



HAL
open science

Genetic loci regulating cadmium content in rice grains

Gareth Norton, Anthony Travis, Panthita Ruang-Areerate, Graeme W. Nicol, Ayotunde Adeosun, Mahmud Hossain, M. Rafiq Islam, Alex Douglas, Adam Price

► **To cite this version:**

Gareth Norton, Anthony Travis, Panthita Ruang-Areerate, Graeme W. Nicol, Ayotunde Adeosun, et al.. Genetic loci regulating cadmium content in rice grains. *Euphytica*, 2021, 217 (3), 10.1007/s10681-020-02752-1 . hal-03411978

HAL Id: hal-03411978

<https://hal.science/hal-03411978>

Submitted on 3 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Genetic loci regulating cadmium content in rice grains

Gareth J. Norton · Anthony Travis · Panthita Ruang-areerate ·
Graeme W. Nicol · Ayotunde A. Adeosun · Mahmud Hossain · M. Rafiq Islam ·
Alex Douglas · Adam H. Price

Received: 27 January 2020 / Accepted: 15 December 2020 / Published online: 10 February 2021
© The Author(s) 2021

Abstract It has been estimated that up to 90% of human exposure to cadmium is through food, and that cadmium within rice grains can be a major contributor to that dietary source. In this study genome wide association mapping was conducted on the Bengal and Assam Aus Panel (BAAP) of rice to identify quantitative trait loci and candidate genes for lowering grain cadmium. Field experiments were conducted over two years under two different irrigation systems: continually flooded and alternate wetting and drying (AWD). There was significant effects of water treatment,

genotype, and genotype by water treatment interaction. Importantly, AWD increased grain cadmium, on average, by 49.6% and 108.8% in year 1 and 2 respectively. There was between 4.6 and 28 fold variation in cadmium concentration. A total of 58 QTLs were detected but no loci are clearly specific to one water regime despite approximately 20% of variation attributable to genotype by water regime interaction. A number of QTLs were consistent across most water treatments and years. These included QTLs on chromosome 7 (7.23–7.61, 8.93–9.04, and 29.12–29.14 Mbp), chromosome 5 (8.66–8.72 Mbp), and chromosome 9 (11.46–11.64 Mbp). Further analysis of the loci on chromosome 7 (8.93–9.04 Mbp), identified the candidate gene *OsNRAMP1*, where cultivars with a deletion upstream of the gene had higher concentrations of cadmium compared to the cultivars that did not have the deletion. The distribution of alleles within the BAAP suggest this QTL is easily detected in this population because it is composed of *aus* cultivars. Local genome cluster analysis suggest high Cd alleles are uncommon, but should be avoided in breeding.

Supplementary Information The online version contains supplementary material available at (<https://doi.org/10.1007/s10681-020-02752-1>).

G. J. Norton · A. Travis · P. Ruang-areerate ·
A. A. Adeosun · A. Douglas · A. H. Price (✉)
School of Biological Sciences, University of Aberdeen,
Aberdeen AB24 3UU, UK
e-mail: a.price@abdn.ac.uk

P. Ruang-areerate
National Omics Center, National Science and Technology
Development Agency (NSTDA), Pathum Thani 12120,
Thailand

G. W. Nicol
Laboratoire Ampère, École Centrale de Lyon, Université
de Lyon, 36 avenue Guy de Collongue,
69134 Ecully Cedex, France

M. Hossain · M. R. Islam
Department of Soil Science, Bangladesh Agricultural
University, Mymensingh, Bangladesh

Keywords Rice · Cadmium · Genome wide
association genetics · Alternate wetting and drying

Introduction

Cadmium in rice grains has been highlighted as an issue, specifically for people that consume large quantities of rice, with a proportion of people in Bangladesh and Sri Lanka consuming rice with unsafe amounts of cadmium (Meharg et al. 2013). It is estimated that for non-smokers 90% of cadmium exposure is through food, while for smokers it is 50%. (FAO/WHO 2001; UNEP 2008). Cadmium can accumulate in several organs, including the liver and kidneys, and in humans can result in itai-itai disease (Kobayashi et al. 2009).

One of the main driving factors for cadmium uptake by plants is its availability in soils. Under anaerobic soil management, for example cultivation of rice in a flooded field, cadmium is present as immobile cadmium sulphide (Khaokaew et al. 2011; Hashimoto and Yamaguchi 2013) or cadmium carbonates (Khaokaew et al. 2011). However, in aerobic soils the cadmium is present as the Cd^{2+} ion. It is the Cd^{2+} that is readily taken up by plants. As freshwater for rice irrigation is considered unsustainable in some parts of the world, rice cultivation under different water management strategies may result in an increase in cadmium availability to rice plants. It has been demonstrated that cadmium concentration in rice grains increases when plants are grown under the water saving irrigation practice of alternate wetting and drying (AWD) (Norton et al. 2017a; Li et al. 2019). AWD is an irrigation practise where the standing water in the fields is allowed to drain until either a specific soil water potential is reached or the water height is at a specific level below the soil surface (normally between 10 and 20 cm) (Lampayan et al. 2015). Grain cadmium concentration has been observed to increase by 28–67% in AWD-grown plants compared to plants grown under flooded irrigation practises (Norton et al. 2017a, b).

The uptake, transport, and accumulation of cadmium in plants has been the focus of numerous studies. Key genes involved in the accumulation of cadmium have been reported in rice. For example, *OsHMA3* and *OsHMA2*, which are P1B-type Heavy Metal ATPases, have been demonstrated to be involved in cadmium accumulation. *OsHMA3* is involved in vacuolar sequestration of cadmium (Ueno et al. 2010; Sasaki et al. 2014) while *OsHMA2* facilitates xylem loading of cadmium (Sato-Nagasawa et al. 2012; Takahashi

et al. 2012a). Another key family of genes involved in cadmium accumulation in plants is the Natural Resistance Associated Macrophage Proteins (NRAMP). *OsNRAMP1*, *OsNRAMP2*, and *OsNRAMP5* have been shown to be involved in cadmium transport in rice (Sasaki et al. 2012; Takahashi et al. 2011; Zhao et al. 2018). Other genes involved in uptake and transport include the Low Cadmium gene *OsLCD* (Uraguchi et al. 2011; Shimo et al. 2011), the ATP-binding cassette *ABCG43* (Oda et al. 2011), *CAL1* (Luo et al. 2018), the rice iron-regulated transporters (*OsIRT*) 1 and 2 (Nakanishi et al. 2006), the Low Affinity Cation Transporter *OsLCT 1* (Uraguchi et al. 2011), and *OsZIP1* (Nakanishi et al. 2006). Genes that function as cadmium transporters also function as transporters of other elements (e.g. iron, zinc, and manganese) (Yoneyama et al. 2015), including *OsIRT 1* and 2 (Buglio et al. 2002; Ishimaru et al. 2006), *OsNRAMP5* (Ishimaru et al. 2012), and *OsHMA3* (Ueno et al. 2010; Miyadate et al. 2011).

Mutant, knockout, and overexpression studies targeting the above genes have been shown to alter the accumulation of cadmium in rice grains; examples include studies by Ishikawa et al. (2012), Ueno et al. (2010), and Miyadate et al. (2011). However, these can produce other undesired impacts in terms of alteration of the accumulation of other elements. In a study where *OsNRAMP5* was mutated, it resulted in rice grains with a nearly undetectable concentration of cadmium (Ishikawa et al. 2012), but there was also a decrease in manganese of more than tenfold within the shoots and three fold in the grains compared to the control plants (Ishikawa et al. 2012). In another study where *OsNRAMP5* was knocked out it resulted in plants that accumulate less manganese in the roots and less manganese and iron in the shoots (Ishimaru et al. 2012).

There are significant genotypic differences in grain cadmium concentration in rice. These genotypic differences have enabled the identification of quantitative trait loci (QTL) for cadmium concentration in rice plants. These studies have found numerous QTLs for cadmium concentration in the roots, shoots, and grains of rice plants. Recent advances in genetic mapping in plants have meant that genome wide association (GWA) studies are feasible in rice (Zhao et al. 2011). To date, five GWA studies has been conducted in rice for grain cadmium all in the last two years (Zhang et al. 2018; Zhao et al. 2018; Lui et al.

2019; Tan et al. 2020; Yan et al. 2019) using different populations mostly comprised of *indica* or *japonica* subpopulations or a mixture of both. In respective order they identified 61, 20, 17, 19 and 11 QTLs for grain Cd in rice.

The Bengal and Assam Aus Panel (BAAP), which has 2 million SNPs for GWA mapping, has been specifically created to assess the genetic diversity within the *aus* subpopulation of rice, and allow traits influenced by flowering time to be better studied (see Norton et al. 2018). To emphasise the features of the BAAP; (1) population structure, which is a hindrance to association studies, is reduced (but not eliminated) in the BAAP by using one subpopulation; (2) the *aus* subpopulation is recognised to be understudied (Kim et al. 2016) yet contains wide trait variation for abiotic stress resistance, (3) the influence of weather or soil conditions at flowering time on grain-related traits has been minimised by producing a population with a small flowering window (Norton et al. 2018); (4) the BAAP has been successfully used to map QTLs for grain arsenic that would not have been detected in *indica*, *japonica* or a mixed *indica/japonica* population (Norton et al. 2019).

This study utilises the Bangladesh and Assam Aus Panel (BAAP) (Norton et al. 2018), grown in the field in two consecutive years under two water management strategies—continually flooded (CF) conditions and AWD. The objectives are to identify QTLs and genes that might be useful in breeding lower grain Cd, and to determine if AWD influences the QTLs such that breeding might need to be different for different irrigation regimes.

Methods

Population and field screen

The population used in this study was the Bengal and Assam Aus Panel (BAAP) (Norton et al. 2018). The population consists of 266 landraces identified as belonging to the *aus* subpopulation, in addition to the OryzaSNP panel (McNally et al. 2009) and a number of elite Bangladeshi cultivars (Norton et al. 2018). The 266 *aus* cultivars from the population have been sequenced, and a 2 million SNP database constructed (Norton et al. 2018).

The BAAP was grown in the field in Mymensingh, Bangladesh in the Boro (dry season) of 2013 and 2014. The population was screened under both AWD and continuously flooded (CF). Full details of the field screening is given in Norton et al. (2017a) and Norton et al. (2018). In both years the seeds were initially grown in a nursery bed and transplanted to the field plot after 44 days in 2013 and 51 days in 2014. The plants were transplanted into eight experimental plots (four control and four AWD), two plants per hill with a distance of 20 cm between each hill in a row and a 20 cm distance between each row of 2 m length. Rice accessions were planted in single rows, with a check cultivar BRRI Dhan 28 transplanted into each alternate row. After transplanting the plots were flooded. For the four CF plots the surface water was kept at a depth of between 2 and 5 cm above the soil surface during the vegetative stage and reproductive stage (13th April 2013). For the four AWD plots plastic perforated tubes (pani pipe) were placed across the plots to monitor the water depth. The aim was to allow water to drain/percolate naturally from the AWD plots until the average depth of the water was 15 cm below the soil surface. At that point the plots were irrigated to bring the water depth to between 2 and 5 cm above the soil surface (Norton et al. 2017a). AWD was applied only from 14 days after transplanting until flowering. Once the cultivars had flowered and the grains matured, the grain and shoots from every cultivar was hand harvested from the six central hills of each row. This resulted in one sample for each of the four AWD and one sample for each of the four control plots.

Cadmium analysis

Grain cadmium concentration was determined as described in Norton et al. (2017a). Briefly, rice grains were de-husked and oven dried, followed by microwave digestion with concentrated nitric acid and hydrogen peroxide as described in Norton et al. (2012). Total cadmium analysis was performed by inductively coupled plasma–mass spectroscopy (ICP–MS). Trace element grade reagents were used for all digests, and for quality control replicates of certified reference material rice flour [NIST 1568b]) were used; blanks were also included. All samples and standards contained 10 $\mu\text{g L}^{-1}$ indium as the internal standard.

Genome-wide association (GWA) mapping

Prior to GWA mapping the mean grain cadmium concentration for each treatment and year were calculated only for accessions with values for at least three replicates. GWA mapping was conducted using “PIQUE” (Parallel Identification of QTL’s Using EMMAX, <https://github.com/tony-travis/PIQUE>) to pre-process genotype and phenotype data and then run association analysis for each phenotype in parallel using EMMAX (Kang et al. 2010) as described in Norton et al. (2018).

Clump analysis of significant SNPs

SNPs were binned together into peaks using a sliding window based on the decay of linkage disequilibrium (LD) using the PLINK (Purcell et al. 2007) command—clumpp1 0.000001—clump-p2 0.00001—clump-r2 0.3—clump-kb 243. Note the command—clumpp1 0.000001 indicates a very high P value threshold ($P < 1 \times 10^{-6}$) was employed only to keep the number of detected QTLs manageable. For every SNP with $P < 0.000001$, pairwise r^2 -values were calculated between surrounding SNPs that fell within 243 kb; any two SNPs meeting this criteria that also shared an $r^2 \geq 0.3$ were clumped into bins. Binning was done for each experiment. Clumps with less than two significant SNPs ($P < 0.00001$) were discarded. The LD (Linkage Disequilibrium) decay value of 243 kb was selected as this is the average LD decay for the population (Norton et al. 2018).

Determination of local linkage disequilibrium (LD)

Local LD decay was estimated as r^2 (VanLiere and Rosenberg 2008) using PLINK (Purcell et al. 2007). LD decay within a 1 Mb genomic region centred on the most significant SNP was determined by calculating all pairwise r^2 values between SNPs and binning LD decay according to distance.

Assignment of genotype clusters within candidate regions

A cluster analysis was conducted on the SNPs identified in the QTL defined by clump. Specifically, an alignment of 2002 SNP positions from 260

genotypes was analysed with a simple pairwise comparison distance approach using MEGA7 (Kumar et al. 2016). Each sequence was compared in a pairwise manner with individual nucleotide positions containing a degeneracy (i.e. where different bases were present on two alleles) excluded from the analysis on a pairwise basis. Distances were calculated using the number of differences method (Nei and Kumar 2000) and cladograms constructed using the neighbour-Joining method (Saitou and Nei 1987). Bootstrap support was calculated from 100 replicates (Felsenstein. 1985) and genotypes clustered based on a 95% node support value. Clusters with less than 5 accessions are not reported.

Determination of deletion on chromosome 7 for each accession

As the accessions sequenced in this study have only been sequenced to an average depth of $6.5 \times$ the alignment of each individual accession to the reference sequence can lead to errors in base calling. Therefore, an approach focussed on sequencing density, rather than base calling, was performed to evaluate if accessions had deletion upstream of LOC_Os07g15370. The deletion upstream of LOC_Os07g15370 was identified in the 3000 genome dataset (Wang et al. 2018), starting at genome position 8,964,600 bp on chromosome 7 and ending at 8,964,988 bp (total of 388 bp). The number of sequences aligned to the 500 bp regions flanking (upstream and downstream) of the deletion and the deletion for each accession was determined. The ratio (number of counts for the deletion divided by 388/number of counts in the flanking regions divided by 1000) was used to identify accessions with a deletion and those without.

Statistical analysis

All statistical analyses were performed using the statistical software Minitab v.17 (State College PA) and SigmaPlot v.13 (Systat Software Inc., CA) and significance reported at $\alpha < 0.05$. Analysis of variance across treatments in each year was only conducted on those cultivars present in both treatments ($n = 218$ in year 1 and 207 in year 2).

Results and discussion

Grain cadmium concentration

The mean grain cadmium concentrations recorded for each cultivar in this study are reported in supplementary Table 1. Grain cadmium concentration was significantly different between the plants grown under AWD and CF in both years (Fig. 1, Table 1). AWD increased grain cadmium, on average by 49.6% and 108.8% in year 1 and 2 respectively, compared to the plants grown under CF (Fig. 1, Table 2). Across all experiments there were genotypic differences for grain cadmium, with genotypic differences explaining between 47.5 and 84.1% of the variation observed in the grain cadmium phenotype depending on treatment and year (Table 1). In both years, highly significant genotype by treatment interactions were observed for grain cadmium concentration, explaining 19.2% and 23.7% of the variation (Table 1).

It has been widely reported that AWD alters the accumulation of a range of elements in rice grains, including cadmium (Norton et al. 2017a; Chou et al. 2016; Linqvist et al. 2015; Wang et al. 2014; Yang et al. 2009). For a small sub sample of this population (22 accessions of which 19 are aus and used in the GWAS reported here) the cadmium data reported has been published along-side 16 other grain elements and with results from one additional field site (Norton et al. 2017b). The increased number of accessions reported here has increased the range of grain cadmium concentrations observed in the accessions under both

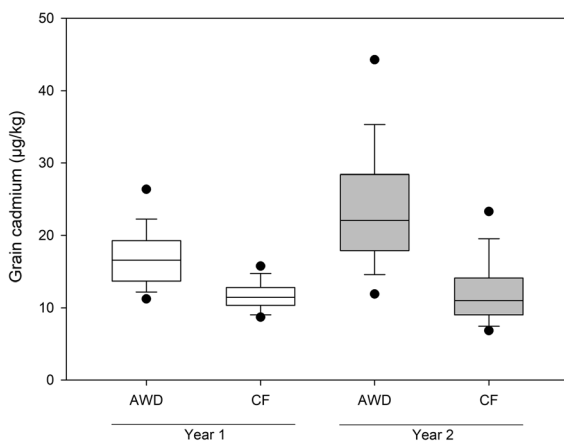


Fig. 1 Difference between the accessions in the BAAP grown under AWD or CF field management in year 1 and year 2

AWD and CF (Table 2). When comparing the relationships of grain cadmium for the cultivars grown under the two different water treatments it is clear that there are a number of cultivars that accumulate high concentrations of cadmium under AWD compared to when grown under CF (Fig. 2). In year 1 these included Aus 37, Rata Boro, and Gouerisail, which accumulated 5.3, 2.8, and 3.1 fold higher cadmium under AWD compared to CF (Fig. 2a). In year 2, Rata Boro, Gouerisail, and Dom Sulfid accumulated 5.5, 5.1, and 4.3 fold higher cadmium under AWD (Fig. 2b). It should be noted that in year 2 Aus 37 had the second highest grain cadmium of the plants grown under AWD in year 1, however it is not presented in Fig. 2, as it was not grown in the CF treatment in year 2. The cultivar Tepa Boro 508 is highlighted in Fig. 2b because it had very high grain cadmium concentration in both AWD and CF in year 2, being three-four times higher than the concentrations in year 1 (Supplementary Table 1). Eight cultivars were consistently found to have low grain cadmium concentrations (Table 3), with ASSAM 4(BORO) and KALIBORO 26 being in the lowest 10% for grain cadmium across all experiments. The observed significant genotype by treatment interaction for grain cadmium (Table 1) has implications for selective breeding for low cadmium cultivars, as these results suggest that different accessions accumulate cadmium through different genetic mechanisms under different water management regimes. In contrast, when exploring other traits, for example grain yield (Norton et al. 2018) and grain arsenic concentration (Norton et al. 2019), it was observed that there would be no specific requirement to breed for these traits under AWD because genotype by treatment interaction was minimal. This is not true for grain cadmium.

The population used in this study is made up of rice accessions from the *aus* subpopulation of rice (Travis et al. 2015; Norton et al. 2018) and therefore does not have the high degree of stratification of some GWA mapping panels, which have representatives from across the different rice subgroups (Zhao et al. 2011). However, there is a degree of structure to the population, with five discernible *aus* subgroups (Norton et al. 2018). There is a significant difference in grain cadmium concentration between these different subgroups for AWD ($P = 0.027$, $df = 4$, $F = 2.88$) and CF ($P < 0.001$, $df = 4$, $F = 5.87$) in year 2 (Supplementary Fig. 1), accounting for 7.2% and 18.1% of the

Table 1 Two-way ANOVA results (F-values) for treatment, genotype and the treatment × genotype interaction for the grain cadmium concentrations across 2013 and 2014

| Year (n) | Trait | 2-way ANOVA F-value | | |
|------------|----------|---------------------|---------------|----------------------------------|
| | | Treatment | Genotype | Treatment × genotype interaction |
| 2013 (218) | Grain Cd | 817 (10.8%) | 13.4 (46.9%) | 6.95 (23.7%) |
| 2014 (207) | Grain Cd | 545 (10.3%) | 10.57 (45.7%) | 4.40 (19.2%) |

Numbers in brackets are the proportion of the variation explain by that factor. All factors are significant at $P < 0.001$

Table 2 Descriptive statistics for grain cadmium concentration measured in 2013 and 2014 for the BAAP accessions grown under AWD and CF

| Year | Treatment | Grain Cd ($\mu\text{g}/\text{kg}$) | | | | | |
|------|-----------|--------------------------------------|-----|-------|--------|------|------|
| | | N | Min | Max | Median | Mean | SD |
| 2013 | AWD | 243 | 9.6 | 142.7 | 16.6 | 17.8 | 10.2 |
| | CF | 226 | 7.0 | 32.1 | 11.5 | 11.9 | 2.8 |
| 2014 | AWD | 220 | 8.9 | 251.1 | 22.1 | 26.1 | 22.2 |
| | CF | 209 | 5.2 | 57.1 | 11.0 | 12.5 | 6.1 |

observed variation. In year 1 there was no significant difference in grain cadmium for the accessions belonging to the different subgroups. Generally, *aus* subgroup 3 has lower grain cadmium concentrations while subgroups 2 and 4 have higher concentrations.

Cadmium GWA mapping

SNP-trait associations were detected for grain cadmium content across the two years of the experiment (Fig. 3). Clump was used to group them for each trait,

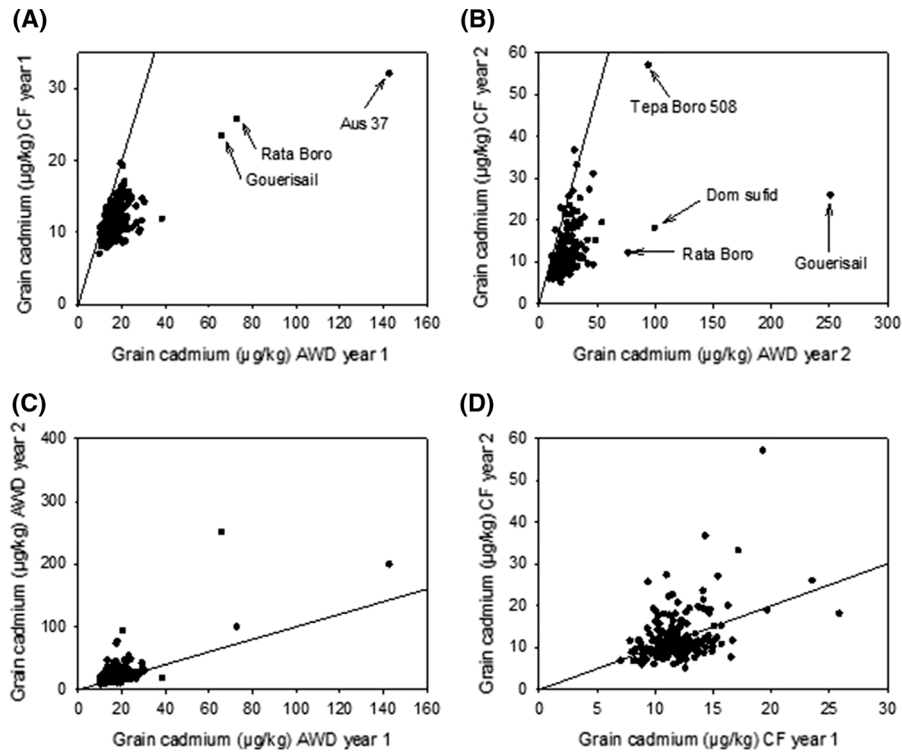


Fig. 2 Correlation between grain cadmium concentration for the BAAP grown under; **a** AWD and CF in year 1, **b** AWD and CF in year 2, **c** AWD in year 1 and AWD in year 2, **d** CF in year 1 and CF in year 2. The line on each graph represents the 1:1 line

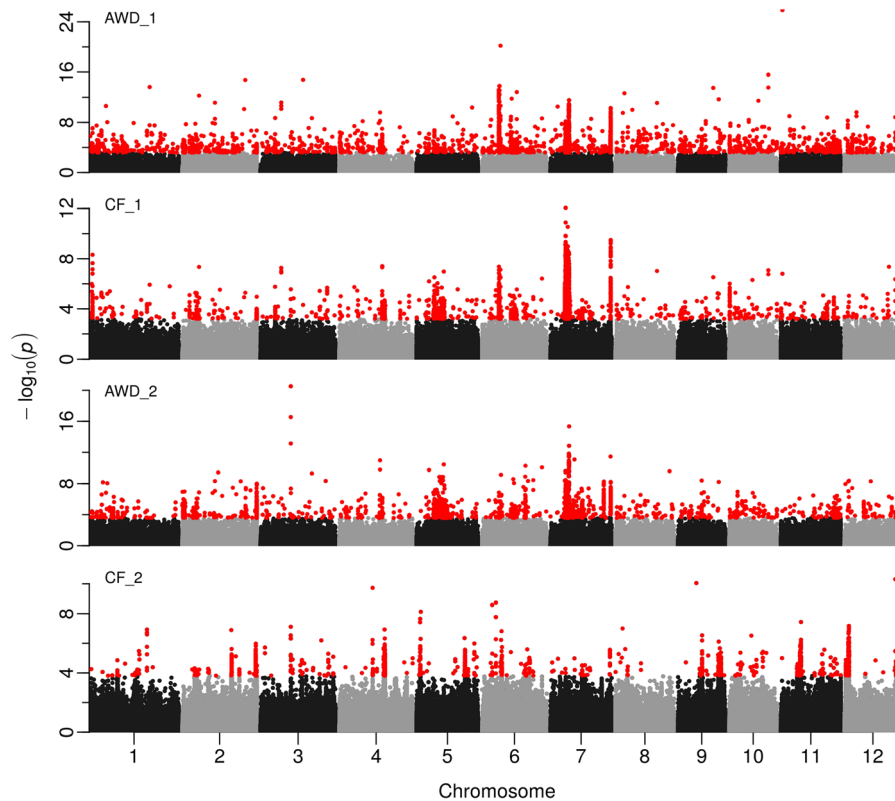
Table 3 Cultivars with consistently low grain cadmium in three or four of the experiments

| BAAP Id | Cultivar name | Year 1 | | Year 2 | |
|---------|----------------|--------|----|--------|----|
| | | AWD | CF | AWD | CF |
| 1 | ASSAM 4(BORO) | X | X | X | X |
| 74 | AUS 364 | X | X | | X |
| 121 | KHAILAORGOABEZ | | X | X | X |
| 136 | KALIBORO 26 | X | X | X | X |
| 151 | KALO BIRA | X | X | | X |
| 222 | Jamir | X | | X | X |
| 243 | Gobir sail | X | X | X | |
| 265 | P 79 | X | X | X | |

An X indicates that that cultivar was in the lowest 10% for grain cadmium concentration in that experiment

revealing 31 groups for AWD1, 24 for CF1, 35 for AWD2 and 13 for CF2 (Supplementary Table 2). These were then combined and further grouped

together as a single QTL if any group was within 243 kb (the global LD for BAAP), revealing a total of 58 QTLs (Supplementary Table 2). Most of these (41) were detected only in one trait (9 for AWD1, 2 for CF1, 22 for AWD2 and 8 for CF2) while 12 were detected in two traits and five were detected in three traits (in each AWD1, CF1 and AWD2). None were detected in all traits. However, it must be noted that a high threshold was used for calling SNP-trait associations ($P < 1 \times 10^{-6}$) which will probably introduce type II error. A probable example can be seen at the bottom of chromosome 7 where all traits have a peak but it is called only for three traits because for CF2 the P value was 9.4×10^{-5} (for SNP at 29.15 Mbp) and therefore not above the threshold used. Even some of the single trait QTLs contained large numbers of SNPs with $P < 0.0001$, (up to 2694 SNPs for AWD2 at 8.17 on chromosome 5). AWD2 was notable for having a very large number of single trait QTLs, while CF2 was notable for having by far the fewest QTLs, only 5 of

**Fig. 3** Manhattan plots from GWA mapping of grain cadmium in field experiments under AWD (Alternate Wetting and Drying) and CF (Continuous Flooding) over years 1 and 2. A

guide-line in blue is shown at $-\log_{10}(0.0001) = 4$. Benjamini-Hochberg adjusted probabilities > 0.1 are highlighted in red. (Color figure online)

which were common to another trait, and none common to two other traits. Examining Fig. 3 also suggests CF2 is different to the other three traits. Previous GWAS studies, using other populations also detected many significant loci for grain cadmium and some co-localize with QTLs in the BAAP. A genetic map of the current results shown with previous studies is given in supplementary Fig. 2. To summarise co-localisation (see supplementary Table 3), Zhang et al. (2018) reported 61 QTLs of which 8 are common with the BAAP, Zhao et al. (2018) reported 20 QTLs of which 4 are common with the BAAP, Lui et al. (2019) reported 17 QTLs of which 1 is common with the BAAP, Tan et al. (2020) reported 19 QTLs of which 3 are common with the BAAP while Yan et al. (2019) reported 11 QTLs of which 2 are common with the BAAP. A striking feature of current and previous Cd QTLs presented in supplementary Fig. 2 is the high number of QTLs, their distribution and relatively small proportion of loci that are common between studies. What distinguishes this study is that here we have detected loci across multiple treatments and years. This has notable benefits when investigating these QTLs further, primarily that this treatment/year stability is essential when breeding rice with reduced cadmium as loci could be of use in a wider array of environments and under a diverse range of management approaches. Five loci are discussed, where co-localised QTLs are detected in at least three of the four of the experiments.

QTL on chromosome 7 between 7.23 and 7.61 Mbp

A QTL was detected for grain cadmium in three out of the four experiments (Supplementary Fig. 2) on chromosome 7 between 7.23 and 7.61 Mbp (as defined by Clump). This region revealed QTLs in the studies of Zhao et al. (2018) and Tan et al. (2020). Cluster analysis in this region revealed 4 clusters, with 12 cultivars from cluster B having higher concentrations of grain cadmium (Fig. 4). This 379 kbp region contains 58 gene models, of which 18 are annotated as either transposons or retrotransposons. Of interest in this region are LOC_Os07g12890 (annotated as a metal cation transporter), *OsZIP8* (Zheng et al. 2018), and LOC_Os07g12900 (annotated as a cadmium/zinc-transporting ATPase *OsHMA3*) (Ueno et al. 2010; Sasaki et al. 2014). *OsZIP8* is a group 1 ZIP

transporter (Zheng et al. 2018) and has been found to be expressed in the roots and panicles of rice, with expression induced under cadmium stress in seedlings (Zheng et al. 2018). In a previous study (using a mapping population derived from Anjana Dhan and Nipponbare) a QTL for cadmium concentration was mapped to this region of chromosome 7, and *OsZIP8* investigated as a potential candidate gene (Ueno et al. 2010). In that study *OsZIP8* showed no difference in expression in the roots of the parents of the population, and the two different alleles of the genes did not have the ability to transport cadmium in yeast experiments; the authors concluded in that study that *OsZIP8* was not the gene responsible for the differential cadmium accumulation observed in the parents (Ueno et al. 2010).

OsHMA3 was the first cloned gene involved in cadmium accumulation in rice (Miyadate et al. 2011; Ueno et al. 2010). *OsHMA3* is located on the tonoplast of the root cells and selectively sequesters cadmium into the vacuole, therefore limiting the translocation of cadmium from the roots to the shoots. Studies have identified two natural non-functional alleles of *OsHMA3*; the cultivars with the non-functional alleles have higher cadmium concentrations in the shoots and grains (Ueno et al. 2010; Yan et al. 2016). The first loss-of-function allele was identified in the *indica* accession Anjana Dhan (Ueno et al. 2010). In a field experiment Anjana Dhan was shown to accumulate higher concentrations of cadmium in the shoots and rice grains compared to Nipponbare (*japonica*), a QTL for cadmium concentration was mapped to a 1.9 Mbp region on chromosome 7 (Ueno et al. 2009) and the functional gene identified as *OsHMA3* (Ueno et al. 2010). The non-functional allele from Anjana Dhan is attributed to an amino acid substitution at the 80th position, changing a His to an Arg (Ueno et al. 2010). When examining this polymorphism in the 3000 rice genome accessions on the SNP-Seek database (www.snpseek.irri.org) (Mansueto et al. 2017; Wang et al. 2018), none of the sequenced *aus* accessions have this polymorphism, and neither do any of the *aus* accessions in the BAAP. To date only one *aus* cultivar has been observed to have this polymorphism (Yan et al. 2016).

A second loss-of-function allele for *OsHMA3* was identified in a *japonica* cultivar (Yelicanghua), which has a predicted amino acid mutation at the 380th position from Ser to Arg. The haplotype had no

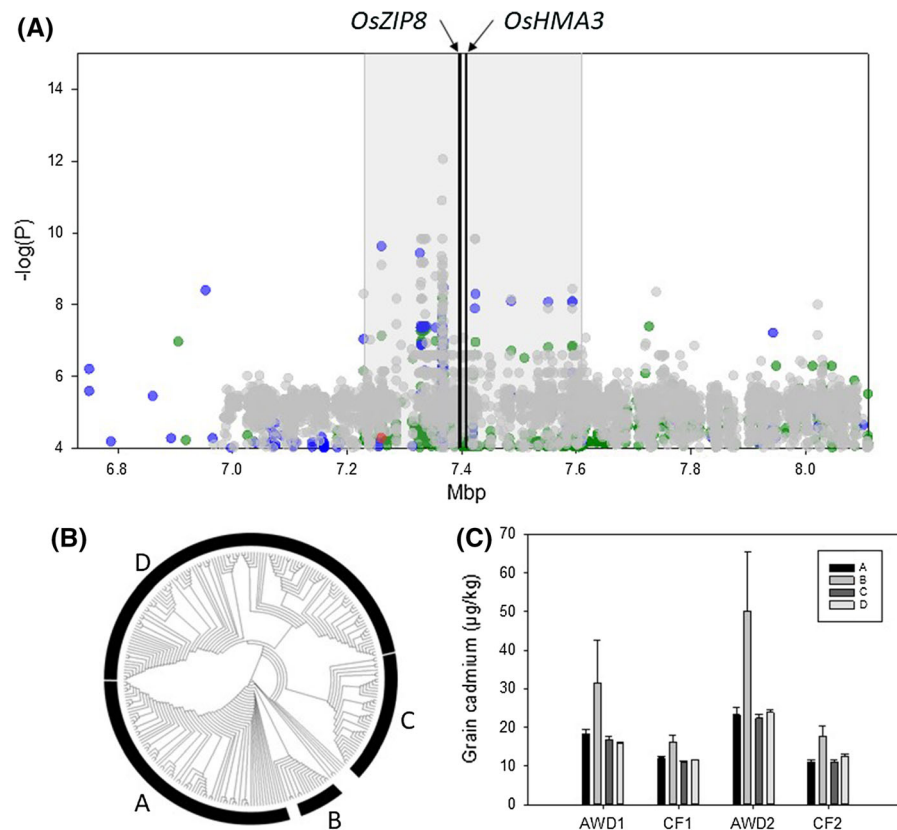


Fig. 4 Significant associations for cadmium traits on chromosome 7 (6.7–8.2 Mb). **a** SNPs associated with the cadmium phenotypes; green symbols = AWD year 1, blue symbols = CF year 1, grey symbols = AWD year 2, red symbols = CF year 2, grey bar indicates the QTL region as defined by Clump, lines indicate the position of two candidate genes. **b** NJT for the

BAAP cultivars for all SNPs located within the Clump region; across the population 71 accessions are in cluster A, 12 in cluster B, 36 in cluster C, and 122 in cluster D. **c** Range of phenotypic variation observed for each cluster across each trait. (Color figure online)

cadmium transport activity when the gene was expressed in yeast, and the allele did not complement a known non-functional allele of *OsHMA3* (Yan et al. 2016). Analysis of the SNPs and sequence data indicates that this is not a polymorphism present in the *aus* accessions in the BAAP. This is supported by the findings of Yan et al. (2016), who only observed this polymorphism in seven accessions from the 950 world rice collection (Huang et al. 2012), with all seven belonging to the temperate *Japonica* subgroup. Another “weak” allele of this gene has recently been described by Sun et al. (2019) but that allele is not present in the BAAP nor in any of the 201 *aus* cultivar in the 3000 rice genome accessions but rather is virtually completely confined to *tropical* and *subtropical japonicas* (based on the 3000 genomes data).

While the *aus* accessions in this population do not have any of the known loss-of-function alleles, *OsHMA3* is still a good candidate gene for this association as both the studies that previously identified the loss-of-function alleles (Ueno et al. 2010; Yan et al. 2016) point to small changes in the DNA sequence having large impacts on the function of this gene. Within the BAAP accessions there are twelve SNPs within the coding regions of *OsHMA3*, five of which are non-synonymous when compared to the Nipponbare sequence. These non-synonymous polymorphisms should be investigated further to see if any cause a change in protein function, and therefore have function effects similar to the previously identified two loss-of-function alleles.

QTL on chromosome 7 between 8.93 and 9.04 Mbp

A QTL was detected for grain cadmium in three out of the four experiments (Fig. 3 and Supplementary Fig. 2) on chromosome 7 between 8.93 and 9.04 Mbp. This region also revealed QTLs in the studies of Ishikawa et al. (2010), Zhao et al. (2018) and Tan et al. (2020). Cluster analysis in this region revealed 3 clusters with cultivars from cluster A (representing only 7 cultivars) having the highest grain cadmium concentrations (Fig. 5). This 115 kbp region contains 20 gene models, of which 7 are annotated as either transposons or retrotransposons. Within this positional candidate gene list is LOC_Os07g15460, which is annotated as metal transporter *Nramp6* and corresponds to *OsNRAMP1* (Mani and Sankaranarayanan 2018). Further analysis of this gene reveals a total of

17 SNPs within *OsNRAMP1* (8,966,023–8,970,882 bp) and these SNPs can be used to classify the accessions into four haplotypes (Fig. 6a). The haplotypes of *OsNRAMP1* correspond to the original clusters generated on SNPs in the whole QTL region: cluster A is haplotype 1, cluster B contains haplotypes 2 and 3, and cluster C is haplotype 4. Out of the 239 accessions where both haplotype and cluster are called, only two accessions don't follow this pattern: BAAP 25 is assigned to cluster C but has haplotype 3 and BAAP 27 is assigned to cluster B but is haplotype 4. There are significant differences in the cadmium concentrations between the different haplotypes in all four experiments, with the cultivars with haplotype 4 having the lowest grain cadmium concentration (Fig. 6b). It must be noted that haplotype 4 is by some way the most common haplotype in the BAAP population. Visual examination of sequence

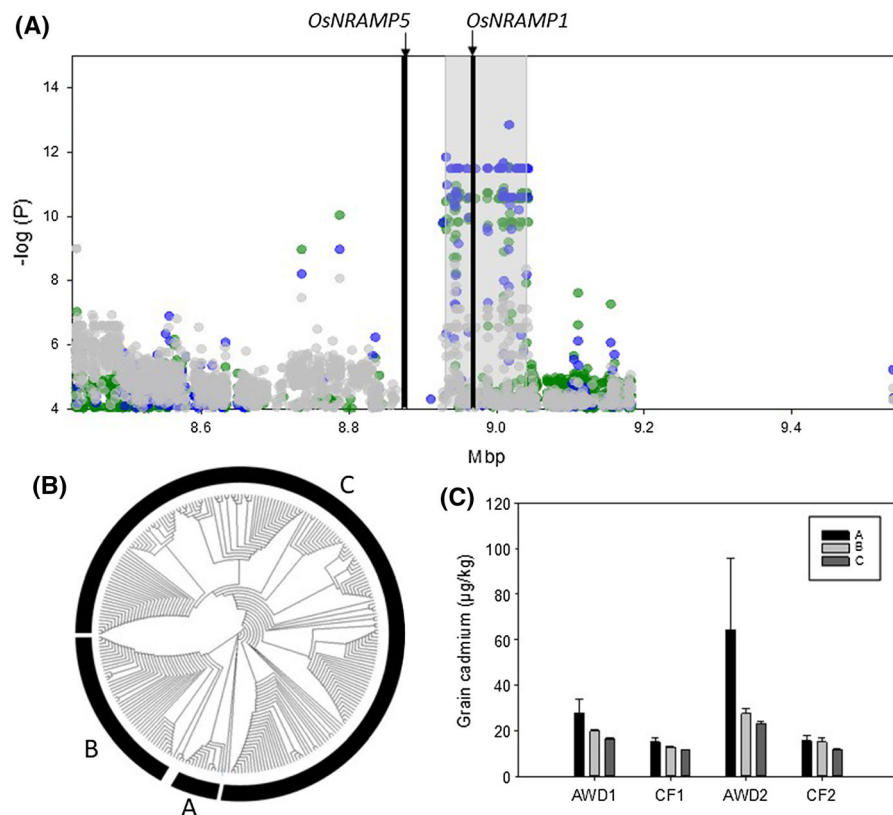


Fig. 5 Significant associations for cadmium traits on chromosome 7 (8.4–9.6 Mb). **a** SNPs associated with the cadmium phenotypes; green symbols = AWD year 1, blue symbols = CF year 1, grey symbols = AWD year 2, red symbols = CF year 2, grey bar indicates the QTL region as defined by Clump, lines

indicate the position of two candidate genes **b** NJT for the BAAP cultivars for all SNPs located within the Clump region; across the population 7 accessions are in cluster A, 43 in cluster B, and 193 in cluster C. **c** Range of phenotypic variation observed for each cluster across each trait. (Color figure online)

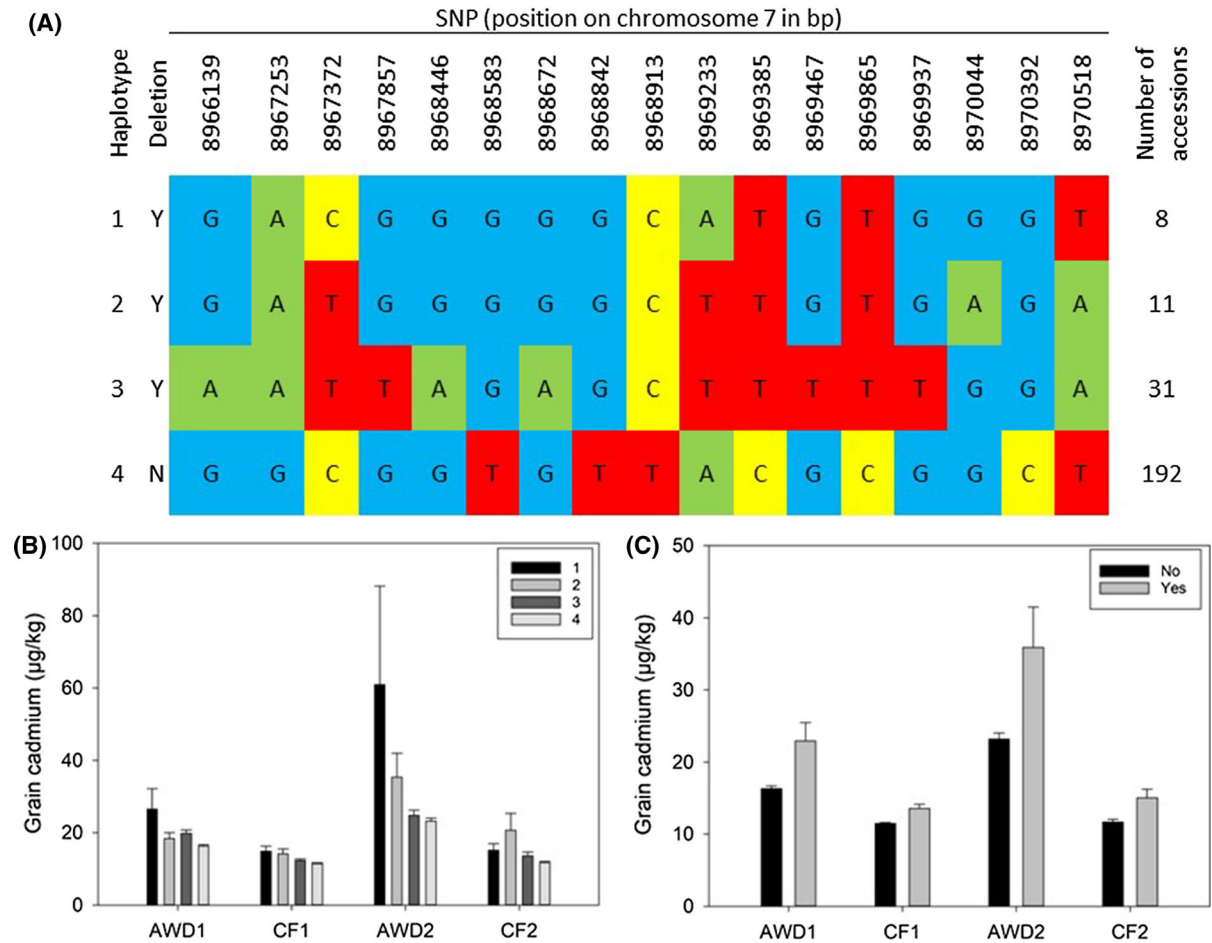


Fig. 6 Haplotypes of *NRAMP1* and presence or absence of upstream deletion. **a** Graphical representation of haplotypes, **b** grain cadmium phenotype of the 4 haplotypes of *Nramp1*,

c grain cadmium phenotype of the accession with and without the deletion upstream of *Nramp1*

alignments using the Integrated Genome Viewer (IGV) suggested that some accessions had a deletion upstream of *OsNRAMP1* (8,964,600 to 8,965,003). Further exploration of the sequence read depth around *OsNRAMP1* (see methods) revealed that this deletion is present in all cultivars from haplotypes 1–3, while cultivars from haplotype 4 do not have the deletion. When the phenotype data is analysed based on the presence or absence of the deletion, the cultivars that do not have the deletion have significantly reduced grain cadmium (Fig. 6c). To understand the distribution of this deletion across the diversity of rice the 3000 genome database (Wang et al. 2018) was examined. Using the informative SNPs used to classify the BAAP into the four different haplotypes, the genotypes from the 3000 genome project were

classified as haplotypes 1–4 or unknown. Haplotype 1 is only found in *indica* and *aus* type rice, haplotype 2 is dominantly *indica* (94.8%) with a small percentage of *aus* (1.5%) and *japonica* (3.6%), haplotype 3 is only present in *indica* and *aus* rice, and haplotype 4 is most common in *japonica* rice (77.8%) and *aus* (12.4%). These results demonstrate that these haplotypes are predominantly separated based on the *japonica* (haplotype 4) and *indica* (haplotypes 1–3) split in rice, however *aus* cultivars have representation in all haplotypes. This observation of balanced allelic variation only in *aus* highlights why *aus*-based populations like BAAP have the potential to reveal some QTLs that are not revealed in other studies.

Initial cloning of *OsNRAMP1* revealed it was predominantly expressed in the roots (Belouchi et al.

1997), while subsequent analysis demonstrated its role in iron transport (Curie et al. 2000) and its important role in cadmium accumulation in rice (Takahashi et al. 2011). *OsNRAMP1* gene expression is increased in the presence of cadmium and arsenic, with significant differences in gene expression between *indica* and *japonica* accessions (Norton et al. 2008; Zhou et al. 2017). When comparing *indica* and *japonica* rice, the *japonica* rice accession, Azucena, has the complete sequence while the *indica* rice accession, Bala, has the deletion in the upstream region. Based on array data Bala has approximately three times the level of expression of *OsNRAMP1* compared to Azucena in rice roots (Norton et al. 2008). Also, *OsNRAMP1* expression in the roots was higher in a high cadmium-accumulating cultivar (Habataki) than a low cadmium-accumulating cultivar (Sasanishiki) regardless of the presence of cadmium, and the amino acid sequence of *OsNRAMP1* showed 100% identity between Sasanishiki and Habataki (Takahashi et al. 2011). Finally, over-expression of *OsNRAMP1* in rice increased cadmium accumulation in the leaves (Takahashi et al. 2011).

However, it must be noted that this QTL is in close proximity to another gene involved in cadmium accumulation in rice: Os07g15370, *OsNRAMP5* (8.87 Mbp). It has been demonstrated that mutational disruption of *OsNRAMP5* dramatically reduces cadmium in rice grains (Ishikawa et al. 2012). However, this gene is not within the candidate region identified by the Clump analysis (Fig. 5) so in this study *OsNRAMP5* is not a candidate gene for this QTL.

Other loci with three colocalised QTLs

A QTL was detected for grain cadmium in three out of the four experiments (Fig. 7a) on chromosome 5 between 8.66 and 8.72 Mbp. This region has not been reported in other studies. Cluster analysis in this region revealed 3 clusters, with cultivars from cluster A having the highest grain cadmium concentrations. This 55.1 kbp region contains 16 genes (9 annotated as either transposons or retrotransposons), however based on current annotation there are no obvious candidate genes for this locus. A QTL was detected for grain cadmium in three out of the four experiments (Fig. 7b) on chromosome 7 between 29.12 and 29.14 Mbp. This region has not been reported in other studies. Cluster analysis in this region revealed 2

clusters, with cultivars from cluster B (the most common cluster) having the lowest grain cadmium concentrations. This 28.8 kbp region contains 4 genes, however based on current annotation there are no obvious candidate genes for this locus. A QTL was detected for grain cadmium in three out of the four experiments (Fig. 7c) on chromosome 9 between 11.46 and 11.64 Mbp. This region has not been reported in other studies. Cluster analysis in this region revealed 2 clusters, however there was no significant difference in grain cadmium concentration between the clusters. This 202.2 kbp region contains 29 genes (11 annotated as either transposons or retrotransposons), however based on current annotation there are no obvious candidate genes for this locus.

This study has demonstrated the complex genetic regulation for the accumulation of cadmium in rice grains. While a genotype by treatment interaction was observed, there were no clear QTLs that corresponded to cadmium concentration for plants grown under AWD compared to those grown under CF. This could be due to the genotype by treatment interaction being driven by a small number of accessions (i.e. those that are distant from the 1:1 line in Fig. 2), however as this number is small it may mean that the criteria used in the GWA mapping cannot detect the potential low frequency alleles responsible for the phenotype. This low power to detect rare alleles (unless the population is large) is a known limitation of GWA studies (Norton et al. 2014).

Conclusion

For grain cadmium concentration five QTLs were detected in three of the four experiments. When exploring the contribution of the different genetic clusters around QTLs to the observed phenotype, it must be noted that in a majority of cases the rare cluster contained the accession with high grain cadmium, meaning that a majority of the tested accessions already had the low grain cadmium alleles. Therefore, for breeding, the important consideration will be to ensure that the high cadmium alleles are selected out. Two of these QTLs are at genes known to be involved in cadmium accumulation. While neither of the previously characterised non-functional alleles of *OsHMA3* were observed in the population, there are a number of sequence polymorphisms worthy of

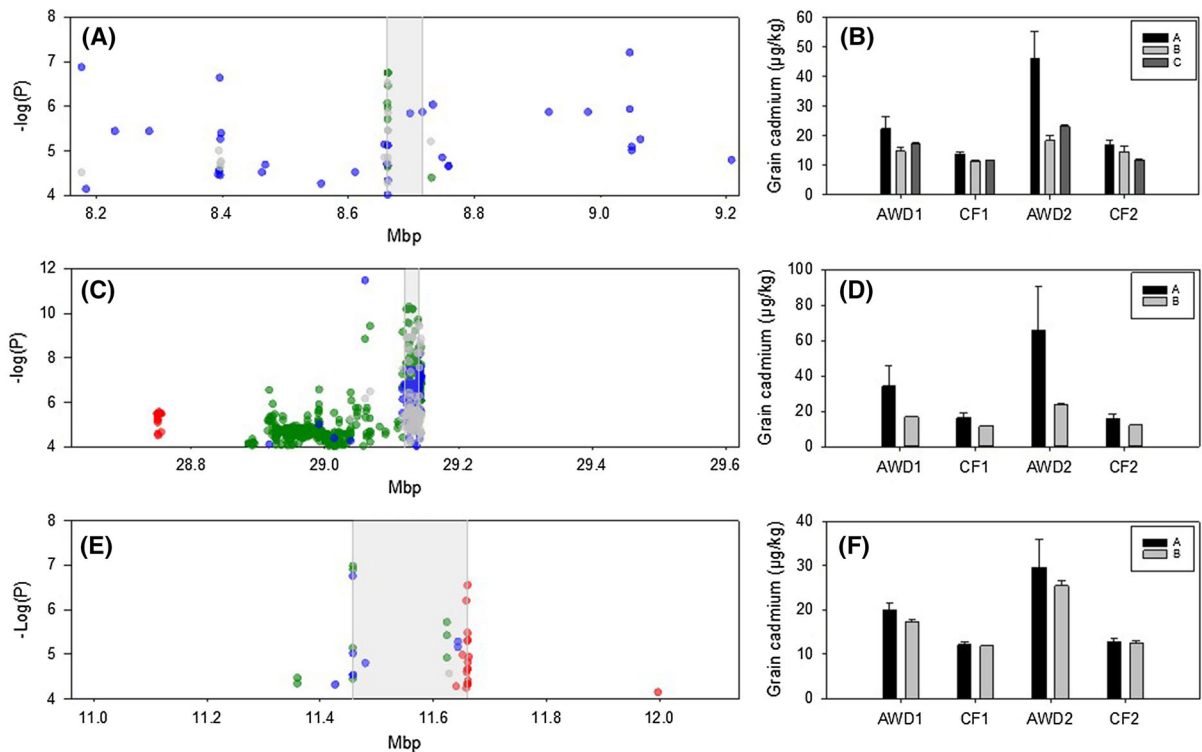


Fig. 7 Significant associations for cadmium traits on three genomic positions. **a** Chromosome 5 (8.18–9.22 Mb), **c** chromosome 7 (28.6–29.6 Mb), **e** chromosome 9 (11.0–12.2 Mb); SNPs associated with the arsenic phenotypes; green symbols = AWD year 1, blue symbols = CF year 1, black symbols = AWD year 2, red symbols = CF year 2. Range of phenotypic variation observed for each cluster across each trait.

b Chromosome 5 (8.18–9.22 Mb) across the population 34 accessions are in cluster A, 8 in cluster B, and 203 in cluster C. **d** chromosome 7 (28.6–29.6 Mb) across the population 12 accessions are in cluster A, and 233 in cluster B. **f** chromosome 9 (11.0–12.2 Mb) across the population 47 accessions are in cluster A, and 200 in cluster B. (Color figure online)

further analysis. Additionally, *OsNRAMP1* appears to be an excellent candidate gene for grain cadmium accumulation in this population, not least since many accessions have a large deletion in the promoter associated with higher cadmium. Finally, the impact that water management has on grain cadmium concentration must be taken into account in the future when water saving irrigation practices are being implemented. Utilisation of water management practices with selected rice cultivars should overcome this. This is supported by the study of Ishikawa et al. (2016) where a low grain cadmium cultivar (using a cultivar with an artificially mutated *OsNRAMP5* gene) in combination with a water management technique resulted in rice with both reduced arsenic and cadmium. Future work should focus on applying this same approach but utilising natural allelic variation.

Acknowledgements Plant material was imported into the UK under import licence IMP/SOIL/19/2014. The authors would like to thank reviewers of the manuscript who helped improve the presentation.

Author contributions Conception: GJN and AHP; Funding Acquisition: GJN and AHP, Managing Field Experiments; MHS and MRI, Data Analysis: GJN, AT, GWN, AD and AHP, Bioinformatics: AAA, GJN, AHP, Manuscript writing: GJN and AHP. Figure Production: PR.

Funding The bulk of this work was supported by the Biotechnology and Biological Sciences Research Council, mostly from Project BB/J003336/1. A small part of the work by AJT was supported by Project BB/N013492/1 (NEWS-India-UK). PR was financially supported by a Royal Thai Government Scholarship and National Science and Technology Development Agency (NSTDA), Thailand and AAA supported by the Elphinstone Scholarship Scheme (University of Aberdeen).

Availability of data and materials The 2 M SNP genotype data of the BAAP is available as a Project called “BAAP” at the

SNP-Seek database (<http://snp-seek.irri.org/>). The phenotype data used for the GWAS is presented in Supplementary Table 1.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there are no conflicts of interest.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Belouchi A, Kwan T, Gros P (1997) Cloning and characterization of the *OsNramp* family from *Oryza sativa*, a new family of membrane proteins possibly implicated in the transport of metal ions. *Plant Mol Biol* 33:1085–1092
- Bughio N, Yamaguchi H, Nishizawa NK, Nakanishi H, Mori S (2002) Cloning an iron regulated metal transporter from rice. *J Exp Bot* 53:1677–1682
- Chou M-L, Jean J-S, Sun G-X, Yang C-M, Hseu Z-Y, Kuo S-F, Tseng H-Y, Yang Y-J (2016) Irrigation practices on rice crop production in arsenic-rich paddy soil. *Crop Sci* 56:422–431
- Curie C, Alonso JM, Le Jean M, Ecker JR, Briat JF (2000) Involvement of NRAMP1 from *Arabidopsis thaliana* in iron transport. *Biochem J* 347:749–755
- FAO/WHO (Food, Agriculture Organization/World Health Organization) (2001) Evaluation of certain food additives and contaminants (fifty-fifth report of the joint FAO/WHO expert committee on food additives). WHO Technical Report Series, No. 901
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Hao X, Zeng M, Wang J, Zeng Z, Dai J, Xie Z, Yang Y, Tian L, Chen L, Li D (2018) A node-expressed transporter OsCCX2 is involved in grain cadmium accumulation of rice. *Front Plant Sci* 9:476
- Hashimoto Y, Yamaguchi N (2013) Chemical speciation of cadmium and sulfur K-edge XANES spectroscopy in flooded paddy soils amended with zerovalent iron. *Soil Sci Soc Am J* 77:1189–1198
- Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W et al (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet* 44:32–39
- Ishikawa S, Abe T, Kuramata M, Yamaguchi M, Ando T, Yamamoto T (2010) A major quantitative trait locus for increasing cadmium-specific concentration in rice grain is located on the short arm of chromosome 7. *J Exp Bot* 61:923–934
- Ishikawa S, Ishimaru Y, Igura M, Kuramata M, Abe T, Senoura T, Hase Y, Arao T, Nishizawa NK, Nakanishi H (2012) Ion-beam irradiation, gene identification, and marker-assisted breeding in the development of low-cadmium rice. *PNAS* 109:19166–19171
- Ishikawa S, Makino T, Ito M, Harada K, Nakada H, Nishida I, Nishimura M, Tokunaga T, Shirao K, Yoshizawa C, Matsuyama M, Abe T, Arao T (2016) Low cadmium rice (*Oryza sativa* L.) cultivar can simultaneously reduce arsenic and cadmium concentrations in rice grains. *Soil Sci Plant Nutr* 62:327–339
- Ishimaru Y, Suzuki M, Tsukamoto T, Suzuki K, Nakazono M, Kobayashi T, Wada Y, Watanabe S, Matsuhashi S, Takahashi M, Nakanishi H, Mori S, Nishizawa NK (2006) Rice plants take up iron as an Fe³⁺-phytosiderophore and as Fe²⁺. *Plant J* 5:335–346
- Ishimaru Y, Kakei Y, Shimo H, Bashir K, Sato Y, Sato Y, Uozumi N, Nakanishi H, Nishizawa NK (2011) A rice phenolic efflux transporter is essential for solubilizing precipitated apoplasmic iron in the plant stele. *J Biol Chem* 286:24649–24655
- Ishimaru Y, Takahashi R, Bashir K, Shimo H, Senoura T, Sugimoto K, Ono K, Yano M, Ishikawa S, Arao T, Nakanishi H, Nishizawa NK (2012) Characterizing the role of rice NRAMP5 in manganese, iron and cadmium transport. *Sci Rep* 2:286
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354
- Khaokaew S, Chaney RL, Landrot G, Ginder-Vogel M, Sparks DL (2011) Speciation and release kinetics of cadmium in an alkaline paddy soil under various flooding periods and draining conditions. *Environ Sci Technol* 45:4249–4255
- Kim HJ, Jung J, Singh N, Greenberg A, Doyle JJ, Tyag W et al (2016) Population dynamics among six major groups of the *Oryza rufipogon* species complex, wild relative of cultivated Asian rice. *Rice* 9:56
- Kobayashi E, Suwazono Y, Dochi M, Honda R, Kido T (2009) Influence of consumption of cadmium-polluted rice or Jinzu River water on occurrence of renal tubular dysfunction and/or Itai-itai disease. *Biol Trace Elem Res* 127:257–268
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874
- Kuramata M, Masuya S, Takahashi Y, Kitagawa E, Inoue C, Ishikawa S, Youssefian S, Kusano T (2009) Novel cysteine-rich peptides from *Digitaria ciliaris* and *Oryza sativa* enhance tolerance to cadmium by limiting its cellular accumulation. *Plant Cell Physiol* 50:106–117
- Lampayan RM, Rejesus RM, Singleton GR, Bouman BAM (2015) Adoption and economics of alternate wetting and

- drying water management for irrigated lowland rice. *Field Crop Res* 170:95–108
- Li C, Carrizo DR, Nakayama Y, Linquist BA, Green PG, Parikh SJ (2019) Impact of alternate wetting and drying irrigation on arsenic uptake and speciation in flooded rice systems. *Agric Ecosyst Environ* 272:188–198
- Linquist BA, Anders MM, Adviento-Borbe MAA, Chaney RL, Nalley LL, da Rosa EFF, van Kessel C (2015) Reducing greenhouse gas emissions, water use, and grain arsenic levels in rice systems. *Glob Change Biol* 21:407–417
- Lee S, An G (2009) Over-expression of OsIRT1 leads to increased iron and zinc accumulations in rice. *Plant Cell Environ* 32:408–416
- Liu X, Chen S, Chen M, Zheng G, Peng Y, Shi X, Qin P, Xu X, Teng S (2019) Association study reveals genetic loci responsible for arsenic, cadmium and lead accumulation in rice grain in contaminated farmlands. *Front Plant Sci* 10:61. <https://doi.org/10.3389/fpls.2019.00061>
- Liu XS, Feng SJ, Zhang BQ, Wang MQ, Cao HW, Rono JK, Chen W, Yang ZM (2019) OsZIP1 functions as a metal efflux transporter limiting excess zinc, copper and cadmium accumulation in rice. *BMC Plant Biol* 19:283
- Luo JS, Huang J, Zeng DL, Peng JS, Zhang GB, Ma HL, Guan Y, Yi HY, Fu YL, Han B, Lin HX, Qian Q, Gong JM (2018) A defensin-like protein drives cadmium efflux and allocation in rice. *Nat Commun* 9:645
- Mani A, Sankaranarayanan K (2018) In silico analysis of natural resistance-associated macrophage protein (Nramp) family of transporters in rice. *Protein J* 37:237–247
- Mansueto L, Fuentes RR, Borja FN, Detras J, Abriol-Santos JM, Chebotarov D et al (2017) Rice SNP-seek database update: new SNPs, indels, and queries. *Nucl Acids Res* 45:D1075–D1081. <https://doi.org/10.1093/nar/gkw1135>
- McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, Stokowski R, Ballinger DG, Frazer KA, Cox DR, Padhukasahasram B, Bustamante CD, Weigel D, Mackill DJ, Bruskiewich RM, Ratsch G, Buell CR, Leung H, Leach JE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *PNAS* 106:12273–12278
- Meharg AA, Norton G, Deacon C, Williams P, Adomako EE, Price A, Zhu Y, Li G, Zhao F-J, McGrath S, Villada A, Sommella A, De Silva PMCS, Brammer H, Dasgupta T, Islam MR (2013) Variation in rice cadmium related to human exposure. *Environ Sci Tech* 47:5613–5618
- Miyadate H, Adachi S, Hiraizumi A, Tezuka K, Nakazawa N, Kawamoto T, Katou K, Kodama I, Sakurai K, Takahashi H, Satoh-Nagasawa N, Watanabe A, Fujimura T, Akagi H (2011) OsHMA3, a P1B-type of ATPase affects root-to-shoot cadmium translocation in rice by mediating efflux into vacuoles. *New Phytol* 189:190–9
- Nakanishi H, Ogawa I, Ishimaru Y, Mori S, Nishizawa NK (2006) Iron deficiency enhances cadmium uptake and translocation mediated by the Fe2⁺ transporters OsIRT1 and OsIRT2 in rice. *Soil Sci Plant Nutr* 52:464–469
- Nei M, Kumar S (2000) *Molecular evolution and phylogenetics*. Oxford University Press, New York
- Norton GJ, Lou-Hing DE, Meharg AA, Price AH (2008) Rice-arsenate interactions in hydroponics: whole genome transcriptional analysis. *J Ex Bot* 59:2267–2276
- Norton GJ, Deacon CM, Xiong L, Huang S, Meharg AA, Price AH (2010) Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. *Plant Soil* 329:139–153
- Norton GJ, Duan G-L, Lei M, Zhu Y-G, Meharg AA, Price AH (2012) Identification of quantitative trait loci for rice grain element composition on an arsenic impacted soil: Influence of flowering time on genetic loci. *Ann Appl Biol* 161:46–56
- Norton GJ, Douglas A, Lahner B, Yakubova E, Guerinot ML et al (2014) Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS ONE* 9:e89685
- Norton GJ, Shafaei M, Travis AJ, Deacon CM, Danku J, Pond D, Cochran N, Lockhart K, Salt D, Zhang H, Dodd IC, Hossain M, Islam MR, Price AH (2017a) Impact of alternate wetting and drying on rice physiology, grain production, and grain quality. *Field Crops Res* 205:1–13
- Norton GJ, Travis AJ, Danku JMC, Salt DE, Hossain M, Islam MR, Price AH (2017b) Biomass and elemental concentrations of 22 rice cultivars grown under alternate wetting and drying conditions at three field sites in Bangladesh. *Food Energy Secur* 6:98–112
- Norton GJ, Travis AJ, Douglas A, Fairley S, De Paiva Alves E, Ruang-areerate P, Naredo MEB, McNally KL, Hossain M, Islam MR, Price AH (2018) Genome wide association mapping of grain and straw biomass traits in the rice Bengal and Assam Aus panel (BAAP) grown under alternate wetting and drying and permanently flooded irrigation. *Front Plant Sci* 9:1223
- Norton G, Travis AJ, Talukdar P, Sumon MH, Rafiqul M, Price AH (2019) Genetic loci regulating arsenic content in rice grains when grown flooded or under alternative wetting and drying irrigation. *Rice* 12:54
- Oda K, Otani M, Uraguchi S, Akihiro T, Fujiwara T (2011) Rice ABCG43 is Cd inducible and confers Cd tolerance on yeast. *Biosci Biotech Biochem* 5:1211–1213
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575
- Satoh-Nagasawa N, Mori M, Nakazawa N, Kawamoto T, Nagato Y, Sakurai K, Takahashi H, Watanabe A, Akagi H (2012) Mutations in rice (*Oryza sativa*) heavy metal ATPase 2 (*OsHMA2*) restrict the translocation of zinc and cadmium. *Plant Cell Physiol* 53:213–224
- Sasaki A, Yamaji N, Yokosho K, Ma JF (2012) Nramp5 is a major transporter responsible for manganese and cadmium uptake in rice. *Plant Cell* 24:2155–2167
- Sasaki A, Yamaji N, Ma JF (2014) Overexpression of OsHMA3 enhances Cd tolerance and expression of Zn transporter genes in rice. *J Exp Bot* 65:6013–6021
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Shim D, Hwang JU, Lee J, Lee S, Choi Y, An G, Martinoia E, Lee Y (2009) Orthologs of the class A4 heat shock

- transcription factor HsfA4a confer cadmium tolerance in wheat and rice. *Plant Cell* 21:4031–4043
- Shimo H, Ishimaru Y, An G, Yamakawa T, Nakanishi H, Nishizawa NK (2011) Low cadmium (LCD), a novel gene related to cadmium tolerance and accumulation in rice. *J Exp Bot* 5:5727–5734
- Sun C, Yang M, Li Y, Tian J, Zhang Y, Liang L, Liu Z, Chen K, Li Y, Lv K, Lian X (2019) Comprehensive analysis of variation of cadmium accumulation in rice and detection of a new weak allele of *OsHMA3*. *J Ex Bot* 70:6389–6400. <https://doi.org/10.1093/jxb/erz400>
- Takahashi R, Ishimaru Y, Senoura T, Shimo H, Ishikawa S, Arao T, Nakanishi H, Nishizawa NK (2011) The OsN-RAMP1 iron transporter is involved in Cd accumulation in rice. *J Exp Bot* 62:4843–4850
- Takahashi R, Ishimaru Y, Shimo H, Ogo Y, Senoura T, Nishizawa NK, Nakanishi H (2012) The OsHMA2 transporter is involved in root-to-shoot translocation of Zn and Cd in rice. *Plant Cell Environ* 35:1948–57
- Takahashi R, Bashir K, Ishimaru Y, Nishizawa NK, Nakanishi H (2012) The role of heavy metal ATPases, HMAs, in zinc and cadmium transport in rice. *Plant Sign Behav* 7:1605–1607
- Tan Y, Sun L, Song Q, Mao D, Zhou J, Jiang Y, Wang J, Fan T, Zhu Q, Huang D, Xiao H, Chen C (2020) Genetic architecture of subspecies divergence in trace mineral accumulation and elemental correlations in the rice grain. *Theor Appl Genet* 133:529–545. <https://doi.org/10.1007/s00122-019-03485-z>
- Travis AJ, Norton GJ, Datta S, Sarma R, Dasgupta T, Savio FL, Macaulay M, Hedley PE, McNally KL, Sumon MH, Islam MR, Price AH (2015) Assessing the genetic diversity of rice originating from Bangladesh. *Assam West Bengal Rice* 8:35
- Ueno D, Koyama E, Kono I, Ando T, Yano A, Ma JF (2009) Identification of a novel quantitative trait locus controlling distribution of Cd between roots and shoots in rice. *Plant Cell Physiol* 50:2223–2233
- Ueno D, Yamaji N, Kono I, Huang CF, Ando T, Yano M, Ma JF (2010) Gene limiting cadmium accumulation in rice. *PNAS* 107:16500–16505
- UNEP (United Nations Environment Programme) (2008) Draft final review of scientific information on cadmium
- Uraguchi S, Kamiya T, Sakamoto T, Kasai K, Sato Y, Nagamura Y, Yoshida A, Kyoizuka J, Ishikawa S, Fujiwara T (2011) Low-affinity cation transporter (OsLCT1) regulates cadmium transport into rice grains. *PNAS* 5:20959–20964
- VanLiere JM, Rosenberg NA (2008) Mathematical properties of the r^2 measure of linkage disequilibrium. *Theor Popul Biol* 74:130–137
- Wang Y, Wei Y, Dong L, Lu L, Feng Y, Zhang J, Pan F, Yang X (2014) Improved yield and Zn accumulation for rice grain by Zn fertilisation and optimized water management. *J Zhenjiang Univ Sci B* 15:365–374
- Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z et al (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557:43–49
- Xu X, Liu X, Ge S et al (2011) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat Biotech* 30:105–111
- Yan J, Wang P, Wang P, Yang M, Lian X, Tang Z, Huang CF, Salt DE, Zhao FJ (2016) A loss-of-function allele of *OsHMA3* associated with high cadmium accumulation in shoots and grain of Japonica rice cultivars. *Plant Cell Environ* 39:1941–54
- Yan H, Xu W, Xie J, Gao Y, Wu L, Sun L, Feng L, Chen X, Zhang T, Dai C, Li T, Lin X, Zhang Z, Wang X, Li F, Zhu X, Li J, Li Z, Chen C, Ma M, Zhang H, He Z (2019) Variation of a major facilitator superfamily gene contributes to differential cadmium accumulation between rice subspecies. *Nat Commun* 10:2562
- Yang JC, Huang D, Duan H, Tan G, Zhang J (2009) Alternate wetting and moderate drying increase grain yield and reduces cadmium accumulation in rice grains. *J Sci Food Agri* 89:1728–1736
- Yoneyama T, Ishikawa S, Fujimaki S (2015) Route and regulation of zinc, cadmium, and iron transport in rice plants (*Oryza sativa* L.) during vegetative growth and grain filling: metal transporters, metal speciation, grain Cd reduction and Zn and Fe biofortification. *Int J Mol Sci* 16:19111–19129
- Yuan L, Yang S, Liu B, Zhang M, Wu K (2012) Molecular characterization of a rice metal tolerance protein, OsMTP1. *Plant Cell Rep* 31:67–79
- Zhang M, Pinson SRM, Tarpley L, Huang XY, Lahner B, Yakubova E, Baxter I, Lou Guerinot M, Salt DE (2014) Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theor Appl Genet* 127:137–165
- Zhang G-M, Zheng T-Q, Chen Z, Wang Y-L, Wang Y, Shi Y-M, Wang C-C, Zhang L-Y, Ma J-T, Deng L-W, Li W, Xu T-T, Liang C-Z, Xu J-L, Li Z-K (2018) Joint exploration of favorable haplotypes for mineral concentrations in milled grains of rice (*Oryza sativa* L.). *Front Plant Sci* 9:447. <https://doi.org/10.3389/fpls.2018.00447>
- Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun* 2:467
- Zhao J, Yang W, Zhang S, Yang T, Liu Q, Dong J, Fu H, Mao X, Liu B (2018) Genome-wide association study and candidate gene analysis of rice cadmium accumulation in grain in a diverse rice collection. *Rice* 21;11(1):61.
- Zheng X, Chen L, Li XF (2018) Arabidopsis and rice showed a distinct pattern in ZIPs genes expression profile in response to Cd stress. *Bot Stud* 59:22. <https://doi.org/10.1186/s40529-018-0238-6>
- Zhou Q, Shao G-S, Zhang Y-X, Dong Q, Wang H, Cheng S-H, Cao L-Y, Shen X-H (2017) The difference of cadmium accumulation between the indica and japonica subspecies and the mechanism of it. *Plant Grow Regul* 81:523–532

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