

SCIENTIFIC REPORTS



OPEN

Genetic connection between cell-wall composition and grain yield via parallel QTL analysis in *indica* and *japonica* subspecies

Zuopeng Xu¹, Shance Li^{2,4}, Changquan Zhang^{1,3}, Baocai Zhang², Kongzhi Zhu¹, Yihua Zhou^{2,4} & Qiaquan Liu¹

Grain yield is a complicated trait, which is highly associated with biomass productivity. The cell wall is a central element of biomass, and its biogenesis contributes to plant architecture and development. However, the genetic link between cell-wall property and grain yield is largely unclear. Here, we report on identification of quantitative trait loci (QTLs) for grain yield-related traits and cell-wall composition with a set of chromosomal segment substitution lines (CSSLs) that were generated by using 9311, an *indica* cultivar as donor, and Nipponbare, a *japonica* cultivar as recipient. Nipponbare and 9311 showed significant differences in grain yield-related traits and cell-wall composition. Genotyping with molecular markers, 125 lines covering 95.6% of the whole genome of 9311 were employed for phenotypic and chemical examinations. Thirty-seven QTLs for grain yield-related traits and nineteen QTLs for cell-wall composition have been identified. In addition to correlation analysis, we found overlapped and closely linked QTLs for two sets of traits. Fine-mapping further narrowed a QTL for cellulose content together with *HD17*, a known QTL for heading date and grain yield, suggesting that plants may regulate cell wall biogenesis and grain yield via related means. Our study provided genetic clues for cloning QTLs for both complicated traits.

Cell walls, consisting of a complex polysaccharide network, encase plant cells and are characteristic plant cellular structures. These structures are enriched in cellulose, hemicelluloses and pectins, which represent the most abundant natural biopolymers on Earth¹. Without cell walls, plant cells lose their shape, indicating fundamental roles on plant morphogenesis and development. Land plants generally possess at least 35 types of cells that are surrounded by different wall structure and properties compatible with their functions in various tissues and across developing stages.

It has been estimated that cell wall synthesis and integration involve more than 1000 genes². Based on characterization of mutants that showed defects in plant growth and mechanical strength, multiple glycosyltransferases (GTs), glycosyl hydrolases (GHs), and other proteins have been found implicated in wall synthesis. It is well known that cellulose synthesis is catalyzed by cellulose synthase subunits belonging to GT2 family at the plasma membrane^{3,4}. Non-cellulosic polysaccharides are synthesized by multiple cellulose synthase-like proteins (CSLs) and GTs, including GT8, GT43, GT47, and GT61 family members, etc^{5–10}. Korrigan (KOR), a GH9 member, and Chitinases1 (CTL1), a GH18/19 protein, are also required for cellulose biosynthesis^{11,12}. Besides GT and GH, other proteins, such as COBRA, affect cellulose synthesis and deposition in *Arabidopsis* and rice^{13,14}. Enzymes involved in substrate synthesis and transportation are also needed for cell-wall biosynthesis^{15,16}. Two classes of proteins, xyloglucan endotransglycosylases (XET) and expansins, modify the cell wall and integrate new

¹Jiangsu Key Laboratory of Crop Genetics and Physiology/Key Laboratory of the Ministry of Education for Plant Functional Genomics, College of Agriculture, Yangzhou University, Yangzhou, 225009, China. ²State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, 100101, China. ³Co-Innovation Center for Modern Production Technology of Grain Crops of Jiangsu Province/Joint International Research Laboratory of Agriculture and Agri-Product Safety of the Ministry of Education, Yangzhou, 225009, China. ⁴University of Chinese Academy of Sciences, Beijing, 100049, China. Zuopeng Xu, Shance Li and Changquan Zhang contributed equally to this work. Correspondence and requests for materials should be addressed to Y.Z. (email: yhzhou@genetics.ac.cn) or Q.L. (email: qqliu@yzu.edu.cn)

polysaccharides during cell expansion^{17,18}. As some components and newly synthesized polysaccharides need to be secreted to the plasma membrane and apoplast, proteins related to cytoskeleton dynamics and vesicle trafficking are required^{19–22}. In addition to cell-wall metabolic processes, transcriptional regulation constitutes another level to modulate cell-wall biosynthesis^{23–25}. Upstream of the transcriptional regulatory network, receptor-like kinases acting as sensors of internal and external stimuli are responsible for perception and transduction of signals to attune downstream processes^{26,27}. Therefore, control of cell wall biogenesis requires the involvement of multiple metabolic and regulatory processes. Cell-wall composition is indeed a highly complex trait that possibly can be dissected as quantitative trait loci (QTL). QTLs for cell-wall properties, such as cell-wall digestibility, lignin content, and sugar composition, have been characterized in several plant species^{28–30}. *CSLF6* is the first cloned cell-wall QTL⁶. As characterization of cell-wall properties is laborious, the wall-related genes identified by QTL mapping are very limited.

Nevertheless, QTL cloning is a powerful approach for identifying major genes for agronomic complex traits, such as grain yield. Rice (*Oryza sativa* L.) is a major staple food crop feeding about three billion people, and the most important trait in rice production is grain yield. *Gn1a*, a major QTL controlling grain number, was cloned based on natural allelic variations³¹. *DEP1*, the locus that governs the panicle architecture and grain number was isolated by QTL positional cloning³². Grain size and grain weight are also key determinants for grain yield. Till now, some QTLs for grain size (GS) and grain weight (GW) control have been cloned. For example, *GS3* encodes a transmembrane protein, *GW2* encodes an E3 ligase, *GW8* encodes SPL16, and *Grain length3.1* (*GL3.1*) is a phosphatase, suggesting that cell-cycle or cell-size regulators are required for grain yield control^{33–36}. Plant height (PH) and heading date (HD) are other determinants of grain yield. The most famous gene for plant height control is *sd1*, which greatly enhanced grain productivity by improving lodging resistance in the 1960s³⁷. But the first cloned QTLs for plant height and heading date are *PH1* and *HD1*^{38,39}, respectively. Moreover, these agronomic traits are often correlated. For example, *GS2/GL2* that controls grain size also affects grain weight^{40,41}. *Ghd7*, a QTL for HD, controls plant height and grain number^{42,43}. In addition to these major loci, numerous minor loci also contribute to grain yield, which might be more useful in genetic breeding. Due to the small effects on genetic variability, the connection between genotype and phenotype is hard to bridge, which impedes cloning of the minor QTLs.

Metabolites that act as either causes or effects of complex traits may build a bridge between genotype and phenotype⁴⁴. As the technical advances progress on high throughput examination of metabolites, comprehensive metabolic profiling and large-scale gene characterization by association mapping have been achieved^{45,46}. Discovery of genetic links between metabolic and phenotypic variation becomes implementable. Correlation analysis of fruit metabolic loci and morphological genes for yield-associated traits in tomato introgression lines has revealed the links between nutritional metabolites and organoleptic traits⁴⁷. By using the same approach, lignin precursors identified by genome wide association study (GWAS) are correlated with plant height⁴⁵. A major locus controlling the trigonelline level identified by metabolic GWAS in 489 rice varieties is identical to a QTL for grain width⁴⁸. Recently, two loci (*Lin5* and *SSC11.1*) regulating sugar content and tomato flavor were found negatively correlated to fruit size⁴⁹. Therefore, discovery of the metabolite-phenotype linkage is a useful approach for interactive functional genomics, especially for cloning minor QTLs for complex traits.

Cell wall is the central element of biomass. It is well known that extended heading date and increased plant height lead to more biomass and consequently promote grain yields. Hence, cell wall synthesis may be a valuable metabolic marker associated with grain yields. However, genetic links between cell-wall composition and grain yield-related traits are unclear.

Here we report on generation of a chromosome segment substitution lines (CSSL) by using Nipponbare (NP), a representative cultivar for *japonica*, as recipient parent and 9311, a representative cultivar for *indica*, as donor parent. Based on the variations between the two rice subspecies *indica* and *japonica*, parallel studies on six grain yield-related traits and cell-wall composition were performed in the CSSLs, and the corresponded QTLs for these traits were mapped. The overlapped QTLs for the two sets of traits have been identified. Fine-mapping of a QTL for cellulose content provided further evidence for the genetic connection between both sets of traits. Our study thus sheds lights on dissecting the machinery on cell-wall biosynthesis and grain-yield improvement in rice.

Results

Analysis of phenotypic variations between NP and 9311. Among the morphological and physiological differences between NP and 9311, biomass abundance is the fundamental one. *Indica* variety 9311 generally has more biomass than *japonica* variety NP, which is in agreement with that 9311 has increased heading date, indicating as a longer vegetative growth period, increased plant height, larger panicles, and more grains than NP (Fig. 1A and B). Another morphological difference between 9311 and NP is the grain shape. Variety 9311 has long and slender grains whereas those of NP are short and wide (Fig. 1C and D). Hence, the ratio of grain length to width is significantly increased in 9311, resulting in increased thousand-grain weight compared to NP (Fig. 1D). To address this difference at a cellular level, we investigated cell shape on glumes by scanning electronic microscopy. Surprisingly, cells in 9311 glumes were shorter in length but wider than those in NP (Fig. 1E and F). Quantification of cell numbers in glumes showed that 9311 had more cells at the longitudinal direction, but fewer cells at the latitudinal direction (Fig. 1F). Therefore, all these differences are likely derived from varied ability of biomass productivity between NP and 9311.

NP and 9311 have different cell-wall composition. As cell walls represent the major element of plant biomass, we investigated whether variations on cell-wall composition exist in the two varieties. Because internodes constitute a central source of rice biomass and contribute directly to plant height and grain traits by linking source and sink, internodes were subjected to cell-wall composition examination. The content of cellulose and xylose, the two major components of cell wall, was significantly increased in 9311 (Fig. 2A). In contrast, galactose

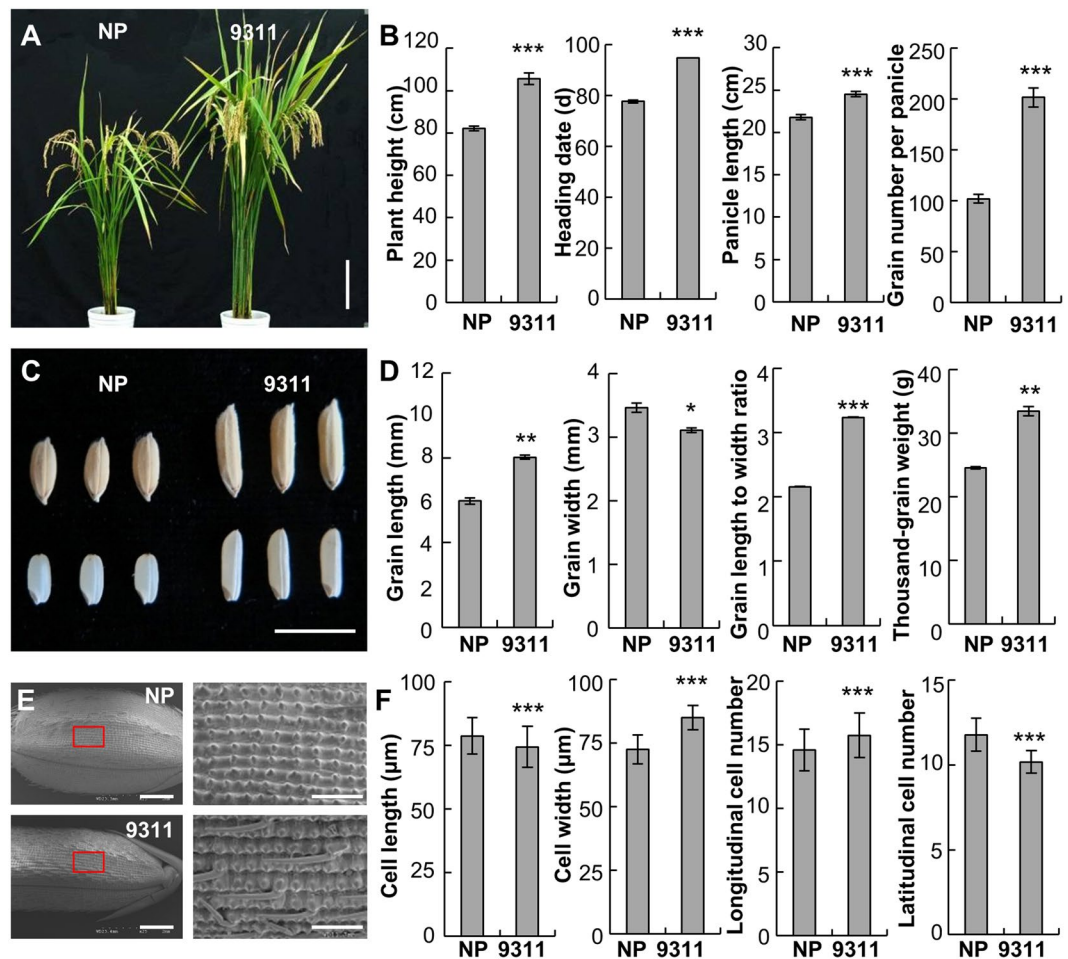


Figure 1. Phenotypic difference between NP and 9311. (A) The mature plants of NP and 9311. (B) Quantification of the morphological phenotypes in NP and 9311. Error bars represent the mean value \pm SD ($n \geq 9$ plants). (C) Grain shape of NP and 9311. (D) Measurement of the grain size and grain weight of NP and 9311. Error bars represent the mean value \pm SD ($n \geq 50$ grains harvested from at least 10 plants). Error bars of thousand-grain weight is the mean value \pm SD ($n = 3$ replicates). (E) Scanning electron microscopic photographs of the outer surface of a glume. The right panels are the enlarged view circled by the rectangles. (F) Quantification of cell length, cell width, and cell number at the longitudinal and latitudinal direction in glume. Error bars represent the mean value \pm SD ($n \geq 500$ cells from 12 seeds). *** $P < 0.001$, ** $P < 0.01$, and * $P < 0.05$ by two-tailed t -test. Bar = 15 cm in (A), 1 cm in (C), and 1 mm for left panel and 250 μ m for right panel in (E).

and glucose that are often present in hemicellulose or amorphous cellulose were reduced in 9311 (Fig. 2A). As arabinosyl residues (Ara) present as side chains on xylans in rice, the ratio of xylose to arabinose (XA) reflects the xylan substitution level. Although the arabinose content was not significantly changed, more xylose content caused higher XA ratios in 9311 (Fig. 2B), indicating that 9311 has lower arabinosyl substitution on xylan compared to NP. Due to a very low level of pectin presenting in rice cell walls, the pectic sugars, rhamnose and fucose, were hard to be detected (Fig. 2C). Therefore, NP and 9311 have different cell-wall composition in the internodes.

Construction of CSSLs. To identify QTL for cell-wall composition and grain yield, we constructed a CSSL population by backcrossing the previously reported 57 CSSLs⁵⁰ with NP and followed by self-crossing. With the aid of molecular marker-assisted selection, a total of 351 CSSLs were obtained. These CSSLs were further genotyped by 357 polymorphic molecular markers that were evenly distributed across the 12 chromosomes. Subsequently, 138 core lines that may cover the whole 9311 genome were selected and subjected to re-sequencing.

Consequently 125 lines that had convincing re-sequencing data were used to construct a graphical genotypic map. As shown in Supplemental Fig. 1, each CSSL contains one to seven substituted segments. In more detail, 53 lines harbored only one substituted segment, 44 lines contained two introgressed segments, and 28 lines had more than two segments (Supplemental Fig. 1). The total length of the 353 substituted segments was around 1999.76 Mb, which is five times of the 9311 genome. Re-sequencing analysis revealed that the length of individual substituted segments ranged from 0.02 to 24.6 Mb, with an average length of 5.6 Mb. Bin mapping showed that five chromosomes, Chr. 2, 3, 6, 7 and 9, were fully covered by the substituted segments, whereas the remaining

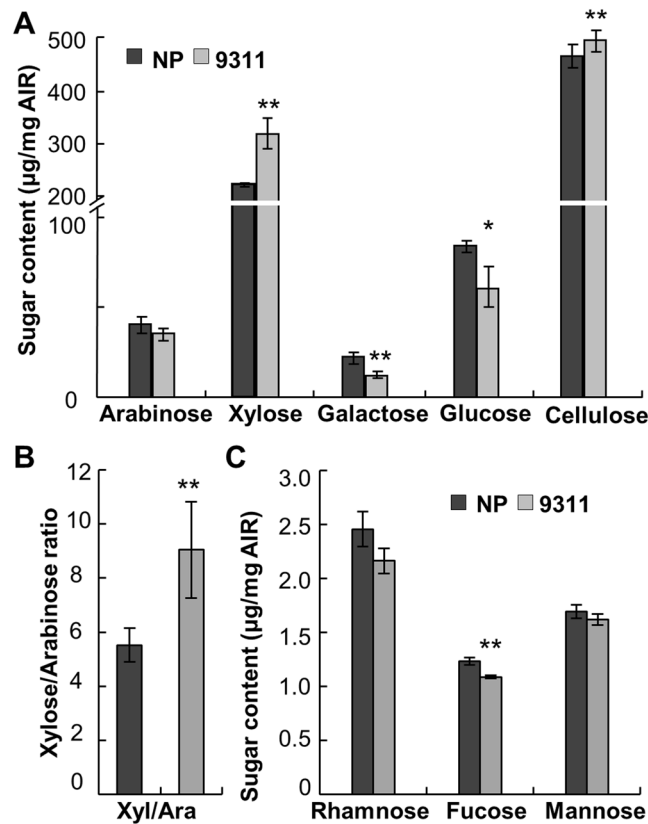


Figure 2. Measurement of the cell-wall composition in NP and 9311. (A) The content of major neutral sugars and crystalline cellulose in wall residues from internodes. (B) The ratio of xylose and arabinose content. (C) The content of minor neutral sugars in wall residues from internodes. Error bars represent the means \pm SD ($n = 4$ replicates using AIR from 10 mature plants). ** $P < 0.01$ and * $P < 0.05$ by two-tailed t -test.

chromosomes had 81.5% to 98.5% coverage. The uncovered regions, shown as seven small gaps, are about 4.4% of the whole 9311 genome. Therefore, this CSSL population covers \sim 95.6% of the 9311 genome, which is suitable for genome-wide surveys.

The CSSLs display a wide range of variation in grain yield-related traits and cell-wall composition.

We investigated six grain yield-related traits, including PH, HD, and grain size/weight in the CSSLs and the two parents. PH, HD, and grain length to width ratio (GLWR) were continuous within the range of two parents, and a few lines were superior to their parents (Supplemental Fig. 2). However, for GL, grain width (GWD), and thousand-grain weight (TGW), more lines were superior to their parents, indicating that loci-interacting effects might exist in these traits (Supplemental Fig. 2). The wide range of phenotypic variation suggested complexity in these agronomic traits.

We also analyzed cell-wall composition of the mature culms from the CSSLs. To guarantee accuracy of examination, we performed variation analysis on the content of four major sugars and cellulose, as that of rhamnose, fucose and mannose is too low to be precisely measured. As shown in Supplemental Fig. 3, the content of the four sugars and cellulose showed a continuous distribution among the CSSLs with transgressive segregation on both sides. The ratio of xylose to arabinose, reflecting arabinoxylan structure, was also continuously distributed. Therefore, cell-wall composition belongs to a complex trait.

Cell-wall composition is correlated to grain yield-related traits.

To determine whether cell-wall composition is correlated with the grain yield-related traits, we performed correlation analysis. The content of cellulose and four sugars and XA ratio showed varied correlation coefficient against six grain yield-related traits (Table 1). PH, one of the most important traits for crop architecture and grain yield, was significantly associated with cellulose content and XA ratio. Similarly, HD that indicates vegetative growth period was associated with XA ratio and the content of cellulose, xylose and galactose (Table 1), demonstrating that biosynthesis of cellulose and xylan is correlated with PH and HD. GWD, GL and TGW are direct determinants of grain yields. Interestingly, XA ratio displayed a high correlation with TGW and GL, while the glucose and galactose content were significantly correlated with TGW and GWD (Table 1). The highest correlation coefficient between cell-wall composition (XA ratio) and grain yield-related traits (TGW) was 0.44, indicating that xylan structure may contribute to TGW and grain yield. As positive controls, 0.85 or 0.6 correlation coefficients were found between the abundance of arabinose and galactose or xylose due to the existence of arabinogalactan and arabinoxylan in rice

| Traits | Ara | Xyl | Gal | Glc | XA | Cel | PH | HD | GL | GWD | GLWR | TGW |
|--------|---------|--------|---------|---------|--------|--------|--------|--------|--------|---------|------|-----|
| Ara | 1 | | | | | | | | | | | |
| Xyl | 0.60** | 1 | | | | | | | | | | |
| Gal | 0.85** | 0.17 | 1 | | | | | | | | | |
| Glc | 0.66** | 0.15 | 0.75** | 1 | | | | | | | | |
| XA | -0.49** | 0.39** | -0.80** | -0.57** | 1 | | | | | | | |
| Cel | 0.01 | 0.20* | -0.15 | -0.13 | 0.28* | 1 | | | | | | |
| PH | -0.05 | 0.14 | -0.20* | -0.18* | 0.23* | 0.34** | 1 | | | | | |
| HD | -0.08 | 0.23* | -0.31** | -0.12 | 0.34** | 0.21** | 0.62** | 1 | | | | |
| GL | -0.08 | 0.16 | -0.18* | -0.19* | 0.27** | 0.08 | 0.33** | 0.28** | 1 | | | |
| GWD | -0.22* | -0.1 | -0.22** | -0.28** | 0.18 | 0.13 | 0.22** | 0.16 | 0.42** | 1 | | |
| GLWR | 0.15 | 0.23* | 0.08 | 0.11 | 0.06 | -0.07 | 0.07 | 0.12 | 0.45** | -0.61** | 1 | |
| TGW | -0.16 | 0.21* | -0.38** | -0.31** | 0.44** | 0.06 | 0.51** | 0.58** | 0.63** | 0.42** | 0.16 | 1 |

Table 1. Correlation coefficient between cell wall components and grain yield-related traits. * $P < 0.05$ and ** $P < 0.01$ by student t test (two tailed). Ara, Arabinose; Xyl, Xylose; Gal, Galactose; Glc, Glucose; XA, Xylose/Arabinose; Cel, Cellulose; PH, Plant height; HD, Heading date; GL, Grain length; GWD, Grain width; GLWR, Grain length-to-width ratio; TGW, Thousand-grain weight.

cell walls. Similarly, around 0.6 correlation coefficients were showed between PH and HD and between GL and TGW (Table 1), which is in agreement with previous reports⁵¹.

Taken together, the cell-wall compositions are correlated with grain yield-related traits in rice.

QTL analysis for grain yield-related traits and cell-wall composition. QTL mapping for the grain yield-related traits and cell-wall composition was carried out based on Bin-map converted from physical map⁵². We found 56 QTLs in total. In detail, nine QTLs for PH were found located on five chromosomes with positive and negative additive effects. The contribution from 9311 indicates a positive additive effect, and that from NP is negative. The variance explained by these QTLs varied from 4.7% to 31.1%. *qPH3-1* and *qPH8* with positive effects were major loci and located on Chr. 3 and Chr. 8, whilst *qPH1-2* with negative effect was located on Chr. 1 (Table 2). We further found that *qPH1-2* and *qPH8* included the known important loci for plant height, such as *sd1*³⁷ and *Ghd8/DTH8*^{53,54}, suggesting that the mapping data are reliable. Among eleven QTLs for HD, three major loci were mapped on Chr.3 and Chr.8, which explained 17.72% to 30.81% phenotypic variance (Table 2). The 30.81% variation is likely due to the presence of *DTH8/Ghd8*^{53,54}, a major QTL for heading date, in the mapping region. The remaining QTLs had minor effects, in which *qHD6-1* included *Hd3a*⁵⁵ and *Hd17*⁵⁶, and *qHD6-2* included *Hd1*³⁸ (Table 2). For grain size or weight control, many minor QTLs with 6–11% explanation were found distributed on seven chromosomes (Table 2). Meanwhile, six major QTLs for GL, TGW, and GLWR located on Chr.2, 3 and 5 and explained 15.31% to 48.9% variance were also identified (Table 2). Within the mapping regions of *qGL3-2* and *qGWD5*, *GS3*, *GL3.2* and *GW5*, the known QTLs for grain size or weight control^{36,57,58}, are located.

Using the same approach, a total of nineteen QTLs for the content of cellulose and four major neutral sugars and XA ratio were identified on nine chromosomes. Probably because cellulose biosynthesis is modulated by multiple genes with small effects, only one QTL for cellulose content (*qCel6*) with positive additive effect was mapped and explained 10.9% phenotypic variance (Table 3). Three QTLs for xylose content were detected, of which *qXyl7* and *qXyl10* were of negative effect and concentrated on Chr. 7 and 10, whilst *qXyl8* that had positive effect was located on Chr. 8. The variance explained by these QTLs varied from ~8% to 13.6% (Table 3). Four QTLs for the arabinose content showing positive effect were mapped on Chr. 4, 7 and 8 with around 6.4%–12.5% variance (Table 3). Similarly, four QTLs for XA ratio were targeted onto Chr. 7 and 8, in which *qXA7-1* and *qXA7-2* are major QTLs as they explained 59.2% and 17.7% phenotypic variance (Table 3). Within five QTLs for the glucose content, *qGlc1-1* and *qGlc11* explained 17.7% and 15.6% phenotypic variance (Table 3). Finally, two QTLs for the galactose content were mapped on Chr. 4 and 8, respectively (Table 3).

By using this CSSL population, QTLs for six grain yield-related traits and cell-wall composition have been parallelly mapped on eleven rice chromosomes with the exception of Chr. 9.

QTLs for cell-wall composition and grain yield are overlapped. Physical mapping of the above loci further showed nicely co-localized QTLs. Firstly, several grain yield-related QTLs were co-localized. QTLs for PH and GLWR overlapped on Chr. 1 (Supplemental Fig. 4). The overlapping could extend to PH and HD on Chr. 3, 6 and 8, and to HD and TGW on Chr. 8, suggesting that plant height and heading date are correlated to grain yield. Within the nineteen QTLs for cell-wall composition, QTLs for the arabinose content were co-localized with XA ratio on Chr. 7 and 8, and with the galactose and xylose content on Chr. 8 (Supplemental Fig. 5), in agreement with the existence of arabinoxylan and arabinogalactan in the rice cell-wall polymers.

Interestingly, QTLs for the cellulose (*qCel6*), glucose (*qGlc1-2*), and galactose (*qGal8*) content, and XA ratio (*qXA7-2*) were overlapped with those for PH (*qPH6-1* and *qPH1-1*), TGW (*qTGW8*) and HD (*qHD6-1* and *qHD7*) (Fig. 3). *qAra7* and *qXA7-1* were adjacent to *qGL7* and *qTGW7*. Similar relationship was found between *qAra8-1* and *qPH8*. This finding indicated that certain genes required for cell-wall biosynthesis may affect agronomic traits. To explore evidence underlying the connection, we investigated the genotypes and phenotypes in the corresponding CSSLs. *qCel6* and *qPH6-1/qHD6-1* were co-targeted within a 2.46-Mb region on Chr. 6. Five substitution lines containing and not containing the 9311 fragment showed consistent tendency in cellulose

| QTLs | Chr. | Bin ^a | LOD | PVE (%) | Add ^b | Position (Mb) |
|-----------------------------|------|------------------|-------|---------|------------------|---------------|
| Plant height | | | | | | |
| <i>qPH1-1</i> | 1 | Bin12 | 9.07 | 8.7 | 11.53 | 24.73–27.63 |
| <i>qPH1-2</i> | 1 | Bin17 | 7.71 | 7.21 | −10.5 | 37.95–40.68 |
| <i>qPH2-1</i> | 2 | Bin20 | 10.05 | 9.82 | 12.25 | 1.1–1.3 |
| <i>qPH2-2</i> | 2 | Bin52 | 6.61 | 6.05 | 3.72 | 28.32–32.53 |
| <i>qPH3-1</i> | 3 | Bin74 | 16.81 | 18.72 | 6.55 | 6.92–10.56 |
| <i>qPH3-2</i> | 3 | Bin97 | 9.18 | 10.19 | 6.24 | 34.98–36.40 |
| <i>qPH6-1</i> | 6 | Bin140 | 7.15 | 6.61 | 4.19 | 1.31–3.77 |
| <i>qPH6-2</i> | 6 | Bin173 | 5.26 | 4.7 | 3.5 | 25.89–26.83 |
| <i>qPH8</i> | 8 | Bin238 | 24.11 | 31.11 | 7.51 | 4.11–4.51 |
| Heading date | | | | | | |
| <i>qHD2</i> | 2 | Bin49 | 7.45 | 1.95 | 1.78 | 26.26–27.2 |
| <i>qHD3-1</i> | 3 | Bin74 | 23.4 | 8.36 | 2.82 | 6.92–10.56 |
| <i>qHD3-2</i> | 3 | Bin97 | 36.62 | 17.72 | 10.49 | 34.98–36.40 |
| <i>qHD5</i> | 5 | Bin130 | 13.45 | 6.87 | 8.17 | 23.29–23.38 |
| <i>qHD6-1</i> | 6 | Bin140 | 20.92 | 7.11 | 2.6 | 1.31–3.77 |
| <i>qHD6-2</i> | 6 | Bin161 | 3.4 | 0.83 | −0.95 | 7.88–9.67 |
| <i>qHD7</i> | 7 | Bin230 | 6.43 | 4.63 | 4.65 | 28.94–29.69 |
| <i>qHD8-1</i> | 8 | Bin234 | 33.81 | 15.78 | 10.61 | 2.18–2.43 |
| <i>qHD8-2</i> | 8 | Bin240 | 41.91 | 30.81 | 15.01 | 4.57–5.36 |
| <i>qHD11</i> | 11 | Bin340 | 13.37 | 4.16 | 3.88 | 20.02–20.03 |
| <i>qHD12</i> | 12 | Bin361 | 3.36 | 0.19 | 1.52 | 6.5–6.78 |
| Grain length | | | | | | |
| <i>qGL3-1</i> | 3 | Bin74 | 4.83 | 17.19 | 0.19 | 6.92–10.56 |
| <i>qGL3-2</i> | 3 | Bin79 | 4.69 | 16.71 | 0.39 | 15.43–17.74 |
| <i>qGL7-1</i> | 7 | Bin212 | 2.55 | 9.48 | 0.21 | 19.36–19.40 |
| <i>qGL7-2</i> | 7 | Bin216 | 2.89 | 10.64 | 0.22 | 21.21–22.67 |
| Grain width | | | | | | |
| <i>qGWD2-1</i> | 2 | Bin35 | 3.88 | 10.89 | −0.22 | 18.92–19.98 |
| <i>qGWD2-2</i> | 2 | Bin55 | 2.84 | 7.83 | −0.15 | 33.75–34.46 |
| <i>qGWD5</i> | 5 | Bin118 | 2.93 | 8.07 | −0.23 | 2.07–5.78 |
| <i>qGWD6</i> | 6 | Bin164 | 2.69 | 7.37 | −0.12 | 11.02–11.46 |
| Grain length to width ratio | | | | | | |
| <i>qGLWR1</i> | 1 | Bin17 | 2.71 | 10.05 | 0.21 | 37.95–40.68 |
| <i>qGLWR2-1</i> | 2 | Bin18 | 5.6 | 9.36 | 0.12 | 0–1.1 |
| <i>qGLWR2-2</i> | 2 | Bin35 | 21.06 | 48.89 | 0.27 | 18.92–19.98 |
| <i>qGLWR3</i> | 3 | Bin79 | 6.45 | 10.97 | 0.1 | 15.43–17.74 |
| <i>qGLWR5</i> | 5 | Bin118 | 8.61 | 15.31 | 0.19 | 2.07–5.78 |
| <i>qGLWR6</i> | 6 | Bin169 | 6.69 | 11.44 | 0.12 | 21.55–21.82 |
| Thousand-grain weight | | | | | | |
| <i>qTGW3</i> | 3 | Bin79 | 4.69 | 16.71 | 0.39 | 15.43–17.74 |
| <i>qTGW7</i> | 7 | Bin216 | 2.89 | 10.64 | 0.22 | 21.21–22.67 |
| <i>qTGW8</i> | 8 | Bin234 | 3.69 | 6.36 | −0.6 | 2.18–2.43 |

Table 2. Identification of QTLs for grain yield-related traits. ^aBin overlapping with the LOD peak of the QTL. ^bThe allelic effect is calculated as the mean effect of replacing NP alleles by 9311 alleles at the QTL. LOD, Logarithm of the odds ratio; PVE, Percentage of the trait variance explained by the QTL.

content and plant height. In detail, the CSSLs possessing the 9311 fragment had increased cellulose content and plant height (Fig. 4A). *qXA7-2*, a major QTL for XA ratio, was mapped at the end of Chr. 7, where *qHD7* is located. Line 88 (N88) containing the 9311 fragment displayed higher XA ratio and longer heading date than NP and N57 that do not harbor the substitution fragment (Fig. 4B). Within the mapping region on Chr. 8, QTLs for thousand-grain weight and *qGal8* were collocated. The substitution lines harboring the 9311 fragment showed increased galactose content and slightly reduced thousand-grain weight (Fig. 4C). Taken together our data suggest that these regions are very likely to harbor QTLs synchronously controlling cell-wall biosynthesis and grain yield.

Fine-mapping of *qCel6*. To find a further support for this conclusion, we generated a mapping population by backcrossing N62, a line containing a 9311 fragment and contributing to high cellulose level, with NP. Based

| Traits | QTLs | Chr. | Bin ^a | LOD | PVE (%) | Add ^b | Position (Mb) |
|------------------------|----------------|------|------------------|-------|---------|------------------|---------------|
| Cellulose | <i>qCel6</i> | 6 | Bin140 | 3.11 | 10.90 | 14.29 | 1.31–3.77 |
| Xylose | <i>qXyl7</i> | 7 | Bin188 | 2.60 | 7.97 | −27.23 | 2.82–2.85 |
| | <i>qXyl8</i> | 8 | Bin258 | 3.02 | 9.65 | 37.95 | 23.01–24.22 |
| | <i>qXyl10</i> | 10 | Bin315 | 4.29 | 13.57 | −49.61 | 14.7–19.14 |
| Arabinose | <i>qAra4</i> | 4 | Bin114 | 2.51 | 6.44 | 4.26 | 33.31–33.41 |
| | <i>qAra7</i> | 7 | Bin226 | 4.67 | 12.47 | 4.25 | 26.32–26.42 |
| | <i>qAra8-1</i> | 8 | Bin234 | 3.82 | 10.04 | 3.33 | 2.18–2.43 |
| | <i>qAra8-2</i> | 8 | Bin261 | 3.64 | 9.54 | 8.91 | 24.64–25.59 |
| Xylose/Arabinose ratio | <i>qXA7-1</i> | 7 | Bin226 | 17.98 | 59.22 | 0.05 | 26.32–26.42 |
| | <i>qXA7-2</i> | 7 | Bin230 | 6.76 | 17.75 | 0.03 | 28.94–29.69 |
| | <i>qXA8-1</i> | 8 | Bin234 | 3.68 | 9.12 | 0.02 | 2.18–2.43 |
| | <i>qXA8-2</i> | 8 | Bin263 | 3.08 | 7.54 | 0.03 | 25.7–26.11 |
| Glucose | <i>qGlc1-1</i> | 1 | Bin8 | 5.29 | 17.72 | 41.60 | 9.44–11.05 |
| | <i>qGlc1-2</i> | 1 | Bin12 | 2.81 | 6.35 | 18.86 | 24.73–27.63 |
| | <i>qGlc2</i> | 2 | Bin44 | 2.55 | 5.75 | 8.16 | 22.86–24.67 |
| | <i>qGlc3</i> | 3 | Bin95 | 3.47 | 7.96 | 12.30 | 31.42–34.98 |
| | <i>qGlc11</i> | 11 | Bin334 | 4.5 | 15.6 | 15.0 | 18.13–18.17 |
| Galactose | <i>qGal4</i> | 4 | Bin105 | 2.81 | 8.64 | 7.23 | 18.60–22.76 |
| | <i>qGal8</i> | 8 | Bin234 | 4.40 | 13.96 | 3.35 | 2.18–2.43 |

Table 3. Identification of QTLs for cell-wall composition. ^aBin overlapping with the LOD peak of the QTL. ^bThe allelic effect is calculated as the mean effect of replacing NP alleles by 9311 alleles at the QTL. LOD, Logarithm of the odds ratio; PVE, Percentage of the trait variance explained by the QTL.

on genotyping of more than 3500 BC₈F₂ plants and examining the cellulose content in the internodes of plants that possess the corresponding genotypes. *qCel6* was narrowed down to a ~800 kb region, in which *HD3a*, *HD3b* and *HD17* controlling heading date and grain yields are located^{55,56} (Fig. 5A and B). Natural variations on these genes between NP and 9311 or the substitution lines were investigated by DNA sequencing. Surprisingly, we have not found any variations in the coding and promoter regions of *HD3a* and *HD3b*. Seven single nucleotide polymorphisms (SNPs) that cause five amino acid substitutions in *HD17* have been identified (Fig. 5C), indicating that the increased cellulose content is likely due to variations on *HD17*. Co-segregation of *qCel6* and *HD17* in the fine-mapping assay provided further evidence for the genetic correlation between cell-wall composition and an agronomic trait.

Discussion

CSSLs that contain one or a few donor segments in each line are widely used for QTL mapping. Since the CSSLs possess the same genetics background, QTL mapping with this population is of high accuracy and could be dissected as single Mendelian factors⁵⁹ although preparation of CSSLs is time-consuming. Here we generated CSSLs using NP as receptor and 9311 as donor. Both parents are representative varieties for two rice subspecies and show a great genetic variance suitable for QTL association analysis. Rice is one of Earth's major staple crops. Improvement of grain yield is the most important issue in rice production. Due to the complexity of grain yield control, a great number of genes have been inferred to be involved. However, although some major QTLs for grain size/weight have been identified^{33–36,60}, the cloned loci with smaller effects are very limited. Recent works have revealed that metabotype-phenotype links are powerful to clone minor QTLs for complex phenotypic traits^{44,45,47}. As cell walls are regarded as basic materials for plant growth, relevant metabolites could be annotated as causes or effects². Therefore, combining QTL mapping for two sets of traits may provide a comprehensive view to understand the mechanism of grain yield control and assigning the biological functions to cell-wall synthesizing genes.

In this study, six phenotypic traits, named as grain yield-related traits, were investigated in the CSSLs. Plant height and heading date are widely studied traits and associated with crop yield^{37,53}. We identified nine QTLs for PH and eleven QTLs for HD, which not only included several previously identified QTLs, such as *SD1*³⁷, *HD1*³⁸, *HD3a*⁵⁵, *HD17*⁵⁶, and *DTH8/GHD8*^{53,54}, according to the mapping region of *qPH1-2*, *qHD6-1*, *qHD6-2*, and *qHD8-1/qPH8*, but also contained several new loci, such as *qPH1-1*, *qHD3-2* and *qHD5*. Similarly, QTLs for grain size and grain weight were investigated, which identified some known QTLs. For example, *qGL3-2/qGLWR3* included *GS3*³⁶ and *GL3.2*⁵⁷, and *qGWD5* contained *GW5*⁵⁸. *qGL7-2* and *qGWD2-1* were adjacent to the previously reported *GL7* and *GW2*, respectively^{33,61}, indicating that our mapping results are reliable. Notably, several unidentified loci were also found, such as *qGL3-1* and *qTGW7*. Although the mapping population was derived from the 57 CSSLs, three previously reported minor QTLs for grain weight⁵⁰ have not been found here, suggesting that mapping minor loci for complex traits is of great challenge.

We parallelly analyzed cell-wall composition in the CSSLs. Consequently, nineteen QTLs for six cell-wall content characteristics were identified. Only one QTL for the cellulose content, *qCel6* that explained 10.9% of phenotypic variation, was found, indicating that cellulose property is affected by many minor loci, consistent with the notion that cellulose synthesis is very complicated and required the involvement of multiple genes^{2–4}.

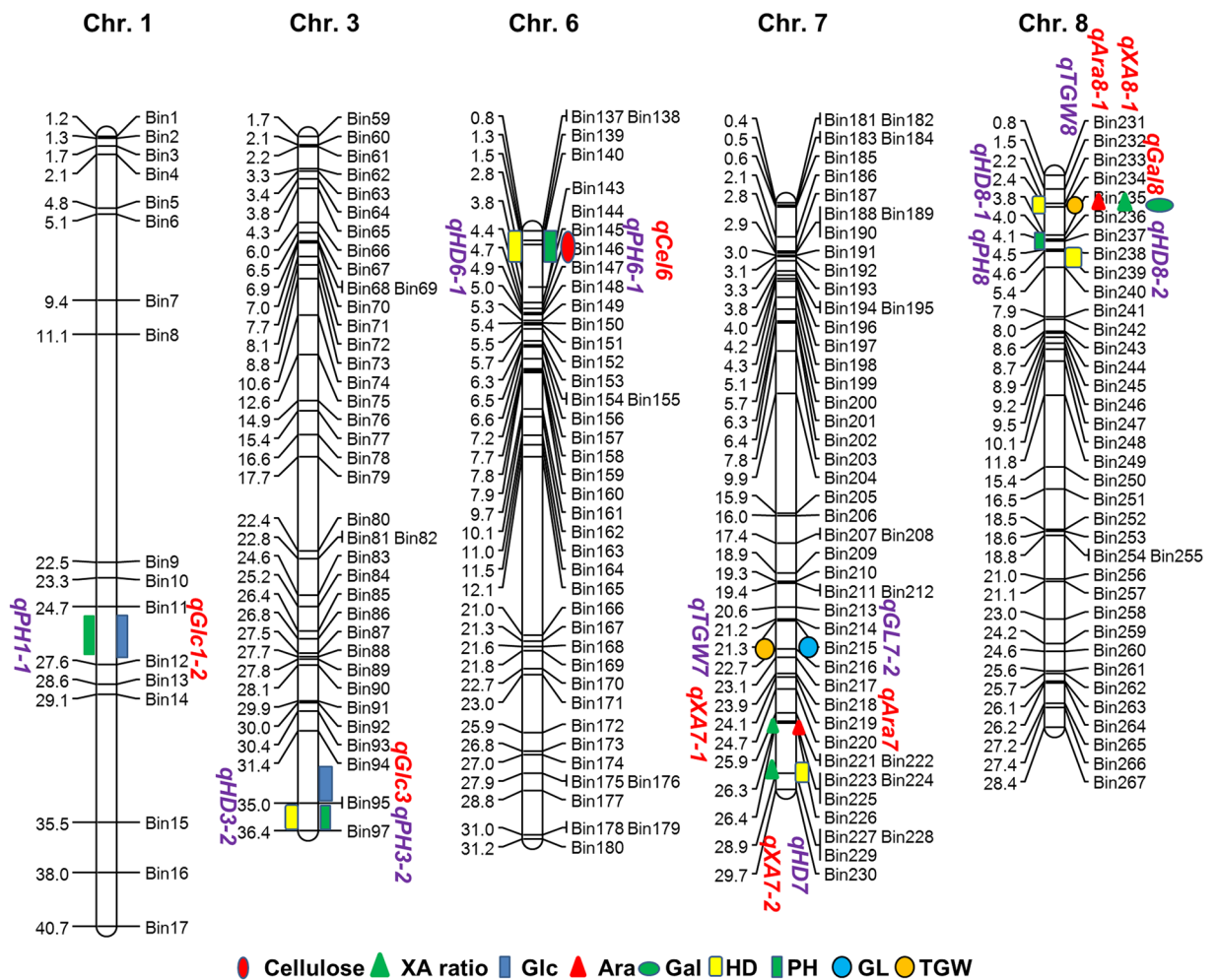


Figure 3. Chromosomal locations of the overlapped and adjacent QTLs for the cell-wall composition and grain yield-related traits. Different symbols indicate the positions where the corresponding QTLs are located. The numbers at the left side of chromosomes indicate the physical locations (Mb).

Interestingly, *qCel6* is very likely to be an uncharacterized cell-wall synthesis related gene, because within the mapping region on Chr. 6, none of known cellulose synthesizing or regulatory genes has been reported. Arabinoxylan is the second most abundant component next to cellulose in rice cell wall. Xylose content and XA ratio are the most important parameters reflecting arabinoxylan abundance and structure. In this study, we identified three QTLs for xylose content and four QTLs for XA ratio. Previous genetic evidence demonstrated that GT43 and GT47 family members are xylan synthases^{8,62,63}. However, we did not find any GT43 and GT47 members located in, or adjacent to, the xylose QTLs, suggesting that these QTLs may represent unknown genes influencing xylan biosynthesis. It is worth to point out that the mapping region of *qXyl10* includes *CESA7*, a key gene for secondary wall cellulose synthesis, in agreement with the elucidated correlation relationship between cellulose and xylan biosynthesis⁸. *OsUAM3*, encoding a putative uridine diphosphate-arabinose mutase, was adjacent to *qXA7-1*, a major QTL for XA ratio that explained up to 59.22% phenotype variation. *OsUAM3* may be a candidate for this locus. Moreover, the nineteen QTLs identified here were different from the previously reported cell-wall QTLs mapped by a CSSL population containing a common wild rice, Yuanj introgressed segment(s) in an *indica* cultivar, Teqing³⁰. The distinction not only indicated the complexity of cell wall biogenesis, but also prompted us to investigate the biological functions behind these genetic clues.

Correlation analysis between different traits has been widely used to ascertain a QTL with pleiotropic functions. Several studies have reported positive correlation between PH and HD, and between GL and TGW/HD^{40,41,43,51}. In this study, we found that the mapping regions for *qPH1-2* and *qGLW1* and for *qHD8-1* and *qTGW8* were overlapped, in agreement with the previous reports that *Ghd7* and *DTH8/Ghd8* are a pleiotropic QTL, which simultaneously regulates flowering time, plant height and grain productivity^{42,54}. Bridging genetic links across metabolites and phenotypic traits by paralleled QTL mapping is effective for annotation of the biological function of a metabolite, and for cloning QTLs with minor effects⁴⁴. Through establishing metabolic profile and GWAS analysis in rice, maize and tomato, several metabolic loci controlling lignin content, trigonelline level, and sugar amount have been cloned and assigned with biological functions, based on the impacts on plant height⁴⁵, grain width⁴⁸ and fruit size⁴⁹. In this study, we adopted the same strategy to determine the genetic

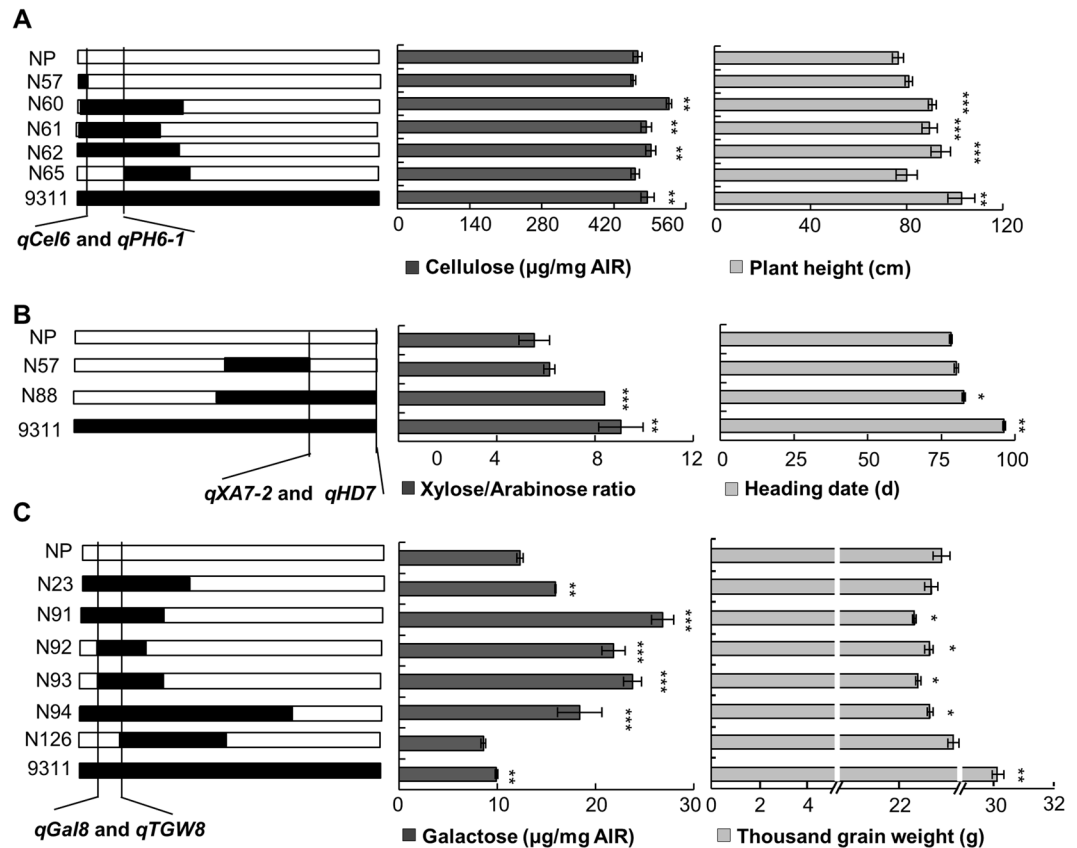


Figure 4. Verification of the overlapped QTLs for cell-wall composition and grain yield-related traits by the CSSLs. (A) The co-localized QTLs on Chr. 6 and examining the cellulose content and plant height in the CSSLs. (B) The co-localized QTLs on Chr. 7 and examining the XA ratio and heading date in the CSSLs. (C) The co-localized QTLs on Chr. 8 and examining the galactose content and thousand-grain weight in the CSSLs. The black and white rectangles in this figure represent 9311 and NP genomic region, respectively. *** $P < 0.001$, ** $P < 0.01$, and * $P < 0.05$ by two-tailed t -test.

connection between cell-wall biosynthesis and grain yield-related traits. Correlation coefficient analysis revealed that all the observed cell-wall composition showed varied degrees of correlation with grain yield-related traits. The cellulose and galactose content was correlated with PH and HD and with TGW and GWD, respectively. Meanwhile, XA ratio was associated with PH and HD, as well as with GL and TGW. The combined QTL analysis revealed several overlapped loci for both sets of traits, providing genetic clues to the association. The inferred genetic links were strengthened by exploring the correlated phenotypes between cell-wall composition (*qCel6*, *qXA7-2* and *qGal8*) and plant height (*qPH6-1*), heading date (*qHD7*) or grain weight (*qTGW8*) in the corresponding CSSLs. Fine-mapping assay narrowed *qCel6* down together with *HD3a*, *HD3b* and *HD17*, QTLs for heading date^{55,56}, within ~800 kb region. Investigation of allelic variations in these genes found seven SNPs in *HD17* between 9311 and NP, which introduce five nonsynonymous mutations, indicating *HD17* as a candidate of *qCel6*. Further studies, such as QTL cloning and transformation verification, are required to ensure the authentic effect of *HD17* on cellulose content. Our study thus provides genetic connection for both sets of traits, which may benefit for understanding the molecular basis of grain yield control, and for assigning biological functions to cell wall synthesizing genes.

Methods

Plant materials. To develop the CSSLs, the previously reported 57 CSSLs⁵⁰ generated by using NP as recipient and 9311 as donor were subjected to a series of backcross with NP and combined with self-crosses. In total, 351 CSSLs were obtained. After genotyping with molecular markers, 138 core lines were chosen and subjected to re-sequencing. Based on the resequencing data, 125 CSSLs that cover 95.6% the 9311 genome were chosen for further phenotyping and QTL mapping.

Field planting. The two parents and 125 lines of CSSLs were grown in the experimental fields located at the campus of Yangzhou University and at the Jiudian County of Yangzhou, Jiangsu Province, China, respectively. N62 and the related backcross population were grown in the experimental fields located at the Jiudian County of Yangzhou, Jiangsu Province and at the Lingshui County of Sanya, Hainan Province, China, respectively. Eighty plants (8×10) of two parents and each CSSL were planted in the plots at 25 cm interval in the horizontal direction and at 15 cm interval in the longitudinal direction.

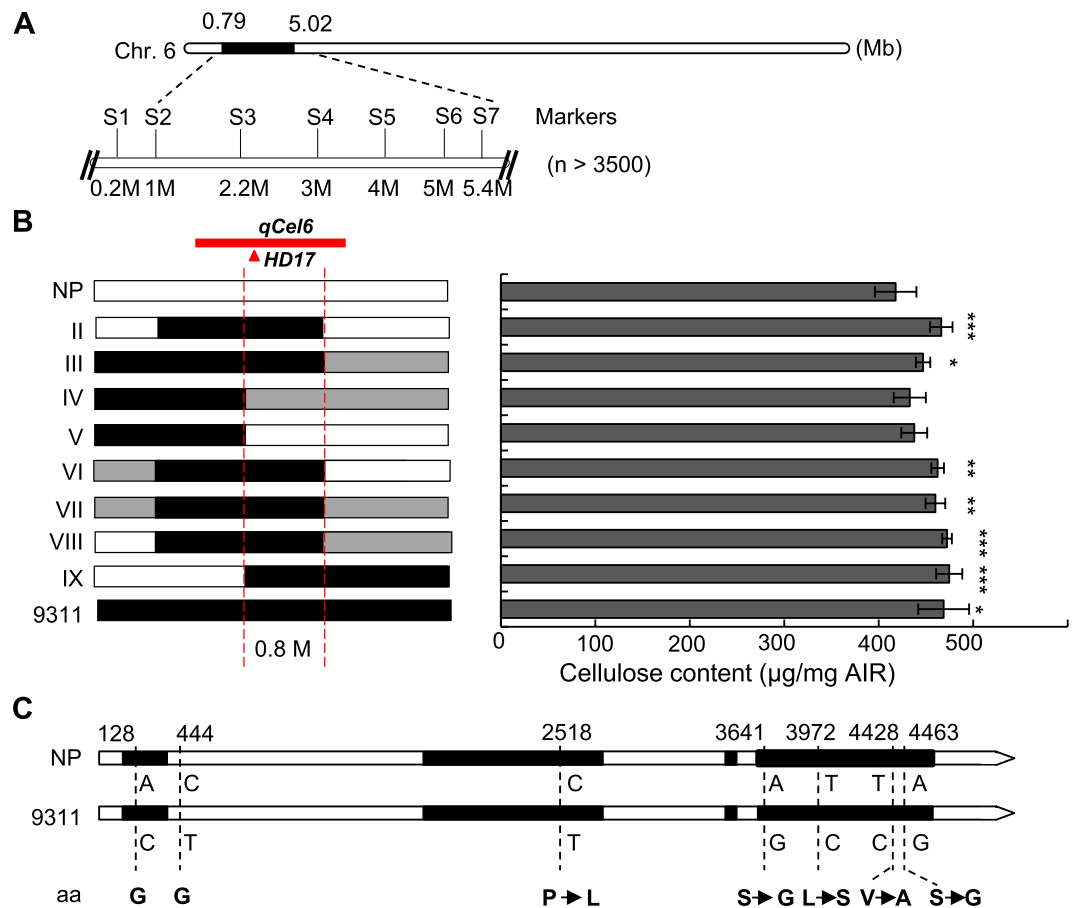


Figure 5. Fine-mapping of *qCel6*. (A) Mapping of *qCel6* in the backcross population. The numbers on chromosome 6 indicate the physical locations (Mb). S1-S7 indicate molecular markers. (B) Genotyping the segregation lines (left) and measuring the cellulose level in the internodes of corresponding lines (right) shown in left. $n = 4$. *** $P < 0.001$, ** $P < 0.01$, and * $P < 0.05$ by two-tailed *t*-test. The black, white and grey rectangles indicate 9311, NP and heterozygous genomic region, respectively. The red bar and red dash lines indicate the QTL mapping and narrowed region of *qCel6*. The red triangle indicates the place *HD17* located. (C) *HD17* gene structure and allelic variations between 9311 and NP. The numbers indicate the place of SNPs located. The arrows indicate the changes of amino acid (aa).

Measurement of grain yield-related traits and data analysis. Heading date of 125 lines and two parents were evaluated by counting the dates from sowing to flowering. Heading date of two parents and each line was recorded when more than 50% individual plants started heading. After maturation, ten representative plants of two parents and each line were subjected to measure the height from ground to the tallest panicle. One month later after seed harvest, seeds derived from 16 individual plants of two parents and each line were pooled and subjected to measure grain size and grain weight. Grain length, grain width, and grain length to width ratio were measured by a camera and analyzed with Smart Grain equipped (Version 1.1). Thousand-grain weight was measured by weighing 200 seeds from two parents and each line for three biological replicates. The data was analyzed using an independent sample *t* test program (SPSS version 15.0). Correlation coefficient between different traits was analyzed by SPSS software (version 15.0).

Scanning electronic microscopy. The outer surface of glumes from mature seeds was sprayed with gold particles and observed with a scanning electron microscope (S-3000N, Hitachi) at an acceleration voltage of 10 kV. At least 500 cells from 12 individual seeds of two parents were measured and quantified.

Cell-wall composition analysis. The cell-wall composition was analyzed as described previously⁶⁴. In brief, the mature 2nd internodes of two parents and CSSLs were ground into powder and treated with 70% ethanol and a mixture of chloroform and methanol (1:1 v/v) twice to prepare alcohol insoluble residues (AIRs). The de-starched AIRs were obtained by treatment with pullulanase M1 (Megazyme) and α -amylase (Sigma) in 0.1 M sodium acetate buffer (pH 5.0) for at least 24 h. After hydrolyzed in 2 M trifluoroacetic acid (TFA) at 121 °C for 90 min, the released sugars were reduced with sodium borohydride (10 mg mL⁻¹ in 1 M ammonium hydroxide). Samples were analyzed by an Agilent 7890 series GC equipped with a 5975 C MS detector (Agilent). Meanwhile, crystalline cellulose content was quantified by anthrone assay⁶⁵ using the remains after TFA treatment. The data were presented as the mean of four biological repeats.

Molecular markers. Molecular markers developed for reconstruction of CSSLs in this study were classified into two types. Rice microsatellite (RM) markers were selected from rice microsatellite maps. The sequence and location information of these markers are available from the Gramene website (<http://www.gramene.org>). Sequence tagged site (STS) markers were developed by using PRIMER PREMIER 5.0 software according to the publicly available rice genome sequence (<http://www.ncbi.nlm.nih.gov/>). The polymorphic markers that were validated in the two parents were used for marker assistant selection.

Resequencing of the CSSLs. Five microgram of genomic DNA from each sample was randomly fragmented by sonication and size-fractionated using electrophoresis. DNA fragments of ~500 bp were then purified. Adapter ligation and DNA clustering were performed according to the procedure from manufacturer. The fragmental samples were sequenced by Illumina HiSeq. 2000 according to the manufacturer instructions (Illumina, San Diego, California, USA). Read sequences were aligned with the reference pseudomolecules (MSU Rice Genome Annotation Project, <http://rice.plantbiology.msu.edu/>) and applied for SNP detection.

QTL mapping. ICIMapping 4.0 software was used to detect the QTLs by the aid of constructing Bin-map in which the genotype of 9311 and NP was assigned as 2 and 0, respectively⁶⁶. The QTL loci were determined by a likelihood ratio test based on single point analysis. QTLs with logarithm of the odds ratio (LOD) larger than 2.5 were considered as real QTLs. The allelic effect was calculated as the mean effect of replacing NP alleles by 9311 alleles at the QTL. The trait value and QTL located chromosome number were indicated. Mapchart (version 2.2) was applied to locate the QTLs on the Bin-map.

Construction of Bin-map. To locate the donor segments in the CSSLs for QTL mapping, Bin-map was constructed based on the above SNP genotyping analyses. The detected SNPs were placed onto the chromosomes according to the physical location. The recombination breakpoints and donor segments were determined using a sliding window approach according to the method described⁵⁰. Finally, 357 Bins were generated. The physical length of bins ranges from 0.01 to 8.92 Mb.

References

1. Cosgrove, D. J. Growth of the plant cell wall. *Nat Rev Mol Cel Bio* **6**, 850–861 (2005).
2. Somerville, C. *et al.* Toward a systems approach to understanding plant cell walls. *Science* **306**, 2206–2211 (2004).
3. Pear, J. R., Kawagoe, Y., Schreckengost, W. E., Delmer, D. P. & Stalker, D. M. Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. *Proc. Natl. Acad. Sci. USA* **93**, 12637–12642 (1996).
4. Arioli, T. *et al.* Molecular analysis of cellulose biosynthesis in Arabidopsis. *Science* **279**, 717–720 (1998).
5. Dhugga, K. S. *et al.* Guar seed β -mannan synthase is a member of the cellulose synthase super gene family. *Science* **303**, 363–366 (2004).
6. Burton, R. A. *et al.* Cellulose synthase-like *CsIF* genes mediate the synthesis of cell wall (1, 3; 1, 4)- β -D-glucans. *Science* **311**, 1940–1942 (2006).
7. Cocuron, J. C. *et al.* A gene from the cellulose synthase-like C family encodes a β -1,4 glucan synthase. *Proc Natl Acad Sci USA* **104**, 8550–8555 (2007).
8. Brown, D. M. *et al.* Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J* **52**, 1154–1168 (2007).
9. Pena, M. J. *et al.* *Arabidopsis irregularxylem8* and *irregular xylem9*: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* **19**, 549–563 (2007).
10. Anders, N. *et al.* Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc Natl Acad Sci USA* **109**, 989–993 (2012).
11. Robert, S. *et al.* An Arabidopsis endo-1,4- β -D-glucanase involved in cellulose synthesis undergoes regulated intracellular cycling. *Plant Cell* **17**, 3378–3389 (2005).
12. Zhong, R., Kays, S. J., Schroeder, B. P. & Ye, Z. H. Mutation of a chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. *Plant Cell* **14**, 165–179 (2002).
13. Schindelman, G. *et al.* *COBRA* encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. *Genes & Development* **15**, 1115–1127 (2001).
14. Li, Y. H. *et al.* *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *Plant Cell* **15**, 2020–2031 (2003).
15. Seifert, G. J. Nucleotide sugar interconversions and cell-wall biosynthesis: how to bring the inside to the outside. *Curr Opin Plant Biol* **7**, 277–284 (2004).
16. Zhang, B. *et al.* Golgi nucleotide sugar transporter modulates cell-wall biosynthesis and plant growth in rice. *Proc Natl Acad Sci USA* **108**, 5110–5115 (2011).
17. Chanliaud, E., De Silva, J., Strongitharm, B., Jeronimidis, G. & Gidley, M. J. Mechanical effects of plant cell wall enzymes on cellulose/xyloglucan composites. *Plant J* **38**, 27–37 (2004).
18. McQueen-Mason, S. & Cosgrove, D. J. Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proc Natl Acad Sci USA* **91**, 6574–6578 (1994).
19. Zhong, R., Burk, D. H., Morrison, W. H. 3rd & Ye, Z. H. A kinesin-like protein is essential for oriented deposition of cellulose microfibrils and cell wall strength. *Plant Cell* **14**, 3101–3117 (2002).
20. Gu, Y. *et al.* Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. *Proc Natl Acad Sci USA* **107**, 12866–12871 (2010).
21. Oda, Y., Iida, Y., Kondo, Y. & Fukuda, H. Wood cell-wall structure requires local 2D-microtubule disassembly by a novel plasma membrane-anchored protein. *Curr Biol* **20**, 1197–1202 (2010).
22. Endler, A. *et al.* A mechanism for sustained cellulose synthesis during salt stress. *Cell* **162**, 1353–1364 (2015).
23. Zhao, Q. Lignification: flexibility, biosynthesis and regulation. *Trends Plant Sci* **21**, 713–721 (2016).
24. Zhong, R., Lee, C. & Ye, Z. H. Evolutionary conservation of the transcriptional network regulating secondary cell-wall biosynthesis. *Trends Plant Sci* **15**, 625–632 (2010).
25. Zhong, R. *et al.* Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell Physiol* **52**, 1856–1871 (2011).
26. Hematy, K. *et al.* A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr Biol* **17**, 922–931 (2007).

27. Xu, S. L., Rahman, A., Baskin, T. I. & Kieber, J. J. Two leucine-rich repeat receptor kinases mediate signaling, linking cell-wall biosynthesis and ACC synthase in Arabidopsis. *Plant Cell* **20**, 3065–3079 (2008).
28. Krakowsky, M., Lee, M. & Coors, J. Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.) II: leaf sheath tissue. *Theor Appl Genet* **112**, 717–726 (2006).
29. Ranjan, P. *et al.* Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in *Populus*. *BioEnergy. Research* **3**, 172–182 (2010).
30. Zhang, S. J. *et al.* Identification of quantitative trait loci affecting hemicellulose characteristics based on cell-wall composition in a wild and cultivated rice species. *Mol Plant* **5**, 162–175 (2012).
31. Ashikari, M. *et al.* Cytokinin oxidase regulates rice grain production. *Science* **309**, 741–745 (2005).
32. Huang, X. *et al.* Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet* **41**, 494–497 (2009).
33. Song, X. J., Huang, W., Shi, M., Zhu, M. Z. & Lin, H. X. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet* **39**, 623–630 (2007).
34. Wang, S. K. *et al.* The *OsSPL16-GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat Genet* **47**, 949–954 (2015).
35. 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).
36. Fan, C. 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* **112**, 1164–1171 (2006).
37. Monna, L. *et al.* Positional cloning of rice semidwarfing gene, *sd-1*: Rice “Green revolution gene” encodes a mutant enzyme involved in gibberellin synthesis. *DNA Research* **9**, 11–17 (2002).
38. Yano, M. *et al.* *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* **12**, 2473–2483 (2000).
39. Ishimaru, K., Ono, K. & Kashiwagi, T. Identification of a new gene controlling plant height in rice using the candidate-gene strategy. *Planta* **218**, 388–395 (2004).
40. Duan, P. G. *et al.* Regulation of *OsGRF4* by *OsmiR396* controls grain size and yield in rice. *Nat Plants* **2**, 15203 (2016).
41. Che, R. H. *et al.* Control of grain size and rice yield by *GL2*-mediated brassinosteroid responses. *Nat Plants* **2**, 15195 (2016).
42. Xue, W. Y. *et al.* Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet* **40**, 761–767 (2008).
43. Weng, X. Y. *et al.* *Grain number, plant height, and headingdate7* is a central regulator of growth, development, and stress response. *Plant Physiol* **164**, 735–747 (2014).
44. Saito, K. Phytochemical genomics—a new trend. *Curr Opin Plant Biol* **16**, 373–380 (2013).
45. Riedelsheimer, C. *et al.* Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc Natl Acad Sci USA* **109**, 8872–8877 (2012).
46. Chen, W. *et al.* Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat Genet* **46**, 714–721 (2014).
47. Schauer, N. *et al.* Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotechnol* **24**, 447–454 (2006).
48. Chen, W. *et al.* Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals. *Nat Commun* **7**, 12767 (2016).
49. Tieman, D. *et al.* A chemical genetic roadmap to improved tomato flavor. *Science* **355**, 391–394 (2017).
50. Zhang, H. *et al.* Development and high-throughput genotyping of substitution lines carrying the chromosome segments of *indica* 9311 in the background of *japonica* Nipponbare. *J Genet Genomics* **38**, 603–611 (2011).
51. Yuan, Y. *et al.* Identification and fine mapping of *qPH6*, a novel major quantitative trait locus for plant height in rice. *Mol Breeding* **35**, 56 (2015).
52. Xu, J. J. *et al.* Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *Bmc Genomics* **11**, 1–14 (2010).
53. Wei, X. J. *et al.* *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol* **153**, 1747–1758 (2010).
54. Yan, W. H. *et al.* A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol Plant* **4**, 319–330 (2011).
55. Monna, L., Lin, H. X., Kojima, S., Sasaki, T. & Yano, M. Genetic dissection of a genomic region for a quantitative trait locus, *Hd3*, into two loci, *Hd3a* and *Hd3b*, controlling heading date in rice. *Theor Appl Genet* **104**, 772–778 (2002).
56. Matsubara, K. *et al.* Natural variation in *Hd17*, a homolog of *Arabidopsis* *ELF3* that is involved in rice photoperiodic flowering. *Plant Cell Physiol* **53**, 709–716 (2012).
57. Xu, F. *et al.* Variations in *CYP78A13* coding region influence grain size and yield in rice. *Plant Cell Environ* **38**, 800–811 (2015).
58. Weng, J. F. *et al.* Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res* **18**, 1199–120 (2008).
59. Nadeau, J. H. & Frankel, W. N. The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nat Genet* **25**, 381–384 (2000).
60. Mao, H. L. *et al.* Linking differential domain functions of the *GS3* protein to natural variation of grain size in rice. *P Natl Acad Sci USA* **107**, 19579–19584 (2010).
61. Wang, Y. X. *et al.* Copy number variation at the *GL7* locus contributes to grain size diversity in rice. *Nat Genet* **47**, 944–948 (2015).
62. Wu, A. M. *et al.* The Arabidopsis *IRX10* and *IRX10-LIKE* glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J* **57**, 718–731 (2009).
63. Faik, A. Xylan biosynthesis: news from the grass. *Plant Physiol* **153**, 396–402 (2010).
64. Huang, D. B. *et al.* A gibberellin-mediated *DELLA-NAC* signaling cascade regulates cellulose synthesis in Rice. *Plant Cell* **27**, 1681–1696 (2015).
65. Updegraff, D. M. Semimicro determination of cellulose in biological materials. *Anal Biochem* **32**, 420–424 (1969).
66. Meng, L., Li, H., Zhang, L. & Wang, J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J* **3**, 269–283 (2015).

Acknowledgements

We thank Yanbao Tian (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) for help with scanning electronic microscope analysis and Prof. Staffan Persson (University of Melbourne) for comments on the manuscript. This study was supported by the National Natural Science Foundation of China (31530051, 31561143008, and 91535105), CAS grants for XDA08010103, and the State Key Laboratory of Plant Genomics (2017A0407-18), and the Jiangsu PAPD and 333 projects.

Author Contributions

Y.Z. and Q.L. conceived the study; Z.X., S.L., C.Z., and K.Z. performed the experiments and analyzed the data; B.Z. reviewed and edited the article; Y.Z. and Z.X. wrote the article.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-12903-5>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017