Article

The strength and pattern of natural selection on gene expression in rice

https://doi.org/10.1038/s41586-020-1997-2

Received: 2 May 2019

Accepted: 13 December 2019

Published online: 12 February 2020

Check for updates

Simon C. Groen[1], Irina Ćalić[2], Zoé Joly-Lopez[1], Adrian E. Platts[1], Jae Young Choi[1], Mignon Natividad[3], Katherine Dorph[1], William M. Mauck III[4], Bernadette Bracken[1,4], Carlo Leo U. Cabral[3], Arvind Kumar[3], Rolando O. Torres[3], Rahul Satija[1,4], Georgina Vergara[3], Amelia Henry[3], Steven J. Franks[2] & Michael D. Purugganan[1,5]✉

Levels of gene expression underpin organismal phenotypes[1,2], but the nature of selection that acts on gene expression and its role in adaptive evolution remain unknown[1,2]. Here we assayed gene expression in rice (*Oryza sativa*)[3], and used phenotypic selection analysis to estimate the type and strength of selection on the levels of more than 15,000 transcripts[4,5]. Variation in most transcripts appears (nearly) neutral or under very weak stabilizing selection in wet paddy conditions (with median standardized selection differentials near zero), but selection is stronger under drought conditions. Overall, more transcripts are conditionally neutral (2.83%) than are antagonistically pleiotropic[6] (0.04%), and transcripts that display lower levels of expression and stochastic noise[7–9] and higher levels of plasticity[9] are under stronger selection. Selection strength was further weakly negatively associated with levels of *cis*-regulation and network connectivity[9]. Our multivariate analysis suggests that selection acts on the expression of photosynthesis genes[4,5], but that the efficacy of selection is genetically constrained under drought conditions[10]. Drought selected for earlier flowering[11,12] and a higher expression of *OsMADS18* (*Os07g0605200*), which encodes a MADS-box transcription factor and is a known regulator of early flowering[13]—marking this gene as a drought-escape gene[11,12]. The ability to estimate selection strengths provides insights into how selection can shape molecular traits at the core of gene action.

To investigate the strength and pattern of selection on gene expression, we assessed transcriptome variation in two rice populations (Supplementary Tables 1–4)—one consisting of 136 varietal group 'Indica' accessions (comprising the indica and circum-aus subgroups) and the other of 84 varietal group 'Japonica' accessions (comprising the japonica and circum-basmati subgroups)—in a field experiment in the Philippines[3]. Replicates of each population, with three individuals per accession, were planted in a continuously wet paddy and a field that imposed intermittent drought (Fig. 1a, Extended Data Figs. 1–3). We used 3′-end mRNA sequencing[14] (Methods) to measure mRNA levels in leaf blades of the 1,320 plants at 50 days after sowing, corresponding to 17 days after withholding water in the dry field. We observed genetic variation in the levels of 15,635 widely expressed transcripts[15] (a broad-sense heritability of about 0.08 to about 0.95, false discovery rate (FDR)-adjusted $q < 0.001$) (Fig. 1b, Extended Data Figs. 2, 3, Supplementary Text, Supplementary Tables 5–8 provide overviews of genetic, environmental and interactive effects).

We focused our analyses on the Indica population, which is the predominant rice population grown globally[3]. We applied phenotypic selection analysis to measure the strength and pattern of selection on the levels of all 15,635 transcripts[4,5], using several complementary approaches. We initially measured total (direct and indirect) selection, and calculated univariate linear (S) and quadratic (C) selection differentials; these differentials estimate directional and stabilizing or disruptive selection, respectively, on the basis of the relationship between the trait value (transcript abundance) and fitness[4,5]. We considered total lifetime fitness through two multiplicative fitness components[16]: (i) flowering success, defined as flowering and producing filled grains before the end of the season[6,11,12] (which was only relevant under drought, owing to stress-related flowering delay and spikelet sterility)[11,12]; and (ii) fecundity, which was quantified as the numbers of filled grains produced (and which was relevant for both fields)[6,11,12] (Fig. 1a, Extended Data Fig. 1, Supplementary Tables 2, 9, Supplementary Notes 1, 2).

In wet conditions, selection on expression appeared to be weak. Transcriptome-wide selection strength was $|S|_{median} = 0.035$, with very few transcripts showing $|S| > 0.1$, which suggests that—for most genes—variation in expression is (nearly) neutral (Fig. 1c); this is similar to the distribution of selection strengths for higher-level organismal traits[4,17]. Directional selection (S) showed an overall bias for stronger and more-prevalent positive selection (a greater fitness with greater expression) than for negative selection (a lower fitness with greater expression) (7,973 versus 7,569 transcripts, with $S_{median} = 0.0361$ (for positive selection) and

[1]Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY, USA. [2]Department of Biological Sciences, Fordham University, New York, NY, USA. [3]International Rice Research Institute, Los Baños, The Philippines. [4]New York Genome Center, New York, NY, USA. [5]Center for Genomics and Systems Biology, NYU Abu Dhabi Research Institute, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates. ✉e-mail: mp132@nyu.edu

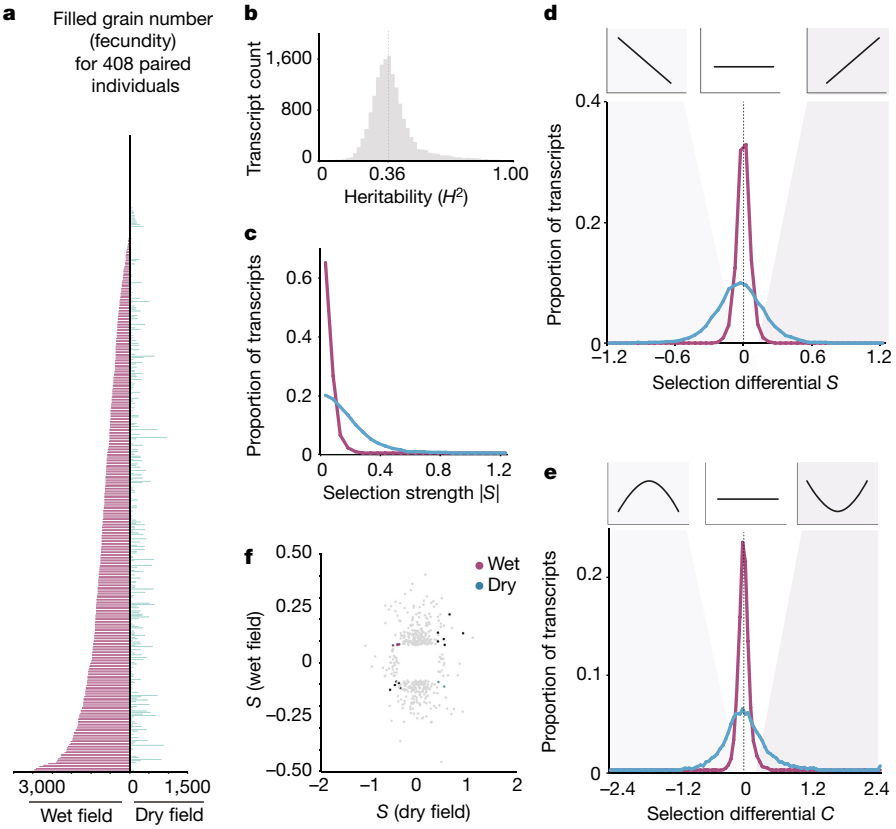


Fig. 1 | The strength and pattern of selection on heritable rice-leaf transcript levels differ across field environments. a, The Indica population showed significant genotype × environment ($G \times E$) variation in fitness as determined by measuring the multiplicative fitness components, fecundity (magenta and green in wet and dry conditions, respectively) and flowering success (zero filled grains indicate no flowering success); variation in flowering success is relevant only under drought conditions. Two-way analysis of variance (ANOVA), $G \times E$ $P = 4.68 \times 10^{-23}$, $n = 136$ accessions. **b**, Distribution of broad-sense heritability (H^2) for transcripts with significant expression polymorphism. Two-way ANOVA, genotype FDR-adjusted $q < 0.001$, $n = 136$ accessions. **c**, The strength of selection $|S|$ on gene expression when considering total lifetime fitness differed between wet (magenta) and dry (blue) conditions. Mann–Whitney *U*-test, two-sided $P < 0.001$, $n = 15{,}542$ transcripts. **d**, Positive directional selection (top right, $n = 7{,}973$ transcripts) was stronger than negative directional selection (top left, $n = 7{,}569$ transcripts) in wet conditions (magenta) (Mann–Whitney *U*-test, two-sided $P = 0.017$), and selection shifted to more extreme values under drought conditions (blue) (Kolmogorov–Smirnov test, two-sided $P < 0.001$, $n = 15{,}542$ transcripts). **e**, Patterns of stabilizing (top left) and disruptive (top right) selection were significantly more extreme under drought conditions. Kolmogorov–Smirnov test, two-sided $P < 0.001$, $n = 15{,}542$ transcripts. **f**, Patterns of conditional neutrality (light grey) and antagonistic pleiotropy (magenta and blue denote transcripts beneficial in wet and dry conditions, respectively) for gene expression. Black indicates transcripts that experienced selection in the same direction in both fields.

$S_{median} = -0.0345$ (for negative selection), respectively; Mann–Whitney *U*-test, $z = 2.38$, $P = 0.0173$). By contrast, *C* was negative (consistent with stabilizing selection) for the majority of transcripts (8,070 transcripts with $C < 0$ and 7,472 transcripts with $C > 0$)—although when *C* was positive, it tended to be stronger (Mann–Whitney *U*-test, $z = -3.28$, $P = 0.001$) (Fig. 1d, e, Supplementary Tables 10, 11). However, none of the transcript levels covaried significantly with fitness, for either *S* or *C*, after Bonferroni correction ($P < 3.2 \times 10^{-6}$). This suggests that—at microevolutionary timescales—variation in gene expression is (nearly) neutral or exhibits very weak stabilizing selection. This contrasts with stronger directional and stabilizing selection at larger evolutionary timescales[18].

Selection was stronger ($|S|_{median} = 0.1367$) under drought conditions than under wet conditions (Mann–Whitney *U*-test, $z = 99.99$, $P < 0.0001$) (Fig. 1c). Although no individual transcript breached the Bonferroni threshold, *S* and *C* exhibit more extreme values under drought conditions, indicating drought-induced shifts in both the strength and pattern of selection (Kolmogorov–Smirnov test, $D = 0.327$ (for *S*) and $D = 0.269$ (for *C*), $P < 0.0001$) (Fig. 1d, e, Extended Data Fig. 4, Supplementary Text show results for fitness components under drought conditions). We examined selection on expression across environments and found patterns of antagonistic pleiotropy (*S* exhibits opposite directionality between environments) for 6 transcripts (about 0.04%) and conditional neutrality (significant *S* in one environment) for 443 transcripts (2.83%) (Fig. 1f). Compared to expectations that are based on chance alone, conditional neutrality appears much more common than antagonistic pleiotropy under our conditions[6] (Supplementary Table 12). This result indicates a general lack of trade-offs at the gene-expression level, and suggests a mechanistic explanation for the lack of yield penalty on drought tolerance in modern rice breeding lines[12].

To identify factors that shape rates of microevolutionary change in gene expression, we performed partial correlation analysis with factors that influence macroevolutionary rates of expression divergence[7,8,19–21] (Supplementary Table 13). We focused on $|S|$ because this value is directly proportional to the response to selection[5], which is a measure of microevolution[22]. Relative expression level and stochastic expression noise were negatively correlated with $|S|$ (Pearson's partial $r < -0.119$, $P < 5.13 \times 10^{-48}$) (Fig. 2a, b, Supplementary Table 14), suggesting fitness is buffered—to some extent—for expression variation in highly expressed genes, as well as for high stochasticity in transcript abundance[9]. However, we observed that accessions with higher genome-wide levels of expression stochasticity tend to have a lower fecundity[23,24] (Spearman's $\rho < -0.174$, $P < 0.05$) (Fig. 2c, Extended Data Fig. 5, Supplementary Table 15). $|S|$ also correlated positively with tissue specificity τ (Pearson's partial $r > 0.024$, $P < 0.01$) (Fig. 2a, b), and for fecundity with expression plasticity (differential gene

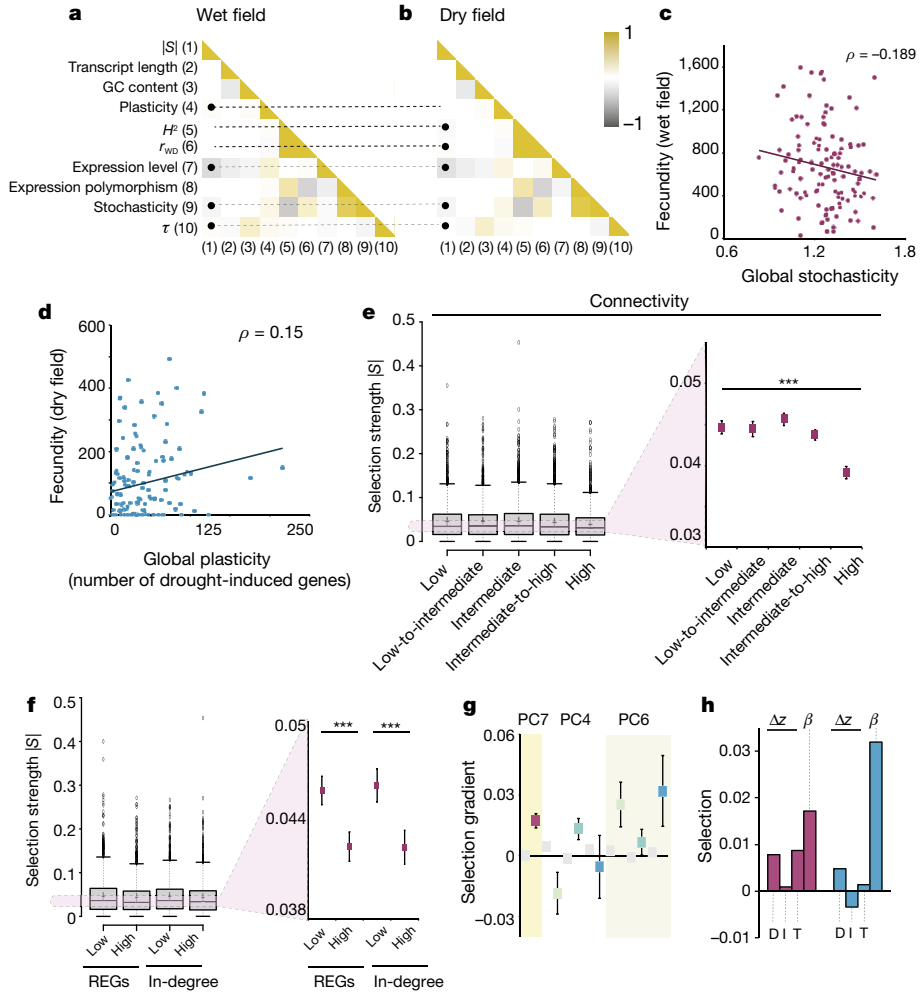


Fig. 2 | Gene-expression level, stochasticity, plasticity, tissue specificity and connectivity influence microevolutionary rates of expression change. **a**, **b**, Partial correlation analyses of factors that negatively (grey) and positively (mustard) influence selection strength $|S|$ on gene expression in wet (**a**) and dry (**b**) conditions. Dots indicate statistical significance for Pearson's partial r correlations; t-test, $P < 0.05$, $n = 14{,}753$ transcripts (Supplementary Table 14). **c**, Global expression stochasticity limits fecundity. Spearman's $\rho = -0.189$, t-test, $P = 0.036$, $n = 123$ accessions. **d**, Global expression plasticity correlates with fecundity under drought conditions. Spearman's $\rho = 0.15$, t-test, $P = 0.041$, $n = 135$ accessions. **e**, $|S|$ is bounded by expression connectivity. Kruskal–Wallis test, two-sided $P = 0.000017$, $n = 12{,}502$ transcripts. Left, box plot with centre line = median, cross = mean, box limits = upper and lower quartiles, whiskers = 1.5 × interquartile range and points = outliers. Right, mean ± s.e.m. **f**, $|S|$ is limited by regulatory constraints, as assessed through numbers of *cis*-regulatory promoter elements (REGs) ($n = 3{,}907$ transcripts; Mann–Whitney U-test, $P = 0.0061$) and transcription factors regulating a gene (in-degree) ($n = 2{,}905$ transcripts; Mann–Whitney U-test, $P = 0.0061$). Left, boxes and whiskers as in **e**. Right, mean ± s.e.m. **g**, Linear (β) (coloured) and quadratic (γ) (grey) selection gradients (± s.e.) on suites of transcripts as principal components (eigengenes). $n = 408$ plants. β values are for total lifetime fitness in wet (magenta) and dry (blue) conditions, and for flowering success (lime) and fecundity (green) under drought conditions. **h**, Prediction of the outcome of selection (Δz) for $PC7_{wet}$ and $PC6_{dry}$ in **g**, indicating that the efficacy of selection under drought is limited (total change (T) lower than β for total lifetime fitness) through genetic constraints (indirect or correlated change (I) and direct change (D) have opposite signs). β values are as in **g** for comparison. Extended Data Tables 1, 2 provide more details. P values are two-sided.

expression between the two environments; Pearson's partial $r > 0.017$, $P < 0.05$) (Fig. 2a, Extended Data Fig. 5). This is consistent with previous reports that tissue specificity can minimize pleiotropic constraints on selection[21], and expression plasticity can affect the efficacy of selection[19,20]. Supporting the importance of plasticity, accessions that induce expression of more genes under drought conditions experience fitness benefits (Spearman's $\rho = 0.15$, $P = 0.041$) (Fig. 2d, Supplementary Table 16).

Gene expression is regulated through networks of transcription factors that interact with *cis*-regulatory DNA elements[9], and these relationships have been shaped by past selection. Highly connected transcripts in regulatory networks should be controlled by more transcription factors[9,25,26] and have evolved to reduce the effects of expression variation on fitness, contributing to robustness[9]. Supporting this hypothesis, fitness was less strongly associated with the expression of genes with higher connectivity (Kruskal–Wallis test, $H \geq 18.94$, $P < 0.001$), numbers of known *cis*-regulatory DNA elements and transcriptional regulators (Mann–Whitney U-test, $z \geq 2.74$, $P < 0.05$) (Fig. 2e, f, Extended Data Fig. 5, Supplementary Table 17).

Because interactive network effects appear to curb the strength of phenotypic selection on gene expression, we hypothesize that genetic correlations between multivariate suites of transcripts may constrain the outcome of selection. We performed dimensional reduction of the transcriptome data using principal component (PC) analysis, and considered the principal components that explain >0.5% of overall variance as suites of transcripts in a multivariate selection analysis[5] (Supplementary Table 18). We estimated linear (β) and quadratic (γ) selection gradients, which together measure the strength and pattern of direct (instead of total) selection on a trait[4,5]. Quadratic selection was generally weak, but PC7 showed significant positive directional selection under wet conditions ($PC7_{wet}$ $\beta = 0.017$, $P = 1.44 \times 10^{-6}$). Under drought conditions, PC6 displayed positive directional selection for flowering success

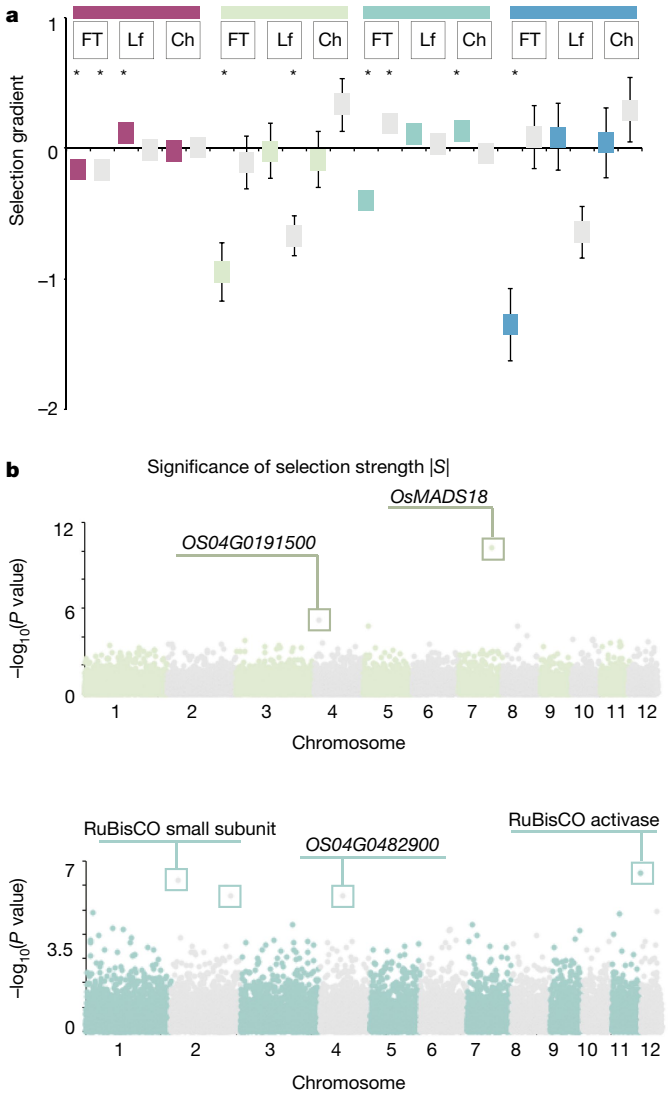


Fig. 3 | Transcripts under selection could affect fitness through regulating early growth vigour and flowering time. a, Wet conditions (magenta) impose stabilizing selection on flowering time (FT) and positive directional selection on growth vigour (leaf area, Lf) (*t*-tests). Drought induces strong, positive flowering-success (*z*-test) and total-lifetime-fitness selection (*t*-test) on early flowering (lime and blue, respectively), and leads to weaker fecundity selection (green) (*t*-test) on chlorophyll concentration (Ch), early flowering and early growth vigour (Supplementary Table 20). Linear (β) and quadratic (γ) selection gradients are denoted by coloured and grey markers, respectively. Mean ± s.e.m., $n = 408$ plants; asterisks indicate selection-gradient significance, two-sided, unadjusted $P < 0.05$. **b**, Two transcripts with significant linear selection differentials ($n = 408$ plants; *z*-test, two-sided, Bonferroni-adjusted $P < 0.05$ for 15,565 transcripts) for flowering success under drought conditions (lime) may promote drought escape through regulating early flowering; absolutized transcript–trait correlations are significant (Pearson's $|r| > 0$, *t*-test, two-sided, unadjusted $P < 0.01$) (Extended Data Fig. 6). Three of four transcripts with significant selection differentials ($n = 408$ plants; *t*-test, two-sided, Bonferroni-adjusted $P < 0.05$ for 15,343 transcripts) for fecundity under drought conditions (green) may affect fitness by influencing photosynthesis and—consequently—early growth vigour; transcript–trait correlations are significantly positive (Pearson's $r > 0$, *t*-test, two-sided, unadjusted $P < 0.01$) (Extended Data Fig. 6, Supplementary Text).

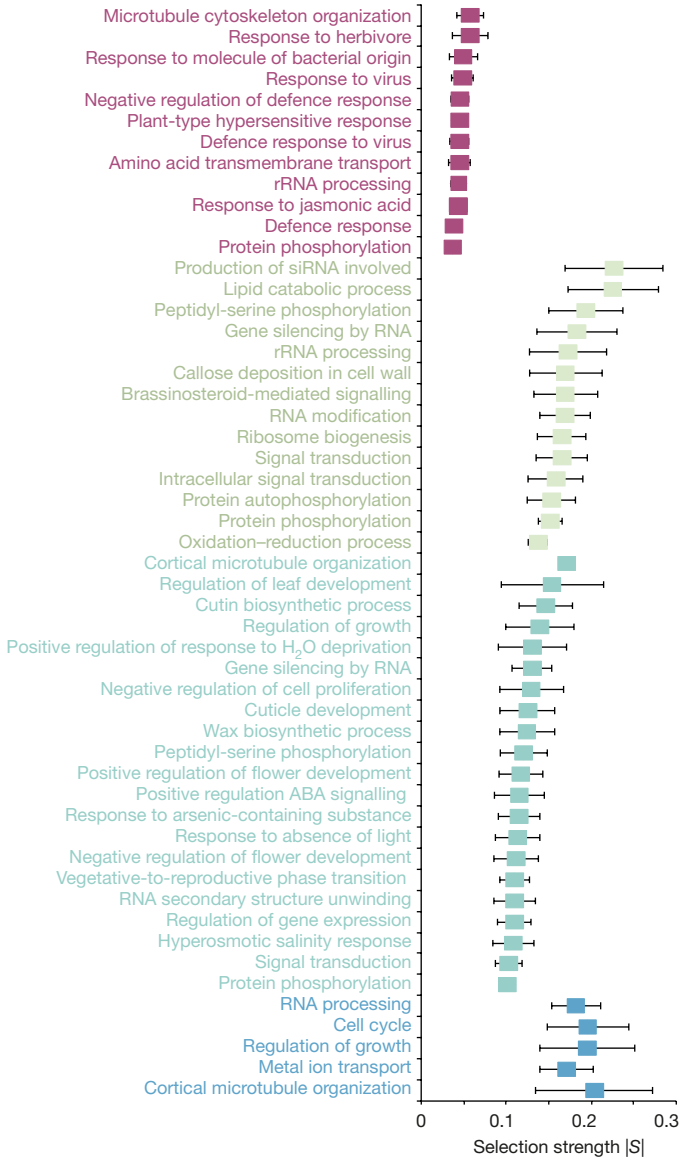


Fig. 4 | Selection targets expression patterns in different biological processes in wet and dry conditions. Biological processes that experience stronger selection appear to be linked to growth and defence for total lifetime fitness in wet conditions (magenta). Under drought conditions, biological processes that experience stronger selection are involved in growth for total lifetime fitness (blue), in early growth vigour and flowering for fecundity (green), and in regulatory processes for flowering success (lime). Only biological processes with $n \geq 20$ transcripts and with significantly higher median selection strengths $|S|_{median}$ than the transcriptome-wide median are shown (nonoverlapping 95% confidence intervals).

($PC6_{dry}\beta = 0.025$, $P = 0.023$), and was marginally non-significant for total lifetime fitness ($\beta = 0.032$, $P = 0.07$) (Fig. 2g, Extended Data Tables 1, 2). Furthermore, fecundity selection under drought conditions was positive for PC4 ($PC4_{dry}\beta = 0.017$, $P = 0.014$), whereas selection for flowering success had the opposite effect—albeit marginally non-significant ($\beta = -0.019$, $P = 0.07$) (Fig. 2g). We can predict the outcomes of selection and evolutionary constraints on gene expression using the breeder's equation[10]. Although the principal components as multivariate suites of transcripts were uncorrelated at the phenotypic level, they genetically covaried given that individual plants were accompanied by two additional genetically identical plants in the population. Despite stronger selection under drought conditions, evolutionary responses to stress were weak owing to constraints (as evidenced by the opposite signs of the direct and indirect responses to selection) that arose from genetic correlations between gene groups (Fig. 2h, Extended Data Table 1).

Gene expression presumably influences fitness through regulating phenological, morphological or physiological traits, and we measured

three of these traits: (i) flowering time, (ii) leaf area and (iii) chlorophyll concentration (all of which display significant genetic variation) (Fig. 3a, Supplementary Tables 2, 19). We find stabilizing selection for flowering time and positive directional selection for leaf area in wet conditions. Drought selected for earlier flowering, and leaf area and chlorophyll concentration experienced positive fecundity selection (Fig. 3a, Supplementary Table 20). We assessed whether selection on these traits could have been driven by selection on suites of transcripts. In the multivariate analysis, translation- and photosynthesis-related gene ontology terms showed loading-value enrichment on principal components with $\beta > 0$ (Supplementary Table 21). Moreover, the levels of several photosynthesis-related transcripts correlated with leaf area, chlorophyll content and fitness (Fig. 3b, Extended Data Fig. 6, Supplementary Tables 10, 11, 22), indicating their expression may increase fitness through promoting growth vigour[11,12] (Supplementary Text). We also ranked biological processes by median selection strengths $|S|$ from the univariate analyses. We observed different rankings between dry and wet conditions (Mann–Whitney U-test, $z = -13.51$, $P < 0.001$) (Fig. 4): plants in wet conditions showed a relatively strong selection on genes related to growth and defence, whereas under drought conditions plants showed a stronger selection associated with genes involved in water deprivation responses, growth and flowering (Fig. 4, Supplementary Table 23).

Flowering time was the trait under strongest selection in drought conditions. Interestingly, expression of only a single gene (*OsMADS18*)—which encoded the transcription factor OsMADS18—was both under selection for flowering success after Bonferroni correction ($S = 0.77$, $P = 5.99 \times 10^{-11}$), and coming close to significance for total lifetime fitness ($S = 0.914$, $P = 3.81 \times 10^{-6}$) (Fig. 3b). Increased expression of *OsMADS18* was tightly linked with early flowering (Extended Data Fig. 6), which has previously been functionally validated[13]. Furthermore, the gene sits in a major quantitative trait locus (QTL) for flowering and yield under drought conditions across *O. sativa*[27,28], and the expression of this gene is also under relatively strong selection for flowering success under drought conditions in our Japonica population (Supplementary Table 24), suggesting *OsMADS18* is an important drought-escape gene[11,12].

To examine the genetic architecture of fitness-related genes, we conducted a genome-wide association study that mapped expression QTLs (eQTLs) for transcripts and expression principal components with significant selection differentials or gradients in our Indica populations[29], using 179,634 randomly sampled single-nucleotide polymorphisms (SNPs)—or about 1 SNP every 2.2 kb. We observe no significant *cis*-eQTLs after Bonferroni correction ($P < 2.78 \times 10^{-7}$). However, *trans*-eQTLs appeared for three of eight transcripts under drought-induced selection (Extended Data Fig. 7, Supplementary Tables 25–27). Although our sample size limits mapping power, these findings suggest *trans*-acting loci have key roles in the expression variation of fitness-related genes[29]. We also mapped fitness component traits, and found no significant QTLs (Supplementary Tables 25–27). Furthermore, taking the top 0.5% of SNPs with the strongest association with fitness, we observed no enrichment for genes with high selection differentials in 100-kb regions surrounding these SNPs ($\chi^2 = 0.088$, $P = 0.77$) (Extended Data Fig. 8, Supplementary Table 27). This suggests that, although there may be strong selection for expression on particular genes, fitness continues to behave (as expected) as a polygenic trait[29].

Gene expression is a fundamental molecular mechanism that is essential for trait development. Previous studies have focused on long-term transcriptome evolution across species[1,2,7,18]; our approach using phenotypic selection analysis demonstrates that measuring the strength and type of ongoing selection on individual genes across the entire genome is possible. However, our study has limitations: we are measuring selection on a snapshot of leaf gene expression, and it would be interesting to see whether selection strength varies across tissues and developmental time points[30]. If so, then the final effect of gene expression on adaptation may arise from the integration of expression over the entire life cycle[30]. Moreover, examining selection across more environments relevant for plants may provide further insights into how gene expression evolves[1,2,30]. Nevertheless, our work opens up the possibility of dissecting the intrinsic and extrinsic factors that drive adaptive evolution via regulated gene expression, providing crucial links between adaptation at the molecular and organismal levels.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-1997-2.

1. Fay, J. C. & Wittkopp, P. J. Evaluating the role of natural selection in the evolution of gene regulation. *Heredity* **100**, 191–199 (2008).
2. Romero, I. G., Ruvinsky, I. & Gilad, Y. Comparative studies of gene expression and the evolution of gene regulation. *Nat. Rev. Genet.* **13**, 505–516 (2012).
3. Wing, R. A., Purugganan, M. D. & Zhang, Q. The rice genome revolution: from an ancient grain to green super rice. *Nat. Rev. Genet.* **19**, 505–517 (2018).
4. Kingsolver, J. G. et al. The strength of phenotypic selection in natural populations. *Am. Nat.* **157**, 245–261 (2001).
5. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).
6. Anderson, J. T., Lee, C. R., Rushworth, C. A., Colautti, R. I. & Mitchell-Olds, T. Genetic trade-offs and conditional neutrality contribute to local adaptation. *Mol. Ecol.* **22**, 699–708 (2013).
7. Lemos, B., Bettencourt, B. R., Meiklejohn, C. D. & Hartl, D. L. Evolution of proteins and gene expression levels are coupled in *Drosophila* and are independently associated with mRNA abundance, protein length, and number of protein–protein interactions. *Mol. Biol. Evol.* **22**, 1345–1354 (2005).
8. Lehner, B. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.* **4**, 170 (2008).
9. MacNeil, L. T. & Walhout, A. J. Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome Res.* **21**, 645–657 (2011).
10. Conner, J. & Via, S. Natural selection on body size in *Tribolium*: possible genetic constraints on adaptive evolution. *Heredity* **69**, 73–83 (1992).
11. Franks, S. J. Plasticity and evolution in drought avoidance and escape in the annual plant *Brassica rapa*. *New Phytol.* **190**, 249–257 (2011).
12. Kumar, A. et al. Breeding high-yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. *J. Exp. Bot.* **65**, 6265–6278 (2014).
13. Fornara, F. et al. Functional characterization of *OsMADS18*, a member of the *AP1/SQUA* subfamily of MADS box genes. *Plant Physiol.* **135**, 2207–2219 (2004).
14. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
15. Ayroles, J. F. et al. Systems genetics of complex traits in *Drosophila melanogaster*. *Nat. Genet.* **41**, 299–307 (2009).
16. Conner, J. Field measurements of natural and sexual selection in the fungus beetle, *Bolitotherus cornutus*. *Evolution* **42**, 736–749 (1988).
17. Hoekstra, H. E. et al. Strength and tempo of directional selection in the wild. *Proc. Natl Acad. Sci. USA* **98**, 9157–9160 (2001).
18. Nourmohammad, A. et al. Adaptive evolution of gene expression in *Drosophila*. *Cell Rep.* **20**, 1385–1395 (2017).
19. Ghalambor, C. K. et al. Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature* **525**, 372–375 (2015).
20. Kenkel, C. D. & Matz, M. V. Gene expression plasticity as a mechanism of coral adaptation to a variable environment. *Nat. Ecol. Evol.* **1**, 0014 (2016).
21. Zhang, L. & Li, W. H. Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Mol. Biol. Evol.* **21**, 236–239 (2004).
22. Hendry, A. P. & Kinnison, M. T. The pace of modern life: measuring rates of contemporary microevolution. *Evolution* **53**, 1637–1653 (1999).
23. Duveau, F. et al. Fitness effects of altering gene expression noise in *Saccharomyces cerevisiae*. *eLife* **7**, e37272 (2018).
24. Jimenez-Gomez, J. M., Corwin, J. A., Joseph, B., Maloof, J. N. & Kliebenstein, D. J. Genomic analysis of QTLs and genes altering natural variation in stochastic noise. *PLoS Genet.* **7**, e1002295 (2011).
25. Plessis, A. et al. Multiple abiotic stimuli are integrated in the regulation of rice gene expression under field conditions. *eLife* **4**, e08411 (2015).
26. Wilkins, O. et al. EGRINs (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. *Plant Cell* **28**, 2365–2384 (2016).
27. Huang, X. et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* **44**, 32–39 (2011).
28. Wang, Y. et al. Background-independent quantitative trait loci for drought tolerance identified using advanced backcross introgression lines in rice. *Crop Sci.* **53**, 430–441 (2013).
29. Liu, X., Li, Y. I. & Pritchard, J. K. *Trans* effects on gene expression can drive omnigenic inheritance. *Cell* **177**, 1022–1034.e6 (2019).
30. Zaidem, M. L., Groen, S. C. & Purugganan, M. D. Evolutionary and ecological functional genomics, from lab to the wild. *Plant J.* **97**, 40–55 (2019).

Methods

Representative studies from the literature were used to determine sample size[11,24,31]. The investigators were blinded to the genetic identity of individuals in the experiment during sampling, sample processing and outcome assessment. The planting order of individuals was randomized according to an alpha lattice design.

Plant material

Plants of 220 *O. sativa* accessions—136 accessions from the Indica varietal group (including the circum-aus and indica subgroups) and 84 accessions from the Japonica varietal group (including the circum-basmati, and temperate, sub-tropical and tropical japonica subgroups) (Supplementary Table 1), consisting of landraces and breeding lines and two additionally replicated checks (accessions IR64 and Sahod Ulan 1)—were selected for the experiment[12,25,26,32–38]. Seeds for all accessions were obtained from the International Rice Genebank Collection at the International Rice Research Institute (IRRI), and from IRRI's Rice Breeding Platform – Breeding for marginal environments.

Establishment of the field experiment

The field experiment was conducted during the 2016 dry season at IRRI in Los Baños, the Philippines. Two to three grams of seed from each of the accessions was sown onto a seed bed on 4 January 2016, and at 17 days after sowing (DAS) seedlings were pulled and transplanted into two different experimental fields. The first, known as UJ (14° 008′ 41.5″ N, 121° 015′ 53.8″ E), remained flooded as a wet paddy field environment. The second, known as UR and located in a rain-out shelter, (14° 008′ 33.3″ N, 121° 016′ 03.4″ E), was maintained flooded until 33 DAS, at which time irrigation was stopped and the field was drained to initiate the drought-stress treatment. This dry field was rewatered by flooding at 53, 64 and 91 DAS to let the plants experience intermittent drought throughout the remainder of the season.

The experiments were arranged in an alpha lattice design with each accession planted in 3 replicates with 1 plant per hill in single 2-m rows with 0.2-m × 0.2-m spacing for a total of 1 focal plant (in the fourth hill) and 9 neighbouring plants per plot. Basal fertilizer was applied at 30 DAS using complete fertilizer (14-14-14) at the rate of 50 kg ha^{-1} each of N_2, P_2O_5 and K_2O. Manual weeding was done regularly in both treatments. Cymbush (1 l ha^{-1}) and Cartap (0.96 kg ha^{-1}) were applied at 37 DAS, and Provado (1.92 l ha^{-1}) was applied at 40 DAS and again at 60 DAS to control insect pests in both treatments.

Soil moisture levels in the dry field were monitored by recording soil water potential using nine tensiometers (Soilmoisture Equipment) installed at a depth of 30 cm in each replicate, and volumetric soil moisture by frequency domain reflectometry (Diviner 2000, Sentek) at 10-cm depth increments through 70-cm PVC tubes installed at 9 locations in the experimental area.

Leaf tissue collection for mRNA sequencing

Leaf sampling was performed at 50 DAS on the focal plant in all plots of the wet and dry fields from 10:00 to 12:00 (4 h after dawn) as previously described[25]. The aim was to collect leaf samples in the shortest amount of time possible to minimize the effects of physiological changes patterned with the circadian rhythm of the plants. Four pairs of technicians were assigned to collect leaves, and the wet and dry fields were sampled simultaneously by different teams working in the same order by replicate and plot.

During collection, two fully expanded leaves were selected for sampling. Approximately 12 cm of leaf length were cut into small pieces and submerged into 4 ml chilled RNALater solution in 5-ml screw-cap tubes. Scissors used for leaf sampling were wiped with 70% ethanol to avoid contamination between plots. The tubes with the collected leaf samples were placed on ice in a styrofoam ice chest, then transferred to a cold room at −4 °C overnight. A total of 1,320 tubes were used for the collections in the wet and dry fields. Leaf samples from each of the 5-ml tubes were then transferred into pairs of 2-ml tubes, then stored at −80 °C. One 2-ml tube of each of the 1,320 pairs was sent to New York University in liquid -nitrogen dry shippers for long-term storage and further processing for mRNA sequencing.

Higher-level trait measurements

A set of physiological, morphological and phenological measurements was conducted to assess individual and genotypic differences in drought response. In both the wet and dry fields, ground cover images were taken from each focal plant at 52 DAS using a high-resolution digital camera at the same height from the ground. Images were processed and analysed using ImageJ software version 1.52 to determine the leaf area (leaf area index or per cent groundcover)[39]. For images in which other green material was present, GNU Image Manipulation Program (GIMP) software version 2.10.0 was used to select the leaves of the designated plant to determine the leaf area index (www.gimp.org). Chlorophyll concentration (chlorophyll content index) (Apogee Instruments) was measured on one leaf of each focal plant at 49 DAS in the dry field, and 50 DAS in the wet field. Flowering time was recorded as the day on which 50% of plants in a plot flowered; these plants included the focal plant and the nine neighbouring plants.

Grain harvesting and processing

To avoid grain loss from shattering, individual panicles were harvested separately from the focal plant in each plot as they reached maturity, for a total of 1,320 plants harvested individually. Filled, partially filled and unfilled grains were sorted and counted with the use of a seed counter (Hoffman Manufacturing) except for seeds with awns, which were counted manually.

Preparation of RNA for library construction

Frozen leaf samples were thawed at room temperature and blotted briefly on a KimWipe for removal of excess RNALater. The leaf tissue was then flash-frozen in liquid nitrogen and pulverized in liquid nitrogen with a pre-cooled mortar and pestle (CoorsTek), and frozen again at −80 °C. Total RNA was extracted from the pulverized bulk tissue using the RNeasy Plant Mini Kit according to manufacturer's protocol (Qiagen). The RNA was quantified on a Qubit (Invitrogen), after which the quality of the RNA was assessed on an Agilent BioAnalyzer (Agilent Technologies). The total RNA preps were then stored at −80 °C in nuclease-free water.

RNA-sequencing library construction and sequencing

Total RNA for each sample was processed individually according to a barcoded, plate-based 3′-end mRNA sequencing (3′ mRNA-seq) protocol that presents a modification of the SMART-seq2 and SCRB-seq protocols[40–42]. In brief, aliquots of total RNA from all samples were transferred individually into wells in 96-well-plates, and diluted to a concentration of 10 ng in a total of 50 µl nuclease-free water. Then, the total RNA was mixed with 5 × Maxima reverse transcription buffer, dNTP mixture, RNase inhibitors (NxGen RNase Inhibitor, Lucigen, at 40 µg $µl^{-1}$) and water. We reverse-transcribed the mRNAs using Superscript II Reverse Transcriptase (Thermo Fisher Scientific), and amplified cDNAs for each sample in individual wells using the Smart-seq2 protocol[41], with a custom modification in which a 12-bp well barcode was included in the 3′ end reverse transcriptase primer using barcoded oligonucleotides from the SCRB-seq protocol[42]. This enabled us to perform multiplexed pooling of 96 samples before library preparation with the Nextera XT DNA sample prep kit (Illumina) and returned 3′-biased cDNA fragments, similar to the Drop-seq protocol[14]. Each library consisted of a pool of 96 sister samples—that is, 48 samples from the wet field environment were matched with samples from the same plot numbers in the dry field environment. We quantified the 14 cDNA libraries on an Agilent BioAnalyzer and sequenced them at 2 × 50

bases on the Illumina NextSeq 500 using the following settings: read 1 was 20 bp (bases 1–12, well barcode; bases 13–20, unique molecular identifier (UMI)), and read 2 (paired end) was 50 bp.

RNA-sequencing data processing

The 3′ mRNA-seq read data were quantified according to the McCarroll Laboratory Drop-seq Cookbook using Drop-seq tools version 1.12 (J. Nemesh and A. Wysoker, https://github.com/broadinstitute/Drop-seq/releases), a wrapper for aligning and parsing both reads and their embedded barcodes with the STAR aligner version 020201. The reference genome used by STAR was Nipponbare IRGSP 1.0 (GCF_001433935.1) including plastids. A reference annotation was generated from Ensembl's IRGSP nuclear *O. sativa* genome annotation (1.0.37) (ftp://ftp.ensemblgenomes.org/pub/plants/release-37/gff3/oryza_sativa) and supplemented with the Refseq Mitochondrial and Chloroplast annotations (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/433/935/GCF_001433935.1_IRGSP-1.0). Metadata were generated with Picard tools version 2.9.0 (https://broadinstitute.github.io/picard/) and Drop-seq tools. The genome and annotations were indexed using STAR (genomeGenerate with options --runThreadN 12 --genomeDir inc_plastids --genomeFastaFiles Oryza_sat_CpMt.fa --sjdbGTFfile1.0.37_all.gtf --sjdbOverhang 49). Where necessary, annotations were converted between RAP-DB and MSU-7 identities using the Rice Annotation Project's conversion table (RAP-MSU_2017-04-14.txt, latest version is at https://rapdb.dna.affrc.go.jp/download/irgsp1.html). For quantification, raw reads were first converted from FASTQ to unaligned BAM format using Picard tools FastqToSam and subsequently processed using the unified script (Drop-seq_alignment.sh) in essentially default mode for a FASTQ starting format. Digital gene-expression profiles were then generated with the DigitalExpression utility, with the expected number of barcodes (indicative of individual samples rather than droplets in our case) set to 96. For quality assurance purposes, the digital gene-expression profiles were output both as UMI count and raw read count matrices with transcripts as rows and samples as columns. The values represent the number of raw reads or UMIs that were detected.

To distinguish sample barcodes arising from beads exposed to total RNA from an individual rice plant, rather than those that corresponded to beads never exposed to RNA, we ordered each of the UMI digital gene-expression matrices from our first 13 libraries by the total number of transcribed elements per barcode, and plotted each barcode in the matrix by the number of transcribed elements from highest to lowest number. As previously described[14], Drop-seq-type data always display a 'knee' at a sample barcode number that is equal to or just under the known number of samples included. All sample barcodes with a number of transcribed elements that was larger than this cutoff were used in downstream analyses, and the remaining sample barcodes were discarded. Samples with RNA-seq data that had to be discarded were replaced by extracting RNA from a back-up sample, and these replacement samples were included in the remaining slots of our fourteenth library.

Data normalization

The aim of normalization is to make expression levels comparable between samples by removing the effect of sequencing depth, and technical sources of heterogeneity (in our case the processing of samples in different libraries) that may confound the signal of interest. To account for differences in the total number of molecules sequenced per library, we normalized UMI counts from each sample by dividing by the total number of UMIs detected in that sample. These numbers were multiplied by 1×10^6 to obtain transcripts per million. This scaling factor largely represents a consistent increase or decrease across all positive values in our dataset. We then merged the 14 pruned digital gene-expression matrices into one super-matrix that contained transcripts-per-million expression data for all 1,320 samples after the low-quality samples had been removed. After this, very lowly transcribed elements (transcript models with a sigma signal < 20) were filtered out, and a relatively strong normalization was applied to the remaining elements in the matrix through invariant set normalization using the DChip utility version 2010.01[43] (Wong Laboratory, https://sites.google.com/site/dchipsoft/). These steps ensured that rarely encountered elements were filtered out and that confounding technical effects were removed. All downstream calculations were performed in log-space, using normalized levels ($\log_2$(normalized transcripts-per-million value + 1)) of transcribed elements that were obtained using the R (version 3.4.3) package edgeR version 3.14[44,45]. To make sure we did not consider transcripts that are relevant only for accessions in the temperate japonica subgroup of which Nipponbare is a representative[38], we kept only transcripts from protein-coding genes on nuclear chromosomes that were detected in at least 10% of individuals across our populations for all subsequent analyses.

Quantitative genetics of gene expression

Expression measures were then processed by ANOVA to partition phenotypic variation. For each gene-expression trait, we fit a mixed-effect general linear model, including a term for accession or genotype (G) as a random factor, field environment (E) as a fixed factor, the $G \times E$ interaction as a random factor and the error variance (ε). The significance of the variance explained by each of the factors was tested using an F-test. In these analyses, we controlled for multiple testing using a FDR-adjusted q value of 0.001[46]. Statistical analyses were carried out using the lme4 package version 1.1 in R[45,47], and were performed separately for the Indica and Japonica populations to control for the major source of population structure in *O. sativa*. We estimated broad-sense heritabilities as $H^2 = 0.5 \times \sigma^2_G / (0.5 \times \sigma^2_G + (0.5 \times \sigma^2_{GE}/e) + (\sigma^2_E/re))$, in which σ^2_G, σ^2_{GE} and σ^2_E are the among-genotype, $G \times E$ and within-genotype variance components (respectively), e is the number of environments and r is the number of replicates per environment. Because the predominant reproductive mode of *O. sativa* is selfing, we applied the factor 0.5 to adjust for the twofold overestimation of additive genetic variance among inbred accessions[31]. We estimated cross-environment genetic correlations as $r_{WD} = \mathrm{cov}_{ij}/\sigma_i\sigma_j$, in which cov_{ij} is the covariance of accession means between a trait as i in the wet and j in the dry field environments, and σ_i and σ_j are the square roots of the among-genotype variance components for the trait in the wet and dry field environments.

Gene set enrichment analysis on differentially expressed transcripts

We performed gene set enrichment analysis (GSEA) to obtain additional biological insight into the transcripts with a significant field-environmental bias (significant E term and $\log_2$-transformed fold change of ≥ 1.5) in their abundance using the PlantGSEA analysis pipeline version 1 at default settings[48].

Univariate phenotypic selection analyses

We measured the strength of selection on gene expression separately for the *O. sativa* Indica and Japonica populations in each of the two field environments. We used univariate regression to estimate the covariance between the expression level of each transcript individually and total lifetime fitness across the populations in the wet and dry fields, as well as the multiplicative fitness components flowering success and fecundity for the populations under drought conditions[4,5,16,49]. To prepare data for univariate selection analysis for total lifetime fitness in the wet field, we removed individuals with zero fecundity fitness (no filled grains produced) from the analysis (59 for Indica and 33 for Japonica), because these presented too few individuals for a selection analysis on flowering success—leaving fecundity fitness as a proxy for total lifetime fitness.

For selection analyses for fecundity fitness, the filled grain number for each individual plant was normalized by dividing by the mean filled

grain number of the population after filtering out individuals with zero fecundity fitness in the previous step: $w' = w_i/\text{mean}(w)$. After this, the abundance values of each transcript across individuals expressing that transcript were standardized by subtracting the population mean abundance of the transcript and dividing by the s.d. of the abundance of that transcript over the population: $z = (x_i - \text{mean}(x))/\text{s.d.}(x)$. Finally, individuals that were severe outliers for the relative abundance of a transcript (±3 s.d.) were removed on a per-transcript basis to satisfy the assumption of normality for the selection analyses. Only transcripts that were expressed in at least 20 individuals in a population were used for analysis using a custom script in Python version 2.7 (Supplementary Note 1).

We performed separate analyses to estimate the strength and direction of selection on gene expression for the fitness component flowering success using univariate logistic regression for each individual transcript (expressed in ≥ 20 individuals) across all individuals in the populations in the dry field environment[50], again using a custom Python script (Supplementary Note 2). Flowering success was defined as a binary state in which individuals were given a value of 1 if they were able to produce at least one filled grain before the end of the growing season, and 0 if not. Because, in our study, flowering success and fecundity are multiplicative fitness components[16], we added up the selection differentials and error estimates for these fitness components under drought conditions to establish selection differentials and error estimates for total lifetime fitness in dry conditions[51].

Both the linear and logistic regression analyses output standardized directional-selection differentials ($S = \text{cov}[w, z]$), and disruptive- or balancing-selection differentials ($C = \text{cov}[w, (z - \text{mean}(z)\,(z - \text{mean}(z))^T)]$) that reflect the total (direct and indirect) selection on the expression level of a gene[4,5].

Multivariate phenotypic selection analyses

For each population, we performed separate dimensional reductions on the transcriptome datasets per field environment through principal component analysis (PCA) using the prcomp function in R[45,52], and conducted multivariate selection analyses for total lifetime fitness in the wet and dry environments, and for flowering success and fecundity in the dry environment[5,49–51]. We calculated both the linear selection gradients, ($\beta = P^{-1}S$), and quadratic selection gradients, ($\gamma = P^{-1}CP^{-1}$), in which P represents the phenotypic variance–covariance matrix of the transcript abundances included as traits[5]. Selection gradients for total lifetime fitness under drought conditions were obtained in the same way as described in 'Univariate phenotypic selection analyses'. The selection gradients reflect the strength and direction of direct selection on a trait. In addition to determining the pattern and strength of selection on gene expression, we also estimated selection differentials and gradients for the three higher-level traits we measured: chlorophyll concentration, flowering time and leaf area.

Factors affecting the strength of selection on gene expression

We performed a series of analyses to assess whether there are factors that might be linked to the heterogeneity of selection strengths between different transcripts. For the Indica population, the covariates of expression level, stochastic noise and polymorphism were directly derived from the transcript expression super-matrix with expression level defined as the grand mean expression level of a transcript in each field environment, expression noise defined as the average variance in the abundance of a transcript between individual replicates of all accessions and expression polymorphism as the population-wide variance between accession mean expression levels. Transcript H^2 and r_{WD} were calculated as described in 'Quantitative genetics of gene expression'.

For nearly all transcripts, information on their length and GC content could be downloaded from the Ensembl Plants BioMart release 43 (https://plants.ensembl.org/biomart/martview) *O. sativa* Japonica IRGSP-1.0 dataset. In addition, we obtained tissue-specific expression data for 29,122 genes in 9 tissues from the EMBL-EBI Expression Atlas, experiment E-MTAB-2039 (https://www.ebi.ac.uk/gxa/experiments/E-MTAB-2039/Results), originally generated in a previous publication[53]. From these data, the tissue specificity index value for each gene was calculated:

$$\tau = \frac{\sum_{i=1}^{n}(1 - x_i)}{n - 1}$$

in which n is the number of tissues and x_i is the normalized expression profile component[54]. For each of the covariates described thus far, information was available for the vast majority of all transcripts (n = 14,753 transcripts or 94.4%) that were included in our phenotypic selection analyses.

The nine covariates did not have irregular distributions and were included in a partial correlation analysis (n = 14,753 transcripts) using the R package corpcor version 1.6.9[45,55]. We started by calculating parametric Pearson product–moment correlations between pairs of all variables for each field environment by selection component combination, after which we estimated the partial correlations by establishing the pseudo-inverse of the resulting correlation matrices[7,56].

Relation between fitness and global gene-expression stochasticity and plasticity

We computed mean values for fecundity across replicates for each accession that were included in the phenotypic selection analyses, and correlated these fitness values with genome-wide (global) measures of gene-expression stochasticity and plasticity (the latter only for the dry field environment). To obtain estimates of global gene expression plasticity, we performed targeted ANOVA for each accession individually by fitting a fixed-effect general linear model, including a term for field environment (E) as a fixed factor, and the error variance (ε). The significance of the variance explained by the environment factor was tested using an F-test. The number of significant drought-induced transcripts at FDR-adjusted $q < 0.05$ for an accession was taken as a proxy for global gene-expression plasticity for that accession[15].

To obtain estimates of global gene-expression stochasticity for accessions in each field environment, we averaged the variance across the three replicate individuals of an accession for all transcripts as previously described[24], after we calculated the level of stochastic noise for each transcript within an accession as σ^2/μ^2, variance divided by the mean squared, known as CV^2 (the squared coefficient of variation)[57]. Our measure of expression stochasticity is corrected by including expression level as a covariate in the analysis, just as it was in the partial correlation analysis. The relation between fitness and global gene-expression plasticity and stochasticity was obtained through computing nonparametric Spearman's rank correlation coefficients.

Network effects on the strength of selection on gene expression

We performed separate analyses on four independent measures of network effects on the strength of selection on gene expression. We obtained measures of within-cluster connectivity from 53 clusters of 17,931 co-expressed transcripts that were previously derived from transcriptome data of 240 samples[25]. These samples were taken in time series from Indica and Japonica accessions growing in wet (irrigated, flooded paddy) and dry (rain-fed) field environments across a dry and a wet season, in the same geographical location as our experiment.

The number of *cis*-regulatory element groups in the promoter regions (from −1 kb to +200 bp relative to the transcription start site) of 3,907 genes that overlapped with genes in our analysis were obtained from the Plant Promoter Database (PPDB) version 3.0[58], which we accessed at http://ppdb.agr.gifu-u.ac.jp/ppdb/cgi-bin/index.cgi. Only those *cis*-regulatory element groups that correspond to known *cis*-elements were included[59]. The median number of *cis*-regulatory element groups per promoter was 5, and we tested whether the expression of genes

with up to 5 *cis*-regulatory element groups in their promoter (n = 2,141 transcripts) experienced stronger selection than the expression of genes with 6 or more *cis*-regulatory element groups in their promoter (n = 1,766 transcripts) by performing a Mann–Whitney U-test.

The number of transcription factors predicted to regulate the abundance of each transcript in a network context (that is, the 'in-degree' of each gene) was obtained from 2,905 transcripts in previously created environmental gene regulatory influence networks[26] that overlapped with transcripts in our analysis. The environmental gene regulatory influence networks were built through combining prior knowledge on experimentally validated or inferred transcription-factor binding preferences[60], with rice gene-expression and chromatin-accessibility data from plants grown in wet and dry conditions[25,26]. We tested whether the level of transcripts predicted to be regulated by one transcription factor (n = 1,505) experienced stronger selection than the level of transcripts predicted to be regulated by more than one transcription factor (n = 1,400) by performing a Mann–Whitney U-test.

GSEA on transcripts under selection

We performed GSEA to obtain additional biological insight into the transcripts in the 5% tails of the distributions of transcripts' loading values on principal components with significant selection gradients, and of the P value distributions of the transcripts' selection differentials for total lifetime fitness in wet and dry conditions, and for fecundity and flowering success under drought conditions using the PlantGSEA analysis pipeline version 1 at default settings[48].

Ranking Gene Ontology biological processes by selection strength

Gene Ontology term annotations were downloaded from Monocots PLAZA 4.0[61,62]. We obtained biological-process Gene Ontology annotations for 11,901 transcripts that overlapped with transcripts in our analyses. To minimize redundancy among Gene Ontology terms, we focused our analysis on biological-process Gene Ontology terms that were represented in the annotations of at least 20 transcripts in our dataset. This resulted in the inclusion of 6,161 transcripts with Gene Ontology biological-process term annotations in our analysis.

We used the median selection strength $|S|$ of all transcripts annotated to be involved in a particular biological process as a proxy for the selection strength $|S|$ on that process. By setting the minimum size per term as at least 20 transcripts and by considering the median $|S|$ for each Gene Ontology term, we not only limited redundancy but also avoided estimates of selection strength per Gene Ontology term being influenced by small group sizes. We tested for rank shifts in the order of biological processes by their median $|S|$ between field environments for total lifetime fitness through conducting Mann–Whitney U-tests (n = 243 biological processes per group). Furthermore, per field environment by fitness component combination, we considered any biological process to be under significantly stronger selection than the transcriptome-wide median (n = 6,161 transcripts) if the median selection strength for a process was removed from the transcriptome-wide median selection strength by at least the 95% confidence interval for the selection strength of that process.

Transcript associations with higher-level traits

We identified transcripts significantly associated ($P < 0.01$) with the three higher-level organismal traits we measured (chlorophyll concentration, flowering time and leaf area) for the Indica population in each field environment by using regression models: $Y = \mu + T + \varepsilon$, in which Y represents the higher-level trait of interest, μ an intercept parameter, T denotes the transcript covariate and ε residual error.

Selection of DNA sequence read data

Raw FASTQ reads from 27 accessions included in the 3K-RG project were downloaded from the Sequence Read Archive (SRA) website under BioProject PRJEB6180[38]. For a further 188 accessions, raw FASTQ reads were downloaded from SRA BioProject accession numbers PRJNA422249 and PRJNA557122[34]. DNA sequence data were available for 215 out of 220 accessions; one accession was a 'filler' accession and its genome was not resequenced, and a further four accessions were replicated checks of two accessions, IR64 and Sahod Ulan 1. Accession numbers and origins of tissue for DNA extraction can be found in Supplementary Table 1. Overall, a total of 1,203,564,772,205 bp (about 1.2 Tbp) were included for downstream analyses.

Reference-genome-based DNA read alignment

FASTQ reads were preprocessed using the bbduk program of BBTools version 37.66 (https://jgi.doe.gov/data-and-tools/bbtools/) for read quality control and adaptor trimming. For bbduk, we used the options: minlen = 25 qtrim = rl trimq = 10 ktrim = r k = 25 mink = 11 hdist = 1 tpe tbo. This trimmed reads below a phred score of 10 on both sides of the reads to a minimum length of 25 bp, trimmed 3′ adapters using a k-mer size of 25 as well as a k-mer size of 11 for ends of reads, allowed one Hamming distance mismatch, trimmed adapters based on overlapping regions of the paired-end reads, and trimmed reads to equal lengths if one of them was adaptor-trimmed. FASTQ reads were aligned to the reference *O. sativa* Nipponbare IRGSP 1.0 genome downloaded from EnsemblPlants release 37 (ftp://ftp.ensemblgenomes.org/pub/plants/). Read alignment was done using the program bwa-mem version 0.7.16a-r1181[63]. Only the 12 pseudomolecules were used as a reference, and the unassembled scaffolds were left out. PCR duplicates during the library preparation step were determined computationally and removed using the Picard tools version 2.9.0.

SNP calling

For each accession, genotype calling for each site was conducted using the GATK HaplotypeCaller engine version 3.8-0-ge9d806836 in the -ERC GVCF mode to output files in the genomic variant call format (gVCF). The gVCF files from each accession were merged together to conduct multi-accession joint genotyping using the GATK GenotypeGVCFs engine. Genotypes were divided into SNP or insertion and deletion (indel) variants and filtered using the GATK bestpractice hard filter pipeline[64]. For SNP variants we excluded regions that overlapped repetitive regions and variants that were within 5 bp of an indel variant. We then used vcftools version 0.1.15 to select SNPs that had at least 80% of sites with a genotype call, and exclude SNPs with minor allele frequency <5% to remove potential false-positive SNP calls arising from sequencing errors or false genotype calls[65]. Because domesticated rice is an inbreeding species, we also implemented a heterozygosity filter for sites that had a heterozygous genotype in more than 5% of the samples using the program vcffilterjdk.jar from the jvarkit suite version 1 (https://figshare.com/articles/JVarkit_java_based_utilities_for_Bioinformatics/1425030). Missing genotypes were imputed and phased using Beagle version 4.1[66]. Finally, we randomly pruned the SNPs by sampling a polymorphic site every 1,000 bp using plink version 1.9[67], leaving a SNP dataset of 179,634 markers.

G-matrix estimation and prediction of short-term phenotypic evolution

A G-matrix consists of the additive genetic variances and covariances of a series of traits, and we assembled one for the principal component axes as eigengenes or reflections of suites of transcripts in our transcriptome data across rice individuals. Although the principal components are—by definition—uncorrelated at the level of the individual replicate plants at which we generated them, they start showing genetic covariances when loading values of replicates of each genotype are averaged. Estimates of additive genetic variance and covariance were obtained using a previously described approach[68]. First, we constructed a kinship matrix from the SNP dataset using the VanRaden method in the R package GAPIT version 3,

a genome association and prediction integrated tool[69,70]. We let GAPIT estimate the contribution of structure between accessions to each trait (principal component) using a variance component model, providing us with the fraction of phenotypic variance explained by the kinship matrix. This fraction (termed pseudo-heritability) resembles the narrow-sense heritability estimated from a pedigree, and serves as an estimate of the additive genetic variance of a trait[71]. We then applied a bivariate genetic model as previously outlined[68] to obtain estimates of the additive genetic covariance between traits and principal components.

We used the G-matrix to predict the outcome of selection on gene expression across one generation (Δz), and assess whether evolutionary constraints were present, by combining it with the linear selection gradients on the principal components in the multivariate breeder's equation: $\Delta z = G\beta$.

Genome-wide association study

We conducted genome-wide association mapping in GAPIT by applying a multi-locus linear mixed model, a model based on EMMA that uses forward–backward stepwise linear mixed-model regression to estimate variance components[72,73]. We included population structure cofactors as well as the kinship matrix described in 'G-matrix estimation and prediction of short-term phenotypic evolution' as a random factor in the model. Structure in our Indica population of 131 different genotypes was inferred with a PCA, and GAPIT used the first four principal components as cofactors (Supplementary Table 26). Significant SNPs were identified using a conservative Bonferroni threshold[74], which was at $P < 2.78 \times 10^{-7}$. Finally, we selected the top approximately 0.5% SNPs (1,000 SNPs) based on P value for association with total lifetime fitness in each environment[75], with the aim of testing whether the 100-kbp windows surrounding these SNPs were enriched for transcripts classed as showing non-neutral microevolutionary selection patterns (selection strength $|S|$ $P < 0.05$). The window size was chosen as a range of 50 kbp at either side of a SNP, which is conservative given an estimated breakdown of linkage disequilibrium in a range of 75–125 kbp in *O. sativa* subgroup indica[38,76–78].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Raw FASTQ reads for 188 accessions with resequenced genomes were downloaded from the SRA under SRA BioProject accession numbers PRJNA422249 and PRJNA557122. Raw FASTQ reads for a further 27 accessions included in the 3K-RG project were downloaded from the SRA under BioProject accession number PRJEB6180. RNA sequence data that support the findings of this study have been deposited under SRA BioProject accession number PRJNA588478. Processed RNA expression count data have been deposited in Zenodo (https://zenodo.org/record/3533431 with DOI 10.5281/zenodo.3533431), alongside a sample metadata file with a key to the RNA sequence data in SRA BioProject accession number PRJNA588478. This key can also be found in Supplementary Table 4. Source Data for Figs. 1–4 and Extended Data Figs. 1–8 are provided with the paper.

Code availability

Selection analyses were run using custom-made scripts in Python version 2.7, which are available in Supplementary Notes 1, 2, and on GitHub in repositories icalic/Linear-regression-analysis (https://github.com/icalic/Linear-regression-analysis.git) and icalic/Logistic-regression-analysis (https://github.com/icalic/Logistic-regression-analysis.git). For all other analyses we used previously developed, publicly available software and code: leaf area was assessed using ImageJ v.1.52 and GIMP v.2.10.0; RNA-seq data were processed and analysed using Drop-seq tools v.1.12, STAR aligner v.020201, Picard tools v.2.9.0, DChip v.2010.01 and R v.3.4.3 packages edgeR v.3.14 and lme4 v.1.1; gene-set enrichment analyses were performed using PlantGSEA v.1; statistical analyses were performed in R v.3.4.3, further using packages lme4 v.1.1 and corpcor v.1.6.9; and genome analyses were performed using bbduk v.37.66, bwa-mem v.0.7.16a-r1181, the GATK GenotypeGVCFs engine v.3.8-0-ge9d806836, vcftools v.0.1.15, jvarkit suite v.1, Beagle v.4.1, plink v.1.9 and GAPIT v.3.

31. Keurentjes, J. J. et al. Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc. Natl Acad. Sci. USA* **104**, 1708–1713 (2007).
32. Caicedo, A. L. et al. Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet*. **3**, e163 (2007).
33. Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S. & McCouch, S. Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**, 1631–1638 (2005).
34. Gutaker, R. M. et al. Genomic history and ecology of the geographic spread of rice. Preprint at bioRxiv https://doi.org/10.1101/748178 (2019).
35. McCouch, S. R. et al. Open access resources for genome-wide association mapping in rice. *Nat. Commun*. **7**, 10532 (2016).
36. McNally, K. L. et al. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl Acad. Sci. USA* **106**, 12273–12278 (2009).
37. Torres, R. O., McNally, K. L., Cruz, C. V., Serraj, R. & Henry, A. Screening of rice genebank germplasm for yield and selection of new drought tolerance donors. *Field Crops Res*. **147**, 12–22 (2013).
38. Wang, W. et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**, 43–49 (2018).
39. Abramoff, M. D., Magalhaes, P. J. & Ram, S. J. Image processing with ImageJ. *Biophoton. Int*. **11**, 36–42 (2004).
40. Bracken, B. Barcoded plate-based single cell RNA-seq. https://www.protocols.io/view/barcoded-plate-based-single-cell-rna-seq-nkgdctw (2018).
41. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098 (2013).
42. Soumillon, M., Cacchiarelli, D., Semrau, S., van Oudenaarden, A. & Mikkelsen, T. S. Characterization of directed differentiation by high-throughput single-cell RNA-seq. Preprint at bioRxiv https://doi.org/10.1101/003236 (2014).
43. Li, C. & Wong, W. H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA* **98**, 31–36 (2001).
44. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
45. R Core Team. R: a language and environment for statistical computing. http://www.R-project.org/ (R Foundation for Statistical Computing, Vienna, 2016).
46. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
47. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw*. **67**, 1–48 (2015).
48. Yi, X., Du, Z. & Su, Z. PlantGSEA: a gene set enrichment analysis toolkit for plant community. *Nucleic Acids Res*. **41**, W98–W103 (2013).
49. Brodie, E. D. III, Moore, A. J. & Janzen, F. J. Visualizing and quantifying natural selection. *Trends Ecol. Evol*. **10**, 313–318 (1995).
50. Janzen, F. J. & Stern, H. S. Logistic regression for empirical studies of multivariate selection. *Evolution* **52**, 1564–1571 (1998).
51. Koenig, W. D., Albano, S. S. & Dickinson, J. L. A comparison of methods to partition selection acting via components of fitness: do larger male bullfrogs have greater hatching success? *J. Evol. Biol*. **4**, 309–320 (1991).
52. Kassambara, A. *Practical Guide to Principal Component Methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra* (STHDA, 2017).
53. Davidson, R. M. et al. Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *Plant J*. **71**, 492–502 (2012).
54. Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).
55. Schäfer, J. & Strimmer, K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol*. **4**, Article32 (2005).
56. Larracuente, A. M. et al. Evolution of protein-coding genes in *Drosophila*. *Trends Genet*. **24**, 114–123 (2008).
57. Keren, L. et al. Noise in gene expression is coupled to growth rate. *Genome Res*. **25**, 1893–1902 (2015)
58. Hieno, A. et al. ppdb: plant promoter database version 3.0. *Nucleic Acids Res*. **42**, D1188–D1192 (2014).
59. Yamamoto, Y. Y. et al. Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics* **8**, 67 (2007).
60. Weirauch, M. T. et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).
61. Proost, S. et al. PLAZA: a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell* **21**, 3718–3731 (2009).
62. Van Bel, M. et al. PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res*. **46**, D1190–D1196 (2018).
63. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997 (2013).

64. Van der Auwera, G. A. et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013).
65. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
66. Browning, B. L. & Browning, S. R. Genotype imputation with millions of reference samples. *Am. J. Hum. Genet*. **98**, 116–126 (2016).
67. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet*. **81**, 559–575 (2007).
68. Tropf, F. C. et al. Human fertility, molecular genetics, and natural selection in modern societies. *PLoS ONE* **10**, e0126821 (2015).
69. Lipka, A. E. et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399 (2012).
70. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci*. **91**, 4414–4423 (2008).
71. Kang, H. M. et al. Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet*. **42**, 348–354 (2010).
72. Kang, H. M. et al. Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709–1723 (2008).
73. Segura, V. et al. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet*. **44**, 825–830 (2012).
74. Bland, J. M. & Altman, D. G. Multiple significance tests: the Bonferroni method. *Br. Med. J.* **310**, 170 (1995).
75. Fournier-Level, A. et al. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86–89 (2011).
76. Huang, X. et al. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet*. **42**, 961–967 (2010).
77. Mather, K. A. et al. The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics* **177**, 2223–2232 (2007).
78. Zhao, K. et al. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun*. **2**, 467 (2011).

Acknowledgements We thank B. U. Principe, P. C. Maturan and L. Holongbayan for assistance with field management, tissue sampling and trait measurements; the staff of IRRI's Climate Unit for providing weather data; Z. Fresquez for help with tissue processing; L. Harshman for assistance with a pilot RNA-seq run; the New York University Center for Genomics and Systems Biology GenCore Facility for sequencing support; and New York University High Performance Computing for supplying computational resources. We are grateful to current and former members of the Purugganan laboratory (particularly J. Flowers, R. Gutaker, A. Plessis, O. Wilkins and M. Zaidem) and the IRRI Strategic Innovation and Rice Breeding research platforms (particularly S. Dixit, A. Kohli, Y. Ludwig, K. McNally, R. Oliva, V. Roman-Reyna and N. Tsakirpaloglou) for insightful discussions; M. Quintana for sharing scripts in R; and S. Zaaijer for codesigning the figures. This work was funded in part by grants from the Zegar Family Foundation, the National Science Foundation Plant Genome Research Program and the NYU Abu Dhabi Research Institute to M.D.P., a fellowship from the Natural Sciences and Engineering Research Council of Canada through Grant PDF-502464-2017 to Z.J.-L., and a fellowship from the Gordon and Betty Moore Foundation/Life Sciences Research Foundation through Grant GBMF2550.06 to S.C.G.

Author contributions M.D.P. conceived and directed the project; M.D.P., G.V., A.H., R.O.T., A.K., and S.C.G. designed and coordinated field experiments; M.N., C.L.U.C., Z.J.-L., J.Y.C., and S.C.G. performed field experiments; K.D., M.N., Z.J.-L., J.Y.C. and S.C.G. processed samples and extracted RNA for sequencing; W.M.M. III and B.B. made RNA-seq libraries; A.E.P. and R.S. designed and ran the bioinformatics workflow for RNA-seq; J.Y.C. and S.C.G. conducted genetic-marker-based analyses; I.Ć., S.C.G., S.J.F. and M.D.P. designed and performed selection analyses; M.N., A.H., J.Y.C., S.C.G. and I.Ć. processed fitness, higher-level-trait and gene-expression data, and performed statistical analyses; and S.C.G., S.J.F. and M.D.P. wrote the manuscript.

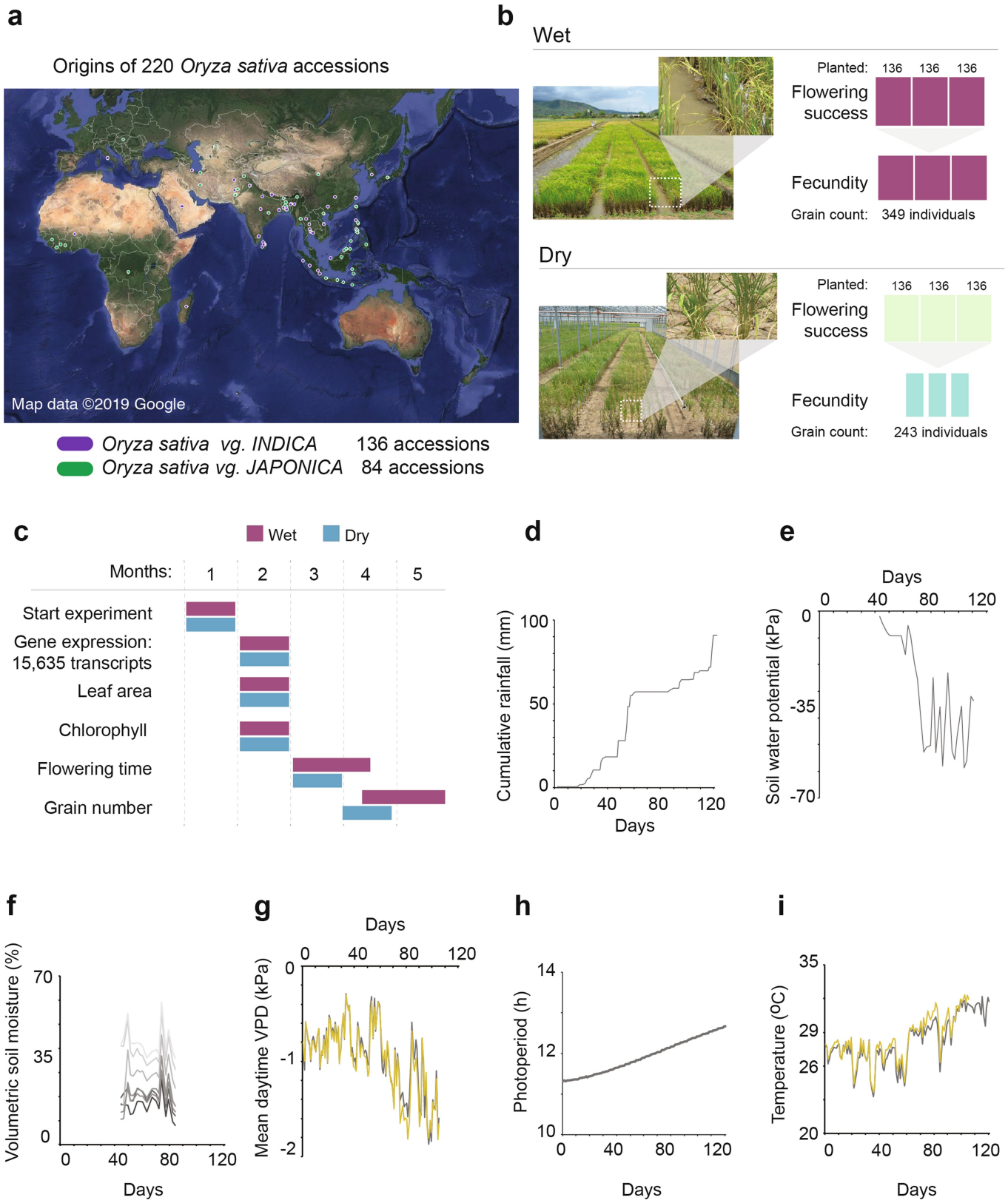
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-1997-2.

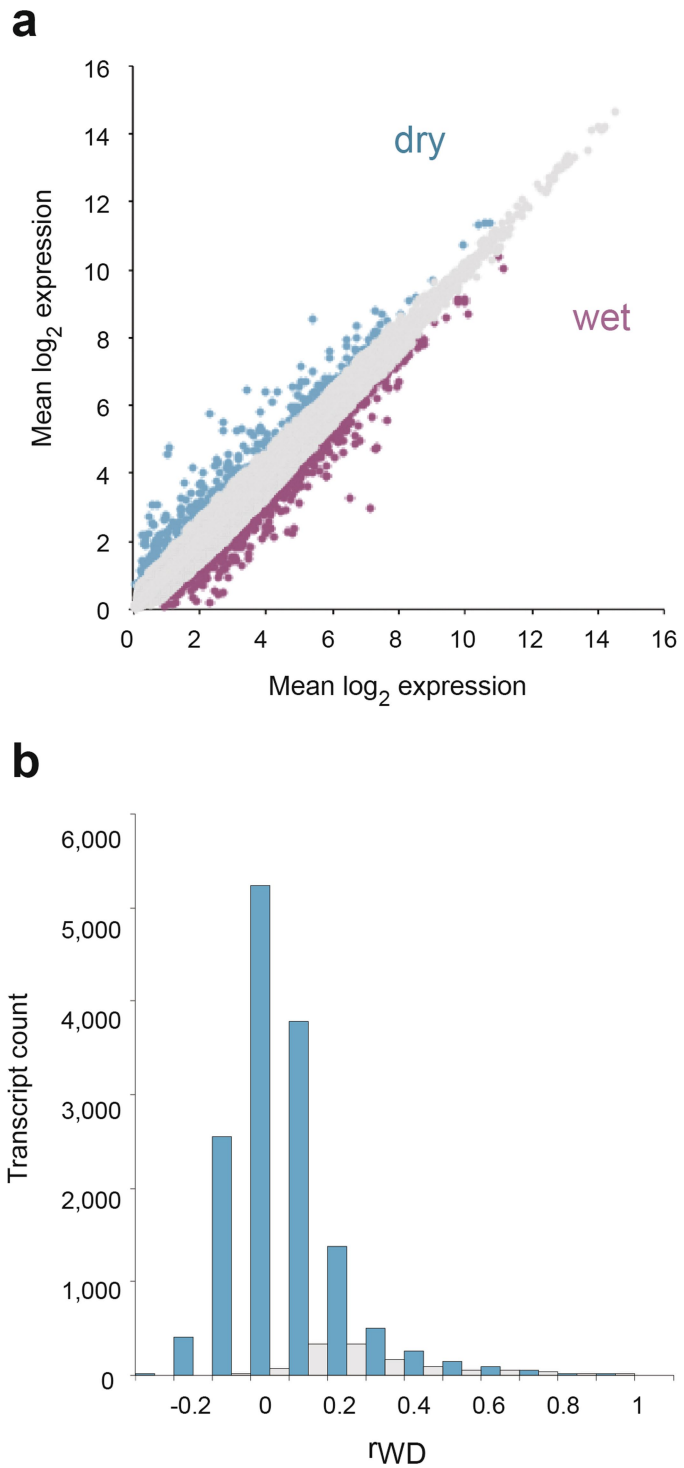
Correspondence and requests for materials should be addressed to M.D.P.

Peer review information *Nature* thanks Anthony Greenberg, Detlef Weigel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

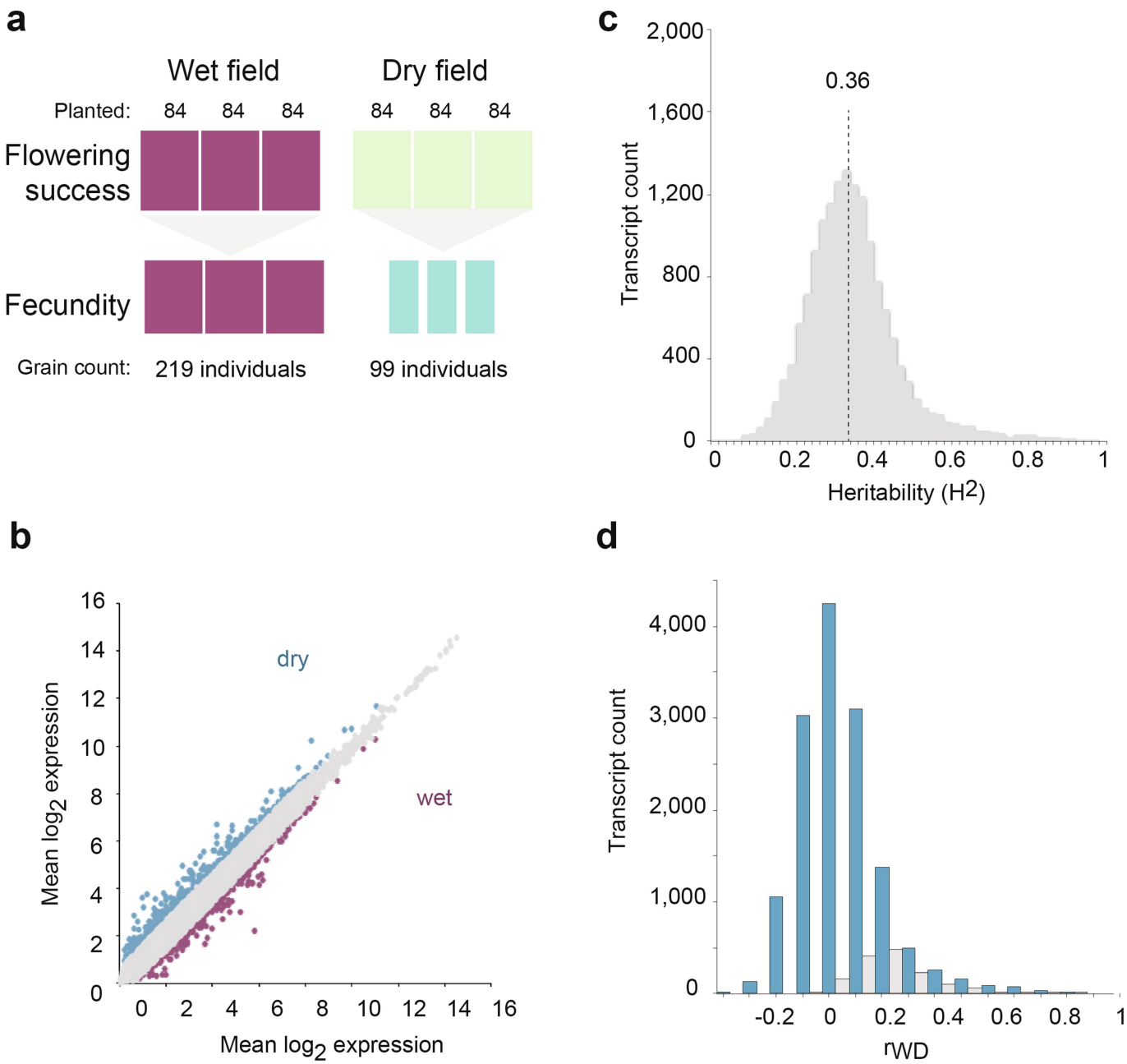


Extended Data Fig. 1 | See next page for caption.

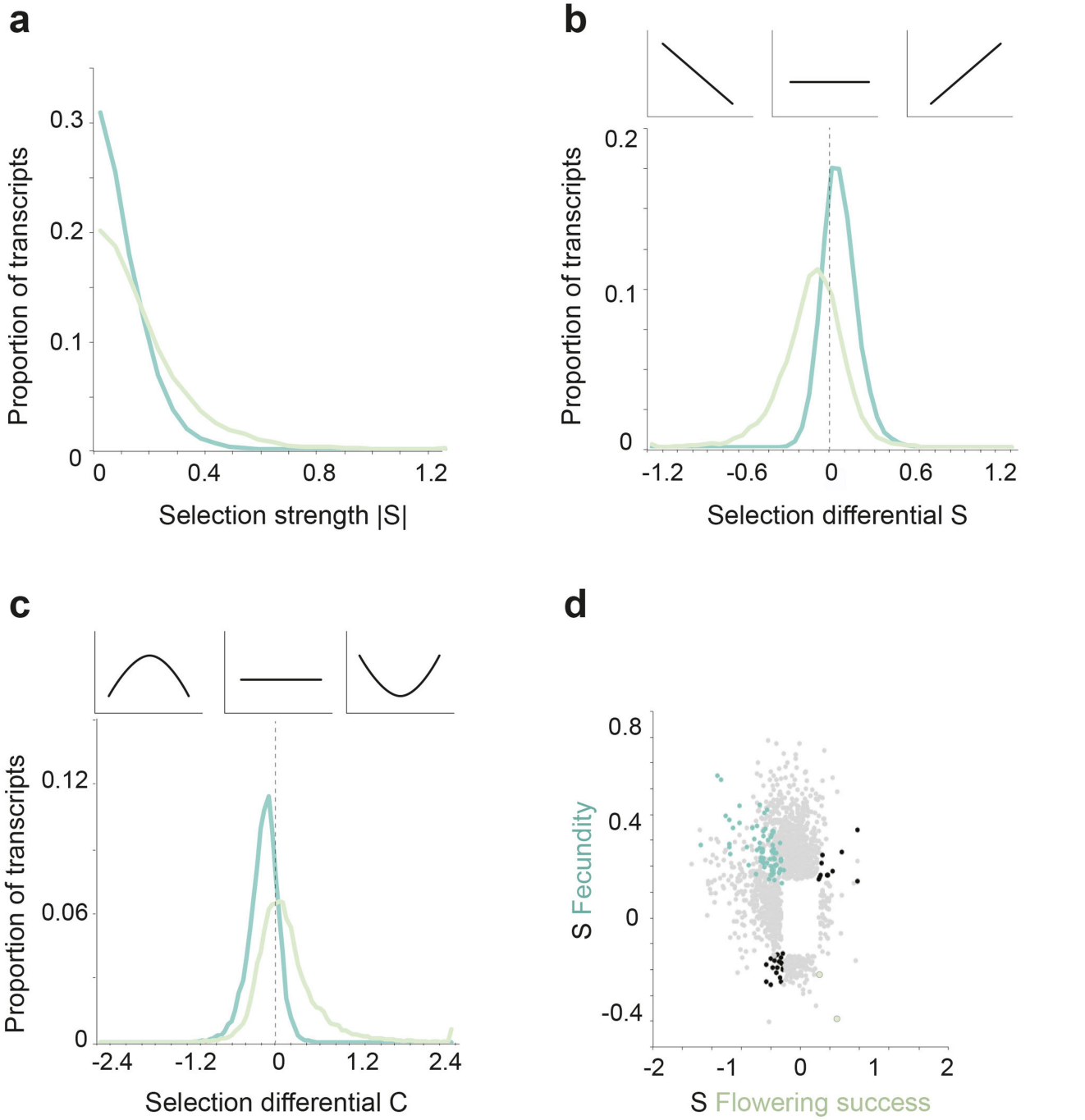
Extended Data Fig. 1 | Experimental setup. a, Geographical origins of 220 *O. sativa* accessions, of which 4 constitute additionally replicated checks (Supplementary Table 1). Seven accessions that are not from Eurasia or Africa are not shown. Varietal group (vg.) Indica accessions are indicated in indigo and vg. Japonica accessions are indicated in jade. Map data ©2019 Google. **b**, Populations of Indica and Japonica accessions (planted in triplicate alongside one another) were monitored for total lifetime fitness in wet (magenta) and dry (blue) fields. Both fields had identical layouts. Numbers reflect Indica populations with 3 × 136 accessions = 408 individuals planted in each field; Extended Data Fig. 3 shows Japonica populations. Under drought conditions, both multiplicative fitness components (flowering success (lime) and fecundity (green)) were relevant (multiplying to total lifetime fitness), but in wet conditions only the latter was relevant (fecundity equating to total lifetime fitness, magenta). **c**, Drought exerts truncating selection on the populations (declining and shifting blue versus magenta bar), and end-of-season was reached earlier under drought conditions. **d**, Cumulative rainfall shows one major rainfall event that caused the rainout shelter over the dry field to close temporarily after the start of the drought treatment and the sampling of leaf tissue for RNA sequencing (>51 DAS). **e**, During the period of flowering (>51 DAS), there was an increasing deficit in soil water potential. **f**, **g**, Patterns of volumetric soil moisture and vapour pressure deficit (VPD) were consistent with the pattern of soil water potential. Lighter shades of grey in **f** indicate deeper layers of soil. Grey and mustard lines in **g** indicate the VPD in the wet and dry field, respectively. **h**, Day length increased over the course of the experiment. **i**, Air temperature generally increased over the course of the experiment (grey and mustard lines indicate the wet and dry field, respectively).



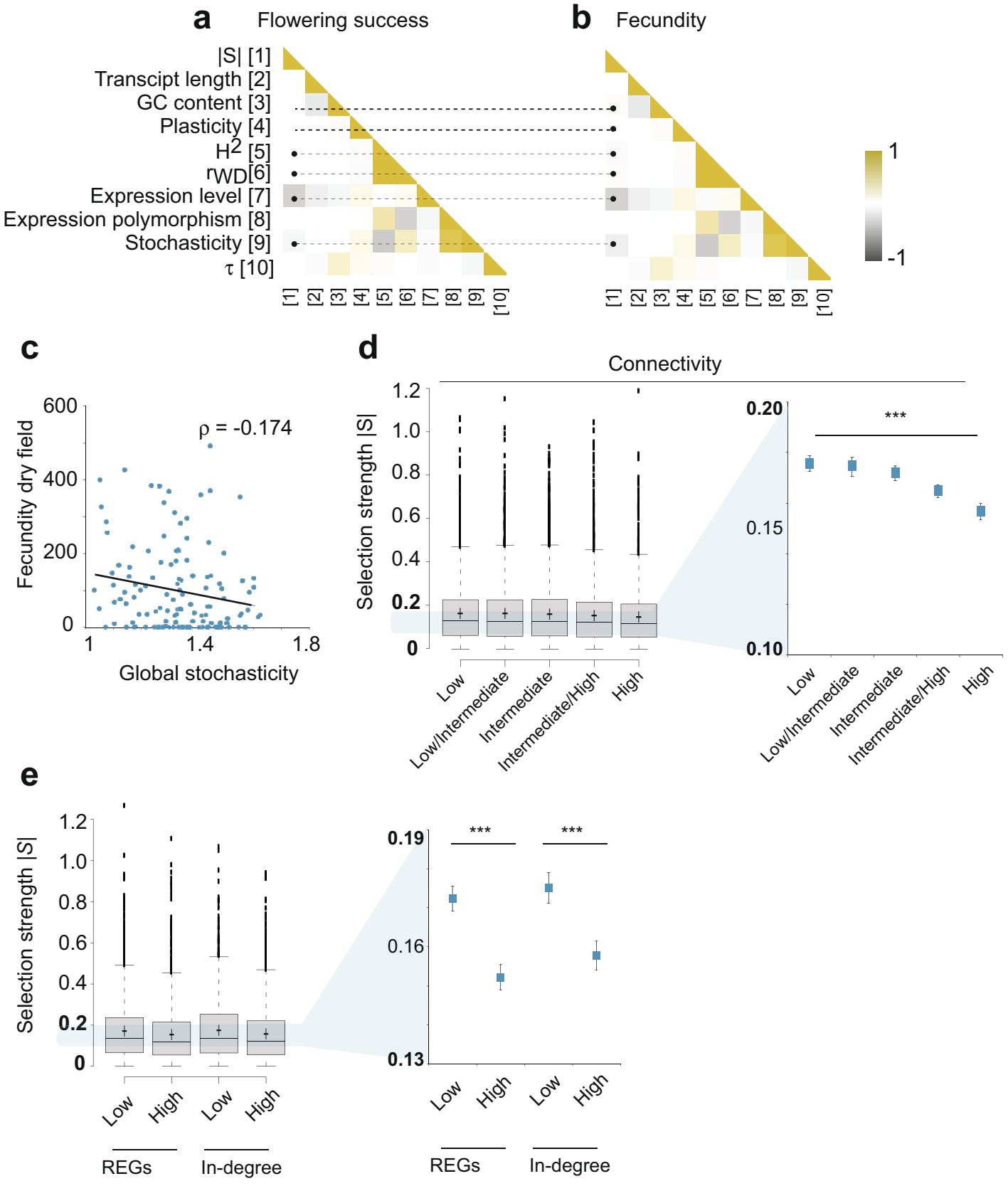
Extended Data Fig. 2 | Systems genetics of gene expression in the Indica populations in wet and dry field environments. a, Environmental bias for transcript expression. Magenta and blue dots represent transcripts showing a 1.5-fold difference in expression between the wet and dry field environments, respectively. ANOVA, Indica environment FDR-adjusted $q < 0.001$, $n = 136$ accessions. **b**, Distribution of cross-environment genetic correlations (r_{WD}) for transcripts showing significant (blue) genotype × environment ($G \times E$) variance. ANOVA, Indica genotype × environment FDR-adjusted $q < 0.001$, $n = 136$ accessions.



Extended Data Fig. 3 | Systems genetics of gene expression in the Japonica populations in wet and dry field environments. a, Monitoring the Japonica populations, with 3 × 84 accessions = 252 individuals planted in both the wet and dry fields, for flowering success, fecundity fitness and total lifetime fitness (legend as in Extended Data Fig. 1b, c). **b**, Environmental bias for transcript expression. Magenta and blue dots represent transcripts showing a 1.5-fold difference in expression between the wet and dry field environments, respectively. ANOVA, Japonica environment FDR-adjusted $q < 0.01$, $n = 84$ accessions. **c**, Distribution of broad-sense heritabilities (H^2) for transcripts with significant expression polymorphism. ANOVA, Japonica genotype FDR-adjusted $q < 0.01$, $n = 84$ accessions. **d**, Distribution of cross-environment genetic correlations (r_{WD}) for transcripts showing significant (blue) genotype × environment ($G \times E$) variance. ANOVA, Japonica genotype × environment FDR-adjusted $q < 0.01$, $n = 84$ accessions.

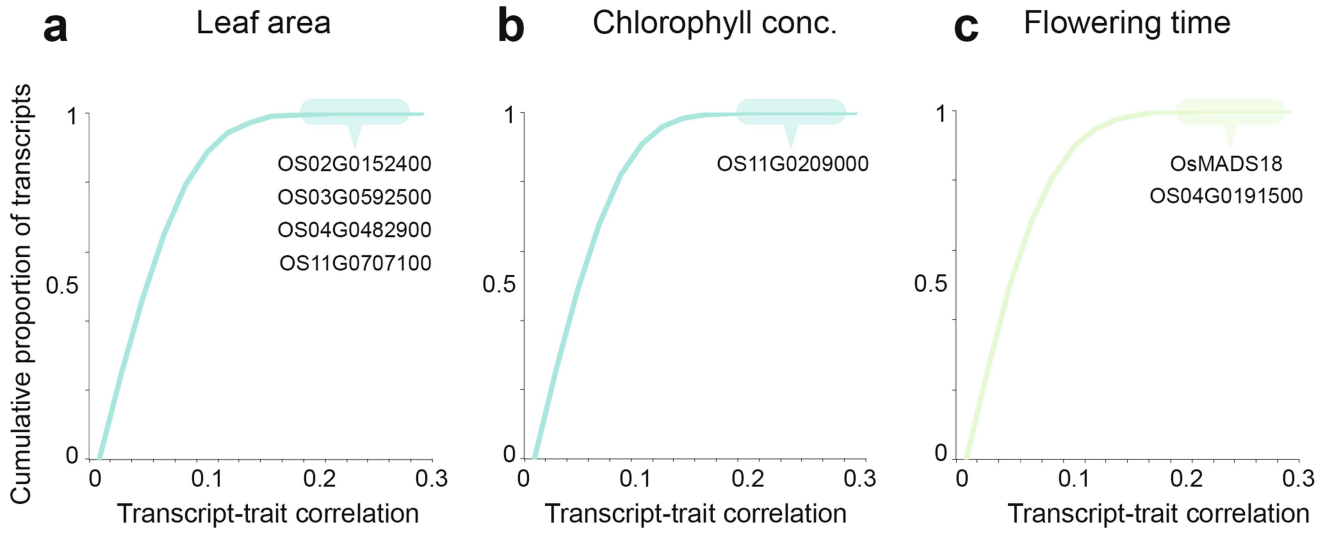


Extended Data Fig. 4 | The strength and pattern of selection on Indica rice-leaf transcript levels under drought conditions differ across fitness components. a, The strength of selection |*S*| on gene expression differed between selection for flowering success (lime), and fecundity (green) in the dry field. Mann–Whitney *U*-test, two-sided $P < 0.001$, $n = 15{,}343$. **b**, Positive directional selection ($n = 11{,}304$) was stronger than negative selection ($n = 4{,}039$) for fecundity under drought (green) (Mann–Whitney *U*-test, two-sided $P < 0.001$), and selection for flowering success showed higher absolute values (Kolmogorov–Smirnov test, two-sided $P < 0.001$, $n = 15{,}343$). **c**, Patterns of quadratic selection differed significantly for the two fitness components. Kolmogorov–Smirnov test, two-sided $P < 0.001$, $n = 15{,}343$. **d**, Patterns of conditional neutrality (light grey) and antagonistic pleiotropy (lime and green for transcripts beneficial for flowering success and fecundity, respectively) for gene expression under drought conditions. Black indicates transcripts that experienced selection in the same direction for both fitness components.

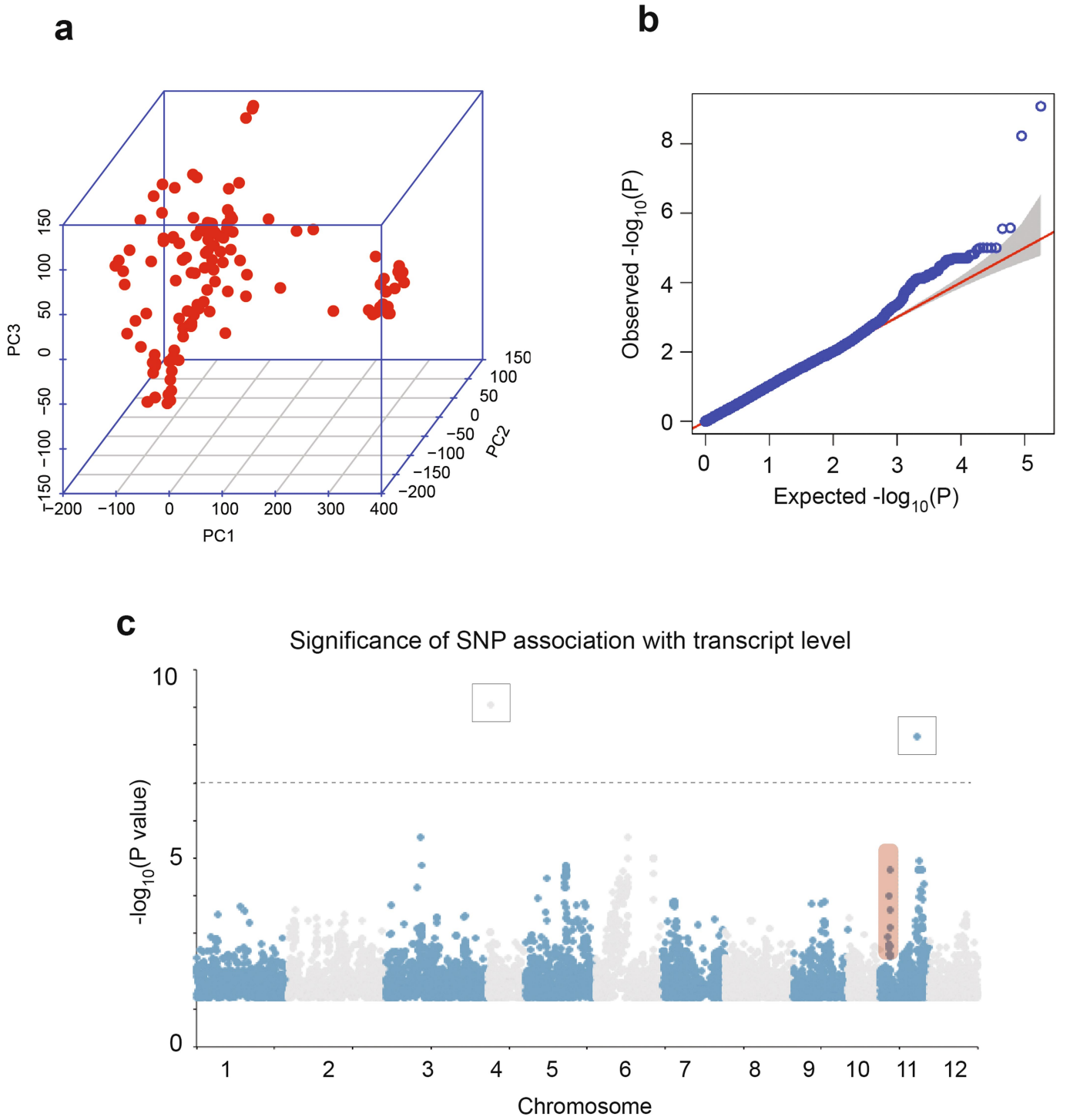


Extended Data Fig. 5 | See next page for caption.

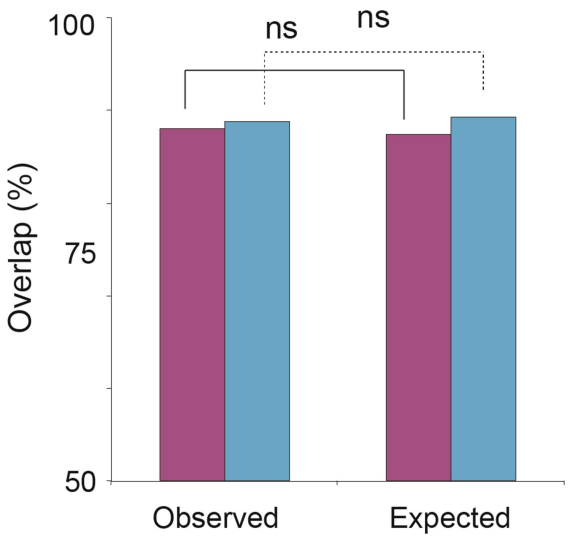
Extended Data Fig. 5 | Stochastic expression noise and transcript connectivity limit the efficacy of selection on gene expression. a, **b**, Partial correlation analyses of factors that negatively (grey) and positively (mustard) influence the strength of selection $|S|$ on gene expression for flowering success (**a**) and fecundity (**b**) fitness in dry conditions. Dots indicate statistical significance of Pearson's partial r (t-test, two-sided $P < 0.05$, $n = 14{,}753$) (Supplementary Table 14). **c**, Global expression stochasticity limits fecundity under drought conditions. Spearman's $\rho = -0.174$, t-test, two-sided $P = 0.042$, $n = 136$ accessions. **d**, As in wet conditions, $|S|$ is bounded by expression connectivity under drought conditions. Kruskal–Wallis test, $P = 0.0008$, $n = 12{,}502$ transcripts. Left, box plot with centre line = median, cross = mean, box limits = upper and lower quartiles, whiskers = 1.5 × interquartile range, points = outliers. Right, mean ± s.e.m. **e**, In dry as well as in wet conditions, $|S|$ is limited by gene regulatory constraints as assessed through the number of *cis*-regulatory elements in the promoter ($n = 3{,}907$ transcripts, Mann–Whitney U-test, two-sided $P = 0.000015$), and the number of transcription factors regulating a gene ($n = 2{,}905$ transcripts, Mann–Whitney U-test, two-sided $P = 0.0027$) illustrated for selection for total lifetime fitness under drought. Left, boxes and whiskers as in **d**. Right, mean ± s.e.m.



Extended Data Fig. 6 | Distributions of transcript–trait correlations for the three higher-level traits measured in the dry field environment. a, Absolute Pearson's correlations |*r*| of transcripts with leaf area (green). $n = 15{,}635$ transcripts. The cloud delineates transcripts (listed) that show significant linear or quadratic selection differentials for fecundity under drought conditions, and significant correlations with leaf area (Supplementary Text). **b**, Absolute Pearson's correlations |*r*| of transcripts with chlorophyll concentration (green). $n = 15{,}635$ transcripts. The cloud delineates a transcript that shows a significant quadratic selection differential for fecundity under drought conditions, and a significant correlation with chlorophyll concentration (Supplementary Text). **c**, Absolute Pearson's correlations |*r*| of transcripts with flowering time (lime). $n = 15{,}635$ transcripts. The cloud delineates transcripts (listed) that show significant linear selection differentials for flowering success under drought conditions, and significant correlations with early flowering (Supplementary Text).



Extended Data Figure 7 | Genome-wide association mapping of the genetic architecture of transcripts that covary significantly with fitness in the Indica population under drought conditions. Three out of eight transcripts are partially controlled by *trans*-eQTLs (illustrated for expression of the glycine-rich family protein-coding gene *Os11g0209000* under drought conditions). Supplementary Table 27 provides results for other transcripts and for expression principal components or eigengenes as suites of transcripts. **a**, PCA of 179,634 SNP markers from the Indica population that were selected for analysis; the three principal components, plus a fourth, were included as cofactors in the multi-locus linear mixed model. **b**, Distribution of expected versus observed *P* values for associations between SNP markers and *Os11g0209000* expression in a *Q*–*Q* plot. $n = 131$ genotypes; multi-locus linear mixed model, two-sided, Bonferroni-adjusted $P < 0.05$ for 179,634 SNP markers. **c**, The Manhattan plot indicates two significant *trans*-eQTL peaks for expression of *Os11g0209000* (gene location indicated with vertical red bar). Only the top approximately 5% of SNPs (10,000 SNPs) are shown.



Extended Data Fig. 8 | Genome-wide association mapping for fitness in the wet and dry field environments. Taking the top approximately 0.5% of SNPs (1,000 SNPs) with the strongest association to total lifetime fitness in the wet (magenta) and dry (blue) field conditions after genome-wide association mapping, we observed no enrichment for transcripts ($n = 809$ and 142 transcripts in the wet and dry fields, respectively) that were expressed in the leaves and had significant linear selection differentials S ($n = 408$ plants, t-test, two-sided, unadjusted $P < 0.05$) among transcripts ($n = 1{,}960$ transcripts in the wet field and $n = 1{,}671$ transcripts in the dry field) from genes in 100-kb regions surrounding these SNPs, compared to transcripts from genes in other genomic regions (χ^2, not significant (ns); two-sided $P = 0.862$ for the wet field and $P = 0.85$ for the dry field). Supplementary Table 27 provides genome-wide association mapping results for total lifetime fitness in wet and dry conditions, and for flowering success and fecundity under drought conditions.

Extended Data Table 1 | Phenotypic selection gradients, G-matrices and outcomes of selection for transcript levels in wet and dry conditions

Principal components	Selection gradients				G-matrix								Δ z		
Wet	γ	*P*	β	*P*	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	Direct	Indirect	Total
PC1	-8.47E-05	0.229	0.001	0.413	0.070	-0.013	-0.352	0.044	-0.108	-0.076	0.149	0.012	6.26E-05	0.002	0.002
PC2	-5.00E-05	0.792	0.001	0.732	-0.013	0	0.041	-0.069	-0.043	0.017	0.016	-0.008	0	6.69E-04	0.001
PC3	4.89E-04	0.169	0.002	0.471	-0.352	0.041	0.252	0.305	0.202	9.34E-05	-0.221	0.269	4.60E-04	-0.004	-0.004
PC4	3.86E-04	0.433	-0.004	0.197	0.044	-0.069	0.305	0.189	0.335	-0.024	-0.009	-0.010	-6.86E-04	-0.001	-0.002
PC5	9.54E-04	0.038	-0.004	0.204	-0.108	-0.043	0.202	0.335	0.086	0.010	-0.112	-0.044	-3.52E-04	-0.003	-0.004
PC6	-1.28E-04	0.814	-0.002	0.622	-0.076	0.017	9.34E-05	-0.024	0.010	0	-0.010	0.004	0	-1.50E-04	-1.50E-04
PC7	**1.78E-04**	**0.755**	**0.017**	**1.44E-06**	**0.149**	**0.016**	**-0.221**	**-0.009**	**-0.112**	**-0.010**	**0.455**	**0.090**	**0.008**	**9.04E-04**	**0.009**
PC8	2.29E-04	0.647	0.007	0.046	0.012	-0.008	0.269	-0.010	-0.044	0.004	0.090	0.077	5.60E-04	0.002	0.003
Dry	γ	*P*	β	*P*	PC1	PC2	PC3	PC4	PC5	PC6	PC7		Direct	Indirect	Total
PC1	-1.09E-04	0.779	0.001	0.820	0.253	-0.138	0.256	-0.019	0.047	0.015	-0.068		3.64E-04	0.005	0.006
PC2	-2.88E-04	0.747	-0.005	0.636	-0.138	0	0.199	-0.090	-0.041	-0.059	0.071		0	0.002	0.002
PC3	-2.79E-04	0.880	0.017	0.223	0.256	0.199	0	-0.273	0.378	-0.171	0.103		0	-0.003	-0.003
PC4	0.003	0.379	-0.006	0.722	-0.019	-0.090	-0.273	0.154	-0.007	0.035	-0.079		-0.001	-0.003	-0.004
PC5	0.002	0.555	0.004	0.816	0.047	-0.041	0.378	-0.007	0	-0.135	0.057		0	0.002	0.002
PC6	**0.002**	**0.556**	**0.032**	**0.070**	**0.015**	**-0.059**	**-0.171**	**0.035**	**-0.135**	**0.150**	**-0.060**		**0.005**	**-0.003**	**0.001**
PC7	0.001	0.691	0.002	0.897	-0.068	0.071	0.103	-0.079	0.057	-0.060	0.260		0.001	3.57E-05	0.001

The selection gradients describing nonlinear (*γ*) and linear (*β*) selection on principal components of genome-wide transcript abundance, the matrix of additive genetic variances and covariances of these principal components (G-matrix), and the outcome of selection (Δz) for total lifetime fitness in wet and dry conditions. $n = 408$ plants, *t*-test, two-sided, adjusted $P < 0.05$.

Extended Data Table 2 | Phenotypic selection gradients on transcript levels for flowering success, fecundity and lifetime fitness in dry conditions

Dry	Flowering Success				Fecundity				Total			
Principal components	γ	*P*	β	*P*	γ	*P*	β	*P*	γ	*P*	β	*P*
PC1	8.18E-06	0.973	0.002	0.618	-1.17E-04	0.409	-0.001	0.811	-1.09E-04	0.779	0.001	0.820
PC2	-2.24E-04	0.693	-0.005	0.443	-6.38E-05	0.845	1.27E-04	0.971	-2.88E-04	0.747	-0.005	0.636
PC3	4.18E-05	0.971	0.008	0.352	-3.21E-04	0.646	0.009	0.088	-2.79E-04	0.880	0.017	0.223
PC4	**0.004**	**0.062**	**-0.019**	**0.070**	**-0.001**	**0.221**	**0.013**	**0.014**	**0.003**	**0.379**	**-0.006**	**0.722**
PC5	0.002	0.271	0.003	0.788	-2.67E-04	0.799	0.001	0.862	0.002	0.555	0.004	0.816
PC6	**0.002**	**0.186**	**0.025**	**0.023**	**-0.001**	**0.371**	**0.007**	**0.301**	**0.002**	**0.556**	**0.032**	**0.070**
PC7	0.002	0.316	-0.008	0.486	-0.001	0.492	0.011	0.125	0.001	0.691	0.002	0.897

The selection gradients describing nonlinear (γ) and linear (β) selection on principal components of genome-wide transcript abundance for flowering success, fecundity and total lifetime fitness in dry conditions. n = 408 plants, z-test for flowering success and t-test for fecundity, two-sided, unadjusted $P < 0.05$.

nature research

Corresponding author(s): Michael D. Purugganan

Last updated by author(s): Nov 14, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The genomic data used were downloaded from SRA BioProjects PRJEB6180, PRJNA422249, and PRJNA557122. The reference genome used by STAR was Nipponbare IRGSP 1.0 (GCF_001433935.1) including plastids. A reference annotation was generated from Ensembl's IRGSP nuclear O. sativa genome annotation (1.0.37, ftp://ftp.ensemblgenomes.org/pub/plants/release-37/gff3/oryza_sativa) and supplemented with the Refseq Mitochondrial and Chloroplast annotations (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/433/935/GCF_001433935.1_IRGSP-1.0).
Data analysis	Leaf area was assessed using ImageJ v1.52 and GIMP v2.10.0. RNA-seq data were processed and analyzed using Drop-seq tools v1.12, STAR aligner v020201, Picard tools v2.9.0, DChip v2010.01 and R v3.4.3 packages edgeR v3.14 and lme4 v1.1. Selection analyses were run using custom-made scripts in Python version 2.7 that are available here in Supplementary Notes 1 and 2, and on GitHub in repositories icalic/Linear-regression-analysis and icalic/Logistic-regression-analysis. Gene-set enrichment analyses were performed using PlantGSEA v1. Statistical analyses were performed in R v3.4.3, further using packages lme4 v1.1 and corpcor v1.6.9. Genome analyses were performed using bbduk 37.66, bwa-mem v0.7.16a-r1181, the GATK GenotypeGVCFs engine v3.8-0-ge9d806836, vcftools v0.1.15, jvarkit suite v1, Beagle v4.1, plink v1.9 and GAPIT v3.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw phenotype data are available in the Supplementary Table and Source Data files. Raw RNA sequence data that support the findings of this study have been

deposited in the SRA under BioProject accession number PRJNA588478, and are publicly available. Files with processed RNA read count data and sample metadata are publicly available from Zenodo (https://zenodo.org/record/3533431 with DOI 10.5281/zenodo.3533431). A key to the RNA sequence data in SRA BioProject accession number PRJNA588478 can be found in Supplementary Table 4 as well as in the metadata file available from Zenodo.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Representative studies from the literature were used to determine sample size. These studies are cited in the manuscript.
Data exclusions	In the selection analyses, individuals that were severe outliers for the relative abundance of a transcript/trait (± 3 SD) were removed on a per-transcript/trait basis to satisfy the assumption of normality. This was a pre-established criterion.
Replication	All accessions were successfully planted and measured in biological triplicate in both field environments.
Randomization	Allocation of individual plants into experimental groups was random, and the planting order of individuals was randomized according to an alpha lattice design.
Blinding	The investigators were blinded to the genetic identity of individuals in the experiment during sampling, sample processing, and outcome assessment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging