nature communications

Article

https://doi.org/10.1038/s41467-024-53611-9

Wild rice *GL12* synergistically improves grain length and salt tolerance in cultivated rice

Received: 16 March 2024

Accepted: 16 October 2024

Published online: 01 November 2024

Check for updates

Yanyan Wang[1,2,3,8], Wenxi Chen[1,2,8], Meng Xing[1,2], Jiaqiang Sun[1], Shizhuang Wang[1], Ziyi Yang[1], Jingfen Huang[1,2], Yamin Nie[1], Mingchao Zhao[2,4], Yapeng Li[2,4], Wenlong Guo[1,2], Yinting Wang[1], Ziyi Chen[1], Qiaoling Zhang[1,2], Jiang Hu[2,5], Yunhai Li[6], Ke Huang[6], Xiaoming Zheng[1,2], Leina Zhou[1], Lifang Zhang[1], Yunlian Cheng[1], Qian Qian[1,5,7] ✉, Qingwen Yang[1,2] ✉ & Weihua Qiao[1,2] ✉

The abounding variations in wild rice provided potential reservoirs of beneficial genes for rice breeding. Maintaining stable and high yields under environmental stresses is a long-standing goal of rice breeding but is challenging due to internal trade-off mechanisms. Here, we report wild rice *GL12*^{W} improves grain length and salt tolerance in both *indica* and *japonica* genetic backgrounds. *GL12*^{W} alters cell length by regulating grain size related genes including *GS2*, and positively regulates the salt tolerance related genes, such as *NAC5*, *NCED3*, under salt stresses. We find that a G/T variation in *GL12* promoter determined its binding to coactivator GIF1 and transcription factor WRKY53. GIF1 promotes *GL12*^{W} expression in young panicle and WRKY53 represses *GL12*^{W} expression under salt stresses. The G/T variation also contributes to the divergence of *indica* and *japonica* subspecies. Our results provide useful resources for modern rice breeding and shed insights for understanding yield and salt tolerance trade-off mechanism.

The domestication of cultivated rice is considered as one of the most important events in human agricultural history[1]. Currently cultivated rice (*Oryza sativa*) is a staple food that feeds more than half of the world's population[2].To meet the growing demand for rice production, it is essential to cultivate high-yielding and environmentally resilient rice varieties. However, simultaneous improvement of yield and biotic/abiotic resistance remains a huge challenge for rice breeders because yield and stress tolerance are typically negatively correlated[2,3].

Grain size and shape are major determinants for grain yield and quality, and have been selected during early rice domestication[4]. Grain size and shape are quantitative traits regulated by a series of quantitative trait loci (QTLs). In the past ten years, many QTLs/genes regulating grain size have been identified through classical mutant analysis or map-based cloning. They are involved in several important genetic pathways, mediated by phytohormones, G-proteins, proteasomal degradation, protein kinases, and transcriptional factors[4–8]. Thus far, several genes for grain size, such as *GW2*, *GL3.1*, *GS5*, *GW8*, *GS2/GL2*, *GS3*, *GL7/GW7* and *GS9*, have been identified in Asian rice[9–15]. *GS3* was the characterized QTL that regulates grain length (GL), encoding a transmembrane protein. The natural variation in *GS3* contributes to grain-length differences between *indica* and *japonica* varieties[10,16]. *GL3.1* is a serine/threonine phosphatase that belongs to the PPKL family. It regulates grain length by mediating cell cycle progression through affecting the phosphorylation status of cell cycle

[1]State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China. [2]National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya 572024, China. [3]Nanjing Institute of Agricultural Sciences in Jiangsu Hilly Area, Jiangsu Academy of Agricultural Sciences, Nanjing, China. [4]Cereal Crop Institute, Hainan Agricultural Academy Sciences, Haikou, China. [5]China National Rice Research Institute, Hangzhou 310006, China. [6]Key Laboratory of Seed Innovation, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [7]Yazhouwan National Laboratory, Sanya, China. [8]These authors contributed equally: Yanyan Wang, Wenxi Chen. ✉e-mail: qianqian@caas.cn; yangqingwen@caas.cn; qiaoweihua@caas.cn

proteins, such as Cyclin-T1;3, thereby controlling grain yield[11,12,17]. Transcription factor *OsMADS1*, protein phosphatase *GL3-1/GL3.1/OsPPKL1* and transcriptional regulator *GS9* have been reported to control grain length by regulating cell proliferation in the grain hull[11,12,17–19]. *GS2* encodes transcription factor *OsGRF4*, which regulates grain size mainly by promoting cell expansion and slightly promoting cell proliferation. *OsGRF4* is regulated by microRNA396. The *GS2* allele with mutations in the targeting site of *OsmiR396* in *OsGRF4* disrupts the repression effects of *OsmiR396* on *OsGRF4*, leading to increased expression of *OsGRF4*, and resulting in large grains and high yield. *OsGRF4* physically interacts with transcription coactivators Os*GIF1*/2/3[5,11,20,21]. However, the genetic and molecular mechanisms by which these factors control grain size still remain to be further elucidated.

As the direct progenitor of modern rice, common wild rice (*Oryza rufipogon* Griff.) has an important role in cultivated rice breeding and been considered as valuable resources for improving rice yield and environmental resilience. Cultivated crops, such as wheat, maize, and rice, have larger seeds than their wild relatives[22]. Almost all cultivated rice varieties have larger seeds than *rufipogon*. Although the overall phenotypic characteristics of wild rice, especially yield-related traits, are inferior to those of cultivated rice, studies have show that genes in wild rice are potentially useful for yield-related trait improvement[23]. However, so far there are no positive regulation alleles for yield reported from wild rice. Cultivated rice is also a salt-sensitive plant, wild rice harbors especially genes controlling adaptation to unfavorable environments which were lost during domestication. The exploration and utilization of salt tolerance genes in wild rice resources is the main strategy of rice salt tolerance breeding. However, the elite resistant traits often linked with inferior yield traits in wild rice. Therefore, the dissection of wild rice QTLs/genes that synergistically increase yield and salt tolerance would be crucial for rice breeding.

In our previous study, we identified a *QTL-qGL12* from wild rice, which regulates grain length in cultivated rice[24]. Here, we find that the overexpression of the wild rice $GL12^{W}$ allele simultaneously increases grain length and salt tolerance. A G/T variation in the *GL12* promoter region, interacting with transcription factor regulator *OsGIF1* (*Growth-regulating factor interacting factor 1*) and transcription factor *OsWRKY53*, is involved in the grain length and salt tolerance regulating pathways. In addition, it is found that $GL12^{W}$ functions in different rice genetic backgrounds including both *indica* and *japonica*. Our results provide resources for rice breeding and insights for synergistic mechanism of rice yield and salt tolerance.

Results

Identification and confirmation of the wild rice *qGL12* candidate gene

To identify wild rice genes underlying the natural variations in grain length (GL) and 1000-grain weight (TGW), we constructed a set of chromosome segment substitution lines (CSSLs) with an *indica* variety '9311' as the recurrent parent and the *O.rufipogon* as the donor parent. One major QTL for GL, *qGL12*, was identified across six environments, and mapped using molecular markers on chromosome (Chr.) 12[25]. One CSSL line, CSSL41, which has a long GL phenotype and carries the QTL-*qGL12*, was selected for further fine mapping (Supplementary Fig. 1a). An F_2 population was constructed from a cross between CSSL41 and 9311 (Fig. 1a). Using bulked-segregant analysis (BSA), and F_3 and F_4 homozygous recombinant analysis, *qGL12* was finally narrowed to a-16-kb interval between molecular markers WR12.4 and RM28586, which contains three predicted ORFs (Supplementary Fig. 1b, Fig. 1a, b). Among them, *LOC_Os12g39640* encodes a MYB transcription factor and sequence analysis revealed 11 single nucleotide polymorphisms (SNPs) and six insertion-deletion (InDels) polymorphism in the 1.8-kb promoter region and four SNPs in the coding region causing amino acid change between wild rice and 9311 (Supplementary Fig. 2a). *LOC_Os12g39650* encodes a tubulin and *LOC_Os12g39660* encodes a calcium-transporting ATPase, and there were eight and 17 SNPs between wild rice and 9311 in these two alleles respectively (Supplementary Fig. 3). We designed a cleaved amplified polymorphic sequence (CAPS) marker associated with wild rice *LOC_Os12g39640* allele (Supplementary Fig. 2b), and a near-isogenic line (NIL) of *qGL12* was developed with marker-assisted selection (Fig. 1a, c). We tested the expression levels of these three genes in young panicles of the NIL and 9311, among them, expression levels of *LOC_Os12g39640* and *LOC_Os12g39650* were significantly higher in the NIL than in 9311 (Fig. 1d). Compare with 9311, the NIL showed significantly increased GL and TGW, with 8.1% and 12.9%, respectively (Fig. 1f, g). To identify the candidate gene, we overexpressed (OE) these three genes from wild rice in 9311 background individually, both of *LOC_Os12g39640*-OE and *LOC_Os12g39650*-OE showed significantly increased GL (5.1% and 2.2%, respectively) than 9311, and *LOC_Os12g39640*-OE have much more increased GL than that of *LOC_Os12g39650*-OE (Fig. 1e–g). Given *LOC_Os12g39650* is a tubulin gene, the *LOC_Os12g39640* gene is the most likely candidate gene. Moreover, we knocked out (KO) *LOC_Os12g39640* in the NIL background using CRISPR/Cas9 technology. The GL and TGW of NIL-KO plants decreased significantly, and the overexpression of *LOC_Os12g39640* from 9311 did not improve GL in 9311 genetic background[24] (Supplementary Fig. 4), which confirmed that *LOC_Os12g39640* was the candidate gene and therefore termed *GL12*.

In addition, we generated transgenic lines in *japonica* variety 'Zhonghua11 (ZH11)' background, which overexpressed *GL12/LOC_Os12g39640* alleles from wild rice ($GL12^{W}$) and 9311 ($GL12^{9311}$), respectively. The $GL12^{W}$-OE lines showed significantly increase in GL and TGW than wild type ZH11 while $GL12^{9311}$-OE lines had no difference (Fig. 1f–j, Supplementary Fig. 5). We also overexpressed $GL12^{W}$ in 'Nipponbare (NIP)' background and the OE lines showed significantly increase in GL and TGW than that of wild type (Supplementary Fig. 6d, e). These data demonstrated that $GL12^{W}$ has stable positive effect on GL in different genetic backgrounds.

$GL12^{W}$ improves grain length by reprograming genes expression and altering cell length

GL12 encodes a MYB transcription factor, which belongs to a family with more than 2,00 members in rice[26,27]. From the NCBI database, we performed phylogenetic analysis and revealed that GL12 was clustered with MYB transcription factor gene family in rice and other species, closest with *japonica* homologous gene (Supplementary Fig. 7). The $GL12^{W}$::GFP fusion protein was localized in the nucleus in tobacco epidermal cells and rice protoplasts (Fig. 2a). These results indicated that the GL12 protein could function in the nucleus and was consistent with the role of GL12 as a transcription factor.

To identify the molecular mechanism and downstream genes regulated by $GL12^{W}$, we first performed RNA-sequencing analysis using young panicles from NIL and NIL-KO plants. As a result, 8079 differentially expressed genes (DEGs) were identified with the stringent criteria ($|\log 2\ (FC)| > 1$, and $FDR < 0.05$). Among these DEGs, 4615 were upregulated and 3464 were downregulated in the NIL compared with the NIL-KO lines. Gene Ontology identified that $GL12^{W}$ mostly participated in biological processes, and KEGG pathway enrichment analysis revealed that $GL12^{W}$ affected the expression of genes involved in metabolism, genetic information processing, environmental information processing, organismal systems, and cellular processes (Supplementary Fig. 8). In addition, we found that several grain shape related genes were downregulated in young panicles of the NIL-KO relative to the NIL (Fig. 2b), including previously reported *GS2/GRF4* genes[11]. We also confirmed the expression levels of grain shape related genes in the NIL and 9311, *GS2*, *GRF3*, and *GRF6* had significant higher expression levels in the NIL than that of 9311 (Fig. 2c). These results indicated that $GL12^{W}$ might regulate grain shape by affecting the expression of GRFs.

https://doi.org/10.1038/s41467-024-53611-9

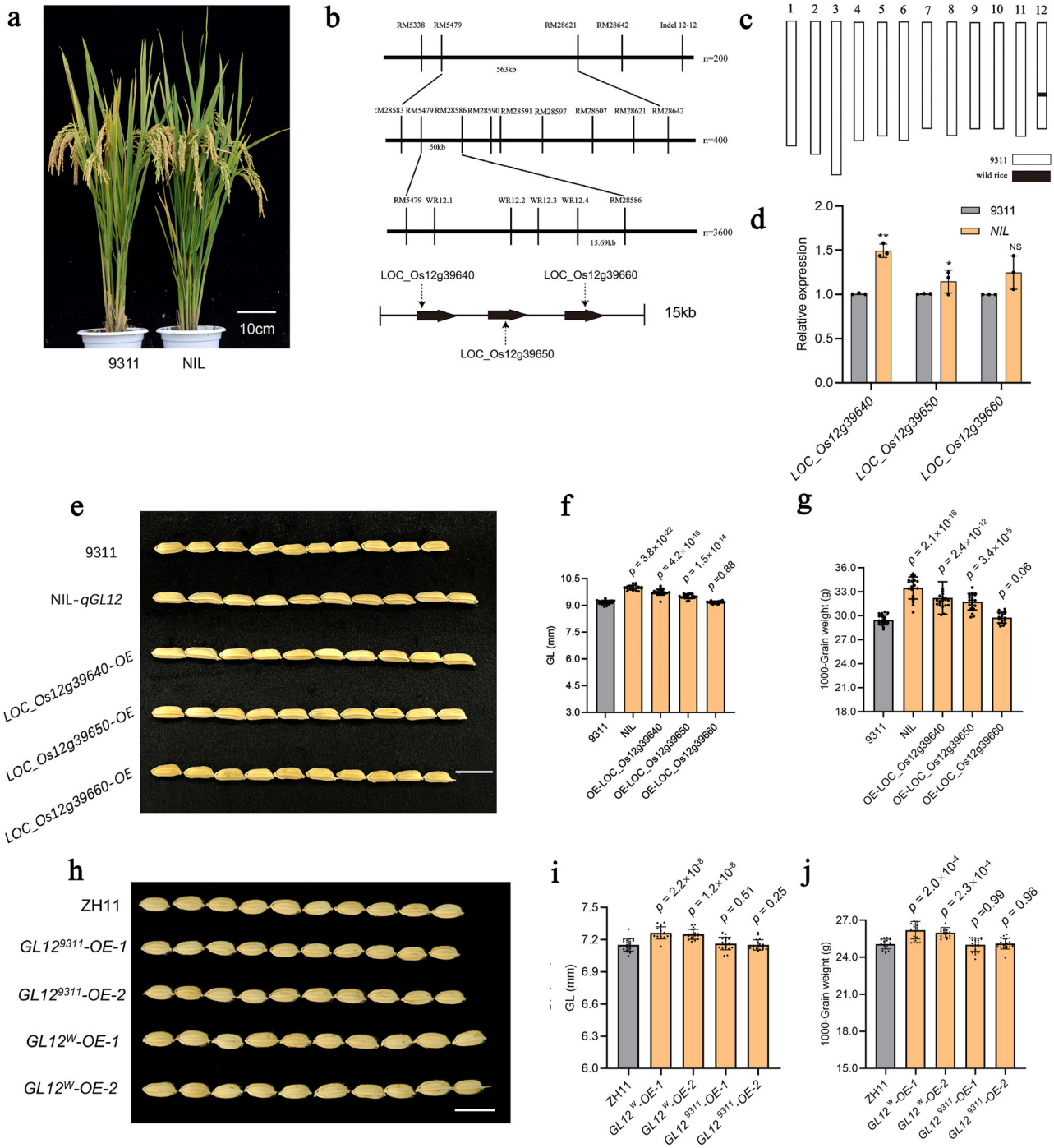

Fig. 1 | Wild rice *GL12*^{W} increases the grain length and weight in different genetic background. a The near isogenic line (NIL) of *qGL12* and recurrent parent 9311. **b** Fine mapping of *qGL12*. **c** Schematic diagram showing 9311 and wild rice chromatin. **d** Expression analyses of three genes in the *qGL12* interval. The 2–5 cm young panicles were used for RT-qPCR. The relative abundances in 9311 plants were set to be one. Values given are means ± SD ($n = 3$) Statistical analysis was performed by two-tailed Student's *t*-test $^*p \le 0.05$, $^{**}p \le 0.01$, NS no significant difference. The *p*-values were 3.7×10^{-4},0.028 and 0.086, respectively. **e** Grains of 9311, NIL, and overexpression lines of three genes in *qGL12*. Bar = 1 cm. **f**, **g** Statistical analysis of grain length (GL) and 1000-grain weight (TGW) of lines in (**e**). Data were given as means ± SD (n = 21/20/22/21/19, 21/20/22/21/19 biological replicates). One-tailed Student's *t*-test was used to generate the *P* values. **h** Grains of *GL12* overexpression lines in Zhonghua11(ZH11) genetic background. Bar = 1 cm. **i**, **j** Statistical analysis of grain length and 1000-grain weight of lines in (**h**). Data were given as means ± SD (n = 21/20/22/21/19, 21/20/22/21/19 biological replicates). One-tailed Student's *t*-test was used to generate the *P* values. Source data are provided as a Source Data file.

The rice spikelet hull, which is coordinately controlled by cell proliferation and cell expansion, was proposed to restrict the rice grain shape[4]. To further understand the function of *GL12* in grain shape, we investigated the spikelet hulls of 9311, *GL12*^{W}-OE and NIL before fertilization. The results verified that the longer GL in the NIL and *GL12*^{W}-OE was the result of cell expansion in spikelet hull (Fig. 2d, e). Furthermore, we examined the cross-section of spikelet hulls, and showed that the outer parenchyma cell layer in *GL12*^{W}-OE and NIL plants were longer and had substantially longer spikelet perimeter than the corresponding layer in 9311 plants (Fig. 2f). Cell division and expansion in eukaryotic organisms are controlled by a highly conserved basic cell cycle machinery[28,29]. To investigate the relationship of

https://doi.org/10.1038/s41467-024-53611-9

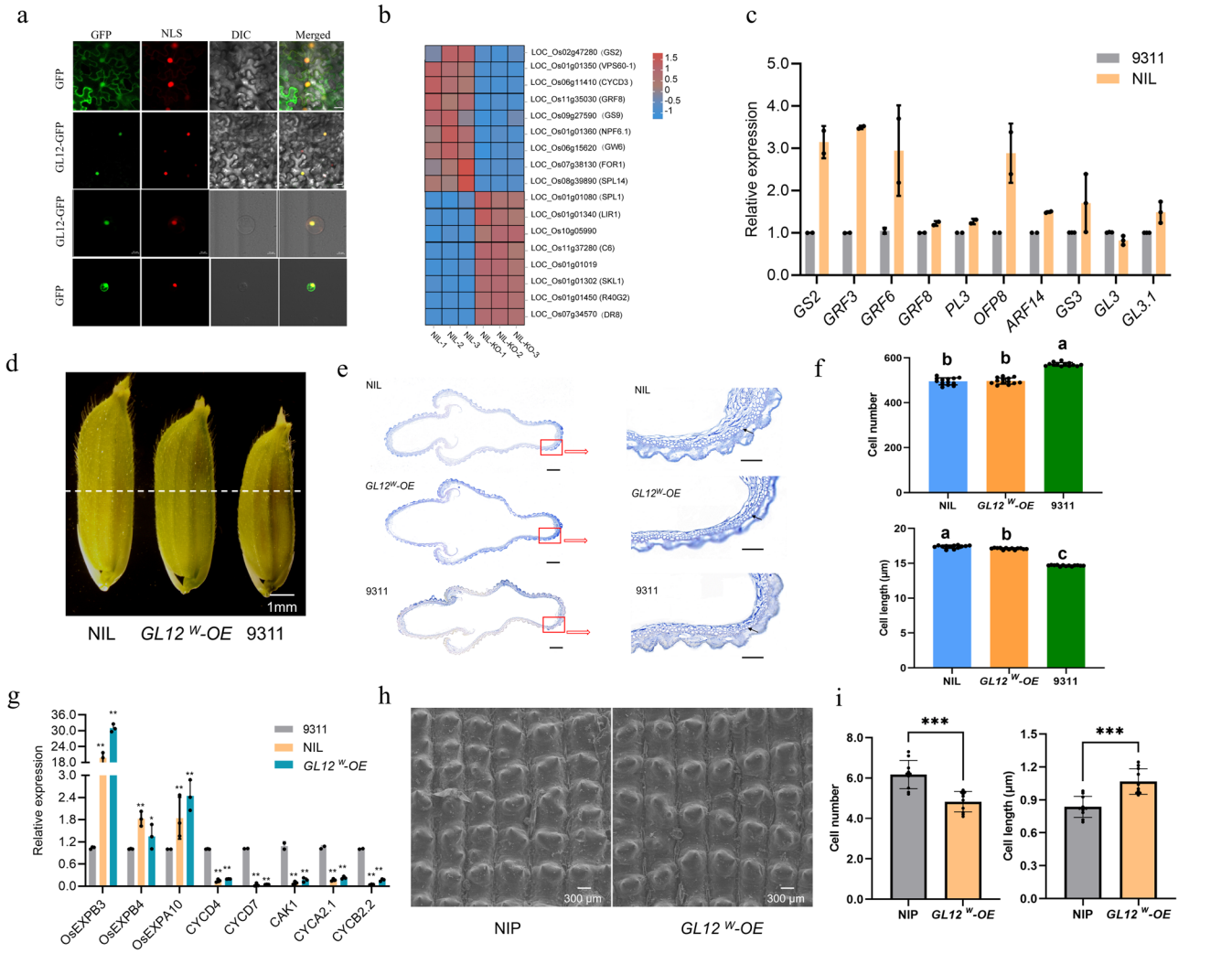

Fig. 2 | *GL12*^{W} regulated grain size related genes and improves grain length by altering cell length. a Subcellular localization of GL12 (35S::GL12-GFP) in tobacco leaf cells and rice protoplasts. The fusion construct 35S::GFP/RFP marker, and GL12 (35S::GL12-GFP) / RFP marker were co-transformed into rice protoplasts and tobacco leaf cells, respectively. Scale bar = 10 μm. **b** Heat map diagram of the expression levels for rice seed size related genes. Data were from transcriptome using 2–5 cm young panicles. The heat map was drawn according to FPKM values. Columns and rows in the heat map represent samples and genes, respectively. Color scale indicates fold changes in gene expression. **c** Transcription levels of the grain size related genes were confirmed through qRT-PCR analysis with three biological replicates in 9311 and NIL. *OsActin* was used as a control. Data show means ± SD. **d**–**f** Spikelet hulls before anthesis (**d**). The white dashed line indicates the sites of the cross-sections shown in (**e**). The images on the right panel showed the close-up views of the boxed region in (**e**), Scale bar = 100 μm. The red arrows indicate the lower epidermal cells. Comparisons of spikelet of cell number, and cell length between 9311, NIL and *GL12*^{W} transgenic plants were shown in (**f**), letters a–c in indicate the different means according to Duncan's multiple range test ($P < 0.05$). **g** Quantitative real-time PCR analysis of cell cycle-related and expansion gene expression in 9311, NIL and *GL12*^{W}-OE lines. *OsActin* was used as a control. Data show means ± SD of three biological replicates. Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, **$p \le 0.01$, NS no significant difference). Scanning electron microscopic analysis of the lemma of NIP and *GL12*^{W}-OE lines (**h**), and statistical analysis of longitudinal cell number and cell length in the lemmas of spikelet hulls (**i**) $n = 10$ view of 10 areas in (**i**). Statistical analysis was performed by two-tailed Student's *t*-test (***$p < 0.001$). Source data are provided as a Source Data file.

GL12 with genes regulating plant cell cycle and expansion, we analyzed the expression of eight related genes in the NIL and *GL12*^{W}-OE relative to 9311 by real-time PCR. The transcription levels of five genes, *CYCD4*, *CYCD7*, *CAK1*, *CYCA2.1* and *CYCB2.2*, were greatly reduced in *GL12*^{W}-OE and NIL as compared to 9311 plants. In contrast, the expression of *OsEXPB3*, *OsEXPB4* and *OsEXPA10* were significantly elevated in the NIL and *GL12*^{W}-OE (Fig. 2g). Additionally, we checked the grain hulls of *GL12*^{W}-OE NIP lines and NIP wild type using a scanning electron microscope. There were significantly fewer longitudinal cells in spikelet hulls of *GL12*^{W}-OE than in NIP; however, the cells of spikelet hulls of *GL12*^{W}-OE were significantly longer than those of Nip (Fig. 2h, i). Thus, *GL12*^{W} functions putatively as a positive regulator upstream of cell expansion genes, and its overexpression may result in an increase in cell expansion in longitudinal direction. These findings suggested that *GL12*^{W} functions as a regulator of both cell proliferation and cell expansion.

The native promoter enhances *GL12*^{W} expression in young panicle and increases grain length

Compared with *GL12*^{W}, *GL12*^{9311} coding region has four mutant sites and has no function on GL improvement (Supplementary Fig. 2a). There were 17 SNPs/InDels in the promoters of *GL12* between 9311 and wild rice (Supplementary Fig. 2a). We performed real-time PCR to detect the expression of *GL12* in young panicles of 9311 and the NIL, and found that the transcription levels of *GL12* in the NIL were significantly higher in all panicle development stages than those in 9311 (Fig. 3a). We then performed the transient expression assays in *Nicotiana benthamiana* leaves, and found that the promoter of *GL12*^{W} (*Pro*^{W}) had stronger

https://doi.org/10.1038/s41467-024-53611-9

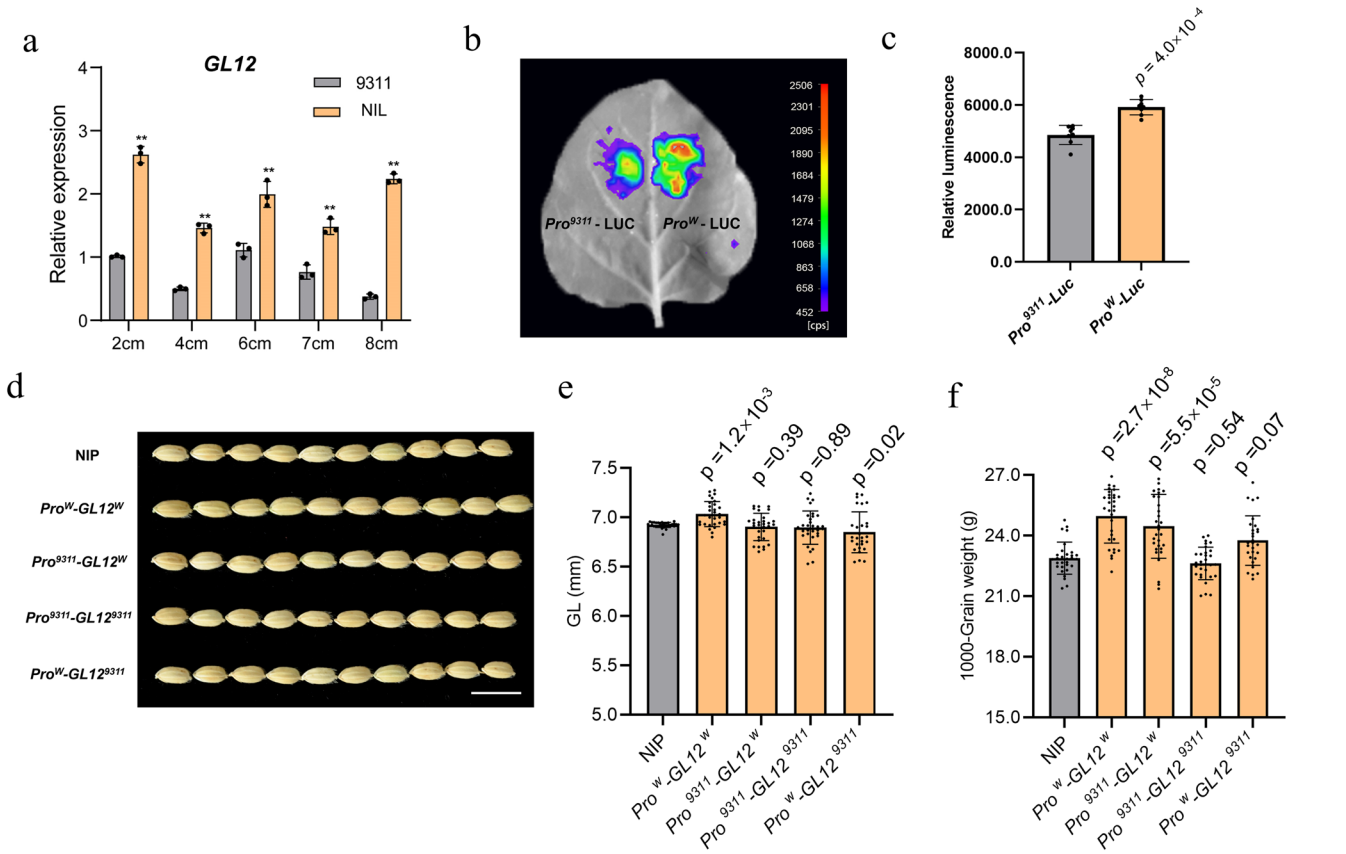

Fig. 3 | Promoter activity assays for the *GL12* alleles in 9311 and wild rice. **a** Relative quantitative real-time PCR analysis of *GL12* expression in young panicles of 2, 4, 6, 7, and 8 cm in 9311 and NIL. *OsActin* was used as a control. qRT-PCR analysis with three biological replicates. Data showed means ± SD. (The *p*-values were 2.9×10^{-5}, 3.5×10^{-5}, 2.7×10^{-3},$1.7 \times 10^{-3}$and 3.4×10^{-6}, respectively). **b** Transient expression assays in *Nicotiana benthamiana* leaves showing the promoter activities of *GL12*, and the quantification of the relative luminescence intensities were shown in (**c**) The quantification of the relative luminescence intensities (Data show means ± SD, $n = 8$). Statistical analysis was performed by one-tailed Student's *t*-test. **d** Grains of the transgenic lines in Nipponbare background. *GL12* alleles form the wild rice and 9311 were driven by both promoters. Bar = 1 cm. Statistical analysis of grain length (**e**) and grain weight (**f**) in transgenic plants ($n = 30/31/31/29$). Data were given as means ± SD. One-tailed Student's *t*-test was used to generate the *P* values. Source data are provided as a Source Data file.

activity than that of $GL12^{9311}$ (Pro^{9311}) (Fig. 3b, c). Furthermore, we generated transgenic plants in the NIP background that harbored the promoter and coding regions of *GL12* from 9311 and NIL, respectively. We found that the GL of Pro^{W}:$GL12^{W}$ transgenic lines were much longer than those of NIP and the Pro^{9311}:$GL12^{W}$ transgenic lines and NIP (4.4% and 3.5% increase than NIP and Pro^{9311}:$GL12^{W}$, respectively) (Fig. 3d–f), demonstrating that $GL12^{W}$ has a stronger effect on GL under its native promoter.

The G/T variation in *GL12* promotes its binding by GIF1

The *OsGIF1* (GRF-interacting factor) gene in rice, which encodes a transcription coactivator, has been reported to interact with the transcription factor GS2/OsGRF4 and positively promote grain size in rice[30–32]. Firstly, we detected the physical interaction between GL12 and OsGIF1. The yeast two-hybrid assay (Y2H) and firefly luciferase complementation experiment showed that both $GL12^{W}$ and $GL12^{9311}$ can interact with OsGIF1 protein (Fig. 4a, b). To determine whether *GL12* was regulated by GIF1, we investigated the expression levels of *GL12* in the *GIF1* overexpression and CRISPR/Cas9 knock out lines in ZH11 background. Compared to ZH11, the *GL12* expression levels in *GIF1-OE* lines were significantly improved while no significantly difference was found in *GIF1-KO* lines (Fig. 4c). Furthermore, we generated four constructs with truncated promoters of Pro^{W} and Pro^{9311} of *GL12* alleles from NIL and 9311, respectively (Fig. 4d). The yeast one-hybrid assay results demonstrated that GIF1 could only bind the promoter of *Pro-2*, and GIF1 bind the Pro^{W}*-2* much stronger than Pro^{9311}*-2*. (Fig. 4d; Supplementary Fig. 9). Unexpectedly, the truncation of Pro^{W}*-2* only harbored one SNP between the wild rice and 9311, which generated a G/T substitution at the -1117th position (Fig. 4d). These results were further supported by the electrophoresis mobility shift assay (EMSA) that used biotin-labeled probes containing the G/T SNP in the Pro^{W}*-2* (Fig. 4e). These results indicated that the G/T site contributed to the binding of GIF1 to the promoter of *GL12* and that was a direct target regulated by *GIF1*. Moreover, we knocked out the *OsGIF1* in NIL background by CRISPR/Cas9 technology, the grain length of *GIF1*-KO-NIL significantly decreased (Supplementary Fig. 10).

$GL12^{W}$ improves salt tolerance in rice

Besides $GL12^{W}$ could increase the grain length and grain weight, we found that $GL12^{W}$ could also improve salt tolerance. After extensive salt treatment (10 days of 150 mM NaCl treatment, recovered for seven days), both the NIL and $GL12^{W}$-OE showed significant higher salt tolerance than that of 9311, and the NIL showed much stronger salt tolerance than $GL12^{W}$-OE lines. The survival rate of the NIL, $GL12^{W}$-OE lines, and 9311 were 80%, 60% and no more than 30%, respectively. And the fresh weight and dry weight of the NIL were significantly higher than that of 9311 and $GL12^{W}$-OE lines. Meanwhile the NIL-KO line showed hypersensitivity to salt stress (Fig. 5a, b). Moreover, the overexpression of $GL12^{W}$ in the NIP genetic background also significantly improved salt tolerance, compared with NIP wild type plants (Supplementary Fig. 11). There was no significant difference in salt tolerance between $GL12^{9311}$-OE lines and 9311 (Supplementary Fig. 12). We then investigated the

https://doi.org/10.1038/s41467-024-53611-9

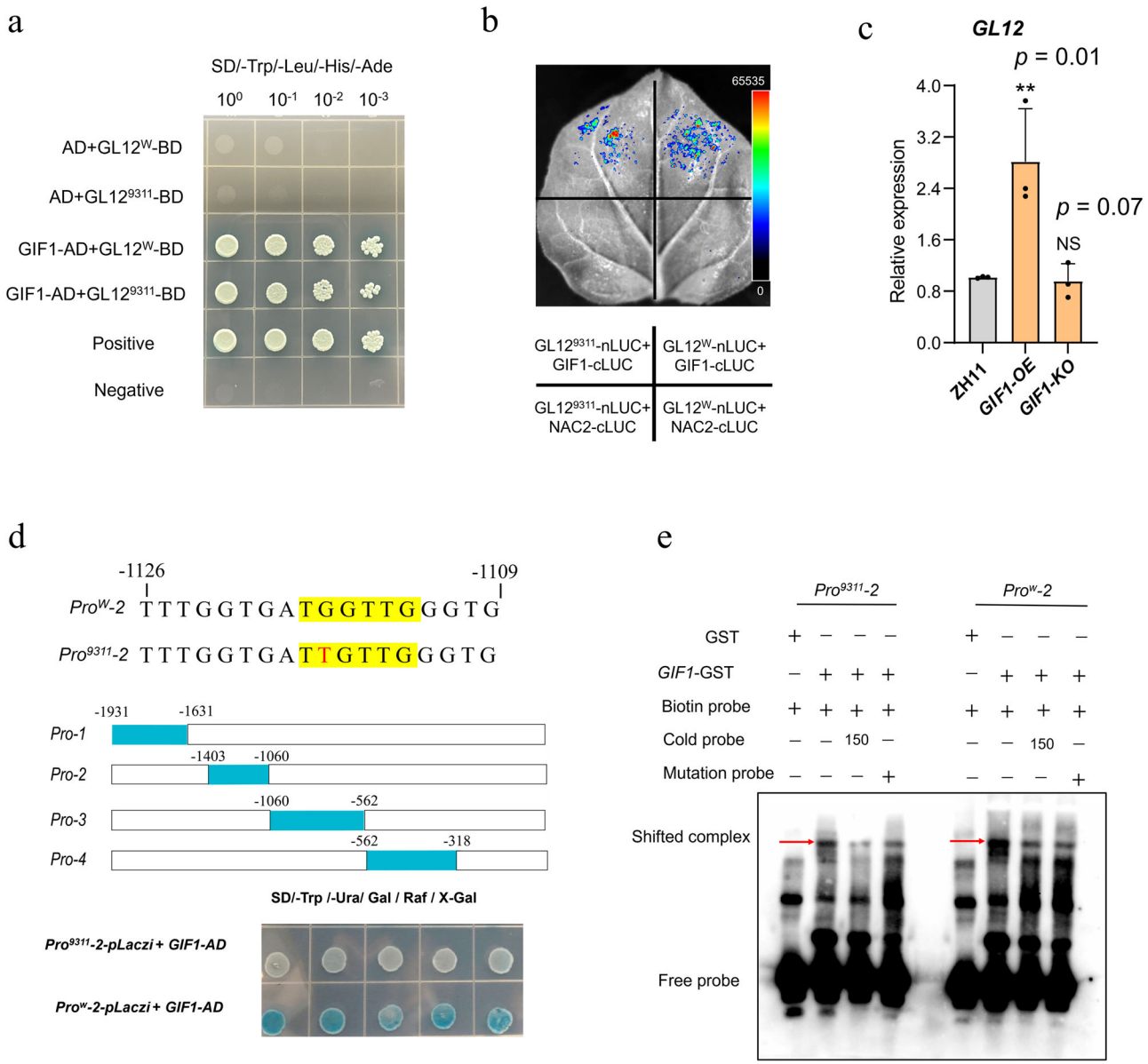

Fig. 4 | The G/T variation in *GL12* promotes its binding by GIF1. a Yeast two-hybrid assay showed that GIF1 physically interacted with GL12^{W} and GL12^{9311}. **b** LCI assays in *N. benthamiana* leaves showing the GIF1 and GL12 interaction. *N. benthamiana* leaves were transformed by injection of *Agrobacterium* GV3101 cells harboring GL12-nLUC and GIF1-cLUC plasmids. Luciferase complementation was observed for the GL12-nLUC and cLUC- GIF1 combination, the negative control was GL12-nLUC + NAC2-cLUC, *NAC2* was a nuclear localization gene, whereas no obvious signal was observed for the negative controls. **c** The *GL12* expression levels in ZH11 and *GIF1* overexpression (*GIF1*-OE) and knockout (*GIF1*-KO) lines. *OsActin* was used as a control. qRT-PCR analysis with three biological replicates. Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, **$p \le 0.01$, NS no significant difference). Data showed means ± SD. **d** Schematic illustration of the four truncations of *GL12* promoters. The MYB motif was highlighted in yellow and the G/T variation was marked in red. Yeast one-hybrid assays were shown. *GIF1* only interacted with *Pro^{W}-2* truncation. Other yeast one-hybrid results were shown in Supplementary Fig. 9. **e** EMSA assays showed that GIF1 bound to the *Pro^{W}-2* truncation. EMSA assays were repeated three times. GST-GIF1 fusion protein was expressed in *E. coli*. The probes were labeled with biotin. Source data are provided as a Source Data file.

GL12 expression pattern under salt treatment. Before salt treatments, *GL12* had higher basic expression levels in leaves of 9311 than that of NIL; after salt treatment, the *GL12* expression in 9311 was induced rapidly compared with that in the NIL (Fig. 5c). These results indicated that *GL12^{W}* and *GL12^{9311}* have different expression patterns under salt treatment. Moreover, we analyzed the expressions of salt tolerance-related genes in the NIL and 9311. Compared with 9311, the *NCED3* and *NAC5* in the NIL were significantly induced after salt treatments (Fig. 5d, e). We further determined the effects on salt tolerance of *OsGIF1* and *GS2*, and found that the *GIF1-OE* lines showed higher salt tolerance while *GIF1*-KO lines were salt sensitive compared with wild type ZH11 (Supplementary Fig. 13). The *GS2*-NIL line, which had increased *GS2* expression levels[11], also showed higher salt tolerance than its genetic background ZH11 (Supplementary Fig. 13). These results indicated that *GL12^{W}* also played an important role in regulating the salt stress response in rice.

The G/T variation in *GL12* promoter determines its repression by WRKY53

Compared with *GL12^{9311}*, the *GL12^{W}* expression was higher in young panicles but lower in leaves under salt stress (Figs. 3a, 5c). We postulated that *GL12^{W}* might be regulated by another transcription factor. Because *OsWRKY53* was reported to be a key regulator of salt tolerance in rice[33], we investigated whether *GL12* was involved in the *OsWRKY53*

https://doi.org/10.1038/s41467-024-53611-9

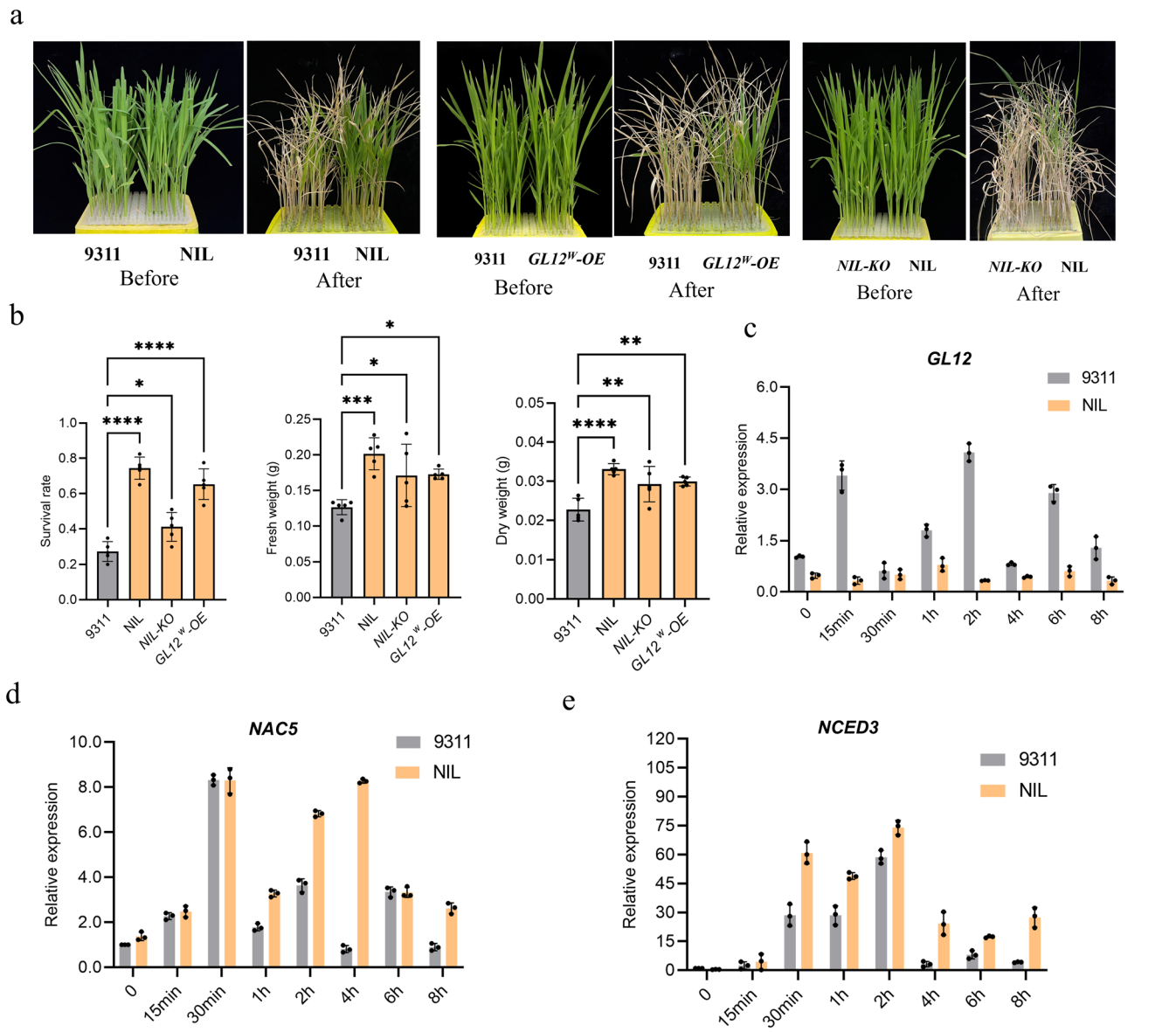

Fig. 5 | *GL12*^{W} improves salt tolerance in rice. a Photos of 9311, NIL and *GL12*^{W} overexpression lines before and after salt treatments. 20-day-old seedlings were treated with 150 mM NaCl for10 days and recovered for seven days. **b** Statistical analysis of survival rates and fresh and dry weight of seedlings in (**a**). Five biological replicates were used to determine the survival rate. Data were means ± SD ($n = 5$). Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, **$p \le 0.01$, NS no significant difference). **c** *GL12* expression analysis in 9311 and NIL. 20-day-old seedlings were treated with 150 mM NaCl for indicated time. qRT-PCR analysis with three biological replicates. Data were means ± SD. Expression analysis of rice salt tolerance related genes *NAC5* (**d**) and *NCED3* (**e**) in 9311 and NIL, 20-day-old seedlings were treated with 150 mM NaCl. *OsActin* was used as a control. qRT-PCR analysis with three biological replicates. Data showed means ± SD. Source data are provided as a Source Data file.

regulated salt tolerance pathway. The *OsWRKY53*-OE and CRISPR/Cas9 knockout lines in *japonica* variety 'LongGeng 11' (LG11) background were used to determine the *GL12* expression levels. The *GL12* expression was significantly increased in the *WRKY53*-KO lines and significantly decreased in the *WRKY53*-OE lines, compared with wild type LG11 (Fig. 6a), which was consistent with the *WRKY53*-KO lines showed stronger salt tolerance[33]. To study the interaction between WRKY53 and GL12, we analyzed the promoter region of *Pro-1* to *Pro-4* of *GL12*^{W} and *GL12*^{931}. Consistent with GIF1, the Y1H experiments demonstrated that WRKY53 could only bind the truncated promoter *Pro-2* (Fig. 6b; Supplementary Fig. 9), which harbored only one G/T variation between the wild rice and 9311 (Fig. 4d). Interestingly, Y1H and EMSA proved that the binding ability of *Pro*^{W}*-2* with WRKY53 was stronger than that of *Pro*^{9311}*-2* (Fig. 6c). We further tested the capacity of WRKY53 protein on driving the expression of *GL12* using dual-luciferase assays. The promoter of *GL12* was fused to the luciferase reporter gene and co-introduced into rice protoplast with constructs driving the expression of WRKY53 (Fig. 6d). In the presence of WRKY53, the luciferase (LUC) activity driven by the promoter of *GL12*^{W} was largely reduced than that of *GL12*^{9311} (Fig. 6e, f). These results strongly suggested that WRKY53 regulated salt tolerance by repressing the *GL12*^{W} expression.

Natural variations in GL12 contribute to the divergence of *indica* and *japonica* subspecies

To identify whether *GL12* has been artificially selected during rice domestication, we further analyzed the genomic sequences of this gene in rice germplasm resources from the 3K Rice Genomes Project[34,35], the Rice Super Pangenome Information Resource Database (RiceSuperPIRdb)[36] and our own re-sequencing data. Firstly, we compared the nucleotide diversity (PI-value) of the 100-kb region flanking

https://doi.org/10.1038/s41467-024-53611-9

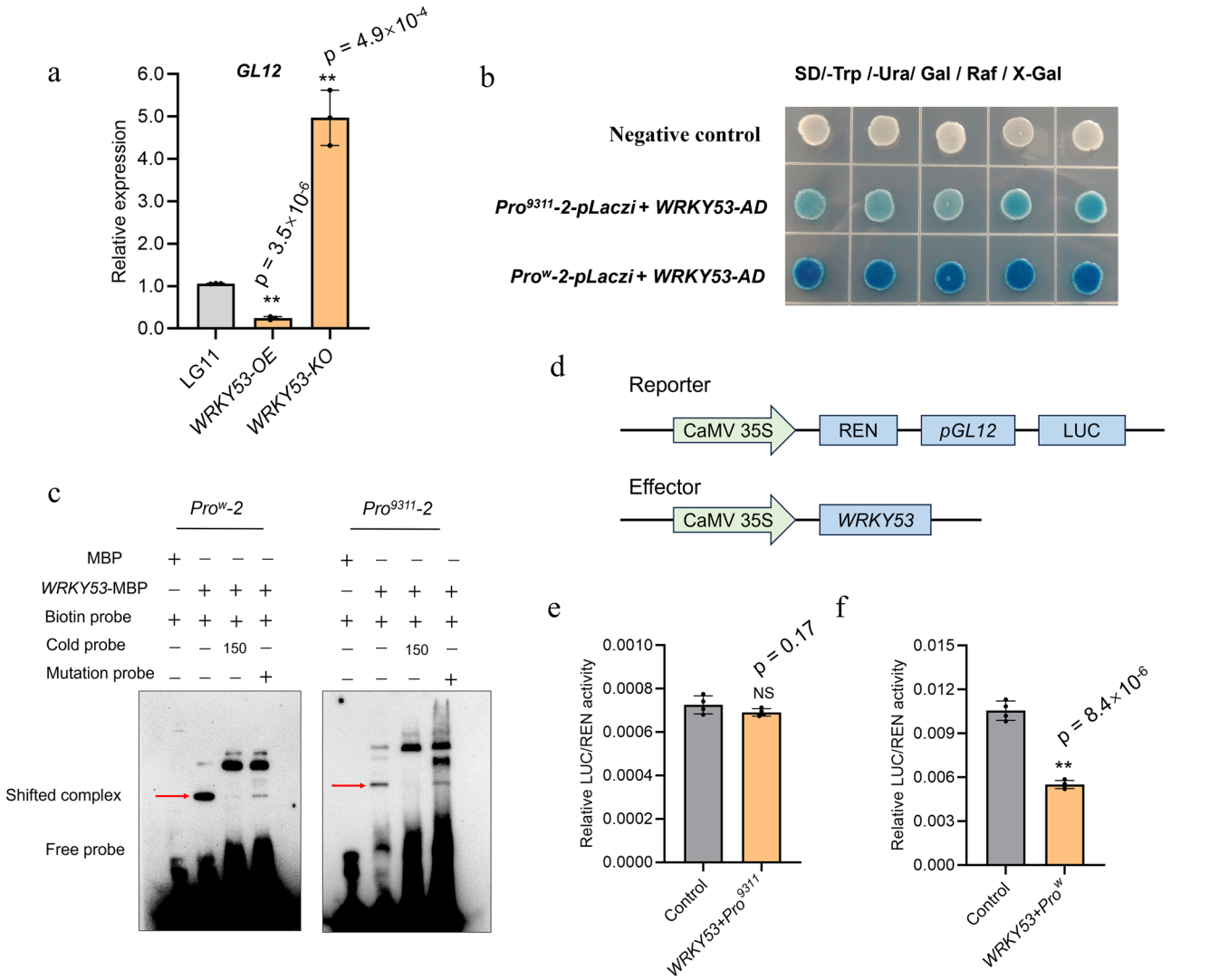

Fig. 6 | The G/T variation in *GL12* determines its repression by WRKY53. a The *GL12* expression levels in Longgeng11 (LG11) wild type, and WRKY53 overexpression and knockout lines. Young leaves from 20-day-old seedlings were used for RNA extraction and Real-time PCR analysis. *OsActin* was used as a control. qRT-PCR analysis with three biological replicates. Data showed means ± SD. Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, **$p \le 0.01$, NS no significant difference). **b** Yeast one-hybrid assays proved that the transcription factor WRKY53 only bound to the *Pro-2*. Other yeast one-hybrid assay results were shown in Supplementary Fig. 9. **c** EMAS assays showed that WRKY53 bound to *Pro^{W}-2*. EMSA assays were repeated three times. MBP-WRKY53 fusion protein was expressed in *E. coli*. The probes were labeled with biotin. **d** Schematic diagrams of the effectors and reporters used in the dual-LUC experiments. Transient expression assays in *N. benthamiana* protoplasts showing that WRKY53 could reduce the transcriptional activity of GL12^{W} (**f**), but not the GL12^{9311} (**e**). The activities of firefly luciferase (LUC) and renilla luciferase (REN) were determined 16 h post transformation. The relative luciferase activities in different samples were calculated by normalizing the LUC values against REN. LUC/REN analysis with five biological replicates. Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, **$p \le 0.01$; NS, not significant). Data show means ± SD. Source data are provided as a Source Data file.

GL12 with our re-sequencing data (56 *indica*, 51 *japonica*, and 52 wild rice accessions). The results showed that the PI values of the *GL12* locus in wild rice were significantly higher than those in *indica* cultivars, but no significant difference with those *japonica* cultivars. In addition, the PI ratio in *indica* to that in wild rice was significantly lower than in *japonica* to that in wild rice in the *GL12* locus (Fig. 7a, b).

Furthermore, with the data from the 3K Rice Genome Project and rice super pangenome data, we found that the G/T substitution (1117-bp upstream of the start codon) in *GL12* was unbalanced between the two cultivated rice subspecies, the majority of *indica* varieties were haplotype-T at the G/T site whereas most of *japonica* varieties were haplotype-G (Supplementary Fig. 14). We further investigated the role of G/T variation in the promoter of *GL12* on GL using the data from 3K Rice Genomes Project of the *indica* varieties. Among them, 15 and 894 *indica* varieties carried the G and T types, respectively. The *indica* varieties with the haplotype-G exhibited longer grains than those with the haplotype-T (Fig. 7c). The haplotype classification results showed that the 228 rice varieties could be divided into a total of 20 haplotypes (Fig. 7d). Interestingly, the rice varieties within the blue dashed line are all Hap-T at promoter -1117 position, while those within the red dashed line are Hap-G. Among them, most *japonica* are Hap-G, *indica* are Hap-T, and wild rice are distributed in both haplotypes. And the *Aus* rice tends towards the *indica* type. All these findings suggested the genetic differentiation of *GL12* between *japonica* and *indica* subspecies, which exhibited significantly reduced nucleotide diversity in *indica* compared with wild rice, suggesting a strong artificial selection in the *GL12* locus of *indica*.

We proposed a possible working model for the network and showed that *GL12* played an essential role in grain length and salt tolerance, which was synergistically regulated by GIF1 and WRKY53 at the G/T variation site (Fig. 7e).

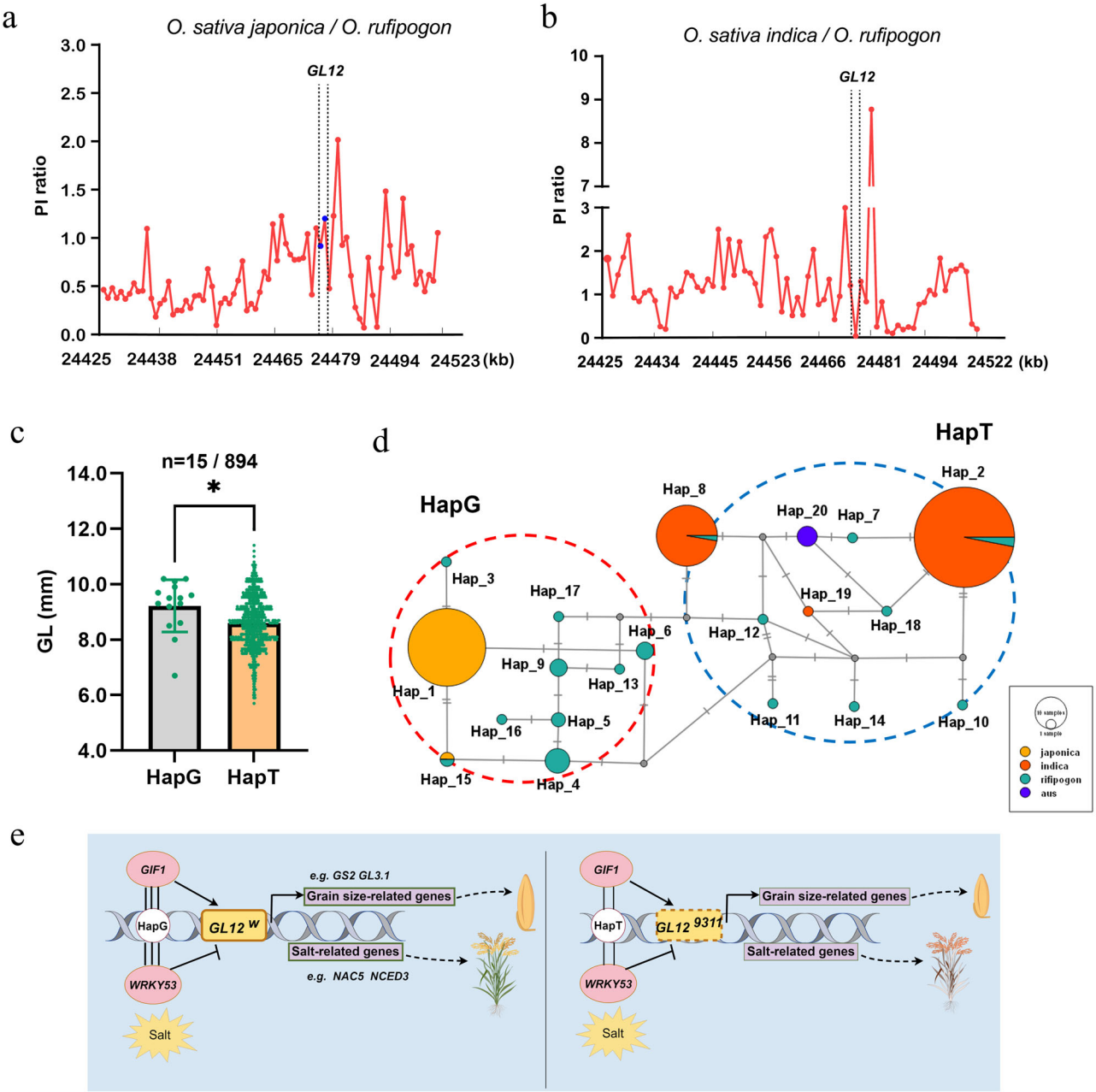

Fig. 7 | Natural variations in *GL12* contribute to the divergence of *indica* and *japonica* subspecies. a, **b** The relative ratio of nucleotide diversity analysis in the complete chromosome 12 of cultivated and common wild rice. The *GL12* locus and its flanking genomic regions were marked. Dotted line area indicates a selective sweep of *GL12* locus in *indica*. **c** Distribution of grain length for the two haplotypes in *indica* rice accessions, $n = 15/894$ means the accessions number of HapG and HapT. The haplotypes and number of accessions from the 3K Rice Genomes Project dataset were shown. Data show means ± SD. Statistical analysis was performed by two-tailed Student's *t*-test (*$p \le 0.05$, *p*-value was 0.0104). **d** Haplotype networks generated from the promoter sequence of *GL12* in both cultivated rice and various groups of wild rice accessions. The circle size is proportional to the sample size. Grey spots on the lines indicate the mutational steps between two haplotypes. **e** A proposed working model for the role of *GL12* in the regulation of rice grain length and salt tolerance. Wild rice $GL12^{W}$ is a functional and $GL12^{9311}$ is a nonfunctional allele. The G/T variation in *GL12* promoter determined its binding with transcription coactivator GIF1 and transcription factor WRKY53. GIF1 promotes $GL12^{W}$ expression levels in young panicles, and then regulates the downstream grain size related genes, such as *GS2*. WRKY53 represses $GL12^{W}$ expression under salt stresses, and $GL12^{W}$ also regulates rice salt response genes, such as *NCED3* and *NAC5*. Therefore, the wild rice allele $GL12^{W}$ synergistically improves both the rice grain length and salt tolerance through the regulation by GIF1 and WRKY53. The image was drawn by Figdraw. Source data are provided as a Source Data file.

Discussion

Early attempts to identify positive yield-related genes proved unsuccessful because most of the agronomic traits, especially yield-related characters of wild rice were inferior to those of cultivated rice. In this study, we identified wild rice gene $GL12^{W}$ that improved rice GL and salt tolerance, which is synergistically regulated by GIF1 and WRKY53. The wild rice CSSL population we development in our previous study has greatly facilitated the discovery of unknown useful genes in wild rice[37]. We identified the $GL12^{W}$ gene using CSSL population and found it improves both the GL and salt tolerance synergistically in cultivated rice.

Compared with $GL12^{9311}$, $GL12^{W}$ improved the GL, TGW and salt tolerance in both *indica* and *japonica* genetic backgrounds (Figs. 1, 5; Supplementary Figs. 6, 9). We demonstrated that $GL12^{W}$ regulated

global gene expression, including *GS2*, and improves the GL by altering cell length (Fig. 2). We found a G/T variation at 1,117-bp upstream of the start codon of *GL12* playing an essential role in binding with *OsGIF1* and *OsWRKY53*, which positively and negatively regulated the *GL12* expression, respectively (Figs. 3, 6). Previous reports demonstrated that *OsGIF1* interacted with *GS2* to control grain-size predominantly by increasing cell expansion in the grain hull[11,20,30]. But whether these two genes involving in rice salt tolerance remains unknown, and no reports on whether GIF1 interacted with DNA and regulates gene expression directly. Our results showed the GIF1 physically interacted with GL12 but no differences were found between $GL12^{W}$ and $GL12^{9311}$ (Fig. 4a, b). GIF1 interacted with $GL12^{W}$ promoter and the G/T site determined its binding to GIF1 and increased the $GL12^{W}$ expression levels in young panicle (Fig. 4c–e). We also found that *GIF1* and *GS2* positively regulated salt tolerance in rice (Supplementary Fig. 13). Our results provided insights into the *GIF1-GL12*^{W}*-GS2* module how to regulate the grain length and salt tolerance. Nevertheless, further studies are needed to explore the detailed mechanism and regulatory network of this module.

The NIL of $GL12^{W}$ showed longer GL and stronger salt tolerance than $GL12^{W}$-OE lines. In addition, our transgenic experiments demonstrated that $GL12^{W}$ had a stronger effect on the GL under its native promoter (Fig. 3d–f). Interestingly, compared with $GL12^{9311}$, $GL12^{W}$ has higher expression levels in young panicle but lower expression levels under salt stresses, and the latter might be caused by the repression of WRKY53 (Fig. 6). The expression levels of $GL12^{W}$ synergistically regulated by GIF1 and WRKY53, reach an optimal balance in NIL, that explains NIL had much longer grain length and higher salt tolerance than $GL12^{W}$-OE lines. We further demonstrated that the G/T variation contributed to the divergence of *indica* and *japonica* subspecies. *GL12* might have been directionally selected in *indica* but not in *japonica* during rice domestication (Fig. 7d). We investigated the agronomic and yield-related traits of 9311, NIL and $GL12^{W}$-OE lines, $GL12^{W}$ has improved plant height, panicle length, but significantly decreased the grain number per panicle, same results were found in $GL12^{W}$-OE lines in NIP background (Supplementary Fig. 15). Unfortunately, the $GL12^{W}$ did not improve yield significantly under both normal and salt stress conditions. This indicated that $GL12^{W}$ might have a negatively effect on the grain number per spike, which can explain why the wild rice $GL12^{W}$ has not been selected from wild rice to *indica* varieties during the rice artificial selection breeding.

Improving the salt tolerance without reducing yield is one of the major goals for modern rice breeding. The wild rice $GL12^{W}$ gene, regulated by GIF1 and WRKY53, provides the endogenous cues for wild rice gene utilization and further paves the way for future investigation of the detailed biochemical mechanisms that synergistically improve both the grain length and salt tolerance. In summary, the discovery of this gene not only helps understand the mechanism for improving grain length and environmental resilience simultaneously for rice improvement but also provides the useful genetic resources for rice breeding programs.

Methods

Plant materials and growth conditions

Rice (*O. sativa* L.) materials include two wild-type *japonica* rice varieties, ZH11 and NIP, and one *indica* rice 9311. The overexpression lines were derived from ZH11 and NIP, and the knockout mutants were from NIL. All rice plants were cultivated in a natural paddy condition at Beijing and Sanya in Hainan province, China. Standard cultivation practices were used, and pests and diseases were managed in rice cultivation. Mature and dry seeds were used for measuring grain length, width, and TGW by image analysis method using SC-E software (Hangzhou Wanshen Detection Technology, Hangzhou, China). The plant height, panicle length, tiller number, and grain number were obtained from the measurement of the main culm. All agronomic traits and grain size measurements were based on at least 20 individuals, which were randomly selected.

Vector construction and rice transformation

To construct the overexpression vector, the full-length coding sequence of *GL12* was amplified from the wild rice and 9311 and cloned into a pBWA(V)HS vector. Genes were under the control of double 35S promoter in overexpression lines. The Pro^{9311}*::*$GL12^{9311}$, Pro^{9311}*::*$GL12^{W}$, Pro^{W}*::*$GL12^{9311}$ and Pro^{W}*::*$GL12^{W}$ transgenic lines also used the pBWA(V)HS vector, but 35S promoter fragment of the pBWA(V)HS vector were replaced by the *GL12* promoter from the wild rice or 9311 respectively, the construct vectors were introduced into the calli of NIP and by *Agrobacterium* mediated transformation. To generate knockout mutants, two *GL12* site-specific target sequences were selected in the fifth exon and introduced into the CRISPR/Cas9 system, and the plasmid were transformed into $GL12^{W}$-NIL to generate the *GL12* knockout mutants. Knockout lines of OsGIF1 in the $GL12^{W}$-NIL background were also generated by CRISPR/Cas9 technology.

RNA extraction and RT-PCR

Total RNA was extracted from young leaves and panicles using the RNA Easy Fast Plant Tissue Kit (TIANGEN, Beijing, China) and reverse-transcribed into the first-strand cDNA with a PrimeScript RT Reagent Kit (Invitrogen, Carlsbad, CA). Real-time qPCR experiments were performed using a SuperReal PreMix Plus (SYBR Green) (TIANGEN) on an Applied Biosystems 7500 Real-Time PCR system following the manufacturer's instructions. All the qRT-PCR analyses were performed on at least three independent biological replicates. Rice *ACTIN* was used as an internal control for all the qRT-PCR analyses. Primers for qRT-PCR are listed in Supplementary Data 1.

Subcellular localization

To generate the 35S::*GL12*-GFP fusion, the full length cDNA of *GL12* without the stop codon was cloned into vector pCAMBIA 1300. The empty 35S: GFP, 35S::*GL12*-GFP and NLS-RFP plasmids were transformed into *Agrobacterium tumefaciens* strain GV3101, and infiltrated into 3-week-old *N. benthamiana* leaves. The *Agrobacteria* harboring the corresponding construct and NLS-RFP were co-infiltrated into *N. benthamiana* leaves. At 2–4 day after infiltration, fluorescent signals in leaves were detected by laser confocal microscopy (Carl Zeiss, LSM780, Germany).

Histological analysis

The fresh young spikelet hulls were fixed using FAA (40% formaldehyde: 50% alcohol: 5% acetic acid) and embedded in Paraplast Plus (Sigma, Germany) then cut into 10-μm-thick sections. Cross sections were observed by light microscopy (Olympus, Japan), and the cell number and cell area in the outer parenchyma cell layer of hulls were measured using ImageJ software[38]. The outer surfaces of the spikelet hulls were observed using a scanning electron microscope (Gemini-SEM 300, Carl Zeiss, Germany). Cell size and cell number were measured with ImageJ software.

Dual-Luciferase transcriptional activity assay

To construct the Pro^{9311}*::LUC* and Pro^{W}*::LUC* plasmids, we amplified a 2 kb promoter sequence from 9311 and wild rice and cloned them into the pGreen0800II-LUC vector, respectively; the resulting vectors were used as reporters. The full-length cDNA of *WRKY53* and *GIF1* were cloned into the pGreen0800II-62-SK vector to construct *35S: WRKY53* and *35S: GIF1*, which were used as the effector. Transient transactivation assays were performed using the tobacco (*N. benthamiana*) leaf system. The LUC and REN activities were measured by the Dual-Luciferase Reporter Assay System (Promega, USA), and the LUC activity was normalized to REN activity. Primers used for these constructs are listed in Supplementary Data 1.

Yeast one-hybrid assay

The full-length cDNA of *WRKY53* and *GIF1* were cloned into the pB42AD vector to generate prey constructs. Approximately 2 kb of the *GL12* promoter from wild rice and 9311 were cloned into the lacZ (β-galactosidase) reporter plasmid pLACZi as bait constructs. Yeast strain EGY48 was used for transformation. All the procedures were performed according to the manufacturer's protocol (Clontech, USA).

Luciferase complementation imaging (LCI) Assay

The LCI assays for the interaction between *GL12* and *GIF1* were performed in *N. benthamiana* leaves. The full-length *GL12* and *GIF1* coding regions were fused with the N-terminal and C-terminal parts of the luciferase reporter gene, respectively. The nLUC-*GL12* and cLUC-*GIF1* constructs were transformed into *Agrobacterium* GV3101 cells and co-infiltrated into *N. benthamiana* leaves. One same subcellular localization gene, *NAC2*, was used as the negative control[39,40]. The LUC activity was analyzed using Tanon 5200 plant imaging system (Tanon Science & Technology, China).

Electrophoretic mobility shift assay

Electrophoretic mobility shift assay (EMSA) was used to examine the binding of the WRKY53 and GIF1 proteins to the motif in the promoter regions of Pro^{9311} and Pro^{W}, respectively. We synthesized 5′- biotin labeled DNA fragments containing binding sites as probes and used unlabeled DNA fragments of the same sequence for competitive reactions. The MBP-OsWRKY53 and GST-GIF1 fusion protein were expressed in *E. coli* pMal-c2x at 16 °C for 20–24 h in the presence of 0.1 mM isopropyl β-D-1-thiogalactopyranoside. The LightShift™ Chemiluminescent EMSA Kit was used to perform EMSA according to the manufacturer's instructions (Thermo Fisher Scientific, USA). The binding reactions were incubated for 30 min at room temperature and then resolved by electrophoresis on 6% (w/v) native polyacrylamide gels in 0.5×TBE buffer. Probe sequences are shown in Supplementary Data 1.

Yeast two-hybrid assay

Matchmaker Gold Yeast Two-Hybrid System (Clontech, USA) was used for Y2H assays. The CDS of *GL12* was cloned in prey vector pGBKT7, and the *GIF1* were cloned into the bait vector pGADT7. The prey and bait vectors were transformed into the Y2H Gold yeast. All the procedures were performed according to the manufacturer's protocol (Clontech, USA). Relevant primers are listed in Supplementary Data 1.

Haplotype and evolutionary analyses

For the Hap-G and Hap-T distribution analysis, 1772 (596 *japonica*, 1176 *indica*) and 909 (79 *japonica*, 191 *indica*, 16 *rifipogon*) accessions from the 3K Rice Genomes Project and rice pan-genome database[36], respectively, were used (Supplementary Data 2 and 3). The nucleotide diversity (π) of 52 wild rice, 56 *indica*, and 51 *japonica* accessions were calculated in 50-kb sliding windows using VCFtools[41] and the ratio of genetic diversity in all accessions was calculated. The genomic sequences and grain length of 909 cultivated *indica* rice were obtained from the 3K Rice Genomes Project database. The grain length of two types of promoters of *GL12* with T or G haplotype was compared using *t*-test. The geographical distribution of 909 cultivated rice were obtained from the 3K Rice Genomes Project database and marked on map using R-version (Team, 2014). The genotype of 219 accessions including wild rice were from rice pan-genome database[36]. Fifteen SNPs and five InDels of *GL12* in the promoter were aligned with ClustalW. Haplotype frequency data were processed with DnaSP5[42] and visualized Median-joining networks were generated by Popart[43].

Reporting summary

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

Data availability

The data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. Source data are provided with this paper.

References

1. Khush, G. What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol. Biol.* **59**, 1–6 (2005).
2. Qian, Q. et al. Breeding high-yield superior quality hybrid super rice by rational design. *Nat. Sci. Rev.* **3**, 283–294 (2016).
3. Bailey, S. et al. Genetic strategies for improving crop yields. *Nature* **575**, 109–118 (2019).
4. Li, N., Xu, R. & Li, Y. Molecular networks of seed size control in plants. *Annu. Rev. Plant Biol.* **70**, 435–463 (2019).
5. Li, N. & Li, Y. Signaling pathways of seed size control in plants. *Curr. Opin. Plant Biol.* **33**, 23–32 (2016).
6. Sun, S. et al. A G-protein pathway determines grain size in rice. *Nat. Commun.* **9**, 851 (2018).
7. Xu, R. et al. Control of grain size and weight by the OsMKKK10-OsMKK4-OsMAPK6 signaling pathway in rice. *Mol. Plant* **11**, 860–873 (2018).
8. Zuo, J. & Li, J. Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu. Rev. Genet.* **48**, 99–118 (2014).
9. Chen, R. et al. A de novo evolved gene contributes to rice grain shape difference between indica and japonica. *Nat. Commun.* **14**, 5906 (2023).
10. Fan, C. et al. *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* **112**, 1164–1171 (2006).
11. Hu, J. et al. A rare allele of *GS2* enhances grain size and grain yield in rice. *Mol. Plant* **8**, 1455–1465 (2015).
12. Qi, P. et al. The novel quantitative trait locus *GL3.1* controls rice grain size and yield by regulating Cyclin-T1;3. *Cell Res.* **22**, 1666–1680 (2012).
13. Song, X. et al. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **39**, 623–630 (2007).
14. Wang, S. et al. The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* **47**, 949–954 (2015).
15. Wang, S. et al. Control of grain size, shape and quality by *OsSPL16* in rice. *Nat. Genet.* **44**, 950–954 (2012).
16. Mao, H. et al. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl Acad. Sci. USA* **107**, 19579–19584 (2010).
17. Zhang, X. et al. Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc. Natl Acad. Sci. USA* **109**, 21534–21539 (2012).
18. Liu, Q. et al. G-protein βγ subunits determine grain size through interaction with MADS-domain transcription factors in rice. *Nat. Commun.* **9**, 852 (2018).
19. Zhao, D. et al. *GS9* acts as a transcriptional activator to regulate rice grain shape and appearance quality. *Nat. Commun.* **9**, 1240 (2018).
20. Che, R. et al. Control of grain size and rice yield by *GL2*-mediated brassinosteroid responses. *Nat. Plants* **2**, 15195 (2015).
21. Sun, P. et al. *OsGRF4* controls grain shape, panicle length and seed shattering in rice. *J. Integr.* **58**, 836–847 (2016).
22. Editorial. Flourish with the wild. *Nat. Plants* **9**, 373–374 (2023).
23. Huang, R. et al. Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci.* **18**, 218–226 (2013).
24. Wang, Y. et al. A novel QTL *GL12* from wild rice increases grain length and weight in cultivated rice. *Plant Biotechnol. J.* **21**, 1513–1515 (2023).

25. Qi, L. et al. Fine mapping and identification of a novel locus *qGL12.2* control grain length in wild rice (*Oryza rufipogon* Griff.). *Theor. Appl. Genet.* **131**, 1497–1508 (2018).
26. Du, H. et al. Biochemical and molecular characterization of plant MYB transcription factor family. *Biochem. Biokhimiia* **74**, 1–11 (2009).
27. Muthuramalingam, P. et al. Global integrated genomic and transcriptomic analyses of MYB transcription factor superfamily in C3 model plant *Oryza sativa* (L.) unravel potential candidates involved in abiotic stress signaling. *Front. Genet.* **13**, 946834 (2022).
28. De Veylder, L., Beeckman, T. & Inzé, D. The ins and outs of the plant cell cycle. *Nat. Rev. Mol. Cell Biol.* **8**, 655–665 (2007).
29. Dewitte, W. & Murray, J. The plant cell cycle. *Annu. Rev. Plant Biol.* **54**, 235–264 (2003).
30. Duan, P. et al. Regulation of *OsGRF4* by OsmiR396 controls grain size and yield in rice. *Nat. Plants* **2**, 15203 (2015).
31. He, Z. et al. *OsGIF1* positively regulates the sizes of stems, leaves, and grains in rice. *Front. Plant Sci.* **8**, 1730 (2017).
32. Li, S. et al. The *OsmiR396c-OsGRF4-OsGIF1* regulatory module determines grain size and yield in rice. *Plant Biotechnol. J.* **14**, 2134–2146 (2016).
33. Yu, J. et al. Genome-wide association studies identify OsWRKY53 as a key regulator of salt tolerance in rice. *Nat. Commun.* **14**, 3550 (2023).
34. Wang, C. et al. Towards a deeper haplotype mining of complex traits in rice with RFGB v2.0. *Plant Biotechnol. J.* **18**, 14–16 (2020).
35. Wang, W. et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**, 43–49 (2018).
36. Shang, L. et al. A super pan-genomic landscape of rice. *Cell Res.* **32**, 878–896 (2022).
37. Qiao, W. et al. Development and characterization of chromosome segment substitution lines derived from *Oryza rufipogon* in the genetic background of *O. sativa spp. indica* cultivar 9311. *BMC Genomics* **17**, 580 (2016).
38. Rueden, C. T. et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* **18**, 529 (2017).
39. Zhou, Z. Y. et al. Luciferase complementation assay for protein-protein interactions in plants. *Curr. Protoc. Plant Biol.* **3**, 42–50 (2018).
40. Mao, C. et al. Overexpression of a NAC-domain protein promotes shoot branching in rice. *New Phytol.* **176**, 288–298 (2007).
41. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
42. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452 (2009).
43. Leigh, J. & Bryant, D. PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 110–1116 (2015).

Acknowledgements

This work was supported by the National Key R&D Program of China (2021YFD1200100, and 2021YFD1200501) to W.H.Q., Hainan key R&D program of China (ZDYF2024XDNY165) to Q.W.Y., Hainan province science and technology talent project (KJRC2023B21) to Q.W.Y., and the Agricultural Science and Technology Innovation Program of Chinese Academy of Agricultural Sciences. We thank Prof. Qingyun Bu (Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences) for the *WRKY53* genetic materials, and Dr. Peng Zhang (Plant Breeding Institute, The University of Sydney) for revising this manuscript.

Author contributions

W.H.Q., Q.W.Y. and Q.Q. conceived the project. Y.Y.W. performed most of the experiments and analyzed the data. W.X.C. participated in the EMSA experiments, mutants investigation, and genetic analysis. J.Q.S. supervised the EMSA, LCI and transgenic experiments. Y.H.L. and K.H. performed the GIF1 transgenic lines and mutants. J.H. provided GS2 related genetic materials. M.X., S.Z.W., Z.Y.Y., J.F.H., W.L.G., Y.M.N., Y.T.W., Z.Y.C., Y.P.L. and M.C.Z. participated in field management and trait investigation, Q.L.Z., L.F.Z., Y.L.C., X.M.Z. and L.N.Z. participated in field management and logistic works. Q.W.Y. and Q.Q. supervised the project. W.H.Q. and Y.Y.W. wrote and revised the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53611-9.

Correspondence and requests for materials should be addressed to Qian Qian, Qingwen Yang or Weihua Qiao.

Peer review information *Nature Communications* thanks Bin Han, Huazhong Shi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.