Oryza CLIMtools: An Online Portal for Investigating Genome-Environment Associations in Rice

Running head: Oryza CLIMtools

Ángel Ferrero-Serrano[1*], David Chakravorty[1], Kobie J. Kirven[2] & Sarah M. Assmann[1*]

[1]Biology Department, Pennsylvania State University, 208 Mueller Laboratory, University Park, PA, 16802, USA
[2]Intercollege Graduate Degree Program in Bioinformatics and Genomics, Pennsylvania State University

*Authors for correspondence:
Ángel Ferrero-Serrano, aaf11@psu.edu, 814-863-9578
Sarah M. Assmann, sma3@psu.edu, 814-863-9579

Abstract

Elite crop varieties display an evident mismatch between their current distributions and the suitability of the local climate for their productivity. To this end, we present Oryza CLIMtools, (https://gramene.org/CLIMtools/oryza_v1.0/) the first pan-genome prediction of climate-associated genetic variants in a crop species. This resource consists of interactive web-based databases that allow the user to: i) explore the local environment and its interaction with natural existing genetic variation in local rice varieties (landraces) in South-Eastern Asia, and; ii) investigate the environment × genome associations for 658 Indica and 283 Japonica rice landrace accessions included in the 3K Rice Genomes Project and previously collected from their geo-referenced local environments. We exemplify the value of these resources, identifying an interplay between flowering time and temperature in the local environment that is facilitated by adaptive natural variation in *OsHD2* and disrupted by maladaptive variation in *OsSOC1*. Prior QTL analysis has suggested the importance of heterotrimeric G proteins in the control of agronomic traits. Accordingly, we analyzed the climate associations of the different heterotrimeric G protein subunits. We identified a coordinated role of G proteins in adaptation to the prevailing Potential Evapotranspiration gradient and their regulation of key agronomic traits including plant height, seed, and panicle length. We conclude by highlighting the prospect of targeting heterotrimeric G proteins to produce crops that are climate-change-ready.

Keywords: *Oryza sativa*, rice, G x E, adaptation, flowering time, GWAS, *OsHD2, OsSOC1,* heterotrimeric G protein*, OsRGA1, d1, OsGS3, OsDEP1, OsGGC2*.

Introduction

Rice is arguably the most critical food source on the planet, feeding over half of the world's population (Bin Rahman and Zhang, 2023). The Green Revolution was crucial to promoting food security worldwide in the late twentieth century despite the concurrent exponential growth of the world's population (Dalrymple, 1986). As farmers began to prioritize high-yielding crops, they often abandoned local crop varieties (Frankel, 1974), resulting in the loss of genetic diversity over time that characterizes modern rice varieties (Khoury et al., 2022). Selective breeding practices in crops have primarily prioritized high-yielding but genetically uniform varieties, which can reduce adaptive potential (Østerberg et al., 2017). Therefore, as crop diversity decreases, so does their resilience to extreme weather events, such as heatwaves, droughts, and floods, which are becoming more frequent due to the already noticeable effects of climate change and are predicted to decrease crop yields globally (Food and Agriculture Organization of the United Nations et al., 2018). To promote global food security, crop production can be stabilized by increasing genetic diversity (Renard and Tilman, 2019). A promising approach is to

restore to modern elite varieties the lost genetic diversity that is present in local varieties (landraces), adapted to the local climate, and thereby foster improved yield stability in diverse conditions.

The 3,000 (3K) Rice Genome Project provides a resource for high-quality sequencing data for 3,010 rice genomes, obtained using Illumina-based Next-Generation sequencing (3,000 rice genomes project, 2014; Wang et al., 2018). Indica and Japonica are the two main subspecies of Asian cultivated rice and are the products of two separate domestication events from the ancestral species *Oryza rufipogon* (Kato, 1930; Sweeney and McCouch, 2007). The 3K dataset provides information on the genetic diversity of 658 Indica and 283 Japonica landraces grown in South-Eastern Asia (Figure 1), for which their collection sites are known and their population structure and history have been described (Gutaker et al., 2020). By studying local landraces, researchers can explore the factors that have shaped rice diversity and identify genetic variants and traits associated with the local environment that are suggestive of local adaptation. This information can aid the development of new high-yielding rice varieties better adapted to anthropogenically altered climate change, to improve food security and sustainability.

Our recent work in locally-adapted accessions of *Arabidopsis thaliana,* collected from their local habitats, identified genome-wide associations between their natural genetic variation and a myriad of geo-environmental variables that define each local habitat. Arabidopsis CLIMtools provides a resource for the community to explore these associations (Ferrero-Serrano and Assmann, 2019; Ferrero-Serrano et al., 2022). However, similar resources have not been available for any crop species. Here, we present the first resource to provide gene x phenotype x environment associations to assist rice research and breeding programs. We integrate datasets comprising 413 geo-environmental variables that define the local environmental conditions at the collection sites of these 941 landraces and make this information available in OryzaCLIM (https://gramene.org/CLIMtools/oryza_v1.0/OryzaCLIM/). Using this environmental dataset, we establish an adaptive relationship between flowering time and the minimum temperature of the coldest month in Indica and Japonica landraces. Additionally, we present an interactive dataset of genotype–environment associations for 658 Indica and 283 Japonica rice landraces, with detailed information on the genes and variants statistically associated with the 413 environmental variables. These results are publicly available in a database platform, Oryza CLIMtools, available in the Gramene portal (Gramene.org; Tello-Ruiz et al., 2021). Oryza CLIMtools allows users to identify candidate adaptive genetic variants related to a given environmental variable of interest as codified in Oryza CLIMGeno and, conversely, to identify environmental variables in Oryza GenoCLIM associated with variation at any genetic locus of interest. We demonstrate the use of these tools, revealing an interplay among plant height, panicle length, and seed length that is associated with the Potential Evapotranspiration (PET) in the local environment and depends on both natural variation in the heterotrimeric G

protein Gα subunit gene, *OsRGA1*, and a SNP that causes a stop codon in the Gγ gene, *OsGS3*.

Oryza CLIMtools has potential to assist the scientific community in ways that include:

- Inform in the design of GWAS analysis for optimal selection of varieties across environmental gradients to uncover candidate adaptive variants.
- Assign relationships between phenotypic variability and environmental variation in this set of sequenced landraces.
- Explore genetic variation in a gene of interest and its association with environmental gradients to inform future studies of the interaction between that gene and the environment.
- Explore genetic variation associated with a particular environmental gradient that, among other uses, may expedite forward genetics research.

Results

We created a comprehensive database of environmental data on rice landraces, to serve as a valuable resource for researchers aiming to enhance rice research by integrating environmental parameters. This suite of online resources facilitates the process of incorporating environmental data into studies, and provides a user-friendly interface for researchers to access and utilize this information. Oryza CLIMtools (Figure 1) comprises three web-based interactive databases, OryzaCLIM, Oryza GenoCLIM, and Oryza CLIMGeno, that allow the user to explore the local environment and its relationship with natural existing genetic variation in local landraces of rice collected from South-Eastern Asia.

Description of the local environment or rice landraces

Oryza CLIMtools is a new web resource that enables users to explore associations of promoter and transcript natural variants in rice landraces (Figure 1A) with 413 geo-environmental variables (Figure 1B, S1) that define the environmental conditions in the areas where they were collected. These environmental variables are collated in our OryzaCLIM tool and characterize the local environment of 658 Indica and 283 Japonica georeferenced and fully sequenced landrace varieties included in the 3K Rice Genome Project ((Gutaker et al., 2020) Figure 1A, Table S1 and S2)). With Oryza CLIMtools, users can easily search for climate-gene associations based on a specific rice gene of interest using Oryza GenoCLIM. Alternatively, users can also search for associations based on any climate parameter of interest using Oryza CLIMGeno (Figure 1C). Of particular note, we also provide riboSNitch prediction, which indicates the likelihood that a variant alters

RNA structure (Halvorsen et al., 2010). As shown by our analyses in Arabidopsis (Ferrero-Serrano et al., 2022), riboSNitches will be a fruitful topic for further investigation.

Natural variation in *OsHD2* and *OsSOC1* influences the flowering time x temperature relationship

To support the use of landraces to uncover the genetic basis of local adaptation, we asked whether there was an association between the phenotypic characteristics of these varieties and the environmental characteristics of the areas where they are grown. Using the environmental dataset we provided in OryzaCLIM, we uncovered a positive relationship in Indica and Japonica landrace varieties between flowering time as documented in publicly available data (Mansueto et al., 2017) and the minimum temperature of the coldest month (Karger et al., 2017) typical of their local environment (Figure 2). This relationship between flowering time and temperature is particularly strong in Japonica landraces (Figure 2F). To uncover the genetic basis of this adaptive trait, we used Oryza CLIMGeno and identified in these Japonica landraces a haplotype consisting of a set of one missense and six intronic covarying SNPs (Figure 2A) in *HEADING DATE 2* (*OsHD2*/*OsPRR37*/*Ghd7.1*/*DTH7*; *Os07g0695100*), and the minimum temperature of the coldest month (Figure 2B). Landraces that contain the version of the haplotype in lower frequency in the Japonica population (minor allele) are typically grown in colder locations and display early flowering. By contrast, late-flowering landraces with the haplotype version more frequent in the Japonica population (major allele) are typically from warmer areas (Figure 2B).

To uncover the genetic basis of flowering time adaptation to local temperature in Indica varieties, we used Oryza CLIMGeno and identified the association of an intronic variant in the floral pathway integrator and MADS domain transcription factor *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*OsSOC1/AGL20/OsMAD50*; *Os03g0122600*; Figure 2C) with the minimum temperature of the coldest month in Indica landraces (Figure 2D). The minor allele of *OsSOC1* is found in landraces that are typically grown in warmer areas and exhibit early flowering (Figure 2D). This relationship disrupts the overall positive correlation between flowering time and temperature in the local environment, and thereby appears maladaptive. From these findings, we hypothesize that the early flowering phenotype associated with the minor allele of *OsSOC1* may have been selected to favor the stability of flowering time and thus sustained yields across multiple locations. To test this hypothesis, we calculated the difference in flowering time for 576 Indica landraces with publicly available records resulting from two different planting locations (Figure 2E), one in Hainan (China) and the other in Los Baños (Philippines; Mansueto et al., 2017; Wang et al., 2020). The difference in flowering time between locations was calculated by subtracting the number of days it took landraces to flower in Hainan (China) from the days to flower in Los Baños

(Philippines). In support of the conclusion that later flowering is typical of landraces that originate from warmer climates, we found that most landraces flowered later at the warmer and lower latitude of Los Baños, relative to those in Hainan (Figure 2F). Additionally, we found that the flowering time of those landrace varieties harboring the minor *OsSOC1* variant remained remarkably stable when planted in these two different locations (Figure 2F), in agreement with our hypothesis that the minor *OsSOC1* allele may be maladaptive and yet selected to confer stability in flowering time.

Genetic variation in *OsRGA1*, the α subunit of the rice heterotrimeric G protein, is associated with a gradient in Potential Evapotranspiration

Potential Evapotranspiration (PET) is defined as a measure of the ability of the atmosphere to remove water through evapotranspiration given unlimited moisture (Hargreaves and Samani, 1982). Therefore, PET provides an estimate of the amount of water-use required for crop growth and can be used to help determine the appropriate amount of water that different varieties require to optimize productivity in a given location. Using bioclimatic variables derived from monthly precipitation and temperature records (Karger et al., 2017) that we collated in OryzaCLIM, we determined that for both Japonica and Indica landraces, PET during the coldest quarter was co-associated with other temperature variables. For example, areas with higher PET during the coldest quarter of the year are typically warmer areas with higher precipitation regimes (Figure 3A).

Heterotrimeric G proteins are signal transduction complexes that consist of Gα, Gβ, and Gγ subunits. In the classical model, upon activation of an associated G protein coupled receptor, the Gα subunit exchanges GDP for GTP, resulting in dissociation from the Gβγ dimer. Both Gα and Gβγ interact with intracellular effectors to evoke downstream signaling cascades. The system is returned to the inactive state when the intrinsic GTPase activity of Gα hydrolyzes GTP to GDP, thereby restoring the affinity of Gα for Gβγ and resulting in an inactive heterotrimeric complex (McCudden et al., 2005). The G protein heterotrimer has been proposed to play a major role in regulating agronomic traits and stress responses (Botella, 2012; Cui et al., 2020). Our previous work has specifically implicated the rice Gα subunit, OsRGA1, in a number of agronomically-relevant traits, including plant architecture, drought sensitivity, light use-efficiency, and mesophyll conductance (Ferrero-Serrano and Assmann, 2016; Ferrero-Serrano et al., 2018; Zait et al., 2021).

We, therefore, used Oryza GenoCLIM to explore possible associations between previously uncharacterized natural variation in *OsRGA1* and the local environment of rice landraces. In Japonica landraces, we identified a correlation between a haplotype, consisting of six covarying intronic SNPs and eight upstream SNPs in *OsRGA1* (*Os05g0333200*; Figure 3B), and PET during the coldest quarter (Figure 3C; Title and

Bemmels, 2018). The haplotype in lower frequency is found in association with areas with lower PET, while the haplotype in higher frequency is typically grown in areas with higher PET (Figure 3C).

To ascertain the ancestral origin of this GxE association that we find in cultivated rice landraces in a wild relative, we identified the genetic variation present in 446 imputed sequences of *Oryza rufipogon* with known geographical origin using OryzaGenome 2.1 (Huang et al., 2012; Kajiya-Kanegae et al., 2021). We determined that five of the six intronic SNPs uncovered in *OsRGA1* associated with PET in Japonica landraces are also present in *Oryza rufipogon*, although not in strict covariation. We extracted the mean potential evapotranspiration (PET) that defines the local environment of these 446 wild rice lines and confirmed a similar pattern of association between natural genetic variation in *OsRGA1* and PET. Namely, *O. rufipogon* varieties that harbor these minor variants in *OsRGA1* are native to areas that experience lower PET. Conversely, *O. rufipogon* lines with the major variants experience higher PET (Figure 3D). The fact that genetic variation in *OsRGA1* is associated with a gradient in PET across the local environment in a wild rice relative, and that not only the genetic variation, but also its association with the same environmental variable is conserved in rice landraces supports the hypothesis that these natural variants in *OsRGA1* are adaptive.

In the distribution range of the georeferenced landraces studied here, areas with lower PET typically experience lower temperatures and lower precipitation regimes. We, therefore, hypothesize that landraces grown in climates of lower PET (and lower precipitation; Fig. 3A) should have been selected for drought tolerance. To evaluate this, we utilized the naturally occurring *d1* mutant available in the Japonica cultivar Taichung 65 (T65), with a 2 bp deletion in *OsRGA1* that results in a protein null allele (Oki et al., 2009). *d1* and wild-type rice plants were kept well-watered for 60 days post-germination, after which soil was maintained at 100% (no drought, ND), 45% (moderate drought, MD), or 35% (severe drought, SD), relative water content (RWC). The dramatic drought tolerant phenotype of *d1* is illustrated in Figure 3E. In the absence of drought (ND), both genotypes exhibited similar photosynthetic rates. However, under MD and SD, *d1* photosynthesis exceeds that of wild-type (Figure S2).

A stop codon polymorphism in *OsGS3*, a rice G protein γ subunit, shows an association with a Potential Evapotranspiration gradient and interacts with *OsRGA1* to regulate vegetative and reproductive traits

Given our findings on *OsRGA1*, we next used Oryza GenoCLIM to explore the relationship between genetic variation in other G protein subunits and climate variables. We did not identify strong patterns of interaction with the environment in variants of the

sole rice Gβ subunit gene, *OsRGB1 (Os03g0669200*). However, we uncovered a strong correlation between a polymorphism introducing a premature stop codon to *OsGS3* (*Os03g0407400*), a Gγ subunit of the rice heterotrimeric G protein (Figure 5A), and the same gradient in mean potential evapotranspiration (PET) during the coldest quarter that we identified for *OsRGA1* (Figures 5B and 5C). We used OryzaGenome 2.1 (Kajiya-Kanegae et al., 2021) to explore the genetic variation of *OsGS3* present in 446 imputed sequences of *Oryza rufipogon* with known geographical origin (Huang et al., 2012). However, the position of this variant in *OsGS3* coincides with a region of low coverage for which no information is available, so we could not further investigate the presence of *OsGS3* SNPs in this wild-rice species.

We uncovered a positive correlation between landrace height in Japonica varieties and PET during the coldest quarter of the year (Figure 4D). Shorter plants were typically found in areas with lower PET areas and were more likely to simultaneously possess the minor *OsRGA1* allele and the minor *OsGS3* allele, while taller plants from areas with higher evapotranspiration regimes typically harbor the major alleles of both *OsRGA1* and *OsGS3* (also referred to as *GS3-3*). Landrace varieties from intermediate evapotranspiration regimes typically harbor the minor *OsGS3* allele in the major *OsRGA1* allele background (Figure 4E). Although plant height in the landraces varieties with the minor *OsGS3* allele and major *OsRGA1* allele was significantly shorter than those with the major *OsRGA1* allele and the major *OsGS3* allele ($p < 0.05$), the average height associated with the minor *OsRGA1* allele suggests that *OsRGA1*, not *OsGS3,* is the major regulator of the variation in plant height in a G protein-mediated process across a PET gradient.

We also identified a positive correlation between landrace variety seed length and PET during the coldest quarter (Figure 4F, G). Shorter grains are typical of landraces grown in low PET (drier and colder and less isothermal areas; Figure 4F), and these varieties are more likely to harbor simultaneously the minor *OsRGA1* allele and the minor *OsGS3* allele (Figure 4F, G). Landraces with longer grains, grown in warmer, wetter, and less seasonal areas, typically harbor both the major allele of *OsRGA1* and the major *OsGS3* allele. Landrace varieties from areas with intermediate PET typically display the minor *OsGS3* allele in the major *OsRGA1* allele background. Although the average seed length in landraces with the minor *OsGS3* allele and minor *OsRGA1* allele is significantly shorter than in those with the major *OsRGA1* allele and major *OsGS3* allele ($p < 0.05$), the magnitude of the average seed length associated with the introduction of the minor *OsGS3* allele implicates the Gγ (*OsGS3*), and not the Gα subunit (*OsRGA1*) as the primary regulator of the variation in seed length in this G protein mechanism across a PET gradient.

Genetic variation in the *OsGS3* and *OsDEP1* Gγ subunits in rice interact to determine seed and panicle length, but not plant height across a Potential Evapotranspiration gradient

In the continued evaluation of associations between natural variation in rice G proteins and PET, we highlight another Gγ subunit, this time in Indica varieties. One single non-synonymous variant in *OsDEP1* (Figure 5A) was found in association with mean potential evapotranspiration (PET) during the coldest quarter, the identical environmental parameter that we observed to be associated with *OsRGA1* and *OsGS3* variants in Japonica landraces. We found that the *OsDEP1* minor allele only occurred in combination with the *OsGS3* major allele ('G', or the functional *OsGS3* allele in the Indica population), and landraces with this *OsDEP1*-minor and *OsGS3*-functional combination are typically grown in areas with low PET. Similarly, in Japonica landraces the *OsGS3*-functional allele was associated with lower PET (Figures 4B and C), though it is observed in lower frequency than in the Indica population (Figures 5B and C). Variation in *OsDEP1* is not associated with variation inl grain length and the minor *OsDEP1* allele is only found in short-grain varieties as it is exclusively present in varieties with a functional *OsGS3* allele within the population in our dataset (Figure 5D). As expected, the long-grain non-functional *OsGS3* allele (T) retained its association with the long-grain phenotype in Indica varieties (Figure 5D). This suggests that grain length is regulated by natural variation in *OsGS3* and not *OsDEP1*. Notably, the minor allele in *OsRGA1* (Figure 3B and 5A) was not present in a single Indica landrace, so there was no genetic interaction between *OsRGA1* and *OsGS3* or *OsDEP1* to be observed within the Indica population.

When we evaluated the association of *OsGS3* and *OsDEP1* with panicle length, we found that Indica varieties with the minor *OsDEP1* allele had shorter panicles relative to the major allele, regardless of the presence of *OsGS3* variation in this major *OsDEP1* allele background (Figure 5F). This suggests that panicle length is regulated by natural variation in *OsDEP1* and not *OsGS3*.

To support the observation that variants in the Gα subunit, *OsRGA1*, and not the Gγ subunits regulate plant height, as we observed in *OsRGA1* and *OsGS3* Japonica lines (Figure 4E), we explored the existing variation in plant height across the different *OsGS3*-*OsDEP1* allelic combinations in these Indica landraces. As all Indica landraces harbor the major *OsRGA1* allele, the lack of association between plant height and any combination of *OsGS3*-*OsDEP1* alleles (Figure S3), indicates that neither of these Gγ natural variants influences height in Indica landraces.

Agronomic traits regulated by natural variation in G proteins that affects the nature of their protein-protein interactions

As described above, we used Oryza CLIMtools to uncover the associations of G protein natural variants in Indica and Japonica landraces with the local environment and their regulation of adaptive agronomic traits. Given these results, we synthesized these findings to propose a molecular mechanism by which rice G proteins natural variants regulate signaling, with phenotypic consequencest.

We hypothesize that our observation that natural variation in *OsDEP1* and not *OsGS3* regulates panicle length in Indica, while *OsRGA1* and *OsGS3* do so in Japonica may be explained by: i) the absence of the *OsRGA1* minor haplotype in Indica, and; ii) the different versions of *OsGCC2* that we observe in Indica and Japonica. In *OsGGC2*, there is a set of three variants in the last exon and one in the 3' UTR region that covary, defining a haplotype in which the minor allele corresponds with the Japonica varieties, and the major allele is found in Indica varieties (Figure 6A and B). Given that rice Gγ subunits are known to function both additively and antagonistically (Sun et al., 2018), a resultant difference in *Os*GGC2 function may shift the equilibrium of signaling from other Gβγ dimers, and thereby alter the interactions between subunits.

We have previously shown that a C-terminal truncation of the Arabidopsis ortholog of *Os*GS3, AtAGG3, consisting of only part of the Gγ domain (residues 1-78), is not sufficient to interact with the Arabidopsis Gβ subunit (Chakravorty et al., 2011). Therefore, we hypothesized that the similarly truncated protein encoded by the minor allele of *OsGS3* would also not bind the rice Gβ subunit. The minor allele of *OsGS3* has previously been named *OsGS3-3* (Mao et al., 2010). Using the yeast 2-hybrid assay of protein-protein interaction, we observed that the protein encoded by the *OsGS3-3* allele indeed displays a severely impaired interaction with *Os*RGB1, as evident by weak yeast growth that occurs only in the absence of 3-AT (Figure 3A). By contrast, all full-length Gγ subunits, corresponding to the Taichung 65 sequences (*Os*RGG1, *Os*RGG2, *Os*GS3, *Os*DEP1, and *Os*GGC2) strongly bound *Os*RGB1 as assessed by growth on a semi-quantitative 3-AT concentration curve (Figure 3A). The previously identified truncations resulting from indels in *OsGS3* or *OsDEP1* that retain a full Gγ domain, namely *OsGS3-4* (Mao et al., 2010), *OsGS3-5*, *OsGS3-6*, *OsGS3-7* (Takano-Kai et al., 2013), *Osdep1* (Huang et al., 2009) and *OsDn1-1* (Taguchi-Shiobara et al., 2011; alignments in Figure S5A and S5B), also all bound *Os*RGB1 with similar or increased strengths compared to the corresponding wild-type subunits (Figure 3A). As expected, no Gγ subunits displayed any growth on interaction selective media in combination with the empty vector control (negative control; Figure S5C). We also utilized a yeast 3-hybrid approach to assess the interaction of the different Gβγ dimers with Gα subunits. Consistent with our yeast 2-hybrid data, the truncated *Os*GS3-3 protein did not allow heterotrimer formation with *Os*RGA1 (Figure 3B). The full-length Taichung 65 Gγ subunits all allowed heterotrimer formation with *Os*RGA1 (Figure 4B). The mutant proteins *OsGS3-4*, *OsGS3-5*, *OsGS3-6*, *OsGS3-7*, *Osdep1*, and *OsDn1-1* also allowed heterotrimer formation (Figs 4B). These results indicate that the *Os*GS3-3 allele likely acts as a functional null due to an inability

to complex with other G protein subunits. Contrastingly, the truncations that can bind Gβ may act as dominant negatives, or sequester Gα and Gβ subunits since we have previously shown that the Cys-rich tail of the Arabidopsis ortholog of OsGS3, AtAGG3, is required for Gγ function (Chakravorty et al., 2011). We did not examine the interactions of the proteins encoded by the *OsDEP1* minor allele (*Os*DEP1^{C261Y}) or the *OsGGC2* Indica allele (*Os*GGC2$^{L169W\ R321C}$) with *Os*RGB1 or with *Os*RGA1 as part of Gβγ: the changes encoded within these proteins occur in the Cys-rich tail, a region of Type C Gγ subunits that does not contribute to binding of other heterotrimeric G protein subunits (Chakravorty et al., 2011) and is proposed to reside in the apoplast (Wolfenstetter et al., 2015).

CLIMtools allows broad exploration of the genetic basis of local adaptation in rice

Above, we have demonstrated the utility of Oryza CLIMtools to uncover adaptive associations between phenotype and environmental variables codified in OryzaCLIM. We have used these tools to focus on agronomic roles of variation in heterotrimeric G proteins, but, as exemplified in Figure 7, Oryza CLIMtools can be utilized to investigate many putative mechanisms for potential genetic bases of adaptation to local climates.

Discussion

Rice is the staple food for more than half of the world's population and the main source of calorie intake in emerging and developing economies (Seck et al., 2012). As such, any reduction in global production will have important implications for food security worldwide (Godfray et al., 2010). A major challenge is thus to feed a burgeoning human population while mitigating crop losses incurred as a result of climate change.

The environment is the missing piece of a puzzle that already includes extensive information on the genetic diversity of rice. To date, no comprehensive resource has been available to investigate the relationship between genetic and environmental variation that exists in the local environment of rice. With this in mind, we created OryzaCLIM (https://gramene.org/CLIMtools/oryza_v1.0/OryzaCLIM/), an intuitive online resource available to the community to explore 413 geo-environmental variables that define the local environmental conditions at the collection sites 658 Indica and 283 Japonica sequenced landraces.

Heading date and adaptation to local temperature conditions

Average global temperatures are predicted to increase by 1.5°C between 2030 and 2052 (Delmotte et al., 2018). These rising temperatures pose a significant threat to rice

production (Zhao et al., 2016) and flowering time is, notably, the developmental stage most sensitive to heat stress (Matsui et al., 2001; Jagadish et al., 2010).

Many species in natural settings demonstrate shifts in flowering time associated with ongoing climate change (Fitter and Fitter, 2002). In the case of rice, modern high-yielding cultivars are grown in many areas of the world that diverge enormously in their environmental conditions. The flowering time of these modern cultivars is not optimized for all local environments. In other words, they are not 'climate-smart.' By contrast, local landraces are cultivated varieties that have evolved as a result of both natural and artificial selection (Casañas et al., 2017) and therefore are amenable to mining for adaptations of flowering time to the local environment.

Temperature is a key environmental factor patterning contemporary rice genomic diversity (Gutaker et al., 2020). Indica and Japonica landrace varieties show considerable phenotypic variability in their flowering time (Fig. S4), and here we show that flowering time in both Indica and Japonica landraces is positively correlated with the minimum temperature of the coldest month in their local environment (Figs 3B and 3F). Indeed, supraoptimal temperatures increase spikelet sterility and critically reduce rice yield (Shrestha et al., 2022). Rice is a facultative short-day plant. Many landraces in tropical and subtropical regions have mild photoperiod sensitivity that inhibits flowering under long days and promotes flowering under short days, maximizing the time for vegetative growth before flowering (Zong et al., 2021). At the same time, suboptimal temperatures negatively affect survival and yield (Zhang et al., 2014), and varieties from areas where temperatures during the cold season are suboptimal sacrifice vegetative growth but ensure early flowering and seed production before the onset of low temperatures concurrent with shorter photoperiods.

We describe here for the first time the relationship between flowering time and temperature in the local environment for a fully sequenced rice population and highlight the differential impact of two well-known flowering time genes, *OsHD2* and *OsSOC1,* in Japonica vs. Indica populations. Oryza CLIMtools opens the door to exploring the genetic basis of flowering time adaptation to create climate-smart rice varieties in which flowering time can be tailored to match the typical temperatures in the area of cultivation.

Bring the climate into climate-smart varieties: flowering time and temperature

As crop genetic diversity decreases, so does resilience to unfavorable climate conditions that are becoming more extreme in a climate change scenario (Food and Agriculture Organization of the United Nations et al., 2018). In favorable areas for rice cultivation, farmers target two to three crops per year. To do so, they must grow early flowering varieties that, as we see here, are not necessarily adapted to local conditions. In these cases, the lack of adaptation may result in lower yield per planting, which can be

compensated for by multiple plantings. One example may be planting of landraces harboring the maladaptive variant in *OsSOC1* that we describe here. *OsSOC1* is a well-conserved and key promoter of the transition from the vegetative to the reproductive growth phase in rice. Introducing flowering time stability, as is provided by the maladaptive minor allele in *OsSOC1,* may be farmer-smart in that it allows for multiple plantings per years, but may not be climate-smart due to the lack of adaptation.

We identified a very strong adaptive association of an allele in *OsHD2* and the minimum temperature of the coldest month (Figure 2B). *OsHD2* has been identified as a flowering time QTL that is important for the temperature regulation of photoinduction (Nakagawa et al., 2005). The Arabidopsis ortholog of Os*HD2*, *AtPRR7*, is a key component of the temperature compensation mechanism to maintain a constant circadian rhythm under changing temperatures (Salomé et al., 2010).

Consistent with our findings, previous work has shown that functional alleles of *OsHD2*, known to produce higher yields, are found mainly in warmer southern latitudes, while 'weaker' and early flowering, non-functional alleles of *OsHD2* are found in colder, northern latitudes (Yan et al., 2013; Koo et al., 2013; Li et al., 2015; Ye et al., 2018). Given this finding, we propose *OsHD2* as a key mediator of flowering time adaptation to the local environment. It is also possible to fine-tune flowering time via genome editing the upstream open reading frame of *OsHD2* (Liu et al., 2021). We can envision an approach by which the appropriate flowering time of any given variety can be modulated to match local temperatures, to create climate smart varieties, and maximize their potential yield in any given location.

Genetic variation in heterotrimeric G protein Gα subunit *OsRGA1* is associated with a gradient in Potential Evapotranspiration

Most modern, high-yielding rice varieties rely on a mutant *SD-1* gene, introduced during the Green Revolution, as the basis of their semi-dwarf phenotype (Hargrove and Cabanilla, 1979; Hargrove et al., 1980; Ferrero-Serrano et al., 2019). However, the *qDTY1.1* allele, a QTL for drought tolerance, has been lost in these modern varieties due to its tight linkage in repulsion with the *sd1* allele. Accordingly, the widespread use of *sd1* as a source of dwarfism both reduces the genetic diversity of this trait and results in modern rice cultivars that are typically drought sensitive. This illustrates one of the most documented consequences of genetic erosion in crops: the reduction in adaptability or tolerance to abiotic factors (Khoury et al., 2022).

Drought is one of plants' most widespread and damaging environmental stress factors (Farooq et al., 2012; Gupta et al., 2020) and affects rice particularly as this crop is often acclimated to either rain-fed or fully irrigated systems. For this reason, the

identification of an alternative source of dwarfism that also provides an adaptive advantage to drought can contribute to yield stability in global yield production. In this study, we argue that natural variants in other dwarfing genes are available that are positively associated with drought-related environments.

Following this rationale, we explored the association of natural-occurring variants in the heterotrimeric G protein Gα subunit, *OsRGA1,* and the local environment, particularly drought-related environmental variables. Naturally occurring functional null *OsRGA1* mutants display a characteristic dwarf phenotype (Ashikari et al., 1999; Oki et al., 2009; Ferrero-Serrano and Assmann, 2016; Zait et al., 2021). Loss of function *OsRGA1* dwarf mutant plants display improved drought resistance compared to wild-type plants as a consequence of their increased water use efficiency, increased mesophyll conductance, lower leaf temperature, and increased light use efficiency (Ferrero-Serrano and Assmann, 2016; Ferrero-Serrano et al., 2018; Zait et al., 2021).

Using Oryza GenoCLIM (https://gramene.org/CLIMtools/oryza_v1.0/Oryza_GenoCLIM/), we identified a set of covarying SNPs in the intronic and upstream regions of *OsRGA1* (Figure 3A) in association with a gradient in Potential Evapotranspiration (PET). The minor allele is found in landraces collected from areas with lower PET, and the most frequent version of the haplotype (major allele) is present in landraces collected from areas with higher PET (Figure 3C).

Potential evapotranspiration (Thornthwaite, 1948; Penman and Keen, 1997) is the amount of water that would evaporate and transpire from a hypothetical reference surface under optimal irrigation conditions and is influenced by environmental factors such as temperature, humidity, solar radiation, and wind (Allen, R.G., Pereira, L.S., Raes, D., and Smith, M., 1998). This index provides an estimate of the amount of water-use for growth and can therefore be used to select drought-tolerant varieties. As shown in Figure 3A, areas with lower PET correspond to colder but also more variable temperatures throughout the year and lower precipitation regimes. In other words, lower PET is associated with more challenging and unpredictable environments (Figure 3A). We hypothesize that those landraces from areas of lower PET will have improved abiotic stress tolerance to multiple factors, a trait in which G proteins play a critical role (Wu and Urano, 2018; Wang and Botella, 2022). We illustrate this point by showcasing the improved drought tolerance of the *RGA1* natural mutant, *d1* (Fig. 3E).

The role of the minor *OsRGA1* variant in the adaptation to lower PET areas is supported by our finding that five of the six intronic *OsRGA1* variants that we found in rice landraces are conserved in the ancestral species of rice, its wild rice relative *Oryza rufipogon*, (Huang et al., 2012), where they are also associated with a gradient in PET. This finding in an ancestor of domesticated rice supports the hypothesis that natural variation in *OsRGA1* is adaptive across a PET gradient in South-Eastern Asia.

Genetic variation in heterotrimeric G proteins is associated with a precipitation gradient and the interplay of vegetative and reproductive traits

Heterotrimeric G proteins play crucial roles in relaying signals from external stimuli through cell surface receptors to internal effector proteins. Rice encodes one canonical Gα (*OsRGA1*), one Gβ (*OsRGB1*), and five Gγ (*OsRGG1*, *Os*RGG2, *Os*GS3, *Os*DEP1, and *Os*GGC2) subunits (Ueguchi-Tanaka et al., 2000; Utsunomiya et al., 2011). G proteins have been proposed to play major roles in regulating agronomic traits and stress responses (Botella, 2012; Cui et al., 2020), and the influence of natural variants in the different G protein subunits in defining these phenotypes has been discussed (Botella, 2012; Wu et al., 2022; Ferrero-Serrano and Chakravorty, 2023). Such studies typically describe the roles of individual subunits and their variants in isolation. Here, we reveal how interaction among natural variants in the Gαβγ complex determines agronomic traits and stress responses that are important to maximize fitness in the local environment.

The minor allele in *OsRGA1* is associated with extremely short grains, possibly by impairing *OsDEP1* and *OsGCC2* signaling. Indeed, knocking out either *OsDEP1* or *OsGCC2* results in short plant, panicle and grain phenotypes that are epistatic to the *OsRGA1 d1* mutation (Sun et al., 2018; Chaya et al., 2022). While we did not identify variants of interest in *OsRGB1*, we found different natural variants of the Gγ subunits in association with PET, the same environmental variable associated with the *OsRGA1* variants, suggesting a genetic interaction.

Gγ proteins are divided into three distinct groups according to their C-terminal domains, with the rice genome encoding two shorter Gγ subunits, *OsRGG1* (Type A) and *OsRGG2* (Type B), and three longer Gγ subunits with a Cys-rich tail: *OsDEP1*, *OsGGC2*, and *OsGS3* (Type C; Trusov et al., 2012). A series of indels have been identified in *OsGS3* and *OsDEP1* that confer changes in grain shape (Mao et al., 2010; Takano-Kai et al., 2013) and a dense and erect panicle morphology (Huang et al., 2009; Taguchi-Shiobara et al., 2011), respectively. The minor allele of *OsGS3* in Indica landraces, previously named both *gs3* and *GS3-3*, encodes a nonsense mutation, C55*, within the Gγ domain and was found to cause a long-grain phenotype (Fan et al., 2006; Mao et al., 2010) which we now show is found in areas of higher PET (Figure 4B and C). The *OsGS3-3* variant was introduced from a Japonica ancestor(s) into the Indica gene pool (Takano-Kai et al., 2009; Mao et al., 2010). Indeed, in our dataset, the *OsGS3-3* allele is at a higher frequency in the Japonica population while at a lower frequency in the Indica population (Figures 5 and 6). We show that the *Os*GS3-3 protein is unable to dimerize with Gβ (Figure 6), which would relieve *Os*GS3-mediated competition for Gβ dimerization with the Gγ subunits that promote grain elongation, *Os*DEP1 and *Os*GGC2 (Sun et al., 2018), resulting in a long-grain phenotype. Contrastingly, multiple *OsGS3* mutations that result in a super short-grain phenotype retain the ability to interact with *Os*RGB1, namely,

*Os*GS3-4, *Os*GS3-5, *Os*GS3-6 and *Os*GS3-7 (Figure 6), consistent with the ability of these proteins to partially suppress *Os*DEP1 and *Os*GGC2 dimerization with Gβ, despite presumably not retaining full intrinsic signaling functionality of their own. It should be noted that while we focussed on heterotrimers containing the canonical Gα subunit *OsRGA1* in this study, rice also encodes four proteins belonging to the plant specific class of extra-large Gα proteins, namely *Os*XLG1, *Os*XLG3a, *Os*XLG3b and *Os*XLG4 (Cantos et al., 2023). We, therefore, tested the ability of the rice Gβγ dimers, including those containing mutant Gγ subunits, to interact with the four *Os*XLGs. The full-length Gγ subunits of the Taichung 65 cultivar all allowed heterotrimer formation with some or all Gα subunits. Only the *Os*GS3-3 protein failed to interact with any of the *Os*XLGs (Figure S5D-G), strengthening our conclusion that the *Os*GS3-3 protein is non-functional.

As our observations of genetic interaction between different Gγ subunit variants are consistent with the concept of Gγ competition for dimerization with the Gβ subunit, introgression of different Gγ alleles into elite varieties is likely to have unintended effects on signaling from other heterotrimeric G protein components. As such scenarios of pleiotropism and genetic interactions are not uncommon, we envision that Oryza CLIMtools will facilitate the identification of sources of complementary genetic variation and suitable genetic backgrounds for rational breeding and genetic improvement programs; unlike previous studies, CLIMtools provides this information in the context of environmental adaptations. As an example, we characterized the environmental association of the *OsGS3-3* variant. We discovered that the functional allele is found in areas of lower PET (Figure 4B and C). *OsGS3* is orthologous to *AtAGG3* from Arabidopsis (Chakravorty et al., 2011). Our previous work identified a role of *AtAGG3* in adaptation to cold temperatures and the association of a missense variant in *AtAGG3* with temperature in the coldest month of the year. In that example, the minor variant is associated with colder areas in northern Eurasia (Ferrero-Serrano and Assmann, 2019), consistent with the lower temperatures that are typical of the areas with low PET that are associated with the functional *OsGS3* in rice.

Originally identified as an erect panicle QTL (Kong et al., 2007; Yan et al., 2007), *OsDEP1* encodes another Gγ protein that, like *OsGS3*, is an atypical Gγ subunit with a Cys-rich tail. Natural variation in *OsDEP1* regulates meristematic activity, with a previously described variant resulting in a reduced length of the inflorescence internode, erect and shorter panicles, an increased number of grains per panicle, and a consequent increase in grain yield (Huang et al., 2009). This variant involves the replacement of a 637-bp stretch of the middle of exon 5 with a 12-bp sequence, creating a premature stop codon and the loss of 230 residues from the C terminus (Huang et al., 2009). The variant we identify here is a missense variant that also occurs in the middle of the 5th exon (Figure 5A). We show that, just as for *OsGS3*, variation in *OsDEP1* is also associated with PET (Figure 5). However, based on our analysis of panicle length in Indica landraces, we propose that natural variation in *OsDEP1*, but not *OsGS3*, is responsible for variation in

this phenotype. We reach this conclusion based on the association of the minor *OsDEP1* allele with shorter panicle phenotypes and the major *OsDEP1* allele with longer panicle phenotypes, independent of a functional/non-functional *OsGS3* allele (Figure 5F). However, this conclusion is challenged by our observation of differences in panicle length in Japonica landraces associated with natural variation in *OsRGA1* and OsGS3 (Figure S4). We suggest that this discrepancy may be due to an interaction with the *OsRGA1* haplotype in Japonica that is not possible in Indica, which completely lacks the *OsRGA1* haplotype. Alternatively, there may be a genetic interaction with the different haplotypic versions of *OsGCC2* present in Indica vs. Japonica (Figure 6A).

We recently proposed a mechanism regulated by G proteins, particularly the Gα subunit, that involves a trade-off between growth and stress tolerance that is critical to mediate adaptability (Ferrero-Serrano and Chakravorty, 2023). We hypothesize that landraces from areas with lower PET (typically, at higher latitudes (Figure 3C)), experience a “less predictable” environment (Figure 3A) and trade-off their stress tolerance with a derived cost in growth that is illustrated by the reduced stature associated with the minor *OsRGA1* allele (Figure 4D and E). This point is illustrated not only by the increased drought tolerance of the dwarf *d1* mutant, but also by previous findings relating loss-of-function mutants of Gα with resistance to diverse stresses in other species (Zhang et al., 2008; Nilson and Assmann, 2010; Jangam et al., 2016; Ferrero-Serrano et al., 2018; Cui et al., 2020). We also highlight the association of natural variation in the Gγ subunit with the same environmental gradient we found in association with *RGA1* variants, PET, and in this case, also with reproductive traits that are important to maximize yield. These results highlight the potential of modulating G protein function for vegetative stress tolerance, and the associated interplay between stress tolerance, growth, and yield.

Conclusion

There is an evident mismatch between the current distributions of crop varieties and climate suitability for their production. Breeding programs that introduce genetic diversity from local landraces can help to address this challenge (Hoisington et al., 1999; Fu and Somers, 2009; Zhang et al., 2017; Renard and Tilman, 2019) but will benefit from taking environmental factors into account.

We have demonstrated the utility of Oryza CLIMtools to i) uncover adaptive associations between phenotype and environmental variables that are included in OryzaCLIM. Using this rationale we have demonstrated the association between flowering time and temperature, and the differential importance of two flowering time genes in Japonica vs. Indica in this environmental association. Further, we have ii) demonstrated the use of Oryza CLIMGeno to explore the genetic basis of temperature adaptation and discuss the maladaptive role of *OsSOC1* and the adaptive role of *OsHD2* in regulating this process. Finally, we have iii) exemplified the use of Oryza GenoCLIM to

uncover the climate associations of different G protein subunits, to identify their coordinated role in adaptation to a PET gradient, and their regulation of different agronomic traits to facilitate this adaptive process. With Oryza CLIMtools we provide extensive resources to explore the genetic basis of local adaptation to inform breeding or genetic engineering strategies to improve climate resilience in crops.

The results described in this report highlight the promise of introducing natural genetic variation from landraces into modern rice varieties. The introduction of naturally occurring genetic variation into breeding programs is not new, but the assessment of genotypic interaction with the local environment adds a new dimension to approaches to create climate-smart and climate-change-ready varieties and create a more sustainable food production system.

Acknowledgments

Supported by National Institute of General Medical Sciences of the NIH under award number 5R01GM126079 to SMA and NSF-IOS-2122357 to PCB and SMA. Kobie Kirven acknowledges support from NIH training grant 5T32GM102057. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We thank Prof. D. Ware, Dr. A. Olson, and Prof. P. C. Bevilacqua for helpful discussions, and comments on an early version of this manuscript. We also thank the Gramene Team (https://www.gramene.org/) for hosting Oryza CLIMtools.

Materials and Methods

Extraction of environmental variables

We compiled a comprehensive database of environmental data for 658 Indica and 283 Japonica georeferenced landrace varieties included in the 3K Rice Genome Project (Gutaker et al., 2020) that includes 413 geo-climatic variables describing the local environment at each collection site (Figure 1, S1, and Table S1). We used the Raster package in R (Hijmans et al., 2013) to extract the geospatial raster data for the environmental conditions prevailing at each landrace's collection site. A detailed list of relevant literature references that describe the climate dataset from which we curated the environmental information, and summary statistics for each of these 413 environmental variables can be found in Table S2.

OryzaCLIM

We developed OryzaCLIM, a user-friendly Shiny component (Chang et al., 2015a) of Oryza CLIMtools that enables the analysis of local environmental data for each georeferenced accession (Gutaker et al., 2020) sequenced as part of the 3,000 (3K) Rice Genome Project (Wang et al., 2018). OryzaCLIM incorporates 413 geo-environmental variables that have been codified to facilitate the examination of environmental conditions at the collection sites of 658 Indica and 283 Japonica georeferenced landraces. These data points are represented on an interactive world map that allows users to analyze pairwise environmental conditions for these 941 landrace varieties.

By clicking on any of the data points on the map, users can access information about the selected variety, including its curation details and the value of the chosen environmental variable. OryzaCLIM is designed to be user-friendly and accessible, and the source code and data used to create this tool are freely available for download at GitHub (https://github.com/CLIMtools/OryzaCLIM). The code is licensed under Apache 2.0.

GWAS analysis

We obtained SNP and INDEL data from the 3K genomes (https://snp-seek.irri.org/_download.zul;jsessionid=0E0F5BCFB0EA14EBB926BEBC7B26D6E0) and converted the data from Plink format to VCF format using PLINK 2.0 (Chang et al., 2015b). We then used VCFtools (Danecek et al., 2011) to subset the resulting VCFs into separate files for 658 Indica and 283 Japonica landrace varieties, which were the focus of this study and were used as two separate GWAS panels. The data were filtered to select genetic variants with less than 50% missing data and minimum allele frequency (MAF) ≥ 5%. After this initial filtering, imputation was conducted using BEAGLE v5.4 with default parameters (Browning et al., 2018).

Rice includes two major sub-species: *Oryza sativa ssp. Indica* and *Oryza sativa ssp. Japonica*. Differences between the subspecies are apparent in a number of physiological and morphological traits, genetic variation, and geographical distribution (Khush, 1997; Garris et al., 2005; Huang et al., 2012; Wang et al., 2018). For instance, our exploration of previously characterized agronomic traits (Wang et al., 2020) in the set of landrace varieties used in this study revealed that Indica landrace varieties typically display more culms that are longer, longer leaves that are narrower, shorter stature at the seedling stage, earlier flowering, more panicles, and longer grain than Japonica varieties (Figure S4). Given such marked subspecies differences, we conducted GWAS analysis separately in the Indica and Japonica subpopulations. We used a general linear model containing a correction for population structure using the first five components as the population structure matrix (GLM+PCA; Zhao et al., 2018).

Using the ‘stats’ package in R (R Core Team), we calculated the Benjamini-Hochberg (Benjamini and Hochberg, 1995) and Bonferroni (Holm, 1979) values for

genome-wide significance to evaluate significance thresholds for multiple tests in our GWAS analysis.

To predict the impact of every genetic variant, we used SnpEff (Cingolani et al., 2012) with the gene model annotation from the Ensembl genome release 53 (https://ftp.ensemblgenomes.ebi.ac.uk/pub/plants/release-53/). We focus on SNPs within transcriptional units, including introns, untranslated regions at 5' and 3' UTRs, and 1 kb promoter regions upstream of the most distal transcription start site. We used the "consensus transcript," as described by SnpEff, as the variant with either the longest coding sequence (CDS) if the gene has translated transcripts or the longest cDNA (Hubbard et al., 2009; Cingolani et al., 2012). We defined 'common variants' as those with MAFs≥5% to restrict our analysis to genetic variation that is more likely to be adaptive.

Genomic signatures of selection were calculated using VCFtools (Danecek et al., 2011). We obtained the fixation index, Fst (Wright, 1965; Weir and Cockerham, 1984), on a per-site basis and the genetic group based on Indica vs. Japonica cultivars. We also calculated nucleotide diversity (π) and Tajima's D in 1 kb windows across the genome (Danecek et al., 2011).

RiboSNitch prediction

A riboSNitch is a single nucleotide polymorphism (SNP) that alters the secondary structure of an RNA. To identify riboSNitches in the rice transcriptome, we used the SNPfold software tool (Halvorsen et al., 2010). SNPfold predicts riboSNitches by using a thermodynamic model that calculates the partition function and base pairing probability matrices for the reference and variant sequences defined by any given SNP. SNPfold then compares the column sums of the base-pairing probability matrices by computing the Pearson correlation coefficient. A low SNPfold score, and thus a low Pearson correlation coefficient, indicates a potential structural change in the RNA, i.e. a riboSNitch. We suggest that SNPs are candidate riboSNitches if the correlation coefficient comparing the reference and variant sequences is less than 0.8 (Halvorsen et al., 2010), but in Oryza CLIMtools we provide the specific SNPfold score (correlation coefficient) for each variant that was predicted, so that users can choose their own threshold for designating a SNP as a riboSNtich.

SNPfold was applied to predict the presence of riboSNitches within the 3,914,483 natural genetic variants among the 658 Indica landrace varieties and 2,835,066 genetic variants among the 283 Japonica landrace varieties studied here. Each SNP in these transcriptomes was placed in the context of its surrounding 80 nucleotides (40 nucleotides upstream and 40 nucleotides downstream) and passed through the SNPfold program to determine the impact of the mutation on the local RNA structure. For the Oryza CLIMtools site, correlation coefficients are provided for all SNPs within annotated transcribed regions unless the SNP is located less than 40 nucleotides from a transcript end, in which

case the correlation coefficient was not calculated. SNPs located less than 40 nucleotides from either end of a transcript were designated as Not Applicable in Oryza CLIMtools.

Oryza GenoCLIM

Oryza GenoCLIM is a user-friendly SHINY (Chang et al., 2015a) component of Oryza CLIMtools. GenoCLIM is a searchable database generated with the DT package (Xie, 2016) that provides information for any locus ID for genetic variants found to be significantly associated with each of the 413 geo-environmental variables included in this study. This tool allows the user to search a locus identifier of any given gene, annotated using the Rice Annotation Project (RAP annotation; (Sakai et al., 2013)), and retrieve, if present, any association between the genetic variation (SNPs and INDELs) within that gene and any given environmental variable(s) described in this study.

The database provides climate associations of genetic variants, and the user can query their ‘gene of interest’ within the Indica or Japonica varieties in an interactive database. The results shown can be filtered according to different FDR thresholds and sorted based on their association score or on any of the other population genetic indicators provided for these associations. Additionally, Oryza GenoCLIM provides an plot generated by Plotly (Sievert et al., 2018), which allows users to visualize the results interactively.

For reproducibility, the source code and data used in GenoCLIM are freely available to download on GitHub (https://github.com/CLIMtools/Oryza_GenoCLIM), with the code licensed under Apache 2.0.

Oryza CLIMGeno

Oryza CLIMGeno comprises a software suite of Shiny apps (Chang et al., 2015a) to analyze the relationships between genetic variation and environmental factors in rice. CLIMGeno is built on Zbrowse (Ziegler et al., 2015), a genome browser tool. Using CLIMGeno, users can visualize the top 2,500 genetic variants (SNPs and INDELS) associated with each of the 413 environmental variables analyzed in this study. Specifically, the tool displays the highest-scoring variants in Indica (658) or Japonica (283) rice varieties.

In Oryza CLIMGeno, the 'manage' tab provides users with a dropdown menu to select an environmental variable of interest. Once selected, the tool automatically generates a list of the 2,500 genetic variants (SNPs and INDELs) with the strongest association score for that variable. Users can navigate to the 'data table' tab to select and sort genetic variation based on various details such as position, allele frequency, locus identification, locus description, riboSNitch prediction, or SNP effect.

The 'whole-genome view' tab displays an interactive Manhattan plot that represents the genetic variation associated with the geo-climatic variable previously

selected by the 'whole-genome view' tab. The x-axis shows each of the twelve rice chromosomes, while the y-axis displays the score (negative logarithm of the p-value), depicting the strength of association between genetic variability and any given environmental variable. Users can then filter the genetic variants to display based on their predicted effect using the 'select effect' box. Hovering over the data points provides information on the genetic effect of a given genetic variant.

Clicking on a genetic variant takes users to the 'chromosome view' tab, which displays two Manhattan plots corresponding to the chromosome where the selected variant is located. The top plot allows users to browse genetic variation on that chromosome and zoom in on any given region. In contrast, the bottom plot combines a Manhattan plot with an interactive display and annotation of genes found in the region of interest according to the gene model annotation from the Ensembl genome release (https://ftp.ensemblgenomes.ebi.ac.uk/pub/plants/release-53/). Users can explore the region of the chromosome surrounding the genetic variant(s) of interest by selecting a window size or using the menu on the left. Lastly, the 'annotations table' tab provides an interactive table with annotation information on the region of interest selected on the Manhattan plot, complemented by a graphical annotation display.

For reproducibility, the source code and data used in Oryza CLIMGeno for both Indica and Japonica varieties are freely available to download at GitHub (https://github.com/CLIMtools/Oryza_CLIMGeno_Indica and https://github.com/CLIMtools/Oryza_CLIMGeno_Japonica), with code licensed under Apache 2.0.

Plant growth conditions and physiological characterization of drought tolerance in an *OsRGA1* mutant, *d1*.

Wild-type (WT) and *d1* mutants were grown in a greenhouse in 2 US gallon pots containing Metro-mix 360 potting mixture. Greenhouse temperatures were maintained at 30˚C during the day and 20˚C during the night in a 16:8 day:night cycle with light supplied as natural daylight supplemented with 1000W metal halide lamps (Philips Lighting Co., Somerset, NJ) for the duration of the light cycle. Light intensity was ~500 μmol m^{-2} s^{-1} PPFD. Plants were maintained under well-watered conditions until 60 days after emergence. At that point, we introduced two different drought treatments, which resulted in moderate and severe stress under the chosen treatments. Plants were watered twice daily to maintain at all times, by lysimetry, the three water treatments: well watered (ND, 100% soil relative water content), moderate drought (MD, 45% SRWC) and severe drought (SD, 30% SRWC). There were 7 replicates per water treatment per genotype, for a total of 42 plants. Relative soil water content was kept constant from the start of the drought treatment until seed maturity, by continued twice daily watering to the appropriate weight and continuous monitoring with a Campbell Scientific TDR 100 system (Campbell

Scientific Inc., Logan, Utah, USA) with a custom-made probe of 20 cm length. It was determined that in-between watering events, the relative soil water content of the soil did not vary more than 5%.

For gas exchange measurements, regions of flag leaves were enclosed in the chamber of a portable gas exchange system (LI-COR 6400 IRGA with an integrated 6400-40 leaf chamber fluorometer, Li-COR, Inc., Lincoln, NE, USA). Leaves were measured at the point of maximal width. The leaf chamber temperature was kept constant at 30°C. Airflow in the chamber was adjusted to 300 µmol s^{-1}. Steady-state measurements of the different physiological parameters were performed at 500 µmol m^{-2} s^{-1} PPFD provided by the Li-COR red/blue LED system with blue light accounting for 10% of the total photon flux.

Amplification of G protein subunits

The G protein subunits of rice are: *OsRGA1* (Os05g0333200), *OsXLG1* (Os12g0593000), *OsXLG3a* (Os11g0206700), *OsXLG3b* (Os06g0111400), *OsXLG4* (Os10g0117800), *OsRGB1* (Os03g0669200), *OsRGG1* (Os03g0635100), *OsRGG2* (Os02g0137800), *OsGS3* (Os03g0407400), *OsDEP1* (Os09g0441900) and *OsGGC2* (Os08g0456600). G protein subunit open reading frames were amplified from seedling or flower cDNA of the Japonica Taichung 65 variety and cloned into the pCR™8/GW/TOPO™ entry vector (Thermo). The 5' end of the *XLG4* ORF includes partially repetitive regions of 80+ % GC content and could not be amplified from cDNA. To circumvent this issue, the region of the *XLG4* cDNA encoding amino acids 198-828 was amplified from Taichung 65 cDNA and joined by overlap extension PCR to a gBlock fragment (IDT) of the region encoding amino acids 1-197 that was designed to encode the correct amino acids with reduced GC content. Overlap extension PCR was used to generate the deletions of *OsGS3* and *OsDEP1* previously described elsewhere (*GS3-4*, *OsGS3-5*, *OsGS3-6*, *OsGS3-7*, *Osdep1,* and *OsDn1-1*), while a *OsGS3-3* clone was created by incorporating the C55* mutation into a 3' primer that only amplifies the shortened open reading frame of the *OsGS3-3* allele. The *OsGS3-3* and *OsGS3-4* alleles were originally named by Mao et al. (Mao et al., 2010). We continued this naming convention with the alleles described by Takano-Kai et al. (2013), i.e. *OsGS3-5* corresponds to the allele identified in the Podiwi-A8 cultivar, *OsGS3-6* to the allele from JC73-4, JC157, and ABRI, *OsGS3-7* to the allele from H343 and Tumo-Tumo, and the allele identified in ARC7291 and TAL214 cultivars correspond to the already described *OsGS3-4* allele (Mao et al., 2010) identified in the Chuan 7 cultivar (Fan et al., 2006). *Osdep1* was described by Huang *et al.* (Huang et al., 2009), and *OsDn1-1* by Taguchi-Shiobara *et al.* (Taguchi-Shiobara et al., 2011). Multiple alignments of the protein sequences of the above *OsGS3* and *OsDEP1* alleles were conducted using Clustal Omega (Nguyen et al., 2016).

Yeast protein-protein interaction assays

Sequence-verified entry clones were mobilized into pDEST-GADT7 and pDEST-GBKT7 destination yeast 2-hybrid vectors (Uetz et al., 2006) by Gateway LR recombination. For yeast 3-hybrid, in which the interaction of two proteins is assayed by a yeast 2-hybrid approach in the presence of a third "bridge" protein, *OsRGB1* was amplified with flanking NotI and BamHI sites to clone into the NotI/BglII sites of pBridge MCS2 (Takara) as the "bridge" protein. Gγ subunits were amplified with flanking BglII sites to ligate into the compatible BamHI site (*OsGS3* alleles), or EcoRI and SalI restriction sites (*OsRGG1*, *OsRGG2*, *OsGGC2*), or EcoRI and BamHI restriction sites (*OsDEP1* alleles) to clone into MCS1 of pBridge as GAL4BD fusions. Yeast 2-/3-hybrid assays were performed as previously described (Chakravorty et al., 2015).

Figure and Table Legends

Figure 1. Global landrace distribution and environmental variables defining the local environment of the 941 Indica and Japonica varieties included in this study. A. Geographic distribution of 658 Indica (orange data points) and 283 Japonica (blue data points) landrace varieties with known collection sites sequenced as part of the 3,000 (3K) Rice Genome Project (Wang et al., 2018; Gutaker et al., 2020). **B.** Waffle chart visualization of the types of climate variable descriptors extracted from the collection sites (see panel B). Each square on the chart represents one environmental variable, and each square's color indicates the type of environmental variable. **C.** Oryza CLIMtools provides a suite of online resources to provide gene x phenotype x environment associations to assist rice research and breeding programs.

Figure 2. Natural variation in *HD2* tunes the relationship between heading date and temperature in the local environment; maladaptive natural variation in *SOC1* disrupts it. A. We explored an environmental and phenotypic cline in Japonica landraces that is associated with a variant haplotype in *OsHD2*. **B.** Flowering time in Japonica landraces is positively correlated with the minimum temperature of the coldest month for their local environment (r^2 = 0.42). Raincloud plots illustrate significantly different probability densities of the association of the local environment and flowering time for the major and minor alleles in *OsHD2* depicted in 3A. Dark grey datapoints and distributions represent landraces with the minor allele in *OsHD2*. Coral datapoints and distributions represent landraces with the major allele in *OsHD2*. **C.** We also explored an environmental and phenotypic cline in Indica landraces that is associated with a natural variant in *OsSOC1*. **D.** Flowering time in Indica landraces, similarly to Japonica landraces, is positively correlated with the minimum temperature of the coldest month for their local environment (r^2 = 0.22). Raincloud plots illustrate significantly different probability

densities of the association of the local environment and flowering time for the major and minor alleles of the intronic SNP in *OsSOC1* which is depicted in 3A. Red datapoints and distributions represent landraces with the minor allele in *OsSOC1*. Blue datapoints and distributions represent landraces with the major allele in *OsSOC1*. **E.** Planting sites in Hainan (China) and Los Baños (Philippines) for 576 Indica landraces with flowering time records at both of these different locations between these publicly available datasets retrieved from the Functional and Genomic Breeding (RFGB 2.0; Wang et al., 2020) and SNPseek (Mansueto et al., 2017) resource sites. **F.** Raincloud plot illustrating the significant flowering time stability of Indica landraces harboring the minor allele *OsSOC1* when planted at different locations. The difference in flowering time is calculated simply by subtracting the number of days that landraces took to flower in Hainan (China) from the days to flower in Los Baños (Philippines). The positive values for most landraces indicate that most landraces flowered later at lower latitudes in Los Baños, relative to those in Hainan. For the raincloud plots, in the box plots, the lower and upper boundaries indicate the 25th and 75th percentile, respectively. The whiskers below and above the box indicate the smallest value within 1.5 times the interquartile range (IQR) below the 25th percentile. The whisker above the box indicates the largest value within 1.5 times the IQR above the 75th percentile, respectively. The black line inside the box indicates the median. Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** (p-<0.0001); *** ($0.0001 < p < 0.001$); ** ($0.001 < p < 0.01$); * ($0.01 < p < 0.05$); non-significant comparisons (ns) ($p > 0.05$). The associations between temperature and flowering time were fitted to a linear model. Regression lines are shown in black; gray shading represents 95% confidence intervals.

Figure 3. Haplotypic distribution of SNPs in *OsRGA1* is correlated with the mean potential evapotranspiration (PET) during the coldest quarter of the year that is typical of their local environment. A. We explored an environmental cline in a haplotype defined by six covarying intronic SNPs and eight upstream SNPs in *OsRGA1*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** Raincloud plots illustrate significantly different probability densities of the *association of the local environment* for the major and minor alleles of *OsRGA1*. The map of South-Eastern Asia shows the geographical distribution of the major and minor haplotypic variants of *OsRGA1*. The grayscale color gradient in this map illustrates the cline in PET during the coldest quarter. Purple dots and distributions represent landraces with the major allele in *OsRGA1*. Orange dots represent landraces with the minor allele in *OsRGA1*. **D.** The genetic variation present in 446 imputed sequences of *Oryza rufipogon* reveals that four of the six intronic SNPs uncovered in *OsRGA1* in Japonica landraces are also associated with PET during the coldest quarter of the year in this wild rice species. Blue dots and distributions represent landraces with the major allele of *OsRGA1*. Red dots represent landraces with the minor allele in *OsRGA1*. The same colors depict the haplotypic

distribution on the map of South-Eastern Asia with a color gradient that represents the cline in this drought-related variant. **C.** The assessment of drought tolerance of the *d1* mutant available in the Japonica cultivar Taichung 65 (T65) confirms its drought tolerance (see Figure S2). **** ($p < 0.0001$; non-parametric Wilcoxon test).

Figure 4. Haplotypic distribution of *OsRGA1* and a stop-causing codon SNP in *OsGS3* are correlated with mean potential evapotranspiration (PET) in the local environment and natural variation in plant height and seed length. A. We explored environmental and phenotypic clines that were associated with a haplotypic variant in *OsRGA1* (see Figure S3A) and a stop codon-causing SNP in *OsGS3*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** The map of South-Eastern Asia shows the geographic distribution of the major and minor haplotypic variants of *OsRGA1*. The grayscale color gradient illustrates the cline in PET during the coldest quarter. **C.** Raincloud plots illustrate significantly different probability densities of the *association of* the local environment for the major and minor alleles of *OsRGA1* in the presence/absence of the minor allele *of OsGS3*. **D.** There is a positive correlation between plant height and PET during the coldest quarter. Haplotypic variation in *OsRGA1* determines this correlation and is strongly associated with plants that display shorter stature as shown in Panel **E**, which is consistent with the well-reported dwarf phenotype of *OsRGA1* loss of function mutants (Oki et al., 2009; See Figure S2). **F.** There is also a positive correlation between grain length and PET during the coldest quarter. Haplotypic variation in *OsRGA1* and allelic variation in *OsGS3* determine this correlation and are strongly associated with plants that display shorter grains as shown in Panel **G**, consistent with the well-reported short-grain phenotype of *OsRGA1* and *OsGS3* loss of function mutants (Fan et al., 2006; See Figure S3). Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** ($p < 0.0001$).

Figure 5. Haplotypic distribution of *OsDEP1* missense allele variants and a stop-causing codon SNP in *OsGS3* are correlated with the mean potential evapotranspiration in the local environment and natural variation in seed and panicle length A. We explored environmental and phenotypic gradients associated with landraces defined by a stop codon-causing SNP in *OsGS3* and a missense variant in *OsDEP1*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** The map of South-Eastern Asia shows the geographic distribution of the major and minor haplotypic variants of the different *OsGS3*-*OsDEP1* allelic combinations. The grayscale color gradient illustrates the gradient in PET during the coldest quarter **C.** Raincloud plots illustrate significantly different probability densities of the association of the local environment for the major and minor alleles of *OsGS3* in the presence/absence of the minor allele *of OsDEP1*. Both B and C illustrate the association of the functional (major)

OsGS3 allele and the minor *OsDEP1* allele with lower PET regions. **D.** Raincloud plot illustrates the association of the *OsGS3* variant with grain length and the lack of association of the *OsDEP1* variant with grain length. **E.** Raincloud plot illustrates the association of the *OsDEP1* variant with panicle length and the lack of association of the *OsGS3* variant with panicle length. Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** ($p < 0.0001$).

Figure 6. Genetic variation in heterotrimeric G proteins is associated with a precipitation gradient and the interplay of vegetative and reproductive traits. A. We uncover a set of three covarying variants in the last exon and one in the 3'UTR of OsGGC2 with one haplotype typical of Japonica, and the other typical of Indica varieties, that may affect the equilibrium of the Gγ subunits' competition for dimerization with Gβ. **B.** Cartoon summarizing our hypothesis on how the alternative combination of the different heterotrimeric G protein variants associated with PET in the local environment affects agronomic traits that may have an adaptive value. See Discussion for details. **C.** Yeast 2-hybrid assays between Gβ and Gγ, and **D.** Yeast 3-hybrid assays between Gβγ dimers and the canonical Gα, OsRGA1, were spotted on SC-Trp-Leu media as growth control, and SC-Trp-Leu-Met-His with the concentrations of 3-Amino-1,2,4-triazole (3-AT) indicated above each column. 3-AT is a competitive inhibitor of the *HIS3* reporter gene of the yeast 2-/3-hybrid systems and therefore allows for semi-quantitative assessment of interaction strength, with growth at higher 3-AT concentrations indicating a stronger interaction. All *Os*RGG1, *Os*RGG2, *Os*GS3, *Os*DEP1 and *Os*GGC2 proteins encoded by Gγ alleles were able to complex with OsRGB1 and *Os*RGA1 via *Os*RGB1, except *Os*GS3-3. **E.** Domain schematic of rice Type C Gγ subunits including known truncations. N = N-terminal sequence, Gγ = Gγ-like domain, TM = putative transmembrane domain, Cys-rich = C-terminal Cys-rich domain and dark gray boxes correspond to out of frame translated regions. See Figure S5A and S5B for multiple alignments of protein sequences.

Figure 7. Examples of genotype x environment associations found using Oryza CLIMtools. Example associations for Indica landraces between **A**. an upstream gene variant in *OsSIRTUIN2B* and interannual temperature, **B.** a missense variant in *OsPIDL1* and annual mean temperature, **C.** a missense variant in *Os09g0371100* and soil pH. **D.** a synonymous variant in *OsHSFB1* and the mean temperature of the coldest month. In Japonica landraces, we uncovered an association between **E.** the same upstream variant that we identified in Indica landraces in *OsSIR2b* with the same climate parameter, interannual temperature, **F.** a 3'UTR variant in *OSCHIT1,* and the ultraviolet index during the spring season, **G.** another 3'UTR variant, here in *OsPPI-PFK* and temperature seasonality, and **H.** a missense variant in *OsNCED3* and latitude.

Figure S1. Correlations among the environmental variables presented in this study for the set of 413 geo-environmental variables extracted for the local environment of the 658 Indica and 283 Japonica georeferenced landrace varieties depicted in this study. The environmental variables in this plot are ordered by hierarchical clustering. White indicates a lack of correlation, with a continuous color scale ranging from dark blue (strong positive correlation) to dark red (strong negative correlation).

Figure S2. *OsRGA1* loss of function mutant (*d1*) displays greater drought tolerance and distinctly shorter panicles and seeds than wild-type (Taichung 65). A. In the absence of drought (ND), *d1* displays similar photosynthetic rates relative to wild-type (WT), but under medium drought (MD) and severe drought (SD), *d1* photosynthesis exceeds that of wild-type **B.** *d1* plants exhibit longer panicles and **C.** longer seeds.

Figure S3. Genetic variation in *OsGS3* and *OsDEP1* is not associated with a differential plant height. Raincloud plots illustrate that the introduction of *OsGS3* and *OsDEP1* variants in Indica varieties is not associated with differential plant height.

Figure S4. Genetic variation in *OsRGA1* and *OsGS3* is associated with differences in panicle length. Raincloud plots illustrate that the introduction of *OsRGA1* and *OsGS3* variants in Japonica varieties is associated with differential panicle length. Four asterisks (****) indicate a *p-value* less than 0.0001 using the non-parametric Wilcoxon test.

Figure S5. *Os*GS3-3 does not interact with heterotrimeric G protein subunits. Multiple sequence alignments of previously described naturally occurring truncations of **A.** *Os*GS3 and **B.** *Os*DEP1. **C.** Empty vector-Gγ yeast 2-hybrid panel (negative control for Figure 6C), demonstrating a lack of substantial growth which would indicate aberrant autoactivation in the assay. Yeast 3-hybrid assays between Gβγ dimers and the extra-large Gα subunits, **D.** *Os*XLG1, **E.** *Os*XLG3a, **F.** *Os*XLG3b and **G.** *Os*XLG4. Assays were carried out in parallel with the Gβγ-*Os*RGA1 assay depicted in Figure 6D. **H.** Gβγ-empty vector yeast 3-hybrid panel (negative control for Figures 6D, S5D, S5E, S5F and S5G). Only *Os*RGG1-*Os*RGB1 displayed weak growth on 0 mM 3-AT medium, and therefore no autoactivation was observed that would alter the interpretation of the interactions observed for the various heterotrimeric G protein combinations.

Figure S6. The two major rice sub-species, Indica and Japonica, display distinct phenotypic characteristics. Data corresponding to the phenotypic characterization of agronomic traits of Japonica and Indica landrace varieties were retrieved from the Functional and Genomic Breeding online resource (RFGB 2.0; Wang et al., 2020). Raincloud plots combine box plots, raw jittered data, and split-half violins. Jittered data points depict individual Indica (orange) and Japonica (blue) landrace varieties. In the box

plots, the lower and upper boundaries indicate the 25th and 75th percentile, respectively. The whiskers below and above the box indicate the smallest value within 1.5 times the interquartile range (IQR) below the 25th percentile. The whisker above the box indicates the largest value within 1.5 times the IQR above the 75th percentile, respectively. The black line inside the box indicates the median. Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** (p-<0.0001); *** (p-<0.0001 < p < 0.001); ** (0.001 < p < 0.01); * (0.01 < p < 0.05); non-significant comparisons (ns) (p > 0.05). The associations between temperature and flowering time were fitted to a linear model. Regression lines are shown in black; gray shading represents 95% confidence intervals.

Table S1: Environmental data descriptors and summary statistics.

Table S2: Environmental data corresponding to the 413 geo-environmental variables that define the local environment at the collection site of the 658 Indica and 283 Japonica landraces originating from South-Eastern Asia included in this study.

Literature cited

3,000 rice genomes project (2014). The 3,000 rice genomes project. *Gigascience* **3**:7.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. (1998). *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. Food and Agriculture Organization of the United Nations (FAO Irrigation and Drainage Paper No. 56).

Ashikari, M., Wu, J., Yano, M., Sasaki, T., and Yoshimura, A. (1999). Rice gibberellin-insensitive dwarf mutant gene *Dwarf 1* encodes the α-subunit of GTP-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* **96**:10284–10289.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**:289–300.

Bin Rahman, A. N. M. R., and Zhang, J. (2023). Trends in rice research: 2030 and beyond. *Food and Energy Security* **12**:e390.

Botella, J. R. (2012). Can heterotrimeric G proteins help to feed the world? *Trends Plant Sci.* **17**:563–568.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* **103**:338–348.

Cantos, C. F., dePamphilis, C. W., and Assmann, S. M. (2023). Extra-large G proteins have extra-large effects on agronomic traits and stress tolerance in maize and rice.

Trends Plant Sci. Advance Access published May 7, 2023, doi:10.1016/j.tplants.2023.04.005.

Casañas, F., Simó, J., Casals, J., and Prohens, J. (2017). Toward an evolved concept of landrace. *Front. Plant Sci.* **8**:145.

Chakravorty, D., Trusov, Y., Zhang, W., Acharya, B. R., Sheahan, M. B., McCurdy, D. W., Assmann, S. M., and Botella, J. R. (2011). An atypical heterotrimeric G-protein γ-subunit is involved in guard cell K -channel regulation and morphological development in *Arabidopsis thaliana*. *Plant J.* **67**:840–851.

Chakravorty, D., Gookin, T. E., Milner, M. J., Yu, Y., and Assmann, S. M. (2015). Extra-large G proteins expand the repertoire of subunits in Arabidopsis heterotrimeric G protein signaling. *Plant Physiol.* **169**:512–529.

Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. (2015a). shiny: Web Application Framework for R. R package version 1.6.0. https://cran.r-project.org/web/packages/shiny/index.htm.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015b). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**:7.

Chaya, G., Segami, S., Fujita, M., Morinaka, Y., Iwasaki, Y., and Miura, K. (2022). *OsGGC2*, Gγ subunit of heterotrimeric G protein, regulates plant height by functionally overlapping with *DEP1* in Rice. *Plants* **11**.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., and Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. **6**:80–92.

Cui, Y., Jiang, N., Xu, Z., and Xu, Q. (2020). Heterotrimeric G protein are involved in the regulation of multiple agronomic traits and stress tolerance in rice. *BMC Plant Biol.* **20**:90.

Dalrymple, D. G. (1986). *Development and spread of high-yielding rice varieties in developing countries*. Int. Rice Res. Inst.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., et al. (2011). The variant call format and VCFtools. *Bioinformatics* **27**:2156–2158.

Delmotte, V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Okia, W. M., Péan, C., Pidcock, R., et al. (2018). Global warming of 1.5◦ C. An IPCC Special Report on the impacts of global warming of 1.5◦ C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change,

sustainable development, and efforts to eradicate poverty. In *Report of the Intergovernmental Panel on Climate Change*.

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X., and Zhang, Q. (2006). *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* **112**:1164–1171.

Farooq, M., Hussain, M., Wahid, A., and Siddique, K. H. M. (2012). Drought stress in plants: an overview. In *Plant responses to drought stress: from morphological to molecular features* (ed. Aroca, R.), pp. 1–33. Berlin, Heidelberg: Springer Berlin Heidelberg.

Ferrero-Serrano, Á., and Assmann, S. M. (2016). The α-subunit of the rice heterotrimeric G protein, *RGA1*, regulates drought tolerance during the vegetative phase in the dwarf rice mutant *d1*. *J. Exp. Bot.* **67**:3433–3443.

Ferrero-Serrano, Á., and Assmann, S. M. (2019). Phenotypic and genome-wide association with the local environment of Arabidopsis. *Nat Ecol Evol* **3**:274–285.

Ferrero-Serrano, Á., and Chakravorty, D. (2023). Plants and heterotrimeric G proteins: Expect the unexpected. *Mol. Plant* **16**:506–508.

Ferrero-Serrano, Á., Su, Z., and Assmann, S. M. (2018). Illuminating the role of the Gα heterotrimeric G protein subunit, RGA1, in regulating photoprotection and photoavoidance in rice. *Plant Cell Environ.* **41**:451–468.

Ferrero-Serrano, Á., Cantos, C., and Assmann, S. M. (2019). The role of dwarfing traits in historical and modern agriculture with a focus on rice. *Cold Spring Harb. Perspect. Biol.* **11**.

Ferrero-Serrano, Á., Sylvia, M. M., Forstmeier, P. C., Olson, A. J., Ware, D., Bevilacqua, P. C., and Assmann, S. M. (2022). Experimental demonstration and pan-structurome prediction of climate-associated riboSNitches in Arabidopsis. *Genome Biol.* **23**:101.

Fitter, A. H., and Fitter, R. S. R. (2002). Rapid changes in flowering time in British plants. *Science* **296**:1689–1691.

Food and Agriculture Organization of the United Nations, United Nations International Children's Emergency Fund, World Health Organization, World Food Programme, and International Fund for Agriculture Development (2018). *The state of food security and nutrition in the world 2018: building climate resilience for food security and nutrition*. Food & Agriculture Org.

Frankel, O. H. (1974). Genetic conservation: our evolutionary responsibility. *Genetics* **78**:53–65.

Fu, Y.-B., and Somers, D. J. (2009). Genome-wide reduction of genetic diversity in wheat breeding. *Crop Sci.* **49**:161.

Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**:1631–1638.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science* **327**:812–818.

Gupta, A., Rico-Medina, A., and Caño-Delgado, A. I. (2020). The physiology of plant responses to drought. *Science* **368**:266–269.

Gutaker, R. M., Groen, S. C., Bellis, E. S., Choi, J. Y., Pires, I. S., Bocinsky, R. K., Slayton, E. R., Wilkins, O., Castillo, C. C., Negrão, S., et al. (2020). Genomic history and ecology of the geographic spread of rice. *Nat Plants* **6**:492–502.

Halvorsen, M., Martin, J. S., Broadaway, S., and Laederach, A. (2010). Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet.* **6**:e1001074.

Hargreaves, G. H., and Samani, Z. A. (1982). Estimating potential evapotranspiration. *J. Irrig. Drain. Div.* **108**:225–230.

Hargrove, T. R., and Cabanilla, V. L. (1979). The impact of semidwarf varieties on Asian rice-breeding programs. *Bioscience* **29**:731–735.

Hargrove, T. R., Coffman, W. R., and Cabanilla, V. L. (1980). Ancestry of improved cultivars of Asian rice 1. *Crop Sci.* **20**:721–727.

Hijmans, R. J., Van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., and Shortridge, A. (2013). Raster: Geographic Data Analysis and Modeling. R package version 2.0. https://cran.r-project.org/web/packages/raster/raster.pdf

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J. M., Skovmand, B., Taba, S., and Warburton, M. (1999). Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* **96**:5937–5943.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scand. Stat. Theory Appl.* **6**:65–70.

Huang, X., Qian, Q., Liu, Z., Sun, H., He, S., Luo, D., Xia, G., Chu, C., Li, J., and Fu, X. (2009). Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat. Genet.* **41**:494–497.

Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., et al. (2012). A map of rice genome variation reveals the origin of

cultivated rice. *Nature* **490**:497–501.

Hubbard, T. J. P., Aken, B. L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., et al. (2009). Ensembl 2009. *Nucleic Acids Res.* **37**:D690-7.

Jagadish, S. V. K., Muthurajan, R., Oane, R., Wheeler, T. R., Heuer, S., Bennett, J., and Craufurd, P. Q. (2010). Physiological and proteomic approaches to address heat tolerance during anthesis in rice (*Oryza sativa L.*). *J. Exp. Bot.* **61**:143–156.

Jangam, A. P., Pathak, R. R., and Raghuram, N. (2016). Microarray analysis of rice *d1* (*RGA1*) mutant reveals the potential role of G-protein alpha Subunit in regulating multiple abiotic stresses such as drought, salinity, heat, and cold. *Front. Plant Sci.* **7**:11.

Kajiya-Kanegae, H., Ohyanagi, H., Ebata, T., Tanizawa, Y., Onogi, A., Sawada, Y., Hirai, M. Y., Wang, Z.-X., Han, B., Toyoda, A., et al. (2021). OryzaGenome2.1: database of diverse genotypes in Wild Oryza Species. *Rice* **14**:24.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**:1–20.

Kato, S. (1930). On the affinity of the cultivated varieties of rice plants, Oryza sativa L. Journal of the Department of Agriculture, Hokkaido Imperial University, 22(3), 121-198.

Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., de Haan, S., Engels, J. M. M., Guarino, L., Hoban, S., Mercer, K. L., Miller, A. J., et al. (2022). Crop genetic erosion: understanding and responding to loss of crop diversity. *New Phytol.* **233**:84–118.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. In *Oryza: from molecule to plant* (ed. Sasaki, T.) and Moore, G.), pp. 25–34. Dordrecht: Springer Netherlands.

Kong, F. N., Wang, J. Y., Zou, J. C., Shi, L. X., De Jin, M., Xu, Z. J., and Wang, B. (2007). Molecular tagging and mapping of the erect panicle gene in rice. *Mol. Breed.* **19**:297–304.

Koo, B.-H., Yoo, S.-C., Park, J.-W., Kwon, C.-T., Lee, B.-D., An, G., Zhang, Z., Li, J., Li, Z., and Paek, N.-C. (2013). Natural variation in *OsPRR37* regulates heading date and contributes to rice cultivation at a wide range of latitudes. *Mol. Plant* **6**:1877–1888.

Li, X., Liu, H., Wang, M., Liu, H., Tian, X., Zhou, W., Lü, T., Wang, Z., Chu, C., Fang, J., et al. (2015). Combinations of *Hd2* and *Hd4* genes determine rice adaptability to Heilongjiang Province, northern limit of China. *J. Integr. Plant Biol.* **57**:698–707.

Liu, X., Liu, H., Zhang, Y., He, M., Li, R., Meng, W., Wang, Z., Li, X., and Bu, Q. (2021). Fine-tuning flowering time via genome editing of upstream open reading frames of *Heading Date 2* in rice. *Rice* **14**:59.

Mansueto, L., Fuentes, R. R., Borja, F. N., Detras, J., Abriol-Santos, J. M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A., et al. (2017). Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Res.* **45**:D1075–D1081.

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., Li, X., and Zhang, Q. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. U.S.A.* **107**:19579–19584.

Matsui, T., Omasa, K., and Horie, T. (2001). The difference in serility due to high temperatures during the flowering period among Japonica-rice varieties. *Plant Prod. Sci.* **4**:90–93.

McCudden, C. R., Hains, M. D., Kimple, R. J., Siderovski, D. P., and Willard, F. S. (2005). G-protein signaling back to the future. *Cell. Mol. Life Sci.* **62**:551–577.

Nakagawa, H., Yamagishi, J., Miyamoto, N., Motoyama, M., Yano, M., and Nemoto, K. (2005). Flowering response of rice to photoperiod and temperature: a QTL analysis using a phenological model. *Theor. Appl. Genet.* **110**:778–786.

Nguyen, K., Guo, X., and Pan, Y. (2016). *Multiple Biological Sequence Alignment: Scoring Functions, Algorithms and Evaluation*. John Wiley & Sons.

Nilson, S. E., and Assmann, S. M. (2010). The alpha-subunit of the Arabidopsis heterotrimeric G protein, *GPA1*, is a regulator of transpiration efficiency. *Plant Physiol.* **152**:2067–2077.

Oki, K., Inaba, N., Kitano, H., Takahashi, S., Fujisawa, Y., Kato, H., and Iwasaki, Y. (2009). Study of novel *d1* alleles, defective mutants of the α subunit of heterotrimeric G-protein in rice. *Genes Genet. Syst.* **84**:35–42.

Østerberg, J. T., Xiang, W., Olsen, L. I., Edenbrandt, A. K., Vedel, S. E., Christiansen, A., Landes, X., Andersen, M. M., Pagh, P., Sandøe, P., et al. (2017). Accelerating the domestication of new crops: feasibility and approaches. *Trends Plant Sci.* **22**:373–384.

Penman, H. L., and Keen, B. A. (1997). Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **193**:120–145.

R Core Team *(2020).* R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Renard, D., and Tilman, D. (2019). National food production stabilized by crop diversity. *Nature* **571**:257–260.

Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C.-C., Iwamoto, M., Abe, T., et al. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* **54**:e6.

Salomé, P. A., Weigel, D., and McClung, C. R. (2010). The role of the Arabidopsis morning loop components *CCA1*, *LHY*, *PRR7*, and *PRR9* in temperature compensation. *Plant Cell* **22**:3650–3661.

Seck, P. A., Diagne, A., Mohanty, S., and Wopereis, M. C. S. (2012). Crops that feed the world 7: Rice. *Food Security* **4**:7–24.

Shrestha, S., Mahat, J., Shrestha, J., K C, M., and Paudel, K. (2022). Influence of high-temperature stress on rice growth and development. A review. *Heliyon* **8**:e12651.

Sievert, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M., and Despouy, P. (2018). Plotly for R. https://plotly-r.com/.

Sun, S., Wang, L., Mao, H., Shao, L., Li, X., Xiao, J., Ouyang, Y., and Zhang, Q. (2018). A G-protein pathway determines grain size in rice. *Nat. Commun.* **9**:1–11.

Sweeney, M., and McCouch, S. (2007). The complex history of the domestication of rice. *Ann. Bot.* **100**:951–957.

Taguchi-Shiobara, F., Kawagoe, Y., Kato, H., Onodera, H., Tagiri, A., Hara, N., Miyao, A., Hirochika, H., Kitano, H., Yano, M., et al. (2011). A loss-of-function mutation of rice *DENSE PANICLE 1* causes semi-dwarfness and slightly increased number of spikelets. *Breed. Sci.* **61**:17–25.

Takano-Kai, N., Jiang, H., Kubo, T., Sweeney, M., Matsumoto, T., Kanamori, H., Padhukasahasram, B., Bustamante, C., Yoshimura, A., Doi, K., et al. (2009). Evolutionary history of *GS3*, a gene conferring grain length in rice. *Genetics* **182**:1323–1334.

Takano-Kai, N., Jiang, H., Powell, A., McCouch, S., Takamure, I., Furuya, N., Doi, K., and Yoshimura, A. (2013). Multiple and independent origins of short seeded alleles of *GS3* in rice. *Breed. Sci.* **63**:77–85.

Tello-Ruiz, M. K., Naithani, S., Gupta, P., Olson, A., Wei, S., Preece, J., Jiao, Y., Wang, B., Chougule, K., Garg, P., et al. (2021). Gramene 2021: harnessing the power of comparative genomics and pathways for plant research. *Nucleic Acids Res.* **49**:D1452–D1463.

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geogr. Rev.* **38**:55–94.

Title, P. O., and Bemmels, J. B. (2018). ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological

niche modeling. *Ecography* **41**:291–307.

Trusov, Y., Chakravorty, D., and Botella, J. R. (2012). Diversity of heterotrimeric G-protein γ subunits in plants. *BMC Res. Notes* **5**:608.

Ueguchi-Tanaka, M., Fujisawa, Y., Kobayashi, M., Ashikari, M., Iwasaki, Y., Kitano, H., and Matsuoka, M. (2000). Rice dwarf mutant *d1*, which is defective in the α subunit of the heterotrimeric G protein, affects gibberellin signal transduction. *Proc. Natl. Acad. Sci. U.S.A.* **97**:11638–11643.

Uetz, P., Dong, Y.-A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., Rajagopala, S. V., Roupelieva, M., Rose, D., Fossum, E., et al. (2006). Herpesviral protein networks and their interaction with the human proteome. *Science* **311**:239–242.

Utsunomiya, Y., Samejima, C., Takayanagi, Y., Izawa, Y., Yoshida, T., Sawada, Y., Fujisawa, Y., Kato, H., and Iwasaki, Y. (2011). Suppression of the rice heterotrimeric G protein β-subunit gene, *RGB1*, causes dwarfism and browning of internodes and lamina joint regions. *Plant J.* **67**:907–916.

Wang, Y., and Botella, J. R. (2022). Heterotrimeric G protein signaling in abiotic stress. *Plants* **11**.

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R. R., Zhang, F., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**:43–49.

Wang, C.-C., Yu, H., Huang, J., Wang, W.-S., Faruquee, M., Zhang, F., Zhao, X.-Q., Fu, B.-Y., Chen, K., Zhang, H.-L., et al. (2020). Towards a deeper haplotype mining of complex traits in rice with RFGB v2.0. *Plant Biotechnol. J.* **18**:14–16.

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistic for the analysis of population structure. *Evolution* **38**:1358–1370.

Wolfenstetter, S., Chakravorty, D., Kula, R., Urano, D., Trusov, Y., Sheahan, M. B., McCurdy, D. W., Assmann, S. M., Jones, A. M., and Botella, J. R. (2015). Evidence for an unusual transmembrane configuration of *AGG3*, a class C Gγ subunit of Arabidopsis. *Plant J.* **81**:388–398.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**:395–420.

Wu, T.-Y., and Urano, D. (2018). Genetic and systematic approaches toward G protein-coupled abiotic stress signaling in plants. *Front. Plant Sci.* **9**:1378.

Wu, T.-Y., Krishnamoorthi, S., Boonyaves, K., Al-Darabsah, I., Leong, R., Jones, A. M., Ishizaki, K., Liao, K.-L., and Urano, D. (2022). G protein controls stress readiness by modulating transcriptional and metabolic homeostasis in *Arabidopsis thaliana* and *Marchantia polymorpha*. *Mol. Plant* Advance Access published

November 1, 2022, doi:10.1016/j.molp.2022.10.020.

Xie, Y. (2016). DT: A wrapper of the JavaScript library 'DataTables'. https://cran.r-project.org/package=DT

Yan, C.-J., Zhou, J.-H., Yan, S., Chen, F., Yeboah, M., Tang, S.-Z., Liang, G.-H., and Gu, M.-H. (2007). Identification and characterization of a major QTL responsible for erect panicle trait in japonica rice (*Oryza sativa L.*). *Theor. Appl. Genet.* **115**:1093–1100.

Yan, W., Liu, H., Zhou, X., Li, Q., Zhang, J., Lu, L., Liu, T., Liu, H., Zhang, C., Zhang, Z., et al. (2013). Natural variation in *Ghd7.1* plays an important role in grain yield and adaptation in rice. *Cell Res.* **23**:969–971.

Ye, J., Niu, X., Yang, Y., Wang, S., Xu, Q., Yuan, X., Yu, H., Wang, Y., Wang, S., Feng, Y., et al. (2018). Divergent *Hd1*, *Ghd7*, and *DTH7* alleles control heading date and yield potential of Japonica rice in Northeast China. *Front. Plant Sci.* **9**.

Zait, Y., Ferrero-Serrano, Á., and Assmann, S. M. (2021). The α subunit of the heterotrimeric G protein regulates mesophyll CO_2 conductance and drought tolerance in rice. *New Phytol.* **232**:2324–2338.

Zhang, W., He, S. Y., and Assmann, S. M. (2008). The plant innate immunity response in stomatal guard cells invokes G-protein-dependent ion channel regulation. *Plant J.* **56**:984–996.

Zhang, Q., Chen, Q., Wang, S., Hong, Y., and Wang, Z. (2014). Rice and cold stress: methods for its evaluation and summary of cold tolerance-related quantitative trait loci. *Rice* **7**:24.

Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., and Song, B.-H. (2017). Back into the wild-Apply untapped genetic diversity of wild relatives for crop improvement. *Evol. Appl.* **10**:5–24.

Zhao, C., Piao, S., Wang, X., Huang, Y., Ciais, P., Elliott, J., Huang, M., Janssens, I. A., Li, T., Lian, X., et al. (2016). Plausible rice yield losses under future climate warming. *Nat. Plants* **3**:16202.

Zhao, H., Mitra, N., Kanetsky, P. A., Nathanson, K. L., and Rebbeck, T. R. (2018). A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). *Stat. Appl. Genet. Mol. Biol.* **17**.

Ziegler, G. R., Hartsock, R. H., and Baxter, I. (2015). Zbrowse: an interactive GWAS results browser. *PeerJ Comput. Sci.* **1**:e3.

Zong, W., Ren, D., Huang, M., Sun, K., Feng, J., Zhao, J., Xiao, D., Xie, W., Liu, S., Zhang, H., et al. (2021). Strong photoperiod sensitivity is controlled by cooperation

and competition among *Hd1*, *Ghd7* and *DTH8* in rice heading. *New Phytol.* **229**:1635–1649.

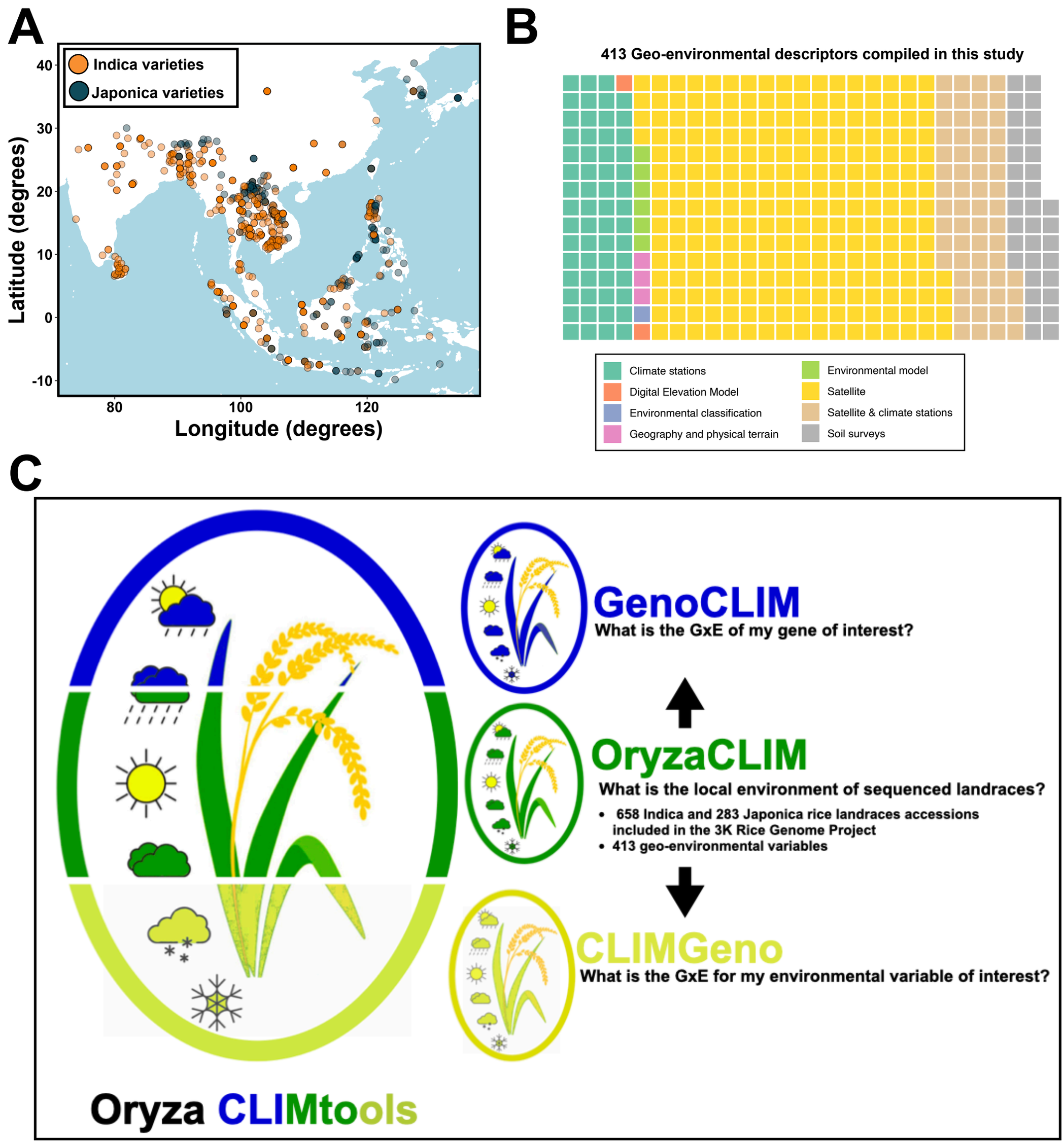

Figure 1. Global landrace distribution and environmental variables defining the local environment of the 941 Indica and Japonica varieties included in this study. A. Geographic distribution of 658 Indica (orange data points) and 283 Japonica (blue data points) landrace varieties with known collection sites sequenced as part of the 3,000 (3K) Rice Genome Project (Wang et al., 2018; Gutaker et al., 2020). **B.** Waffle chart visualization of the types of climate variable descriptors extracted from the collection sites (see panel B). Each square on the chart represents one environmental variable, and each square's color indicates the type of environmental variable. **C.** Oryza CLIMtools provides a suite of online resources to provide gene x phenotype x environment associations to assist rice research and breeding programs.

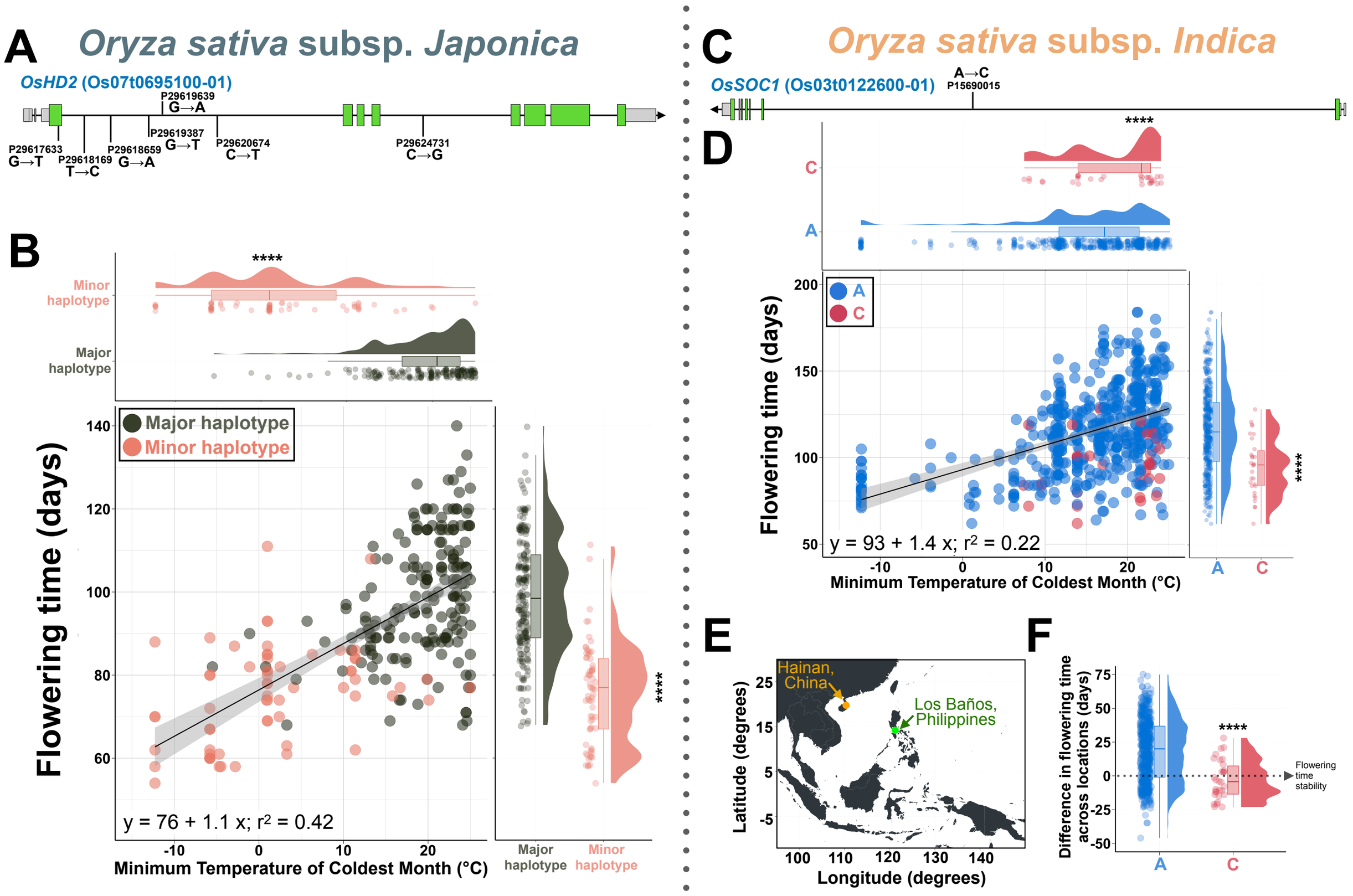

Figure 2. Natural variation in *HD2* tunes the relationship between heading date and temperature in the local environment; maladaptive natural variation in *SOC1* disrupts it. A. We explored an environmental and phenotypic cline in Japonica landraces that is associated with a variant haplotype in *OsHD2*. **B.** Flowering time in Japonica landraces is positively correlated with the minimum temperature of the coldest month for their local environment (r^2 = 0.42). Raincloud plots illustrate significantly different probability densities of the association of the local environment and flowering time for the major and minor alleles in *OsHD2* depicted in 3A. Dark grey datapoints and distributions represent landraces with the minor allele in *OsHD2*. Coral datapoints and distributions represent landraces with the major allele in *OsHD2*. **C.** We also explored an environmental and phenotypic cline in Indica landraces that is associated with a natural variant in *OsSOC1*. **D.** Flowering time in Indica landraces, similarly to Japonica landraces, is positively correlated with the minimum temperature of the coldest month for their local environment (r^2 = 0.22). Raincloud plots illustrate significantly different probability densities of the association of the local environment and flowering time for the major and minor alleles of the intronic SNP in *OsSOC1* which is depicted in 3A. Red datapoints and distributions represent landraces with the minor allele in *OsSOC1*. Blue datapoints and distributions represent landraces with the major allele in *OsSOC1*. **E.** Planting sites in Hainan (China) and Los Baños (Philippines) for 576 Indica landraces with flowering time records at both of these different locations between these publicly available datasets retrieved from the Functional and Genomic Breeding (RFGB 2.0; Wang et al., 2020) and SNPseek (Mansueto et al., 2017) resource sites. **F.** Raincloud plot illustrating the significant flowering time stability of Indica landraces harboring the minor allele *OsSOC1* when planted at different locations. The difference in flowering time is calculated simply by subtracting the number of days that landraces took to flower in Hainan (China) from the days to flower in Los Baños (Philippines). The positive values for most landraces indicate that most landraces flowered later at lower latitudes in Los Baños, relative to those in Hainan. For the raincloud plots, in the box plots, the lower and upper boundaries indicate the 25th and 75th percentile, respectively. The whiskers below and above the box indicate the smallest value within 1.5 times the interquartile range (IQR) below the 25th percentile. The whisker above the box indicates the largest value within 1.5 times the IQR above the 75th percentile, respectively. The black line inside the box indicates the median. Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** (p-<0.0001); *** (0.0001 < p < 0.001); ** (0.001 < p < 0.01); * (0.01 < p < 0.05); non-significant comparisons (ns) (p > 0.05). The associations between temperature and flowering time were fitted to a linear model. Regression lines are shown in black; gray shading represents 95% confidence intervals.

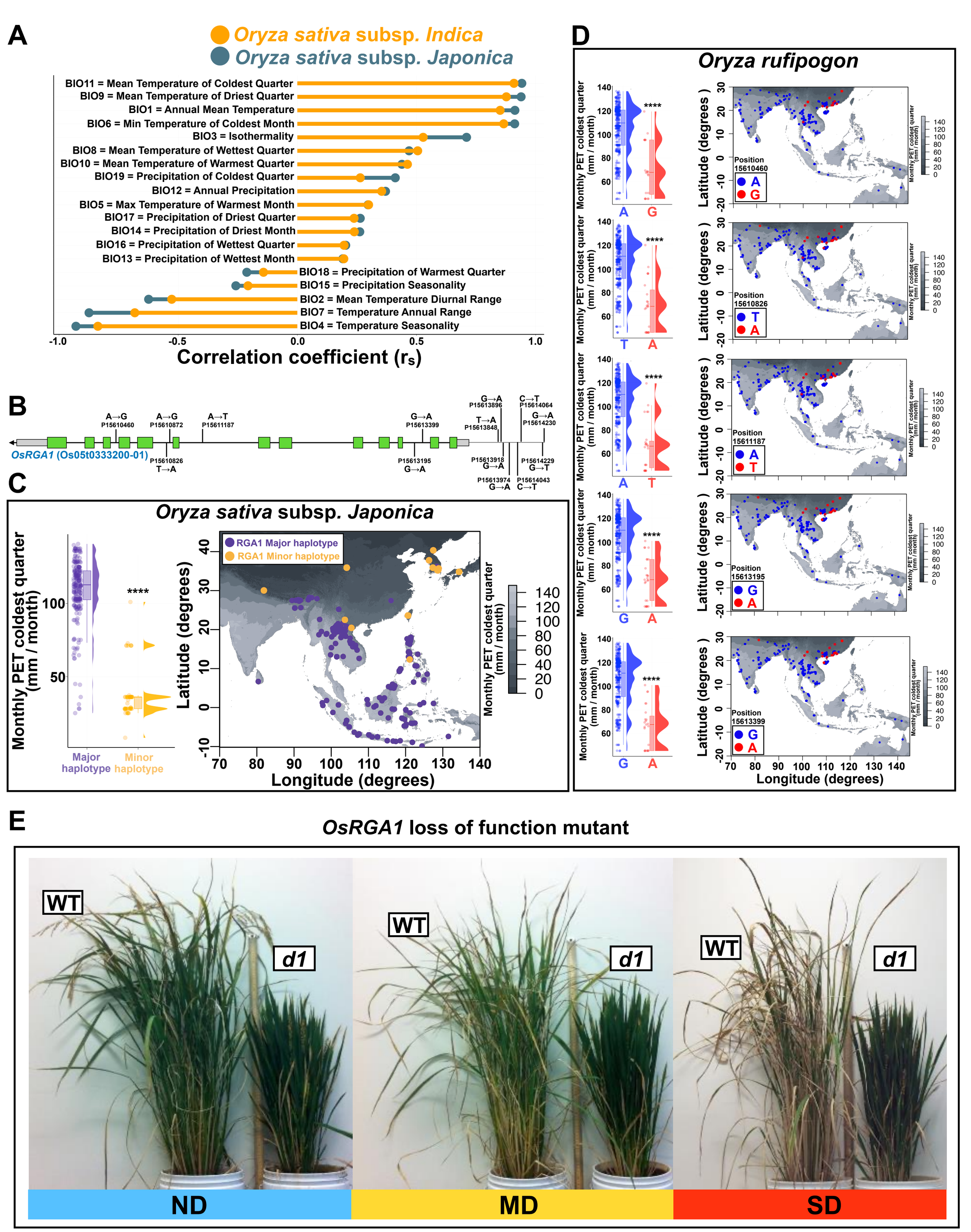

Figure 3. Haplotypic distribution of SNPs in *OsRGA1* is correlated with the mean potential evapotranspiration (PET) during the coldest quarter of the year that is typical of their local environment. A. We explored an environmental cline in a haplotype defined by six covarying intronic SNPs and eight upstream SNPs in *OsRGA1*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** Raincloud plots illustrate significantly different probability densities of the *association of the local environment* for the major and minor alleles of *OsRGA1*. The map of South-Eastern Asia shows the geographical distribution of the major and minor haplotypic variants of *OsRGA1*. The grayscale color gradient in this map illustrates the cline in PET during the coldest quarter. Purple dots and distributions represent landraces with the major allele in *OsRGA1*. Orange dots represent landraces with the minor allele in *OsRGA1*. **D.** The genetic variation present in 446 imputed sequences of *Oryza rufipogon* reveals that four of the six intronic SNPs uncovered in *OsRGA1* in Japonica landraces are also associated with PET during the coldest quarter of the year in this wild rice species. Blue dots and distributions represent landraces with the major allele of *OsRGA1*. Red dots represent landraces with the minor allele in *OsRGA1*. The same colors depict the haplotypic distribution on the map of South-Eastern Asia with a color gradient that represents the cline in this drought-related variant. **C.** The assessment of drought tolerance of the *d1* mutant available in the Japonica cultivar Taichung 65 (T65) confirms its drought tolerance (see Figure S2). **** ($p < 0.0001$; non-parametric Wilcoxon test).

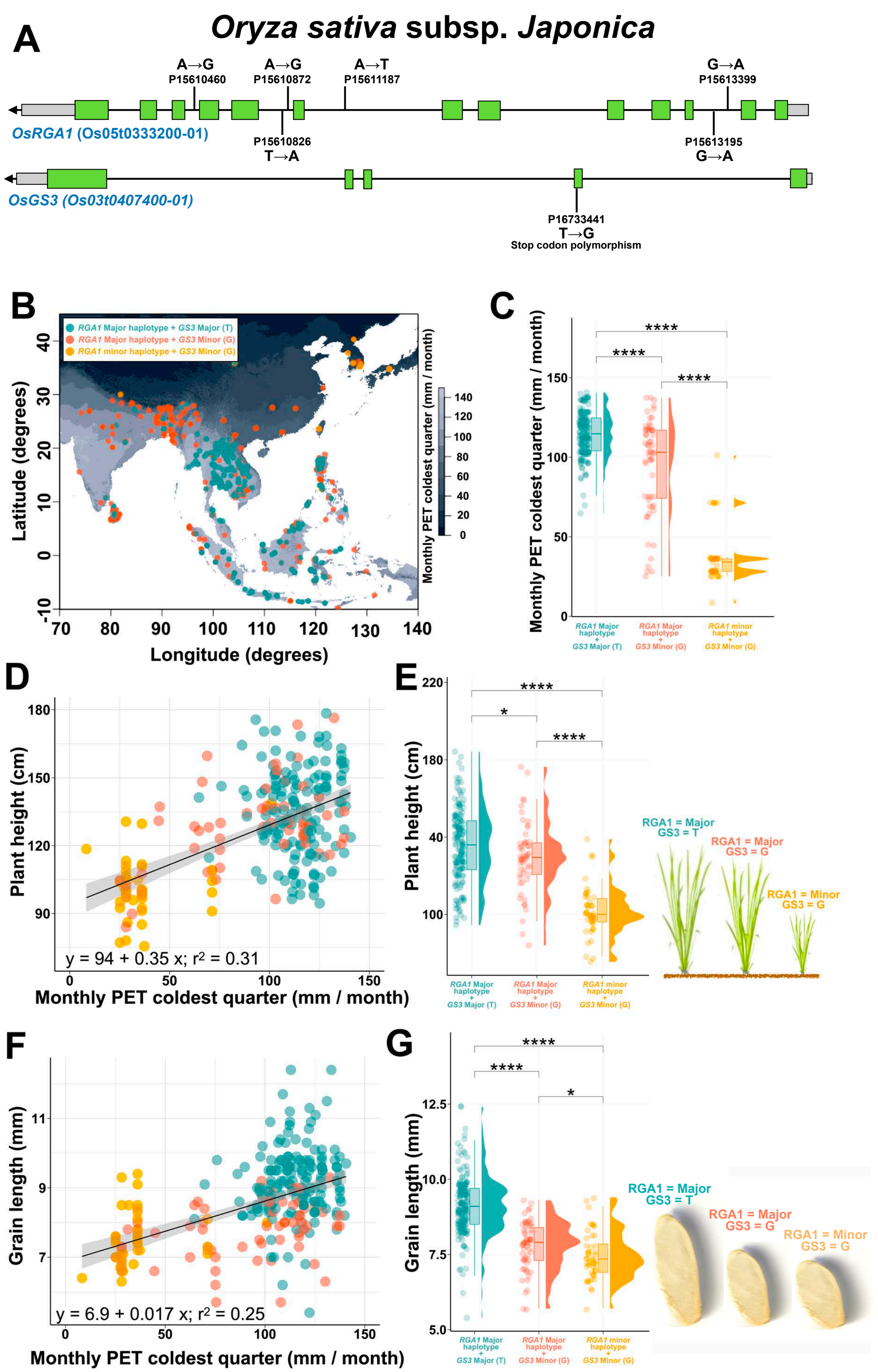

Figure 4. Haplotypic distribution of *OsRGA1* and a stop-causing codon SNP in *OsGS3* are correlated with mean potential evapotranspiration (PET) in the local environment and natural variation in plant height and seed length. A. We explored environmental and phenotypic clines that were associated with a haplotypic variant in *OsRGA1* (see Figure S3A) and a stop codon-causing SNP in *OsGS3*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** The map of South-Eastern Asia shows the geographic distribution of the major and minor haplotypic variants of *OsRGA1*. The grayscale color gradient illustrates the cline in PET during the coldest quarter. **C.** Raincloud plots illustrate significantly different probability densities of the *association of* the local environment for the major and minor alleles of *OsRGA1* in the presence/absence of the minor allele *of OsGS3*. **D.** There is a positive correlation between plant height and PET during the coldest quarter. Haplotypic variation in *OsRGA1* determines this correlation and is strongly associated with plants that display shorter stature as shown in Panel **E**, which is consistent with the well-reported dwarf phenotype of *OsRGA1* loss of function mutants (Oki et al., 2009; See Figure S2). **F.** There is also a positive correlation between grain length and PET during the coldest quarter. Haplotypic variation in *OsRGA1* and allelic variation in *OsGS3* determine this correlation and are strongly associated with plants that display shorter grains as shown in Panel **G**, consistent with the well-reported short-grain phenotype of *OsRGA1* and *OsGS3* loss of function mutants (Fan et al., 2006; See Figure S3). Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** ($p < 0.0001$).

Oryza sativa subsp. *Indica*

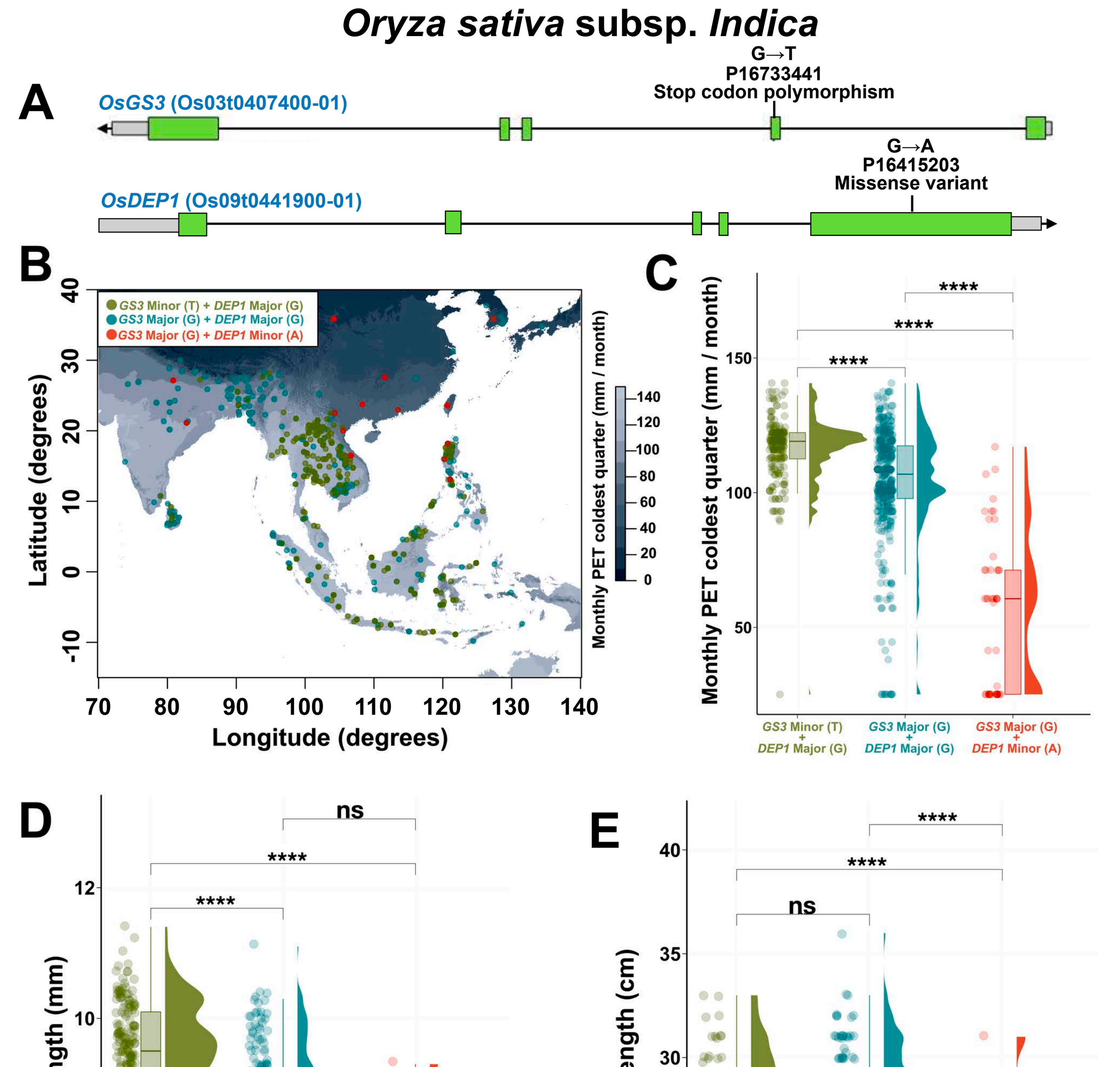

Figure 5. Haplotypic distribution of *OsDEP1* missense allele variants and a stop-causing codon SNP in *OsGS3* are correlated with the mean potential evapotranspiration in the local environment and natural variation in seed and panicle length A. We explored environmental and phenotypic gradients associated with landraces defined by a stop codon-causing SNP in *OsGS3* and a missense variant in *OsDEP1*. Grey indicates UTRs, green indicates exons, and black indicates introns. **B.** The map of South-Eastern Asia shows the geographic distribution of the major and minor haplotypic variants of the different *OsGS3*-*OsDEP1* allelic combinations. The grayscale color gradient illustrates the gradient in PET during the coldest quarter **C.** Raincloud plots illustrate significantly different probability densities of the association of the local environment for the major and minor alleles of *OsGS3* in the presence/absence of the minor allele *of OsDEP1*. Both B and C illustrate the association of the functional (major) *OsGS3* allele and the minor *OsDEP1* allele with lower PET regions. **D.** Raincloud plot illustrates the association of the *OsGS3* variant with grain length and the lack of association of the *OsDEP1* variant with grain length. **E.** Raincloud plot illustrates the association of the *OsDEP1* variant with panicle length and the lack of association of the *OsGS3* variant with panicle length. Pairwise nonparametric Wilcoxon tests were conducted to assess differences between alleles. **** ($p < 0.0001$).

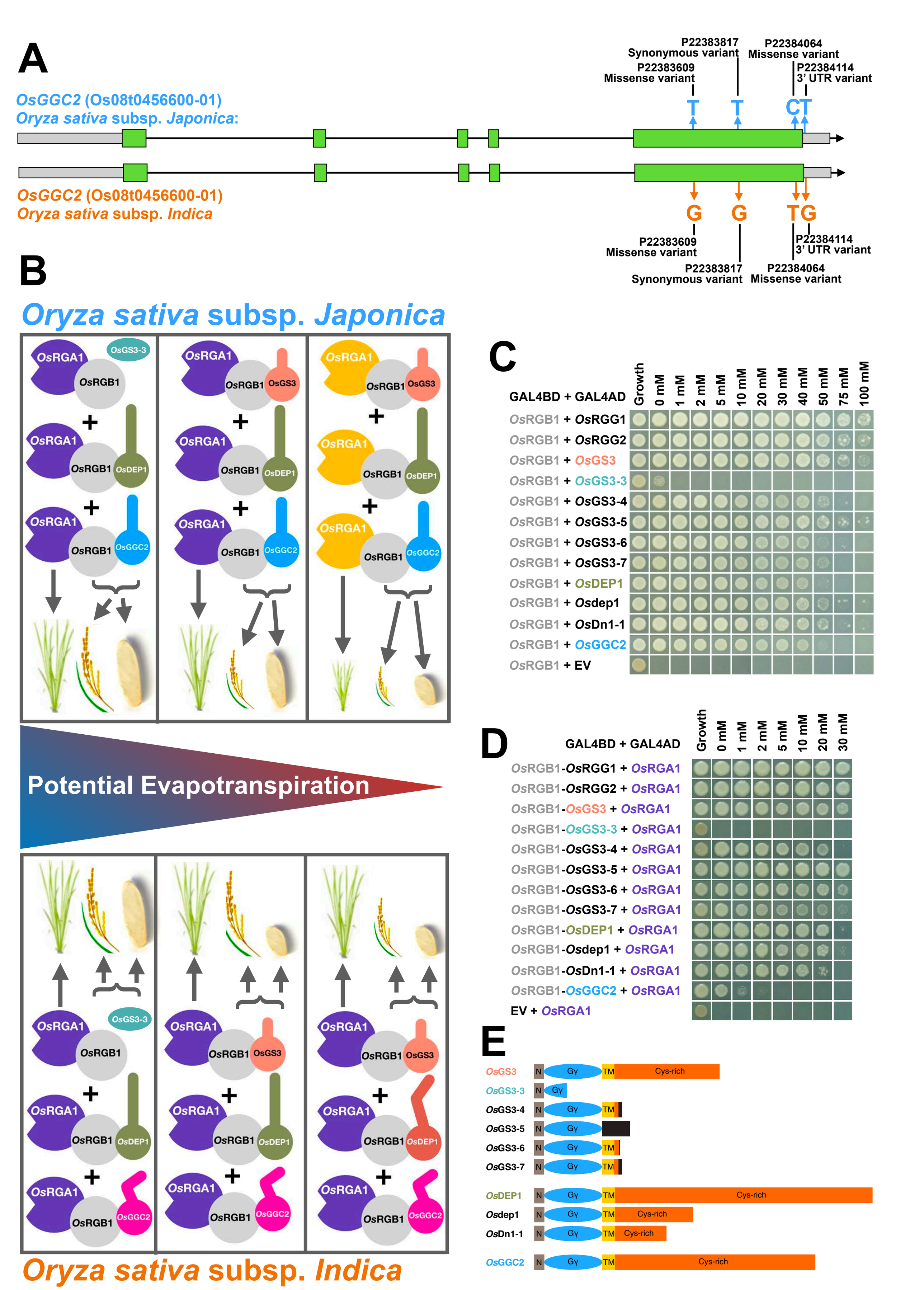

Figure 6. Genetic variation in heterotrimeric G proteins is associated with a precipitation gradient and the interplay of vegetative and reproductive traits. A. We uncover a set of three covarying variants in the last exon and one in the 3'UTR of OsGGC2 with one haplotype typical of Japonica, and the other typical of Indica varieties, that may affect the equilibrium of the Gγ subunits' competition for dimerization with Gβ. **B.** Cartoon summarizing our hypothesis on how the alternative combination of the different heterotrimeric G protein variants associated with PET in the local environment affects agronomic traits that may have an adaptive value. See Discussion for details. **C.** Yeast 2-hybrid assays between Gβ and Gγ, and **D.** Yeast 3-hybrid assays between Gβγ dimers and the canonical Gα, OsRGA1, were spotted on SC-Trp-Leu media as growth control, and SC-Trp-Leu-Met-His with the concentrations of 3-Amino-1,2,4-triazole (3-AT) indicated above each column. 3-AT is a competitive inhibitor of the *HIS3* reporter gene of the yeast 2-/3-hybrid systems and therefore allows for semi-quantitative assessment of interaction strength, with growth at higher 3-AT concentrations indicating a stronger interaction. All *Os*RGG1, *Os*RGG2, *Os*GS3, *Os*DEP1 and *Os*GGC2 proteins encoded by Gγ alleles were able to complex with OsRGB1 and *Os*RGA1 via *Os*RGB1, except *Os*GS3-3. **E.** Domain schematic of rice Type C Gγ subunits including known truncations. N = N-terminal sequence, Gγ = Gγ-like domain, TM = putative transmembrane domain, Cys-rich = C-terminal Cys-rich domain and dark gray boxes correspond to out of frame translated regions. See Figure S5A and S5B for multiple alignments of protein sequences.

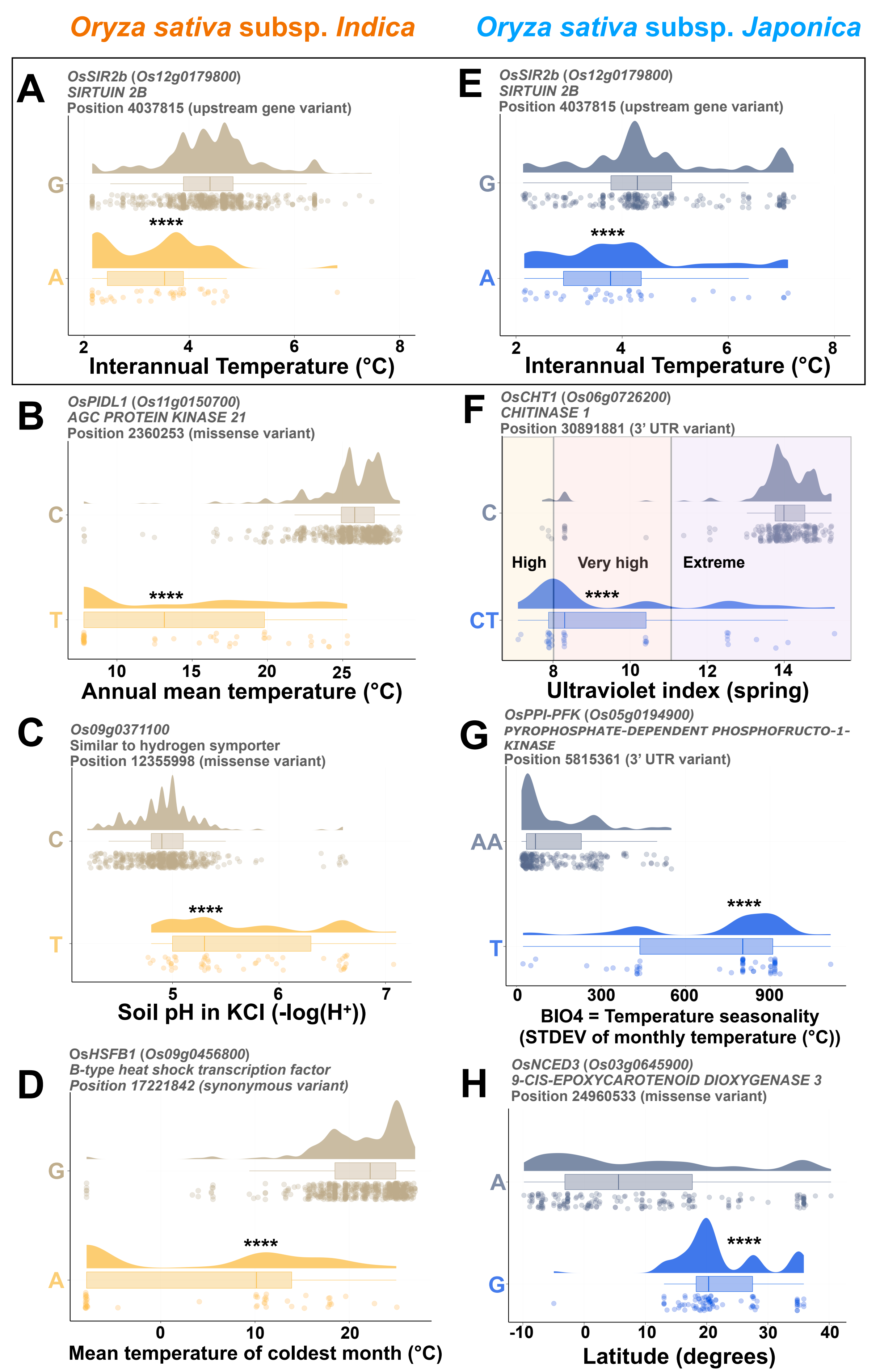

Figure 7. Examples of genotype x environment associations found using Oryza CLIMtools. Example associations for Indica landraces between **A**. an upstream gene variant in *OsSIRTUIN2B* and interannual temperature, **B.** a missense variant in *OsPIDL1* and annual mean temperature, **C.** a missense variant in *Os09g0371100* and soil pH. **D.** a synonymous variant in *OsHSFB1* and the mean temperature of the coldest month. In Japonica landraces, we uncovered an association between **E.** the same upstream variant that we identified in Indica landraces in *OsSIR2b* with the same climate parameter, interannual temperature, **F.** a 3'UTR variant in *OSCHIT1,* and the ultraviolet index during the spring season, **G.** another 3'UTR variant, here in *OsPPI-PFK* and temperature seasonality, and **H.** a missense variant in *OsNCED3* and latitude.