## Plant material and genotypic data

The study used data evaluated for a diverse *indica-aus* rice panel of 280 accessions, of which 245 represent the four major genetic subgroups belonging to *indica* genetic background and 35 to *aus* genetic background (Additional file 1: Table S4). They were selected from the 3000 accessions recently re-sequenced within the framework of the Rice Genome Project (Li et al. 2014), for their potential to breeding programs targeting rainfed lowland and upland drought environments in South and South-East Asia. In the selected panel, 215 accessions are landraces originating mainly from Asia and 65 accessions are improved lines. Seeds of the accessions were obtained from the IRRI gene bank.

The genotypic data for the 280 accessions were obtained from the International Rice Informatics Consortium (IRIC) database for the 3,000 rice genomes project (http://oryzasnp.org/iric-portal). The raw genotypic data extracted from the database contained 962k SNPs. The filtering for missing data (≤ 20%), minor allele frequency (MAF) ≥ 2% and rate of heterozygosity (Ho) ≤ 5% led to a working set of 215,250 SNPs, referred to as 215k set. This panel and the associated genotypic data were previously described in Bhandari et al. (2019).

## Phenotyping of population

### Experimental design and crop management

Six experiments (Additional file 1: Table S1) were conducted in the 2014 wet season (WS) and 2015 dry season (DS) at IRRI (14.18°N, 121.25°E), Philippines. In each season, the experiment was conducted under control conditions or non-stress experiment (LL\_N) in lowland (under flooded, puddled, transplanted and anaerobic conditions) while the reproductive-stage drought-stress experiments were conducted in lowland and upland (under direct-sown, non-puddled, non-flooded and aerobic conditions in leveled fields) environments, referred as LL\_S and UL\_S, respectively. The LL\_N experiments were established in augmented randomized complete block design in single-row plots with 5 m row length. The LL\_S and UL\_S experiments were established in a α-lattice design with two replications in single or two-row plots with 5 m row length in lowland and 2–3 m row length in upland. The crop management practices were as described in Kumar et al. (2014).

### Drought application procedure

RS-drought phenotyping was as described in Kumar et al. (2014). Briefly, in the LL\_S experiments, the field was drained 30 days after transplantation and irrigation was withheld to impose the RS-drought stress. Stress was continued until severe leaf rolling was observed in at least 75% of the accessions and water table depth remained below 100cm for more than 2 weeks. Fields were thereafter re-irrigated (flash-flooding -WS and sprinklers - DS) and the water was drained after 24 hours to impose a subsequent cycle of drought stress. This cyclic pattern was implemented until harvest. In the UL\_S experiments, where the crop was established by direct-seeding, RS-drought stress was initiated 45 days after sowing, by withholding sprinkler irrigation until the soil water tension fell below –50 kPa at 30 cm depth. Thereafter, sprinkler-irrigation and subsequent drainage after 24 hours for the imposition of drought stress were done in a cyclic pattern till harvest.

### Traits measured

For each experiment, days to 50% flowering (DTF, in days), plant height (PH, in cm, the average for 3 measurements per plot), panicle length (PL, in cm, the average for 3 measurements per plot), flag leaf area (FlgLA, in cm2, the average for 3 measurements per plot), dry biomass at maturity (BMDW, in kg ha-1), number of effective panicles (NBP), grain yield (GY, in kg ha-1), 1000-grain weight (TGW, in g) and spikelet fertility (SPKFT, in percentage) were measured in individual plots and harvest index (HI) was calculated as GY/BMDW. Details of measurement procedures of each trait are given in Additional file 1: Table S5.

### Analysis of phenotypic data for each trait

For each trait from each of the six experiments, best linear unbiased predictors (BLUP) were estimated using the restriction maximum likelihood method (REML) in the PROC MIXED procedure of SAS v9.0 (SAS Institute Inc., 2002). Within a season, the performance of a genotype was modeled as *Yij = μ + ßi + cj + αi + εij* for augmented randomized complete block design where Yij is the phenotype of the ith genotype in jth block, μ the overall mean, ßi the block effect which was considered as random, cj the checks effect in jth block which was considered as fixed, αi the random effect of the ith genotype and εij is the residual considered as a random effect. We constructed two variables- “checks” and “genotypes” variables in both WS and DS. Checks refer to the control genotypes included additionally in the experiment to compare the performance of genotypes being tested and were used to recover the block effects. For α-lattice design, genotype performance was modeled as *Yijk = μ + αi + rj + bkj + εijk* where Yijk is the phenotype of the ith genotype in kth block of jth replicate, μ the overall mean, αi is the genotype effect considered as random, rj is the replicate effect considered as fixed, bkj is the random effect of the kth block within jth replicate and εijk is the residual considered as a random effect.

The variance components were estimated using the REML method to extract Yadj (*μ* + Yij(k)) values for each genotype which were used in GWAS for analysis at both individual experiment level and combined analysis for each environment- lowland non-stress, lowland stress, and upland stress, to detect genomic regions associated with traits of interest. For each of the studied trait, the broad-sense heritability was estimated using the formula

H² = σ2g / σ2p

where σ²g is the genotypic variance obtained from the experimental data (assuming only additive genetic variance among accessions) and the phenotypic variance is σ2p= σ2g + σ2e/r, where σ2e is the residual variance obtained from the ANOVA and r is the number of replication.

The corrplot package in R (R. v.1.2.5001) was used to estimate the correlation among the measured traits.

The drought susceptibility index (DSI) was calculated for DTF, PH, BMDW, GY, TGW and SPKFT. Drought intensity (DI) was calculated according to (Lazar et al. 1995) as follows-

DI = 1 – YD / YN

Where YD is the average all genotypes for a given trait under drought stress, while, the YN is the average of all genotypes for the same trait under normal condition. The drought susceptibility index (DSI) was estimated for each genotype and calculated according to (Lazar et al. 1995) as follows-

 

Where XD is the mean performance of each genotype for a given trait under drought stress, while, the XN is the mean performance of each genotype for the same trait under normal condition.

## Methods for characterizing the population

### Experimental evaluation

Multi-dimensional analysis of the phenotypic data by FDA was performed on phenotypic data (280 accessions x 10 trait variables x 6 experiments) to estimate the pairwise Fisher distance between the experiments using the XLSTAT package [Internet] 2012. (<http://www.xlstat.com/en/products-solutions/pro.html>). Using mean grain yield as criterion, each experiment was re-classified based on the grain yield reduction compared to the control-lowland-non-stress experiment (Kumar et al. 2009) (Additional file 1: Table S1).

### Genetic structure

The genetic diversity among the 280 accessions was studied with the working set of 215k markers using the Neighbor-joining (NJ) clustering method in TASSEL 5 (Bradbury et al. 2007) and visualization using FigTree v1.4.3 (Rambaut and Drummond 2016). The population structure was assessed using ADMIXTURE v.1.3.0 (Alexander et al. 2009) and results visualized using R/pophelper (Francis 2017) package for 280 accessions and 215k SNPs. Series of models for K value ranging from 2-8 were run with 5 fold cross-validation to prime the main algorithm- QuasiNewton for convergence acceleration. Accuracy and precision were ensured by performing 20 runs for each value of K and the optimal number of clusters was determined by the K value with the least cross-validation (CV) error. Principal components (PC) explaining genetic variation were estimated using R/GAPIT 3.0 package (Lipka et al. 2012). The estimated population structure covariates (Q) and kinship matrix (K) were used to improve the statistical power of the GWAS models used.

### Pairwise linkage disequilibrium (LD)

LD between SNP loci at the individual chromosomal level was calculated and plotted by computing r² estimators between all pairs of SNP markers using the PopLDdecay (Zhang et al. 2019).

## Methods for identifying associations at the population level

In our study, we implemented GWAS with MLM, SUPER and Farm-CPU methods using R/GAPIT 3.0 package and visualization of circular manhattan and qq plots using rMVP package (0.99.17) (<https://github.com/xiaolei-lab/rMVP>). The false positives in GWAS study were corrected using “Bonferroni Correction” factor. Using the Bonferroni multiple test correction (0.05/215,250; at 5% level of significance), the calculated threshold value was 2.32 × 10−7. Only the MTAs that exceed the threshold value and which were consistent across multi-locus methods- SUPER and Farm-CPU methods have been reported in this study. To detect seasonal variations, we explored two p-value thresholds (1e-6 and 1e-4).

The percent phenotypic variance (PV) explained by all significant SNPs detected in each environment and season was output from all models used in the study. PV explained by each significant SNP was calculated as the squared correlation between the phenotype and genotype of the SNP.

## Candidate genes discovery

The candidate genes were searched within the 200-kb region around (100 kb upstream and 100 kb downstream) the detected significant SNP. The literature searches were also performed using QTARO and MSU databases (<http://qtaro.abr.affrc.go.jp> and [http://rice.plantbiology.msu.edu](http://rice.plantbiology.msu.edu/)) to identify the earlier reported QTLs present in the LD region.

## Selection of accessions as potential donors in breeding programs

Promising accessions were selected from the population based on yield advantage over non-stress condition in WS for both lowland and upland stress environments and over checks in each environment in DS. The premise was to identify a set of accessions that can be incorporated in breeding programs for drought tolerance under both lowland and upland environments with the advantage of early flowering and short plant type under RS drought.

These selected accessions were analyzed for allelic effect using 101 (on 94 unique loci with 4 having colocalisation for multiple traits) significant MTAs validated from database and earlier reported literature for grain yield QTLs. Allelic variation was studied for effect of allelic contribution to trait mean for DTF, PH and GY under LL\_S and UL\_S in both seasons. Five classes of loci were established – three based on presence of major allele in all seven accessions contributing to phenotypic performance for tolerance under LL\_S + UL\_S (class I abbreviated as cl-I); under UL\_S only (class II abbreviated as cl-II) and under LL\_S only (class III abbreviated as cl-III) while the fourth class (cl-IV) contained loci with minor allele associated to phenotypic performance for tolerance under both LL\_S + UL\_S. The fifth class (cl-V) consisted of loci with neither the major nor minor allele associated to phenotypic performance for tolerance under RS drought in the selected accessions. Further, validation of phenotypic-based selection of each accession was done by computing the percentage composition of favorable alleles in the set of 94 loci.