

# The novel quantitative trait locus *GL3.1* controls rice grain size and yield by regulating Cyclin-T1;3

Peng Qi<sup>1,\*</sup>, You-Shun Lin<sup>2,\*</sup>, Xian-Jun Song<sup>1,\*</sup>, Jin-Bo Shen<sup>2</sup>, Wei Huang<sup>1</sup>, Jun-Xiang Shan<sup>1</sup>, Mei-Zhen Zhu<sup>1</sup>, Liwen Jiang<sup>2</sup>, Ji-Ping Gao<sup>1</sup>, Hong-Xuan Lin<sup>1</sup>

<sup>1</sup>National Key Laboratory of Plant Molecular Genetics and National Center for Plant Gene Research (Shanghai), Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China; <sup>2</sup>School of Life Sciences, Centre for Cell and Developmental Biology, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

Increased crop yields are required to support rapid population growth worldwide. Grain weight is a key component of rice yield, but the underlying molecular mechanisms that control it remain elusive. Here, we report the cloning and characterization of a new quantitative trait locus (QTL) for the control of rice grain length, weight and yield. This locus, *GL3.1*, encodes a protein phosphatase kelch (PPKL) family — Ser/Thr phosphatase. *GL3.1* is a member of the large grain WY3 variety, which is associated with weaker dephosphorylation activity than the small grain FAZ1 variety. *GL3.1*-WY3 influences protein phosphorylation in the spikelet to accelerate cell division, thereby resulting in longer grains and higher yields. Further studies have shown that *GL3.1* directly dephosphorylates its substrate, Cyclin-T1;3, which has only been rarely studied in plants. The downregulation of Cyclin-T1;3 in rice resulted in a shorter grain, which indicates a novel function for Cyclin-T in cell cycle regulation. Our findings suggest a new mechanism for the regulation of grain size and yield that is driven through a novel phosphatase-mediated process that affects the phosphorylation of Cyclin-T1;3 during cell cycle progression, and thus provide new insight into the mechanisms underlying crop seed development. We bred a new variety containing the natural *GL3.1* allele that demonstrated increased grain yield, which indicates that *GL3.1* is a powerful tool for breeding high-yield crops.

**Keywords:** QTL; cell cycle; yield; phosphatase; Cyclin T

*Cell Research* (2012) 22:1666-1680. doi:10.1038/cr.2012.151; published online 13 November 2012

## Introduction

The global population is expected to reach 9.1 billion by 2050 [1], which would require each hectare of land to feed 43 people [2]. However, various factors, such as water scarcity, soil salinity, disease, climate change and reduced arable land area will intensify food shortages in the next 50 years [3-5]; therefore, improvements in crop yields are urgently required. Rice (*Oryza sativa* L.) is one of the most important crops worldwide, and it is a major staple in Asia, which contains 66.67% of the

world's population. Improved rice yields could prevent global food shortages. Since the middle of the 20th century, yield increases in the dominant short-stature rice cultivars, which were developed in Asia during the Green Revolution, have stagnated [6], and new methods are urgently required to help breeders to continue to increase rice yields. The characterization of the functions of genes that control agricultural traits in rice may provide new opportunities for breeding. Rice yield depends on three major components: grain weight, grain number per panicle and panicle number per plant (tiller number). Several genes associated with these traits (or quantitative trait loci, QTLs) have been isolated recently [7-21]; however, the pathways regulated by these genes have not been characterized. As QTLs control most rice yield traits, the identification of these genes may facilitate the breeding of high-yield rice varieties.

Post-translational modifications, such as phosphoryla-

\*These three authors contributed equally to this work.

Correspondence: Hong-Xuan Lin<sup>a</sup>, Ji-Ping Gao<sup>b</sup>

<sup>a</sup>E-mail: hxlin@sibs.ac.cn

<sup>b</sup>E-mail: jpgao@sibs.ac.cn

Received 30 August 2012; revised 12 September 2012; accepted 12 September 2012; published online 13 November 2012

tion, ubiquitination, glycosylation and acetylation, can alter protein function. One-third of all proteins are either phosphorylated or dephosphorylated, and phosphorylation regulates many cellular events/activities, such as enzymatic activity, substrate specificity, protein subcellular localization, and protein turnover/stability [22]. Of the nine amino acids that can be phosphorylated, serine, threonine and tyrosine are the major amino acids that affect protein functions [23]. Phosphatases can be divided into groups based on their substrate specificities: serine/threonine-specific phosphoprotein phosphatases (PPP), metal-independent protein phosphatases (PPM; also known as PP2Cs), tyrosine-specific phosphotyrosine phosphatases (PTP) and dual (serine/threonine and tyrosine)-specific phosphatases (DSP) [24]. The PPP family can be further divided into several subfamilies based on substrate specificity and pharmacological properties: protein phosphatase 1 (PP1), PP2 (including PP2A and PP2B), PP4, PP5, PP6, PP7 and PPP kelch (PPKL) [25-27]. The PPKL protein family contains an N-terminal kelch-repeat domain and a C-terminal phosphatase domain, and these enzymes are more closely related to PP1 than to PP2A on the basis of conserved residues [28, 29]. PPKLs only exist in plants, green algae and several *Alveolata* species [30], and they include *Plasmodium falciparum* PfPPA [31], BSU1 [32] and *Arabidopsis* BSU1-like protein 1 (BSL1), BSL2 and BSL3 [30, 33]. BSU1 is important in the brassinolide signaling pathway [34]. BSU1 receives a signal from BRI1 and inhibits the activity of BIN2; subsequently, BZR1, which is suppressed by BIN2, induces the expression of BR-response genes. However, the functions of other PPKLs remain poorly characterized.

In the present study, we cloned and characterized a new QTL, *GL3.1*, which regulates grain length and yield in rice. *GL3.1* is a novel serine/threonine phosphatase that belongs to the PPKL family. *GL3.1* regulates grain length by mediating cell cycle progression through affecting the phosphorylation status of cell cycle proteins, such as Cyclin-T1;3, thereby controlling grain yield. These findings provide insight into seed development and establish a new tool for improved crop breeding.

## Results

### *Fine-mapping of a new QTL, GL3.1, which regulates rice grain yield*

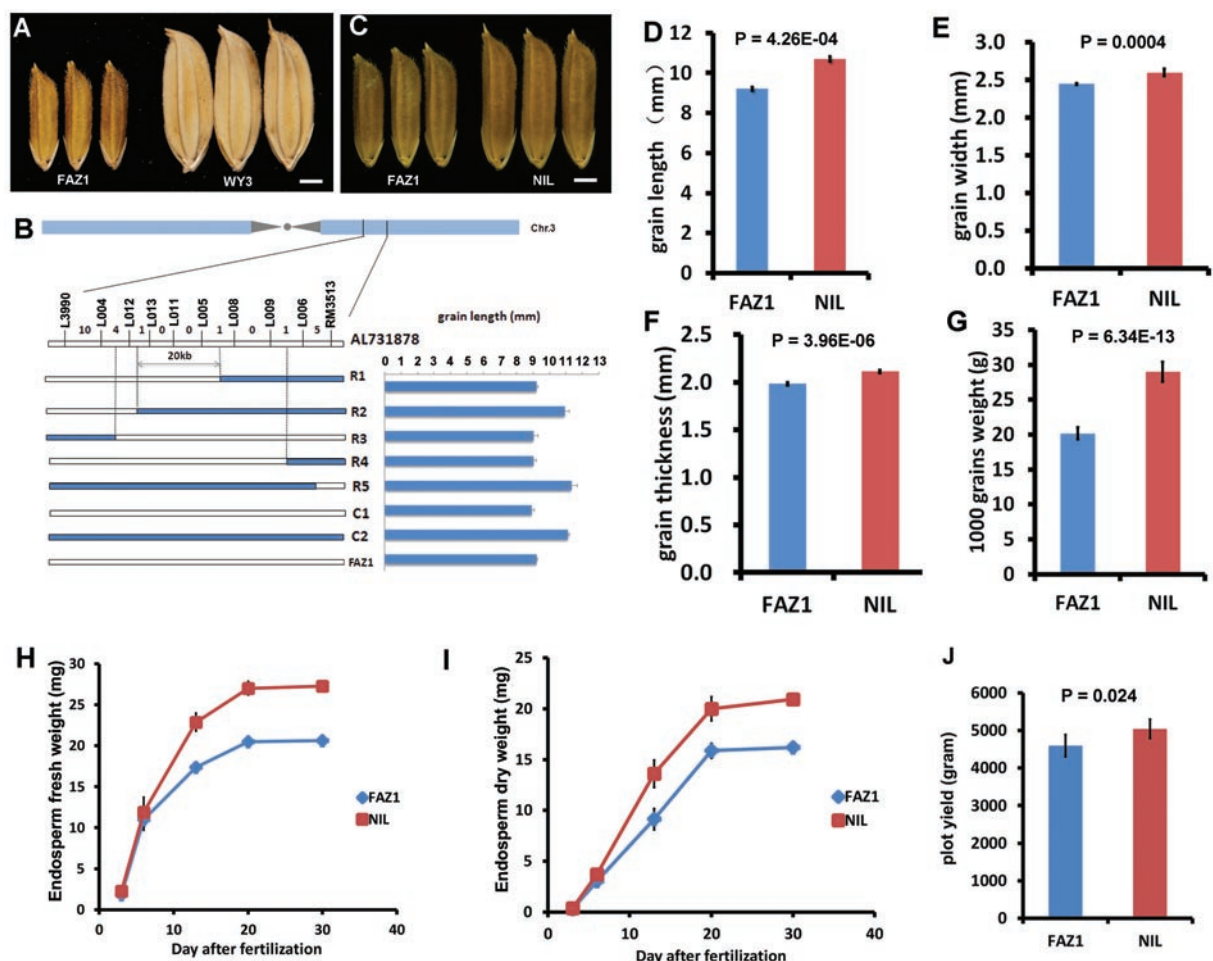
Fengaizhan-1 (FAZ1) and Waiyin-3 (WY3) rice varieties were selected as parents to map the QTLs that affect grain length. FAZ1 is a small grain *indica* variety (1 000-grain weight:  $20.18 \pm 0.89$  g), whereas WY3 is a larger grain *japonica* variety (1 000-grain weight:  $43.40 \pm 0.92$  g; Figure 1A). We fine-mapped a new major QTL

(*GL3.1*) for grain length to a 20-kb region between the L012 and L008 markers on chromosome 3 (25 036 192 bp to 25 060 567 bp at chromosome 3) (Figure 1B), which is distinct from other previously reported QTLs [35-38]. This region contains two genes: *Os03g44510*, which is a predicted transposon that was excluded from further analysis because the transcript was not detected in both parents, and the predicted phosphatase *Os03g44500*, which was expressed in both parents and considered as the *GL3.1* candidate.

Based on the mapping results, we developed a near-isogenic line (NIL) from BC4F2 generations that contained a 30-kb WY3 chromosomal region at the *GL3.1* locus in a FAZ1 genetic background (Figure 1C; Supplementary information, Figure S1A-S1C). NIL had longer grains (+16.1%) than FAZ1 ( $10.71 \pm 0.13$  mm vs  $9.22 \pm 0.09$  mm), but there were no significant differences in grain width or thickness (Figure 1D-1F), plant height or tiller number (Supplementary information, Figure S1D-S1E). NIL had a significantly greater 1 000-grain weight than FAZ1 (+43.5%; Figure 1G) and reduced grain number per main panicle (21.3%, Supplementary information, Figure S1F). NIL exhibited an increase in the milk filling rate (Figure 1H-1I) and higher expression of milk filling-related genes (Supplementary information, Figure S1G). The plot grain yield was significantly increased in NIL (+ 11.1%; Figure 1J); however, the grain quality was not affected, as the packing density of starch granules was similar in the mature seeds of NIL and FAZ1 (Supplementary information, Figure S1H-S1I), and the chalky grain percentage and protein and amylose contents were similar between the NIL and FAZ1 grains (Supplementary information, Figure S1J-S1L). We crossed NIL with Huanghuazhan, which is a relatively high-yield elite *indica* variety that is widely cultivated in Southern China, and subsequently backcrossed the F1 generation with Huanghuazhan to obtain a Huanghuazhan (*GL3.1*) variety that exhibited a longer and heavier grain (Supplementary information, Figure S2A-S2E) and a higher grain yield than Huanghuazhan under field conditions (Supplementary information, Figure S2F). These findings confirmed that *GL3.1* potentially increases grain yield.

### *Confirmation of GL3.1 function*

*GL3.1* contains a 3 012-bp open reading frame (ORF) that encodes 21 exons and 20 introns. The FAZ1 *GL3.1* allele contains 4-bp differences when compared with the WY3 *GL3.1* allele (FAZ1 to WY3: 1092C-A, 1495C-T, 2643A-G, 2838T-C), which results in two amino acid substitutions (364 aspartic acid — glutamic acid (364D-E), 499 histidine — tyrosine (499H-Y); Figure 2A). We



**Figure 1** Map-based cloning of *GL3.1*. **(A)** FAZ1 and WY3 rice grains. Scale bar, 2 mm. **(B)** Genetic location of *GL3.1* and progeny testing of fixed recombinant plants. The numbers indicate the number of recombinant plants between *GL3.1* and molecular markers. R1-R5, five recombinant lines; C1, control 1 (FAZ1 homozygous); C2, control 2 (WY3 homozygous). Grain length (mean  $\pm$  SD) of R2, R5 and C2 is greater than R1, R3, R4 and C1 ( $n = 10$  plants). The white and blue areas represent FAZ1 and WY3 homozygous chromosomal segments, respectively. **(C)** FAZ1 and NIL grains. Scale bar, 2 mm. **(D-G)** Comparison of FAZ1 and NIL grain length ( $n = 20$ ) **(D)**, grain width ( $n = 20$ ) **(E)**, grain thickness ( $n = 20$ ) **(F)** and 1 000-grain weight ( $n = 20$ ) **(G)**. **(H, I)** Time course of FAZ1 and NIL endosperm fresh weight **(H)** and endosperm dry weight **(I)** ( $n = 40$  plants). **(J)** FAZ1 and NIL plot yield (plot area: 9.6 m<sup>2</sup>) ( $n = 2$ ). A Student's *t*-test was used to generate the *P* values in **D-G** and **J**.

sequenced *GL3.1* in several large grain varieties and detected the japonica variety Nanyangzhan with a truncated *GL3.1* allele as well as Jizi1560 and Jizi1581, which contained 15 additional amino acids at the C-terminus compared with FAZ1 and WY3. At the positions 364 and 499, Jizi1560 and Jizi1581 exhibited the same amino acid substitutions as WY3 (Supplementary information, Figure S3). Transgenic rice plants were generated to determine whether *GL3.1* controls grain length. FAZ1 and WY3 failed to regenerate shoots from the callus, and therefore we used the small-grain japonica variety Zhonghua 11, which was easily regenerated [39]. We generated constructs containing the full-length *GL3.1*

ORFs from FAZ1 or WY3 under the CaMV 35S promoter. Some of the obtained transgenic lines that overexpressed the WY3 *GL3.1* allele showed an increased grain length (*GL3.1*-WY3), whereas the grain length was not changed in all lines overexpressing the FAZ1 *GL3.1* allele (*GL3.1*-FAZ1; data not shown). Only the *GL3.1*-WY3 line, which expressed relatively high levels of *GL3.1*-WY3, exhibited increases in grain length (Figure 2B-2D, Supplementary information, Figure S4A-S4B), which confirms that *GL3.1* controls grain length. *GL3.1*-FAZ1 RNA interference (RNAi) and antisense transgenic plants were generated; however, no phenotypic changes in grain length were observed (data not shown). We also

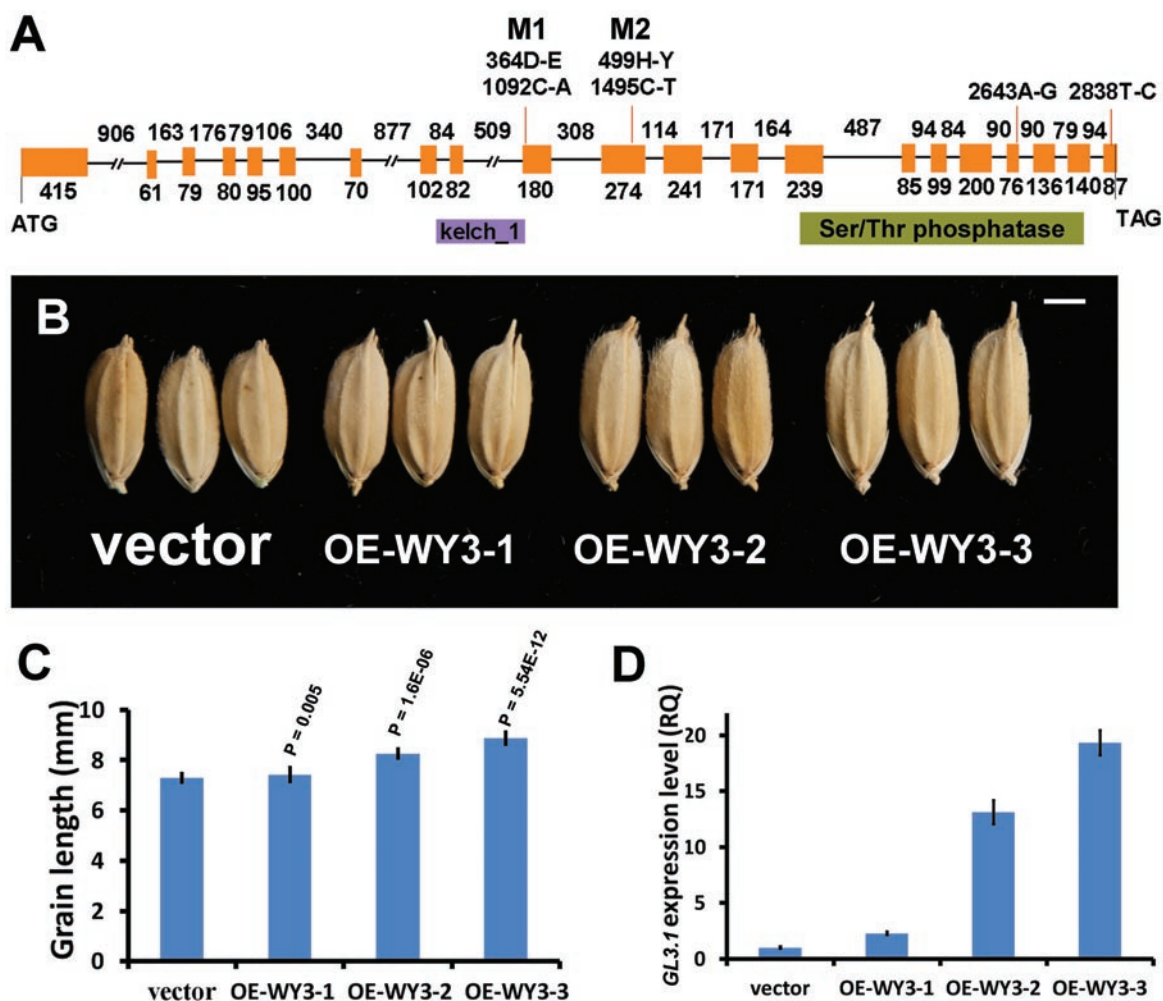
observed that *GL3.1* was downregulated but not completely suppressed in these lines; therefore, we hypothesized that these lines retained adequate *GL3.1* function, as *GL3.1* was abundantly expressed.

*GL3.1* is predicted to encode a Ser/Thr phosphatase of unknown function and with two predicted domains: a Kelch\_1 protein interaction domain and a Ser/Thr phosphatase domain (Figure 2A). Transgenic plants were generated to investigate the effect of the *GL3.1* point substitutions *GL3.1*-M1 (364E, 499H) and *GL3.1*-M2 (364D, 499Y) in FAZ1 and WY3. Both transgenic lines exhibited significant increases in grain length (Supplementary information, Figure S4C-S4H). Similar to the *GL3.1*-WY3 transgenic lines, only high levels of *GL3.1*-

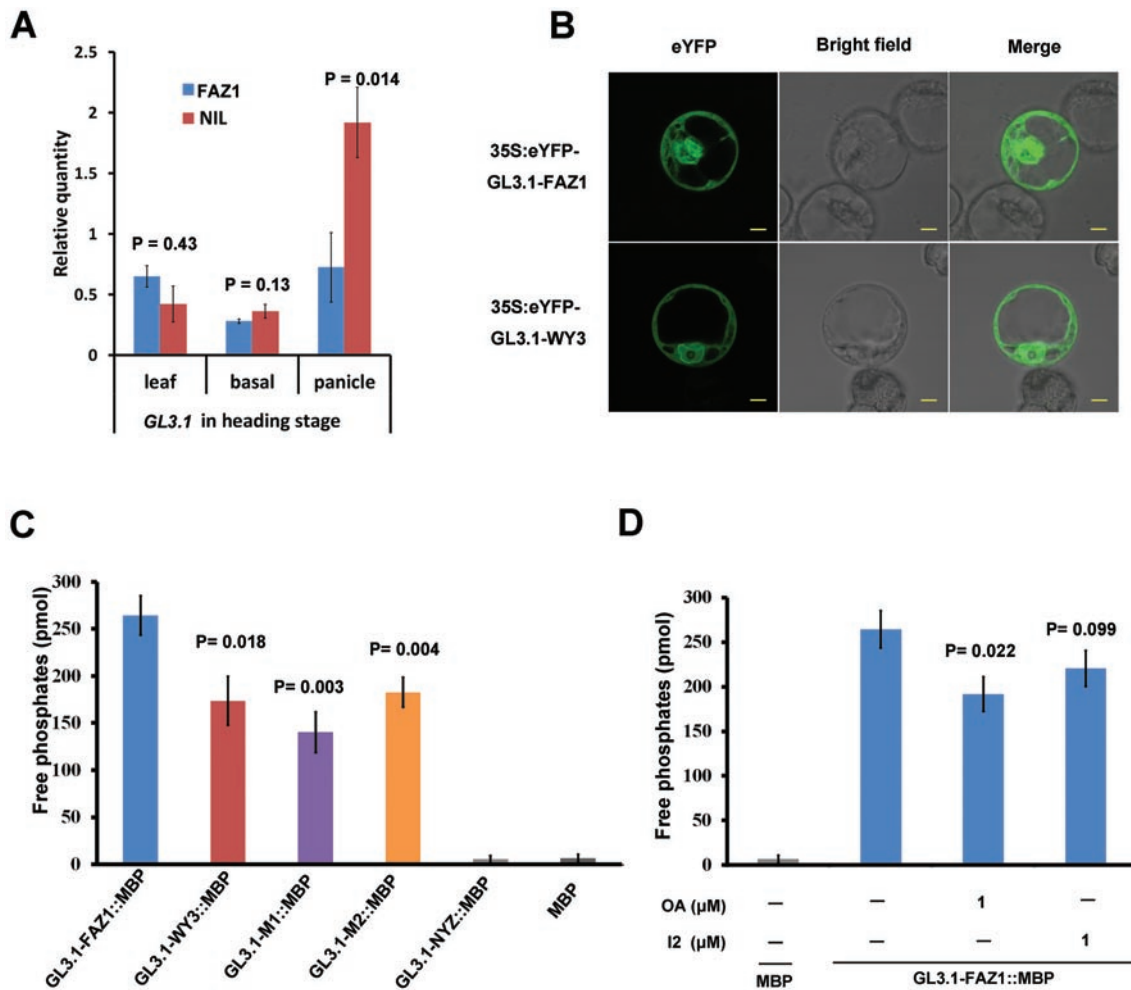
M1 or *GL3.1*-M2 overexpression led to enhanced grain length. These results suggest that the 364D-E and 499H-Y substitutions both influence the function of *GL3.1*.

*GL3.1* functions as a Ser/Thr phosphatase

Although some nucleotides were different between the promoters of FAZ1 and WY3, the *GL3.1* expression pattern remained similar (Supplementary information, Figure S5). *GL3.1* was expressed in all organs and developmental stages tested in FAZ1 and NIL (Supplementary information, Figure S6A-S6C). Notably, expression of *GL3.1* was higher in the panicle of NIL at the heading stage than in the panicle of FAZ1 (Figure 3A) and lower in the calluses from FAZ1 and NIL, which primarily con-



**Figure 2** Transgenic analysis of *GL3.1*. **(A)** *GL3.1* gene structure. Orange blocks indicate exons; black lines indicate introns; numbers indicate the intron (top) and exon (below) sizes (bp); purple and green blocks indicate conserved domains. **(B)** Grain phenotypes of *GL3.1*-WY3-overexpressing transgenic lines. Scale bar, 2 mm. **(C)** Grain length of *GL3.1*-WY3-overexpressing transgenic lines ( $n = 10$ ). **(D)** Real-time PCR analysis of *GL3.1* expression in *GL3.1*-WY3-overexpressing transgenic lines; the expression levels have been normalized to those of ubiquitin and are expressed relative to the vector ( $n = 3$ ). A Student's *t*-test was used to generate the *P* values in **C**.



**Figure 3** Expression pattern and molecular function of *GL3.1*. **(A)** Real-time PCR analysis of *GL3.1* expression in FAZ1 and NIL during the heading stage; the expression levels have been normalized to ubiquitin and are expressed relative to FAZ1 seedling leaves ( $n = 3$ ). **(B)** Subcellular location of *GL3.1*. *GL3.1* was fused to enhanced yellow fluorescent protein (eYFP), under the control of the 35S promoter, and transformed into *Arabidopsis* protoplasts before examination using confocal laser-scanning microscopy. Scale bar, 10  $\mu\text{m}$ . **(C)** The enzymatic activity of *GL3.1*. *GL3.1*-FAZ1, *GL3.1*-WY3 and *GL3.1* from Nanyangzhan (*GL3.1*-NYZ) was examined using MBP fusions. The amount (pmol) of free phosphates released by *GL3.1*-FAZ1::MBP, *GL3.1*-WY3::MBP, *GL3.1*-M1::MBP, *GL3.1*-M2::MBP and MBP was recorded ( $n = 3$ ); the  $P$  value was generated by comparing *GL3.1*-WY3::MBP, *GL3.1*-M1::MBP and *GL3.1*-M2::MBP with *GL3.1*-FAZ1::MBP. **(D)** Pharmacological testing of *GL3.1* phosphatase activity ( $n = 3$ ). The  $P$  value was generated through a comparison of the *GL3.1*-FAZ1::MBP (with inhibitors) with *GL3.1*-FAZ1::MBP (without inhibitors). A Student's  $t$ -test was used to generate the  $P$  values in **A**, **C** and **D**.

sist of dividing cells (Supplementary information, Figure S6D). *GL3.1* in both parents was detected throughout the entire cell (Figure 3B). Purified *GL3.1*-FAZ1 and *GL3.1*-WY3 dephosphorylated myelin basic protein (MyBP; a standard substrate) *in vitro*, which demonstrates that *GL3.1* is a functional Ser/Thr phosphatase (Figure 3C); however, *GL3.1*-FAZ1 exhibited higher activity than *GL3.1*-WY3. In addition, *GL3.1* from Nanyangzhan (*GL3.1*-NYZ) did not show dephosphorylation ability (Figure 3C). *GL3.1* was insensitive to both okadaic acid

(OA) and Inhibitor 2 (Figure 3D), which suggests that *GL3.1* may encode a novel type of PPKL, as the PPKL and PP1 enzymes are generally sensitive to Inhibitor 2 [27], whereas the PPKL family member BSU1 is sensitive to OA [32]. Furthermore, we observed that the phosphatase domain of *GL3.1* was also not sensitive to these two inhibitors (Supplementary information, Figure S6E). A comparative analysis of phosphatases that are typically sensitive to OA revealed that 929G in *GL3.1* conferred resistance to OA (Supplementary information, Figure

S7). According to a previous study [40] in rats, PP2A $\alpha$  is sensitive to OA, but the Y267G mutant of this protein is resistant to OA, which indicates that Y267G is a key mutation for OA resistance. Alignment analysis revealed that 267Y in rat PP2A $\alpha$  corresponds to 929G in *GL3.1*, which suggests that *GL3.1* harbors a Y to G mutation at this key site that may be responsible for the OA insensitivity of *GL3.1*.

#### *GL3.1* regulates spikelet hull cell division

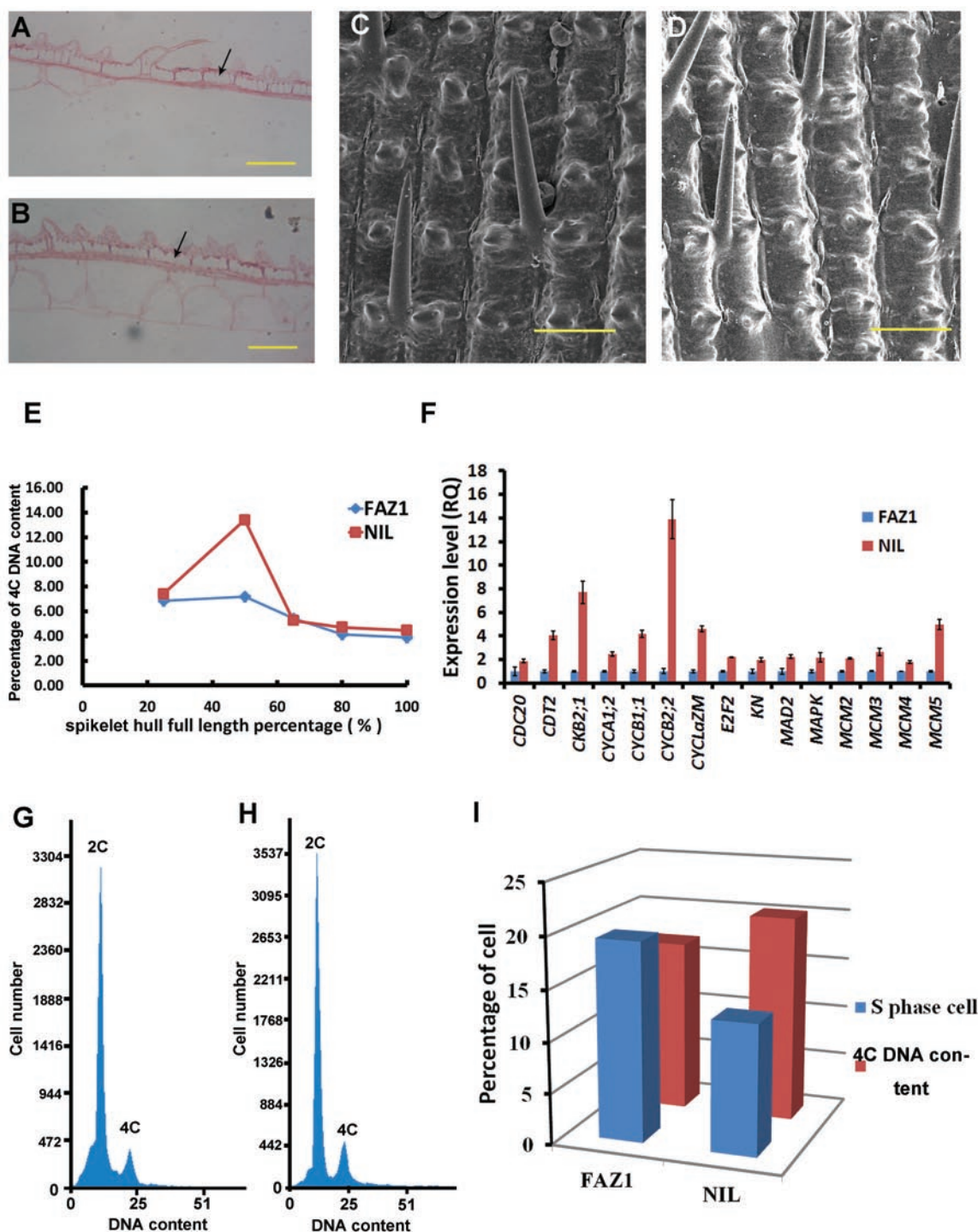
We measured the lemma cell length to determine whether *GL3.1* regulates grain length. There were no significant differences in cell length at the central point of the lemma in the vertical orientation (Figure 4A-4B, Supplementary information, Figure S6F) and the distance between tubercles at the outer spikelet hull surface (Figure 4C-4D, Supplementary information, Figure S6G), which indicates that cell length is not responsible for the difference in grain length between FAZ1 and NIL. As the FAZ1 and NIL life cycles and heading days are similar, we hypothesized that an increased rate of cell division may be responsible for the longer spikelet hull in NIL. Therefore, we assessed the cell division rate during different developmental stages of the spikelet hull in the two parents. The detection points were set at the stages when the spikelet hull length reached 25%, 50%, 65%, 80% and 100% of the full spikelet hull length in FAZ1 and NIL, and the percentage of cells with 4C DNA content in the spikelet hull as well as the cell lengths at the central zone of the spikelet hull at these points were also recorded (Figure 4E, Supplementary information, Figure S6F). At the five detection points, the cell length in the vertical orientation was not different between FAZ1 and NIL. Notably, at 50% of full spikelet hull length, the percentage of cells with a 4C DNA content was significantly higher in NIL than in FAZ1 (Figure 4E). Consistent with this, the expression of cell cycle-related genes was significantly higher in the NIL than in the FAZ1 spikelet (Figure 4F). Therefore, we propose that rapid cell division occurs in NIL during spikelet hull development. Furthermore, we synchronized cells from FAZ1 and NIL using hydroxycarbamide, which blocks cell division at the G1/S boundary. 8 h after release from hydroxycarbamide, the expression of Histone H4 was maximal in FAZ1 and NIL (Supplementary information, Figure S6H), which suggests that the cells from FAZ1 and NIL had entered the S phase. In addition, a higher percentage of cells with 4C DNA content and a lower percentage of cells in S phase were observed in NIL when compared with FAZ1 (Figure 4G-4I), which implies that more cells from NIL completed DNA duplication. We also observed that the maximal expression of *CYCD4;1*, which was

expressed from early G2 phase to M phase, was earlier in NIL (28 h after release) than in FAZ1 (32 h after release) (Supplementary information, Figure S6I), which implies faster entry into the G2 phase in NIL cells. Furthermore, we synchronized cells from FAZ1 and NIL using nocodazole, which blocks cell division at the G2/M boundary. The expression of *CYCD3;1*, which is specifically expressed at the G1/S stage, remained the same between FAZ1 and NIL (Supplementary information, Figure S6J), which implies that the transformation from the G2 phase to G1 phase was not different between the two parents. These results suggest that the transformation from G1 to G2 may be accelerated in NIL. Thus, our results collectively demonstrate that the *GL3.1*-WY3 allele increases the rate of cell division during spikelet hull development compared with the *GL3.1*-FAZ1 allele, which results in a longer spikelet hull.

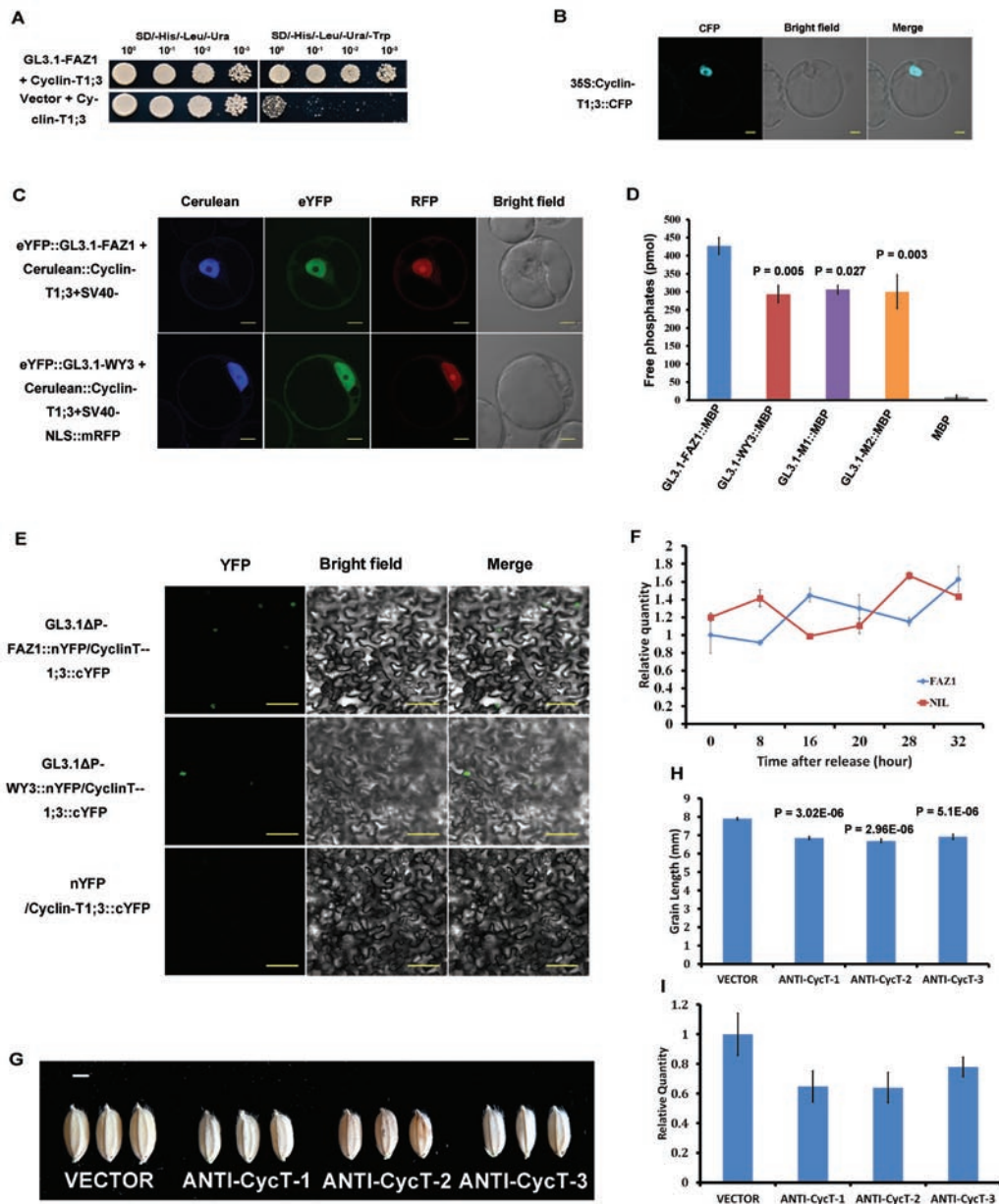
#### *GL3.1* interacts with Cyclin-T1;3 to regulate grain length

To determine the mechanism underlying the *GL3.1*-mediated regulation of grain size, a yeast two-hybrid system was used to screen a cDNA library constructed from Zhonghua 11 spikelets using *GL3.1*-FAZ1 as bait. We identified 23 interacting proteins, of which Os11g05850, annotated as Cyclin-T1;3, was selected for further analysis (Figure 5A). The expression of Cyclin-T1;3 was localized to the nucleus of *Arabidopsis* protoplasts (Figure 5B), which was consistent with the expression pattern observed in humans. When *GL3.1* was co-expressed with Cyclin-T1;3, increased accumulation of *GL3.1* was observed in the nuclei from both parents when compared with expression without Cyclin-T1;3 (Figures 5C and 3B). We confirmed that *GL3.1* dephosphorylated Cyclin-T1;3 *in vitro*. *GL3.1*-FAZ1 exhibited stronger Cyclin-T1;3 dephosphorylation activity than *GL3.1*-WY3, *GL3.1*-M1 and *GL3.1*-M2 (Figure 5D), which was consistent with the effects observed for the common substrate MyBP (Figure 3C). As the kelch-repeat domain has demonstrated potential for protein interactions [41], we used a bimolecular fluorescence complementation (BiFC) assay to identify such interactions. *GL3.1* $\Delta$ P, which contains a kelch-repeat domain, interacted with Cyclin-T1;3 in the absence of the *GL3.1* Ser/Thr phosphatase domain (Figure 5E). In addition, Cyclin-T1;3 was constitutively expressed in various tissues and organs in a pattern similar to that of *GL3.1* (Supplementary information, Figure S8A-S8D). These results demonstrate that Cyclin-T1;3 interacts with *GL3.1* and is dephosphorylated through *GL3.1*.

Real-time PCR was performed to analyze the expression of *Cyclin-T1;3* after cell synchronization. In contrast to the high expression levels observed at 16 h and 32 h



**Figure 4** *GL3.1* alters spikelet hull cell division to regulate grain length. (A, B) Longitudinal sections of FAZ1 (A) and NIL (B) lemma before flowering. Scale bar, 100  $\mu$ m. (C, D) Scanning electron microscopic analysis of the outer spikelet hull surfaces of FAZ1 (C) and NIL (D). Scale bar, 100  $\mu$ m. (E) Flow cytometric analysis of FAZ1 and NIL spikelet hulls at five developmental stages ( $n = 50$ ). The percentage of cells with 4C DNA content was determined as a proportion of the total number of isoploid and tetraploid nuclei. (F) Quantitative real-time PCR analysis of cell cycle-related gene expression in FAZ1 and NIL spikelets during the heading stage. The expression of these genes in the spikelet of NIL was normalized to that of *ubiquitin* and expressed relative to FAZ1 spikelets at this stage ( $n = 3$ ). (G-H) Flow cytometric analysis showing that the cells of FAZ1 (G) contain less 4C DNA than those of NIL (H). (I) Percentage of cells in S phase, and the number of cells containing 4C DNA content in FAZ1 and NIL at 8 h after release.



**Figure 5** GL3.1 and Cyclin-T1;3 interact. **(A)** Yeast two-hybrid assay confirmed the interaction between GL3.1-FAZ1 and Cyclin-T1;3. *GL3.1-FAZ1* was fused with the bait in PPC86, and *Cyclin-T1;3* was cloned into PEG202-GL3.1-FAZ1 with the prey. **(B)** Nuclear localization of Cyclin-T1;3. *Cyclin-T1;3* was fused to CFP, transformed into *Arabidopsis* protoplasts and observed using confocal laser-scanning microscopy. Left panel: CFP; middle panel: bright field; right panel: merge. Scale bar, 10  $\mu$ m. **(C)** Co-expression of eYFP::GL3.1-FAZ1 or eYFP::GL3.1-WY3 with Cerulean::Cyclin-T1;3 and SV40 nuclear localization signal (NLS) were fused with mRFP in *Arabidopsis* protoplasts. Scale bar, 10  $\mu$ m. **(D)** GL3.1 dephosphorylates the phosphorylated form of Cyclin-T1;3::GST. The amount (pmol) of free phosphate released by GL3.1-FAZ1::MBP, GL3.1-WY3::MBP, GL3.1-M1::MBP, GL3.1-M2::MBP and MBP was recorded ( $n = 3$ ). The  $P$  value was obtained through a comparison of GL3.1-WY3::MBP, GL3.1-M1::MBP and GL3.1-M2::MBP with GL3.1-FAZ1::MBP. **(E)** BiFC assay confirmed the interaction between the N-terminus of GL3.1 and Cyclin-T1;3. Truncated GL3.1 lacking the GL3.1 phosphatase domain ( $\Delta$ P) from both parents was fused with the N-terminus of YFP, whereas Cyclin-T1;3 was fused with the C-terminus of YFP. Scale bar, 50  $\mu$ m. **(F)** Real-time PCR analysis of *Cyclin-T1;3* expression in FAZ1 and NIL cells after release; the expression levels were normalized to those of ubiquitin and are expressed relative to those of FAZ1 cells at 0 h after release ( $n = 3$ ). **(G)** Grain phenotypes of *Cyclin-T1;3* antisense transgenic lines. Scale bar, 2 mm. **(H)** Grain length of *Cyclin-T1;3* antisense transgenic lines ( $n = 10$ ). **(I)** Real-time PCR analysis of *Cyclin-T1;3* expression in *Cyclin-T1;3* antisense transgenic lines; the expression levels were normalized to those of *ubiquitin* and are expressed relative to the vector ( $n = 3$ ). A Student's  $t$ -test was used to generate the  $P$  values in **D** and **H**.



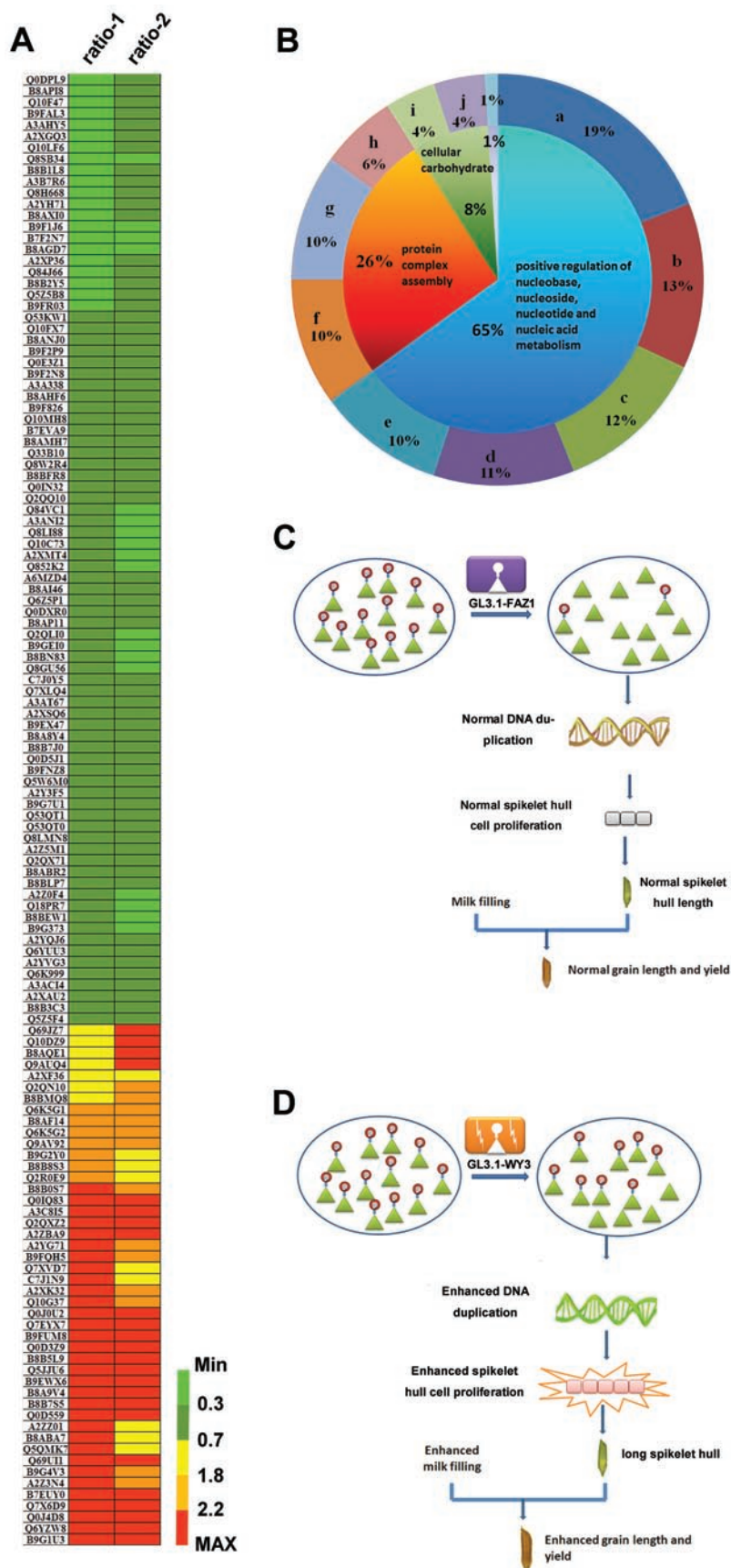
in FAZ1 cells, Cyclin-T1;3 showed increased expression at 8 h and 28 h in NIL cells (Figure 5F). At these two specific points, the cells were entering the S and G2 phases, respectively, which indicates that Cyclin-T1;3 may be involved in cell cycle control. Moreover, we used transgenic rice plants to determine whether Cyclin-T1;3 influences grain size. No obvious phenotype was observed when we overexpressed Cyclin-T1;3 in Zhonghua 11. However, antisense strands of Cyclin-T1;3 resulted in smaller grain sizes in the transgenic plants as well as reduced expression of Cyclin-T1;3 (Figure 5G-5I). Thus, our data indicate that Cyclin-T1;3 is involved in the *GL3.1*-mediated regulation of grain length.

#### *GL3.1 is a widespread gene that influences protein phosphorylation in vivo*

A quantitative proteomic analysis using two-dimensional difference gel electrophoresis (2-D DIGE) indicated that 21 proteins were differentially expressed in FAZ1 and NIL, of which 18 were upregulated in NIL (Supplementary information, Table S1). Mass spectrometry revealed that the 21 proteins were associated with cellular metabolic processes. Interestingly, actin was upregulated in NIL, which is consistent with the observation that *GL3.1* influences the rate of cell proliferation. The phosphopeptides from the young spikelets of FAZ1 and NIL were enriched on the TiO<sub>2</sub> beads and quantified using iTRAQ, which confirmed that *GL3.1* is a phosphatase. 556 phosphopeptides were detected, and 464 of these molecules were quantified. Proteins showing a 1.5-fold difference between FAZ1 and NIL and demonstrating the same trend when quantified using two different labelling systems were chosen for further analysis. At least 130 proteins demonstrated a different phosphorylation status between FAZ1 and NIL during spikelet development (Figure 6A and Supplementary information, Table S2). Gene ontology analysis revealed that these proteins are primarily involved in processes related to nucleic acid metabolism and protein complex assembly (Figure 6B and Supplementary information, Table S3). The molecular functions of these proteins include nucleotide binding and the activities of phosphotransferases, helicases and the RNA polymerase II transcription factor. These results strongly suggest that *GL3.1* could influence DNA duplication. Thus, we propose that *GL3.1* regulates the expression and phosphorylation of a variety of genes involved in metabolism and cell division. Further phylogenetic analyzes based on genomic BLAST searches demonstrated the widespread existence of *GL3.1* in plants (Supplementary information, Figure S9), which suggests that *GL3.1* has an important conserved function in plants.

## Discussion

In this study, we cloned *GL3.1*, which is a QTL that controls rice grain length. The natural *GL3.1* allele increased grain length, weight and yield. Interestingly, we did not observe significant changes in the other traits in NIL, which potentially indicates that *GL3.1* primarily regulates spikelet hull development. Therefore, *GL3.1* may be a useful gene for breeding to achieve a higher yield. *GL3.1* was identified as a novel type of Ser/Thr phosphatase that belongs to the PPKL subfamily. Little is known about the PPKL family, which is only distributed in the Viridiplantae. It is thought that a common progenitor captured an ancestral PP1-like enzyme containing a kelch-repeat domain via endosymbiosis [42, 43], and subsequently the PPKLs evolved independently in algae and plants. The enzymatic properties of PPKLs are closely related to those of PP1; however, high concentrations of either Inhibitor 2 or OA did not inhibit the phosphatase activity of *GL3.1*. Therefore, *GL3.1* represents a novel type of PPKL phosphatase. In contrast, kelch-repeat proteins have been well characterized and are widely conserved from *Caenorhabditis elegans* to humans. Kelch-repeat motifs typically form conserved  $\beta$ -propeller tertiary structures and contain potential protein-protein interaction sites [41]. Kelch-repeat proteins regulate many cell processes through associations with the actin cytoskeleton [44, 45]; regulate transcription, cell morphology and organization [46]; or exert extracellular functions [47]. We demonstrated that the fragment containing the kelch domain of *GL3.1* interacts with Cyclin-T1;3 and observed that actin and other cell cycle genes were upregulated in NIL. In the present study, Cyclin-T1;3 was identified as a substrate of *GL3.1* in the grain length regulation pathway. Thus, we propose that Cyclin-T1;3 influences the spikelet hull cell cycle during its development. It is generally accepted that Cyclin T1 controls the elongation phase of RNA polymerase II through the formation of a P-TEFb complex with CDK9 [48]; however, other studies have suggested that Cyclin T1 may affect cell proliferation [12, 49]. We observed that the expression of *Cyclin-T1;3* in FAZ1 was delayed for 4 h to 8 h compared with that in NIL. *Cyclin-T1;3* may play an important role in cell division because its expression was associated with the percentage of cells with a 4C DNA content at 8 h after release. However, the similarity of the expression pattern to that of *CYCD4;1* indicates that *Cyclin-T1;3* may also be involved in G2-phase control. Furthermore, we observed that the phosphorylation of some proteins related to RNA polymerase II was different between FAZ1 and NIL during spikelet development, which implies that *Cyclin-T1;3* involve-



**Figure 6** GL3.1 influences protein phosphorylation status as a potential model for grain size control. **(A)** Differences in the quantities of phosphoproteins in FAZ1 and NIL. iTRAQ was used to determine the protein quantities. The proteins from FAZ1 were labelled with 114 and 116, whereas the proteins from NIL were labelled with 115 and 117. Ratio-1 corresponds to 115/114, while ratio-2 corresponds to 117/116. **(B)** The proteins mentioned in **A** are involved in various biological processes. The number of proteins associated with each biological process was calculated in *REIVGO* (revigo.irb.hr) using the  $-\log_{10}$  *P*-value method. **a** to **j** represent distinct biological processes. **a**, positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism; **b**, transcription initiation from RNA polymerase II promoter; **c**, positive regulation of biological processes; **d**, transcription from RNA Polymerase II promoter; **e**, transcription initiation, DNA-dependent; **f**, protein complex assembly; **g**, protein complex biogenesis; **h**, macromolecular complex subunit organization; **i**, cellular carbohydrate metabolism; **j**, carbohydrate metabolism. **(C, D)** Potential model for the regulation of grain size through GL3.1. GL3.1 dephosphorylates substrates, such as Cyclin-T1;3, which promotes spikelet hull cell proliferation. In the FAZ1 allele, most phosphorylated substrates (green triangles) are dephosphorylated through GL3.1-FAZ1, which results in normal spikelet hull development **(C)**. In the WY3 allele, the weaker dephosphorylation activity of *GL3.1-WY3* may result in the accumulation of phosphorylated substrates, thereby increasing the rate of cell division **(D)** and accelerating milk filling to result in a longer grain.

ment in the *GL3.1*-mediated pathway controls cell division.

*GL3.1-WY3* differs from *GL3.1-FAZ1* by two amino acid substitutions, induces a longer grain length as a result of an increased number of cells in the spikelet hull, and accelerates milk filling, which leads to increased grain weight and yield. *GL3.1* influenced grain length through an increased rate of cell division during spikelet hull development, as evidenced by the upregulation of cell cycle-related genes and an increase in the proportion of cells with 4C DNA content; however, there was no effect on cell size. Protein phosphorylation and dephosphorylation critically regulate the cell cycle [50]. The dephosphorylation activity of *GL3.1-FAZ1* was higher than that of *GL3.1-WY3*, *GL3.1-M1* and *GL3.1-M2*; therefore, we hypothesize that the ability of *GL3.1* to dephosphorylate substrates, such as Cyclin-T1;3, may regulate cell proliferation during spikelet development. We propose that the high activity of *GL3.1-FAZ1* dephosphorylates substrates, such as Cyclin-T1;3, to promote cell proliferation, which leads to normal cell proliferation during spikelet hull development (Figure 6C). For the *WY3* allele, the weaker dephosphorylation activity of *GL3.1-WY3* resulted in an accumulation of phosphorylated substrates that promoted cell proliferation during spikelet development and increased the grain length (Figure 6D). Only transgenic lines expressing high levels of *GL3.1-WY3* had an increased grain size, which indicates that a significant increase in the expression of the low-activity *GL3.1-WY3* allele would result in competition for substrates with the abundant endogenous *GL3.1-FAZ1*, which has increased dephosphorylation activity. As a result, higher quantities of phosphorylated substrates accumulate in *GL3.1-WY3*-overexpressing transgenic plants, which leads to accelerated cell division and eventually results in an elongated grain. Thus, we propose that *Cyclin-T1;3* is one of the substrates in this pathway. In summary, we hypothesize that the larger grain size in lines overexpressing *GL3.1-WY3*, *GL3.1-M1* and *GL3.1-M2* that contain high levels of exogenous *GL3.1* may result from dominant negative effects.

Efficient breeding requires a thorough understanding of the molecular and genetic factors that regulate seed development. Although considerable research has been performed to this end, an investigation of the relationship among cloned QTLs for grain size is needed. According to a previous study, *GW2* and *GW5* both function in the ubiquitin-proteasome pathway and control grain size through the regulation of cell division; thus, *GW2* and *GW5* may be involved in the same pathway for grain size control [7-9]. However, *GL3.1* and *GS3* encode different types of proteins [10], and whether they function

in the same pathway remains unknown. Determining the relationship between these QTLs would be useful for understanding grain development. However, *GL3.1* is similar to *BSU1*, and whether *GL3.1* plays a role in BR responses needs to be further investigated. In this study, we cloned a new QTL, *GL3.1*, characterized its function, and further identified the *GL3.1*-mediated pathway that regulates grain size and yield. This results obtained in this study contribute to the understanding of the molecular and genetic factors involved in seed development and provide a new approach for improving crop grain yield.

## Materials and Methods

### Plant materials

The large grain *WY3* of the *japonica* variety and the small grain *FAZ1* of the *indica* variety were chosen as parents for mapping the genes that control grain length. *WY3* was crossed with *FAZ1* to produce F1 plants, and subsequently, the F1 plants were backcrossed with *FAZ1* as the recurrent parent to generate BC1F1. Repetitive backcrossing and marker-assisted selection produced some plants where the region surrounding *GL3.1* was heterozygous, but nearly all other regions were homozygous for *FAZ1*. We used segregating populations for the fine- and high-resolution mapping of *GL3.1*. From the BC4F2 generations, we developed a nearly isogenic line (NIL) for *GL3.1*, which contained a 30-kb region from the *WY3* chromosomal region including the *GL3.1* locus in the *FAZ1* genetic background. In total, 107 markers were used to screen the genetic background of NIL.

### Fine-mapping of *GL3.1*

We used 179 BC2F2 plants for the fine-mapping of *GL3.1* at chromosome 3. Subsequently, 5 542 BC3F2 plants were used to further fine-map *GL3.1*. The phenotypes of the grains in selected recombinants were further confirmed using BC3F3 progeny. To further determine the location of *GL3.1*, we developed markers based on the sequence of the PAC clone (AL731878). The molecular marker primers are listed in Supplementary information, Table S4. The candidate genes for *GL3.1* from *FAZ1* and *WY3* were sequenced and compared.

### Trait measurement

The grain length, width, thickness and 1 000-grain weight were measured when the plants were completely matured. For the trait measurements, we chose grains from the top part of the largest panicle. To determine grain length, width and thickness, 20 grains were measured. To determine the 1 000-grain weight, 100 grains were weighed, and the measurement was repeated three times. The entire spikelet hull, including the lemma and palea, were included in this measurement. The endosperm was dried at 90 °C overnight and the dry weight was measured. For the plot yield tests, all plants in the plots were collected and threshed (the plants at the border of plots were removed before harvest). After most of the impurities were removed, the grains were weighed, and the weight was recorded as W1. Subsequently, 5 kg of grain from each plot was dried at 42 °C for 1 week. When the water content (WC) was below 17% (measured using a seed moisture analyzer), the grains

were weighed, and the weight was recorded as W2. The standard water content of the rice grains was 13.5%. The plot yields were determined from  $W3 = W1 \times W2 \times (1 - WC) / 5 / (1 - 13.5\%)$ . We repeated the yield tests for FAZ1 and NIL twice. We did not replicate the tests for the yield of Huanghuazhan and Huanghuazhan (*GL3.1*) because we used a larger area (about 110 m<sup>2</sup>) to test the yield.

#### RNA extraction and real-time PCR

Total RNA was isolated from different tissues of FAZ1, WY3 and NIL plants during different stages using the RNeasy Plant mini kit (Qiagen). The spikelet tissues included the palea and lemma. RNase-free DNase I treatment was performed to remove any genomic DNA contamination. First-strand cDNA was synthesized according to a published method [39]. Real-time PCR was performed on an ABI 7300 using the SYBR Green method. The cycling conditions were 30 s at 95 °C followed by 40 cycles of amplification (95 °C for 5 s and 60 °C for 31 s). The ubiquitin gene was used as an endogenous control to normalize the detected gene expression. To determine the number of transgenic copies, we used DNA as a template and *SPS* as an endogenous control to examine the expression of Hyg. The data were analyzed according to the  $\Delta\Delta C_t$  method. Each analysis was repeated three times. The primers used for real-time PCR are listed in Supplementary information, Table S4.

#### Prokaryotic expression and purification

The full-length *GL3.1* ORFs from FAZ1 and WY3 were amplified using specific primers (Supplementary information, Table S4), digested using EcoRI and Sall and cloned into pMAL-c4x (NEB). Mutated *GL3.1*-M1 and *GL3.1*-M2 were constructed as previously described [20] and also cloned into pMAL-c4x. The constructs were transformed into the bacterial strain TB1 (NEB). The transformed cells were cultured at 37 °C overnight and transferred into fresh medium at a 1:50 dilution. After 2 h, and once the OD<sub>600</sub> of the cell culture had reached 0.3, 1 mM isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) was added into the medium to induce protein expression. The cells were cultured for 16 h at 20 °C and subsequently lysed using an ultrasonic crusher. After centrifugation at 16 000 $\times$  g for 30 min at 4 °C, the supernatant was collected for protein purification. The fusion proteins were purified on amylose resin (NEB) [51].

The ORF of *Cyclin-T1;3* was amplified using specific primers, and the products were digested using *Sall* and *NotI* and cloned into pGEX4T. The construct was transformed into Rosetta (DE3) bacteria. The transformed cells were cultured at 37 °C overnight and transferred into fresh medium at a 1:50 dilution. Protein expression was induced after 2 h as described above. The cells were cultured for 5 h at 37 °C. The supernatant containing the fusion protein was collected as described above [51] and purified using glutathione Sepharose (GE).

#### Yeast two-hybrid assay

*GL3.1*-FAZ1 was cloned into the vector PEG202 (Invitrogen) with bait, and a rice cDNA library from the young spikelet of Zhonghua 11 was constructed into PPC86 with prey. These two constructions were co-transformed with the reporter vector psh18-34 into the yeast strain EGY48. The yeast colonies were cultured on -Leu/-His/-Ura SD media, and after two days, the colonies were printed onto -Leu/-His/-Ura/-Trp SD media to screen for positive

colonies. The genes encoding proteins that interact with *GL3.1* were characterized through sequencing of the positive colonies. *Cyclin-T1;3* was cloned into PPC86 and co-transformed with PEG202-*GL3.1*-FAZ1 and psh18-34 into the EGY48 to confirm the interaction. The overnight yeast cultures were diluted to an OD<sub>600</sub> of 0.5, serially diluted and printed onto either -Leu/-His/-Ura SD or -Leu/-His/-Ura/-Trp SD media.

#### Subcellular localization, BiFC

The coding sequences of *GL3.1* from FAZ1 and WY3 were fused with yellow fluorescent protein (YFP) or eYFP; *Cyclin-T1;3* was fused with CFP or Cerulean; the SV40 nuclear localization signal was fused with mRFP; and all of these fusion proteins were transcribed from the 35S promoter. These plasmids were co-transformed or separately transformed into the protoplast of *Arabidopsis* using an established method [52]. The subcellular and co-localization of these proteins was evaluated using a confocal laser-scanning microscope (FluoView FV1000; Olympus).

For the BiFC tests, *GL3.1* $\Delta P$  was fused with the N-terminus of YFP and *Cyclin-T1;3* was fused with the C-terminus of YFP. The resulting plasmids were transformed into *Agrobacterium* strain GV3101. After culturing for 36 h, the thallus of *GL3.1* $\Delta P$ -nYFP and *Cyclin-T1;3*-cYFP was co-injected into the young leaves of tobacco. Transformed tobacco plants were cultured for 72 h at 22 °C, and subsequently, the leaves were observed using confocal laser-scanning microscopy (LAM510; Carl Zeiss).

#### Vector construction and rice transformation

We used the Site-Directed Mutagenesis Kit (Stratagene) to generate the fragments of *GL3.1*-M1 and *GL3.1*-M2. The ORFs of *GL3.1*-FAZ1, *GL3.1*-WY3, *GL3.1*-M1 and *GL3.1*-M2 were cloned into pHB [53], which contains two CaMV 35S promoters, for plant overexpression; the reversed *GL3.1*-FAZ1 ORF was cloned into pHB to generate *GL3.1*-antisense transgenic lines. For *GL3.1*-RNAi constructs, a *GL3.1*-FAZ1 250 bp fragment was also cloned into the plant RNAi vector Tck303. The ORFs of *Cyclin-T1;3* were cloned into pHB for plant overexpression, and the reversed *Cyclin-T1;3* ORF was cloned into pHB to generate *Cyclin-T1;3* antisense transgenic lines. All constructs were confirmed through sequencing and transformed into the small grain *japonica* variety Zhonghua 11 using an *Agrobacterium*-mediated method [7].

#### Synchronization cultures

The brown grains of FAZ1 and NIL were placed onto N6D2 medium to induce calluses. These calluses were separated into small balls (2 mm in diameter) and placed onto N6D2 medium containing 5 mM hydroxycarbamide or 0.3  $\mu$ M nocodazole for 18 h. The calluses were washed three times with liquid N6D2 and again placed on N6D2 medium. Five calluses were collected for RNA isolation and flow cytometric analysis at each detection point.

#### Nucleus isolation and assessment of ploidy

After the spikelet hull length (when the length reached the 5 detection points mentioned earlier) was measured, the spikelet hulls at different developmental stages were soaked separately in the nuclear isolation mixture (Beckman). In total, 50 spikelet hulls were selected and chopped with a sharp blade. After filtering through a 40- $\mu$ m nylon riddle, the suspension was loaded into the Mofol from Beckman (for flow cytometric analysis) to assess the

ploidy of the nuclei. For each test, the ploidy of 10 000 to 20 000 nuclei was recorded. The numbers of isoploid and tetraploid nuclei were recorded, and subsequently, the percentage of cells with 4C DNA content was calculated from the number of cells with 4C DNA content divided by the total number of isoploids and cells with 4C DNA content. The proportions of G1, G2 and S phase cells were analyzed using FCS Express 4 (De novo Software).

#### Enzyme activity assessment

Prior to the enzyme activity tests, different forms of GL3.1, Cyclin-T1;3 and MyBP were phosphorylated using protein kinase A (NEB) at 30 °C overnight according to the manufacturer's instructions. Considering that the activity of many phosphatases is regulated through phosphorylation, GL3.1 was also phosphorylated using the same method. After phosphorylation, 50 ng GL3.1::MBP or MBP was incubated with 1 µg of Cyclin-T1;3 or MyBP at 30 °C for 30 min, and the enzyme activity was measured according to the manufacturer's instructions (Serine/Threonine Phosphatase Assay System, Promega). The inhibitors OA (NEB) and Inhibitor 2 (NEB) were added when necessary. The optical densities of the samples were recorded using Synergy 2 (BioTek).

#### DIGE and protein identification

Young spikelets from FAZ1 and NIL (panicle length less than 5 cm) were chosen as test materials for DIGE. We used methods previously described [54], with slight modifications. The test materials were powdered in liquid nitrogen, and subsequently, the proteins were extracted using an ultrasonic crusher. After centrifugation at 12 000 r.p.m. for 45 min at 4 °C, the supernatant was collected and the protein components were quantified using the Bradford method. For DIGE, 50 ng of proteins were labelled with 400 pmol of dye; proteins from FAZ1 and NIL were labelled separately with Cy3 and Cy5. Total protein was labelled with Cy2 as an endogenous control. For each assay, 50 ng of protein was loaded onto each gel. The images were analyzed using DeCyder 2D 6.5 (GE Healthcare). Matrix-assisted laser desorption ionization-time of flight/time of flight (MALDI-TOF/TOF) mass spectrometry (MS) was used for protein identification.

#### Isolation, identification and quantification of phosphoproteins

Proteins from young spikelets (panicle length less than 5 cm) of FAZ1 and NIL were isolated according to the method used for DIGE and digested in buffer (10 mM dithiothreitol, 50 mM IAA, 2 U trypsin buffer) at 37 °C for 18 h. Then, the peptides were labelled according to the illustration of iTRAQ (ABI company); the peptides from FAZ1 were labelled with 114 and 116, whereas peptides from NIL were labelled with 115 and 117 [55]. TiO<sub>2</sub> beads were used to collect the phosphorylated peptides. After elution from the beads, the peptides were analyzed using nano-liquid chromatography (LC)-MS/MS, and the resulting data were analyzed at [www.bioinfo.cau.edu.cn/agriGO](http://www.bioinfo.cau.edu.cn/agriGO).

#### Acknowledgments

We would like to thank Mr Xiao-Yan Gao, Mr Xiao-Shu Gao and Mr Wei-fang Jiang for assistance with transmission electron microscopy, confocal laser scanning microscopy and flow cytometry, respectively. This work was supported by the Min-

istry of Science and Technology of China (2012CB944800 and 2012AA10A302), the National Natural Science Foundation of China (31130071 and 31121063), the Ministry of Agriculture of China (2009ZX08009-102B), the Chinese Academy of Sciences (KSCX3-EW-N-01), the Research Grants Council of Hong Kong (CUHK466011 and CUHK2/CRF/11G) and the CAS-Croucher Foundation.

#### References

- 1 Tilman D, Fargione J, Wolff B, *et al.* Forecasting agriculturally driven global environmental change. *Science* 292:281-284.
- 2 Hibberd JM, Sheehy JE, Langdale JA. Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. *Curr Opin Plant Biol* 2008; 11:228-231.
- 3 Khush GS. Green revolution: preparing for the 21st century. *Genome* 1999; 42:646-655.
- 4 Khush GS. What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol Biol* 2005; 59:1-6.
- 5 Zhang Q. Strategies for developing Green Super Rice. *Proc Natl Acad Sci USA* 2007; 104:16402-16409.
- 6 Kropff M, Cassman K, Peng S, Matthews R, Setter T. Quantitative understanding of yield potential. In: Cassman K, ed. *Breaking the Yield Barrier*. Los Baños, Philippines: International Rice Research Institute, 1994:21-38.
- 7 Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet* 2007; 39:623-630.
- 8 Shomura A, Izawa T, Ebana K, *et al.* Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet* 2008; 40:1023-1028.
- 9 Weng J, Gu S, Wan X, *et al.* Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res* 2008; 18:1199-1209.
- 10 Mao H, Sun S, Yao J, *et al.* Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc Natl Acad Sci USA* 2010; 107:19579-19584.
- 11 Fan C, Xing Y, Mao H, *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* 2006; 112:1164-1171.
- 12 Li Y, Fan C, Xing Y, *et al.* Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat Genet* 2011; 43:1266-1269.
- 13 Wang E, Wang J, Zhu X, *et al.* Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat Genet* 2008; 40:1370-1374.
- 14 Ashikari M, Sakakibara H, Lin S, *et al.* Cytokinin oxidase regulates rice grain production. *Science* 2005; 309:741-745.
- 15 Xue W, Xing Y, Weng X, *et al.* Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet* 2008; 40:761-767.
- 16 Huang X, Qian Q, Liu Z, *et al.* Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet* 2009; 41:494-497.
- 17 Li X, Qian Q, Fu Z, *et al.* Control of tillering in rice. *Nature* 2003; 422:618-621.
- 18 Jiao Y, Wang Y, Xue D, *et al.* Regulation of OsSPL14 by

- OsmiR156 defines ideal plant architecture in rice. *Nat Genet* 2010; **42**:541-544.
- 19 Miura K, Ikeda M, Matsubara A, et al. OsSPL14 promoters panicles branching and higher grain productivity in rice. *Nat Genet* 2010; **42**:545-549.
- 20 Jin J, Huang W, Gao JP, et al. Genetic control of rice plant architecture under domestication. *Nat Genet* 2008; **40**:1365-1369.
- 21 Tan L, Li X, Liu F, et al. Control of a key transition from prostrate to erect growth in rice domestication. *Nat Genet* 2008; **40**:1360-1364.
- 22 Olsen JV, Blagoev B, Gnäd F, et al. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006; **127**:635-648.
- 23 Cohen P. The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem Sci* 2000; **25**:596-601.
- 24 Farkas I, Dombradi V, Miskei M, Szabados L, Koncz C. Arabidopsis PPP family of serine/threonine phosphatases. *Trends Plant Sci* 2007; **12**:169-176.
- 25 Denu JM, Stuckey JA, Saper MA, Dixon JE. Form and function in protein dephosphorylation. *Cell* 1996; **87**:361-364.
- 26 Janssens V, Goris J. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem J* 2001; **353**:417-439.
- 27 Luan S. Protein phosphatases in plants. *Annu Rev Plant Biol* 2003; **54**:63-92.
- 28 Goldberg J, Huang HB, Kwon YG, Greengard P, Nairn AC, Kuriyan J. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* 1995; **376**:745-753.
- 29 Barton GJ, Cohen PT, Barford D. Conservation analysis and structure prediction of the protein serine/threonine phosphatases. Sequence similarity with diadenosine tetraphosphatase from *Escherichia coli* suggests homology to the protein phosphatases. *Eur J Biochem* 1994; **220**:225-237.
- 30 Kutuzov MA, Andreeva AV. Protein Ser/Thr phosphatases with kelch-like repeat domains. *Cell Signal* 2002; **14**:745-750.
- 31 Li JL, Baker DA. A putative protein serine/threonine phosphatase from *Plasmodium falciparum* contains a large N-terminal extension and five unique inserts in the catalytic domain. *Mol Biochem Parasitol* 1998; **95**:287-295.
- 32 Mora-Garcia S, Vert G, Yin Y, Cano-Delgado A, Cheong H, Chory J. Nuclear protein phosphatases with Kelch-repeat domains modulate the response to brassinosteroids in *Arabidopsis*. *Genes Dev* 2004; **18**:448-460.
- 33 Kerk D, Bulgrien J, Smith DW, Barsam B, Veretnik s, Grib-skov M. The complement of protein phosphatase catalytic subunits encoded in the genome of *Arabidopsis*. *Plant Physiol* 2002; **129**:908-925.
- 34 Kim TW, Guan S, Sun Y, et al. Brassinosteroid signal transduction from cell-surface receptor kinases to nuclear transcription factors. *Nat Cell Biol* 2009; **11**:1254-1260.
- 35 Aluko G, Martinez C, Tohme J, et al. QTL mapping of grain quantity traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. *Theor Appl Genet* 2004; **109**:630-639.
- 36 Li JM, Michael T, Susan RM. Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. *Genetics* 2004; **168**:2187-2195.
- 37 Liu TM, Mao DH, Zhang SP, Xu CG, Xing YZ. Fine mapping *SPP1*, a QTL controlling the number of spikelets per panicle, to a BAC clone in rice (*Oryza sativa*). *Theor Appl Genet* 2009; **118**:1509-1517.
- 38 Wang JK, Wan XY, Jose C, Jonathan C, Weng JF, Zhai HQ, Wan JM. QTL mapping of grain length in rice (*Oryza sativa* L.) using chromosome segment substitution lines. *Genet Res* 2006; **88**:93-104.
- 39 Ren ZH, Gao JP, Li LG, et al. A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat Genet* 2005; **37**:1141-1146.
- 40 Shigeto K, Hiroshi S, Teruo A, Minoru T, Takash S, Minako N. Analysis by *in vitro* mutagenesis of PP2A $\alpha$  okadaic acid responsive sequences. *Biochem and Biophys Res Commun* 1995; **214**:518-523.
- 41 Adams J, Kelso R, Cooley L. The kelch repeat superfamily of proteins: propellers of cell function. *Trends Cell Biol* 2000; **10**:17-24.
- 42 Weber AP, Linka M, Bhattacharya D. Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot Cell* 2006; **5**:609-612.
- 43 Cavalier-Smith T. Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* 1999; **46**:347-366.
- 44 Eichinger L, Bomblies L, Vandekerckhove J, Schleicher M, Gettemans J. A novel type of protein kinase phosphorylates actin in the actin-fragmin complex. *EMBO J* 1996; **15**:5547-5556.
- 45 Kim IF, Mohammadi E, Huang RC. Isolation and characterization of IPP, a novel human gene encoding an actin-binding, kelch-like protein. *Gene* 1999; **228**:73-83.
- 46 Philips J, Herskowitz I. Identification of Kellp, a kelch domain-containing protein involved in cell fusion and morphology in *Saccharomyces cerevisiae*. *J Cell Biol* 1998; **143**:375-389.
- 47 Duke-Cohan JS, Gu J, McLaughlin DF, et al. Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions. *Proc Natl Acad Sci USA* 1998; **95**:11336-11341.
- 48 Herrmann CH, Mancini MA. The Cdk9 and cyclin T subunits of TAK/P-TEFb localize to splicing factor-rich nuclear speckle regions. *J Cell Sci* 2001; **114**:1491-1503.
- 49 Moiola C, De Luca P, Gardner K, Vazquez E, De Siervi A. Cyclin T1 overexpression induces malignant transformation and tumor growth. *Cell Cycle* 2010; **9**:3119-3126 ()
- 50 De Wulf P, Montani F, Visintin R. Protein phosphatases take the mitotic stage. *Curr Opin Cell Biol* 2009; **21**:806-815.
- 51 Sun SY, Chao DY, Li XM, et al. OsHAL3 mediates a new pathway in the light-regulated growth of rice. *Nat Cell Biol* 2009; **11**:845-851.
- 52 Miao Y, Jiang L. Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. *Nat Protoc* 2007; **2**:2348-2353.
- 53 Mao J, Zhang YC, Sang Y, Li QH, Yang HQ. A role for *Arabidopsis* cryptochromes and COP1 in the regulation of stomatal

- opening. *Proc Natl Acad Sci USA* 2005; **102**:12270-12275.
- 54 Shi S, Chen W, Sun W. Comparative proteomic analysis of the *Arabidopsis* cb11 mutant in response to salt stress. *Proteomics* 2011; **11**:4712-4725.
- 55 Ross PL, Huang YN, Marchese JN, *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004; **12**:1154-1169.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)