



ARTICLE

Genome-wide association mapping for grain manganese in rice (*Oryza sativa* L.) using a multi-experiment approach

Panthita Ruang-areerate^{1,2} · Anthony J. Travis¹ · Shannon R. M. Pinson³ · Lee Tarpley⁴ · Georgia C. Eizenga³ · Mary Lou Guerinot⁵ · David E. Salt⁶ · Alex Douglas¹ · Adam H. Price¹ · Gareth J. Norton¹

Received: 10 April 2020 / Revised: 6 November 2020 / Accepted: 6 November 2020 / Published online: 24 November 2020
© The Author(s), under exclusive licence to The Genetics Society 2020

Abstract

Manganese (Mn) is an essential trace element for plants and commonly contributes to human health; however, the understanding of the genes controlling natural variation in Mn in crop plants is limited. Here, the integration of two of genome-wide association study approaches was used to increase the identification of valuable quantitative trait loci (QTL) and candidate genes responsible for the concentration of grain Mn across 389 diverse rice cultivars grown in Arkansas and Texas, USA, in multiple years. Single-trait analysis was initially performed using three different SNP datasets. As a result, significant loci could be detected using the high-density SNP dataset. Based on the 5.2 M SNP dataset, major QTLs were located on chromosomes 3 and 7 for Mn containing six candidate genes. In addition, the phenotypic data of grain Mn concentration were combined from three flooded-field experiments from the two sites and 3 years using multi-experiment analysis based on the 5.2 M SNP dataset. Two previous QTLs on chromosome 3 were identified across experiments, whereas new Mn QTLs were identified that were not found in individual experiments, on chromosomes 3, 4, 9 and 11. *OsMTP8.1* was identified in both approaches and is a good candidate gene that could be controlling grain Mn concentration. This work demonstrates the utilisation of multi-experiment analysis to identify constitutive QTLs and candidate genes associated with the grain Mn concentration. Hence, the approach should be advantageous to facilitate genomic breeding programmes in rice and other crops considering QTLs and genes associated with complex traits in natural populations.

Associate editor: Yuan-Ming Zhang

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-020-00390-w>) contains supplementary material, which is available to authorised users.

✉ Panthita Ruang-areerate
panthita.rua@nstda.or.th

¹ Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen AB24 3UU, UK

² National Omics Center, National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand

³ USDA-ARS Dale Bumpers National Rice Research Center, Stuttgart, AR 72160, USA

⁴ Texas A&M System AgriLife Research Center, Beaumont, TX 77713, USA

⁵ Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA

⁶ Future Food Beacon of Excellence and the School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK

Introduction

Genome-wide association (GWA) mapping is a powerful approach to identify genetic loci associated with complex traits in natural populations. The approach has been successfully applied in plants such as *Arabidopsis thaliana* (Atwell et al. 2010; Baxter et al. 2010), maize (Kump et al. 2011) and rice (Huang et al. 2010, 2012; Zhao et al. 2011; Norton et al. 2014; McCouch et al. 2016; Bettembourg et al. 2017) for identifying important agronomic, disease resistance and ionomic (the elemental composition of biological samples) loci. GWA mapping for ionomic traits in plants has been commonly used to perform QTL analysis related on a single trait in individual experiments (Atwell et al. 2010; Baxter et al. 2010; Li et al. 2010; Norton et al. 2014, 2018; Dimkpa et al. 2016; Yang et al. 2018). However, these studies have not always identified QTLs for a trait from several experiments. There are several reasons why a number of ionomic QTLs have not been consistently detected, including a different range of phenotypic values in

experiments due to large environmental effects such as different geographic locations and climate data (temperature and humidity) (Pinson et al. 2015; Huang and Salt 2016), and complex genetic architecture such as distributed allelic variants, each of which have small effects (Korte and Farlow 2013). These reasons may reduce the power of statistical association tests in GWA mapping using single-trait analysis in only one experiment.

Recently, multi-trait approaches have been developed to improve QTL detection by increasing the statistical power with correlated traits and multiple experiments (Korte et al. 2012; Lippert et al. 2014; Zhou and Stephens 2014; Loh et al. 2015). Zhou and Stephens (2014) developed a genome-wide efficient mixed-model association (GEMMA) software for testing multiple traits for each single genetic marker with a multivariate linear mixed model (mvLMM), which controls population stratification and accounts for relatedness between individuals. It was developed from the efficient mixed-model association (EMMA) algorithm for single-trait analysis (Kang et al. 2008) with which identified QTLs could be compared with multi-trait analysis based on the EMMA algorithm.

Central to effective GWA mapping is the population that is used. The Rice Diversity Panel 1 (RDP1) is a rice panel representing the broad range of rice varieties from >70 countries (Eizenga et al. 2014). This panel was initially genotyped using 44,100 SNPs (Zhao et al. 2011), subsequently it was genotyped with 700,000 SNPs (McCouch et al. 2016) and in the latest iteration 5,231,435 SNPs have been imputed on this panel by comparing the 700,000 SNPs with whole-genome sequence data of the 3000 sequenced rice cultivars (Wang et al. 2018).

Manganese (Mn) is an essential trace element for plants and humans. It is an important cofactor or activator of many enzymes, and is involved in photosynthesis in plants (Marschner 1995; Soetan et al. 2010). Mn deficiency in plants can cause a reduction in growth and yield (Marschner 1995; Hebbert et al. 2005), whereas if the Mn concentration is elevated, it can become toxic to plants (Ducic and Polle 2005; Millaleo et al. 2010). Mn homeostasis in the embryo is required for efficient seed germination (Eroglu et al. 2017). Mn deficiency in humans is rare; however, it can lead to a range of health impacts, including severe birth defects, impaired reproductive functions, skeletal defects and asthma (Bashir et al. 2013), while overexposure can lead to neurological disorders (Crossgrove and Zheng 2004; O'Neal and Zheng 2015). The recommended daily intake of Mn for an adult is 2.3 mg day⁻¹, while the recommended tolerable levels are 11 mg day⁻¹ (Institute of Medicine 2001). Rice grain concentrations of Mn are variable, but a recent dietary study of rice consumers in West Bengal, India, demonstrated that rice alone can contribute between 0.82 and 4.21 mg day⁻¹ for an adult, 35.7–183% of the recommend daily amount of Mn (Halder et al. 2020).

The mechanisms of Mn uptake, transport, accumulation and detoxification have been studied in plants (Ducic and Polle 2005; Millaleo et al. 2010; Socha and Guerinet 2014). For rice, a number of Mn-transporter genes have been identified such as natural resistance-associated macrophage protein 3 (*OsNRAMP3*), *OsNRAMP5* and *OsNRAMP6* (Ishimaru et al. 2012; Sasaki et al. 2012; Yang et al. 2013, 2014; Peris-Peris et al. 2017), yellow stripe-like protein 2 (*OsYSL2*) and *OsYSL6* (Koike et al. 2004; Ishimaru et al. 2010; Sasaki et al. 2011) and cation diffusion facilitator/metal tolerance protein 8.1 (*OsMTP8.1*), *OsMTP8.2*, *OsMTP9* and *OsMTP11* (Chen et al. 2013, 2016; Ueno et al. 2015; Takemoto et al. 2017; Zhang and Liu 2017; Ma et al. 2018). In addition to these genes, a number of studies have identified QTLs of grain Mn concentration in rice grains based on biparental mapping (Stangoulis et al. 2007; Lu et al. 2008; Ishikawa et al. 2010; Norton et al. 2010, 2012a; Du et al. 2013; Zhang et al. 2014). For example, QTLs have been detected on chromosomes 1 (Stangoulis et al. 2007; Lu et al. 2008), 1, 2, 7 and 12 (Ishikawa et al. 2010), 10 and 11 (Norton et al. 2010), 3, 5, 7, 8 and 9 (Norton et al. 2012a), 3, 6, 8 and 9 (Du et al. 2013) and 2, 3, 4, 6, 7, 8, 11 and 12 (Zhang et al. 2014).

Knowledge of natural genetic variation that regulates Mn concentration in grains among rice landraces and cultivars is limited. To address this, we conducted GWA mapping of grain Mn concentration in the RDP1 from four-field experiments in Arkansas and Texas, USA. The aims of this study were (1) to compare the impact of increasing marker density on detecting loci in GWA mapping for grain Mn concentration, and (2) to identify QTLs and candidate genes associated with grain Mn concentration across experiments in multiple locations and years using both single-trait and multi-experiment GWA analyses.

Materials and methods

Sample data

A total of 389 rice accessions from the RDP1 (Zhao et al. 2011; McCouch et al. 2016) consisting of 57 *aus* (AUS), 78 *indica* (IND), 100 *temperate japonica* (TEJ), 96 *tropical japonica* (TRJ) and 14 *aromatic* (ARO) as well as 44 admixtures were used in this study (Supplementary Table S1). There were two major varietal groups, *Indica* (AUS and IND) and *Japonica* (TEJ, TRJ and ARO) (McCouch et al. 2016).

The experimental design, planting methods and rice growth conditions were described in Norton et al. (2012b). Briefly, the RDP1 was grown in two locations under either flooded or unflooded cultivation. The locations were Stuttgart, Arkansas (USDA-ARS Dale Bumpers National Rice

Research Center) and Beaumont, Texas (Texas A&M Agri-Life Research Center), USA. In Arkansas, the rice cultivars were grown in the same location using nearby fields in 2006 and 2007, the field layout in both years was a randomised complete block design (RCBD) with two replications, with identical field management practices wherein fields were flooded when plants were at the five-leaf stage then drained before harvest (datasets referred to as ArFl06 and ArFl07, respectively). Seeds of each cultivar were sown with a seed drill to ~2-cm depth in a single row 5-m long with spacing of 25 cm between the plants and 50 cm between the rows. Fields were flush-irrigated twice before a permanent flood was applied to the fields ~2–3 weeks after seedling emergence. In Texas, three replications of the RDP1 accessions were grown in 2009 using two different water-treatment conditions: flooded and unflooded (datasets referred to as TxFl09 and TxUnfl09, respectively), with all other field management practices remaining the same. The experiment was set up in a RCBD. Five seeds per cultivar were drill-seeded 2-cm deep into 13-cm length lines, hereafter called hillplots. Five hillplots were planted per row with 61 cm between hillplots within each field row, and 25 cm between rows.

Accessions were represented by one hillplot per replication. The 10-cm depth flood was applied when the average plant height was ~18 cm and maintained until harvest, whereas the unflooded treatment received regular flush irrigations (once or twice a week) to keep the root zone damp but aerated.

For the Arkansas field experiments, rice grains for three plants per row for each of the two replications were collected. Seed collection was done by hand and threshed with an Almaco small-bundle thresher to obtain the seeds for the grain Mn determination. For the Texas field experiment, 20 fully mature seeds per hillplot were dehulled, from which three seeds were randomly selected for analysis of grain Mn.

The concentrations of Mn were determined in the harvested grains using inductively coupled plasma mass spectrometry (ICP-MS) described in Norton et al. (2012b, 2014) and Pinson et al. (2015). In brief, three whole grains of dehusked rice (c. 0.05 g) were digested with 1.0 ml of concentrated nitric acid and heated. The temperature was ramped up from ambient to 110 °C over a period of 12 h. An internal standard of indium (final concentration of 20 µg l⁻¹) was added to each sample. Samples were diluted to 10.0 ml and analysed on a PerkinElmer (Waltham, MA, USA) Elan DRCe ICP-MS for Mn. To control for drift, the samples were combined and used as a matrix-matched standard and measured every nine samples.

Phenotypic analysis

Phenotypic variances for Mn concentrations were calculated and parsed using two-way ANOVA conducted in R (version

3.3.0) (R Core Team 2016). Across the two field locations (Arkansas and Texas), four-field experiments were conducted, designated as ArFl06, ArFl07, TxFl09 and TxUnfl09. Across the Arkansas experiments (ArFl06 and ArFl07), the phenotypic variance was parsed into proportions estimated by genotypes, years, and interaction between genotype and year. For the Texas experiments (TxFl09 and TxUnfl09), the phenotypic variance was parsed into genotypes, water treatments and genotype-by-water-treatment interaction effects.

The average Mn concentration (Supplementary Table S1) of each accession per experiment and treatment was used for the GWA mapping. Prior to GWA mapping, the trait data were visualised to assess normality.

Genotypic data and analysis

The rice accessions in the RDP1 have three publicly available SNP datasets consisting of 36,901 (44 K) SNPs (Zhao et al. 2011), 700,000 (700 K) SNPs (McCouch et al. 2016) and 5,231,435 (5.2 M) SNPs (Wang et al. 2018). The 44 K and 700 K SNP datasets were generated by genotyping using 44 K SNP array and High-Density Rice Array (HDRA), respectively (Zhao et al. 2011; McCouch et al. 2016), whereas the 5.2 M SNP dataset, which contains no missing data, was generated by imputing from the set of the intersection of 700 K and 18 M SNPs (missing data <5% and minor allele frequency (MAF) > 1%) with 4.8 M SNPs of the 3000 Rice Genome Project (Wang et al. 2018).

The SNPs in each dataset were initially filtered using PLINK version 1.9 (Chang et al. 2015), whereby SNPs were removed when the percentage of missing genotype data for a single SNP exceeded 20% (the 5.2 M SNP dataset had no missing data, due to being imputed) and MAF was less than 5%.

GWA mapping with single-trait analysis

GWA mapping was performed using the three SNP datasets based on LMMs from EMMAX (version beta-7Mar2010) (Kang et al. 2010) using the PIQUE (Parallel Identification of QTLs using EMMAX) pipeline (<https://github.com/tony-travis/PIQUE>). Phenotype–genotype association was analysed for all accessions (ALL) and four subpopulations (AUS, IND, TEJ and TRJ) in the four-field experiments. Due to low accession numbers (<30) from the *aromatic* (ARO) subpopulation and the mixed genetic background of admixtures, these accessions were not analysed as separate subpopulations. Population structure was estimated by performing a principal component analysis on the informative SNP data and the eigenvectors for the first four principal components were included in the model as fixed effects for the analysis of the whole (ALL) population (Price et al. 2010) (note: population structure was not

included in the analysis of the subpopulations). Relatedness (K matrix) between accessions was estimated by calculating pairwise identity-by-state (IBS) using the SNP data and was included in the models as random effects. For the ALL population GWA analyses, relatedness was estimated using the accessions that had phenotype data in each experiment. For the subpopulation analyses, relatedness was estimated using accessions from each subpopulation for which phenotype data were collected. The significance threshold for the association between SNP and traits was set at P value < 0.0001 , a value previously used for this population (Famoso et al. 2011; Norton et al. 2014). To further filter these SNPs for false-discovery rates (FDR), the P values calculated by the GWA mapping analysis were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg 1995). To be reported as a SNP significantly associated with the trait, the SNP had to both meet the P value < 0.0001 and meet the criteria of a 5% FDR. Manhattan plots were used to visualise SNP positions on chromosomes with $-\log_{10}(P)$ and Q-Q plots were used to visualise observed versus expected value probabilities using the qqman package in R version 3.3.0 (Turner 2014).

GWA mapping with multi-experiment analysis

Multi-experiment analysis of GWA mapping for grain Mn concentration for the three flooded-field experiments (ArFI06, ArFI07 and TxFI09) was performed. For this analysis, each environment was viewed as one trait. A total of 303 rice accessions (all accessions common among the three experiments) were used for the analysis with the 3,430,260 filtered SNPs (MAF $> 0.05\%$) using the mvLMM in the GEMMA version 0.97 (Zhou and Stephens 2014). The mvLMM accounts for both population stratification and relatedness among samples to control confounding factors. The eigenvectors of the first four principal components were calculated using the smartpca programme in EIGENSOFT (Patterson et al. 2006) and included in the model as fixed effects. One eigen decomposition of the centred relatedness matrix (the n -by- n relatedness matrix; n = the number of samples) for random effects was computed from all filtered SNPs using the relatedness matrix function in GEMMA. The null hypothesis is that SNP effects of a single SNP in all experiments are zero, whereas the alternative hypothesis is nonzero effects of at least one SNP tested by a Wald test. P values of all association tests were presented with Manhattan plots and observed P values against expected P values were presented by Q-Q plots using the qqman package in R (Turner 2014). The guideline of reliability for significant SNPs was 0.0001 (Famoso et al. 2011; Norton et al. 2014). SNPs were also tested to a 5% FDR based on the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995), as previously described.

Clustering significant SNPs and comparing QTLs on rice chromosomes

The grouping function CLUMP was used in PLINK to define candidate regions in the ALL analysis based on the 5.2 M SNP. Index SNPs were identified with P value < 0.0001 (Norton et al. 2014) and neighbouring SNPs were clumped with P value < 0.01 (default value) and squared allele frequency correlation (r^2) > 0.5 (applying the criteria from Butardo et al. (2017) based on the 700 K SNP dataset) with the index SNPs of each peak within 500 kb, which was the LD-decay average of all accessions in the RDP1 (Zhao et al. 2011). The candidate regions/QTLs were then mapped and compared with previously reported QTLs based on physical genome positions on the 12 rice chromosomes.

Local linkage disequilibrium decay analysis

To determine LD blocks in subpopulations that supported the significant peaks in the ALL analysis, a subset of the 5.2 M SNP data surrounding (1 Mbp) a significant peak was extracted using PLINK. Two methods were used: (1) local LD decay was estimated at $r^2 = 0.2$, where r^2 values were calculated using PLINK and estimated by binning the average r^2 values of 10-kb windows (Biscarini et al. 2016; Norton et al. 2018); (2) r^2 values in each SNP pair in each region 500 kb upstream and downstream were calculated and visualised as a local Manhattan plot against a LD heatmap using the LD.heatmap package in R version 3.3.0 (Shin et al. 2006), and then LD blocks were estimated using $r^2 \geq 0.6$ (high LD) (Ripke et al. 2014; Yano et al. 2016).

Candidate gene identification

Within each candidate region, positional genes were identified based on genes identified in the Rice Genome Annotation Project (version 7, <http://rice.plantbiology.msu.edu>). Retrotransposons and transposon genes were excluded. Genes located within candidate regions were examined and used to identify potential positional functional candidate genes, e.g., genes involved in the uptake, transport and accumulation of elements, associated with Mn. In addition, protein sequences (<http://rice.plantbiology.msu.edu>) of the list of candidate genes that were not matched with genes previously related to Mn were compared with protein database using BLASTp (<https://blast.ncbi.nlm.nih.gov>) to investigate gene-sequence homology with other species, in which genes were reported and characterised with functions involving Mn. In addition to gene validation, the gene expression profiles across a range of rice organs and tissues of all identified candidate genes obtained from RiceXPro (<http://ricexpro.dna.affrc.go.jp>, Sakai et al. 2013) were used to confirm the validity of candidate genes.

Differential gene expression of candidate genes was determined based on the gene expression analysis conducted by Campbell et al. (2020). This dataset is transcriptomic data from shoots of young plants from 91 accessions from the RDP1. Initially, the data were screened to identify which of the proposed candidate genes were expressed. Low- and high-grain Mn accessions were identified based on being in the highest 20% and lowest 20% for grain Mn concentration, for the three flooded experiments. Then only low- or high-grain Mn concentration accessions were selected for further analysis if they were low or high in at least two of the experiments. A total of 14 accessions were identified as high and 18 identified as low-grain Mn accessions for which transcript data were available (Supplementary Table S2). The expression of candidate genes was examined for evidence of differential expression based on this grouping. An ANOVA was used to determine if the gene expression was different between the two groups.

Estimation of phenotypic variance explained by significant SNPs

To determine the effect size of the QTLs, two approaches were taken. Either the smallest *P*-value/index SNPs or the most significant SNP located in candidate genes based on the 5.2 M SNP dataset was analysed. The proportion of pheno-

typic variance explained by each SNP was estimated using linear models, correcting for population structure and contrasting with the population-structure effects for all accessions (Zhao et al. 2011). ANOVA was used to contrast the linear models. For subpopulations, the phenotypic variance distribution of a significant SNP was estimated using a simple linear model without correcting for population structure.

Effect sizes by index SNPs in multi-experiment analysis

Effect sizes in each index SNP of QTLs newly identified in multi-experiment analysis were observed in the individual experiments estimated by the mvLMM model.

Results

Variation of grain Mn concentration in the RDP1

In Arkansas, grain Mn concentration for the accessions in 2006 and 2007 ranged from 21.4 to 62.7 mg kg⁻¹ and from 20.6 to 68.5 mg kg⁻¹, with means of 34.6 and 40.8 mg kg⁻¹, respectively (Fig. 1a and Table 1). There were significant differences ($P < 2 \times 10^{-16}$, *df* = 321) in grain Mn concentration among genotypes, years and a significant interaction

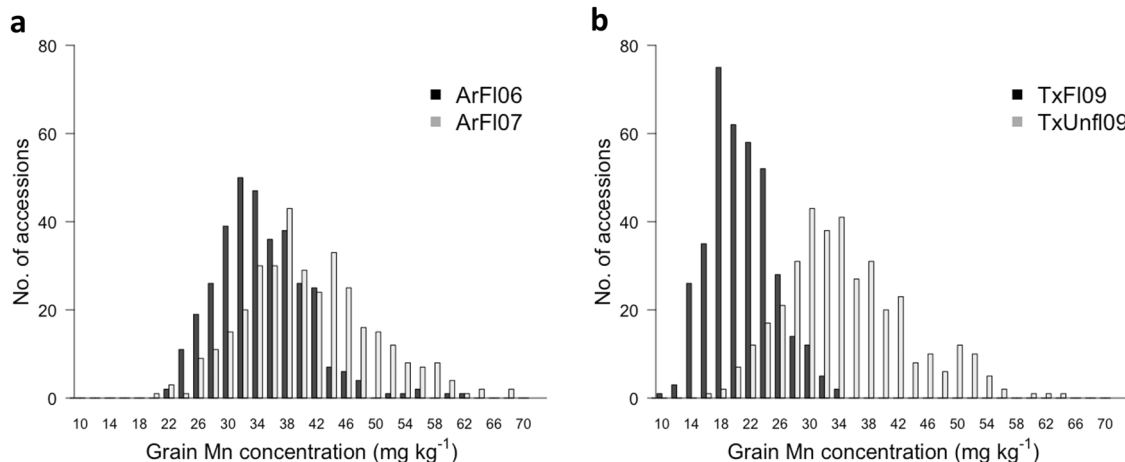


Fig. 1 Grain Mn distributions in all rice accessions. **a** Distribution of grain Mn concentration in Arkansas under flooded condition in 2006 and 2007. **b** Distribution of grain Mn concentration in Texas under flooded and unflooded conditions in 2009.

Table 1 Grain Mn concentration (mg kg⁻¹) in the RDP1 accessions at each field experiment.

Site	Year	Condition	No. of accessions	Range	Mean	SD	CV (%)
Arkansas	2006	Flooded	342	21.4–62.7	34.6	6.2	17.8
	2007	Flooded	349	20.6–68.5	40.8	8.6	21.1
Texas	2009	Flooded	373	10.6–33.5	20.9	4.3	20.7
	2009	Unflooded	370	16.4–63.8	34.8	8.4	24.2

SD standard deviation, *CV* coefficient of variation

between years and genotypes that explained 39, 15 and 16% of the phenotypic variance, respectively. In Texas, grain Mn concentration for the accessions in 2009 under flooded and unflooded conditions ranges from 10.6 to 33.5 mg kg⁻¹ and from 16.4 to 63.8 mg kg⁻¹, with means of 20.9 and 34.8 mg kg⁻¹, respectively (Fig. 1b and Table 1). The grain Mn concentration under the unflooded condition was significantly higher (1.7 times, $P < 2 \times 10^{-16}$, $df = 367$) than the concentration under the flooded condition. The rice grain Mn concentrations were affected by genotypes, water treatments and their interaction, which explained 14%, 61% and 11% of the phenotypic variance, respectively.

To compare the grain Mn accumulation among subpopulations, only those subpopulations with at least 30 accessions (AUS, IND, TEJ and TRJ) were studied. There was a significant difference in grain Mn concentration among the subpopulations (Fig. 2). In the three flooded-field experiments, the *Japonica* (TEJ and TRJ) subgroups had higher average grain Mn concentration than the *Indica* (AUS and IND) subgroups. In contrast, the TRJ subpopulation had the lowest average grain Mn concentration in TxUnf109.

The accessions screen at these field sites are known to vary in the length of time to heading (Norton et al. 2012b); therefore, a correlation analysis was conducted to determine if there was a relationship between heading date and grain manganese concentration. For Arkansas 2007 and flooded experiment in Texas, there was no correlation between flowering time and grain Mn concentration. However, at the Arkansas 2006 experiment, there was a significant weak positive correlation ($r = 0.235$, $P < 0.001$) between grain manganese and flowering time, while at the Texas unflooded field site, there was a significant weak negative correlation ($r = -0.278$,

$P < 0.001$) between grain manganese concentration and flowering time.

Density of SNPs among all accessions and subpopulations

To obtain high-quality SNPs in each SNP dataset, SNPs were filtered with genotype missing >20% and MAF < 0.05 (Supplementary Table S3). After SNP filtering, for example, average SNP density of 11.40, 0.99 and 0.11 kb per SNP was observed for the 44 K, 700 K and 5.2 M SNP datasets, respectively, for the ArFI06 dataset.

For subpopulations, it is noteworthy that the final number of filtered SNPs was lower in the TEJ and TRJ subpopulations compared to the AUS and IND subpopulations (Supplementary Table S3). For example, the SNP density in the TEJ subpopulation was 1 SNP per 0.41 kb, whereas the SNP density in the IND subpopulation was 1 SNP per 0.17 kb, when using the 5.2 M SNP dataset.

Single-trait GWA mapping for grain Mn concentration

Using the three SNP datasets, GWA mapping for grain Mn concentration was performed for all accessions (Fig. 3a and Supplementary Figs. S1–S3) and for the four subpopulations using the 5.2 M SNP dataset only (Supplementary Figs. S4–S7) in the four-field experiments.

Increasing the SNP density increased the number of significant SNPs associated with the trait in analyses of all accessions and in subpopulation analysis (Fig. 3a, Supplementary Figs. S1–S7 and Supplementary Table S4). For example, no significant SNPs for grain Mn in the ALL analysis in ArFI06 were identified using the 44 K dataset,

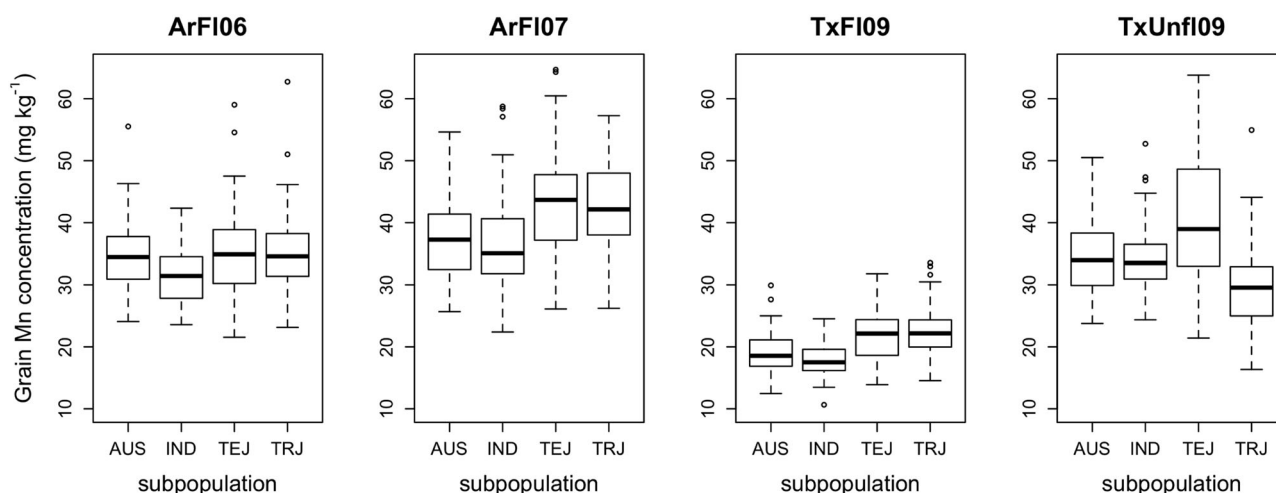
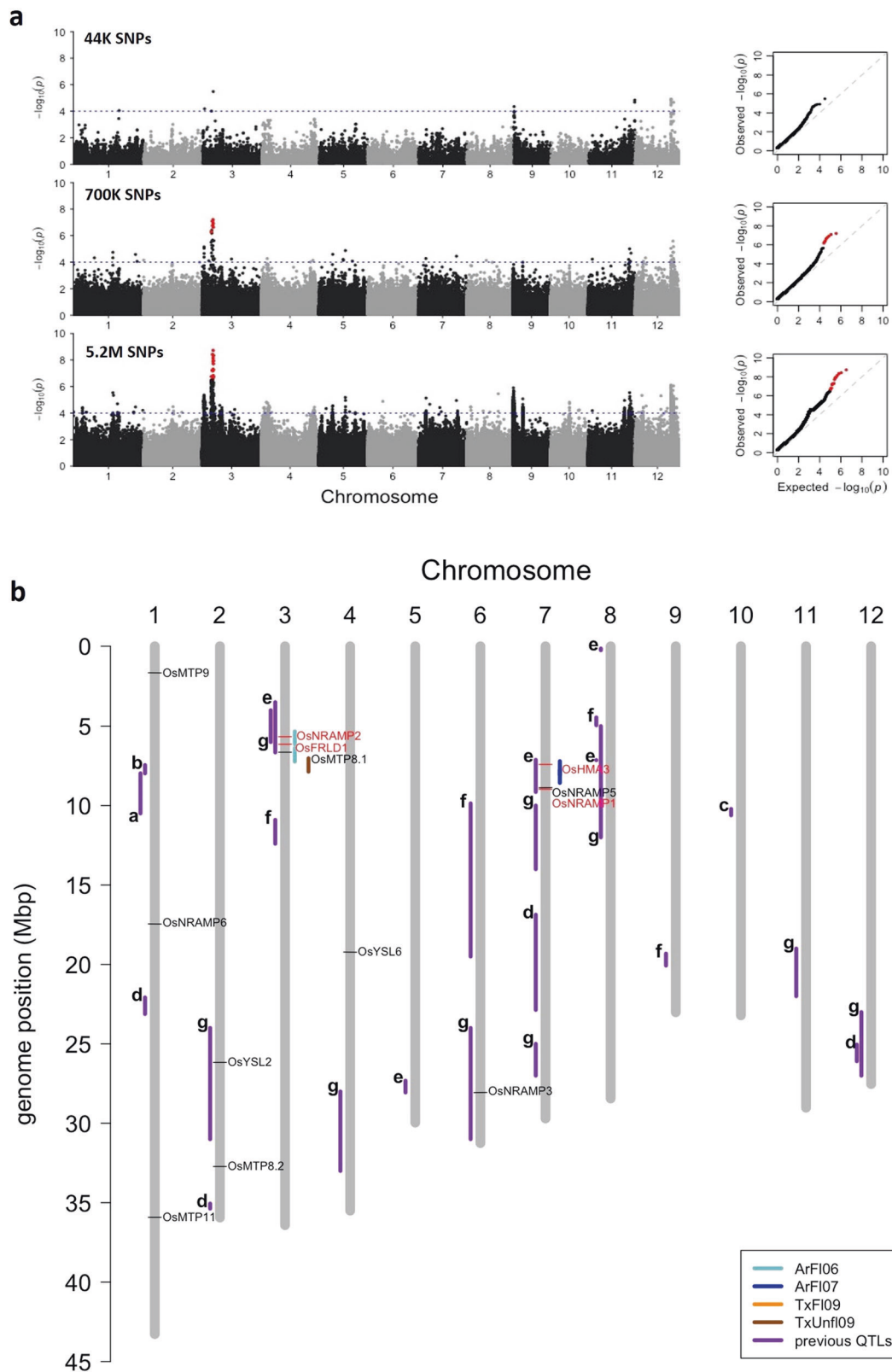


Fig. 2 Distribution of grain Mn concentration in rice in four subpopulations in four-field experiments. The horizontal black bar is the median of grain Mn concentration.



while 6 and 16 SNPs were significant using the 700 K and 5.2 M SNP datasets, respectively (Supplementary Table S4). In addition, there were no significant SNPs associated

with Mn accumulation in several subpopulations based on the 44 K SNP dataset, whereas a number of significant SNPs were identified based on the 700 K and 5.2 M SNP

◀ **Fig. 3 Genome-wide association-mapping results for grain Mn concentration in rice using single-trait analysis in all accessions grown in Arkansas under flooded condition in 2006.** **a** Manhattan (left) and Q-Q (right) plots are presented for the 44 K, 700 K and 5.2 M SNP datasets. The blue horizontal line represents the $-\log_{10}(P)$ threshold at 4. The red dot indicates SNP loci that passed 5% FDR. **b** Location of QTLs associated with grain Mn concentration in rice for four-field experiments based on the 5.2 M SNP dataset using single-trait analysis and previously reported QTLs. The four-field experiments are ArFI06—light blue, ArFI07—blue, TxFI09—orange and TxUnfl09—brown. Previous reported QTLs are displayed in purple with the letter indicating the study they were detected in; a = Stangoulis et al. (2007), b = Lu et al. (2008), c = Norton et al. (2010), d = Ishikawa et al. (2010), e = Norton et al. (2012a), f = Du et al. (2013) and g = Zhang et al. (2014). Known Mn-transporter locations are indicated by horizontal black lines, whereas the locations of candidate genes are indicated by horizontal red lines.

datasets. For example, in the TEJ subpopulation in ArFI07, no significant SNPs were detected using the 44 K SNP dataset, while 6 and 11 significant SNPs for grain Mn were detected using the 700 K and 5.2 M SNP datasets, respectively (Supplementary Table S4).

Identification of grain Mn QTLs and candidate genes based on single-trait analysis

Based on the high-density SNP dataset (5.2 M SNPs), a number of candidate regions/QTLs in the four experiments (Supplementary Table S5) were mapped on rice chromosomes and compared with previously reported QTLs (Fig. 3b). QTLs were further focused on when SNPs within QTLs passed the 5% FDR (Table 2). Consequently, there were three QTLs on chromosome 3 and two QTLs on chromosome 7 that were significantly associated with grain Mn concentration under flooded and unflooded conditions that met the criteria (Table 2). Based on overlap regions from the CLUMP analysis in experiments, these QTLs on chromosome 3 were at 5.33–6.14, 6.39–7.23 and 7.02–7.87 Mbp. For two of these QTL regions, there are a number of candidate genes, including LOC_Os03g11010 (*OsNRAMP2*), LOC_Os03g11734 (*OsFRDL1*) and LOC_Os03g12530 (*OsMTP8.1*). On chromosome 7, the two overlapping QTL regions were at 7.21–8.06 and 7.78–8.57 Mbp. For the first of the two QTL regions, there was a good candidate gene: LOC_Os07g12900 (*OsHMA3*) but this gene is outside the candidate region for the second. The expression profiles of all candidate genes under normal growth conditions were obtained from the RiceXPro database (Supplementary Figs. S8–S11).

All four candidate genes mentioned above were identified as being expressed in shoots (Campbell et al. 2020). Of these four genes, two of the candidate genes (*OsMTP8.1* and *OsHMA3*) were found to be differentially expressed between the low-grain Mn and high-grain Mn accessions (Supplementary Fig. S12a, b). The expression of

Table 2 Significant QTLs on chromosomes 3 and 7, and candidate genes for grain Mn concentration in all rice accessions under flooded and unflooded conditions based on $P < 0.0001$ and passing 5% FDR using the 5.2 M SNP dataset.

Chr.	Loc	Experiment	Index SNP id	Position	Minor/major allele	MAF	P value	Estimated candidate region	Clumped size (kb)	MSU id	Gene name
3	1	ArFI06	mlid0018058877	5,804,440	T/C	0.11	1.77E-07	5,326,994–6,138,356	811.36	LOC_Os03g11010	<i>OsNRAMP2</i>
	2	ArFI06	mlid0018228403	6,819,735	T/C	0.13	1.87E-09	6,387,194–7,233,912	846.72	LOC_Os03g11734	<i>OsFRDL1</i>
	3	TxUnfl09	mlid0018352659	7,507,792	T/A	0.12	1.22E-08	7,019,729–7,866,465	846.74	LOC_Os03g12530	<i>OsMTP8.1</i>
7	1	ArFI07	mlid0048560313	7,594,101	A/G	0.14	1.80E-08	7,206,339–8,057,921	851.58	LOC_Os07g12900	<i>OsHMA3</i>
	2	ArFI07	mlid0048735332	8,256,487	A/C	0.15	1.44E-08	7,780,036–8,568,541	788.51	–	–

All QTLs are concordant with previously reported QTLs of which references are given in Fig. 3.

LOC_Os03g12530 (*OsMTP8.1*) was higher in the accessions with low-grain Mn compared to the accessions with high-grain Mn, while the expression of LOC_Os07g12900 (*OsHMA3*) was higher in the accessions identified as having high-grain Mn compared to the low-grain Mn accessions.

Identification of grain Mn QTLs in subpopulations and candidate genes based on single-trait analysis

Due to the complex population structure in the RDP1, the estimation of linkage disequilibrium (LD) decay for single QTL across the whole panel is difficult. Therefore, to estimate the size of QTL regions based on LD, QTL analysis was conducted in the individual subpopulations (AUS, IND, TEJ and TRJ). In QTLs that were detected for the whole population and one of the subpopulations, the subpopulation analysis was used to estimate local LD.

For grain Mn QTLs in subpopulations, one significant QTL on chromosome 7 was identified in only the TEJ subpopulation based on the 5.2 M SNP dataset (Fig. 4a, Supplementary Figs. S4–S7 and Supplementary Table S6) that was concordant with significant SNPs on chromosome 7 in all analyses at the 5% FDR. To determine the accurate genomic position of the QTLs, local LD was analysed with two approaches, LD decay and LD heatmap. The QTL was identified at ~8.26 Mbp in the TEJ subpopulation (Fig. 4a). The average local LD decay between 7 and 9 Mbp on chromosome 7 was high at >1 Mbp ($r^2 > 0.2$) (Fig. 4b). The result was concordant with LD heatmap that showed a large LD block at approximately 1.23 Mbp from 7.64 to 8.87 Mbp ($r^2 \approx 0.6$) (Fig. 4c). One candidate gene, *OsNRAMP5* (~8.87 Mbp), was found to be located within the QTL. *OsHMA3* at 7.40 Mbp, which was identified as a candidate gene for the QTL detected here in the ALL analysis, is just before this block, while *OsNRAMP1*, which is at 8.97 Mbp, is just after it (Fig. 4c). In this QTL, the significant SNP mldid0048878287 (8.78 Mbp, $P = 8.11E^{-07}$), which located close to *OsNRAMP5* and *OsNRAMP1*, explained approximately 8 and 29% of phenotypic variance in ALL and TEJ, respectively. Rice accessions with the TT genotype at this SNP had high Mn accumulation in grains compared to the rice accessions with the CT and CC genotypes (Fig. 4d). The expression profile of *OsNRAMP5* and *OsNRAMP1* under normal growth conditions was obtained from the RiceXPro database (Supplementary Figs. S13, S14). Both *OsNRAMP5* and *OsNRAMP1* were identified as being expressed in shoots of rice plants (Campbell et al. 2020). Of these two genes, *OsNRAMP1* was found to be differentially expressed between the low-grain Mn and high-grain Mn accessions (Supplementary Fig. S12c). The expression of LOC_Os07g15460 (*OsNRAMP1*) was higher in the accessions identified as having grain Mn compared to the low-grain Mn accessions.

Multi-experiment GWA mapping for grain Mn concentration and candidate genes

To increase the power of GWA mapping, a single GWA mapping was conducted for grain Mn concentration of 303 accessions for the three flooded-field experiments (ArFI06, ArFI07 and TxFI09) based on the 5.2 M SNP dataset (MAF > 0.05, 3,430,260 filtered SNPs) using the mvLMM in the GEMMA software. A total of 64 SNPs were significantly associated with grain Mn concentration. Eight QTLs across the 12 rice chromosomes were identified (Fig. 5a and Supplementary Table S7). Two of these QTLs on chromosome 3, 5.97–6.95 and 6.63–7.51 Mbp, including *OsFRDL1* and *OsMTP8.1* (Figs. 3b, 5b and Supplementary Table S7), were consistent with the QTLs identified based on single-trait analysis. However, a total of 6 QTLs for grain Mn not detected by single-trait analysis were identified using multi-experiment analysis (Fig. 5 and Supplementary Table S7). The six QTLs of interest were at 1.16–1.38 Mbp on chromosome 3, 2.40–3.33 and 3.41–4.27 Mbp on chromosome 4, 0.39–1.00 Mbp on chromosome 9 and 11.39–12.30 and 25.61–25.62 Mbp on chromosome 11. All of these QTLs were novel for grain Mn concentration.

Comparison of the effect sizes of index SNPs for the putative QTLs in each experiment estimated by the mvLMM showed that they were various (Table 3). For example, the QTL on chromosome 3 had similar small positive SNP effects in all experiments, whereas the two QTLs on chromosome 4 had negative SNP effects in ArFI07 compared to other experiments.

Discussion

This study has identified QTLs for grain Mn in rice. Some of those co-localise with previously identified QTLs and known genes involved in Mn accumulation in rice, while some are novel putative QTLs. One of the key objectives of QTL mapping is the identification of stable QTLs (e.g., those are detected in multiple environments). Using a multi-experiment GWA mapping approach, we have been able to identify these stable QTLs.

The environmental factors (different years, locations and water-management treatments) and genetic composition of the accessions affected the concentration of Mn in rice grains. In Arkansas, the average grain Mn concentrations between 2006 and 2007 were significantly different and year explained ~15% of the phenotypic variance. In Texas, the Mn concentration in grains under non-flooded condition significantly increased when compared to the rice cultivation under flooded condition with flooding explaining ~61% of the variation. This is in agreement with Pinson et al. (2015) who reported that water-treatment effects had a

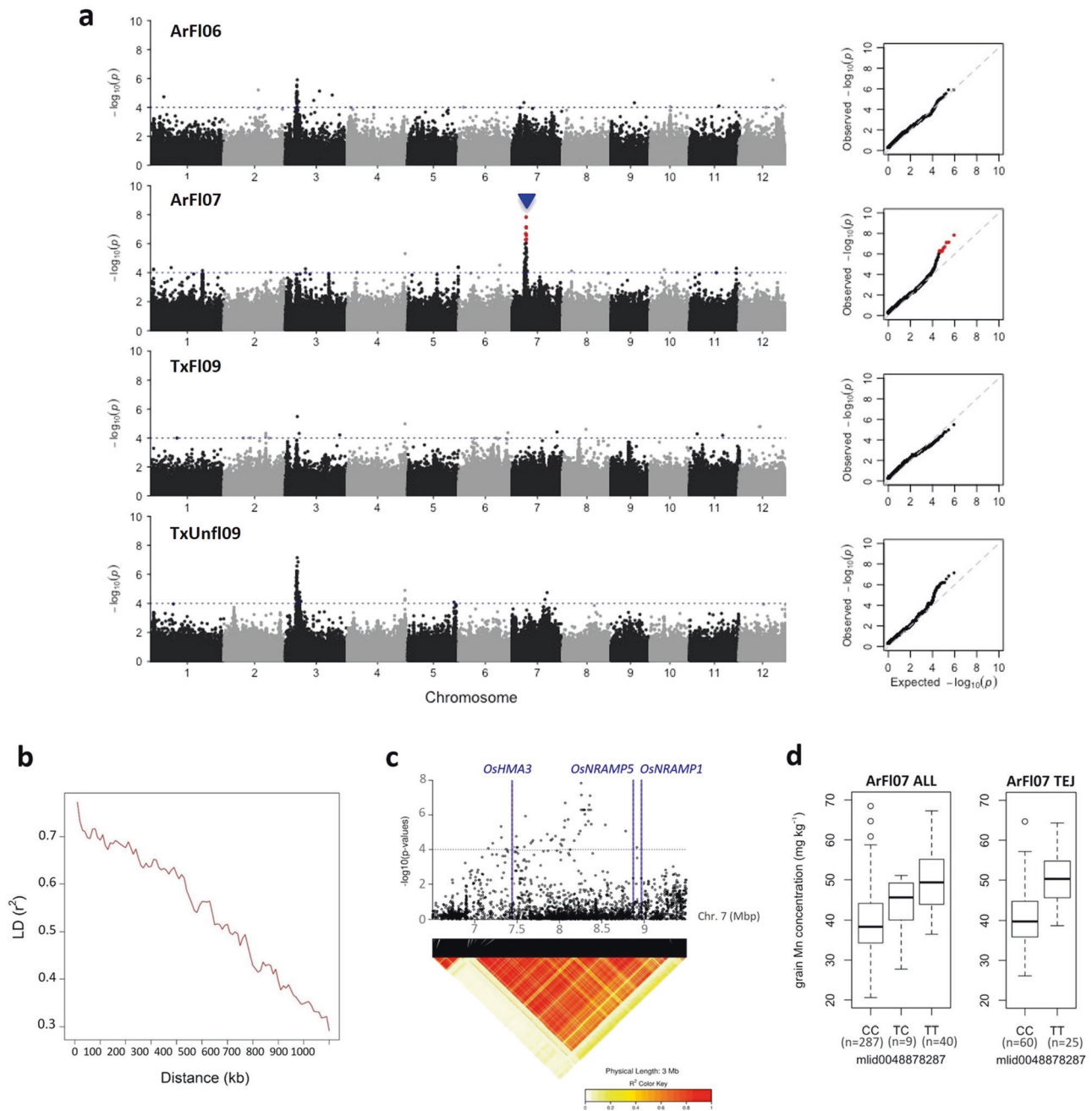


Fig. 4 Genome-wide association-mapping results for grain Mn concentration in the *temperate japonica* subpopulation based on the 5.2 M SNP dataset, as well as local linkage disequilibrium analysis and SNP allele effects. **a** Manhattan (left) and Q-Q (right) plots are presented in four-field experiments as ArFI06: Arkansas flooded 2006, ArFI07: Arkansas flooded 2007, TxFI09: Texas flooded 2009 and TxUnfl09: Texas unflooded 2009. The blue horizontal line represents the $-\log_{10}(P)$ threshold at 4. The red dot indicates SNP loci

that passed 5% FDR. **b** LD decay and **c** local Manhattan plot (top), as well as LD heatmap (bottom) for grain Mn concentration in ArFI07 on chromosome 7 at 7–9 Mbp and 6.5–9.5 Mbp, respectively. **d** The effect of SNP alleles on grain Mn concentration for the QTL on chromosome 7 with the SNP mli0048878287 (8,781,883 bp) in all accessions (left) and the *temperate japonica* subpopulation (right) in ArFI07.

higher impact for element accumulation in rice grains than year effects, and the average grain Mn concentration under unflooded condition was greater than the average grain Mn concentration under flooded condition among 1763 rice accessions grown in Texas in 2007 and 2008. Senewiratne

and Mikkelsen (1961) found that Mn concentration in rice leaves was 7.7-fold higher under unflooded condition as compared with flooded condition. One genetic factor that could have an influence on grain element concentrations is flowering time. As this population comprised a wide range

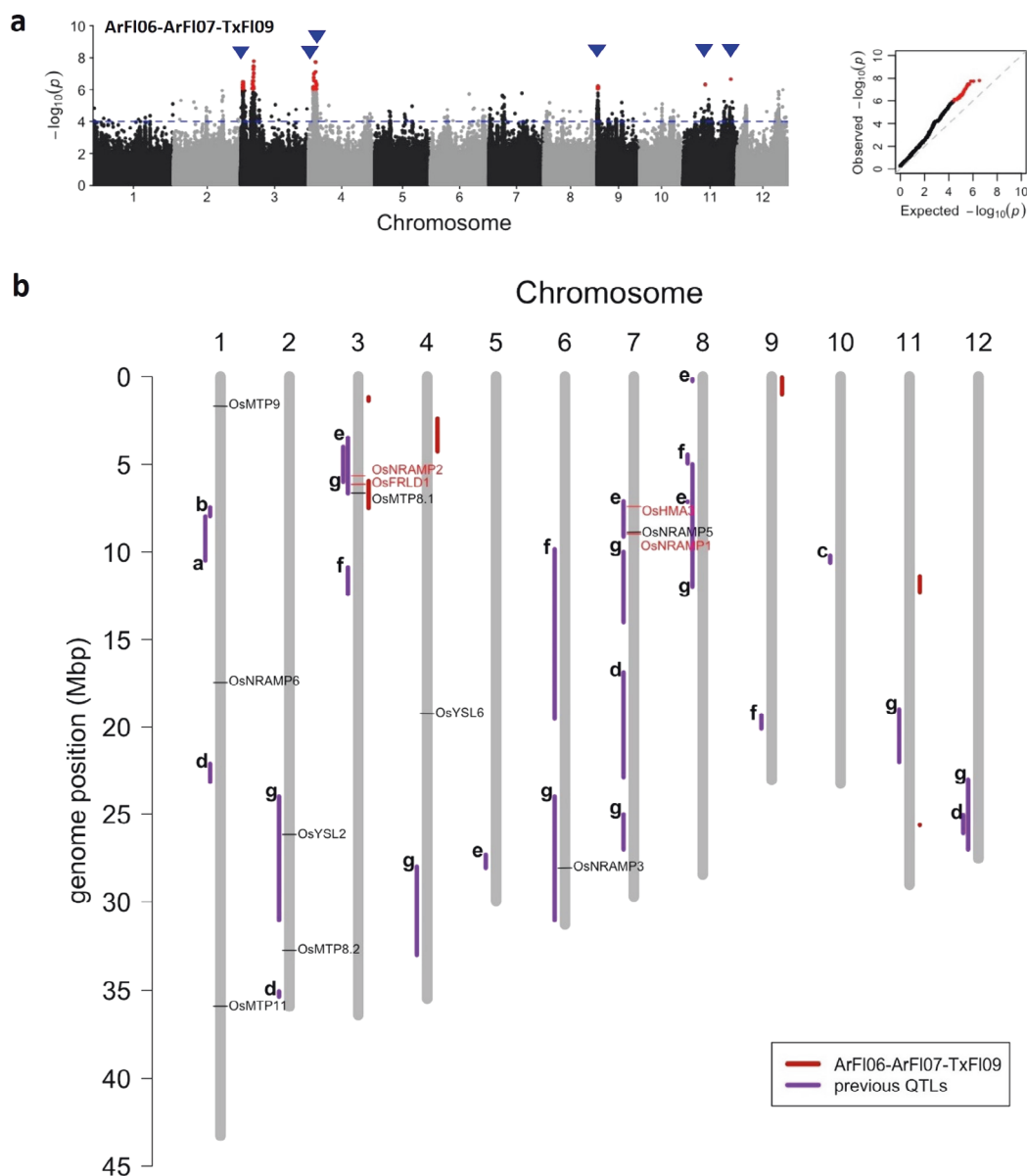


Fig. 5 Genome-wide association mapping results for grain Mn concentration in rice based on the 5.2 M SNP dataset using multi-experiment analysis. **a** Manhattan (left) and Q-Q (right) plots are presented. The blue horizontal line represents the $-\log_{10}(P)$ threshold at 4. The red dot indicates SNP loci that passed 5% FDR. The blue arrows point to new QTLs based on multi-experiment analysis. **b** Location of QTLs associated with grain Mn concentration in rice based on the 5.2 M SNP dataset using multi-experiment analysis and

previously reported QTLs. QTLs in ArFI06–ArFI07–TxFI09 are presented in red. Previous reported QTLs are displayed in purple with the letter indicating the study they were detected in; a = Stangoulis et al. (2007), b = Lu et al. (2008), c = Norton et al. (2010), d = Ishikawa et al. (2010), e = Norton et al. (2012a), f = Du et al. (2013) and g = Zhang et al. (2014). Known Mn-transporter locations are indicated by horizontal black lines, whereas the locations of candidate genes are indicated by horizontal red lines.

of different accessions, the flowering window (the time from the first accession flowering to the last) is quite large (Norton et al. 2012b). During this time, the environmental conditions can change, which may affect the availability and therefore the accumulation of manganese. However, in this study, only at two sites were relationships between flowering time and grain manganese concentration observed, and in both cases, the relationships explained only a small component of the variation.

The genetic differences among subpopulations also affected grain Mn concentrations such as higher grain Mn concentration in the TEJ and TRJ subpopulations grown under flooded conditions compared with the AUS and IND subpopulations (Fig. 2). In another study under flooded conditions, *Japonica* subgroup accessions had higher Mn concentrations in their rice grains than *Indica* subgroup accessions (Yang et al. 2018). Pinson et al. (2015) have shown that although water-management treatments had a

Table 3 New putative QTLs for grain Mn concentration based on the 5.2 M SNP dataset using multi-experiment analysis.

Chr.	Index	SNP id	Position	Minor/major allele	MAF	P value	Allele effect			Estimated candidate region	Clumped size (kb)
							ArFI06	ArFI07	TxFI09		
3		mlid0017254736	1,253,009	T/C	0.24	3.43E-07	0.0066	0.0180	0.0075	1,164,504–1,377,789	213.29
4		mlid0025444660	2,891,380	T/C	0.17	1.03E-07	0.0043	-0.0262	0.0085	2,403,945–3,330,179	926.23
4		mlid0025744000	3,767,588	C/G	0.14	1.89E-08	0.0065	-0.0014	0.0105	3,411,664–4,267,559	855.90
9		mlid0060879641	504,844	A/G	0.40	6.17E-07	-0.0113	-0.0397	-0.0135	38,664–1,004,262	965.60
11		mlid0074604290	11,886,926	G/A	0.39	4.80E-07	0.0062	-0.0320	-0.0095	11,391,970–12,295,345	903.38
11		mlid0078299980	25,621,708	A/C	0.05	2.24E-07	0.0094	0.0340	-0.0026	25,619,048–25,621,708	2.76

high impact, genetic backgrounds in the 1763 rice accessions were a major factor for grain element accumulation in both flooded (average broad-sense heritability (H^2) of 16 elements: 0.49, Mn: 0.58) and unflooded (average H^2 : 0.57, Mn: 0.70) conditions.

The efficiency of GWA mapping depends on several factors such as the proportion of variation explained by genotype (heritability), the underlying population structure within the panel, sample size and marker density. McCouch et al. (2016) suggested that increasing SNP density increases the ability to detect genetic loci. In this study, we tested the impact of marker density while using the same rice accessions, the same phenotype data and the same statistical modelling approach to account for population structure and kinship (Kang et al. 2010). We demonstrated that the number of markers covering the genome affects the efficacy of GWA mapping (Fig. 3a and Supplementary Figs. S1–S3). Our results revealed that higher marker density increases the number of significant loci associated with the trait. A similar observation has been made in a recent study, where GWA mapping for root cone angle in rice was conducted using 15,000 and 300,000 SNPs (Bettembourg et al. 2017). Wang et al. (2018) also showed that increasing from 700 K to 4.8 M SNPs in GWA mapping for the grain amylose content in 326 *indica* accessions provided increased confidence in QTLs, as well as revealing new ones. In the present study, increasing marker density improved the identification of genetic loci using GWA mapping. However, at this stage, it is unknown what the optimal marker density for GWA mapping in rice is.

From single-trait analyses using 5.2 M SNPs, five QTLs on chromosomes 3 and 7 were found to affect grain Mn concentration, and the QTL sizes ranged from 789 to 852 kb (Table 2). Some of these QTLs co-localise with QTLs previously identified in rice (Fig. 3b). For example, Norton et al. (2012a) detected grain Mn QTLs using the Bala × Azucena mapping population, located similarly with QTLs detected in this study (chromosome 3 at approximately 3.49–6.65 Mbp and chromosome 7 at ~7.12–9.14 Mbp). Zhang et al. (2014) detected QTLs for grain Mn under

flooded growing conditions in a TeQing-into-Lemont backcross introgression population on chromosome 3 (4–6 Mbp) and chromosome 7 (10–14 Mbp), with the chromosome 7 QTL being identified also in an independent population of Lemont × TeQing recombinant inbred lines. To further narrow down a QTL on chromosome 7, Liu et al. (2017) characterised a major QTL for grain Mn accumulation in recombinant inbred lines from the cross of 93-11 (low-grain Mn) with PA64s (high-grain Mn) grown in two environments. A major QTL located on the short arm of chromosome 7 was fine-mapped between two markers (L8857 and L8906), a 49.3-kb region encompassing the known Mn transporter, *OsNRAMP5* (Liu et al. 2017). Recently, Shrestha et al. (2018) conducted GWA mapping for shoot Mn toxicity in 271 RDPI accessions based on the 700 K SNP dataset. Numerous significant SNPs were identified in a large region on the top of chromosome 7. Although they did not report an exact QTL size, both *OsNRAMP5* and *OsNRAMP1* were identified as candidate genes. These results reveal that our identified QTLs based on GWA mapping with the high SNP density were smaller than the comparable genomic regions when using other mapping approaches. As a result, the identified QTLs contained a smaller number of positional candidates, which means the identification of genes underpinning the QTLs should be easier.

Within the grain Mn QTLs on chromosomes 3 and 7, six genes are proposed as contributing to the natural variation observed in grain Mn concentration in the RDPI. The candidate genes were highly expressed in roots, shoots, reproductive organs or embryo and endosperm tissues (Supplementary Figs. S8–S11 and S13, S14). On chromosome 3, three candidate genes were identified as *OsNRAMP2* (LOC_Os03g11010), *OsFRDL1* (LOC_Os03g11734) and *OsMTP8.1* (LOC_Os03g12530). While the function of *OsNRAMP2* is unknown in rice, *OsNRAMP2* has high structural similarity with an Mn transporter from *Eremococcus coleocola* (Mani and Sankaranarayanan 2018). *OsNRAMP2* in rice is also an orthologous gene with *AtNRAMP2* in *Arabidopsis* (Thomine et al. 2000) that is a trans-Golgi

network-localised Mn transporter in roots under Mn deficiency (Gao et al. 2018). *OsFRDL1* is a good candidate gene as a knockout of *OsFRDL1* in rice resulted in lower leaf Fe concentration, and higher accumulation of Zn and Mn in leaves of rice (Yokosho et al. 2009). *OsMTP8.1* has been shown to be involved in Mn homeostasis achieved by sequestering excess Mn into vacuoles of rice (Chen et al. 2013, 2016), and to be an orthologous gene with *AtMTP8* involving in the localisation of Mn and Fe in *Arabidopsis* seeds (Chu et al. 2017). On chromosome 7, there were three candidate genes: *OsHMA3* (LOC_Os07g12900), *OsNRAMP5* (LOC_Os07g15370) and *OsNRAMP1* (LOC_Os07g15460). *OsHMA3* is a known tonoplast-localised transporter for Zn and Cd in rice roots, but it is reported that the overexpression of *OsHMA3* affected Mn concentration in roots and shoots (Sasaki et al. 2014). *OsNRAMP5* is a major transporter for Mn as well as for Fe and Cd in rice (Ishimaru et al. 2012; Sasaki et al. 2012; Yang et al. 2014; Liu et al. 2017). Although *OsNRAMP1* is an Fe transporter that is involved in Cd accumulations in rice (Takahashi et al. 2011), a phylogenetic analysis of *NRAMP* sequences in plants showed that *OsNRAMP1* was most similar to *OsNRAMP5* (Vatansever et al. 2016). Sheartha et al. (2018) also identified a QTL for Mn toxicity in rice using GWA mapping that encompassed both *OsNRAMP1* and *OsNRAMP5*. Therefore, *OsNRAMP1* is possibly involved in Mn transport or cross-talk between Fe and Mn homeostasis (Vatansever et al. 2016).

Due to genetic similarity within the TEJ subpopulation, local LD for the identified QTL on chromosome 7 was analysed and estimated to define their candidate regions. The average LD decay from 7 to 9 Mbp in the TEJ subpopulation was high at >1 Mbp with the threshold of $r^2 = 0.2$ (Fig. 4b). To determine if this large LD decay was specific to the TEJ subpopulation, the LD decay in the other subpopulations was determined (Supplementary Fig. S15). The LD decays across the other subpopulations with only the average LD decay in the IND subpopulation being lower. In addition to LD heatmap, the estimated LD distance in the region (9017 SNPs at 6.5–9.5 Mbp) in the TEJ subpopulation was 1.23 Mbp from 7.64 to 8.87 Mbp, indicating few historical recombination events. It was similar to a large LD block in the AUS (23,041 SNPs) subpopulation, whereas several LD blocks in the IND (13,731 SNPs) and TRJ (6513 SNPs) subpopulations were observed (Supplementary Fig. S15).

For multi-experiment analysis, conducting GWA mapping for grain Mn concentration with the phenotypic values of the three flooded-field experiments (ArFI06, ArFI07 and TxFI09) using the mvLMM, there were two QTLs that had previously been detected in the single-site analysis and six newly identified QTLs (Fig. 5). Similarly, Korte et al. (2012) re-analysed the flowering time data of Li et al. (2010) in 459 *A. thaliana* accessions grown over two seasons in each of two

different locations using MTMM (Multi-trait mixed model) to reveal new QTLs. Three detected loci were involved in the differential flowering response to different environments that were not detected in the individual screens. Indeed, multi-trait analysis is an efficient tool for detecting loci/QTLs associated with multiple traits, because of the increased power obtained from additional data from correlated traits or a single trait in multiple experiments (Korte et al. 2012; Zhou and Stephens 2014). Thus, this approach should be used to identify stable QTLs, and is potentially beneficial in terms of GWAS of complex traits. The validation of the new QTLs could be further studied for identification of candidate genes underlying these QTLs that may contribute to the ultimate grain Mn concentration in rice.

While gene expression data were not collected for the plants grown in this experiment, recently transcriptomic analysis for 91 of the RDP1 accessions was conducted (Campbell et al. 2020). This database consists of gene expression data from shoots, and can be used to determine if genes are differentially expressed between accessions. For candidate genes discussed, all were found to be expressed in shoots with *OsMTP8.1*, *OsHMA3* and *OsNRAMP1* differentially expressed between the low- and high-grain Mn accessions (Supplementary Fig. S12). Differential gene expression means that these genes are very good candidates for the trait as this expression difference could be driving the QTLs. However, future analysis of gene expression between low and high Mn-accumulating accessions during grain filling will give a further insight into the role these genes play in the Mn accumulation in the grain.

Conclusion

This study uses data from multiple field experiments (locations, years and irrigation treatments) to conduct GWA mapping for a grain elemental trait, Mn concentration, in rice. We have demonstrated that multi-experiment analysis has a number of potential benefits, including the identification of QTLs not detected in individual analyses. Future study would be required to validate these genes, and identify the alleles that are responsible for variation in Mn accumulation in rice grains.

Acknowledgements This research was partly supported by the US National Science Foundation, Plant Genome Research Program (grant #IOS 0701119 to DES, MLG and SRMP) and The US National Institutes of Health (grant 2P4ES007373 to MLG and DES). PR is a PhD student funded by the Thai Government Scholarship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

References

- Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y et al. (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465:627–631
- Bashir K, Takahashi R, Nakanishi H, Nishizawa NK (2013) The road to micronutrient biofortification of rice: progress and prospects. *Front Plant Sci* 4:15
- Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E et al. (2010) A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter *AtHKT1;1*. *PLoS Genet* 6:e1001193
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300
- Bettembourg M, Dardou A, Audebert A, Thomas E, Frouin J, Guiderdoni E et al. (2017) Genome-wide association mapping for root cone angle in rice. *Rice* 10:45
- Biscarini F, Cozzi P, Casella L, Riccardi P, Vattari A, Orasen G et al. (2016) Genome-wide association study for traits related to plant and grain morphology, and root architecture in temperate rice accessions. *PLoS One* 11:e0155425
- Butardo VM, Anacleto R, Parween S, Samson I, de Guzman K, Alhambra CM et al. (2017) Systems genetics identifies a novel regulatory domain of amylose synthesis. *Plant Physiol* 173:887–906
- Campbell MT, Du Q, Liu K, Sharma S, Zhang C, Walia H (2020) Characterization of the transcriptional divergence between the subspecies of cultivated rice (*Oryza sativa*). *BMC Genom* 21:394
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7
- Chen Z, Fujii Y, Yamaji N, Masuda S, Takemoto Y, Kamiya T et al. (2013) Mn tolerance in rice is mediated by MTP8.1, a member of the cation diffusion facilitator family. *J Exp Bot* 64:4375–4387
- Chen X, Li J, Wang L, Ma G, Zhang W (2016) A mutagenic study identifying critical residues for the structure and function of rice manganese transporter *OsmTP8.1*. *Sci Rep* 6:32073
- Chu HH, Car S, Socha AL, Hindt MN, Punshon T, Guerinot ML (2017) The *Arabidopsis* MTP8 transporter determines the localization of manganese and iron in seeds. *Sci Rep* 7:11024
- Crossgrove J, Zheng W (2004) Manganese toxicity upon over-exposure. *NMR Biomed* 17:544–553
- Dimkpa SON, Lahari Z, Shrestha R, Douglas A, Gheysen G, Price AH (2016) A genome-wide association study of a global rice panel reveals resistance in *Oryza sativa* to root-knot nematodes. *J Exp Bot* 67:1191–1200
- Du J, Zeng D, Wang B, Qian Q, Zheng S, Ling HQ (2013) Environmental effects on mineral accumulation in rice grains and identification of ecological specific QTLs. *Environ Geochem Health* 35:161–170
- Ducic T, Polle A (2005) Transport and detoxification of manganese and copper in plants. *Braz J Plant Physiol* 17:103–112
- Eizenga GC, Ali ML, Bryant RJ, Yeater KM, McClung AM, McCouch SR (2014) Registration of the rice diversity panel 1 for genomewide association studies. *J Plant Regist* 8:109–116
- Eroglu S, Giehl RFH, Meier B, Takahashi M, Terada Y, Ignatyev K et al. (2017) Metal tolerance protein 8 mediates manganese homeostasis and iron reallocation during seed development and germination. *Plant Physiol* 174:1633–1647
- Famoso AN, Zhao K, Clark RT, Tung C-W, Wright MH, Bustamante C et al. (2011) Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genet* 7:e1002221
- Gao H, Xie W, Yang C, Xu J, Li J, Wang H et al. (2018) NRAMP2, a trans-golgi network-localized manganese transporter, is required for *Arabidopsis* root growth under manganese deficiency. *N Phytologist* 217:179–193
- Halder D, Saha JK, Biswas A (2020) Accumulation of essential and non-essential trace elements in rice grain: possible health impacts on rice consumers in West Bengal. *India Sci Total Env* 706:135944
- Hebbert CA, Pedas P, Schjoerring JK, Knudsen L, Husted S (2005) Genotypic differences in manganese efficiency: field experiments with winter barley (*Hordeum vulgare* L.). *Plant Soil* 272:233–244
- Huang XY, Salt DE (2016) Plant ionomics: from elemental profiling to environmental adaptation. *Mol Plant* 9:787–797
- Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y et al. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961–967
- Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q 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
- Institute of Medicine (2001) Dietary reference intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. National Academies Press, Washington DC
- Ishikawa S, Abe T, Kuramata M, Yamaguchi M, Ando T, Yamamoto T et al. (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
- Ishimaru Y, Masuda H, Bashir K, Inoue H, Tsukamoto T, Takahashi M et al. (2010) Rice metal-nicotianamine transporter, *OsYSL2*, is required for the long-distance transport of iron and manganese. *Plant J* 62:379–390
- Ishimaru Y, Takahashi R, Bashir K, Shimo H, Senoura T, Sugimoto K et al. (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 S-Y, Freimer NB et al. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42:348–354
- Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ et al. (2008) Efficient control of population structure in model organism association mapping. *Genetics* 178:1709–1723
- Koike S, Inoue H, Mizuno D, Takahashi M, Nakanishi H, Mori S et al. (2004) *OsYSL2* is a rice metal-nicotianamine transporter that is regulated by iron and expressed in the phloem. *Plant J* 39:415–424
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29
- Korte A, Vilhjálmsson BJ, Segura V, Platt A, Long Q, Nordborg M (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat Genet* 44:1066–1071
- Kump KL, Bradbury PJ, Wissner RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA et al. (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat Genet* 43:163–168
- Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO (2010) Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proc Natl Acad Sci* 107:21199–21204
- Lippert C, Casale FP, Rakitsch B, Stegle O (2014) LIMIX: genetic analysis of multiple traits. *bioRxiv*
- Liu C, Chen G, Li Y, Peng Y, Zhang A, Hong K et al. (2017) Characterization of a major QTL for manganese accumulation in rice grain. *Sci Rep* 7:17704

- Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM et al. (2015) Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* 47:284–290
- Lu K, Li L, Zheng X, Zhang Z, Mou T, Hu Z (2008) Quantitative trait loci controlling Cu, Ca, Zn, Mn and Fe content in rice grains. *J Genet* 87:305–310
- Ma G, Li J, Li J, Li Y, Gu D, Chen C et al. (2018) OsMTP11, a trans-golgi network localized transporter, is involved in manganese tolerance in rice. *Plant Sci* 274:59–69
- 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
- Marschner H (1995) Mineral nutrition of higher plants. Academic Press, London
- McCouch SR, Wright MH, Tung C-W, Maron LG, McNally KL, Fitzgerald M et al. (2016) Open access resources for genome-wide association mapping in rice. *Nat Commun* 7:10532
- Millaleo R, Reyes-Diaz M, Ivanov AG, Mora ML, Alberdi M (2010) Manganese as essential and toxic element for plants: transport, accumulation and resistance mechanisms. *J Soil Sci Plant Nutr* 10:476–494
- 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, Douglas A, Lahner B, Yakubova E, Guerinot ML, Pinson SRM 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, Duan GL, Lei M, Zhu YG, Meharg AA, Price AH (2012a) 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, Pinson SRM, Alexander J, McKay S, Hansen H, Duan GL et al. (2012b) Variation in grain arsenic assessed in a diverse panel of rice (*Oryza sativa*) grown in multiple sites. *N Phytologist* 193:650–664
- Norton GJ, Travis AJ, Douglas A, Fairley S, Alves EDP, Ruangareerate P et al. (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
- O'Neal SL, Zheng W (2015) Manganese toxicity upon overexposure: a decade in review. *Curr Environ Heal Rep* 2:315–328
- Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2:e190
- Peris-Peris C, Serra-Cardona A, Sánchez-Sanuy F, Campo S, Ariño J (2017) Two NRAMP6 isoforms function as iron and manganese transporters and contribute to disease resistance in rice. *Mol Plant-Microbe Interact* 30:385–398
- Pinson SRM, Tarpley L, Yan W, Yeater K, Lahner B, Yakubova E et al. (2015) Worldwide genetic diversity for mineral element concentrations in rice grain. *Crop Sci* 55:294–311
- Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 11:459–463
- R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA et al. (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–427
- Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y et al. (2013) Rice annotation project database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol* 54:e6
- 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
- Sasaki A, Yamaji N, Xia J, Ma JF (2011) OsYSL6 is involved in the detoxification of excess manganese in rice. *Plant Physiol* 157:1832–1840
- 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
- Senewiratne ST, Mikkelsen DS (1961) Physiological factors limiting growth and yield of *Oryza sativa* under unflooded conditions. *Plant Soil* 14:127–146
- Shin JH, Blay S, McNeney B, Graham J (2006) LDheatmap: an R function for graphical display of pairwise linkage disequilibrium between single nucleotide polymorphisms. *J Stat Softw* 16:c03
- Shrestha A, Dziwornu AK, Ueda Y, Wu LB, Mathew B, Frei M (2018) Genome-wide association study to identify candidate loci and genes for Mn toxicity tolerance in rice. *PLoS One* 13:e0192116
- Socha AL, Guerinot ML (2014) Mn-euvering manganese: the role of transporter gene family members in manganese uptake and mobilization in plants. *Front Plant Sci* 5:106
- Soetan K, Olaiya CO, Oyewole OE (2010) The importance of mineral elements for humans, domestic animals and plants: a review. *Afr J Food Sci* 4:200–222
- Stangoulis JCR, Huynh BL, Welch RM, Choi EY, Graham RD (2007) Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content. *Euphytica* 154:289–294
- Takahashi R, Ishimaru Y, Nakanishi H, Nishizawa NK (2011) Role of the iron transporter OsNRAMP1 in cadmium uptake and accumulation in rice. *Plant Signal Behav* 6:1813–1816
- Takemoto Y, Tsunemitsu Y, Fujii-Kashino M, Mitani-Ueno N, Yamaji N, Ma JF et al. (2017) The tonoplast-localized transporter MTP8.2 contributes to manganese detoxification in the shoots and roots of *Oryza sativa* L. *Plant Cell Physiol* 58:1573–1582
- Thomine S, Wang R, Ward JM, Crawford NM, Schroede JI (2000) Cadmium and iron transport by members of a plant metal transporter family in Arabidopsis with homology to Nramp genes. *Proc Natl Acad Sci* 97:4991–4996
- Turner SD (2014) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*
- Ueno D, Sasaki A, Yamaji N, Miyaji T, Fujii Y, Takemoto Y et al. (2015) A polarly localized transporter for efficient manganese uptake in rice. *Nat Plants* 1:15170
- Vatansever R, Filiz E, Ozyigit II (2016) In silico analysis of Mn transporters (NRAMP1) in various plant species. *Mol Biol Rep* 43:151–163
- Wang DR, Agosto-Pérez FJ, Chebotarov D, Shi Y, Marchini J, Fitzgerald M et al. (2018) An imputation platform to enhance integration of rice genetic resources. *Nat Commun* 9:3519
- Yang M, Lu K, Zhao F-J, Xie W, Ramakrishna P, Wang G et al. (2018) Genome-wide association studies reveal the genetic basis of ionic variation in rice. *Plant Cell* 30:2720–2740
- Yang M, Zhang W, Dong H, Zhang Y, Lv K, Wang D et al. (2013) OsNRAMP3 is a vascular bundles-specific manganese transporter that is responsible for manganese distribution in rice. *PLoS One* 8:e83990
- Yang M, Zhang Y, Zhang L, Hu J, Zhang X, Lu K et al. (2014) OsNRAMP5 contributes to manganese translocation and distribution in rice shoots. *J Exp Bot* 65:4849–4861
- Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P, Hu L et al. (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat Genet* 48:927–934
- Yokosho K, Yamaji N, Ueno D, Mitani N, Ma JF (2009) OsFRDL1 is a citrate transporter required for efficient translocation of iron in rice. *Plant Physiol* 149:297–305

- Zhang M, Liu B (2017) Identification of a rice metal tolerance protein OsMTP11 as a manganese transporter. *PLoS One* 12:e0174987
- Zhang M, Pinson SRM, Tarpley L, Huang XY, Lahner B, Yakubova E et al. (2014) Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theor Appl Genet* 127:137–165
- Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH et al. (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun* 2:467
- Zhou X, Stephens M (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat Methods* 11:407–409