





ARTICLE

<https://doi.org/10.1038/s41467-019-10544-y>

OPEN

# Variation of a major facilitator superfamily gene contributes to differential cadmium accumulation between rice subspecies

Huili Yan <sup>1,6</sup>, Wenxiu Xu<sup>1,6</sup>, Jianyin Xie<sup>2,6</sup>, Yiwei Gao <sup>3</sup>, Lulu Wu<sup>1,4</sup>, Liang Sun<sup>5</sup>, Lu Feng<sup>1</sup>, Xu Chen<sup>1,4</sup>, Tian Zhang<sup>1,4</sup>, Changhua Dai<sup>1,4</sup>, Ting Li <sup>1,4</sup>, Xiuni Lin<sup>3</sup>, Zhanying Zhang<sup>2</sup>, Xueqiang Wang<sup>2</sup>, Fengmei Li<sup>2</sup>, Xiaoyang Zhu<sup>2</sup>, Jinjie Li<sup>2</sup>, Zichao Li <sup>2</sup>, Caiyan Chen<sup>5</sup>, Mi Ma<sup>1</sup>, Hongliang Zhang<sup>2</sup> & Zhenyan He<sup>1</sup>

Cadmium (Cd) accumulation in rice grain poses a serious threat to human health. While several transport systems have been reported, the complexity of rice Cd transport and accumulation indicates the necessity of identifying additional genes, especially those that are responsible for Cd accumulation divergence between *indica* and *japonica* rice subspecies. Here, we show that a gene, *OsCd1*, belonging to the major facilitator superfamily is involved in root Cd uptake and contributes to grain accumulation in rice. Natural variation in *OsCd1* with a missense mutation Val449Asp is responsible for the divergence of rice grain Cd accumulation between *indica* and *japonica*. Near-isogenic line tests confirm that the *indica* variety carrying the *japonica* allele *OsCd1*<sup>V449</sup> can reduce the grain Cd accumulation. Thus, the *japonica* allele *OsCd1*<sup>V449</sup> may be useful for reducing grain Cd accumulation of *indica* rice cultivars through breeding.

<sup>1</sup>Key Laboratory of Plant Resources, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China. <sup>2</sup>Key Lab of Crop Heterosis and Utilization of Ministry of Education, Beijing Key Lab of Crop Genetic Improvement, China Agricultural University, Beijing 100193, China. <sup>3</sup>College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing 100083, China. <sup>4</sup>University of Chinese Academy of Sciences, Beijing 100049, China. <sup>5</sup>Key Laboratory of Agro-Ecological Processes in Subtropical Region, Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, China. <sup>6</sup>These authors contributed equally: Huili Yan, Wenxiu Xu, Jianyin Xie. Correspondence and requests for materials should be addressed to H.Z. (email: [zhangl@cau.edu.cn](mailto:zhangl@cau.edu.cn)) or to Z.H. (email: [hezhenyan@ibcas.ac.cn](mailto:hezhenyan@ibcas.ac.cn))

Cadmium (Cd) is a toxic heavy metal and can lead to Cd-related diseases such as renal tubular dysfunction and bone disease. Rice is a major component of diet for over half of the world's population. The accumulation of Cd is a serious threat to human being since it can be concentrated in body through the food chain<sup>1</sup> and the biological half-life is estimated to be nearly 30 years<sup>2</sup>. Molecular genetic tools have been urgently sought to develop low-Cd rice to reduce potential health risks.

Generally, Cd is first absorbed from soil by roots, and then translocated into shoots, and accumulates into grains in final<sup>3</sup>. Transport systems, especially the genetic components locating on the membrane, play crucial roles in the Cd accumulation processes in rice. In root exodermis and endodermis, OsNramp1<sup>4,5</sup> and OsNramp5<sup>6–11</sup>, which locate on the plasma membrane, are transporters for Cd uptake. The heavy metal ATPase, OsHMA3<sup>12–15</sup>, functions in vacuolar sequestration of Cd. In root stele, OsHMA2<sup>16–18</sup>, an efflux-type metal transporter on the plasma membrane, is suggested to facilitate Cd loading into the xylem. CAL1, a defensin-like protein preferentially expressed in root exodermis and xylem parenchyma cells, drives long-distance Cd transport via xylem vessels<sup>19</sup>. And in shoot node, OsLCT1<sup>20,21</sup>, a plasma membrane-localized efflux transporter, is responsible for Cd intervascular transfer.

By over-expression of the functional OsHMA3 or the disruption of OsLCT1 and OsHMA2, the grain Cd content can be reduced to some extent, which gives a clue that rice grain Cd contents can be reduced through regulating transporters expression. However, the rice Cd transport mechanism is complex and grain Cd accumulation is a result of cooperative interactions among multiple cells and issues. It is, therefore, worthwhile to explore more genetic loci involving in Cd accumulation of rice grain.

Rice varieties exhibit substantial genetic variation with respect to Cd accumulation ability<sup>22–24</sup>, which is a valuable resource for dissecting functional alleles and genetic improvement. OsHMA3 shows different Cd transport ability<sup>13</sup> and OsNRAMP1 expresses in a different level among two cultivars<sup>4</sup>. By using chromosome segment substitution lines (CSSLs) carrying segments of *indica* rice Kasalath in a background of the *japonica* rice Koshihikari, several CSSLs have significantly lower Cd concentrations than that of Koshihikari<sup>25</sup>. However, current understanding of the genetic basis of rice Cd accumulation diversity remains at the level of the identification of several quantitative trait loci (QTLs)<sup>25–27</sup>. The natural allelic variations responsible for rice varietal differences have not been fully explored and the genetic basis of grain Cd accumulation differences remains unknown.

More recently, genome-wide association studies (GWAS) have been successfully used for identifying genes and alleles underlying several agronomic traits<sup>28–32</sup>, while the association study of rice grain Cd accumulation has not been attempted. Here, we show that a gene belonging to major facilitator superfamily, *OsCd1*, is associated with divergence in rice grain Cd accumulation. *OsCd1* locates at the plasma membrane of root and is involved in the grain Cd accumulation. Additionally, a single-nucleotide mutation, SNP22, diverges between *indica* and *japonica* and alters Cd transport ability of *OsCd1*. Interestingly, the natural variation *OsCd1*<sup>V449</sup> in *japonica*, which is associated with a reduced Cd transport ability and decreased grain Cd accumulation, shows a potential value in low-Cd rice selection.

## Results

### *OsCd1* is associated with grain cadmium accumulation in rice.

In this study, we combined data from GWAS, gene annotation in GO Slim database and the yeast assay to look for genes associated

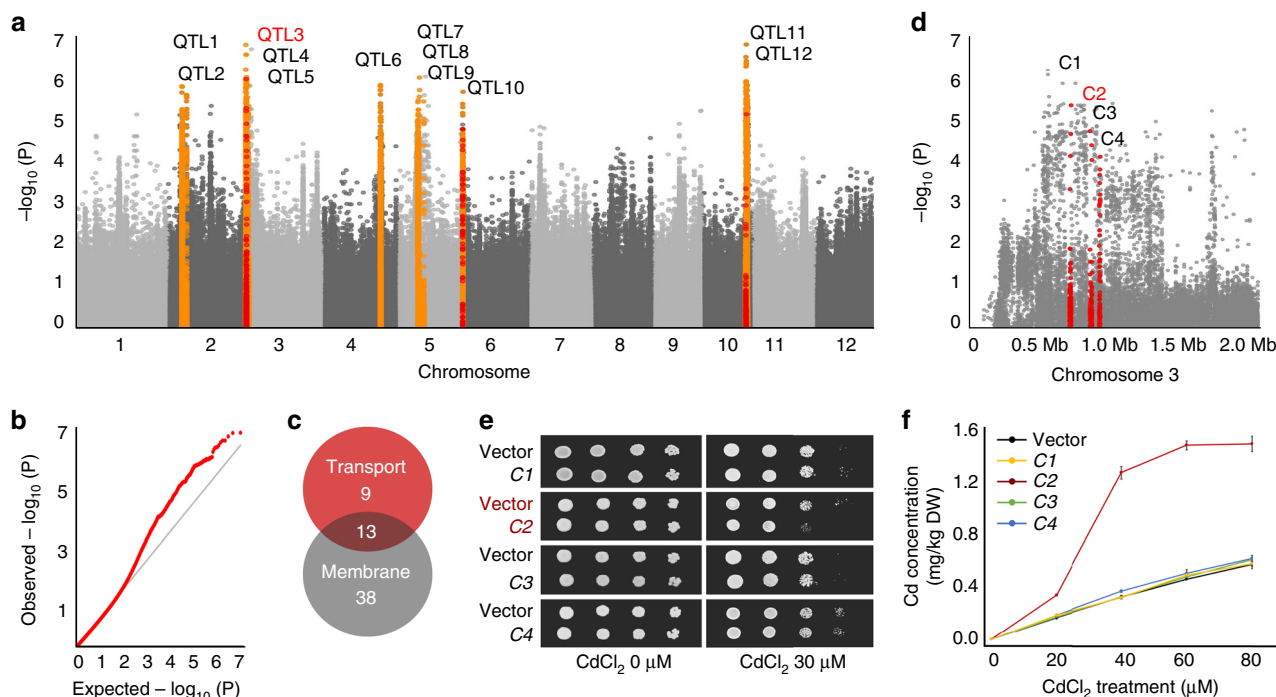
with rice grain Cd accumulation. A set of 127 rice cultivars including 41 *japonica* and 86 *indica* from mini-core collection of rice in China<sup>33</sup> and other wide-spread regions was used for GWAS of grain Cd concentration (Supplementary Fig. 1a, b). After testing rice varieties with a Cd treatment assay, we found that *indica* varieties could be phenotypically distinguished from *japonica* varieties by their significantly higher grain Cd accumulation (Supplementary Fig. 1c and Supplementary Data 1). Using 3,291,150 single-nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) >0.05 covering the whole rice genome, we performed GWAS to identify the genetic loci associating with grain Cd accumulation. Under the compressed general mixed linear model ( $P < 1 \times 10^{-5}$ , MLM, threshold derived from 1000 permutation test) (Fig. 1a, b), 12 QTLs were identified as significantly associated with grain Cd accumulation (Supplementary Table 1). According to the rice genome annotation project website (MSU-RGAP), 494 genes were annotated to be located in the candidate locus (Supplementary Data 2). To identify the membrane transporter conferring Cd accumulation in rice, we then performed the gene ontology (GO) Slim analysis to select candidate genes associated with the keyword membrane and transport for further research. Altogether, 13 candidate genes located on chromosome 2, 3, 4, 6, and 10 were found to be associated both with transport and membrane in GO Slim annotation (Fig. 1c and Supplementary Data 3). Four candidate genes were selected on QTL3 of chromosome 3, which explained ~20.7% of the phenotypic variation (Fig. 1d), and transformed into *S. cerevisiae* to evaluate whether these candidate genes were involved in Cd transport.

Spot assays were tested with and without 30  $\mu$ M CdCl<sub>2</sub> and the growth rate was normalized by non-transgenic strains. As a result, only yeast transformed C2 (*LOC\_Os03g02380*; GenBank accession AP014959.1) displayed sensitivity to Cd (Fig. 1e). Under 24 h Cd exposure, expression of C2 enhanced Cd accumulation in a Cd dose-dependent manner while the other three showed no significant difference with the control (Fig. 1f). These results indicated that C2 was the most likely candidate transporter gene in QTL3 and it was christened *OsCd1*.

*OsCd1* encoded a protein belonging to the MFS with 12 bilayer-spanning domains and the signature motif of MFS, G(93)slaD(97)kqG(100)rkR(103)<sup>17</sup>, located in the second cytoplasmic loop (Supplementary Figure 2a, b). There were 149 MFS superfamily members in rice and several of them were predicted to transport various substrates (Supplementary Data 4). A phylogenetic tree was constructed for *OsCd1* along with all the MFS proteins in rice (Supplementary Fig. 3). It was clearly shown that the proteins facilitating the same substrate transport shared a closer phylogenetic relationship, while *OsCd1* clusters separately formed an unknown-function clade. Together with the yeast spot assay results, we deduced that *OsCd1* may form a new-function clade in MFS involving in Cd transport.

### *OsCd1* altered cadmium uptake and grain accumulation in rice.

Considering the possible function of *OsCd1* in Cd transport, we transformed the *CRISPR/Cas9* constructs in rice callus and regenerated the transgenic plants to investigate the role of *OsCd1* in rice. Three Cas9-positive lines (*oscd1-cr1*, *oscd1-cr2*, and *oscd1-cr3*), which had either an insertion or deletion of one or few bases at the target sequences and disrupted the Cd transport ability of *OsCd1* (Supplementary Fig. 4), were selected for further research (Fig. 2a). In vegetative growth stage, the growth without CdCl<sub>2</sub> treatment was not affected by the mutation of *OsCd1*. After being treated with 1  $\mu$ M CdCl<sub>2</sub> for 20 days, the *CRISPR-oscd1* lines displayed a better growth (Fig. 2b) and the Cd concentration both in root and shoot were much lower than that in the wild-type rice (Fig. 2c, d). We then grew them in the field until



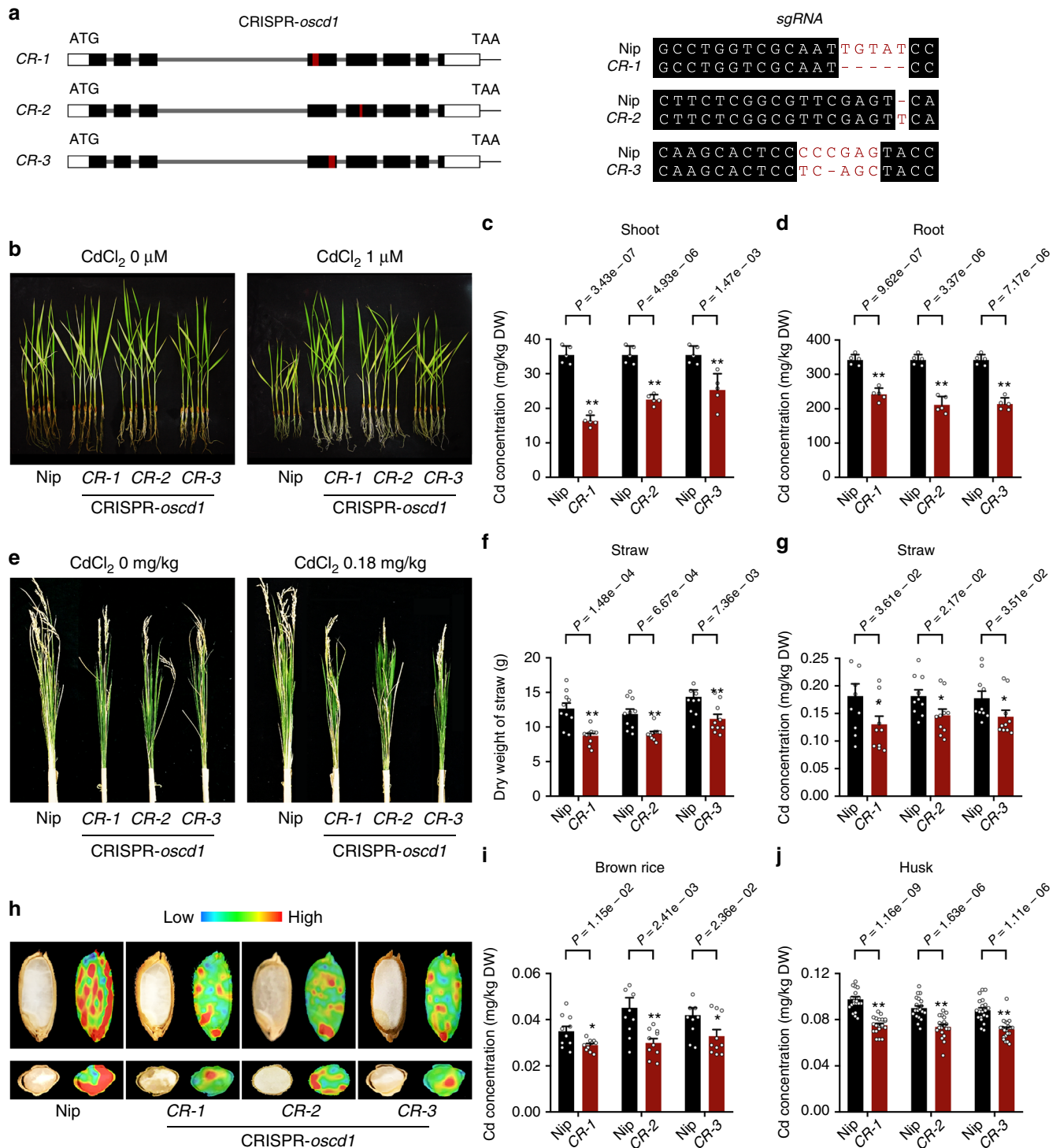
**Fig. 1** GWAS of the grain Cd accumulation. **a** Manhattan plots for grain Cd accumulation in diverse population. The orange pots indicated the 12 QTLs were identified as significantly associated with grain Cd accumulation. The red dots indicated the 14 candidate genes associated both with transport and membrane in GO Slim annotation. **b** The QQ plots for GWAS MLM model. **c** The numbers of genes annotated with transport and membrane. **d** The genome-wide association signals for grain Cd accumulation were shown in the region of 0–2 Mb on chromosome 3 (x-axis). The locations of four candidate genes *C1* (*LOC\_Os03g02150*), *C2* (*LOC\_Os03g02380*), *C3* (*LOC\_Os03g02390*) and *C4* (*LOC\_Os03g02480*) were indicated with red pot respectively. **e** Dilution-series spot assays of yeast treated with and without 30  $\mu\text{M}$   $\text{CdCl}_2$ . **f** Cd accumulation in wild-type W303 (black) and W303 expressing *C1* (yellow), *C2* (red), *C3* (green) and *C4* (blue) treated with  $\text{CdCl}_2$  for 24 h. Error bars indicate standard deviation. Data from the transgenic lines were designed with three replications and data points for all biological replicates are shown. Source data of Fig. 1f are provided as a Source Data file

ripening to investigate the physiological role of *OsCdl* in grain Cd accumulation. At harvest, both the growth and the yield were affected by the disruption of *OsCdl* (Fig. 2e). The biomass of the straw (Fig. 2f) and the filled spikelet were reduced in the mutant lines due to decreased fertility (Supplementary Fig. 5). As to the Cd concentration, the result showed that there was a decrease Cd accumulation in the straw of three *CRISPR-osc1* lines compared with the wild type (Fig. 2g). The synchrotron radiation microscopic X-ray fluorescence (SR- $\mu\text{XRF}$ ) scanning was further used to reveal the spatial disparity in the distribution of Cd throughout the grain in situ. As it is shown in Fig. 2h, Cd enriched both in the endosperm and aleurone layer, and its concentration was in a lower level in *CRISPR-osc1* lines than the wild type. Moreover, the Cd concentration was significantly decreased in brown rice and husk in three *CRISPR-osc1* lines, which was in well agreement with the SR- $\mu\text{XRF}$  results above (Fig. 2i, j). The concentration of other micronutrients, including Zn, Mn, and Fe, were also detected in the *CRISPR-osc1* lines. Compared with the wild-type rice, no difference in the concentration of Zn and Fe was observed but Mn, which is important in photosynthetic oxygen evolution in chloroplasts of plants, displayed a significant decrease in the *CRISPR-osc1* lines (Supplementary Fig. 6).

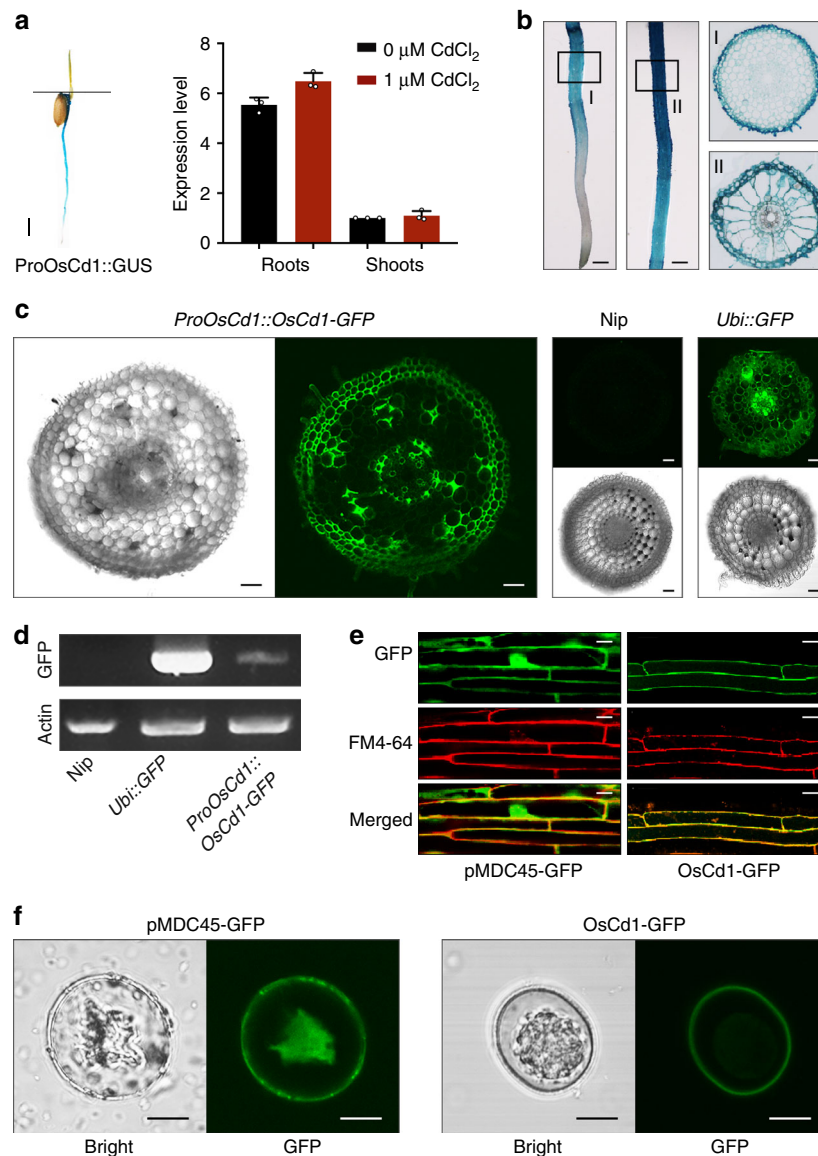
To determine the expression profile of *OsCdl* in rice, we analyzed the expression level from root and shoot at vegetative growth stage in rice via qPCR. The results showed that *OsCdl* was mainly detected in roots, which was consistent with the GUS histochemical assay, and not induced by  $\text{CdCl}_2$  (Fig. 3a). The cell-specific expression of *OsCdl* in rice root tissues was further analyzed using the GUS and GFP reporter fused to the *OsCdl* promoter. The promoter activity of *OsCdl* was mainly detected in

the root exodermis, parenchyma in cortex, endodermis and stele cells (Fig. 3b–d and Supplementary Fig. 7). We then constructed the GFP fusion protein to examine the subcellular location of *OsCdl* in rice. The green fluorescence was mainly localized in the cytosol and nucleus of cells expressing GFP alone while the *OsCdl* fused with GFP was only observed at the periphery of the cells both in root cells and protoplast. The merged images of GFP and plasma membrane marker FM4-64 further confirmed the plasma membrane subcellular localization of *OsCdl* (Fig. 3e, f). Together with the *CRISPR-osc1* results, it is suggested that the *OsCdl*, which was a plasma membrane protein in root, may be involved in Cd uptake and contributed to grain Cd accumulation in rice.

**SNP22 altered the cadmium transport ability of *OsCdl*.** The full-length sequences of 127 rice cultivars were used to investigate functional allelic variations in *OsCdl* locus. A total of 24 SNPs were identified in the *OsCdl* genomic region: 3 SNPs in the 5'UTR, 3 SNPs in the exon, 17 SNPs in the intron and 1 SNP in the 3'UTR, respectively (Supplementary Data 5). In specific, only the SNP22 (g. 846037 T>A), a missense mutation in eighth exon, resulted in a negative amino acid valine corresponding to a neutral one aspartic (a p. Val449Asp substitution) (Fig. 4a). *OsCdl* was a typical MFS protein with 12 transmembrane  $\alpha$ -helices. Additionally, there were three extra intracellular helical domains (ICHs) on the intracellular side, one at the C terminus (ICH1) and the other two (ICH2 and ICH3) locating between the amino- and C-terminal TM bundles. The amino acid substitution V449D located on the C-terminal and around the ICH1 from the cytoplasmic face of the plasma membrane (Fig. 4b).



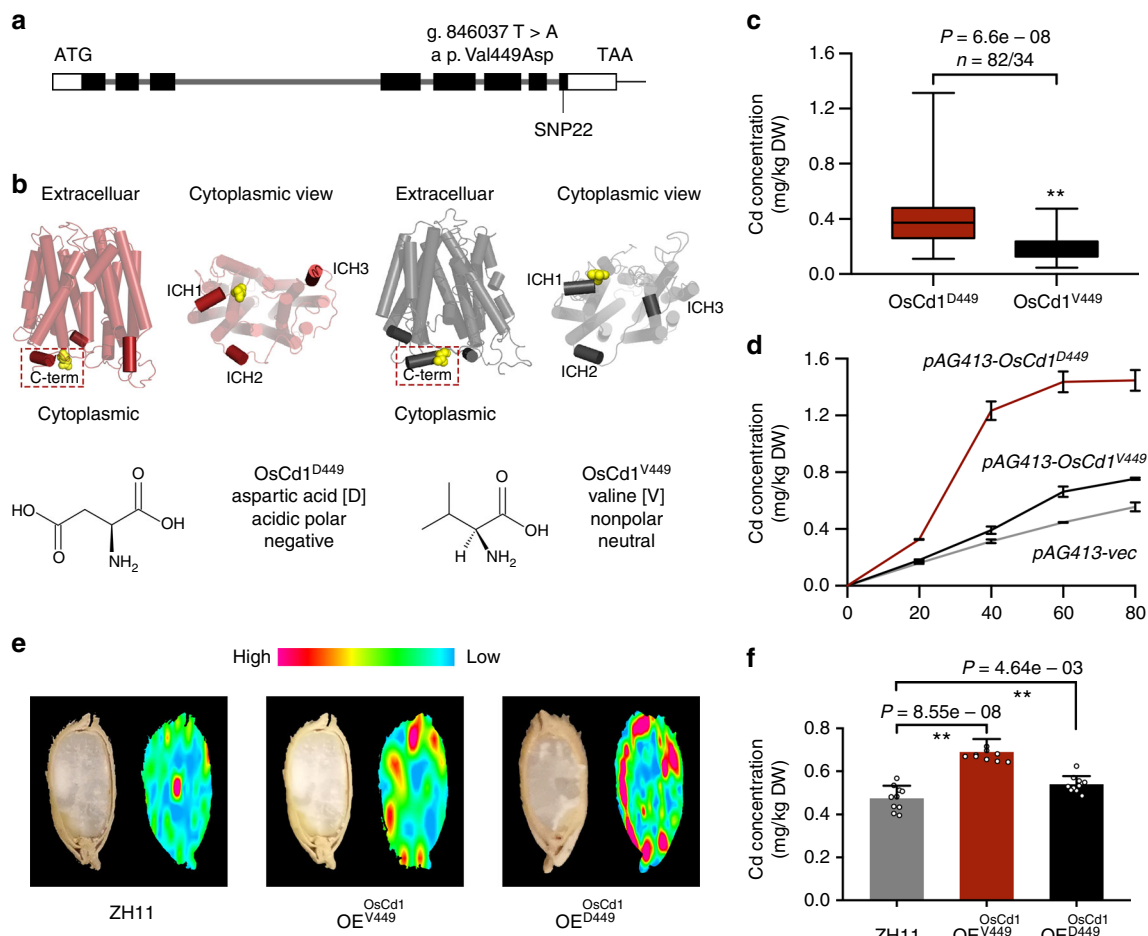
**Fig. 2** *OsCd1* contributes to the Cd uptake and grain accumulation in rice. **a** Left: Schematic diagram of three *CRISPR-OsCd1* lines. Black rectangles represented exons of *OsCd1*, red rectangle represented the exon of target sequence. Right: Sequences of *CRISPR-OsCd1* alleles *CR-1*, *CR-2*, and *CR-3*. sgRNA targets sequences were showed and deletions are indicated by red dashes. **b** Growth of wild-type rice and three *CRISPR-OsCd1* lines after treating with and without 1  $\mu\text{M}$   $\text{CdCl}_2$  for 20 days. **c**, **d**, **g**, **i**, **j** Concentration of Cd in the shoot (**c**), root (**d**), straw (**g**), brown rice (**i**) and husk (**j**). **e** Growth of wild-type rice and three *CRISPR-OsCd1* lines after treated with and without 0.18  $\text{mg kg}^{-1}$  Cd in pot-test. **f** Dry weight of straw wild-type rice and three *CRISPR-OsCd1* lines after treated with 0.18  $\text{mg kg}^{-1}$  Cd in pot-test. **h** SR- $\mu\text{XRF}$  images of Cd distribution in the longitudinal (upper) and latitudinal (lower) sections of rice grain. The emission intensity of each pixel was normalized using the beam intensity as reference. The *CRISPR-OsCd1* lines were shown in red and the wild-type line was shown in black. Error bars indicate standard deviation in the hydroponic experiment and standard error of mean in pot-test. Statistical comparison was performed by one-side t-test ( $*P < 0.05$  and  $**P < 0.01$ ). All data were compared with Nipponbare and designed with at least five replications and data points for all biological replicates are shown. Source data of Fig. 2c, d, f, g, i, and j are provided as a Source Data file



**Fig. 3** Expression pattern of *OsCd1* in rice. **a** Tissue-dependent expression of *OsCd1* at vegetative stage. Left: GUS histochemical assay of the transgenic plant with the GUS reporter driven by the *OsCd1* promoter. Right: the expression level of *OsCd1* in root and shoot with and without 1  $\mu\text{M}$   $\text{CdCl}_2$  treatment via qPCR. Bars = 2 mm. **b** Tissue-specific localization of *OsCd1* in the root of rice using the GUS reporter driven by the *OsCd1* promoter. Bars = 0.5 mm. **c** Tissue-specific localization of *OsCd1* in the root of rice using the GFP driven by the *OsCd1* promoter in cross frozen sections. Nip, negative control; Ubi::GFP, positive control. Bars = 10  $\mu\text{m}$ . **d** The GFP transcriptional level shown by RT-PCR. Upper, GFP; Down, Actin. **e** Subcellular location of *OsCd1* in the root of rice observed by confocal laser scanning microscopy. The GFP fluorescence, fluorescence of FM4-64 and overlay of FM4-64 and pMDC45-GFP (left), *OsCd1*-GFP (right) were shown, respectively. Bars = 20  $\mu\text{m}$ . **f** Subcellular location of *OsCd1* in root protoplasts. Bars = 10  $\mu\text{m}$ . Source data of Fig. 3a and d are provided as a Source Data file

To find the haplotype between SNP22 with rice grain Cd accumulation, we then conducted the haplotype analysis for SNP22 and grain Cd accumulation in 127 rice cultivars. The results showed that the *OsCd1* can be classified into two genotypes based on the SNP22: *OsCd1*<sup>D449</sup> with the A on SNP22 ( $n = 82$ ) and *OsCd1*<sup>V449</sup> with the T on SNP22 ( $n = 34$ ). Statistically, lines with *OsCd1*<sup>D449</sup> have an  $\sim 2$ -folds grain Cd concentration compared to those with *OsCd1*<sup>V449</sup> ( $P = 6.68 \times 10^{-8}$ ) (Fig. 4c) but with no obvious difference in expression levels and subcellular location (Supplementary Fig. 8 and 9). We further analyzed the Cd concentration in *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> transgenic yeast strains to investigate if the variation caused by SNP22 might alter the Cd transport ability of *OsCd1*. As shown in Fig. 4d and Supplementary Fig. 10,

*OsCd1*<sup>D449</sup> resulted in a higher Cd concentration compared to *OsCd1*<sup>V449</sup> under Cd treatment. Meanwhile, the expression level showed that there was no significant difference between the *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> transgenic lines with or without Cd treatment (Supplementary Fig. 11). To determine whether SNP22 led to alteration of grain Cd accumulation in rice, we generated transgenic lines by overexpressing the *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> in a ZH11 background. Both the *OE-OsCd1*<sup>V449</sup> and *OE-OsCd1*<sup>D449</sup> transgenic lines accumulated more Cd compared to ZH11. In addition, the transgenic lines of *OE-OsCd1*<sup>V449</sup> showed a significantly less Cd accumulation ability in rice grain than *OE-OsCd1*<sup>D449</sup> (Fig. 4e, f). We then conducted the functional complementation assay by crossing both NILs with the *OsCd1*<sup>D449</sup> and *OsCd1*<sup>V449</sup> to the *CRISPR-osc1* knockout lines



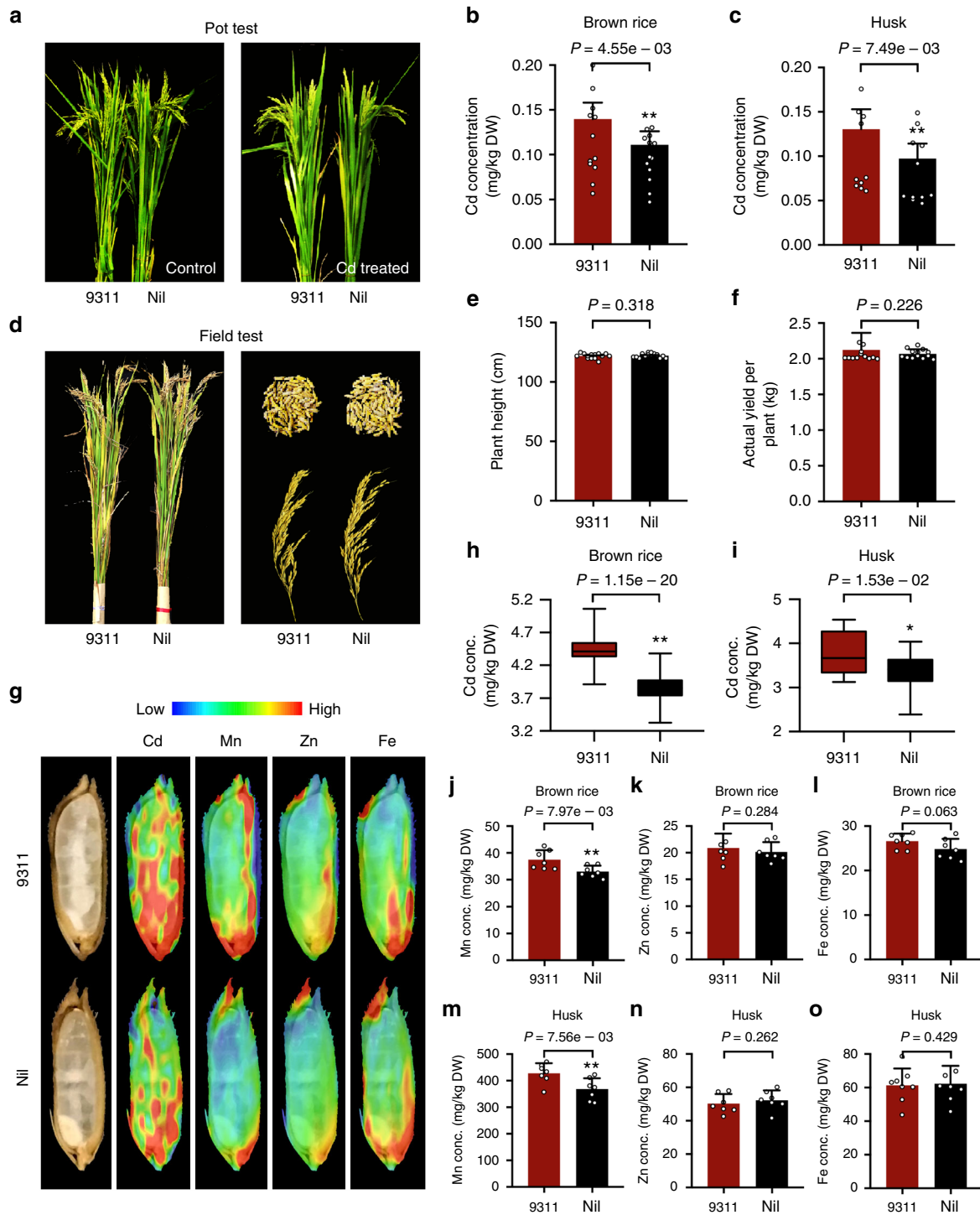
**Fig. 4** Functionally important amino acids in *OsCd1*. **a** Schematic diagram of gene structure and SNPs in *OsCd1*. **b** Predicted three-dimensional structural model of *OsCd1*. The structures of *OsCd1*<sup>D449</sup> (red) were viewed parallel (left one) and perpendicular (right one) to the membrane, respectively, and so for *OsCd1*<sup>V449</sup> (gray). Residue Asp449 in *OsCd1*<sup>D449</sup> (red) and Val449 in *OsCd1*<sup>V449</sup> (gray) were labeled at the cytoplasmic end respectively. ICH indicated intracellular helix. All structure figures were prepared with PyMol. **c** The grain Cd distribution of each genotype of *OsCd1* among rice natural variations group treated with 0.18 mg kg<sup>-1</sup> in pot-test; *n* denoted the number of genotypes belonging to each group. In each box plot, the bold horizontal line indicates the median, the edges of the box represent the first and third quartiles, and whiskers extend to span a 1.5 interquartile range from the edges. **d** Cd accumulation in W303 expressing *OsCd1*<sup>D449</sup> (red) and *OsCd1*<sup>V449</sup> (black) yeast treated with 20, 40, 60, 80 μM CdCl<sub>2</sub> for 24 h. Data from the transgenic lines were designed with three replications. **e** SR-μXRF images of Cd distribution of ZH11, OE-*OsCd1*<sup>V449</sup>, and OE-*OsCd1*<sup>D449</sup> lines in the rice grain. The emission intensity of each pixel was normalized using the beam intensity as reference. **f** Concentration of Cd in the brown rice of ZH11 (gray), OE-*OsCd1*<sup>V449</sup> (black) and OE-*OsCd1*<sup>D449</sup> (red) lines treated with 1.8 mg kg<sup>-1</sup> in the field. Error bars indicate standard deviation. Statistical comparison was performed by one-side *t*-test (\**P* < 0.05 and \*\**P* < 0.01). Data from the transgenic lines were designed with ten replications and data points for all biological replicates are shown. Source data of Fig. 4c, d and f are provided as a Source Data file

and treated the F1s with 1.8 mg kg<sup>-1</sup> Cd in the field. At harvest, we measured the Cd concentration in the grain and it was observed that a higher level in F1s with *OsCd1*<sup>D449</sup> than that of F1s with *OsCd1*<sup>V449</sup> (Supplementary Figure 12). Together with the results in yeast systems, we concluded that the SNP22, resulting in a change of encoded amino acid, altered the Cd transport ability of *OsCd1*.

***OsCd1*<sup>V449</sup> introgression reduced grain cadmium accumulation.** Considering that *OsCd1*<sup>V449</sup> exhibits lower Cd transport ability than that of *OsCd1*<sup>D449</sup>, we generated near-isogenic line (NIL) by the introgression *OsCd1*<sup>V449</sup> of Nipponbare (*japonica*) into the background of 9311 (*indica* with *OsCd1*<sup>D449</sup> genotype) for further research (Supplementary Figs. 13a and b). The effect of the allele makeup at *OsCd1* on grain Cd accumulation was firstly investigated in 0.18 mg kg<sup>-1</sup> treated pot-grown rice at greenhouse. As it is shown in Supplementary Fig. 13c and d, 9311

displayed higher Cd concentration in rice grain than the female parent Nipponbare lines. With the introgression of *OsCd1*<sup>V449</sup>, the Cd accumulation of NIL was significantly reduced both in brown rice and husk compared with the 9311 backgrounds with no apparent difference in growth (Fig. 5a–c).

To evaluate the effect of *OsCd1*<sup>V449</sup> allele in vitro, we selected the Cd-polluted fields at Changsha (Hunan Province, China, E112°, N28°)<sup>18</sup> with 1.8 mg kg<sup>-1</sup> Cd to perform the field trial. According to the newly published National Standard of the People's Republic of China: Soil environment quality Risk control standard of soil contamination of agricultural land (GB15618-2018), the risk intervention values for Cd is set to 1.5 mg kg<sup>-1</sup> (pH ≤ 5.5) or 2.0 mg kg<sup>-1</sup> (5.5 < pH ≤ 6.5)<sup>34</sup>. The experimental level of Cd used in this field is representative of a general situation of Cd-contaminated paddy fields in China. At harvest, there was no apparent difference in plant growth between the NIL and 9311. Moreover, the filled spikelet and actual yield per plot were also not affected by the introgression of *OsCd1*<sup>V449</sup> (Fig. 5d–f).



**Fig. 5** *OsCd1<sup>V449</sup>* introgression reduces grain Cd accumulation. **a** Growth of 9311 and NIL after treated with or without 0.18 mg kg<sup>-1</sup> Cd in pot-test. **b, c** Concentration of Cd in the brown rice and husk. Error bars indicate standard error of mean; statistical analysis was inspected using one-side paired *t*-test (\**P* < 0.05 and \*\**P* < 0.01). **d** Growth and grain morphologies of 9311 and NIL after treated with 1.8 mg kg<sup>-1</sup> Cd in the field. **e** Plant height of 9311 and NIL. **f** Actual yield per plot of 9311 and NIL. **g** SR-μXRF images of Cd, Mn, Zn, Fe distribution in the longitudinal sections of rice grain. The emission intensity of each pixel was normalized using the beam intensity as reference. **h-o** Concentrations of Cd, Mn, Zn, Fe in the brown rice (**h, j-l**) and husk (**i, m-o**). 9311 were shown in red and the NIL was shown in black. In each box plot, the bold horizontal line indicates the median, the edges of the box represent the first and third quartiles, and whiskers extend to span a 1.5 interquartile range from the edges. Error bars indicate standard deviation. Statistical analysis was inspected using one-side *t*-test (\**P* < 0.05 and \*\**P* < 0.01). All data were compared with 9311 and designed with at least five replications and data points for all biological replicates are shown. Source data of Fig. 5b, c, e, f, and h-o are provided as a Source Data file

Notably, the grain Cd concentration was detected to be reduced in NIL (Fig. 5h, i). According to the SR- $\mu$ XRF results, the Cd distribution in whole grain of NIL, especially the edible part endosperm, were significantly less than that in 9311 (Fig. 5g). For other micronutrients, except for a slightly decrease of Mn, no obvious change in the concentration of Zn and Fe was observed between 9311 and NIL (Fig. 5g, j–o). Additionally, when *OsCd1*<sup>V449</sup> was introduced into GUICHAO-2, another *indica* cultivar widely cultivated in south China, the grain Cd accumulation was also substantially decreased (Supplementary Fig. 14). These results demonstrated that the allelic effect of *OsCd1*<sup>V449</sup> contributed to the reduction of Cd in rice and, furthermore, may have a widely potential application value in rice genetic improvement.

### *OsCd1* diverged between *indica* and *japonica* subspecies.

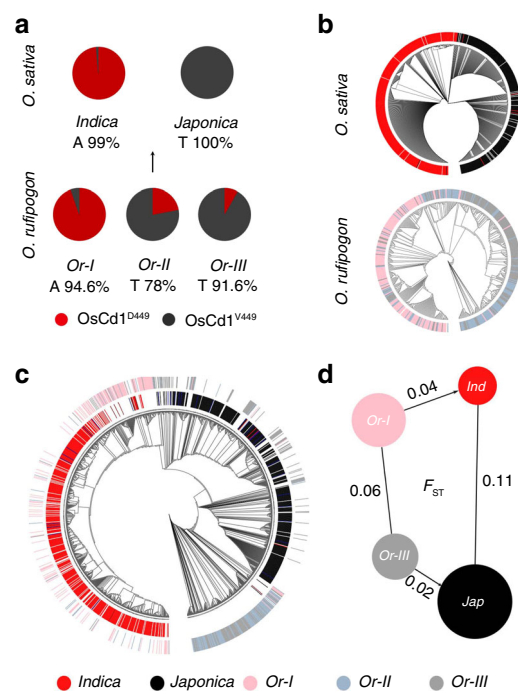
Interestingly, it was also observed that the *OsCd1* in most of the *indica* accessions was with *OsCd1*<sup>D449</sup> (A at SNP22 site) and that in most of the *japonica* accessions was with the *OsCd1*<sup>V449</sup> (T at SNP22 site) (Supplementary Table 2). In special, the grain Cd accumulation level was significantly lower in the *indica* cultivars with *OsCd1*<sup>V449</sup> than that of *OsCd1*<sup>D449</sup> (Supplementary Fig. 15), indicating that the Cd accumulation of *indica* cultivars with *OsCd1*<sup>V449</sup> also fitted the regular pattern for *japonica* rice. To explore the phylogenetic relationship of this allele, we then used whole-genome sequencing data for a large panel of accessions<sup>35</sup> containing 446 *O. rufipogon* accessions and 950 *O. sativa* varieties to determine the ancestral states of the SNP22 in *OsCd1*. Phylogenetic analysis using the 950 cultivar rice accessions showed that all the *japonica* accessions have nucleotide T (genotype *OsCd1*<sup>V449</sup>) at SNP22 while 99% of the *indica* accessions have nucleotide A (genotype *OsCd1*<sup>D449</sup>). Moreover, both the nucleotide A and T in SNP22 retained in its ancestor *O. rufipogon* (Fig. 6a, Supplementary Table 3).

Based on phylogenetic tree and PCA analysis, it was demonstrated that *OsCd1* diverged between *indica* and *japonica* subspecies in cultivated rice and between subgroup Or-I and Or-III in wild rice (Fig. 6b). Further analysis revealed that *OsCd1* in *indica* was descended from *O. rufipogon* Or-I, which mainly distributes in South and Southeast Asia, while that in *japonica* was from Or-III, which mainly distributes in China<sup>19</sup> (Fig. 6c and Supplementary Figure 16). The level of population differentiation ( $F_{st}$ ) was estimated to be 0.06 in *O. rufipogon* and increased to 0.11 in *O. sativa* during domestication (Fig. 6d). No significant Tajima's *D* values were observed for *japonica* (Tajima's *D*-1.16) and *indica* (Tajima's *D*-1.32) cultivars, indicating that the divergence of *OsCd1* between *indica* and *japonica* did not subject to selection and derived from the divergence of their progenitor gene pools hundreds of thousands of years ago.

### Discussion

Reducing rice grain Cd accumulation is an important task for decreasing risks to human health. Rice absorbed Cd from soil by roots and finally transferred then into grain via transport systems. Thus, identifying the genetic components underlying Cd transport is of great importance to reduce Cd accumulation in rice grain. More recently, GWAS has been successfully used for dissecting loci in several important agronomic traits<sup>28–32</sup> and was a promising strategy for finding genetic variants underlying the grain Cd accumulation in rice.

In this study, we analyzed the rice grain Cd content in 127 varieties and performed GWAS to dissect the valuable genetic components underlying rice grain Cd accumulation (Fig. 1a–d). As a result, QTL3, which showed the greatest contribution to the phenotypic variation, was selected for further research. Totally,



**Fig. 6** Genotypes and phylogenetic analysis of *OsCd1*. **a** The spectrum of allele frequencies at the causal polymorphisms of *OsCd1* in *O. rufipogon* (Or-I, Or-II and Or-III) and *O. sativa* (*japonica* and *indica*). **b** Phylogenetic tree of the full population calculated from the SNPs of *OsCd1* in *O. sativa* (*indica* and *japonica* subspecies, up) and *O. rufipogon* (Or-I, Or-II and Or-III, down). **c** Phylogenetic tree of the full population calculated from the SNPs of *OsCd1* in *O. rufipogon* (outer ring) and *O. sativa* (inner ring). **d** Illustration of genetic diversity and population differentiation in *O. rufipogon* and *O. sativa*. The  $F_{st}$  values between the groups were indicated. ind, *indica*; jap, *japonica*. Or-I, Or-II and Or-III are colored in pink, light blue and gray respectively; *indica* and *japonica* subspecies are in red and black, respectively

126 predicted genes with the annotation, such as transporter and transcription factor, were identified and may be the possible functional components in QTL3 (Supplementary Data 2). Using a composite bioinformatics tool-box method, *OsCd1*, which belongs to the major facilitator superfamily (MFS), was detected to be associated with rice grain Cd accumulation on the whole-genome scale and confirmed to mediate Cd influx into cell in yeast system (Fig. 1e, f).

MFS is the largest group of secondary active membrane transporters, and its members transport a diverse range of substrates<sup>36</sup>. Recently, MFS proteins were reported to be functioned as transporters of some metal elements such as *Arabidopsis* zinc (Zn) transporter ZIF1<sup>37</sup> and *Arabidopsis* cesium (Cs) and potassium (K) transporter ZIFL2<sup>38</sup>, whereas the role of MFS proteins in Cd transport was still not clear. Here, we characterized the MFS protein *OsCd1* as a possible Cd transporter gene in rice and the disruption of *OsCd1* resulted in a decrease of Cd accumulation in root, shoot, and grain (Fig. 2). Furthermore, sub-cellular and tissue localization showed that *OsCd1* was a plasma membrane protein and mainly expressed in the root exodermis, cortex, endodermis, and stele cells tissue (Fig. 3). We could then speculate that *OsCd1* may mediate the Cd uptake in rice root and, ultimately, contribute to grain Cd accumulation. Together with the results in transgenic yeast and *CRISPR-osc1* lines, we thus deduced that *OsCd1* may form a new-function clade in MFS involving in Cd transport. The exact role of *OsCd1* in Cd accumulation, however, needed to be further explored.



The identification of *OsCd1* enables us to reduce Cd accumulation in rice grain by transgenic techniques. However, the disruption of *OsCd1* is likely to result in a decrease of growth at harvest (Fig. 2e and Supplementary Fig. 5). As it was reported previously, *OsCd1* (*ASL*) shares 93.2% similarity to the gene *ASY*, which regulates early shoot development in rice. The absence of *ASY* showed various shoot abnormalities at the early vegetative stage. *OsCd1* (*ASL*) exists in the rice genome and is supposed to have redundant functions. For the reason that the *OsCd1* (*ASL*) expression level increased at later stages, it was speculated that *OsCd1* (*ASL*) may be involved in the shoot development of later stages<sup>39</sup>. In addition, it was also observed that *OsCd1* influences the Mn accumulation while the exact mechanism needs to be further investigated (Supplementary Fig. 6). Mn is a required trace mineral for all known living organisms and functions as cofactors for a large variety of enzymes with many functions<sup>40,41</sup> and the disruption of *OsCd1* was likely to influence the rice growth at harvest.

Considering the fact that the disruption of *OsCd1* has a negative impact on plant growth and reduces yield at harvest, it is thus not an optimal strategy to establish low-Cd rice by simply knocking out *OsCd1*, while selection of superior alleles of *OsCd1* may be a more effective strategy for genetic improvement of rice. We then conducted research to excavate natural variation allele in *OsCd1*. Based on the sequences, we identified a non-synonymous mutation (SNP22), locating at the ICH1 of C-terminal (Fig. 4a, b), could cause a charged amino acid D449 substituting to an uncharged one V449. Typically, MFS protein possesses 12 transmembrane  $\alpha$ -helices with both their N- and C-terminal located in the cytoplasm<sup>42–46</sup>. During the transport reaction cycle, some amino acids on N- and C-terminal form hydrogen bonds with the substrate and translocate the substrate via the “rocker-switch” mechanism<sup>42,47,48</sup>. N- and C-terminal, especially the charged or polar residues in the ICH domain, are likely to have a critical role to close the transporter on the intracellular side during the alternating access cycle<sup>49</sup>. When the V449 was substituted to D449 in SNP22, the polarity change caused a relatively stronger Cd transport ability in *OsCd1*<sup>D449</sup> compared with *OsCd1*<sup>V449</sup> (Fig. 4d). The yeast assay, together with the complementary experiments with over-expression promoter indicated that the Cd transport ability varied in two alleles (Fig. 4e, f). Considering native promoter, which is a better strategy for functional complementation, was not used, we thus crossed the alternative alleles to the *CRISPR-oscd1* lines and the Cd transport ability variation was further confirmed (Supplementary Fig. 12). The difference in Cd transport ability, finally, endowed the *OsCd1*<sup>D449</sup> rice cultivars with a higher grain Cd accumulation ability than the cultivars with *OsCd1*<sup>V449</sup> (Fig. 4c).

In addition, no tissue expression level and subcellular location differences were observed between the *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> (Supplementary Figs. 8 and 9). We thus regarded SNP22 in *OsCd1* as a critical allele contributing to the divergence in rice grain Cd accumulation and tried to apply the functional allele *OsCd1*<sup>V449</sup> for marker-assisted breeding.

Two rice cultivars, 9311 and GUICHAO-2, which showed a relatively higher grain Cd content compared with Nipponbare and Guihuahuang, were selected as the background of NIL. When the *OsCd1*<sup>V449</sup> allele was introduced into them, lower grain Cd accumulation NILs were obtained (Fig. 5b, c, h, i and Supplementary Fig. 14). Meanwhile, it is noteworthy that, different from *CRISPR-oscd1* lines, which resulted in a complete loss of gene function, the NIL with lower grain Cd accumulation has no apparent difference in plant growth and spikelet fertility compared with 9311, though the Mn accumulation in NIL was slightly decreased (Fig. 5d–f and Supplementary Fig. 14).

This suggests *OsCd1*<sup>V449</sup> a great potential application in low-Cd rice cultivation.

Phylogenetic and population genetic analyses demonstrated the natural variation in *OsCd1* diverged between *indica* and *japonica* subspecies: the *japonica* accessions with genotype *OsCd1*<sup>V449</sup> displayed a relatively lower Cd accumulation in rice grain than the *indica* accessions with genotype *OsCd1*<sup>D449</sup> (Fig. 6). This was different with *OsNramp5*, a major Cd transporter in rice. In *OsNramp5*, the frequency of the two missense variation SNPs in *OsNramp5* showed almost no difference in *indica* and *japonica*, indicating it contributed little to the Cd divergence between *indica* and *japonica* rice (Supplementary Fig. 17). In Asia, the cultivated rice is mainly classified into *indica* and *japonica* subspecies. *Indica* cultivars, which are the major food and widely grown in the South and Southeast Asia, generally exhibit higher grain Cd accumulation than *japonica* cultivars<sup>50,51</sup> (Supplementary Fig. 1c). Since Cd is one of the most serious soil heavy metal pollutant in these regions<sup>52,53</sup>, the ingestion of *indica* rice grain brings an unneglectable risk of Cd to people living there. In the previous study, three *osnramp5-1* rice mutants were developed by carbon ion-beam irradiation strategy for marker-assisted selection of low-Cd cultivars<sup>11</sup>. However, the parent of these three rice mutants is Koshihikari, which is the main cultivar in Japan but not in China and other Southeast Asian countries. *OsCd1*<sup>V449</sup> exists in most of the *japonica* cultivars which makes it a wider parent selection range in low-Cd cultivar breeding. By integrating the low-Cd allele *OsCd1*<sup>V449</sup> into *indica* cultivars, the grain Cd accumulation had been successfully reduced on the Cd-contaminated paddy fields in Hunan province, where is considered as a typical Cd-contaminated area in China and has suffered from Cd contamination for many years. This finding, along with the function characterizing of natural variation of *OsCd1*, emphasizes the potential application value of *OsCd1*<sup>V449</sup> in low-Cd rice cultivation, especially in *indica* cultivars.

## Methods

**GWAS analysis and gene annotation.** A diverse worldwide collection of 127 *O. sativa* accessions including both landraces and elite varieties was obtained from a core/mini-core collection of *Oryza sativa* L. for GWAS analysis<sup>33</sup>. Each accession was germinated and planted in cultivation pot treated with 1.8 mg kg<sup>-1</sup> Cd (Beijing, China; 39.9°N, 116.3°E). The grain was collected after harvest to test the Cd concentration of each genotype. In total, we used 3,291,150 SNPs with a minor allele frequency of >0.05 to carry out GWAS. Association analysis using a mixed model was performed with the GAPIT software package. The top three principal components were used as fixed effects, and the matrix of genetic similarity based on simple SNP matching coefficients was used to model the variance-covariance matrix of the random effect. The analyses were performed in GAPIT and the parameters for each trait were optimized automatically.

Permutation tests were used to help define the genome-wide significant *P*-value threshold<sup>54</sup>. For all the examined traits, we reshuffled the original phenotype data, and then performed association analysis using GAPIT with the same parameters. There ought to be no real associations between the SNPs and the simulated phenotypes, so all the SNPs passing the permutation threshold should be positives. After conducting 1000 permutation analyses, we clearly found that the permutation threshold varied across genomic region and some of the association signals even passing the significant cutoff 10<sup>-5</sup>. Thus, we adopted a threshold *P* = 10<sup>-5</sup> at genome-wide level.

For QTL and gene annotation, we firstly selected the SNP signal that passing the threshold at *P* = 10<sup>-5</sup>, to check if the distance to the adjacent SNP was less than 170 kb<sup>32</sup>. We then merged it into the same QTL, all the genes locating in the QTL region were predicted by the Rice Genome Annotation Project (MSU-RGAP, Nipponbare version 6.1, <http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/>) and annotated by gene ontology (GO) Slim annotation with the keyword membrane and transport.

**Primers.** The sequences for all primers used in this study are listed in Supplementary Table 4.

**Cadmium and manganese transport activities assay in yeast.** RNA was extracted from *CRISPR-oscd1*, 9311 and Nipponbare accessions. For candidate gene selection assay, the coding sequence of *C1*, *C2*, *C3*, and *C4* was amplified from 9311

cDNA. The haplotypes *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> were amplified from 9311 and Nipponbare cDNA, respectively. The *oscd1-CR1*, *oscd1-CR2* and *oscd1-CR3* were amplified from *CRISPR-oscd1*, respectively. The sequences were then cloned into *pAG413GAL* vector to construct yeast expression vector. The constructed vectors along with the empty vector *pAG413GAL* were transformed into the wild-type *Saccharomyces cerevisiae* W303 (*MATa his3-1 met15-0 trp1-1 ura3*, Laboratory preservation), *Δycf1* (*MATa ura3 leu2 his3-1 met15-0 YDR135c::kanMX4*, Laboratory preservation) and *Δsmf1* (*MATa his3-1 leu2 met15-0 ura3 YOL122c::kanMX4*, Chen Caiyan's Laboratory preservation).

Polyethylene glycol (PEG)/LiAc-based method was used for preparing and transforming competent yeast cells<sup>55</sup>. Yeast cells were grown at 30 °C in YPAD medium and then treated with 100 mM LiAc to make yeast competent cells. The mixture of ~200 ng μl<sup>-1</sup> plasmid, 240 μl PEG solutions, 26 μl 1 M LiAc and 10 μl Carrier DNA were added into the yeast competent cells and incubated at 30 °C for 30 min. Then the mixture was heat shocked at 42 °C for 20 min. Yeast transformants were cultured in synthetic defined (SD) medium without histidine at 30 °C for 48 h.

For plate growth tests, yeast transformants were diluted to an OD<sub>600</sub> of 1.0, 0.1, 0.01, and 0.001 step by step with sterile water. Then the yeast transformants were spotted on SD-His plates (with 2% (w v<sup>-1</sup>) galactose/glucose) with or without 30 μM CdCl<sub>2</sub> or 5 mM EGTA, respectively. The yeast carrying empty vector was used as control. Plates were incubated for 2–3 d at 30 °C for phenotype observation. All the assays were performed at least three times.

For the Cd and Mn concentration determination, the yeast transformants were cultured overnight with liquid SD medium (with 2% galactose) to OD<sub>600</sub> = 1 with 200 rpm at 30 °C. Then the yeast transformants were treated with 20, 30, 40, 60, 80 μM CdCl<sub>2</sub> for 24 h or 5 mM EGTA for 48 h, respectively. The yeast strain carrying empty vector was used as control. All the yeast samples were collected and washed with sterile water for three times. Then all the samples were dried at 80 °C in an oven for further Cd and Mn concentration determination. All the assays were performed at least three times.

**3D structure modeling and protein structure analysis.** The *OsCd1* sequence was analyzed by Phyre2<sup>56</sup> and secondary structure was displayed using TMRPres2D<sup>57</sup> to determine the position of the transmembrane helices. Three-dimensional model of *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> were constructed using the Phyre2 and MODELLER software<sup>58</sup>. Three high-resolution protein structures GlpT (PDB entry 1pw4), YajR (PDB entry 3wdo), and POT (PDB entry 4iky) were used as templates simultaneously in the comparative modeling procedure. Three-dimensional models were visualized within PyMOL<sup>59</sup>.

**Phylogenetic reconstruction of *OsCd1*.** In total 149 protein sequence of MFS members from rice were obtained from Transport Database (<http://www.membranetransport.org/transportDB2/index.html>). Multiple-sequence alignment was optimized by MUSCLE. The phylogeny of *OsCd1* in the MFS was constructed by MEGA7<sup>60</sup> using the neighbor-joining method with No. of differences model, pairwise deletion for missing data and 1000 bootstrap pseudoreplicates. The software EvolView<sup>61</sup> was used for visualizing the phylogenetic trees.

**CRISPR-*oscd1* mutant lines and function complementary assay.** The sgRNA-Cas9 and *pUN1301-OsCd1* plant expression vectors were constructed<sup>62</sup>. The oligos used in constructing the sgRNA vectors for *OsCd1* are listed in Supplementary Table 5. The rice variety Nipponbare (*Oryza sativa* L. ssp. *japonica*) was used as the hosts in agrobacterium-mediated transformation for *CRISPR-oscd1* mutant and ZH11 (*Oryza sativa* L. ssp. *japonica*) was used as the hosts in agrobacterium-mediated transformation for *OE-OsCd1* lines (Hangzhou Bioglo Co., LTD). The transgenic rice lines were grown in the paddy field in Hainan (18.48° N, 110.02° E, China), during normal rice-growing seasons. For *CRISPR-oscd1* mutant lines, mature seeds were collected from T<sub>0</sub> plants and germinated for genotyping. Total DNA was isolated from transgenic plants and PCR was performed to amplify the genomic region surrounding the CRISPR target sites using the specific primers. The PCR fragments were directly sequenced by Sanger method to identify mutations. The *CRISPR-oscd1* lines were then used for crossing with NIL and 9311 respectively and the T<sub>1</sub> plants was the functional complementary experiments. For the *OE-OsCd1* lines, the GUS marker was used for the selection of pure rice lines in T<sub>2</sub> plants and then the pure rice lines were used for further experiments.

**Hydroponic experiments.** To conduct the hydroponic experiments in vegetative stage, rice seeds of wild-type rice and *CRISPR-oscd1* lines were sown in a germination box with water in the dark. Then the 3-days rice were transferred to Kimura B solution containing (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, MgSO<sub>4</sub>·7H<sub>2</sub>O, KNO<sub>3</sub>, Ca(NO<sub>3</sub>)<sub>2</sub>·4H<sub>2</sub>O, and KH<sub>2</sub>PO<sub>4</sub> and the micronutrients MnCl<sub>2</sub>·4H<sub>2</sub>O, H<sub>3</sub>BO<sub>3</sub>, (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O, ZnSO<sub>4</sub>·7H<sub>2</sub>O, CuSO<sub>4</sub>·5H<sub>2</sub>O, and Fe-EDTA, pH 5.6. After growing in Kimura B solution 5 days, 1 μM CdCl<sub>2</sub> were added and the seedlings were treated with CdCl<sub>2</sub> for 20 days. The plants were cultivated in the growth chamber at a temperature of 30 °C and the nutrient solution was replaced every 2 d. Then the shoot and root were collected separately with five replications to determine the Cd concentration. All the assays were performed at least three times.

**Pot and field experiments.** Pot experiments were carried out at the greenhouse of Institute of Botany, Beijing. Three-week seedlings of Nipponbare, *CRISPR-oscd1* lines, 9311 and NILs pre-cultured hydroponically were transplanted to the pot with 0.18 mg kg<sup>-1</sup> Cd in May and harvested at September. All experiments were repeated with 30 biological replicates.

Field tests of 9311 and the NIL were performed during May to September in 2017 and 2018. All the cultivars were planted in the Cd-polluted paddy fields at Institute of Subtropical Agriculture, CAS. The soil Cd concentration was 1.80 mg kg<sup>-1</sup>, pH 5.4. Randomized complete block design was applied in the field experiments. Each sample was assigned with three replicates. All the seedlings were transplanted at a spacing of 17 cm within rows and the distance between two rows was 20 cm. Following method was used to measure the agronomic traits including plant height, filled spikelet rate and grain yield per plant after harvest. Plant height was the length of main tiller; Filled spikelet rate was calculated as follows: filled grains/(unfilled grains + filled grains) × 100%; The grain yield was determined by weighing all the grains of one plant after drying at 50 °C in a lab oven for 8 h. The grain metal concentration was determined as described follows.

**Determination of metal accumulation using ICP-MS.** The content of total Cd, Zn, Mn, and Fe in the yeast and plant samples were determined by Inductive Coupled Plasma Emission Spectrometer (ICP-OES) (iCAP6300, Thermo Electron Corp., MA, USA). All samples were dried at 80 °C for 6 h. Dried samples were digested using 2 ml of concentrated HNO<sub>3</sub> overnight. Sample digestion was carried out by heating block at 200 °C for 8 h. After cooling, the digested solution was diluted to 15 ml with deionized water and filtered through 0.22 μm cellulose acetate membranes filters. All digestions were performed in triplicate. The blank HNO<sub>3</sub> was used as the negative control and the certified standard material sample (CRM rice; GBW100348) was used as the positive control. All the assays were performed at least three times.

**In situ analyses using synchrotron X-ray fluorescence.** Metal distribution in brown rice was evaluated using micro-X-ray fluorescence (SR-μXRF) method, which performed at 4W1B beamline, Beijing synchrotron Radiation Facility. The longitudinal and latitudinal sections of rice samples were prepared for SR-μXRF analysis. The incident X-ray energy is 15 keV and the spot size diameter is 50 μm. Si (Li) solid state detector was used to detect X-ray fluorescence emission lines. Step-mode with 100×100 μm was used to acquire the two-dimensional mapping of rice grain samples and each step's live time is 30 s. After sample mapping, the data processing was performed using PyMca package<sup>63</sup> and Origin 8.

**Quantitative real-time PCR.** Nipponbare varieties, 20 *japonica* and 20 *indica* cultivars were treated with and without 1 μM CdCl<sub>2</sub> for 24 h, respectively. Total RNA of the root, shoot and the whole plant was extracted using TRIzol reagent (Invitrogen) and then treated with DNase I. After DNaseI digest, ~2 μg RNA was converted to cDNA and the product was used as templates for the qPCR. Water and the cDNA without DNaseI digestion were used as negative control. qPCR was performed on an ABI 7500 instrument using SYBR Green PCR Master mix following the manufacturer's instructions (Takara, RR420). *Oshiston* gene were used as internal standards, and normalized relative expression was calculated by the ΔΔCt (cycle threshold) method. The primers used in qPCR are listed in Supplementary Table 4.

**Tissue expression assay.** For histochemical analysis of GUS activity, the upstreaming 2.5 kb genomic fragment of *OsCd1* was regarded as the *OsCd1* promoter and cloned into pCambia1391 to generate *OsCd1*<sub>promoter::GUS</sub> vector. The vectors were then transformed into Nipponbare and the transgenic rice were sampled to detect GUS expression. To determine the activity of GUS, the samples were incubated into the solution containing 1 mM X-gluc, 10% methanol, 0.5% Triton X-100 and 50 mM Na<sub>3</sub>PO<sub>4</sub> at 37 °C for 24 h. After being stained, 75% ethanol was applied to remove chlorophyll. Resin semi-thin sections were prepared for image collection to determine GUS activity.

For the GFP observation, the *OsCd1*'s promoter was cloned into pMDC45 and then the *OsCd1*<sub>promoter::GFP</sub> vector was then transformed into Nipponbare. The transgenic rice root tissues were prepared into the frozen sections and then the confocal laser scanning microscope (FluoView FV1000, Olympus) was used to observe the location of *OsCd1*. Nipponbare was used as the negative control and Ubi::GFP transgenic rice was used as the positive control.

**Subcellular-localization assay.** The CDs of *OsCd1*<sup>V449</sup> and *OsCd1*<sup>D449</sup> were fused in frame with GFP via cloning into the binary vector pMDC45. The resulting vectors pMDC45-*OsCd1*<sup>V449</sup> were transformed into Nipponbare (*Oryza sativa* L. ssp. *japonica*) and used as hosts in agrobacterium-mediated transformation method. For FM4-64 staining, rice root was transferred to 2.5 μM FM4-64 diluted in 1/2 MS medium for 3 min. Besides, the protoplasts were isolated from root and the pMDC45-*OsCd1*<sup>V449</sup> and pMDC45-*OsCd1*<sup>D449</sup> were transiently expressed in the root protoplasts for GFP observation. The subcellular localization and co-localization were evaluated using a confocal laser scanning microscope (FluoView FV1000, Olympus).

**Population genetics analysis.** 12 kb of the sequence centered on *OsCd1* in 446 *O. rufipogon* accessions and 950 *O. sativa* varieties were obtained from rice HapMap3 data set (<http://server.ncgr.ac.cn/RiceHap3/>). SNPs in the 12-kb region were used for the *OsCd1* variety phylogenetic reconstruction analysis by MEGA7 using the neighbor-joining method with No. of differences model, pairwise deletion for missing data and 1000 bootstrap pseudoreplicates. Principal component analysis of the SNPs was performed using the software TASSEL. The level of population differentiation ( $F_{st}$ )<sup>64</sup> and Tajima's  $D$ <sup>65</sup>. Statistic was calculated by a custom PERL script using a 1000 bp window. The  $F_{st}$  between Or-I and Or-III in *O. rufipogon* and between *japonica* and *indica* of *O. sativa* was calculated, respectively.

**NIL construction.** *Oryza sativa* near-isogenic line (NIL) were selected from a BC<sub>5</sub>F<sub>2</sub>, which was generated by crossing between *japonica* variety Nipponbare × *indica* variety 9311 then backcrossing to 9311, respectively. 120 SSR markers distributed evenly throughout 12 chromosomes were used for identification and selection of the candidate lines containing the target donor segment. The size of the introgression fragment in the NIL was about 1 Mb between RM3413 and RM3372.

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

### Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The datasets generated and analyzed during the current study are available from the corresponding author upon request. The source data underlying Figs. 1f, 2c, d, f, g, i, j, 3a, d, 4c, d, f, 5b, c, e, f, and h–o, and Supplementary Figs. 1c, 4b, 5b, 6a–c, 9, 10b, c, 11, 12b, 13c, d, 14b, and 15 are provided as a Source Data file.

Received: 20 December 2017 Accepted: 14 May 2019

Published online: 12 June 2019

### References

- Tsukahara, T. et al. Rice as the most influential source of cadmium intake among general Japanese population. *Sci. Total Environ.* **305**, 41–51 (2003).
- Phillipp, R. Cadmium in the environment and cancer registration. *J. Epidemiol. Community Health* **34**, 151–151 (1980).
- Uraguchi, S. & Fujiwara, T. Rice free ground for cadmium-free cereals. *Curr. Opin. Plant Biol.* **16**, 328–34 (2013).
- Takahashi, R. et al. The OsNRAMP1 iron transporter is involved in Cd accumulation in rice. *J. Exp. Bot.* **62**, 4843–50 (2011).
- Takahashi, R., Ishimaru, Y., Nakanishi, H. & Nishizawa, N. K. Role of the iron transporter OsNRAMP1 in cadmium uptake and accumulation in rice. *Plant Signal. Behav.* **6**, 1813–1816 (2011).
- Yang, M. et al. OsNRAMP5 contributes to manganese translocation and distribution in rice shoots. *J. Exp. Bot.* **65**, 4849–4861 (2014).
- Tang, L. et al. Knockout of OsNramp5 using the CRISPR/Cas9 system produces low Cd-accumulating *indica* rice without compromising yield. *Sci. Rep.* **7**, 14438 (2017).
- Takahashi, R. et al. From laboratory to field: OsNRAMP5-knockdown rice is a promising candidate for Cd phytoremediation in paddy fields. *Plos ONE* **9**, e98816 (2014).
- Ishimaru, Y. et al. Characterizing the role of rice NRAMP5 in manganese, iron and cadmium transport. *Sci. Rep.* **2**, 286 (2012).
- Ishimaru, Y., Bashir, K., Nakanishi, H. & Nishizawa, N. K. OsNRAMP5, a major player for constitutive iron and manganese uptake in rice. *Plant Signal. Behav.* **7**, 763–766 (2012).
- Ishikawa, S. et al. Ion-beam irradiation, gene identification, and marker-assisted breeding in the development of low-cadmium rice. *Proc. Natl Acad. Sci. USA* **109**, 19166–19171 (2012).
- Yan, J. et al. A loss-of-function allele of OsHMA3 associated with high cadmium accumulation in shoots and grain of *Japonica* rice cultivars. *Plant Cell Environ.* **39**, 1941–1954 (2016).
- Miyadate, H. et al. OsHMA3, a P1B-type of ATPase affects root-to-shoot cadmium translocation in rice by mediating efflux into vacuoles. *New Phytol.* **189**, 190–199 (2011).
- Kumagai, S. et al. Functional analysis of the C-terminal region of the vacuolar cadmium-transporting rice OsHMA3. *Febs Lett.* **588**, 789–794 (2014).
- Sasaki, A., Yamaji, N. & Feng, M. J. Overexpression of OsHMA3 enhances Cd tolerance and expression of Zn transporter genes in rice. *J. Exp. Bot.* **65**, 6013–6021 (2014).
- Yamaji, N., Xia, J., Mitaniueno, N., Yokosho, K. & Feng, M. J. Preferential delivery of zinc to developing tissues in rice is mediated by P-type heavy metal ATPase OsHMA2. *Plant Physiol.* **162**, 927–939 (2013).
- Takahashi, R. et al. The OsHMA2 transporter is involved in root-to-shoot translocation of Zn and Cd in rice. *Plant Cell Environ.* **35**, 1948–1957 (2012).
- Satohngasawa, N. et al. Mutations in rice (*Oryza Sativa*) heavy metal ATPase 2 (OsHMA2) restrict the translocation of zinc and cadmium. *Plant Cell Physiol.* **53**, 213 (2012).
- Luo, J. S. et al. A defensin-like protein drives cadmium efflux and allocation in rice. *Nat. Commun.* **9**, 645 (2018).
- Uraguchi, S. et al. Low-affinity cation transporter (OsLCT1) regulates cadmium transport into rice grains. *Proc. Natl Acad. Sci. USA* **108**, 20959–20964 (2011).
- Uraguchi, S., Kamiya, T., Clemens, S. & Fujiwara, T. Characterization of OsLCT1, a cadmium transporter from *indica* rice (*Oryza sativa*). *Physiol. Plant.* **151**, 339–347 (2014).
- Morishita, T., Fumoto, N., Yoshizawa, T. & Kagawa, K. Varietal differences in cadmium levels of rice grains of *japonica*, *indica*, *javanica*, and hybrid varieties produced in the same plot of a field. *Soil Sci. Plant Nutr.* **33**, 629–637 (1987).
- Liu, J. et al. Interaction of Cd and five mineral nutrients for uptake and accumulation in different rice cultivars and genotypes. *Field Crop. Res.* **83**, 271–281 (2003).
- He, J., Zhu, C., Ren, Y., Yan, Y. & Jiang, D. Genotypic variation in grain cadmium concentration of lowland rice. *J. Plant Nutr. Soil Sci. -Z.* **169**, 711–716 (2006).
- Ishikawa, S., Ae, N. & Yano, M. Chromosomal regions with quantitative trait loci controlling cadmium concentration in brown rice (*Oryza sativa*). *New Phytol.* **168**, 345–350 (2005).
- Kato, H. et al. Structural diversity and evolution of the Rf-1 locus in the genus *Oryza*. *Heredity* **99**, 516–524 (2007).
- Ueno, D. et al. A major quantitative trait locus controlling cadmium translocation in rice (*Oryza sativa*). *New Phytol.* **182**, 644–653 (2009).
- Si, L. et al. OsSPL13 controls grain size in cultivated rice. *Nat. Genet.* **48**, 447–456 (2016).
- Ellinor, P. T. et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat. Genet.* **44**, 670–675 (2012).
- Zuo, W. et al. A maize wall-associated kinase confers quantitative resistance to head smut. *Nat. Genet.* **47**, 151–157 (2015).
- Huang, X. et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* **44**, 32–U53 (2012).
- Huang, X. et al. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* **42**, 961–U76 (2010).
- Zhang, H. et al. A core collection and mini core collection of *Oryza sativa* L. in China. *Theor. Appl. Genet.* **122**, 49–61 (2011).
- Wang, L. et al. A review of soil cadmium contamination in China including a health risk assessment. *Environ. Sci. Pollut. Res.* **22**, 16441–16452 (2015).
- Huang, X. et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497–501 (2012).
- Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1–34 (1998).
- Haydon, M. J. & Cobbett, C. S. A novel major facilitator superfamily protein at the tonoplast influences zinc tolerance and accumulation in *Arabidopsis*. *Plant Physiol.* **143**, 1705–1719 (2007).
- Remy, E. et al. The Major Facilitator Superfamily Transporter ZIFL2 Modulates Cesium and Potassium Homeostasis in *Arabidopsis*. *Plant Cell Physiol.* **56**, 148–162 (2015).
- Somers, I. I. & Shive, J. W. The iron-manganese relation in plant metabolism. *Plant Physiol.* **17**, 582–602 (1942).
- Hibara, K. I. et al. Abnormal shoot in youth, a homolog of molybdate transporter gene, regulates early shoot development in rice. *Am. J. Plant Sci.* **30**, 1–9 (2013).
- Homann, P. H. Studies on the manganese of the chloroplast. *Plant Physiol.* **42**, 997–1007 (1967).
- Dang, S. et al. Structure of a fucose transporter in an outward-open conformation. *Nature* **467**, 734–U130 (2010).
- Pedersen, B. P. et al. Crystal structure of a eukaryotic phosphate transporter. *Nature* **496**, 533–536 (2013).
- Tao, Y. et al. Structure of a eukaryotic SWEET transporter in a homotrimeric complex. *Nature* **527**, 259–263 (2015).
- Kumar, H. et al. Structure of sugar-bound LacY. *Proc. Natl Acad. Sci. USA* **111**, 1784–1788 (2014).
- Madej, M. G., Sun, L., Yan, N. & Kaback, H. R. Functional architecture of MFS D-glucose transporters. *Proc. Natl Acad. Sci. USA* **111**, E719–E727 (2014).

47. Abramson, J. et al. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610–615 (2003).
48. Jiang, D. et al. Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proc. Natl Acad. Sci. USA* **110**, 14664–14669 (2013).
49. Lu, J. M. Y. & Bush, D. R. His-65 in the proton-sucrose symporter is an essential amino acid whose modification with site-directed mutagenesis increases transport activity. *Proc. Natl Acad. Sci. USA* **95**, 9025–9030 (1998).
50. Sun, L. et al. Genetic diversity, rather than cultivar type, determines relative grain Cd accumulation in hybrid rice. *Front. Plant Sci.* **7**, 1407 (2016).
51. Arao, T. & Ae, N. Genotypic variations in cadmium levels of rice grain. *Soil Sci. Plant Nutr.* **49**, 473–479 (2003).
52. Duan, Q., Lee, J., Liu, Y., Chen, H. & Hu, H. Distribution of heavy metal pollution in surface soil samples in China: A graphical review. *Bull. Environ. Contam. Toxicol.* **97**, 303–309 (2016).
53. Liu, X., Tian, G., Jiang, D., Zhang, C. & Kong, L. Cadmium (Cd) distribution and contamination in Chinese paddy soils on national scale. *Environ. Sci. Pollut. Res.* **23**, 17941–17952 (2016).
54. Churchill, G. A. & Doerge, R. W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971 (1994).
55. Roberts, C. Methods in yeast genetics. *Methods Med. Res.* **3**, 37–50 (1950).
56. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
57. Spyropoulos, I. C., Liakopoulos, T. D., Bagos, P. G. & Hamodrakas, S. J. TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics* **20**, 3258–3260 (2004).
58. Fiser, A. & Sali, A. MODELLER: Generation and refinement of homology-based protein structure models. *Macromol. Crystallogr.* **374**, 461–491 (2003).
59. DeLano, W. L. & Lam, J. W. PyMOL: A communications tool for computational models. *Abstr. Pap. Am. Chem. Soc.* **230**, U1371–U1372 (2005).
60. Tamura, K. et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
61. Zhang, H., Gao, S., Lercher, M. J., Hu, S. & Chen, W.-H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res.* **40**, W569–W572 (2012).
62. Shan, Q., Wang, Y., Li, J. & Gao, C. Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc.* **9**, 2395–2410 (2014).
63. Sole, V. A., Papillon, E., Cotte, M., Walter, P. & Susini, J. A multiplatform code for the analysis of energy-dispersive X-ray fluorescence spectra. *Spectrosc. Acta Pt. B-At. Spectr.* **62**, 63–68 (2007).
64. Wright, S. The interpretation of population-structure by f-statistics with special regard to systems of mating. *Evolution* **19**, 395–420 (1965).
65. Tajima, F. Statistical-method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics* **123**, 585–595 (1989).

## Acknowledgements

This work was supported by the grants from Chinese Academy of Sciences (XDA08010406), the Ministry of Science and Technology of China (2015FY111300, 2017YFD0800901), China Postdoctoral Science Foundation (2018M64152) and Innovation Training Programs for Undergraduates, CAS (201717001751). We thank Fengqin Dong for help in the resin semi-thin section, Shumin Zhang and Wei Liu for metal concentration determination and Jingquan Li for microscopy experiments from Plant Science Facility of the Institute of Botany, Chinese Academy of Sciences for their excellent technical assistance. The  $\mu$ -XRF beam time was granted by 4W1B beamline of Beijing Synchrotron Radiation Facility, Institute of High Energy Physics, Chinese Academy of Sciences. The staff members of 4W1B are acknowledged for their technical support.

## Author contributions

Z.H. and H.Z. designed the research and M.M. and C.C. contributed to the experiments. H.Y., W.X., Y.G., L.W., X.C., C.D., T.Z., X.L., T.L. and L.F. conducted the research. J.X., Z.Z., W.X., Q.W., M.L., Y.Z., J.L. and C.L. performed the GWAS analysis. H.Y., W.X., and L.S. contributed to the field management. H.Y. and Y.G. analyzed data. Z.H. and H.Y. wrote the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-10544-y>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information:** *Nature Communications* thanks David Salt and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

**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) 2019