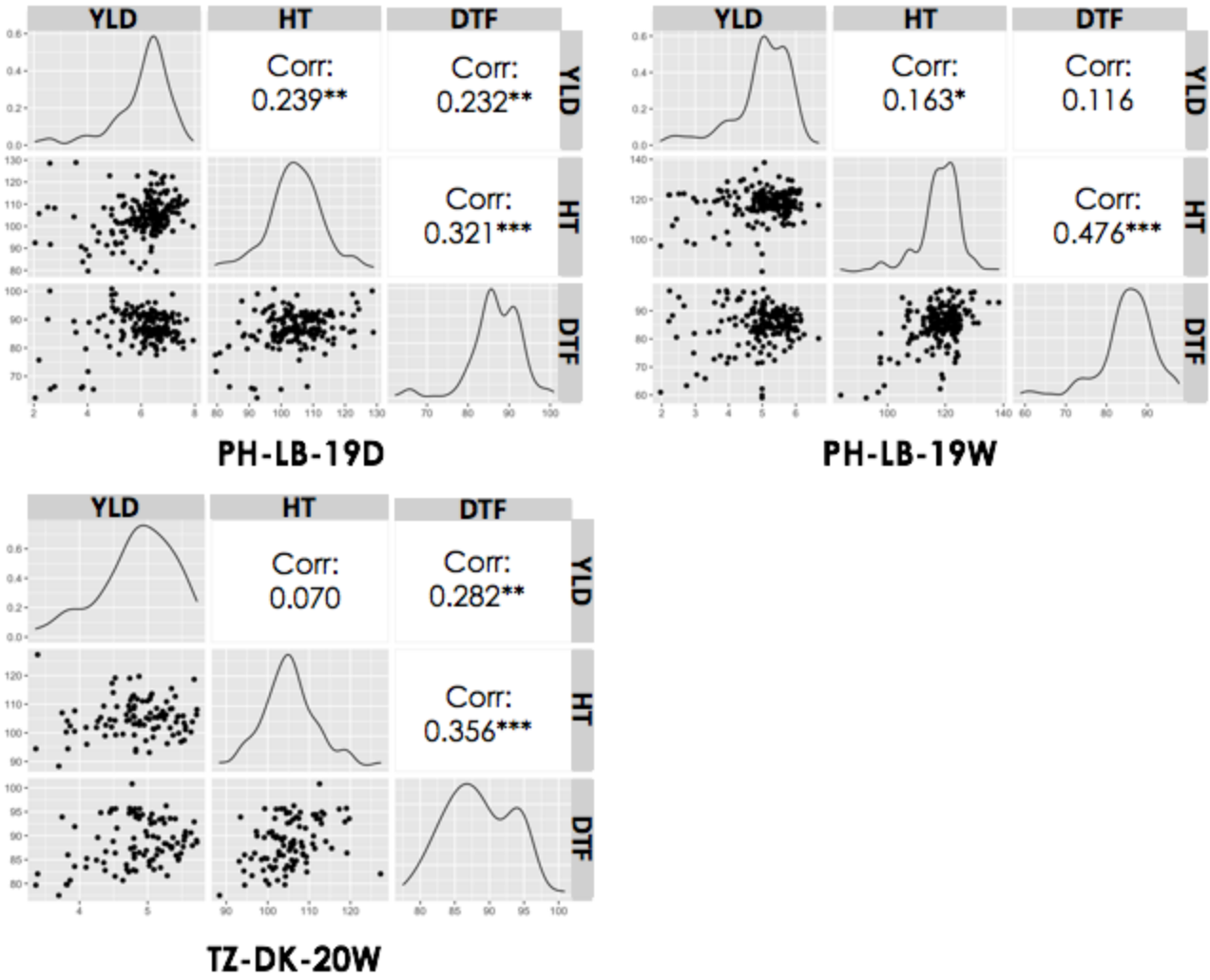

Fig. S5 Continued.

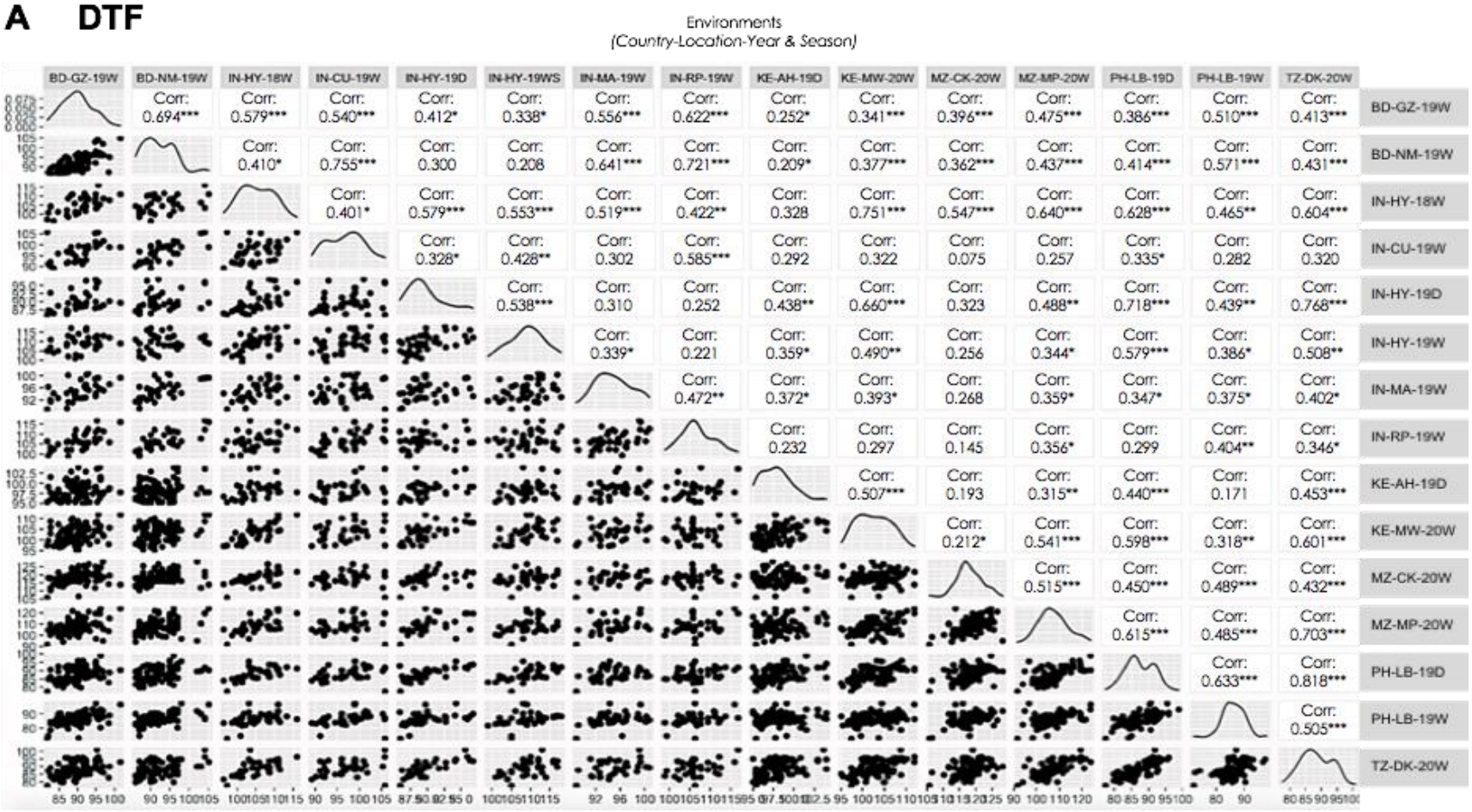

Fig. S6 The scatter plot, histogram, and correlation between environments for DTF trait (A), HT trait (B) and YLD trait (C). The names of environments are abbreviated with the format of country-location-year and season. At BD (Bangladesh), the trial environments were conducted at GZ (Gazipur), and NM (Nizmawna). At IN (India) conducted at HY (Hyderabad), CU (Cuttack), MA (Maruteru), and RP (Raipur). At KE (Kenya) conducted at AH (Ahero), and MW (Mwea). At MZ (Mozambique) conducted at CK (Chokwe), and MP (Maputo). At PH (Philippines) conducted at LB (Los Banos). At TZ (Tanzania) conducted at DK (Dakawa). 18W: Wet season in 2018.

19D: Dry season in 2019. 19W: Wet season in 2019. 20W: Wet season in 2020. *, **, ***: statistically significant at 5% (p<0.05), 1% (p<0.01), and 0.1% (p<0.001).

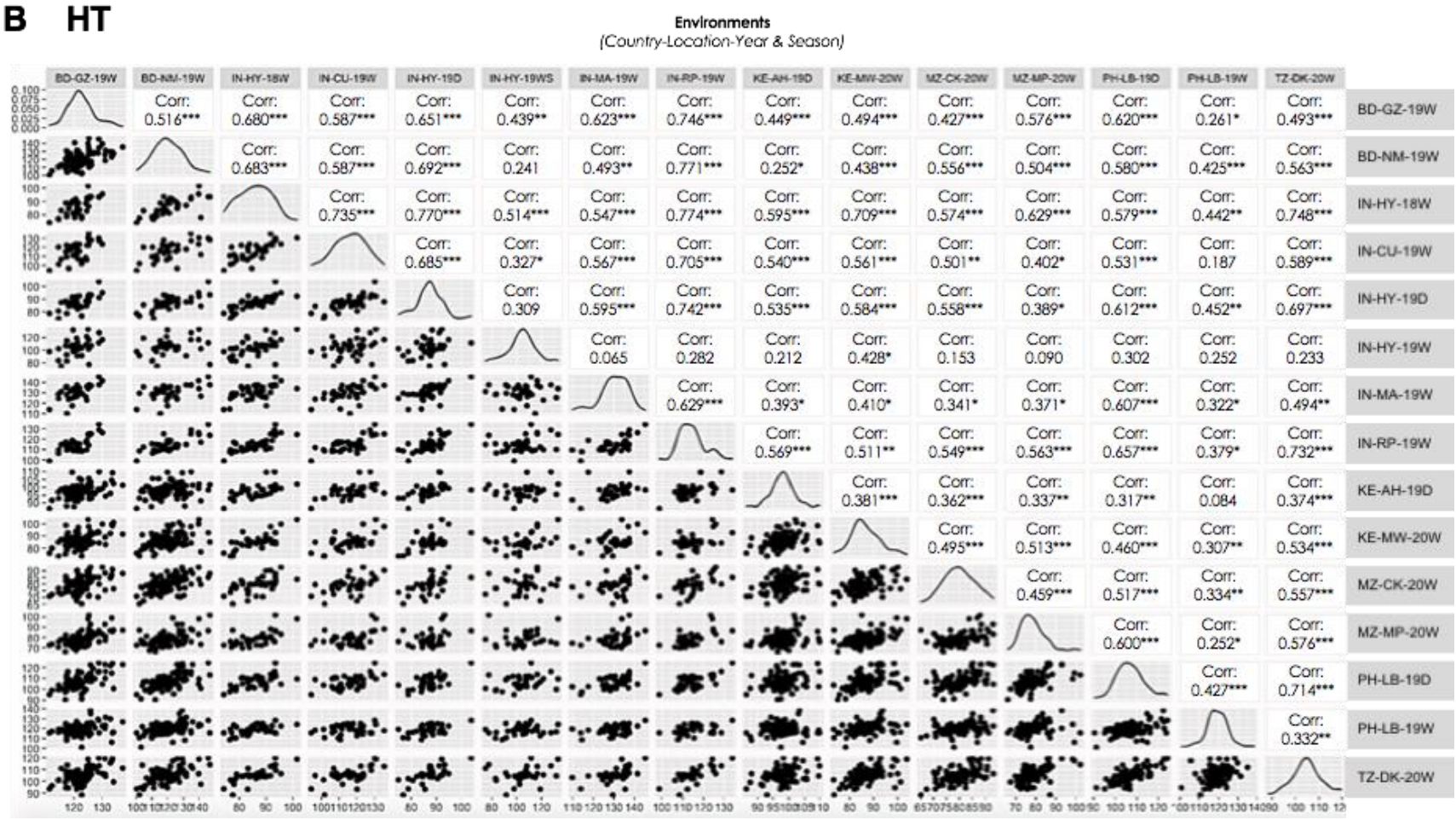

Fig. S6 Continued.

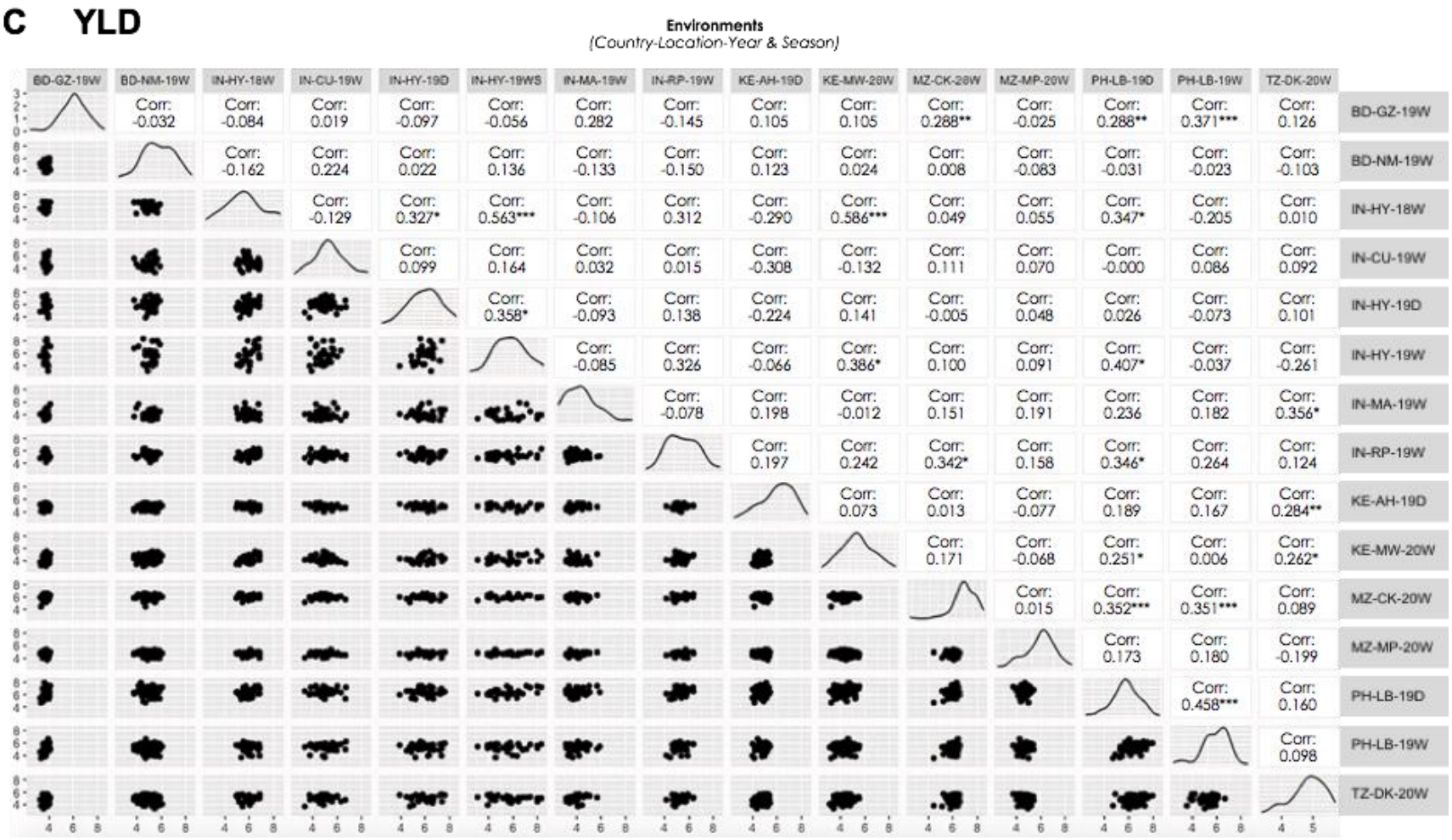

Fig. S6 Continued.

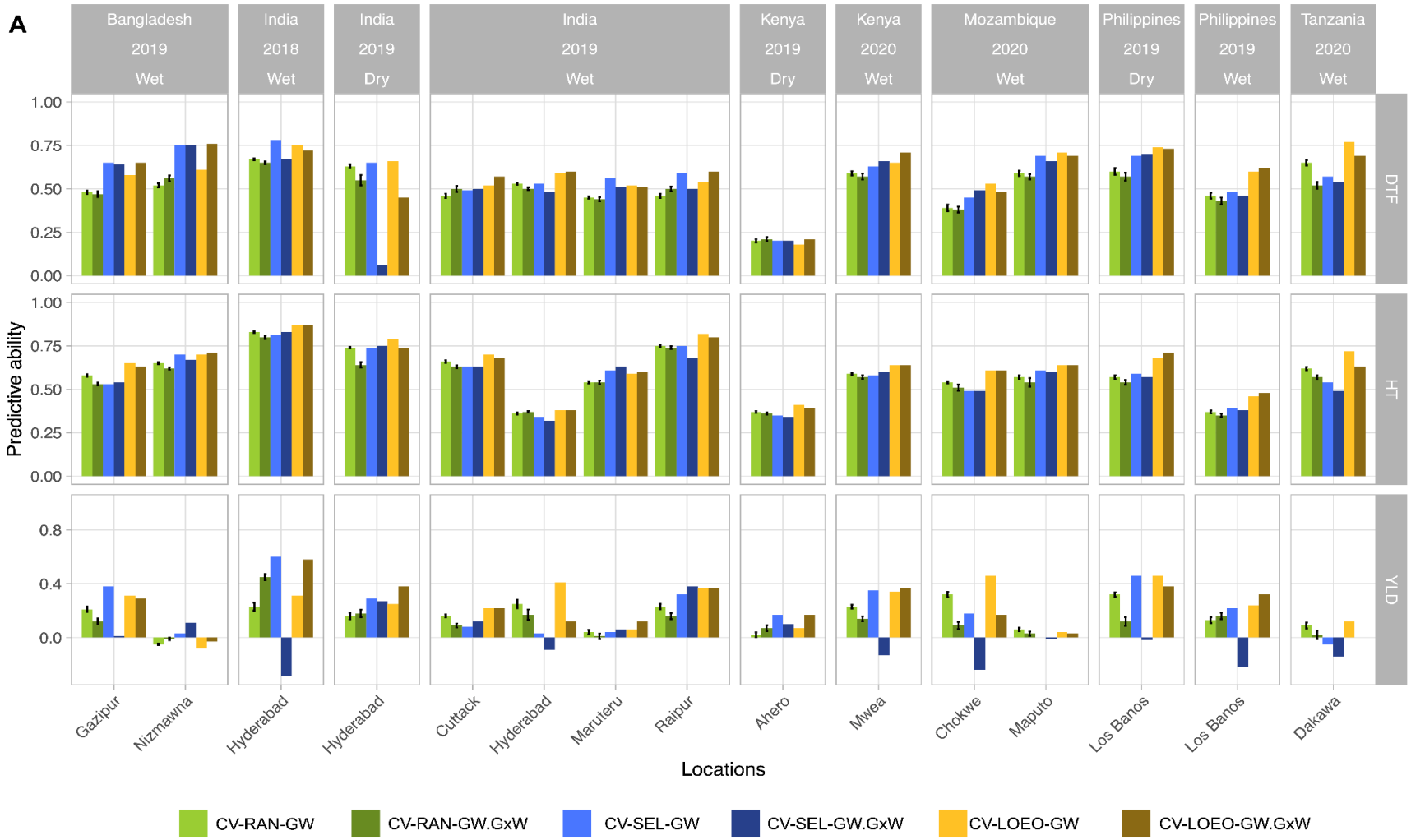

Fig. S7 Comparison of predictive abilities between the three cross-validation scenarios. (A) models with genotypes and environmental covariates main effects (GW) and their interactions (GW-GxW); (B) models combined all main effects (GEW) and their interactions (GEW-GxE-GxW).

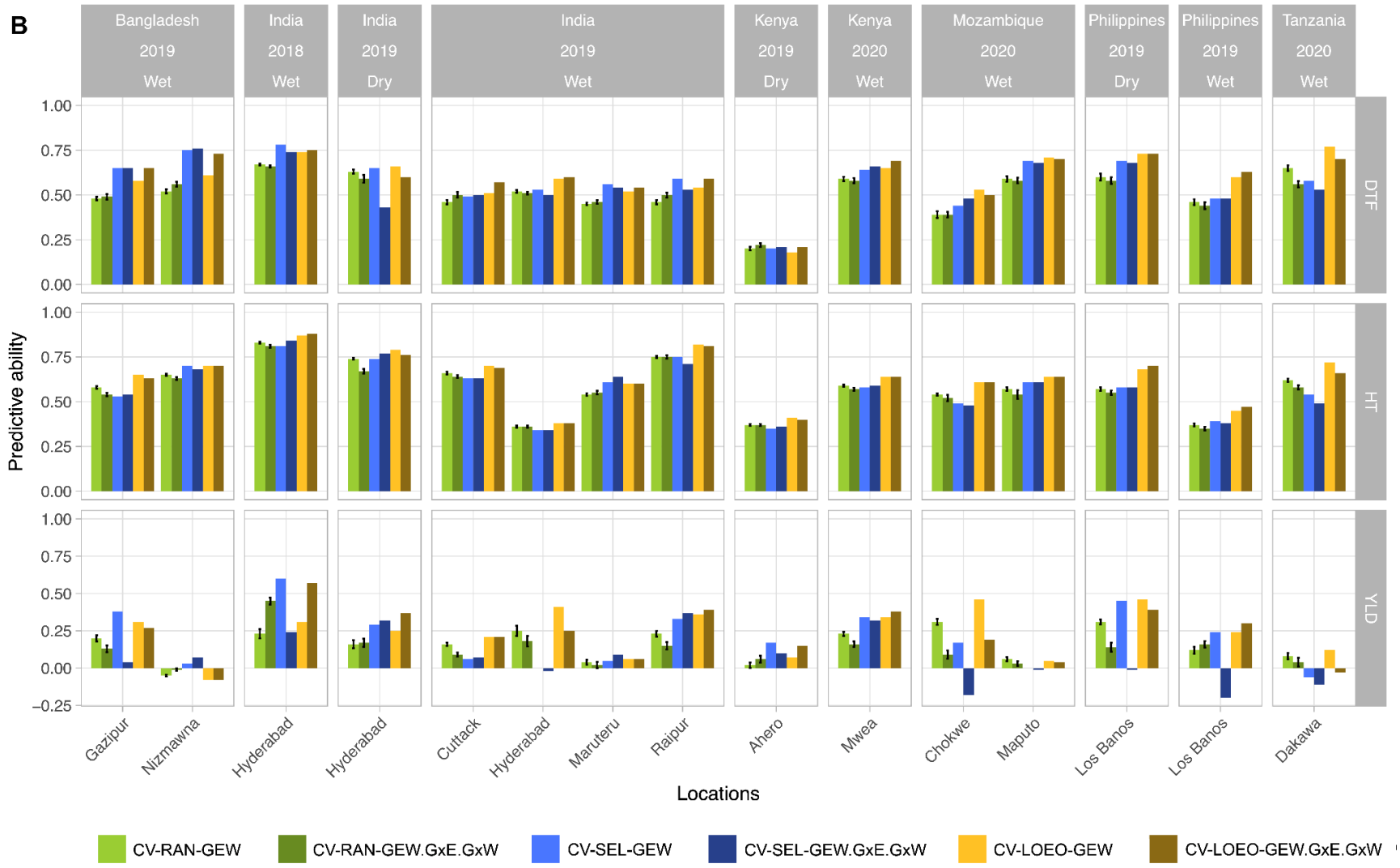

Fig. S7 Continued.