nature communications

Article

https://doi.org/10.1038/s41467-024-51145-8

Variation in a Poaceae-conserved fatty acid metabolic gene cluster controls rice yield by regulating male fertility

Received: 13 January 2024

Accepted: 31 July 2024

Published online: 06 August 2024

Check for updates

Chenkun Yang[1,2,7], Shuangqian Shen[2,7], Chuansong Zhan[2,7], Yufei Li[1,3,7], Ran Zhang[1], Yuanyuan Lv[2], Zhuang Yang[1], Junjie Zhou[1], Yuheng Shi[2], Xianqing Liu[1], Jianxin Shi[4,5], Dabing Zhang[4,5,8], Alisdair R. Fernie[6] & Jie Luo[1,2]

A wide variety of metabolic gene clusters exist in eukaryotic genomes, but fatty acid metabolic gene clusters have not been discovered. Here, combining with metabolic and phenotypic genome-wide association studies, we identify a major locus containing a six-gene fatty acid metabolic gene cluster on chromosome 3 (*FGC3*) that controls the cutin monomer hydroxymonoacylglycerols (HMGs) contents and rice yield, possibly through variation in the transcription of *FGC3* members. We show that HMGs are sequentially synthesized in the endoplasmic reticulum by *OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2* and subsequently transported to the apoplast by *OsABCG22* and *OsLTPL82*. Mutation of *FGC3* members reduces HMGs, leading to defective male reproductive development and a significant decrease in yield. OsMADS6 and OsMADS17 directly regulate *FGC3* and thus influence male reproduction and yield. *FGC3* is conserved in Poaceae and likely formed prior to the divergence of *Pharus latifolius*. The eukaryotic fatty acid and plant primary metabolic gene cluster we identified show a significant impact on the origin and evolution of Poaceae and has potential for application in hybrid crop breeding.

Unlike in bacteria, where genes associated with the same metabolic pathway are often clustered to form operons, in eukaryotes, most of the non-homologous genes involved in the same metabolic pathway are scattered throughout the genome and only some non-homologous genes in the metabolic pathways of fungi and plants function by aggregating to form gene clusters[1]. A metabolic gene cluster (MGC) is a cluster of at least three genes encoding different enzymes involved in the same metabolic pathway that controls the accumulation of metabolites. Most of the reported MGCs are for secondary metabolites (such as alkaloids and terpenes)[2], and fatty acid MGCs have not yet been discovered, nor have primary MGCs *in planta*.

MGCs typically consist of a core enzyme and several modifying enzymes, both of which are synthetases[1]. The biosynthesis and transport of metabolites are two key steps in their biological function, but few transporter members are currently involved in gene clusters[3]. Exploration of transporters in MGCs will both enrich the composition of these gene clusters and deepen our understanding of transmembrane transport, spatiotemporal distribution and metabolite signal transmission. Identical or similar (or induced) expression patterns between members of MGCs suggest that specific transcription factors may co-regulate members of the gene clusters[4,5]. However, few transcription factors directly regulating MGC have been identified.

[1]School of Breeding and Multiplication (Sanya Institute of Breeding and Multiplication), Hainan University, Sanya, China. [2]Yazhouwan National Laboratory, Sanya, China. [3]Hainan Seed Industry Laboratory, Sanya, China. [4]Joint International Research Laboratory of Metabolic & Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. [5]Yazhou Bay Institute of Deepsea Sci- Tech, Shanghai Jiao Tong University, Sanya, China. [6]Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany. [7]These authors contributed equally: Chenkun Yang, Shuangqian Shen, Chuansong Zhan, Yufei Li. [8]Deceased: Dabing Zhang. e-mail: jie.luo@hainanu.edu.cn

The epidermis of most organisms is covered by the cuticle, and as one of its main components, cutin is composed primarily of long-chain hydroxy fatty acids of 16 and 18 carbons[6,7]. Hydroxy fatty acids and hydroxymonoacylglycerol (HMG), which are fatty acid metabolites, account for 30%–60% of cutin monomers (Supplementary Fig. 1). Cutin monomer synthesis involves long-chain acyl coenzyme A synthase (LACS), cytochrome P450 enzymes (CYP450), and glycerol-3-phosphate acyltransferase (GPAT), but the reaction order is inconclusive and may vary by species[8–13]. We previously reported the functions of fatty acid reductase (OsFAR2) and OsCYP704B2 in rice anther cutin biosynthesis[14,15]. In addition to controlling water loss, the cuticle also influences floral organ development by affecting organ differentiation, thus impacting crop yields[16,17]. Transcription factors such as AP2, HD-ZIP, and MYB have been shown to regulate cuticle synthesis[18–22]. Notably, while the MADS family of transcription factors broadly regulates plant flower organ development[23], its regulation of the cuticle remains unexplored.

In this work, considering the crucial role of primary metabolites in regulating growth and development, we identify a eukaryotic fatty acid and plant primary MGC (involved in both synthesis and transport) controlling rice yield by regulating anther fertility. The regulation of this MGC and the underlying evolutionary mechanisms are further analysed, providing a foundation and genetic resources for the ecological adaptation and hybrid breeding of Poaceae crops.

Results

mGWAS and pGWAS co-locate fatty acid metabolic gene cluster

To assess lipid variation in rice, we collected brown rice samples from a diverse global collection of 533 rice (*Oryza sativa*) accessions (Supplementary Data 1) and performed high-throughput quantification of lipid metabolites (Supplementary Data 2) using a widely targeted metabolomics analysis strategy[24,25]. We found that free fatty acids (FFAs) and hydroxy fatty acids (HFAs) were enriched in some varieties, almost all of which were *japonica* (Supplementary Fig. 2). A metabolic genome-wide association study (mGWAS), revealed that a major locus at 3.6 Mb on chromosome 3 was mapped by these fatty acids (Fig. 1a and Supplementary Fig. 3a–f). Interestingly, a phenotypic genome-wide association study (pGWAS) revealed that this locus

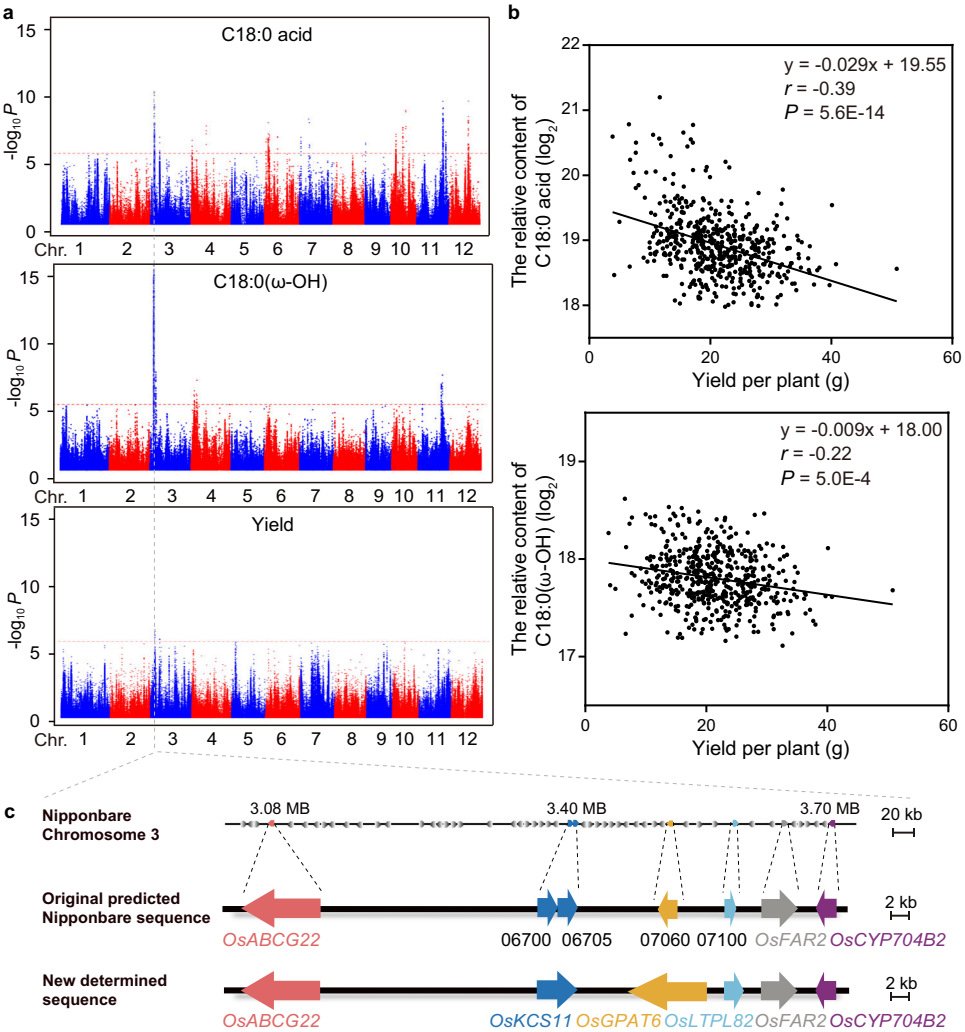


Fig. 1 | Mapping the candidate of *FGC3*. **a** The stearate (C18:0 acid) and 18-hydroxystearic acid (C18:0(ω-OH)) metabolites of the rice seed population and the rice agronomic traits yield per plant (Yield) were co-localized to the members of *FGC3*. **b** Correlations between the relative contents of C18:0 acid or C18:0(ω-OH) and yield per plant in rice populations. Pearson correlation coefficients were used for correlation analysis, and two-sided Student's *t* tests were used to calculate significance, $p \le 0.05$ was considered a significant correlation. **c** Distribution of *FGC3* members on chromosome 3. Genomic organization of the candidate gene on chromosome 3 in the original predicted Nipponbare genome compared with the newly determined genome sequence. The panel shows the genomic organization of the candidate genes, *OsABCG22* (light red), *OsKCS11* (dark blue), *OsGPAT6* (yellow), *OsLTPL82* (light blue), *OsFAR2* (gray), and *OsCYP704B2* (purple). Genes are denoted without "*Os*" and chromosome number. The gradient gray squares represent the intervening genes. See also Supplementary Fig. 5. Source data are provided as a Source Data file.

was co-localized with rice yield per plant (Yield; Fig. 1a and Supplementary Fig. 3g) and spikelet fertility (Supplementary Fig. 3h). We further revealed strong correlations between the fatty acid content and yield per plant (Fig. 1b and Supplementary Fig. 4). Combining the metabolic pathways of fatty acid and gene annotation of this locus, five major candidate genes, encoding fatty acid reductase (*OsFAR2*, *Os03g07140*), glycerol 3-phosphate acyltransferase (*OsGPAT6*, *Os03g07060*), cytochrome oxidase (*OsCYP704B2*, *Os03g07250*), and 3-ketoacyl-CoA synthase (*OsKCS11*, *Os03g06700* and *Os03g06705*) were identified (Fig. 1c). In addition, we also identified two annotated genes at this locus: one encoding an ABC-type transporter protein (*OsABCG22*, *Os03g06139*) and another encoding a lipid transporter protein (*OsLTPL82*, *Os03g07100*; Fig. 1c). By comparing the genomic regions between *indica* Minghui63 and *japonica* Nipponbare, we found that the full-length *OsKCS11* transcript should contain both *Os03g06700* and *Os03g06705* annotated sequences, and the misannotated *Os03g07060* and *Os03g07100* sequences were corrected and named *OsGPAT6* and *OsLTPL82*, respectively (Fig. 1c and Supplementary Fig. 5). We hypothesized that these six genes (*OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82*) individually or collectively control fatty acid metabolism and rice yield. Analysis of spatiotemporal expression patterns revealed that the candidate genes were predominantly co-expressed in the rice panicle (Supplementary Fig. 6), and further GUS staining revealed their specific expression in anthers (Supplementary Fig. 7), which was consistent with the basic characteristics of the gene clusters. Since the six genes that may be responsible for fatty acid metabolic reaction steps are relatively closely distributed in the rice genome, approximately 620 kb, we tentatively defined them as the *fatty acid metabolic gene cluster on chromosome 3* (*FGC3*).

Functional analysis of *FGC3* synthesis members

To investigate the metabolic roles of the synthesis members (*OsFAR2*, *OsKCS11*, *OsGPAT6* and *OsCYP704B2*), these candidate genes were independently cloned. Subcellular localization indicated that OsFAR2 was located mainly in plastids, whereas OsKCS11, OsCYP704B2 and OsGPAT6 were located in the endoplasmic reticulum, indicating their respective functional regions (Supplemental Fig. 8).

The recombinant protein of *OsFAR2* was expressed in an *E. coli* prokaryotic expression system and incubated with a palmiltoyl-acyl carrier protein (C16:0 ACP), and its reaction product was identified as palmityl alcohol (C16:0 alcohol, **1**) using an authentic standard (Fig. 2a and Supplementary Fig. 9a, reaction I). When OsFAR2 was incubated with palmiltoyl coenzyme A (C16:0 CoA), the production of C16:0 alcohol was also detected, but at a lower rate than reaction I (Fig. 2a, reaction II). These results indicate that OsFAR2 can catalyze the synthesis of C16:0 alcohol in vitro, which can be used for synthesizing the C16:0 required for downstream metabolic reactions.

A yeast expression system was then utilized to investigate the in vitro enzyme activities of OsKCS11, OsCYP704B2 and OsGPAT6, the predicted membrane-localized proteins. We found that OsKCS11 activated the fatty acid C16:0 with coenzyme A (CoA) to produce C16:0 CoA **2** (identified with a commercial standard; Fig. 2b and Supplementary Fig. 9b, reaction III). When the hydroxy fatty acid C16:0(ω-OH) was used as a substrate, a compound was produced with a mass corresponding to the precursor ion $[M+H]^+$ at *m/z* 1022.3434 (Fig. 2b and Supplementary Fig. 9c, reaction IV) and 16-hydroxypalmitoyl CoA (C16:0(ω-OH) CoA, **3**) was identified via a chemically synthesized standard[26]. These results indicated that OsKCS11 can use both fatty acids and hydroxy fatty acids as substrates. We further revealed that OsCYP704B2 displayed activities against both C16:0 and C16:0 CoA, yielding C16:0(ω-OH) **4** and C16:0(ω-OH) CoA **3**, respectively (Fig. 2c and Supplementary Fig. 9d, reactions V and VI). Finally, we found that OsGPAT6 can react with both C16:0 CoA and C16:0 (ω-OH) CoA and transfer an acyl group to G3P, producing 2-palmitoyl glycerol (MG16:0, **5**) and 16-hydroxyhexadecanoylglycerol (HMG16:0, **6**; Fig. 2d and Supplementary Fig. 9e, f, reactions VII and VIII), respectively. These results suggest that OsKCS11, OsCYP704B2 and OsGPAT6 catalyze successive metabolic reactions. This was indeed the case when these proteins were incubated with C16:0, which resulted in the production of C16:0 CoA as an intermediate product and HMG16:0 as the terminal product (Fig. 2e). These results also suggest two possible synthetic routes for HMG biosynthesis, namely OsCYP704B2-OsKCS11-OsGPAT6 and OsKCS11-OsCYP704B2-OsGPAT6, which form a diverging pathway (Fig. 2f). Considering that cutin is composed primarily of long-chain hydroxy fatty acids of 16 and 18 carbons[6,7,12], we further used C18:0 fatty acids as a substrate for in vitro reactions with OsKCS11, OsCYP704B2 and OsGPAT6, which can catalyze successive metabolic reactions, resulting in the production of HMGs (Supplemental Figs. 10 and 11). These findings suggest that *FGC3* can indeed catalyze the generation of fatty acids with 16 and 18 carbons into cutin monomer HMG metabolites.

When *OsKCS11*, *OsCYP704B2*, and *OsGPAT6* were transiently expressed in *Nicotiana benthamiana*, we observed significantly increased accumulation of HMGs, such as HMG16:0, HMG 16:1, HMG 18:0, HMG 18:1 and HMG18:2, compared to that in the control group (Supplementary Fig. 12). Taken together, these results suggest that *OsFAR2*, *OsKCS11*, *OsGPAT6* and *OsCYP704B2* form a cluster controlling the biosynthesis of cutin monomer HMGs.

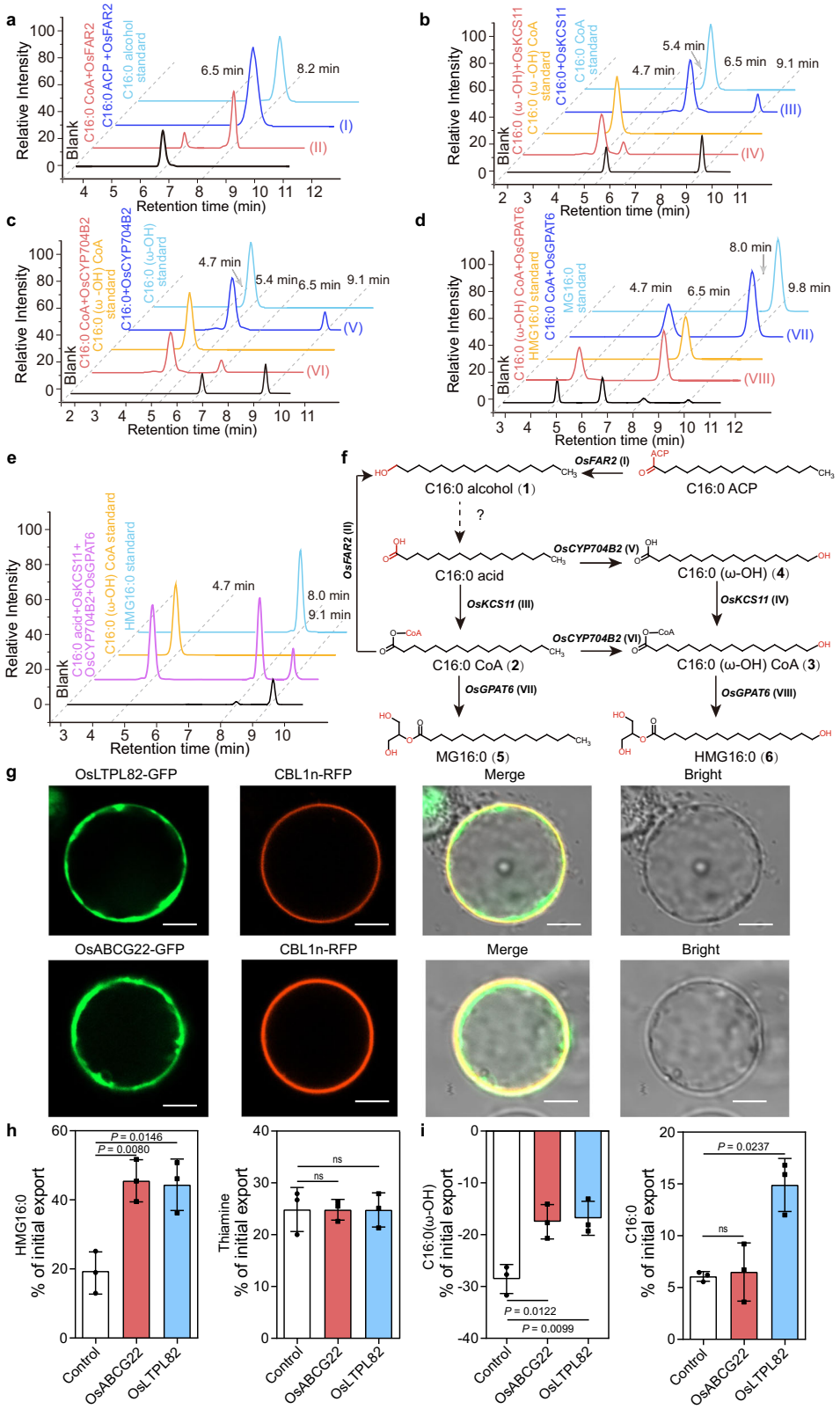
Functional analysis of *FGC3* transporter members

After synthesis, cutin monomers may be transported from the endoplasmic reticulum across the cell membrane to the cell wall by adenosine triphosphate-binding cassette transporter proteins (ABCGs) and lipid transporter proteins (LTPs)[9]. Structural prediction with the TMHMM server v2.0 revealed that both proteins have typical transmembrane structures (Supplementary Fig. 13a). Domain analysis of the *FGC3* candidates revealed that OsABCG22 has ABC_ATP_DarD and ABC2_membrane domains, and that OsLTPL82 contains an AAI domain, which is conserved for transporters (Supplementary Fig. 13b).

Subcellular localization indicated that OsLTPL82 and OsABCG22 were indeed mainly localized at the cell membrane (Fig. 2g). To investigate whether OsABCG22 and OsLTPL82 exhibit export activity for cutin monomers, the *N. benthamiana* protoplast system was used[27,28]. Since [^{13}C]-HMG16:0 was unavailable, we generated the above stable isotope products via biochemical reactions of the OsKCS11, OsCYP704B2, and OsGPAT6 proteins with the commercially available palmitic acid-1-(^{13}C) ([^{13}C]-C16:0). *N. benthamiana* cells transfected with either functional OsABCG22, or OsLTPL82 were then loaded with [^{13}C]-HMG16:0, and the unrelated substrate thiamine-4-methyl-^{13}C-thiazol-5-yl-$^{13}C_3$ ([$^{13}C_4$]-thiamine) was used as a control. Compared with the control, which presented 19% initial export, *N. benthamiana* cells expressing either OsABCG22 or OsLTPL82 presented 48% and 44% initial export, respectively, suggesting that OsABCG22 and OsLTPL82 significantly increased export. However, no enhancement was observed when labeled thiamine was loaded, indicating that both OsABCG22 and OsLTPL82 exhibited transport activities for HMGs (Fig. 2h). These results suggest that OsABCG22 and OsLTPL82 are membrane localized cutin monomer HMG transporters.

To further investigate the specificity of OsABCG22 and OsLTPL82, we synthesized [^{13}C]-C16:0(ω-OH) utilizing the biochemical reaction of OsCYP704B2 with [^{13}C]-C16:0 as a stable isotope-labeled product, a typical cutin component. We then performed transport assays with [^{13}C]-C16:0(ω-OH), and the results showed that the expression of OsABCG22 and OsLTPL82 significantly reduced the net import of labeled C16:0(ω-OH) in the control, suggesting the export activities of both proteins (Fig. 2i). Similarly, we observed enhanced release of [^{13}C]-C16:0 in the presence of OsLTPL82 compared with the control, but not with OsABCG22 (Fig. 2i). Together, these results indicate that,

https://doi.org/10.1038/s41467-024-51145-8



compared with OsABCG22, OsLTPL82 has broader export activity towards cutin monomers.

OsMADS6 and OsMADS17 regulate *FGC3*

Cutin monomer synthesis is influenced not only by structural genes but also by transcription factors. Two candidate transcription factors, OsMADS6 and OsMADS17, which were screened using co-expression analysis (Supplementary Fig. 6), were highly correlated with *FGC3* synthesis members (Fig. 3a). We found that the promoters of *FGC3* members contain different numbers of MADS box family binding motifs (Fig. 3b). The subcellular localization results indicated that both OsMADS6 and OsMADS17 were localized in the nucleus, which is consistent with their roles as transcription factors (Supplementary Fig. 14a). Dual luciferase assays revealed that both OsMADS6 and

Fig. 2 | Functional analysis of *FGC3*. **a** Results of the enzyme activity of the recombinant OsFAR2 protein in vitro. **(I)** Extracted ion chromatograms (EICs) for C16:0 alcohol **1** (*m/z* $[M\text{-}H]^-$ = 241.2531, product). **(II)** EIC for C16:0 alcohol **1** (*m/z* $[M\text{-}H]^-$ = 241.2531, product) and C16:0 COA (*m/z* $[M+H]^+$ = 1006.3527, substrate). Negative control reactions were carried out with purified GST protein (Blank). **b**–**d** Results of the enzyme activity of the OsKCS11, OsCYP704B2, and OsGPAT6 yeast microsomes in vitro. **(III)** EIC for C16:0 COA **2** (*m/z* $[M+H]^+$ = 1006.3527, product) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate). **(IV)** EIC for C16:0(ω-OH) CoA **3** (*m/z* $[M+H]^+$ = 1022.3476, product) and C16:0(ω-OH) (*m/z* $[M\text{-}H]^-$ = 271.2273, substrate). **(V)** EIC for C16:0(ω-OH) **4** (*m/z* $[M\text{-}H]^-$ = 271.2273, product) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate). **(VI)** C16:0(ω-OH) CoA **3** (*m/z* $[M+H]^+$ = 1022.3476, product) and C16:0 COA (*m/z* $[M+H]^+$ = 1006.3527, substrate). **(VII)** EIC for MG 16:0 **5** (*m/z* $[M+H]^+$ = 331.2848, product) and C16:0 COA (*m/z* $[M+H]^+$ = 1006.3527, substrate). **(VIII)** EIC for HMG16:0 **6** (*m/z* $[M+H]^+$ = 347.2789, product) and C16:0(ω-OH) CoA (*m/z* $[M+H]^+$ = 1022.3476, substrate). **e** Results of enzyme activity in vitro after the yeast microsomes OsKCS11, OsCYP704B2, and OsGPAT6 were incubated with palmitic acid. EIC for HMG16:0 **6** (*m/z* $[M+H]^+$ = 347.2789, product), C16:0(ω-OH) CoA **3** (*m/z* $[M+H]^+$ = 1022.3476, product) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate). For (**b**–**e**), Negative control reactions were carried out with microsomal preparations from recombinant yeast transformed with 'empty' pESC-URA (Blank). **f** Synthetic route diagram for the cutin monomer HMGs. **g** Results of *OsLTPL82* and *OsABCG22* subcellular localization in *Nicotiana benthamiana* leaves. CBL1n-RFP was used as a plasma membrane marker. Bars = 10 μm. **h**, **i** HMG16:0, thiamine, C16:0(ω-OH) and C16:0 transport activities in *N. benthamiana* protoplasts, expressing OsABCG22 (red), OsLTPL82 (blue) and the vector control (white) were measured. All products were validated by comparison with authentic standards (see Supplementary Fig. 9). The error bars represent the standard errors of three experimental repeats. Data are presented as the means ± SD. Two-sided unpaired Student's *t* tests. *P* values are indicated. Source data are provided as a Source Data file.

OsMADS17 activated the promoters of *FGC3* synthesis members, but for transport members, only OsMADS6 significantly activated *OsLTPL82* (Fig. 3c). The ChIP-qPCR results revealed that OsMADS6 and OsMADS17 can bind to all *FGC3* member promoters in vivo (Fig. 3d). Furthermore, a yeast one-hybrid assay demonstrated that OsMADS6 and OsMADS17 could directly bind to the *OsKCS11* promoter (Supplementary Fig. 14b).

To further validate the functions of OsMADS6 and OsMADS17 in rice, mutant plants harboring these two transcription factors were successfully constructed. In *osmads17* mutants, the expression levels of *FGC3* members were all significantly decreased, whereas in *osmads6* mutants, only the expression levels of *OsKCS11*, *OsCYP704B2*, *OsGPAT6* and *OsLTPL82* were significantly decreased (Fig. 3e). These results indicate that both OsMADS6 and OsMADS17 are involved in the transcriptional regulation of *FGC3*. When *OsMADS6* and *OsMADS17* were transiently expressed in *N. benthamiana* leaves, the contents of HFAs and HMGs significantly increased compared with that in the control, confirming that *OsMADS6* and *OsMADS17* positively regulate the synthesis of hydroxylated cutin monomers (Fig. 3f). Thus, *OsMADS6* and *OsMADS17* regulate both the biosynthesis and transport of cutin monomer HFAs and HMGs by directly regulating *FGC3*.

FGC3 controls rice yield by regulating male fertility

Physiological functions were investigated by using CRISPR/Cas9 gene editing or mutant plants with a T-DNA insertion and stably transformed *FGC3* members, and *OsMADS6* and *OsMADS17* were used for phenotypic observations (Supplementary Fig. 15). All the mutant plants presented abnormal floral organ development (Fig. 4 and Supplementary Fig. 16). Compared with the wild type plants, the mutant plants presented smaller and paler anthers (Fig. 4a, c, e, g, i, k and Supplementary Fig. 16a, d) and significantly fewer viable mature pollen grains (Fig. 4b, d, f, h, j, l and Supplementary Fig. 16b, e). Further statistical analysis revealed that, compared with the control, the *FGC3* mutants generally presented a 30% to 62% decrease in yield per plant, whereas the *osmads6* mutants presented a massive 98% to 99% decrease in yield (Fig. 4m and Supplementary Fig. 16c, f). These findings indicate that *FGC3* controls rice yield by regulating male fertility.

Further TB staining revealed that the anthers of *FGC3*, *OsMADS6* and *OsMADS17* mutant plants had defects in the epidermis (Supplementary Fig. 17). To demonstrate that HMGs are indeed transported to the apoplast epidermis of the cell wall, we used Sudan Red 7B to stain HMGs from anthers and sectioned them for cytological observation. The results indicated that the anther surface and pollen exine of the wild-type plants were more intense stained, whereas the anther surface and pollen exine of the mutants (*osfar2*, *oscyp704b2*, *osgpat6*, *oskcs11*, *osltpl82* and *osabcg22*) were unable to stain or showed much weaker staining, suggesting that HMGs are transported to the anther surface and pollen exine and that mutation of any members of *FGC3* resulted in the loss of the lipid layer (including HMGs) in these tissues (Supplemental Fig. 18). Analysis of anther HMG content and cutin levels in the mutant plants revealed that they were significantly lower than those in their wild type counterparts (Supplementary Figs. 19 and 20). These results indicate that mutation of any member of *FGC3* or the transcription factors *OsMADS6* or *OsMADS17* causes alterations in the composition and content of cutin, monomers and abnormal development of anthers and pollen grains, thus resulting in a significant decrease in yield.

Natural variation and ecological distribution of *FGC3*

To explore potential determinants of natural yield variation, we investigated possible functional polymorphisms in *FGC3*. We observed 24, four, six, 11, one and eight possible functional SNPs within the promoters of *OsGPAT6*, *OsCYP704B2*, *OsKCS11*, *OsFAR2*, *OsLTPL82* and *OsABCG22*, respectively, that were significantly correlated with the palmitate and stearate contents and rice yield per plant (Supplementary Figs. 21 and 22, Supplementary Data 3). These results suggest that natural variations in *FGC3* gene expression may lead to variations in HMG content and yield. On the basis of the nucleotide polymorphisms identified in these gene promoter regions, the 533 rice accessions could be divided into two main haplotypes, and there were significant differences in stearate content and rice yield per plant between HapA and HapB, with HapB having higher stearate levels but lower yields than HapA (Fig. 5a, Supplementary Table 1). Intriguingly, most HapA varieties were found to be *indica*, whereas all HapB varieties were *japonica* (Fig. 5b). We next tested whether differences in the expression of *FGC3* synthesis members contributed to variations in stearate content and rice yield per plant. The qRT-PCR results revealed that the expression levels of *OsFAR2*, *OsGPAT6* and *OsCYP704B2* were significantly correlated with the stearate content and yield per plant (Fig. 5c and Supplementary Fig. 23a), and significantly different between haplotypes (Supplementary Fig. 23b). These results suggest that the natural variation in stearate content and yield is determined by the overall transcription of *FGC3* synthesis members.

Given that the cuticle plays a central role in restricting water loss, we queried the average annual relative humidity of regions where some Chinese varieties are located[16,17,29]. Since 96% of the *indica* varieties are located in areas where the humidity is greater than 70%, we analysed the geographical distribution of the remaining *japonica* varieties, and the results revealed that most HapB varieties grow in areas where the humidity is greater than 70% (Fig. 5d). Moreover, the Jap^B allele frequency increased from 0% at <60% humidity to 18.8% at 60%–70% humidity and then to 55.3% at >70% humidity (Fig. 5e). Further analyses revealed significant differences in relative humidity (RH), stearate content and rice yield per plant between Jap^A and Jap^B (Fig. 5f–h). We conducted experiments to verify the effect of humidity on pollen viability. We selected the two haplotype varieties with the

https://doi.org/10.1038/s41467-024-51145-8

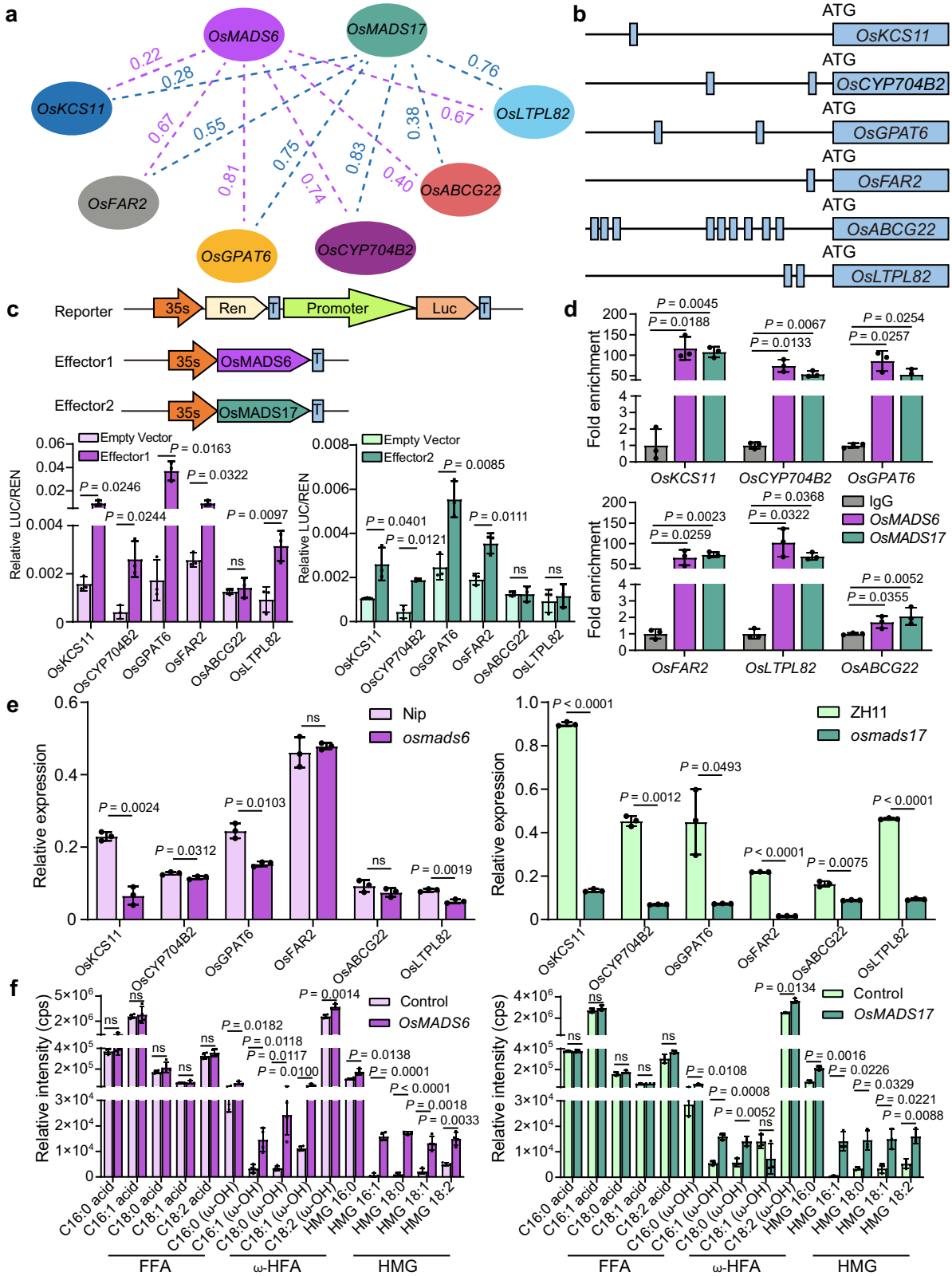


Fig. 3 | OsMADS6 and OsMADS17 regulate *FGC3*. **a** Co-expression analysis of *OsMADS6* and *OsMADS17* with *FGC3* members. Pearson correlation coefficient values were calculated between the expression levels of each pair of genes. **b** Promoter analysis of *FGC3* members. Diagram of the *OsKCS11*, *OsCYP704B2*, *OsGPAT6*, *OsFAR2*, *OsABCG22* and *OsLTPL82* promoter regions, showing the relative positions of the MADS box family binding motif (blue rectangles). **c** Dual luciferase assays demonstrated that OsMADS6 and OsMADS17 could activate *FGC3* synthesis members. **d** ChIP-qPCR results showing that OsMADS6 and OsMADS17 can bind to *FGC3* member promoter regions using young panicles. The fold enrichment was calculated relative to the negative control IgG antibody. **e** Analysis of *FGC3* member expression in *osmads6* and *osmads17* mutant plants. **f** Accumulation of cutin monomer in *Nicotiana benthamiana* leaves transiently overexpressing *OsMADS6* and *OsMADS17*. For (**c**–**e**), n = 3 biological replicates. For (**f**), n = 4 biological replicates. Data are presented as the means ± SD. Two-sided unpaired Student's *t* tests. *P* values are indicated. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-51145-8

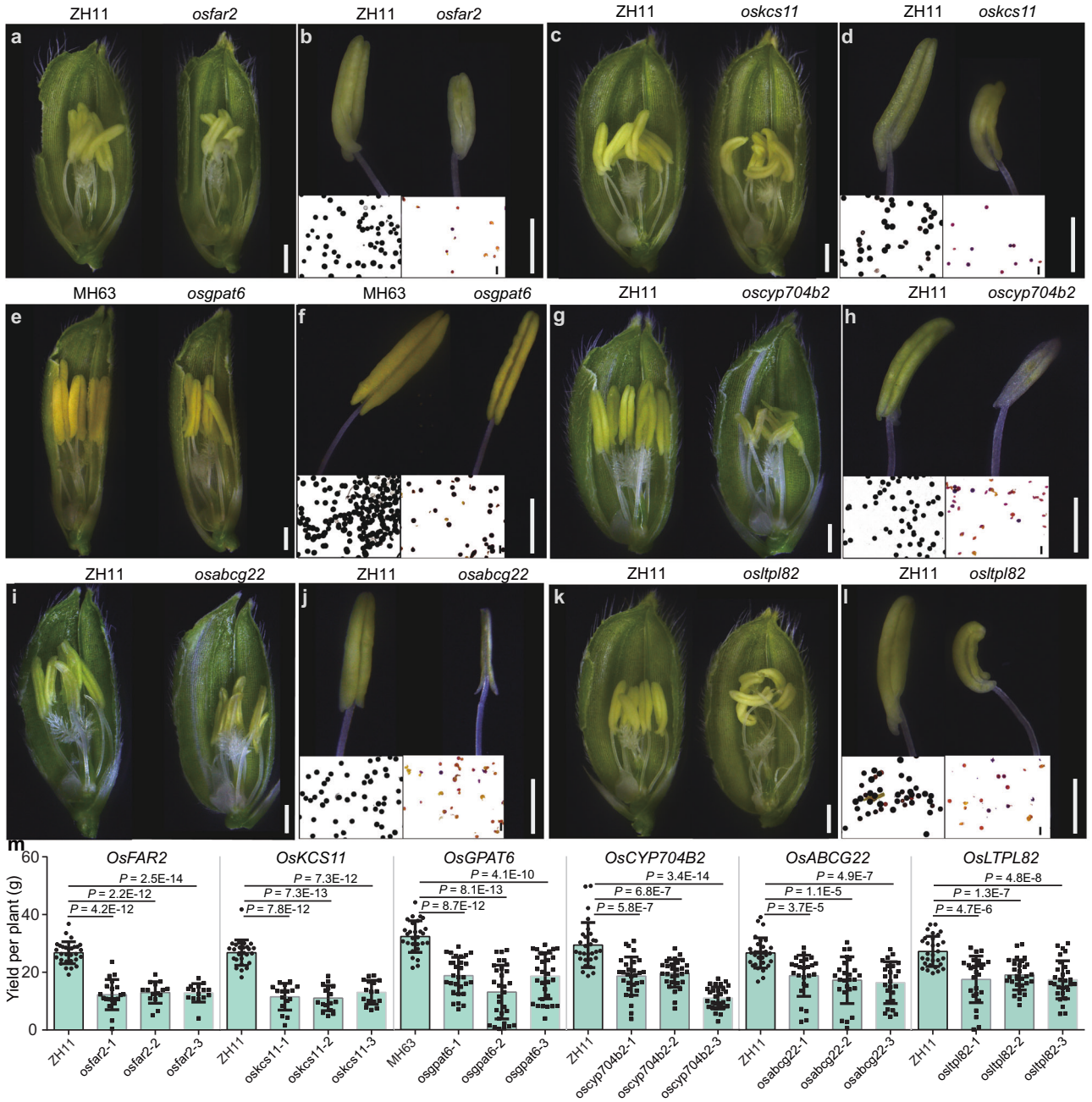


Fig. 4 | Phenotype analysis of *FGC3*. Anther phenotype and yield per plant in crispr transgenic plants of *FGC3* members *OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82*. **a**, **c**, **e**, **g**, **i**, **k**, Wild-type spikelet (left) and mutant spikelet (right) after half of the lemma and palea were removed. Bars = 1 mm. **b**, **d**, **f**, **h**, **j**, **l**, Wild-type (left) and mutant (right) anthers, bars = 1 mm, and wild-type (left) and mutant (right) pollen grains after stained with I_2-KI solution, bars = 25 μm. **m** Yield per plant statistics of wild type and mutant. For (**m**), n = 15–30 biological replicates. Data are presented as the means ± SD. Two-sided unpaired Student's *t* tests. *P* values are indicated. Source data are provided as a Source Data file.

closest background among the 533 *O. sativa* accessions for testing, these varieties presented functional variation. A phylogenetic tree was constructed using all the published SNPs of 533 rice populations[30], and two haplotype varieties of the same or closer clade were selected as the closest background rice varieties (Supplemental Fig. 24).When freshly shed pollen grains were placed on a glass slide and maintained at 27–32 °C and 30–60% RH, the pollen grains of Jap^B (containing varieties C111, C134, C144, C185 and W319) shrank immediately, becoming fully dehydrated within one minute, whereas the pollen grains of Jap^A (containing varieties C130, C137, C179, W127 and W326) adhered to one another, produced exudate and shrank over a period of approximately six minutes (Supplemental Fig. 25a). In contrast, when the RH was greater than 70%, the pollen grains of Jap^B also adhered to each other and were able to remain for some time to prevent rapid water loss, becoming completely dehydrated after approximately six minutes, whereas the pollen grains of Jap^A behaved similarly for approximately twelve minutes (Supplementary Fig. 25b). In addition, the pollen fertility of the two main haplotype plants was tested in the I_2-KI solution, and the results revealed that the pollen fertility of HapB was significantly lower than that of HapA (Supplementary Fig. 26). These results indicated that the varieties with high HMG contents are able to remain viable for longer periods of time and are fertile at both high and low humidities, whereas those with low HMG contents suffer from rapid dehydration and the loss of pollen viability at low humidities, and

https://doi.org/10.1038/s41467-024-51145-8

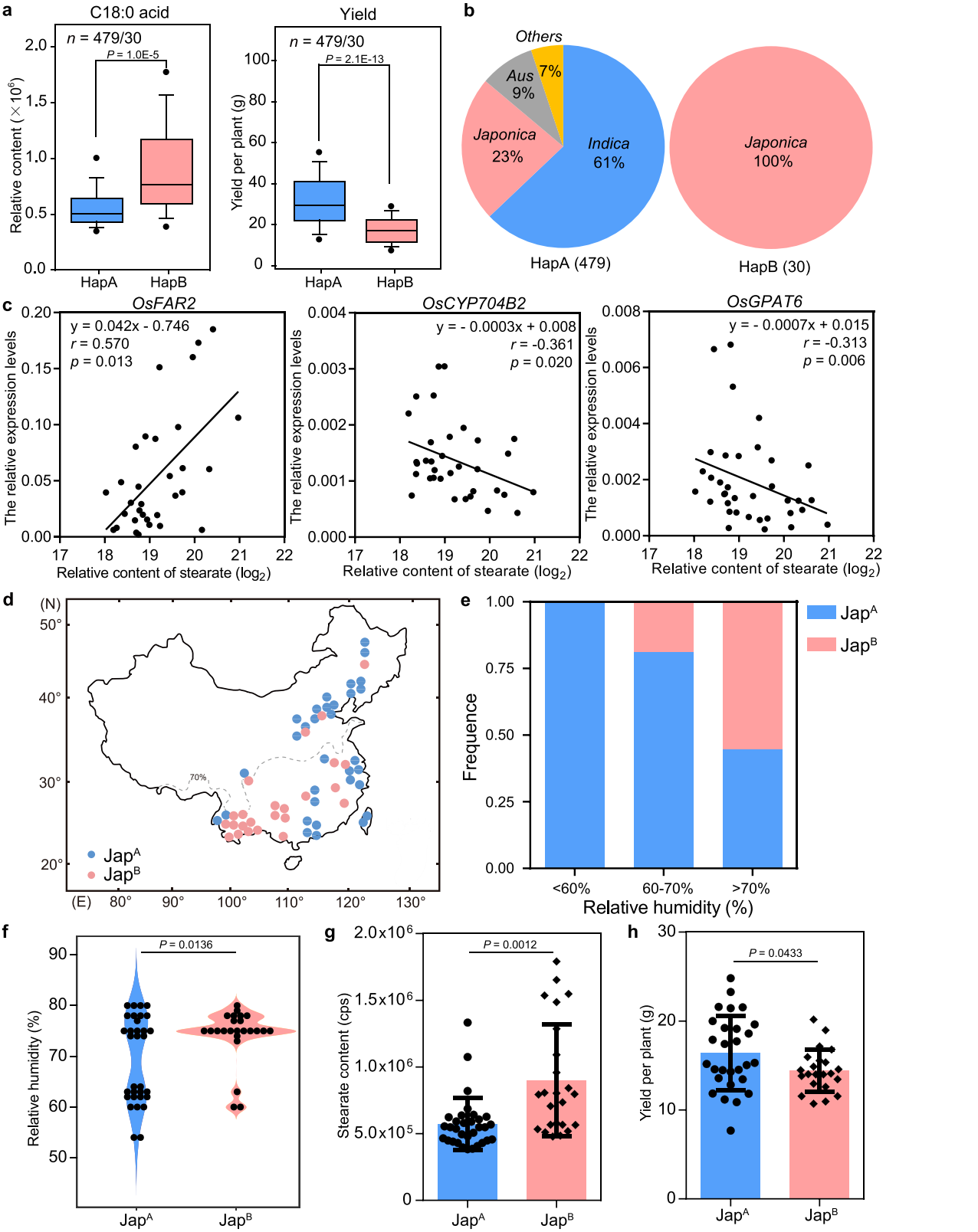


Fig. 5 | Natural variation and ecological distribution of *FGC3*. **a** Genotyping of stearate and yield per plant trait at the population level. In box plots, the box limits indicate the 25th and 75th percentiles, the whiskers indicate the full range of the data, and the center line indicates the median. **b** Statistical results of the distribution of different haplotypes in subspecies at the population level. **c** Correlations between the relative contents of stearate and the transcription levels of *OsFAR2*, *OsCYP704B2*, and *OsGPAT6* in rice varieties. **d** Geographic distribution the two main alleles of *FGC3* relative to the average annual relative humidity in China. The distribution of *japonica* varieties relative to humidity data is shown in the figure, with 32 varieties belonging to Jap^A and 24 varieties belonging to Jap^B. The dashed gray line represents the threshold of 70% relative humidity. **e** Allele distribution of the varieties containing the A or B allele associated with relative humidity. **f** Violin plot of average annual humidity distribution in the two alleles. **g**, **h** Box plot of stearate content and yield per plant distribution in two alleles. For (**f**–**h**), n = 32 and 24 biologically independent samples, respectively. Data are presented as the means ± SD. Two-sided unpaired Student's *t* tests. *P* values are indicated. Source data are provided as a Source Data file.

are able to remain fertile at higher humidities for some time to prevent rapid dehydration but with significantly reduced yields. These findings suggest that *FGC3* might impact rice ecological adaptation.

Evolution and conservation of *FGC3*

To gain further insights into the evolutionary context of *FGC3*, we performed an extensive BLAST search for homologs in the Phytozome and NCBI databases using the protein sequences of *FGC3* members as baits. We found that *FGC3* members have genes with high sequence identity (identity > 50%) in most monocotyledons and that *FGC3* members are partially clustered in some non-poaceae monocots (Supplementary Fig. 27). For example, orthologous *AcoFAR2*, *AcoCYP704B2*, *AcoLTPL82*, and *AcoABCG22* are distributed on the same chromosome in *Ananas comosus*, whereas *AcoGPAT6* and *AcoKCS11* are distributed on different chromosomes, resulting in the possibility that HMGs may not be efficiently and rapidly synthesized. However, all members of *FGC3* are tightly distributed on the same chromosome in Poaceae, indicating that *FGC3* is conserved in Poaceae (Fig. 6a). To better understand the homologous genes of the two transcription factors, extensive BLAST searches were performed for orthologues of *OsMADS6* and *OsMADS17*. *OsMADS6* and *OsMADS17* have orthologues in dicots, but they are still more homologous to species in the Poaceae (Supplementary Fig. 28). Divergent time analyses suggest that *FGC3* probably formed prior to the divergence of *Pharus latifolius* from other grasses approximately 90–130 million years ago and then formed in Oryzoideae, Pooideae and Panicoideae. To verify the biochemical function of *FGC3* in Poaceae, we cloned OsKCS11, OsGPAT6 and OsCYP704B2 orthologues from *Pharus latifolius*, *Setaria italica* and *Brachypodium distachyon* (Fig. 6a). In vitro enzymatic activity assays demonstrated that all the above proteins have the same biochemical function across these species (Fig. 6b–d). The three proteins in each species can produce HMG16:0 from C16:0 (Fig. 6e), indicating that *FGC3* is indeed conserved in Poaceae.

The *FGC3* identified in this study comprises the HMGs synthesizing the genes *OsFAR2*, *OsCYP704B2*, *OsKCS11* and *OsGPAT6*, and the transport genes *OsABCG22* and *OsLTPL82*, and a working model of their role in cutin metabolism was proposed (Fig. 7). De novo synthesized and esterified fatty acyl-ACPs are first reduced by OsFAR2 to fatty alcohols in plastids, which then produce fatty acids via unidentified genes, and then diffuse into the endoplasmic reticulum. HMGs can subsequently be produced in the ER through sequential reactions via either of two main synthetic routes, namely, OsCYP704B2-OsKCS11-OsGPAT6 or OsKCS11-OsCYP704B2-OsGPAT6. The HMGs are subsequently transported to the apoplast epidermis of the cell wall by OsABCG22 and OsLTPL82, where cutin is deposited.

Discussion

The present study identifies HMG components in rice anther cuticles, a class of metabolites previously described in *Stephanotis floribunda*, *Prunus laurocerasus* and *Citrus aurantium*[31]. These findings help elucidate cutin monomer metabolism, improve the network of HMG synthesis, transport and regulation, and lay a foundation for further study of male reproductive development. In particular, it was reported that the the MADS family gene regulates male reproductive development in rice through cutin monomer metabolism. The anther surface and pollen exine are covered with lipid layers that consist of fatty acid derivative hydroxy fatty acids and HMGs, which are critical for male reproductive development, protecting the male gametophytes from dehydration and facilitating subsequent pollen-stigma communication and adhesion, and thus any disruption of these layers often leads to microspore abortion and male sterility[6,7,12]. By identifying the functions of genes in *FGC3* and further optimizing the components of this gene cluster, related genes can be engineered to help produce crops resistant to low humidity environmental stress and expand crop planting areas.

With a focus on metabolic research, the study of MGCs has made significant strides in recent years. However, further exploration and expansion are necessary to deepen our understanding of the identification and physiological function of primary MGCs, as well as their composition and regulation. Here, we efficiently identified a eukaryotic fatty acid and plant primary MGC. Compared with the reported MGCs with only biosynthetic genes (core and modifier enzymes), *FGC3* comprises four biosynthesis members (*OsFAR2*, *OsKCS11*, *OsGPAT6* and *OsCYP704B2*) and two transporters (*OsABCG22* and *OsLTPL82*), which collectively belong to six different gene families and can control both biosynthesis and transport processes. In addition, the metabolites synthesized by MGCs usually play important roles in the biotic/abiotic stress response or have medicinal value[1,2,32]. Primary metabolites, such as sugars, fatty acids and amino acids, are the material basis for growth, development, and reproduction[2]. It has previously been reported that MGCs of fatty acid-related metabolites, such as the β-diketone gene cluster in wheat and barley, are responsible for wax synthesis and may affect yield, especially under drought conditions[33]. Another reported gene cluster is the highly modified fatty acid falcarindiol gene cluster in tomatoes, which has been shown to be involved in the defense response of pathogenic bacteria[34]. In this study, we found that *FGC3* regulates anther development in rice, which fits the canonical description of the function of primary metabolites[35,36], thus providing the substantive case of primary MGC *in planta*.

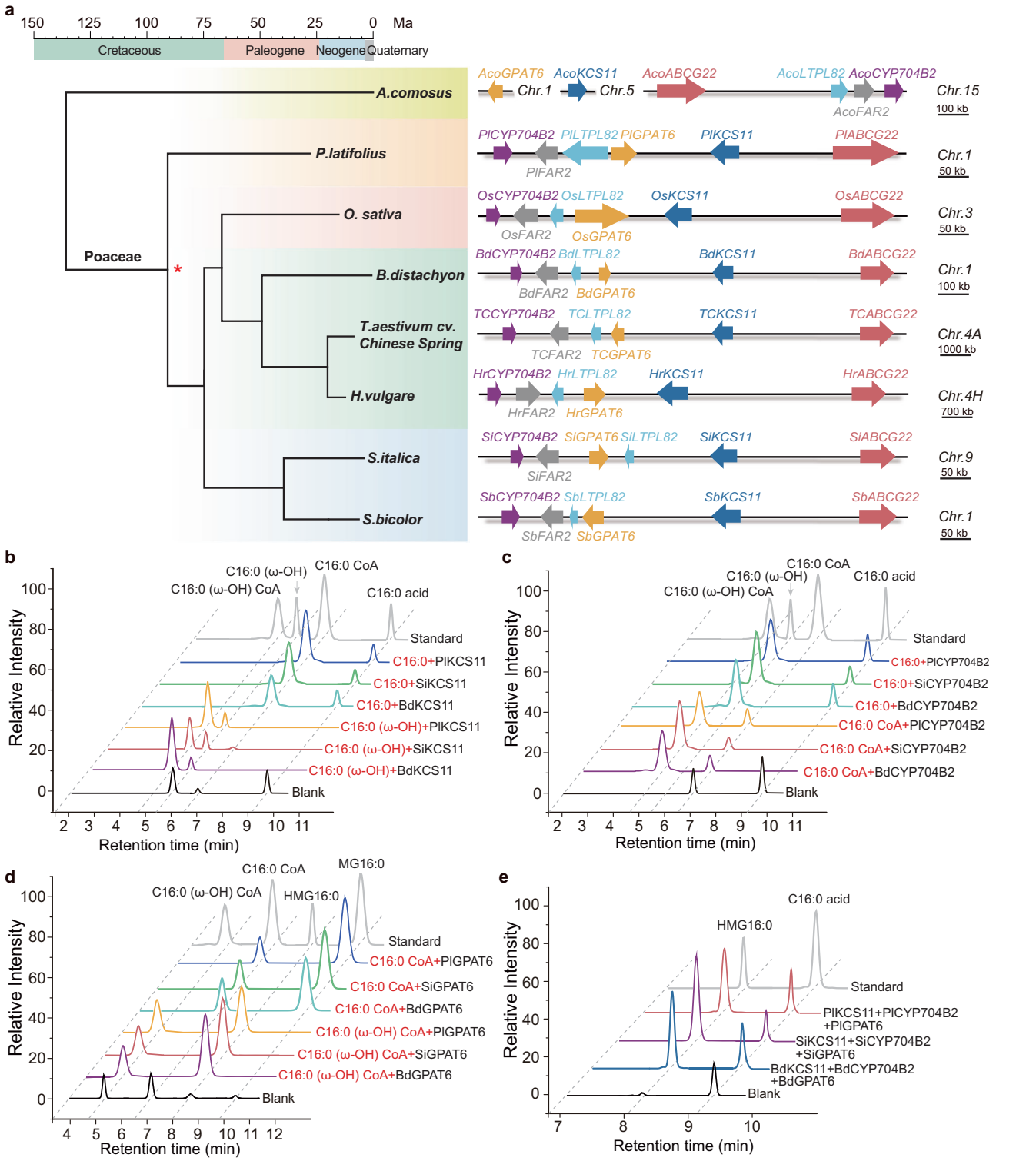
The Poaceae represents the fifth largest angiosperm family, with approximately 770 genera and 12,000 species including rice, wheat, reeds, and sugarcane, and is the most closely related to human life, and the most common crop resource[37]. This family includes the major food crops on which we survive, such as rice, wheat, maize, sorghum, and a variety of forage grasses, as well as other economically important species such as reeds, sugarcane, and bamboo. The *FGC3* identified in this study probably formed prior to the divergence of *Pharus latifolius*, the earliest group in Poaceae to evolve a spikelet structure, whose unique floral morphology is considered an ancestral trait of Poaceae and is closely related to crop yield. The conservation of *FGC3* in Poaceae suggests that it may have significant impacts on the origin and evolution of Poaceae spikelets. Our results provide a theoretical basis for investigating the origin and evolution of spikelets and provide opportunities for manipulating cereal crops to increase yield.

Our study revealed that *FGC3* HapB rice varieties are male fertile only under high humidity conditions due to their low cutin monomer content, and may exhibit male sterility at low humidity (RH < 60%). Given that *FGC3* causes the humidity-sensitive genic male sterility (HGMS) trait, this provides a option for two-line hybrid breeding in rice. There are different types of environment-sensitive genic male sterility (EGMS), such as thermo-sensitive genic male sterility (TGMS), photoperiod-sensitive genic male sterility (PGMS) and HGMS. Currently, PGMS and TGMS lines have a wide range of applications in hybrid breeding, but sudden and strong environmental fluctuations may be encountered during hybrid seed production, resulting in false-positive hybrid seeds[29,38]. In contrast, HGMS is unaffected by temperature and photoperiod, and thus can potentially be used for hybrid rice breeding in areas where the RH is less than 60%, such as Urumqi, Xinjiang, China. Since *FGC3* is conserved in Poaceae, the HGMS identified here may be applicable to major crops and other species.

Methods

Plant material and growth conditions

A diverse global collection of 533 *O. sativa* accessions including both landraces and elite varieties was obtained[39]. Information about the accessions, such as variety name, country of origin, longitude and latitude origin and subpopulation identity is listed in Supplementary Data 1. The plants were grown in a randomized complete-block design (including two rows of each accession and ten plants in each row) with



three replicates. Mature seeds were randomly collected from 10 out of the 20 plants and pooled for metabolic profiling. Rice plants examined under field conditions were grown during the normal rice-growing seasons at the Experimental Station of Huazhong Agricultural University (Wuhan, China). All seeds were planted in a seedbed in mid-May and subsequently transplanted to the field in mid-June. The planting density was at 16.5 cm between individual plants within each row, while a distance of 26 cm between rows was maintained. Field management practices, including irrigation, fertilizer application and pest control strictly adhered to conventional agricultural norms. The yield per plant was measured as the weight of all the filled grains of the plant. To prevent loss from overripening, each accession was harvested individually at its maturity. The yield per plant trait was determined for each accession in 533 rice populations, and the average yield per plant for each accession was used for the pGWAS.

Preparation of metabolic samples

Rice samples were extracted as previously reported[24,25]. Briefly, the samples were first freeze-dried and then ground into powder using a mixer mill (MM400, Retsch) for 1.5 min at 30 Hz. Subsequently, 100 mg

Fig. 6 | Study of the evolution and conservation of *FGC3*. **a** Study of the evolution and conservation of *FGC3*. Phylogeny and time-scale evolution of the grass family. Estimated divergence times for the grass family on the basis of *FGC3* member (*OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82*) genes from seven grasses representing three subfamilies and one outgroup, revealing the origin of this family in the Cretaceous and the divergence of its major lineages. The red star represents the key evolutionary node. The orientation of the *H. vulgare* and *O. sativa* chromosomal segments have been reversed in order to clarify the resemblance to the other MGCs. **b**–**d** In vitro results of orthologous proteins of OsKCS11, OsGPAT6 and OsCYP704B2 in *Pharus latifolius*, *Setaria italica* and *Brachypodium distachyon*, respectively. The extracted ion chromatograms (EICs) is consistent with that in Fig. 2. In short, (**b**), EIC for C16:0 COA **2** (*m/z* $[M+H]^+$ = 1006.3527, product, RT 6.5 min) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate, RT 9.1 min). EIC for C16:0(ω-OH) CoA **3** (*m/z* $[M+H]^+$ = 1022.3476, product, RT 4.7 min) and C16:0(ω-OH) (*m/z* $[M\text{-}H]^-$ = 271.2273, substrate, RT 5.4 min). **c** EIC for C16:0(ω-OH) **4** (*m/z* $[M\text{-}H]^-$ = 271.2273, product, RT 5.4 min) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate, RT 9.1 min). C16:0(ω-OH) CoA **3** (*m/z* $[M+H]^+$ = 1022.3476, product, RT 4.7 min) and C16:0 COA (*m/z* $[M+H]^+$ = 1006.3527, substrate, RT 6.5 min). **d** EIC for MG 16:0 **5** (*m/z* $[M+H]^+$ = 331.2848, product, RT 9.8 min) and C16:0 COA (*m/z* $[M+H]^+$ = 1006.3527, substrate, RT 6.5 min). EIC for HMG16:0 **6** (*m/z* $[M+H]^+$ = 347.2789, product, RT 8.0 min) and C16:0(ω-OH) CoA (*m/z* $[M+H]^+$ = 1022.3476, substrate, RT 4.7 min). **e** Results of in vitro enzyme activity of orthologous proteins of OsKCS11, OsGPAT6 and OsCYP704B2 in *Pharus latifolius*, *Setaria italica* and *Brachypodium distachyon* after incubation with palmitic acid. EIC for HMG16:0 **6** (*m/z* $[M+H]^+$ = 347.2789, product, RT 8.0 min) and C16:0 (*m/z* $[M\text{-}H]^-$ = 255.2324, substrate, RT 9.1 min). Negative control reactions were carried out with microsomal preparations from recombinant yeast transformed with 'empty' pESC-URA (Blank).

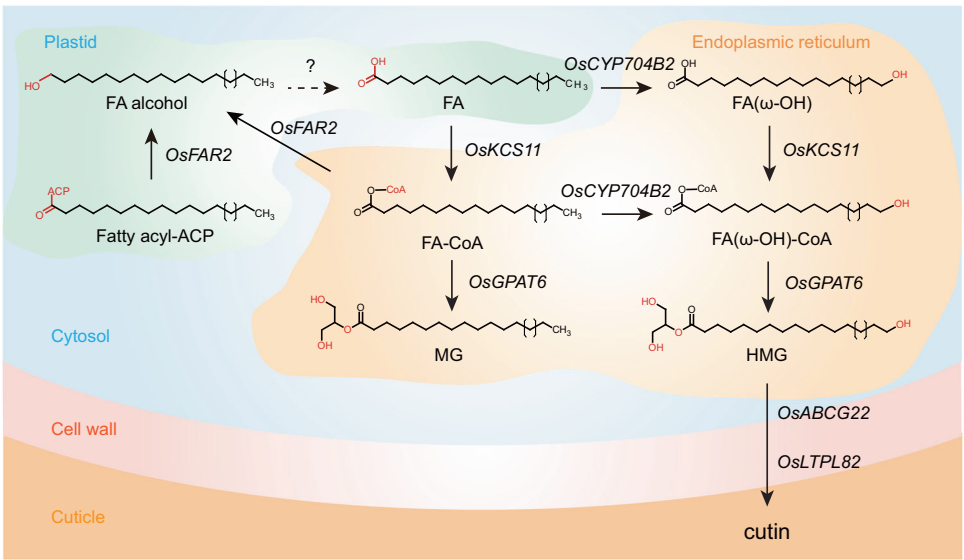


Fig. 7 | A proposed model of the biosynthesis pathway for cutin monomer. De novo synthesized and esterified fatty acyl-ACPs are first reduced by OsFAR2 to fatty alcohols in plastids, which then produce fatty acids via unidentified genes, and then diffuse into the endoplasmic reticulum (ER). Cutin monomer HMGs can subsequently be produced in the ER through sequential reactions via either of two main synthetic routes, namely, OsCYP704B2-OsKCS11-OsGPAT6 or OsKCS11-OsCYP704B2-OsGPAT6, which are subsequently transported to the apoplast epidermis of the cell wall by OsABCG22 and OsLTPL82, where cutin is deposited. The black arrows represent the steps verified in this article, and dashed arrows represent unknown step. FA, fatty acid; MG, monoacylglycerol; HMG, hydroxymonoacylglycerol.

of the resulting powder was weighed and extracted overnight at 4 °C with pure methanol. The extracts were then centrifuged for 10 min at 10,000 × *g*. The supernatants were collected, followed by filtration (SCAA-104, 0.22 μm pore size; ANPEL Shanghai, China, www.anpel.com.cn/) before LC–MS analysis. To establish a specific metabolic database of brown rice, samples were mixed into multiple samples for widely targeted metabolomics analysis. For rice anther sample analysis, we measured cutin monomer precursors before cutin polymerization.

Metabolome analysis

For metabolome analysis, samples were analysed via an ultrahigh performance liquid chromatography-tandem mass spectroscopy (UPLC-MS/MS)-based targeted method combined with a nontargeted metabolic profiling method[25]. The UPLC analytical conditions were as follows: column, shim-pack GISS C18 (pore size 1.9 μm, length 2.1 × 100 mm); solvent system, water (0.04% acetic acid): acetonitrile (0.04% acetic acid); gradient program, 0 min, 5% B; 11.0 min, 100% B; 15.0 min, 100% B; 15.5 min, 5% B; 20.0 min, 5% B; flow rate, 0.4 mL/min; temperature, 40 °C; injection volume: 2 μL. Nontargeted metabolic profiling analysis via LC-ESI-QqTOF-MS/MS (TripleTOF 5600 + , Applied Biosystems, USA) was performed in TOF MS scans mode. Data were acquired using the following settings: ion source gas 1, 55 psi; ion source gas 2, 55 psi; curtain gas, 35 psi; interface heater temperature, 550 °C; ion spray voltage floating, 5500 V. The scanning mass range was from *m/z* 100–1000 with an accumulation time of 0.10 s in TOF MS mode. For IDA, the CEs were set at 10, 30, and 50 eV, survey scans were acquired in 0.08 s and as many as 10 product ion scans were collected. The total cycle time was fixed to 0.79 s. Targeted metabolic profiling analysis was performed using a scheduled multiple reaction monitoring (MRM) via LC-ESI-QQQ-MS/MS (LCMS-8060, SHIMADZU, Japan). The ESI source operation parameters were as follows: nebulizing gas flow, 3 L/min; heating gas flow, 10 L/min; interface temperature, 500 °C; DL temperature, 250 °C; heat block temperature, 400 °C; drying gas flow, 10 L/min. The total loop time was fixed at 1.00 s.

Recombinant protein expression, purification and enzyme assay

The cDNAs of the targeted *OsFAR2* were recombined into the modified pGEX-6p-1 expression vector (Novagen, Darmstadt, Germany) with a glutathione S-transferase (GST) tag via the ClonExpress MultiS One Step Cloning Kit (Vazyme, Nanjing, China). These expression vectors were transformed into *E. coli* BL21 (DE3) (Weidi Biotechnology, Shanghai, China) for heterologous protein expression[40]. Single

colonies were selected and cultured in Luria–Bertani liquid medium containing ampicillin until the A_{600nm} reached 0.6–0.8, followed by induction with isopropyl-β-D-thiogalactoside (with a final concentration of 0.1 mM) for 16 h at 20 °C. The cells were harvested (6000 × *g* for 10 min) and disrupted by a high-pressure cracker. GST-fused target proteins were purified with glutathione Sepharose 4B agarose (GE Healthcare, Boston, USA) and confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). All the purified proteins were stored at -80 °C until further use.

The in vitro enzymatic assays for OsFAR2 were performed in a 100 μL total volume containing 50 μM C16:0 ACP or C16:0 CoA substrates, and 500 ng of purified protein in Tris-HCl buffer (100 mM, pH 7.5) was incubated at 37 °C for 30 min. The reaction mixtures were stopped by adding 50 μL of ice-cold methanol and then centrifuged at 4 °C, 13,000 × *g* for 10 min. The supernatant was then filtered through a 0.2 μm filter (Millipore, Burlington, USA) before targeted LC-MS analysis.

Gene expression in yeast and enzyme assays

The purified PCR products of *OsCYP704B2*, *OsKCS11*, and *OsGPAT6* were subsequently cloned and inserted into the pESC-URA expression vector and sequenced for errors. The constructed *OsGPAT6* vectors were transformed into the yeast gat1Δ strain (EUROSCARF, Germany) and the *OsCYP704B2* and *OsKCS11* vectors were transformed into the yeast strain WAT11 via the lithium acetate method following the manufacturer's protocol (Zymo Research, catalog no. T2001). Yeast cultures were grown and microsomes were prepared as previously reported with some modification[41]. In particular, the recombinant cells were first cultured in SC minimal medium containing 2% glucose at 30 °C. For protein induction, the cells were collected and resuspended in synthetic complete medium yeast minimal medium containing 2% galactose instead of glucose and cultured at 30 °C for 2 days. The cells were harvested via centrifugation and broken with glass beads (0.45 mm in diameter, Sigma) in 50 mM Tri-HCl buffer, pH 7.5, containing 1 mM EDTA and 600 mM sorbitol. The cells were broken down via a mixing mill (model MM 400, Retsch). The homogenate was centrifuged for 60 min at 12,000 × *g* and the resulting supernatant was centrifuged for 90 min at 120,000 × *g*. The pellet consisting of microsomal membranes was resuspended in 100 mM Tris-HCl, pH 7.5, 1 mM EDTA and 20% (v/v) glycerol and stored at −80 °C for long-term storage.

The in vitro enzymatic assays for *OsGPAT6* were performed in a shaking incubator (120 r.p.m.) at 30 °C for 4 h in 100 μL of 20 mM Tris-HCl, pH 7.0, containing 1 mg of total microsomal proteins, 400 μM G3P, 1 mM DTT, 5 mM EDTA, 2.5 mg/mL BSA and 200 μM substrate. The in vitro enzymatic assays for *OsCYP704B2* were performed in a shaking incubator (120 r.p.m.) at 30 °C for 4 h in 100 μL of 20 mM Tris-HCl, pH 7.5, containing 1 mg of total microsomal proteins, 500 mM NADPH and 200 μM substrate. The in vitro enzymatic assays for *OsKCS11* were performed in a shaking incubator (120 r.p.m.) at 30 °C for 4 h in 100 μL of 100 mM Bis-Tris-propane, pH 7.6, containing 1 mg of total microsomal proteins, 10 mM $MgCl_2$, 5 mM ATP, 2.5 mM DTT, 1 mM CoA and 200 μM substrate. The reactions were stopped by the addition of 50 μL of ice-cold methanol and vortexing. Negative control reactions were carried out with microsomal preparations from recombinant yeast transformed with 'empty' pESC-URA. The total protein content was estimated by measuring ultraviolet absorbance at 280 nm on a NanoDrop ND-1000 spectrophotometer.

In vitro metabolic analysis of samples

The metabolic profiling analysis of the enzyme assays was performed using scheduled multiple reaction monitoring (MRM) via an LC-ESI-QQQ-MS/MS system (LCMS-8060, SHIMADZU, Japan). The UPLC analytical conditions were as follows: column, shim-pack GISS C18 (pore size 1.9 μm, length 2.1 × 100 mm); solvent system, water (10 mM ammonium acetate): acetonitrile; gradient program, 0 min, 20% B; 10.0 min, 100% B; 15.0 min, 100% B; 15.1 min, 20% B; 18.0 min, 20% B; flow rate, 0.3 mL/min; temperature, 32 °C; injection volume: 2 μL. The ESI source operation parameters were as follows: nebulizing gas flow, 3 L/min; heating gas flow, 10 L/min; interface temperature, 500 °C; DL temperature, 250 °C; heat block temperature, 400 °C; drying gas flow, 10 L/min. All of in vitro products were validated by comparison with authentic standards (see Supplementary Figs. 9 and 11).

Transport assays

Transport assays were performed as previously reported[28]. Briefly, leaves of 4-week-old *N. benthamiana* plants were agroinfiltrated with the following constructs: 35 *S*: *OsABCG22* and 35 *S*: *OsLTPL82*. After four days, the protoplasts were prepared, [^{13}C]-HMG16:0, [^{13}C]-C16:0(ω-OH), [^{13}C]-C16:0 or [$^{13}C_4$]-thiamine (control) stable isotope substrates were added to the protoplasts to yield final concentrations of 1 mM each. The mixture was incubated on ice for 15 min for substrate intake. The external metabolites were removed by separating the protoplasts using a 50%/25%/5% Percoll gradient built from bottom to top, and the protoplasts were recovered from the 25%/5% Percoll interphase. For the export experiment, the protoplasts were incubated at 25 °C and aliquots of 100 μL were taken at 0, 5, 10, 15 and 20 min. The aliquots were centrifuged and the supernatants were subjected to metabolite detection. The percentage of initial export was calculated as the metabolite content normalized to the first time point (0 min). The experiments were repeated at least three times, and the data are presented as the means with the standard error of the mean of all the technical replicates.

Genome-wide association study

Sequence data were obtained from the RiceVarMap website (http://ricevarmap.ncpgr.cn)[30]. In total, 4,300,150 high-quality SNPs with a minor allele frequency (MAF) > 0.05 and deletion rate <0.1 in the 533 rice cultivars were selected for GWAS. The GWAS analyses were performed via efficient mixed model association expedited (EMMAX)[42–44]. The kinship (K) matrix was constructed via the emmax-kin program.

RNA extraction and expression analyses

RNA extraction was performed using an RNA extraction kit (TRIzol reagent; TransGen) according to the manufacturer's instructions. Three micrograms of RNA was used to synthesize the first-strand cDNAs in a 20 μL of reaction mixture via the EasyScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen, Beijing, China). Real-time quantitative reverse transcription PCR (qRT-PCR) was performed with a SYBR Premix ExTaq kit (TaKaRa, Tokyo, Japan) on an ABI7500 Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). The rice *OsUBI13* gene was used as an internal reference. The primer sequences are listed in Supplementary Data 4.

Transient expression in *Nicotiana benthamiana*

The predicted full coding sequences of *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsMADS6* and *OsMADS17* were subsequently cloned and inserted into pEAQ-HT-DEST2 via a gateway system[45]. The expression constructs were subsequently introduced into *Agrobacterium tumefaciens EHA105* and cultured overnight. The cells were collected and redissolved in buffer containing 10 mM 2-morpholinoethanesulfonic acid monohydrate (Sigma Aldrich), 10 mM $MgCl_2$, and 200 mM acetosyringone (Sigma Aldrich) and adjusted to A_{600nm} = 1.0. After 2 h of incubation, the suspensions were pressure infiltrated into 4-week-old *N. benthamiana* leaves. The green fluorescent protein containing the pEAQ-HT expression construct was used as a negative control. The infiltrated *N. benthamiana* plants were grown in a growth chamber maintained at 25 °C with 16 h of daylight per day. *N. benthamiana* leaves were collected 72 h after infiltration and immediately frozen in liquid nitrogen. *N.*

benthamiana samples were using an approach that is similar to that described above for metabolite profiling.

Construction of mutant lines

The *osfar2*, *oskcs11*, *osgpat6*, *oscyp704b2*, *osabcg22*, *osltpl82* and *osmads17* mutant constructs were generated via CRISPR-Cas9 technology, and the *osmads6* mutant construct was generated via T-DNA insertion technology. The generated *OsFAR2*, *OsKCS11*, *OsCYP704B2*, *OsABCG22*, *OsLTPL82* and *OsMADS17* constructs were introduced into the calli of Zhonghua11 (ZH11), the *OsGPAT6* constructs were introduced into the calli of Minghui63 (MH63), and the *OsMADS6* constructs were introduced into the calli of Nipponbare (Nip) via *Agrobacterium tumefaciens*-mediated transformation[43].

Subcellular localization

The full-length coding sequences of *OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22*, *OsLTPL82*, *OsMADS6* and *OsMADS17* were amplified via Nipponbare total cDNA as a template, and the PCR product was cloned and inserted into the PJC032 (pH7WGF2.0) vector to generate an OsFAR2-GFP, OsKCS11-GFP, OsGPAT6-GFP, OsCYP704B2-GFP, OsABCG22-GFP, OsLTPL82-GFP, OsMADS6-GFP and OsMADS17-GFP fusion construct driven by the constitutive cauliflower mosaic virus promoter (35 S: OsFAR2-GFP, 35 S: OsKCS11-GFP, 35 S: OsGPAT6-GFP, 35 S: OsCYP704B2-GFP, 35 S: OsABCG22-GFP, 35 S: OsLTPL82-GFP, 35 S: OsMADS6-GFP and 35 S: OsMADS17-GFP). Then transiently expressed in *Nicotiana benthamiana* leaves with 35 S: OsGhd7-RFP via *A. tumefaciens*-mediated infiltration[46]. The fluorescence signal was observed via confocal laser scanning microscopy (FV1000; Olympus, Tokyo, Japan) at 2 d after transformation.

Transient dual luciferase reporter assay

The promoter sequences of *OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82* (2000 bp upstream of these genes) were amplified and cloned and inserted into the modified pH2GW7 vector containing the firefly luciferase (fLUC) gene and the Renilla luciferase gene (rLUC) as reporters, while the *OsMADS6* and *OsMADS17* cDNA was cloned and inserted into the pEAQ-HT-DEST2 vector as an effector. The plasmids were transferred into *A. tumefaciens EHA105* by electroporation and co-infiltrated into *N. benthamiana* leaves. The luciferase activities were measured via the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. The relative reporter gene expression levels are expressed as the ratio of firefly LUC to Renilla luciferase (LUC/REN). Five independent transformations were performed for each sample.

ChIP-qPCR

The ChIP experiment was carried out as previously reported[47]. Briefly, rice tissues (1 g) were cross-linked in 1% formaldehyde (F8775, Sigma-Aldrich) for 30 min, ground into powder in liquid nitrogen, and extracted with lysis buffer (10 mM Tris-HCl, pH 8.0, 0.4 M sucrose, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 10 mM $MgCl_2$, 5 mM β-mercaptoethanol, 1 protease inhibitor cocktail). Next, the mixture was cleaned by filtering through a double layer of Miracloth (Millipore) and centrifuged at 4000 × *g* at 4 °C for 20 min. The final chromatin precipitate was washed at least five times with washing buffer (0.25 M sucrose, 10 mM Tris, pH 8.0, 1% Triton X-100, 10 mM $MgCl_2$, 0.1 mM PMSF, 5 mM β-mercaptoethanol, 1 protease inhibitor cocktail) and suspended in 300 μL of nuclear lysis buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% [w/v] SDS). Then the chromatin was sonicated to 200–500 bp (20 μL of sonicated chromatin were taken out as input) and immune-precipitated with the anti-HA antibody (5 μg, ABclonal, AE008). The immunoprecipitated chromatin was washed sequentially with the following buffers: low salt buffer (150 mM NaCl, 50 mM HEPES-KOH, 1 mM EDTA, 0.1% sodium deoxycholate, 1% Triton X-100, 0.1% SDS), high salt buffer (same to low salt buffer except using 350 mM NaCl), LiCl buffer (10 mM Tris-HCl pH 8.0, 0.5% NP-40, 250 mM LiCl, 0.1% sodium deoxycholate, 1 mM EDTA) and TE buffer (1 mM EDTA, 10 mM Tris-HCl, pH 8.0), and eluted with elution buffer (10 mM EDTA, 50 mM Tris-HCl, pH 7.5, 1% SDS) at 65 °C for 15 min with shaking (800 rpm). The elutes were reverse-crosslinked with 5 M NaCl at 65 °C overnight, purified with chloroform, and precipitated with alcohol. The input and precipitated DNA fragments were checked by RT-qPCR using primers listed in Supplementary Data 4. Three independent biological replicates were performed.

Protein extraction and immunoblotting

Protein extraction was performed as previously reported[48]. Briefly, rice tissues were ground in liquid nitrogen and suspended in extraction buffer (10 mM EGTA, pH 7.5, 1% [v/v] Triton X-100, 10 mM Tris-HCl, 1 mM PMSF, 1 mM DTT, and 250 mM sucrose). The samples were centrifuged at 15,000 × *g* for 15 min at 4 °C and precipitated with cold acetone at −20 °C for 2 h. The protein pellets were collected by 15,000 × *g* centrifugation at 4 °C for 15 min and the precipitates were washed three times with cold acetone, lyophilized, and stored at −80 °C.

Immunoblotting was performed as previously reported[48]. Briefly, proteins were separated by 10–12% SDS-PAGE gel and transferred onto a PVDF membrane. The final blot was probed with anti-HA (1:1000, ABclonal, AE008) and anti-ACTIN antibodies (1:1000, Sanqon Biotech, D191048-0100).

One-hybrid assays in yeast

The full-length cDNA sequences of *OsMADS6* and *OsMADS17* were amplified and fused in frame with the GAL4 activation domain of pGADT7 (Clontech, Shiga, Japan), forming pGADT7-*OsMADS6* and pGADT7-*OsMADS17*. The 2 kb promoter regions of *OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82* were amplified and ligated into the pHIS2 reporter vector. Then, the fusion constructs were cotransformed into Y187 yeast cells (Clontech) grown on synthetic dropout (SD) -Trp/-Leu plates for 3 days at 30 °C and then transferred onto SD -Trp/Leu/-His media supplemented with 25 mM 3-amino-1,2,4-triazole. Yeast strains containing empty pGADT7 in combination with the pHIS2: *OsFAR2*, pHIS2: *OsKCS11*, pHIS2: *OsGPAT6*, pHIS2: *OsCYP704B22*, pHIS2: *OsABCG22* and pHIS2: *OsLTPL82* reporter constructs were used as negative controls.

Histochemical analysis

Histochemical analysis was performed as previously reported[49–51]. Briefly, the mature anther samples were obtained and fixed in 3.7% formaldehyde (5% acetic acid, 50% ethanol, 10% formaldehyde (37%), 35% ddH_2O) by pumping until all the anthers sank to the bottom, and then placed in a refrigerator at 4 °C for at least 48 h. The fixed samples were dehydrated with 70%, 80% and 95% ethanol for 30 min and anhydrous ethanol for 2 h. The samples were specifically stained for lipids using Sudan Red 7B. The samples were stained in 0.1% Sudan Red 7B solution (w/v isopropanol) for 60 min, followed by rinsing with distilled water. After drying and embedding in frozen section medium (OCT, Leica, Germany) and freezing at −25 °C, the samples were cut crosswise into 20 μm thick sections using a Leica microtome. The sections were placed on slides with 50% glycerol (v/v water) and immediately observed and photographed under a microscope.

Phenotypic characterization

Fresh rice spikelets, flowers, and dehiscent anthers were photographed with a LeicaM205A microscope. For pollen viability analysis, both wild-type and mutant anthers were immersed in I_2-KI solution (2% potassium iodide and 0.2% iodine in water) at room temperature and then photographed with a Nikon Eclipse 80i microscope.

Phylogenetic analysis and divergence time estimation

A total of six genes from *FGC3* (*OsFAR2*, *OsKCS11*, *OsGPAT6*, *OsCYP704B2*, *OsABCG22* and *OsLTPL82*) orthologous groups found between all eight species above were identified for phylogenetic analysis. Multiple sequence alignments of individual genes were performed via MAFFT v7.407 using the E-INS-I algorithm and then concatenated with a total length of 32,230 bp[52]. We generated a phylogeny on the alignment as a single partition using maximum likelihood method[53]. The support values of the tree were calculated by running 1000 bootstrap analyses. The relative rates of evolution between *P. latifolius* and core grasses were evaluated with *A. comosus* as the outgroup using the sequence alignment of the above six genes.

To estimate the origin and evolutionary timescale of grasses, we calibrated a relaxed molecular clock with Poaceae and Pooideae phytolith-based age constraints for the phylogeny, and the root age of the tree was set to a maximum age of 150 Ma. The divergence time was estimated using the MCMCtree program in the PAML package v4.9h[54]. The Markov chain Monte Carlo process was run for 2 million iterations with samples drawn every 100 steps following a 10% of the iterations were used as burn-in[36].

Geographic distribution of humidity data

The geographic distribution of the average annual relative humidity in China was obtained from the website of the China Meteorological Data Network (http://data.cma.cn/en)[55].

Reporting summary

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

Data availability

The metabolomics data generated in this study have been deposited in the OMIX, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number OMIX006912. Source data are provided with this paper.

References

1. Nutzmann, H. W., Scazzocchio, C. & Osbourn, A. Metabolic gene clusters in eukaryotes. *Annu. Rev. Genet.* **52**, 159–183 (2018).
2. Zhan, C. S. et al. Plant metabolic gene clusters in the multi-omics era. *Trends Plant Sci.* **27**, 981–1001 (2022).
3. Darbani, B. et al. The biosynthetic gene cluster for the cyanogenic glucoside dhurrin in *Sorghum bicolor* contains its co-expressed vacuolar MATE transporter. *Sci. Rep.* **6**, 37079 (2016).
4. Takos, A. M. & Rook, F. Why biosynthetic genes for chemical defense compounds. *Trends Plant Sci.* **17**, 383–388 (2012).
5. Nutzmann, H. W. et al. Active and repressed biosynthetic gene clusters have spatially distinct chromosome states. *Proc. Natl. Acad. Sci. USA* **117**, 13800–13809 (2020).
6. Wan, X. Y., Wu, S., Li, Z. W., An, X. L. & Tian, Y. H. Lipid metabolism: critical roles in male fertility and other aspects of reproductive development in plants. *Mol. Plant* **13**, 1–29 (2020).
7. Philippe, G. et al. Cutin and suberin: assembly and origins of specialized lipidic cell wall scaffolds. *Curr. Opin. Plant Biol.* **55**, 11–20 (2020).
8. Zhao, L. F., Haslam, T. M., Sonntag, A., Molina, I. & Kunst, L. Functional overlap of long-chain acyl-CoA synthetases in Arabidopsis. *Plant Cell Physiol.* **60**, 1041–1054 (2019).
9. Fich, E. A., Segerson, N. A. & Rose, J. K. C. The plant polyester cutin: biosynthesis, structure, and biological roles. *Annu. Rev. Plant Biol.* **67**, 18.11–18.27 (2016).
10. Yang, W. L. et al. A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, *sn*-2 preference, and evolution. *Plant Physiol.* **160**, 638–652 (2012).
11. Zhu, T. T. et al. Normal structure and function of endothecium chloroplasts maintained by ZmMs33-mediated lipid biosynthesis in tapetal cells are critical for anther development in maize. *Mol. Plant* **13**, 1624–1643 (2020).
12. Pollard, M., Beisson, F., Li, Y. H. & Ohlrogge, J. B. Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* **13**, 236–246 (2008).
13. Dominguez, E., Heredia-Guerrero, J. A. & Heredia, A. Plant cutin genesis: unanswered questions. *Trends Plant Sci.* **20**, 551–558 (2015).
14. Li, H. et al. Cytochrome P450 family member CYP704B2 catalyzes the ω-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell* **22**, 173–190 (2010).
15. Shi, J. et al. Defective pollen wall is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. *Plant Cell* **23**, 2225–2246 (2011).
16. Shi, J. X., Cui, M. H., Yang, L., Kim, Y. J. & Zhang, D. B. Genetic and biochemical mechanisms of pollen wall development. *Trends Plant Sci.* **20**, 741–753 (2015).
17. Berhin, A. et al. The root cap cuticle: a cell wall structure for seedling establishment and lateral root formation. *Cell* **176**, 1367–1378.e1368 (2019).
18. Kannangara, R. et al. The transcription factor WIN1/SHN1 regulates Cutin biosynthesis in *Arabidopsis thaliana*. *Plant Cell* **19**, 1278–1294 (2007).
19. Shi, J. X. et al. SHINE transcription factors act redundantly to pattern the archetypal surface of Arabidopsis flower organs. *PLoS Genet.* **7**, e1001388 (2011).
20. Nadakuduti, S. S. et al. Pleiotropic phenotypes of the sticky peel mutant provide new insight into the role of *CUTIN DEFICIENT2* in epidermal cell function in tomato. *Plant Physiol.* **159**, 945–960 (2012).
21. Wu, R. H. et al. CFL1, a WW domain protein, regulates cuticle development by modulating the function of HDG1, a class IV homeodomain transcription factor, in rice and Arabidopsis. *Plant Cell* **23**, 3392–3411 (2011).
22. Liu, X. Z. et al. The ZmMYB84-*ZmPKSB* regulatory module controls male fertility through modulating anther cuticle-pollen exine trade-off in maize anthers. *Plant Biotechnol. J.* **20**, 2342–2356 (2022).
23. Ohmori, S. et al. *MOSAIC FLORAL ORGANS1*, an *AGL6*-like MADS box gene, regulates floral organ identity and meristem fate in rice. *Plant Cell* **21**, 3008–3025 (2009).
24. Chen, W. et al. A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* **6**, 1769–1780 (2013).
25. Yang, C. K. et al. Rice metabolic regulatory network spanning the entire life cycle. *Mol. Plant* **15**, 258–275 (2022).
26. Kawaguchi, A., Yoshimura, T. & Okuda, S. A new method for the preparation of acyl-CoA thioesters. *J. Biochem.* **89**, 337–339 (1981).
27. Aryal, B. et al. ABCG36/PEN3/PDR8 is an exporter of the auxin precursor, indole-3-butyric acid, and involved in auxin-controlled development. *Front. Plant Sci.* **10**, 899 (2019).
28. Elejalde-Palmett, C. et al. ABCG transporters export cutin precursors for the formation of the plant cuticle. *Curr. Biol.* **31**, 1–22 (2021).
29. Xue, Z. Y. et al. Deficiency of a triterpene pathway results in humidity-sensitive genic male sterility in rice. *Nat. Commun.* **9**, 604 (2018).
30. Zhao, H. et al. RiceVarMap: a comprehensive database of rice genomic variations. *Nucleic Acids Res.* **43**, D1018–1022, (2015).
31. Graca, J., Schreiber, L., Rodrigues, J. & Pereira, H. Glycerol and glyceryl esters of ω-hydroxyacids in cutins. *Phytochemistry* **61**, 205–215 (2002).

32. Shen, S. Q. et al. Elucidation of the melitidin biosynthesis pathway in pummelo. *J Integr. Plant Biol* **65**, 2505–2518 (2023).
33. Hen-Avivi, S. et al. A metabolic gene cluster in the Wheat *W1* and the Barley *Cer-cqu* Loci determines β-Diketone biosynthesis and glaucousness. *Plant Cell* **28**, 1440–1460 (2016).
34. Jeon, J. E. et al. A pathogen-responsive gene cluster for highly modified fatty acids in tomato. *Cell* **180**, 176–187.e119 (2020).
35. Maeda, H. A. Evolutionary diversification of primary metabolism and its contribution to plant chemical diversity. *Front. Plant Sci.* **10**, 881 (2019).
36. Erb, M. & Kliebenstein, D. J. Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. *Plant Physiol.* **184**, 39–52 (2020).
37. Ma, P. F. et al. The Pharus latifolius genome bridges the gap of early grass evolution. *Plant Cell* **33**, 846–864 (2021).
38. Ni, E. et al. OsCER1 regulates humidity-sensitive genic male sterility through very-long-chain (VLC) alkane metabolism of tryphine in rice. *Funct. Plant Biol.* **48**, 461–468, (2021).
39. 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).
40. Shen, S. Q. et al. An Oryza specific hydroxycinnamoyl tyramine gene cluster contributes to enhanced disease resistance. *Sci. Bull.* **66**, 2369–2380 (2021).
41. Zhan, C. S. et al. Selection of a subspecies-specific diterpene gene cluster implicated in rice disease resistance. *Nat. Plants* **6**, 1447–1454 (2020).
42. Li, Y. F. et al. The *NET* locus determines the food taste, cooking and nutrition quality of rice. *Sci. Bull.* **67**, 2045–2049 (2022).
43. Zhan, C. S. et al. Disease resistance conferred by components of essential chrysanthemum oil and the epigenetic regulation of *OsTPS1*. *Sci. China Life Sci.* **66**, 1108–1118 (2022).
44. Zhu, G. T. et al. Rewiring of the fruit metabolome in tomato breeding. *Cell* **172**, 249–261.e212 (2018).
45. Sainsbury, F., Thuenemann, E. C. & Lomonossoff, G. P. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693 (2009).
46. Sun, Y. Y. et al. Natural variation in the *OsbZIP18* promoter contributes to branched-chain amino acid levels in rice. *New Phytol.* **228**, 1548–1558 (2020).
47. Cheng, S. F. et al. WOX11 recruits a histone H3K27me3 demethylase to promote gene expression during shoot development in rice. *Nucleic Acids Res.* **46**, 2356–2369 (2018).
48. He, D. et al. Global proteome analyses of lysine acetylation and succinylation reveal the widespread involvement of both modification in metabolism in the embryo of germinating rice seed. *J. Proteome Res.* **15**, 879–890 (2016).
49. Brundrett, M. C., Kendrick, B. & Peterson, C. A. Efficient lipid staining in plant material with sudan red 7B or fluorol yellow 088 in polyethylene glycol-glycerol. *Biotech. Histochem.* **66**, 111–116 (1991).
50. Hofer, R. et al. The Arabidopsis cytochrome P450 *CYP86A1* encodes a fatty acid ω-hydroxylase involved in suberin monomer biosynthesis. *J. Exp. Bot.* **59**, 2347–2360 (2008).
51. Xu, D. et al. Defective Pollen Wall 2 (*DPW2*) Encodes an acyl transferase required for rice pollen development. *Plant Physiol.* **173**, 240–255 (2017).
52. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Briefings Bioinf.* **9**, 286–298 (2008).
53. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
54. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
55. Guo, J. & Chen, S. Long-term change in relative humidity across China from 1961-2018. *Clim. Res.* **87**, 167–181 (2022).

Acknowledgements

We thank Professor Wan Xiangyuan from the University of Science and Technology Beijing for offering the yeast gat1Δ strain. This work was supported by the Key project of regional joint fund of National Natural Science Foundation (U22A20476 to J.L.), the National Natural Science Foundation of China (32100318 to C.Z. and 32302664 to Y.S.), the Project of Sanya Yazhou Bay Science and Technology City (SCKJ-JYRC-2022-06 to J.L.), the "111" Project (No. D20024 to J.L.), the Postdoctoral Fellowship Program of CPSF (GZB20230187 to C.Y.), the National Postdoctoral Program for Innovative Talents (BX20220097 to S.S.), the China Postdoctoral Science Foundation (2022M710991 to S.S., 2023M730896 to C.Y. and 2023M740955 to Y.L.), the Hainan Provincial Natural Science Foundation of China (321QN184 to C.Z., 323MS019 to S.S., 324QN196 to J.Z. and 324QN209 to C.Y.), the Hainan Provincial Academician Innovation Platform Project (HD-YSZX-202003 to J.L.), and the Hainan University Startup Fund (KYQD(ZR) 1866 to J.L.), the Hainan Province Science and Technology Talent Innovation Project (KJRC2023D10 to Y.L.).

Author contributions

J.L. designed the research and supervised this study. C.Y., S.S., C.Z., Y.L., R.Z., Y-Y.L., Z.Y., J.Z., Y.S., and X.L. performed experiments. C.Y., S.S., C.Z., and Y.L. analyzed data. J.S., D.Z., and A.R.F. contributed valuable suggestions on this study. J.L., C.Y., S.S., and A.R.F. wrote the manuscript. All the authors edited and proofed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51145-8.

Correspondence and requests for materials should be addressed to Jie Luo.

Peer review information *Nature Communications* thanks Ziqiang Liu, Guy Polturak and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.